Pacific Journal of Mathematics

Volume 254 No. 2 December 2011

PACIFIC JOURNAL OF MATHEMATICS

http://pacificmath.org

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

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

The subscription price for 2011 is US $420/year for the electronic version, and $485/year for print and electronic.
Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840, is published monthly except July and August. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

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

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS
at the University of California, Berkeley 94720-3840
A NON-PROFIT CORPORATION
Typeset in LaTeX

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2012

CURVATURES OF SPHERES IN HILBERT GEOMETRY

ALEXANDER BORISENKO AND EUGENE OLIN

We prove that the normal curvatures of hyperspheres, the Rund curvature, and the Finsler curvature of circles in Hilbert geometry tend to 1 as the radii tend to infinity.

1. Introduction

A smooth connected manifold M^n is called a *Finsler* manifold [Bao et al. 2000] if there is a smooth positively homogeneous function $F : TM^n \to [0, \infty)$ on the coordinates in tangent spaces such that the symmetric bilinear form

$$\boldsymbol{g}_y(u, v) = g_{ij}(x, y)u^i v^j : T_x M^n \times T_x M^n \to \mathbb{R}$$

is positively definite for each pair $(x, y) \in TM^n$, where $g_{ij}(x, y) = \frac{1}{2}[F^2(x, y)]_{y^i y^j}$.

Consider a bounded open convex domain U in $\mathbb{R}^n$ with the Euclidean norm $\|\cdot\|$, and let ∂U be a C^3 hypersurface with positive normal curvatures. For a point $x \in U$ and a tangent vector $y \in T_x U = \mathbb{R}^n$, let x_- and x_+ be the intersection points of the rays $x + \mathbb{R}_- y$ and $x + \mathbb{R}_+ y$ with *absolute* ∂U. Then the Hilbert metric is defined as follows:

$$F(x, y) = \tfrac{1}{2}(\Theta(x, y) + \Theta(x, -y)), \tag{1}$$

where

$$\Theta(x, y) = \|y\| \frac{1}{\|x - x_+\|}, \quad \Theta(x, -y) = \|y\| \frac{1}{\|x - x_-\|}$$

are called the Funk metrics on U.

Hilbert geometries are the generalizations of Klein's model of the hyperbolic geometry. Hilbert geometries are also Finsler spaces of constant negative flag curvature -1 [Bao et al. 2000]. The Hilbert metric is invariant under projective transformations of $\mathbb{R}^n$ leaving U bounded.

B. Colbois and P. Verovic [2002] proved that the Hilbert metric is asymptotically Riemannian at infinity. That means that in a given Hilbert geometry the unit sphere

MSC2010: 53B40, 53C60.
Keywords: Hilbert geometry, Finsler geometry, spheres, norman curvature, Rund curvature, Finsler curvature.

of the norm $F(x,\cdot)$ approaches the ellipsoid in C^0 topology as the point x tends to ∂U.

Unlike the Riemannian geometry, in the Finsler geometry there are several definitions of the curvature of a curve.

The normal curvature of a hypersurface in a Finsler space is defined as follows [Shen 2001]. Let $\varphi : N \to M^n$ be a hypersurface in a Finsler manifold M^n. A vector $\boldsymbol{n} \in T_{\varphi(x)}M^n$ is called a normal vector to N at the point $x \in N$ if $\boldsymbol{g_n}(y, \boldsymbol{n}) = 0$ for all $y \in T_x N$. The *normal curvature* $\boldsymbol{k_n}$ at the point $x \in N$ in a direction $y \in T_x N$ is defined as

$$\boldsymbol{k_n} = \boldsymbol{g_n}(\nabla_{\dot c(s)}\dot c(s)|_{s=0}, \boldsymbol{n}), \tag{2}$$

where $\dot c(0) = y$, $c(s)$ is a geodesic in the induced connection on N, and $\boldsymbol{n}$ is the chosen unit normal vector.

For a curve $c(s)$ parametrized by its arc length in M^n, it is possible to define two more curvatures.

The Finsler curvature of $c(s)$ [Finsler 1951; Rund 1959] is defined as

$$\boldsymbol{k}_F(c(s)) = \sqrt{\boldsymbol{g}_{\dot c(s)}(\nabla_{\dot c(s)}\dot c(s), \nabla_{\dot c(s)}\dot c(s))}. \tag{3}$$

The Rund curvature of $c(s)$ [Rund 1959] is defined as

$$\boldsymbol{k}_R(c(s)) = \sqrt{\boldsymbol{g}_{\nabla_{\dot c(s)}\dot c(s)}(\nabla_{\dot c(s)}\dot c(s), \nabla_{\dot c(s)}\dot c(s))}. \tag{4}$$

It is well-known that the normal curvatures of hyperspheres in the hyperbolic space $\mathbb{H}^n$ are equal to $\coth(r)$ and tend to 1 as the radius r tends to infinity. We prove the same property for the Hilbert geometry.

Theorem 1.1. *The normal curvature, the Rund curvature, and the Finsler curvature of the circles centered at the same point in the* 2*-dimensional Hilbert geometry tend to* 1 *as their radii tend to infinity, uniformly at the point of the circle.*

Theorem 1.2. *The normal curvatures of the hyperspheres centered at the same point tend to* 1 *as their radii tend to infinity, uniformly at the point of the hypersphere and in the tangent vector at this point of the hypersphere.*

This can be interpreted as meaning that the Hilbert metric tends to the Riemannian metric of the hyperbolic space in C^2-topology.

2. The choice of the coordinate system

Consider the Hilbert geometry based on a two-dimensional domain U in the Euclidean plane. Fix a point o in the domain U and a point $p \in \partial U$. Since ∂U is a convex curve, it admits the polar representation $\omega(\varphi)$ from the point o such that the point p corresponds to $\varphi = 0$.

Choose the coordinate system on the plane with the origin O at the point p; let the axis x_2 be orthogonal to ∂U at p, x_1 be tangent to ∂U at p, and $U - \{p\}$ lie in the half-plane $x_2 > 0$.

In this section we will construct a projective transformation P of the plane that sends U to $\hat{U}$ and has the following properties:

(1) $P(p) = p$.

(2) The vector $u = (0, 1)$ is orthogonal to $\partial\hat{U}$ at the point p.

(3) The tangent line to $\partial\hat{U}$ at the point p is parallel to the tangent line to $\partial\hat{U}$ at the point corresponding to $\varphi = \pi$.

(4) $\partial\hat{U}$ is the graph of the function $x_2 = \hat{f}(x_1)$ such that $\hat{f}(0) = 0$, $\hat{f}'(0) = 0$, and $\hat{f}''(0) = \frac{1}{2}$ in the neighborhood of p.

We are going to give the explicit expression for this transformation and show that after this transformation the curvature of $\partial\hat{U}$ and the derivatives of f remain uniformly bounded.

The next lemma gives the upper bound on the angle between the radial and normal direction to the convex curve.

Lemma 2.1 [Borisenko 2002]. *Let γ be a closed embedded curve in the Euclidean plane whose curvature is greater than or equal to k. Let o be a point in the interior of the set bounded by γ, ω_0 the distance from o to γ, and φ the angle between the outer normal vector at the point $p \in \gamma$ and the vector op. Then*

$$\cos\angle(u_m, N(m)) \geqslant \omega_0 k. \tag{5}$$

Denote by k and K the minimum and maximum of the curvatures of ∂U. Also, $\omega_0 = \min_\varphi \omega(\varphi)$, $\omega_1 = \max_\varphi \omega(\varphi)$.

Let the length of the chord of U in the direction u equal H, the distance from o to the origin equal ω_u, $\omega_0 \leqslant \omega_u \leqslant \omega_1$, and the angle between u and x_2 equal α.

Step 1. Construct an affine transformation that makes the vector $\overrightarrow{oO}$ parallel to x_2. This transformation sends the points $(0, 0)$ and $(1, 0)$ to themselves, the point $(H \sin\alpha, H \cos\alpha) \in \partial U$ to the point $(0, H)$, and has the expression:

$$\tilde{x}_1 = x_1 - \tan\alpha x_2, \quad \tilde{x}_2 = \frac{x_2}{\cos\alpha}. \tag{6}$$

Denote the image of U as $\tilde{U}$. The point o now has the coordinates $(0, \omega_u)$. Denote by $\tilde{k}$ the minimum of the curvature of $\partial\tilde{U}$ in the $(\tilde{x}_1, \tilde{x}_2)$ coordinate system, and by $\tilde{\omega}_0$ denote the distance from the point $(0, \omega_u)$ to $\partial\tilde{U}$. Note that the eigenvalues of the transformation (6) are equal to 1 and $1/\cos\alpha$. Hence

$$\omega_0 \leqslant \tilde{\omega}_0 \leqslant \frac{1}{\cos\alpha}\omega_0. \tag{7}$$

Lemma 2.1 then implies that the curvature of $\partial\tilde{U}$ remains bounded and separated from zero.

Step 2. Construct the transformation such that the tangent line

$$\tilde{x}_2 = -\tan\beta\tilde{x}_1 + H$$

to $\partial\tilde{U}$ at the point $(0, H)$ will be parallel to the axis $\tilde{x}_1$, where β is the angle between $\tilde{x}_2$ and the normal vector to $\partial\tilde{U}$ at $(0, H)$. This transformation has the expression

$$\bar{x}_1 = \frac{H\tilde{x}_1}{H-\tan\beta\tilde{x}_1}, \quad \bar{x}_2 = \frac{H\tilde{x}_2}{H-\tan\beta\tilde{x}_1}. \tag{8}$$

Denote the image of $\tilde{U}$ as $\bar{U}$.

We can estimate the angle $|\tan\beta|$. Using Lemma 2.1, we have

$$0 \leqslant |\tan\beta| \leqslant \sqrt{\frac{1}{(\tilde{k}^2\tilde{\omega}_0^2)} - 1}. \tag{9}$$

Estimate the curvature $\partial\bar{U}$. Let the curve $\partial\tilde{U}$ be given in the parametric form $r(t) = (\tilde{x}_1(t), \tilde{x}_2(t))$. Then $\partial\bar{U}$ has the parametrization

$$\bar{r}(t) = \frac{Hr(t)}{H-\tan\beta\tilde{x}_1(t)}.$$

Differentiating leads to

$$\bar{r}'(t) = \frac{Hr'(t)}{H-\tan\beta\tilde{x}_1(t)} + \frac{Hr(t)\tan\beta\tilde{x}_1'(t)}{(H-\tan\beta\tilde{x}_1(t))^2},$$

$$\bar{r}''(t) = \frac{2H\tan\beta r'(t)\tilde{x}_1'(t)}{(H-\tan\beta\tilde{x}_1(t))^2} + \frac{2Hr(t)\tan^2\beta\tilde{x}_1'(t)^2}{(H-\tan\beta\tilde{x}_1(t))^3} + \frac{Hr''(t)}{H-\tan\beta\tilde{x}_1(t)} + \frac{Hr(t)\tan\beta\tilde{x}_1''(t)}{(H-\tan\beta\tilde{x}_1(t))^2}.$$

The strict convexity of $\partial\tilde{U}$ implies that $H-\tan\beta\tilde{x}_1(t) \geqslant const > 0$ for each t. This and the compactness argument leads to the maximum of the curvature of $\partial\bar{U}$ being bounded from above for some constant.

If the curve $\partial\tilde{U}$ is the graph $\tilde{x}_2 = f(\tilde{x}_1)$ and $f(0) = f'(0) = 0$, then its curvature at the point $(0, 0)$ after the transformation (8) will not change. Indeed,

$$\tilde{x}_1'(t)^2 + \tilde{x}_2'(t)^2 = \left(\frac{Ht\tan\beta}{(H-t\tan\beta)^2} + \frac{H}{H-t\tan\beta}\right)^2 + \left(\frac{H\tan\beta f(t)}{(H-t\tan\beta)^2} + \frac{Hf'(t)}{H-t\tan\beta}\right)^2,$$

$$\tilde{x}_1'(t)\tilde{x}_2''(t)-\tilde{x}_1''(t)\tilde{x}_2'(t)$$
$$=-\left(\frac{2Ht\tan^2\beta}{(H-t\tan\beta)^3}+\frac{2H\tan\beta}{(H-t\tan\beta)^2}\right)\left(\frac{H\tan\beta f(t)}{(H-t\tan\beta)^2}+\frac{Hf'(t)}{H-t\tan\beta}\right)$$
$$+\left(\frac{Ht\tan\beta}{(H-t\tan\beta)^2}+\frac{H}{H-t\tan\beta}\right)\left(\frac{2H\tan^2\beta f(t)}{(H-t\tan\beta)^3}+\frac{2H\tan\beta f'(t)}{(H-t\tan\beta)^2}+\frac{Hf''(t)}{H-t\tan\beta}\right).$$

We obtain the claim after substituting the equalities $f(0)=f'(0)=0$. So the curvature of $\partial\bar{U}$ at the origin is still separated from zero.

Step 3. Construct a transformation such that the distance from $(0,\omega_u)$ to the origin is equal to 1 and the curvature of $\partial\bar{U}$ at the origin is equal to $\frac{1}{2}$. This transformation has the expression:

$$\hat{x}_1=\frac{\bar{x}_1}{\omega_u},\quad \hat{x}_2=\frac{\bar{x}_2}{2\omega_u^2\bar{k}(0)}. \tag{10}$$

Denote the image of $\bar{U}$ as $\hat{U}$. It is obvious that the curvature of $\partial\hat{U}$ remains bounded.

The announced transformation P is the composition of the transformations (6), (8), and (10), and the following proposition holds:

Proposition 2.2. *There exists a constant C_0 depending on U such that the curvature of $P(\partial U)$ is bounded from above by C_0.*

Let ∂U be the graph of the function $x_2=f(x_1)$ in the initial coordinate system. After the transformation P, $P(\partial U)$ can be considered the graph of the function $x_2=\hat{f}(x_1)$ such that $\hat{f}(0)=0$, $\hat{f}'(0)=0$, and $\hat{f}''(0)=\frac{1}{2}$ in the neighborhood of p.

Finally, estimate the third derivative $\hat{f}'''(0)$. Evidently, under the affine transformations (6) and (10) the third derivative remains bounded. We only need to control $f'''(0)$ at Step 2.

So let the curve $\partial\tilde{U}$ be the graph $\tilde{x}_2=\tilde{f}(\tilde{x}_1)$ and after the transformation (8) we obtain the graph $\bar{f}$. The rules for differentiation lead to

$$\bar{f}'''(0)=\tilde{f}'''(0)-\frac{\tan\beta\tilde{k}(0)}{H}. \tag{11}$$

As ∂U is the compact curve, we obtain:

Proposition 2.3. *There exist constants C_1, C_2 depending on U, such that*

$$C_1\leqslant\hat{f}'''(0)\leqslant C_2.$$

Analogously we can estimate all higher derivatives.

The Hilbert metrics for the domains U and $\hat{U}$ are isometric. Therefore, without loss of generality, we will consider the Hilbert metric for the domain $\hat{U}$ and will denote $\hat{U}$ by U.

3. Series expansions for the metric tensor of the Hilbert metric

From the decomposition of the Hilbert metric through the Funk metrics (1), we conclude

$$\begin{aligned} g_{ij}(x,y) &= \mathrm{F}(x,y)\mathrm{F}_{y^i y^j}(x,y) + \mathrm{F}_{y^i}(x,y)\mathrm{F}_{y^j}(x,y) \\ &= \tfrac{1}{2}\mathrm{F}(x,y)(\Theta_{y^i y^j}(x,y) + \Theta_{y^i y^j}(x,-y)) \\ &\qquad + \tfrac{1}{4}(\Theta_{y^i}(x,y) - \Theta_{y^i}(x,-y))(\Theta_{y^j}(x,y) - \Theta_{y^j}(x,-y)). \end{aligned}$$

Okada's lemma [Shen 2001] for Funk metrics gives the expression of the derivatives of $\Theta(x,y)$ with respect to the coordinates on tangent spaces through the derivatives with respect to the coordinates on U:

$$\Theta(x,y)_{x^k} = \Theta(x,y)\Theta(x,y)_{y^k}.$$

Using this lemma, we can write:

$$\begin{aligned} g_{ij}(x,y) &= \tfrac{1}{2}\mathrm{F}(x,y)\frac{\Theta_{x^i x^j}(x,y)\Theta(x,y) - 2\Theta_{x_i}(x,y)\Theta_{x_j}(x,y)}{\Theta(x,y)^3} \\ &\quad + \tfrac{1}{2}\mathrm{F}(x,y)\frac{\Theta_{x^i x^j}(x,-y)\Theta(x,-y) - 2\Theta_{x_i}(x,-y)\Theta_{x_j}(x,-y)}{\Theta(x,-y)^3} \\ &\quad + \tfrac{1}{4}\left(\frac{\Theta_{x^i}(x,y)}{\Theta(x,y)} - \frac{\Theta_{x^i}(x,-y)}{\Theta(x,-y)}\right)\left(\frac{\Theta_{x^j}(x,y)}{\Theta(x,y)} - \frac{\Theta_{x^j}(x,-y)}{\Theta(x,-y)}\right). \end{aligned} \tag{12}$$

For convenience we will use lower indices x_i for coordinates. Let $\mathrm{F}(x_1,x_2,y_1,y_2)$ be a two-dimensional Hilbert metric and $\Theta(x_1,x_2,y_1,y_2)$ the corresponding Funk metric. Assume that the point (x_1,x_2) is sufficiently close to ∂U. Then we can express ∂U as the graph $x_2 = f(x_1)$ such that $f(0)=0$, $f'(0)=0$, and $f''(0)=\frac{1}{2}$. Consider a point (x_1,x_2) above the graph $x_2 = f(x_1)$. Denote by $r(x_1,x_2,y_1,y_2)$ the distance between the point (x_1,x_2) and the intersection point of the line passing trough (x_1,x_2) in the direction (y_1,y_2) with the curve $x_2 = f(x_1)$. Then

$$\Theta(x_1,x_2,y_1,y_2) = \sqrt{y_1^2+y_2^2}\, r(x_1,x_2,y_1,y_2)^{-1}. \tag{13}$$

Now we obtain the derivatives of $r(x_1,x_2,y_1,y_2)$ on x_1, x_2. The parameter $t(x_1,x_2,y_1,y_2)$ corresponding to the intersection points of the curve $x_2 = f(x_1)$ with the line

$$x_1(t) = x_1 + ty_1, \quad x_2(t) = x_2 + ty_2$$

satisfies the functional equation

$$x_2 + t y_2 = f(x_1 + t(x_1, x_2, y_1, y_2) y_1). \tag{14}$$

Differentiate (14) on x_1, x_2:

$$t_{x_1} y_2 = f'(x_1 + t y_1)(1 + t_{x_1} y_1), \quad 1 + t_{x_2} y_2 = f'(x_1 + t y_1) t_{x_2} y_1. \tag{15}$$

We obtain the explicit expressions for t_{x_1}, t_{x_2}:

$$t_{x_1} = \frac{f'(x_1 + t y_1)}{y_2 - y_1 f'(x_1 + t y_1)}, \quad t_{x_2} = \frac{1}{y_1 f'(x_1 + t y_1) - y_2}. \tag{16}$$

Differentiating (15) leads to

$$\begin{aligned} y_2 t_{x_1 x_1} &= f''(x_1 + t y_1)(1 + y_1 t_{x_1})^2 + f'(x_1 + t y_1) y_1 t_{x_1 x_1}, \\ y_2 t_{x_1 x_2} &= f''(x_1 + t y_1)(1 + y_1 t_{x_1}) y_2 t_{x_2} + f'(x_1 + t y_1) y_1 t_{x_1 x_2}, \\ y_2 t_{x_2 x_2} &= f''(x_1 + t y_1)(y_1 t_{x_2})^2 + f'(x_1 + t y_1) y_1 t_{x_2 x_2}. \end{aligned} \tag{17}$$

We obtain the expressions for the second derivatives of t:

$$\begin{aligned} t_{x_1 x_1} &= \frac{f''(x_1 + t y_1)(1 + y_1 t_{x_1})^2}{y_2 - y_1 f'(x_1 + t y_1)}, \\ t_{x_1 x_2} &= \frac{f''(x_1 + t y_1)(1 + y_1 t_{x_1}) y_1 t_{x_2}}{y_2 - y_1 f'(x_1 + t y_1)}, \\ t_{x_2 x_2} &= \frac{f''(x_1 + t y_1)(y_1 t_{x_2})^2}{y_2 - y_1 f'(x_1 + t y_1)}. \end{aligned} \tag{18}$$

We need the derivatives of $r(x_1, x_2, y_1, y_2)$. By definition,

$$r(x_1, x_2, y_1, y_2) = \sqrt{(y_1 t)^2 + (y_2 t)^2} = \sqrt{y_1^2 + y_2^2}\, t(x_1, x_2, y_1, y_2).$$

Hence $r_{x_k} = t_{x_k}$ and $r_{x_k x_l} = t_{x_k x_l}$.

Now it is possible to calculate the derivatives of the Funk metric. Formula (13) implies

$$\Theta_{x_k} = -\sqrt{y_1^2 + y_2^2}\, \frac{r_{x_k}}{r^2}. \tag{19}$$

After differentiating (19), we obtain

$$\Theta_{x_k x_l} = -\sqrt{y_1^2 + y_2^2}\, \frac{r_{x_k x_l} r^2 - 2 r r_{x_l} r_{x_k}}{r^4} = \sqrt{y_1^2 + y_2^2}\,(2\Theta^3 r_{x_k} r_{x_l} - \Theta^2 r_{x_k x_l}). \tag{20}$$

Finally, from (12) it is possible to obtain the coefficients of the metric tensor. We will need the values of $g_{ij}(x_1, x_2, y_1, y_2)$ at the points $(x_1, x_2) = (0, x_2)$.

3.1. ***Expansions for $g_{ij}(0, x_2, 1, 0)$.*** Note that the strict convexity of ∂U implies that $f'(t(x_1, x_2)) \neq 0$ for $t(x_1, x_2) \neq 0$. Then from (16) we deduce

$$t_{x_1}(0, x_2, 1, 0) = -1, \tag{21}$$

$$t_{x_2}(0, x_2, 1, 0) = \frac{1}{f'(t(0, x_2, 1, 0))}, \tag{22}$$

and from (18)

$$t_{x_1x_1}(0, x_2, 1, 0) = t_{x_1x_2}(0, x_2, 1, 0) = 0, \tag{23}$$

$$t_{x_2x_2}(0, x_2, 1, 0) = -\frac{f''(t(0, x_2, 1, 0))}{f'(t(0, x_2, 1, 0))^3}. \tag{24}$$

Expanding the functional equation (14) in a power series with respect to t as $x_2 \to 0$, we find the expansions of $t(0, x_2, 1, 0)$.

$$x_2 = \tfrac{1}{4}t^2 + \tfrac{1}{6}f'''(0)t^3 + O(t^4). \tag{25}$$

We will find t in expanded form

$$t = A + B\sqrt{x_2} + Cx_2 + Dx_2^{3/2} + O(x_2^2) \tag{26}$$

After substituting (26) into (25) and transposing all members in the left side, we obtain the following system of equations:

$$\begin{aligned} &3A^2 + 2A^3f'''(0) + (6AB + 6A^2f'''(0)B)\sqrt{x_2} \\ &\quad + (-12 + 3B^2 + 6Af'''(0)B^2 + 6AC + 6A^2f'''(0)C)x_2 + (2f'''(0)B^2 \\ &\qquad + 6BC + 12Af'''(0)BC + 6AD + 6A^2f'''(0)D)x_2^{3/2} + O(x_2^2) = 0. \end{aligned} \tag{27}$$

Choose the coefficients A, B, C, and D so that the left side of (27) is $O(x_2^2)$. Equating the coefficients under the powers of x_2 to zero we obtain two expansions for t which correspond to the directions $(1, 0)$ and $(-1, 0)$.

$$t(0, x_2, \pm 1, 0) = \pm 2\sqrt{x_2} - \tfrac{4}{3}f'''(0)x_2 + O(x_2^2). \tag{28}$$

In our case $r = t$, so we get

$$r(0, x_2, 1, 0) = 2\sqrt{x_2} - \tfrac{4}{3}f'''(0)x_2 + O(x_2^2). \tag{29}$$

Later on, all power series will be considered in the neighborhood of 0. The series expansion for the metric F is

$$\begin{aligned}
\mathrm{F}(0, x_2, 1, 0) &= \tfrac{1}{2}\left(\frac{1}{r(0, x_2, 1, 0)} + \frac{1}{r(0, x_2, -1, 0)}\right)\\
&= \tfrac{1}{2}\left(\frac{1}{2\sqrt{x_2} - \frac{4}{3}f'''(0)x_2 + O(x_2^2)} + \frac{1}{2\sqrt{x_2} + \frac{4}{3}f'''(0)x_2 + O(x_2^2)}\right)\\
&= \frac{9}{\sqrt{x_2}(18 - 8f'''(0)^2 x_2) + O(x_2^{3/2})}.
\end{aligned}$$

$$\mathrm{F}(0, x_2, 1, 0) = \frac{1}{2\sqrt{x_2}} + \frac{2f'''(0)^2}{9}\sqrt{x_2} + O(x_2^{3/2}). \tag{30}$$

We will also need the difference:

$$\begin{aligned}
\Theta(0, x_2, 1, 0) - \Theta(0, x_2, -1, 0) &= \frac{1}{r(0, x_2, 1, 0)} - \frac{1}{r(0, x_2, -1, 0)}\\
&= \frac{-6f'''(0) + O(x_2)}{4f'''(0)^2 x_2 - 9 + O(x_2^{3/2})}\\
&= \tfrac{2}{3}f'''(0) + O(x_2).
\end{aligned} \tag{31}$$

From (21), using $r_{x_k} = t_{x_k}$, we get

$$r_{x_1}(0, x_2, 1, 0) = -1. \tag{32}$$

Expand the denominator of (22) with respect to t:

$$\begin{aligned}
r_{x_2}(0, x_2, 1, 0) &= \frac{1}{f'(t(0, x_2, 1, 0)}\\
&= \frac{1}{f''(0)t(0, x_2, 1, 0) + \frac{1}{2}f'''(0)t(0, x_2, 1, 0)^2 + O(t(0, x_2, 1, 0)^3)}.
\end{aligned}$$

Using the fact that $f'(0) = 0$ and $f''(0) = \frac{1}{2}$ and substituting the value of t from (28), we obtain

$$r_{x_2}(0, x_2, 1, 0) = \frac{1}{\frac{1}{2}(2\sqrt{x_2} - \frac{4}{3}f'''(0)x_2) + \frac{1}{2}f'''(0)(2\sqrt{x_2} - \frac{4}{3}f'''(0)x_2)^2 + O(x_2^2)}.$$

$r_{x_2}(0, x_2, -1, 0)$ is analogous. Finally,

$$r_{x_2}(0, x_2, 1, 0) = \frac{1}{\sqrt{x_2}} - \frac{4f'''(0)}{3} + \frac{40f'''(0)^2}{9}\sqrt{x_2} + O(x_2). \tag{33}$$

The second derivative has the form $r_{x_k x_l} = t_{x_k x_l}$. From (23) we obtain

$$r_{x_1 x_1}(0, x_2, 1, 0) = r_{x_1 x_2}(0, x_2, 1, 0) = 0. \tag{34}$$

And (24) implies

$$r_{x_2 x_2}(0, x_2, 1, 0) = -\frac{f''(t(0, x_2, 1, 0))}{f'(t(0, x_2, 1, 0))^3}.$$

We now expand the numerator and denominator in a series with respect to t and use $f'(0)=0$, $f''(0)=\frac{1}{2}$, and (28):

$$\begin{aligned} r_{x_2x_2}&(0,x_2,1,0)\\ &=-\frac{\frac{1}{2}+f'''(0)t(0,x_2,1,0)+\frac{1}{2}f^{(4)}(0)t(0,x_2,1,0)^2+O(t^3)}{\big(f''(0)t(0,x_2,1,0)+\frac{1}{2}f'''(0)t(0,x_2,1,0)^2+O(t(0,x_2,1,0)^3)\big)^3}\\ &=\frac{-\frac{1}{2}-2f'''(0)\sqrt{x_2}+(\frac{4}{3}f'''(0)^2-4f^{(4)}(0))x_2+\frac{16}{3}f'''(0)f^{(4)}(0)x_2^{3/2}+O(x_2^2)}{x_2^{3/2}+4f'''(0)x_2^2+O(x_2^{5/2})}\end{aligned}$$

Thus

$$r_{x_2x_2}(0,x_2,1,0)=-\frac{1}{2x_2^{3/2}}-\frac{2f^{(4)}(0)}{\sqrt{x_2}}+O(1). \tag{35}$$

From (19), (29), (32) we find that

$$\Theta_{x_1}(0,x_2,1,0)=\frac{1}{\big(2\sqrt{x_2}-\frac{4}{3}f'''(0)x_2+O(x_2^2)\big)^2}.$$

Analogously, acting for the vector $(-1,0)$, we get

$$\Theta_{x_1}(0,x_2,\pm1,0)=\pm\frac{1}{4x_2}+\frac{f'''(0)}{3\sqrt{x_2}}+O(1). \tag{36}$$

From (29) and (33) we deduce

$$\Theta_{x_2}(0,x_2,1,0)=-\frac{1/\sqrt{x_2}-4f'''(0)/3+(40f'''(0)/9)\sqrt{x_2}+O(x_2)}{\big(2\sqrt{x_2}-\frac{4}{3}f'''(0)x_2+O(x_2^2)\big)^2},$$

and finally,

$$\Theta_{x_2}(0,x_2,\pm1,0)=-\frac{1}{4x_2^{3/2}}-\frac{f'''(0)^2}{\sqrt{x_2}}+O(1). \tag{37}$$

Using the formulae (20), (29), (33), and (35), we obtain the expression for the second derivatives of the Funk metric:

$$\Theta_{x_2x_2}(0,x_2,\pm1,0)=\frac{3}{8x_2^{5/2}}+\frac{13f'''(0)^2+3f^{(4)}(0)}{6x_2^{3/2}}+O\Big(\frac{1}{x_2}\Big). \tag{38}$$

Finally we can estimate the metric coefficients. From (13), (29), and (36) we get

$$\frac{\Theta_{x_1}(0,x_2,1,0)}{\Theta(0,x_2,1,0)}-\frac{\Theta_{x_1}(0,x_2,-1,0)}{\Theta(0,x_2,-1,0)}=\frac{1}{\sqrt{x_2}}+\frac{4f'''(0)^2}{9}\sqrt{x_2}+O(x_2^{3/2}). \tag{39}$$

It follows from (13), (29), and (37) that

$$\frac{\Theta_{x_2}(0,x_2,1,0)}{\Theta(0,x_2,1,0)}-\frac{\Theta_{x_2}(0,x_2,-1,0)}{\Theta(0,x_2,-1,0)}=\frac{2f'''(0)}{3\sqrt{x_2}}+O(1). \tag{40}$$

Note that

$$\begin{aligned}\Theta_{x_1x_1}(0,x_2,\pm1,0)\Theta(0,x_2,\pm1,0)-2\Theta_{x_1}(0,x_2,\pm1,0)\Theta_{x_1}(0,x_2,\pm1,0)\\ =(2\Theta^3 r_{x_1}r_{x_1}-\Theta^2 r_{x_1x_1})\Theta-2\Theta^2 r_{x_1}\Theta^2 r_{x_1}=0,\end{aligned} \tag{41}$$

since $r_{x_1x_1}=0$, and analogously

$$\begin{aligned}\Theta_{x_1x_2}(0,x_2,\pm1,0)\Theta(0,x_2,\pm1,0)-2\Theta_{x_1}(0,x_2,\pm1,0)\Theta_{x_2}(0,x_2,\pm1,0)\\ =0.\end{aligned} \tag{42}$$

Then from (13), (29), (37), and (38) we get

$$\begin{aligned}\frac{\Theta_{x_2x_2}(0,x_2,\pm1,0)\Theta(0,x_2,\pm1,0)-2\Theta_{x_2}(0,x_2,\pm1,0)\Theta_{x_2}(0,x_2,\pm1,0)}{\Theta(0,x_2,\pm1,0)^3}\\ =\frac{1}{2x_2^{3/2}}+\frac{2f^{(4)}(0)}{\sqrt{x_2}}+O(1).\end{aligned} \tag{43}$$

Finally, using (30), (12), (39), (40), (41), (42), and (43), we obtain the series expansions of the metric tensor of the Hilbert metric.

$$\begin{aligned}g_{11}(0,x_2,1,0)&=\frac{1}{4x_2}+O(1),\\ g_{12}(0,x_2,1,0)&=\frac{f'''(0)}{6x_2}+O(1),\\ g_{22}(0,x_2,1,0)&=\frac{1}{4x_2^2}+\frac{2f'''(0)^2+9f^{(4)}(0)}{18x_2}+O(1).\end{aligned} \tag{44}$$

3.2. *Expansions for $g_{ij}(0,x_2,0,1)$.* The formulae in (16) imply that, at $(0,x_2)$,

$$\begin{aligned}t_{x_1}(0,x_2,0,\pm1)&=0,\\ t_{x_2}(0,x_2,0,\pm1)&=-1,\\ t_{x_1x_2}(0,x_2,0,\pm1)&=t_{x_2x_2}(0,x_2,0,\pm1)=0.\end{aligned}$$

Note that the functions $t(0,x_2,0,\pm1)$ have the representations

$$t(0,x_2,0,-1)=-x_2,\quad t(0,x_2,0,1)=H-x_2.$$

Here H denotes the length of the chord of ∂U in the direction $(0,1)$. Then

$$\Theta(0,x_2,0,-1)=\frac{1}{x_2},\quad \Theta(0,x_2,0,1)=\frac{1}{H-x_2}.$$

Consequently,

$$\mathrm{F}(0, x_2, 0, 1) = \tfrac{1}{2}\Big(\frac{1}{H-x_2} + \frac{1}{x_2}\Big) = \frac{1}{2x_2} + O(1). \tag{45}$$

We can estimate the derivatives of the Funk metrics $\Theta(0, x_2, 0, \pm 1)$. It follows from (19) and (20) that

$$\Theta_{x_2}(0, x_2, 0, -1) = \frac{1}{x_2^2}, \qquad \Theta_{x_2}(0, x_2, 0, 1) = -\frac{1}{(H-x_2)^2}, \tag{46}$$

$$\Theta_{x_2 x_2}(0, x_2, 0, -1) = \frac{2}{x_2^3}, \quad \Theta_{x_2 x_2}(0, x_2, 0, 1) = \frac{2}{(H-x_2)^3}. \tag{47}$$

Using (12), (46), and (47), we get the expansions:

$$\begin{aligned} &g_{12}(0, x_2, 0, 1) = 0, \\ &g_{22}(0, x_2, 0, 1) = \tfrac{1}{4}\Big(\frac{1}{H-x_2} + \frac{1}{x_2}\Big)^2 = \frac{1}{4x_2^2} + O\Big(\frac{1}{x_2}\Big). \end{aligned} \tag{48}$$

We will also need the values $\mathrm{F}(0, x_2, l, \frac{1}{2})$.
We have

$$t\big(0, x_2, -l, -\tfrac{1}{2}\big) = -2x_2 + 2l^2 x_2^2 + O(x_2^3),$$
$$t\big(0, x_2, l, \tfrac{1}{2}\big) = L + O(x_2).$$

Then

$$\begin{aligned} \mathrm{F}\big(0, x_2, l, \tfrac{1}{2}\big) = & \frac{\sqrt{\frac{1}{4} + l^2}}{2\sqrt{\frac{1}{4}t(0, x_2, l, \frac{1}{2})^2 + (lt(0, x_2, l, \frac{1}{2}))^2}} \\ & + \frac{\sqrt{\frac{1}{4} + l^2}}{2\sqrt{\frac{1}{4}t(0, x_2, -l, -\frac{1}{2})^2 + (lt(0, x_2, -l, -\frac{1}{2}))^2}} \\ = & \frac{\sqrt{\frac{1}{4} + l^2}}{2\sqrt{\frac{1}{4} + l^2}}\Big(\frac{1}{t(0, x_2, l, \frac{1}{2})} - \frac{1}{t(0, x_2, -l, -\frac{1}{2})}\Big). \end{aligned}$$

Finally,

$$\mathrm{F}\big(0, x_2, l, \tfrac{1}{2}\big) = \frac{1}{4x_2} + \frac{1}{2L} + O(x_2). \tag{49}$$

4. Proof of the theorems

The Chern–Rund covariant derivative along the curve $c(t)$ in the Finsler space equipped with the Hilbert metric F is given by the formula [Shen 2001]

$$\nabla_{c'(t)} c'(t) = \{c''(t)^i + (\Theta(c(t), c'(t)) - \Theta(c(t), -c'(t))c'(t)^i\}\frac{\partial}{\partial x^i}. \tag{50}$$

For calculating the normal curvature (2), the Finsler curvature (3), and the Rund curvature (4), we need the covariant derivative $\nabla_{\dot{c}(s)}\dot{c}(s)$ of the curve $c(s)$ parametrized by its arc length.

For a given curve $c(t)$, we will denote by the dot the derivative with respect to the arc length s, and by the prime the derivative with respect to t. Then let $t = t(s)$ be the reparametrization. We get

$$\dot{c}(s) = c'(t)t'_s.$$

Using that s in the length parameter, we get

$$1 = \mathrm{F}(c(t), c'(t))t'_s.$$

Hence

$$\dot{c}(s) = \frac{c'(t)}{\mathrm{F}(c(t), c'(t))}.$$

The next step is to calculate $\nabla_{\dot{c}(s)}\dot{c}(s)$.

$$\begin{aligned}\nabla_{\dot{c}(s)}\dot{c}(s) &= \nabla_{c'(t)/\mathrm{F}(c(t),c'(t))}\frac{c'(t)}{\mathrm{F}(c(t), c'(t))}\\ &= \frac{1}{\mathrm{F}(c(t), c'(t))}\left(\nabla_{c'(t)}\left(\frac{1}{\mathrm{F}(c(t), c'(t))}\right)c'(t) + \frac{1}{\mathrm{F}(c(t), c'(t))}\nabla_{c'(t)}c'(t)\right).\end{aligned}$$

According to [Bao et al. 2000],

$$\nabla_{c'(t)}\left(\frac{1}{\mathrm{F}(c(t), c'(t))}\right) = -\frac{\boldsymbol{g}_{c'(t)}(\nabla_{c'(t)}c'(t), c'(t))}{\mathrm{F}(c(t), c'(t))^3}.$$

Then the derivative $\nabla_{\dot{c}(s)}\dot{c}(s)$ has the form

$$\nabla_{\dot{c}(s)}\dot{c}(s) = \frac{1}{\mathrm{F}(c(t), c'(t))^2}\left(\nabla_{c'(t)}c'(t) - \frac{\boldsymbol{g}_{c'(t)}(\nabla_{c'(t)}c'(t), c'(t))}{\mathrm{F}(c(t), c'(t))^2}c'(t)\right).$$

Finally, using (50), we get the formula:

$$\nabla_{\dot{c}(s)}\dot{c}(s) = \frac{c''(t) + c'(t)\left(\Theta(c(t), c'(t)) - \Theta(c(t), -c'(t)) - \dfrac{\boldsymbol{g}_{c'(t)}(\nabla_{c'(t)}c'(t), c'(t))}{\mathrm{F}(c(t), c'(t))^2}\right)}{\mathrm{F}(c(t), c'(t))^2}. \tag{51}$$

As in Section 2 fix a point o in the domain U and a point $p \in \partial U$. The curve ∂U admits the polar representation $\omega(\varphi)$ from the point o such that the point p corresponds to $\varphi = 0$. According to Section 2, we assume that U satisfies the conditions (1)–(4).

Then one can get that $\omega'(0)=0$, $\omega(0)=1$, $\omega''(0)=\frac{1}{2}$, $\omega'(\pi)=0$. Set

$$C=\frac{1+\omega(\pi)}{\omega(\pi)}.$$

In [Borisenko and Olin 2008] the polar function $\rho_r(u)$ of the hypersphere of radius r was obtained:

$$\rho_r(u)=\frac{\omega(-u)\omega(u)(e^{2r}-1)}{\omega(u)+\omega(-u)e^{2r}}. \tag{52}$$

As $r\to\infty$,

$$\omega(u)-\rho_r(u)=\omega(u)\Big(\frac{\omega(u)}{\omega(-u)}+1\Big)e^{-2r}+o(e^{-2r}). \tag{53}$$

From (52) we get that the circle of radius r admits the parametrization

$$c(\varphi)=\left(\frac{\omega(\pi-\varphi)\omega(\varphi)(e^{2r}-1)}{\omega(\varphi)+\omega(\pi-\varphi)e^{2r}}\sin\varphi,\ \frac{\omega(\pi-\varphi)\omega(\varphi)(e^{2r}-1)}{\omega(\varphi)+\omega(\pi-\varphi)e^{2r}}\cos\varphi\right),$$

where $\omega(\varphi)$ is the polar function of ∂U.

Then

$$c'(0)=\frac{\omega(\pi)(e^{2r}-1)}{1+\omega(\pi)e^{2r}}(1,0)=(1-Ce^{-2r}+O(e^{-3r}),0),\quad r\to\infty. \tag{54}$$

The second derivative:

$$c''(0)=\frac{(e^{2r}\omega(\pi)^2(\omega''(0)-1)-\omega(\pi)+\omega''(\pi))(e^{2r}-1)}{(1+e^{2r}\omega(\pi))^2}(0,1),$$

$$c''(0)=\big(0,-\tfrac{1}{2}+O(e^{-2r})\big),\quad r\to\infty. \tag{55}$$

From (53) we get that at the point of the circle the second coordinate is

$$x_2=\omega(0)-\frac{\omega(\pi)\omega(0)(e^{2r}-1)}{\omega(0)+\omega(\pi)e^{2r}}=Ce^{-2r}+O(e^{-3r}). \tag{56}$$

Estimate the derivative $\nabla_{\dot c(0)}\dot c(0)$ using the formulae (51), (31), and (56):

$$\begin{aligned}\Theta(c(0),c'(0))-\Theta(c(0),-c'(0))&=\Theta(0,Ce^{-2r}+O(e^{-3r}),1+O(e^{-2r}),0)\\&\quad-\Theta(0,Ce^{-2r}+O(e^{-3r}),-1+O(e^{-2r}),0)=\tfrac{2}{3}f'''(0)+O(e^{-2r}).\end{aligned}$$

Therefore, formula (50) leads to

$$\begin{aligned}\nabla_{c'(0)}c'(0)&=c''(0)+c'(0)\big(\Theta(c(0),c'(0))-\Theta(c(0),-c'(0))\big)\\&=\big(\tfrac{2}{3}f'''(0),-\tfrac{1}{2}\big)+O(e^{-2r}).\end{aligned} \tag{57}$$

Using (56) and (57) we get

$$\frac{\boldsymbol{g}_{c'(0)}(\nabla_{c'(0)}c'(0),c'(0))}{\mathrm{F}(c(0),c'(0))^2}=\frac{\frac{2}{3}f'''(0)g_{11}-\frac{1}{2}g_{12}}{\mathrm{F}(0,Ce^{-2r}+O(e^{-3r}),1+O(e^{-2r}),0)^2}.$$

Here g_{ij} are calculated at the point $(0,Ce^{-2r}+O(e^{-3r}),1+O(e^{-2r}),0)$. After substituting the values from (30) and (44), we obtain

$$\frac{\boldsymbol{g}_{c'(0)}(\nabla_{c'(0)}c'(0),c'(0))}{\mathrm{F}(c(0),c'(0))^2}=-\frac{f'''(0)}{3}+O(e^{-2r}).$$

Therefore,

$$\nabla_{\dot c(0)}\dot c(0)=\frac{(f'''(0),-\frac{1}{2})+(1,1)O(e^{-2r})}{\mathrm{F}(c(0),c'(0))^2}. \tag{58}$$

Taking into account (30),

$$\nabla_{\dot c(0)}\dot c(0)=(4f'''(0),-2)e^{-2r}+(1,1)O(e^{-3r}).$$

Calculate the Rund curvature (4) using the formulae (56) and (58).

$$\begin{aligned}\boldsymbol{k}_R(r)^2&=\mathrm{F}(c(0),\nabla_{\dot c(0)}\dot c(0))\\&=\frac{\mathrm{F}(0,Ce^{-2r}+O(e^{-3r}),-f'''(0)+O(e^{-2r}),\frac{1}{2}+O(e^{-2r}))}{\mathrm{F}(0,Ce^{-2r}+O(e^{-3r}),1-Ce^{-2r}+O(e^{-3r}),0)^2}.\end{aligned}$$

From (30) and (49) we get

$$\boldsymbol{k}_R(r)^2=1+C\left(\frac{2}{L}-\frac{8f'''(0)^2}{9}\right)e^{-2r}+O(e^{-3r}). \tag{59}$$

Here $L>0$ is the length of the chord ℓ of ∂U in the direction $(f'''(0),-1/2)$. Proposition 2.2 gives the uniform bounds on the curvature of ∂U. Proposition 2.3 claims that the angle between the chord ℓ and x_2 is uniformly separated from $\pi/2$. Thus we conclude that $2/L$ is bounded from above.

Calculate the Finsler curvature (3) using the formulae (56) and (58).

$$\begin{aligned}\boldsymbol{k}_F(r)^2&=\boldsymbol{g}_{\dot c(0)}(\nabla_{\dot c(0)}\dot c(0),\nabla_{\dot c(0)}\dot c(0))\\&=\frac{f'''(0)^2g_{11}-f'''(0)g_{12}+\frac{1}{4}g_{22}}{\mathrm{F}(0,Ce^{-2r}+O(e^{-3r}),1-Ce^{-2r}+O(e^{-3r}),0)^4}.\end{aligned}$$

Here g_{ij} are considered at the point $(0,Ce^{-2r}+O(e^{-3r}),1+O(e^{-2r}),0)$. Finally, from (30) and (44) we obtain that

$$\boldsymbol{k}_F(r)^2=1+C\big(-\tfrac{8}{9}f'''(0)^2+4f^{(4)}(0)\big)e^{-2r}+O(e^{-3r}). \tag{60}$$

Proposition 2.3 gives the uniform bounds on the derivatives of f. Theorem 1.1 is proved.

Note that the normal curvature $\boldsymbol{g_n}(\nabla_{\dot c(s)}\dot c(s), \boldsymbol{n})$ of a hypersurface at the point x depends only on the tangent vector to the curve $c(s)$ at x [Shen 2001]. So in order to obtain the normal curvature of the Hilbert hypersphere S_r centered at o at the point p in the tangent direction w, we consider the normal curvature of the circle $S_r \cap \Pi$ which lies in the plane $\Pi = \operatorname{span}(w, \overrightarrow{op})$.

From (57) we get the normal curvature of the circle of radius r:

$$\boldsymbol{k_n}(r) = \boldsymbol{g_n}(\nabla_{\dot c(0)}\dot c(0), \boldsymbol{n}) = \frac{\boldsymbol{g_n}(c''(0), \boldsymbol{n})}{\mathrm{F}(c(0), c'(0))^2}. \tag{61}$$

Since $g_{12}(0, x_2, 0, 1) = 0$ by (48), it follows that the unit normal vector $\boldsymbol{n}$ to the circle at $(0, x_2)$ is exactly

$$\frac{1}{\mathrm{F}(0, x_2, 0, 1)}(0, -1).$$

Finally, taking into account (30), (56), (55), (45), and (48):

$$\begin{aligned}\boldsymbol{k_n}(r) &= \frac{\frac{1}{2}g_{22}(0, Ce^{-2r} + O(e^{-3r}), 0, 1)}{\mathrm{F}(0, Ce^{-2r}+O(e^{-3r}), 1-Ce^{-2r}+O(e^{-3r}), 0)^2\mathrm{F}(0, Ce^{-2r}+O(e^{-3r}), 0, 1)}\\ &= 1 + C\Big(\frac{1}{H} - \frac{8f'''(0)^2}{9}\Big)e^{-2r} + O(e^{-3r}).\end{aligned} \tag{62}$$

If it is the case that the Euclidean normal curvatures of the hypersurface ∂U are bounded ($k_2 \leqslant k_n \leqslant k_1$) then the curvature of the curve $\partial U' = \partial U \cap \Pi$ is bounded as well. Consider the point $x \in \partial U' \subset \partial U$. Then the curvature $k(x)$ of $\partial U'$ and the normal curvature $k_n(x)$ of ∂U are related as

$$k(x) = \frac{k_n(x)}{\cos\beta}.$$

Here β is the angle between the radial and normal direction to ∂U at x. Using Lemma 2.1 we find that $\omega_0 k_2 \leqslant \cos\beta \leqslant 1$. Hence the curvature of $\partial U'$ is uniformly bounded for all y. Applying Proposition 2.2 for the Hilbert geometry based on U', we get the uniformity of the series expansion (62) which ends the proof of Theorem 1.2.

References

[Bao et al. 2000] D. Bao, S.-S. Chern, and Z. Shen, *An introduction to Riemann–Finsler geometry*, Graduate Texts in Mathematics **200**, Springer, New York, 2000. MR 2001g:53130 Zbl 0954.53001

[Borisenko 2002] A. A. Borisenko, "Convex sets in Hadamard manifolds", *Differential Geom. Appl.* **17**:2-3 (2002), 111–121. MR 2003k:53035 Zbl 1034.53056

[Borisenko and Olin 2008] A. A. Borisenko and E. A. Olin, "Asymptotic properties of Hilbert geometry", *Zh. Mat. Fiz. Anal. Geom.* **4**:3 (2008), 327–345, 443. MR 2010e:53124 Zbl 1167.53066

[Colbois and Verovic 2002] B. Colbois and P. Verovic, "Rigidity of Hilbert metrics", *Bull. Austral. Math. Soc.* **65**:1 (2002), 23–34. MR 2003c:53053 Zbl 1007.53055

[Finsler 1951] P. Finsler, *Über Kurven und Flächen in allgemeinen Räumen*, Birkhäuser, Basel, 1951. MR 13,74d Zbl 0044.37003

[Rund 1959] H. Rund, *The differential geometry of Finsler spaces*, Die Grundlehren der Mathematischen Wissenschaften **101**, Springer, Berlin, 1959. MR 21 #4462 Zbl 0087.36604

[Shen 2001] Z. Shen, *Lectures on Finsler geometry*, World Scientific, Singapore, 2001. MR 2002f:53032 Zbl 0974.53002

Received May 10, 2011.

ALEXANDER BORISENKO
SCHOOL OF MATHEMATICS AND MECHANICAL ENGINEERING, GEOMETRY DEPARTMENT
KHARKOV NATIONAL UNIVERSITY
SVOBODY SQUARE, 4
KHARKOV 61022
UKRAINE
borisenk@univer.kharkov.ua

EUGENE OLIN
SCHOOL OF MATHEMATICS AND MECHANICAL ENGINEERING, GEOMETRY DEPARTMENT
KHARKOV NATIONAL UNIVERSITY
SVOBODY SQUARE, 4
KHARKOV 61022
UKRAINE
evolin@mail.ru
http://geometry.univer.kharkov.ua/~olin/olin.html

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

A FORMULA EQUATING OPEN AND CLOSED GROMOV–WITTEN INVARIANTS AND ITS APPLICATIONS TO MIRROR SYMMETRY

KWOKWAI CHAN

We prove that open Gromov–Witten invariants for semi-Fano toric manifolds of the form $X = \mathbb{P}(K_Y \oplus \mathcal{O}_Y)$, where Y is a toric Fano manifold, are equal to certain 1-pointed closed Gromov–Witten invariants of X. As applications, we compute the mirror superpotentials for these manifolds. In particular, this gives a simple proof for the formula of the mirror superpotential for the Hirzebruch surface F_2.

1. Introduction

Let X be a compact complex n-dimensional toric manifold equipped with a toric Kähler structure ω. Let L be a Lagrangian torus fiber of the moment map associated to (X, ω). Fukaya, Oh, Ohta and Ono define open Gromov–Witten invariants for (X, L) as follows (see [Fukaya et al. 2010a]). Let $\beta \in \pi_2(X, L)$ be a relative homotopy class with Maslov index $\mu(\beta) = 2$. Let $\mathcal{M}_1(L, \beta)$ be the moduli space of holomorphic disks in X with boundaries lying in L and with one boundary marked point representing the class β. A compactification of $\mathcal{M}_1(L, \beta)$ is given by the moduli space $\overline{\mathcal{M}}_1(L, \beta)$ of stable maps from genus 0 bordered Riemann surfaces $(\Sigma, \partial\Sigma)$ to (X, L) with one boundary marked point representing the class β. As shown in their monumental work [Fukaya et al. 2009] by the same authors, $\overline{\mathcal{M}}_1(L, \beta)$ is a Kuranishi space with real virtual dimension n. By Corollary 11.5 in [Fukaya et al. 2010a], there exists a virtual fundamental n-cycle $[\overline{\mathcal{M}}_1(L, \beta)]^{\text{vir}}$. The pushforward of this cycle by the evaluation map $\text{ev} : \overline{\mathcal{M}}_1(L, \beta) \to L$ at the boundary marked point then gives

$$c_\beta = \text{ev}_*\big([\overline{\mathcal{M}}_1(L, \beta)]^{\text{vir}}\big) \in H_n(L, \mathbb{Q}) \cong \mathbb{Q}.$$

By Lemma 11.7 in the same reference, the homology class c_β is independent of the perturbation data (transversal multisections) used to define $[\overline{\mathcal{M}}_1(L, \beta)]^{\text{vir}}$. Hence, c_β is an open Gromov–Witten invariant for (X, L).

MSC2010: primary 53D45, 53D37; secondary 53D12, 14M25, 14N35.
Keywords: Open Gromov–Witten invariants, semi-Fano, toric manifolds, mirror symmetry, Landau–Ginzburg model, superpotential.

Let Y be an $(n-1)$-dimensional toric Fano manifold. Consider the $\mathbb{P}^1$-bundle $X=\mathbb{P}(K_Y\oplus\mathbb{O}_Y)$ over Y, where K_Y denotes the anticanonical bundle of Y. Then X is an n-dimensional toric manifold which is semi-Fano, i.e., the anticanonical bundle K_X^{-1} is nef. Let $h\in H_2(X,\mathbb{Z})$ be the fiber class. Let $\alpha\in H_2(X,\mathbb{Z})$ be an effective class with $c_1(\alpha)=c_1(X)\cdot\alpha=0$. Consider the moduli space $\overline{\mathcal{M}}_{0,1}(X,h+\alpha)$ of genus-0 stable maps to X with one marked point representing the class $h+\alpha$.[1] By [Fukaya and Ono 1999], $\overline{\mathcal{M}}_{0,1}(X,h+\alpha)$ is a Kuranishi space with complex virtual dimension n. The pushforward of the virtual fundamental cycle $[\overline{\mathcal{M}}_{0,1}(X,h+\alpha)]^{\mathrm{vir}}$ by the evaluation map $\mathrm{ev}:\overline{\mathcal{M}}_{0,1}(X,h+\alpha)\to X$ gives a 1-pointed closed Gromov–Witten invariant of X:

$$\mathrm{GW}_{0,1}^{X,h+\alpha}\big([\mathrm{pt}]\big)=\mathrm{ev}_*\big([\overline{\mathcal{M}}_{0,1}(X,h+\alpha)]^{\mathrm{vir}}\big)\in H_{2n}(X,\mathbb{Q})\cong\mathbb{Q},$$

where $[\mathrm{pt}]$ denotes the Poincaré dual of a point in X.

Now let $\iota_0:Y\hookrightarrow X$ be the closed embedding of Y as the zero section of K_Y. The image is a toric prime divisor $D_0=\iota_0(Y)\subset X$. As above, we equip X with a toric Kähler structure ω and fix a Lagrangian torus fiber L in X. Corresponding to D_0 is a relative homotopy class $\beta_0\in\pi_2(X,L)$. More precisely, $\beta_0\in\pi_2(X,L)$ is the class such that $D_i\cdot\beta_0=\delta_{i0}$ for any toric prime divisor D_i in X. The main result of this paper is the following formula.

Theorem 1.1. *For the $\mathbb{P}^1$-bundle $X=\mathbb{P}(K_Y\oplus\mathbb{O}_Y)$ over a toric Fano manifold Y and for any effective class $\alpha\in H_2(X,\mathbb{Z})$ with $c_1(\alpha)=0$, we have*

$$c_{\beta_0+\alpha}=\mathrm{GW}_{0,1}^{X,h+\alpha}\big([\mathrm{pt}]\big).$$

Note that $\beta_0+\alpha\in\pi_2(X,L)$ is a Maslov index two class since $c_1(\alpha)=0$. We will prove this formula in Section 4 by comparing the Kuranishi structures of $\overline{\mathcal{M}}_1(L,\beta_0+\alpha)$ and $\overline{\mathcal{M}}_{0,1}(X,h+\alpha)$.

We can apply this formula to study mirror symmetry. Recall that the mirror of a compact toric n-fold X is given by a Landau–Ginzburg model $(X^\vee,W)$ consisting of a bounded domain $X^\vee\subset(\mathbb{C}^*)^n$ and a holomorphic function $W:X^\vee\to\mathbb{C}$ called the mirror superpotential. In [Fukaya et al. 2010a] (see also [Cho and Oh 2006; Auroux 2007; 2009; Chan and Leung 2010a; 2010b]), it is shown that the mirror superpotential can be expressed as a power series whose coefficients are the open Gromov–Witten invariants defined above. However, when X is non-Fano, these invariants are in general very hard to compute. The only known examples are the mirror superpotentials for the Hirzebruch surfaces F_2 and F_3, first computed in [Auroux 2009] using degeneration methods and wall-crossing formulas. Fukaya, Oh, Ohta and Ono gave a different proof for the F_2 case in [Fukaya et al. 2010b].

[1]In $\overline{\mathcal{M}}_{0,1}$, subscript 0 and 1 denote the genus and number of marked points respectively.

As an immediate application of our formula, we can express the mirror superpotential of $X = \mathbb{P}(K_Y \oplus \mathbb{O}_Y)$ in terms of 1-point closed Gromov–Witten invariants (see Theorem 5.2). In particular, since $F_2 = \mathbb{P}(K_{\mathbb{P}^1} \oplus \mathbb{O}_{\mathbb{P}^1})$ and its Gromov–Witten invariants are easy to compute as it is symplectomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$, this gives a very simple proof of the formula for the mirror superpotential of F_2. See the example in Section 5. Our formula has since then been applied to study mirror symmetry for various classes of toric manifolds. See [Lau et al. 2011; 2010; Chan et al. 2010; Chan and Lau 2010] more details.

The rest of this paper is organized as follows. In Section 2, we briefly review Kuranishi spaces and recall the results that we need in this paper. In Section 3, we establish several preliminary results concerning the toric manifolds $X = \mathbb{P}(K_Y \oplus \mathbb{O}_Y)$. In Section 4 we prove our formula by a direct comparison of Kuranishi structures. In Section 5, we discuss applications of our formula to mirror symmetry.

2. Kuranishi structures

In this section, we briefly review the theory of Kuranishi spaces and recall some of their properties for later use. We follow Appendix A1 in the book [Fukaya et al. 2009]. See also Section 3 in [Fukaya and Ono 1999].

Let $\mathcal{M}$ be a compact metrizable space.

Definition 2.1 [Fukaya et al. 2009, Definitions A1.1, A1.3, A1.5]. A Kuranishi structure on $\mathcal{M}$ of (real) virtual dimension d consists of the following data:

(1) For each point $\sigma \in \mathcal{M}$,
- (1.1) A smooth manifold V_σ (with boundary or corners) and a finite group Γ_σ acting smoothly and effectively on V_σ.
- (1.2) A real vector space E_σ on which Γ_σ has a linear representation and such that $\dim V_\sigma - \dim E_\sigma = d$.
- (1.3) A Γ_σ-equivariant smooth map $s_\sigma : V_\sigma \to E_\sigma$.
- (1.4) A homeomorphism ψ_σ from $s_\sigma^{-1}(0)/\Gamma_\sigma$ onto a neighborhood of σ in $\mathcal{M}$.

(2) For each $\sigma \in \mathcal{M}$ and for each $\tau \in \operatorname{Im} \psi_\sigma$,
- (2.1) A Γ_τ-invariant open subset $V_{\sigma\tau} \subset V_\tau$ containing $\psi_\tau^{-1}(\tau)$.[2]
- (2.2) A homomorphism $h_{\sigma\tau} : \Gamma_\tau \to \Gamma_\sigma$.
- (2.3) An $h_{\sigma\tau}$-equivariant embedding $\varphi_{\sigma\tau} : V_{\sigma\tau} \to V_\sigma$ and an injective $h_{\sigma\tau}$-equivariant bundle map $\hat{\varphi}_{\sigma\tau} : E_\tau \times V_{\sigma\tau} \to E_\sigma \times V_\sigma$ covering $\varphi_{\sigma\tau}$.

Moreover, these data should satisfy the following conditions:

(i) $\hat{\varphi}_{\sigma\tau} \circ s_\tau = s_\sigma \circ \varphi_{\sigma\tau}$.[3]

---

[2]Here and in C2 below, we regard ψ_τ as a map from $s_\tau^{-1}(0)$ to $\mathcal{M}$ by composing with the quotient map $V_\tau \to V_\tau/\Gamma_\tau$.

[3]Here and after, we also regard s_σ as a section $s_\sigma : V_\sigma \to E_\sigma \times V_\sigma$.

(ii) $\psi_\tau = \psi_\sigma \circ \varphi_{\sigma\tau}$.

(iii) If $\xi \in \psi_\tau(s_\tau^{-1}(0) \cap V_{\sigma\tau}/\Gamma_\tau)$ then, in a sufficiently small neighborhood of ξ,

$$\varphi_{\sigma\tau} \circ \varphi_{\tau\xi} = \varphi_{\sigma\xi} \quad \text{and} \quad \hat{\varphi}_{\sigma\tau} \circ \hat{\varphi}_{\tau\xi} = \hat{\varphi}_{\sigma\xi}.$$

The spaces E_σ are called obstruction spaces (or obstruction bundles), the maps $\{s_\sigma : V_\sigma \to E_\sigma\}$ are called Kuranishi maps, and $(V_\sigma, E_\sigma, \Gamma_\sigma, s_\sigma, \psi_\sigma)$ is called a Kuranishi neighborhood of $\sigma \in \mathcal{M}$.

To define virtual fundamental chains, we need Kuranishi spaces with extra structures.

Definition 2.2 [Fukaya et al. 2009, Definitions A1.14, A1.17]. A Kuranishi space is said to have a tangent bundle if the differential of s_σ in the direction of the normal bundle induces a bundle isomorphism

$$ds_\sigma : \frac{\varphi_{\sigma\tau}^* T V_\sigma}{T V_{\sigma\tau}} \cong \frac{E_\sigma \times V_{\sigma\tau}}{\hat{\varphi}_{\sigma\tau}(E_\tau \times V_{\sigma\tau})} \tag{2-1}$$

as Γ_τ-equivariant bundles on $V_{\sigma\tau}$.

For a Kuranishi space with tangent bundle, an orientation consists of trivializations of $\bigwedge^{\mathrm{top}} E_\sigma^* \otimes \bigwedge^{\mathrm{top}} T V_\sigma$ compatible with the isomorphisms (2-1).

We will not give the precise definition of multisections here. See Definitions A1.19, A1.21 in [Fukaya et al. 2009] for details. Roughly speaking, a multisection $\mathfrak{s}$ is a system of multivalued perturbations $\{s'_\sigma : V_\sigma \to E_\sigma\}$ of the Kuranishi maps $\{s_\sigma : V_\sigma \to E_\sigma\}$ satisfying certain compatibility conditions. For a Kuranishi space $\mathcal{M}$ with tangent bundle, there exist (a family of) multisections $\mathfrak{s}$ which are transversal to 0 (Theorem A1.23 in [Fukaya et al. 2009]). Furthermore, suppose that $\mathcal{M}$ is oriented. Let $\mathrm{ev} : \mathcal{M} \to Y$ be a strongly smooth map to a smooth manifold Y, i.e., a family of Γ_σ-invariant smooth maps $\{\mathrm{ev}_\sigma : V_\sigma \to Y\}$ such that $\mathrm{ev}_\sigma \circ \varphi_{\sigma\tau} = \mathrm{ev}_\tau$ on $V_{\sigma\tau}$. Then, using these transversal multisections, one can define the virtual fundamental chain $\mathrm{ev}_*([\mathcal{M}]^{\mathrm{vir}})$ as a $\mathbb{Q}$-singular chain in Y (Definition A1.28 in [Fukaya et al. 2009]).

We will also need the notion of fiber products of Kuranishi spaces. See Appendix A1.2 in [Fukaya et al. 2009] for more details. As before, let $\mathrm{ev} : \mathcal{M} \to Y$ be a strongly smooth map from a Kuranishi space $\mathcal{M}$ to a smooth manifold Y. Suppose that ev is weakly submersive, i.e., each $\mathrm{ev}_\sigma : V_\sigma \to Y$ is a submersion. Let W be another manifold and $g : W \to Y$ be a smooth map. Consider the fiber product

$$\mathcal{X} = \mathcal{M} \times_Y W = \{(\sigma, p) \in \mathcal{M} \times W : \mathrm{ev}(\sigma) = q(p)\}.$$

Definition 2.3 [Fukaya et al. 2009, Definition A1.37]. Let $(\sigma, p) \in \mathcal{X}$ and let $(V_\sigma, E_\sigma, \Gamma_\sigma, s_\sigma, \psi_\sigma)$ be a Kuranishi neighborhood of $\sigma \in \mathcal{M}$. We set

$$V_{(\sigma,p)} = \{(\tau, q) \in V_\sigma \times W : \mathrm{ev}_\sigma(\tau) = g(q)\}.$$

Then $V_{(\sigma,p)}$ is a smooth manifold since ev_σ is a submersion. We also set $E_{(\sigma,p)} = E_\sigma$, $\Gamma_{(\sigma,p)} = \Gamma_\sigma$ and define $s_{(\sigma,p)}$, $\psi_{(\sigma,p)}$ in the obvious way. This defines a Kuranishi neighborhood of $(\sigma, p) \in \mathscr{X}$, and they glue together to give a Kuranishi structure on $\mathscr{X}$.

Lemma 2.4 [Fukaya et al. 2009, Lemma A1.39]. *If the Kuranishi space $\mathcal{M}$ has a tangent bundle, so does the Kuranishi structure on $\mathscr{X}$. If the Kuranishi structure on $\mathcal{M}$ and the manifolds Y, W are all oriented, so is the Kuranishi structure on $\mathscr{X}$.*

Let $\hat{\mathrm{ev}} : \mathscr{X} \to W$ be the projection map. We remark that this is a strongly smooth map. The following lemma is crucial to the proof of our main result.

Lemma 2.5 [Fukaya et al. 2009, Lemma A1.43]. *Suppose that Y and W are oriented and compact without boundary, and $\partial\mathcal{M} = \varnothing$. Then we have*

$$\mathrm{PD}\big(\hat{\mathrm{ev}}_*([\mathscr{X}]^{\mathrm{vir}})\big) = g^*\big(\mathrm{PD}\big(\mathrm{ev}_*([\mathcal{M}]^{\mathrm{vir}})\big)\big),$$

where PD *denotes Poincaré dual.*

3. A class of semi-Fano toric manifolds

Let Y be an $(n-1)$-dimensional toric Fano manifold. Denote by K_Y its canonical line bundle. Consider the $\mathbb{P}^1$-bundle $X = \mathbb{P}(K_Y \oplus \mathcal{O}_Y)$ over Y. In this section, we shall establish some elementary properties of the toric manifold X which will be of use later.

Let $e_1, \dots, e_n$ be the standard basis of a rank n lattice $N \cong \mathbb{Z}^n$, and let

$$N' = \Big\{ v = \sum_{j=1}^{n} v^j e_j \in N \;\Big|\; v^n = 0 \Big\} \cong \mathbb{Z}^{n-1}.$$

Let $N_{\mathbb{R}} = N \otimes_{\mathbb{Z}} \mathbb{R}$ and $N'_{\mathbb{R}} = N' \otimes_{\mathbb{Z}} \mathbb{R}$. Without loss of generality, we can choose the primitive generators of the 1-dimensional cones of the fan Δ in $N_{\mathbb{R}}$ defining X to be

$$v_0 = e_n,\ v_1 = w_1 + e_n,\ \dots,\ v_m = w_m + e_n,\ v_{m+1} = -e_n,$$

where $w_1, \dots, w_m \in N'$ are the primitive generators of the 1-dimensional cones of a fan Δ' in $N'_{\mathbb{R}}$ defining Y.

Lemma 3.1. *Let $h \in H_2(X, \mathbb{Z})$ be the fiber class of the $\mathbb{P}^1$-bundle $X = \mathbb{P}(K_Y \oplus \mathcal{O}_Y)$. Let $\iota_0 : Y \hookrightarrow X = \mathbb{P}(K_Y \oplus \mathcal{O}_Y)$ be the closed embedding which maps Y to the zero section of K_Y. ι_0 induces an embedding $\iota_{0*} : H_2(Y, \mathbb{Z}) \hookrightarrow H_2(X, \mathbb{Z})$. Then we have*

$$H_2^{\mathrm{eff}}(X, \mathbb{Z}) \cong \mathbb{Z}_{\geq 0} h \oplus \iota_{0*} H_2^{\mathrm{eff}}(Y, \mathbb{Z}).$$

*Here, superscript "*eff*" refers to effective classes. Moreover, we have $c_1(h) = 2$ and $c_1(\alpha) = 0$ for any $\alpha \in \iota_{0*} H_2^{\mathrm{eff}}(Y, \mathbb{Z})$. In particular, X is semi-Fano, i.e., the anticanonical bundle K_X^{-1} is nef.*

Proof. Recall that a subset $\mathcal{P} = \{v_{i_1}, \dots, v_{i_p}\} \subset \{v_0, \dots, v_{m+1}\}$ is called a *primitive collection* if for each $1 \le k \le p$, the elements of $\mathcal{P} \setminus \{x_{i_k}\}$ generate a $(p-1)$-dimensional cone in Δ but $\mathcal{P}$ itself does not generate a cone in Δ; see [Batyrev 1991]. The *focus* $\sigma(\mathcal{P})$ of $\mathcal{P}$ is the cone in Δ of the smallest dimension which contains $v_{i_1} + \dots + v_{i_p}$. Let $v_{j_1}, \dots, v_{j_q}$ be the generators of $\sigma(\mathcal{P})$. Then there exists positive integers $n_1, \dots, n_q$ such that

$$v_{i_1} + \dots + v_{i_p} = n_1 v_{j_1} + \dots + n_q v_{j_q}.$$

This is known as a *primitive relation*. Recall that the homology group $H_2(X, \mathbb{Z})$ is given by the kernel of the surjective map $\mathbb{Z}^d \to N,\ E_i \mapsto v_i$, where $\{E_1, \dots, E_d\}$ is the standard basis of $\mathbb{Z}^d$. Also, the effective cone $H_2^{\text{eff}}(X, \mathbb{Z})$ is generated by primitive relations.

In our case, $\mathcal{P}_0 := \{v_0, v_{m+1}\}$ is obviously a primitive collection for Δ. The primitive relation $v_0 + v_{m+1} = 0$ corresponds to the fiber class h of the $\mathbb{P}^1$-bundle $X \to Y$. It is obvious that we have $c_1(h) = c_1(X) \cdot h = 2$.

By Proposition 4.1 in [Batyrev 1991], we have $\mathcal{P} \cap \mathcal{P}_0 = \varnothing$ for any other primitive collection $\mathcal{P} \ne \mathcal{P}_0$. Suppose that $\mathcal{P} \ne \mathcal{P}_0$ is a primitive collection consisting of the elements $v_{i_1} = w_{i_1} + e_n, \dots, v_{i_p} = w_{i_p} + e_n$, where $1 \le i_1 < \dots < i_p \le m$. Then $\mathcal{P}' := \{w_{i_1}, \dots, w_{i_p}\}$ is obviously a primitive collection for the fan Δ' defining Y. Now, let $w_{j_1}, \dots, w_{j_q}$ be the generators of the focus $\sigma(\mathcal{P}')$ of $\mathcal{P}'$. The primitive relation for Δ' is given by

$$w_{i_1} + \dots + w_{i_p} = n_1 w_{j_1} + \dots + n_q w_{j_q}, \tag{3-1}$$

for some $n_1, \dots, n_q \in \mathbb{Z}_{>0}$. Let $\gamma \in H_2^{\text{eff}}(Y, \mathbb{Z})$ be the corresponding effective class. Since Y is Fano, we have $p - n_1 - \dots - n_q = c_1(Y) \cdot \gamma > 0$. In terms of the v_i's, (3-1) becomes a primitive relation

$$v_{i_1} + \dots + v_{i_p} = n_1 v_{j_1} + \dots + n_q v_{j_q} + (p - n_1 - \dots - n_q) v_0$$

for Δ. This corresponds to the class $\alpha := \iota_{0*}(\gamma) \in H_2^{\text{eff}}(X, \mathbb{Z})$, whose Chern number is given by $c_1(\alpha) = p - n_1 - \dots - n_q - (p - n_1 - \dots - n_q) = 0$. □

As usual, denote by $D_0, D_1, \dots, D_m, D_{m+1}$ the toric prime divisors corresponding to the primitive generators $v_0, v_1, \dots, v_m, v_{m+1}$ respectively. Note that $D_0 = \iota_0(Y)$.

Lemma 3.2. *Let $\varphi : \mathbb{P}^1 \to X$ be a nonconstant holomorphic map from $\mathbb{P}^1$ to X.*

(1) *Suppose that $[\varphi(\mathbb{P}^1)] = h + \alpha \in H_2(X, \mathbb{Z})$ for some $\alpha \in \iota_{0*} H_2^{\text{eff}}(Y, \mathbb{Z})$. Then $\varphi(\mathbb{P}^1)$ is contained in one of the toric prime divisors $D_0, D_1, \dots, D_m$.*

(2) *Suppose that $[\varphi(\mathbb{P}^1)] = \alpha \in \iota_{0*} H_2^{\text{eff}}(Y, \mathbb{Z})$. Then $\varphi(\mathbb{P}^1)$ is contained in the toric prime divisor D_0.*

Proof. Suppose that $\varphi : \mathbb{P}^1 \to X$ is a nonconstant holomorphic map with class $h+\alpha$ for some $\alpha \in \iota_{0*} H_2^{\mathrm{eff}}(Y, \mathbb{Z})$. From the proof of the above lemma, we know that the class $h+\alpha$ corresponds to the primitive relation

$$\Bigl(1 - \sum_{i=1}^{m} a_i\Bigr) v_0 + \sum_{i=1}^{m} a_i v_i + v_{m+1} = 0.$$

Moreover, we have $\sum_{i=1}^{m} a_i \geq 1$, and if $\sum_{i=1}^{m} a_i = 1$, then there exists $1 \leq i \leq m$ such that $a_i < 0$. Hence there exists $0 \leq i \leq m$ such that $D_i \cdot \varphi(\mathbb{P}^1) = D_i \cdot (h+\alpha) < 0$. This implies that $\varphi(\mathbb{P}^1)$ is contained in D_i. This proves (1). (2) can be proved in the same way. □

4. Proof of Theorem 1.1

We equip $X = \mathbb{P}(K_Y \oplus \mathbb{O}_Y)$ with a toric Kähler structure ω. Let $L \subset X$ be a Lagrangian torus fiber of the associated moment map. For $i = 0, 1, \ldots, m, m+1$, let $\beta_i \in \pi_2(X, L)$ be the relative homotopy class such that $D_j \cdot \beta_i = \delta_{ij}$. Then each β_i is a Maslov index two class with $\partial\beta_i = v_i$, where $\partial : \pi_2(X, L) \to \pi_1(L)$ is the boundary map, and $\pi_2(X, L)$ is generated by $\beta_0, \beta_1, \ldots, \beta_m, \beta_{m+1}$. Moreover, each β_i is represented by a family of holomorphic disks $\varphi_i : (D^2, \partial D^2) \to (X, L)$. Here, $D^2 = \{z \in \mathbb{C} : |z| \leq 1\}$ is the unit disk.

Fix a nonzero effective class $\alpha \in H_2^{\mathrm{eff}}(X, \mathbb{Z})$ with $c_1(\alpha) = 0$. Let $\overline{\mathcal{M}}_1(L, \beta_0+\alpha)$ be the moduli space of stable maps from genus 0 bordered Riemann surfaces to (X, L) with one boundary marked point representing the class $\beta_0+\alpha$. To simplify notation, we denote $\overline{\mathcal{M}}_1(L, \beta_0+\alpha)$ by $\mathcal{M}^L$. Similarly, we denote by $\mathcal{M}^X$ the moduli space $\overline{\mathcal{M}}_{0,1}(X, h+\alpha)$ of genus 0 stable maps to X with one marked point representing the class $h+\alpha$. We have evaluation maps[4]

$$\mathrm{ev} : \mathcal{M}^L \to L \quad \text{and} \quad \mathrm{ev} : \mathcal{M}^X \to X.$$

By [Fukaya et al. 2009], both $\mathcal{M}^L$ and $\mathcal{M}^X$ are oriented Kuranishi spaces with tangent bundles, and the evaluation maps are both strongly smooth and weakly submersive. The real virtual dimensions of $\mathcal{M}^L$ and $\mathcal{M}^X$ are n and $2n$ respectively. Moreover, since $\mu(\beta_0+\alpha) = 2$, we have $\partial\mathcal{M}^L = \varnothing$ by Corollary 11.5 in [Fukaya et al. 2010a]. It is also well-known that $\mathcal{M}^X$ has no boundary. Hence, they define virtual fundamental cycles

$$\mathrm{ev}_*\bigl([\mathcal{M}^L]^{\mathrm{vir}}\bigr) \in H_n(L, \mathbb{Q}) \quad \text{and} \quad \mathrm{ev}_*\bigl([\mathcal{M}^X]^{\mathrm{vir}}\bigr) \in H_{2n}(X, \mathbb{Q}).$$

Fix a point $p \in L \subset X$. Let $\iota : \{p\} \hookrightarrow L$ (resp. $\iota : \{p\} \hookrightarrow X$) be the inclusion of the point p. We can then apply the construction in Definition 2.3 and Lemma 2.4

[4]By a slight abuse of notation, we use ev to denote both evaluation maps. It should clear from the context which one we are referring to.

to give oriented Kuranishi structures with tangent bundles on the spaces:

$$\mathcal{M}_p^L := \mathcal{M}^L \times_L \{p\},\ \mathcal{M}_p^X := \mathcal{M}^X \times_X \{p\}.$$

Both have real virtual dimension 0. Let $\hat{\text{ev}} : \mathcal{M}_p^L \to \{p\}$, $\hat{\text{ev}} : \mathcal{M}_p^X \to \{p\}$ be the induced (constant) maps. Then we have virtual fundamental cycles

$$\hat{\text{ev}}_*\big([\mathcal{M}_p^L]^{\text{vir}}\big),\ \hat{\text{ev}}_*\big([\mathcal{M}_p^X]^{\text{vir}}\big) \in H_0(\{p\}, \mathbb{Q}) \cong \mathbb{Q}.$$

Now Lemma 2.5 says that

Proposition 4.1. *We have*

$$\text{PD}\big(\hat{\text{ev}}_*\big([\mathcal{M}_p^L]^{\text{vir}}\big)\big) = \iota^*\,\text{PD}\big(\text{ev}_*\big([\mathcal{M}^L]^{\text{vir}}\big)\big),$$
$$\text{PD}\big(\hat{\text{ev}}_*\big([\mathcal{M}_p^X]^{\text{vir}}\big)\big) = \iota^*\,\text{PD}\big(\text{ev}_*\big([\mathcal{M}^X]^{\text{vir}}\big)\big)$$

in $H^0(\{p\}, \mathbb{Q}) \cong \mathbb{Q}$.

Therefore, to prove Theorem 1.1, it suffices to show that $\mathcal{M}_p^L$ and $\mathcal{M}_p^X$ have the same Kuranishi structures.

To do this, we first show that $\mathcal{M}_p^L$ can naturally be identified with $\mathcal{M}_p^X$ as a set. Let us recall the following results proved in [Cho and Oh 2006], which holds for general toric manifolds.

Theorem 4.2 ([Cho and Oh 2006, Theorem 5.2]; see also [Fukaya et al. 2010a, Theorem 11.1]). *Let* (X, ω) *be a toric Kähler manifold and* L *be a Lagrangian torus fiber of its moment map. Let* $D_1, \dots, D_d$ *be all the toric prime divisors in* X *and* $\beta_1, \dots, \beta_d \in \pi_2(X, L)$ *be the relative homotopy classes such that* $D_j \cdot \beta_i = \delta_{ij}$.

(1) *If* $\varphi : (D^2, \partial D^2) \to (X, L)$ *is a holomorphic map from a disk representing a Maslov index two class* $\beta \in \pi_2(X, L)$, *then* $\beta = \beta_i$ *for some* $i \in \{1, \dots, d\}$.

(2) *For* $i = 1, \dots, d$, *let* $\overline{\mathcal{M}}_1(L, \beta_i)$ *be the moduli space of stable maps from genus* 0 *bordered Riemann surfaces to* (X, L) *with one boundary marked point representing the class* β_i. *Then the evaluation map* ev : $\overline{\mathcal{M}}_1(L, \beta_i) \to L$ *is an orientation preserving diffeomorphism. In particular, for any* $p \in L$ *and any* $i \in \{1, \dots, d\}$, *there is a unique* (*up to automorphisms of the domain*) *genus* 0 *bordered stable map whose boundary passes through* p *and whose domain is a disk which represents the class* β_i.

Now, let $\sigma^L = ((\Sigma^L, z), \varphi)$ represent a point in $\mathcal{M}_p^L$. This consists of a genus 0 bordered Riemann surface Σ^L with a boundary marked point $z \in \partial\Sigma^L$ and a stable map $\varphi : (\Sigma^L, \partial\Sigma^L) \to (X, L)$ such that $\varphi(z) = p$.

Proposition 4.3. Σ^L *can be decomposed as* $\Sigma^L = \Sigma_0^L \cup \Sigma_1$, *where* $\Sigma_0^L = D^2$ *is a disk and* Σ_1 *is a genus* 0 *nodal curve, such that the restrictions* $\varphi_0 := \varphi|_{\Sigma_0^L}$ *and* $\varphi_1 := \varphi|_{\Sigma_1}$ *represent the classes* β_0 *and* α *respectively.*

Proof. The Maslov index of $\beta_0 + \alpha$ is $\mu(\beta_0 + \alpha) = 2$ since $c_1(\alpha) = 0$. By Theorem 4.2 (1), there does not exist any nonconstant holomorphic map from a disk to (X, L) with class $\beta_0+\alpha$, so Σ^L must be singular. Decompose Σ^L into irreducible components. Let $\varphi_j : (D^2, \partial D^2) \to (X, L)$ and $\varphi_k : \mathbb{P}^1 \to X$ be the restriction of φ to the disk and sphere components respectively. Then $\beta_0+\alpha = \sum_j[\varphi_j]+\sum_k[\varphi_k]$. Notice that, by the proof of Lemma 3.1, any $\alpha \in H_2(X, \mathbb{Z})$ with $c_1(\alpha) = 0$ cannot be expressed as a $\mathbb{Z}$-linear combination of β_i's with positive coefficients. Hence, there must be only one disk component in Σ. Therefore, we can decompose Σ into $\Sigma_0^L \cup \Sigma_1$, where $\Sigma_0^L = D^2$ is a disk and Σ_1 is a genus 0 nodal curve (i.e., a tree of $\mathbb{P}^1$'s). Now, the restriction $\varphi_0 := \varphi|_{\Sigma_0^L}$ is a nonconstant holomorphic map from $(D^2, \partial D^2)$ to (X, L). By Theorem 4.2 (1) again, the class of φ_0 must be β_0. Hence $\varphi_1 := \varphi|_{\Sigma_1}$ represents α. □

Proposition 4.4. *There exists a unique holomorphic map*

$$\varphi_{m+1} : (D^2, \partial D^2) \to (X, L)$$

representing the class β_{m+1} such that its boundary $\partial\varphi_{m+1} := \varphi_{m+1}|_{\partial D^2}$ is exactly given by $\partial\varphi_0 := \varphi_0|_{\partial D^2}$ with the opposite orientation, where φ_0 is the map obtained in Proposition 4.3

Proof. Let $\varphi_{m+1} : (D^2, \partial D^2) \to (X, L)$ be a holomorphic map representing the class β_{m+1} such that $p \in \varphi(\partial D^2)$. By Theorem 4.2 (2), there exists one and only one such map up to automorphisms of D^2. Consider the moduli space $\overline{\mathcal{M}}_{0,1}(X, h)$ of genus 0 stable maps to X with one marked point which represent the fiber class h. Since $X \to Y$ is a $\mathbb{P}^1$-bundle, the evaluation map $\mathrm{ev} : \overline{\mathcal{M}}_{0,1}(X, h) \to X$ is an isomorphism. Hence, there exists a unique (up to automorphisms of the domain) holomorphic map $\phi : \mathbb{P}^1 \to X$ representing the class h which passes through $p \in L \subset X$. The image of this map is the fiber $C_p \cong \mathbb{P}^1$ of $X \to Y$ which contains p. Now, the intersection $C_p \cap L \cong S^1$ splits the fiber C_p into two disks. This gives two holomorphic maps $\varphi_0' : (D^2, \partial D^2) \to (X, L)$ and $\varphi_{m+1}' : (D^2, \partial D^2) \to (X, L)$ with classes β_0 and β_{m+1} respectively. By Theorem 4.2 (2), they must be the same as φ_0, φ_{m+1} up to automorphisms of D^2. Hence, by composing φ_{m+1} with an automorphism of D^2, which is uniquely determined by φ_0, we get the desired unique holomorphic map representing the class β_{m+1}. □

By Proposition 4.4, we can glue the maps $\varphi : (\Sigma^L, \partial\Sigma^L) \to (X, L)$ and $\varphi_{m+1} : (D^2, \partial D^2) \to (X, L)$ together to give a holomorphic map $\varphi' : \Sigma \to X$ which represents the class $\beta_0 + \beta_{m+1} + \alpha = h + \alpha$, where Σ is the union of Σ^L and D^2 with their boundaries identified in the obvious way. It is easy to see that this map is stable. Hence, $\sigma^X := ((\Sigma, z), \varphi')$ represents a point in $\mathcal{M}^X = \overline{\mathcal{M}}_{0,1}(X, h + \alpha)$ and we have $\mathrm{ev}(\sigma) = p$. This defines a map

$$j : \mathcal{M}_p^L \to \mathcal{M}_p^X, \quad [\sigma^L] \mapsto [\sigma^X].$$

This is well-defined: Any automorphism of $\sigma^L = ((\Sigma^L, z), \varphi)$ acts trivially on the component Σ_0^L because φ is nonconstant on this component. So any representative of $[\sigma^L]$ is mapped to the same isomorphism class in $\mathcal{M}_p^X$. We need to show that j is bijective.

Let $\sigma^X = ((\Sigma, z), \varphi)$ be representing a point in $\mathcal{M}_p^X$. This consists of a genus 0 nodal curve Σ with a marked point $z \in \Sigma$ and a stable map $\varphi : \Sigma \to X$ representing the class $h + \alpha$ such that $\varphi(z) = p$. The following is an analog of Proposition 4.3.

Proposition 4.5. *Σ can be decomposed as $\Sigma = \Sigma_0 \cup \Sigma_1$, where $\Sigma_0 \cong \mathbb{P}^1$ is irreducible, such that the restrictions $\varphi_0 := \varphi|_{\Sigma_0}$ and $\varphi_1 := \varphi|_{\Sigma_1}$ represent the classes h and α respectively.*

Proof. By Lemma 3.2 (1), there does not exist any nonconstant holomorphic map from $\mathbb{P}^1$ to X representing the class $h + \alpha$ whose image is not contained entirely in the toric divisors. Hence, Σ must be singular. Decompose Σ into components $\Sigma = \bigcup_a \Sigma_a$, where each $\Sigma_a \cong \mathbb{P}^1$ is irreducible. Then we have

$$\sum_a [\varphi(\Sigma_a)] = h + \alpha.$$

Since h is primitive, there exists a_0 such that $\varphi(\Sigma_{a_0}) = h + \alpha'$ and $\sum_{a \neq a_0} [\varphi(\Sigma_a)] = \alpha''$ for some $\alpha', \alpha'' \in \iota_{0*} H_2^{\text{eff}}(Y, \mathbb{Z}) \subset H_2^{\text{eff}}(X, \mathbb{Z})$ with $\alpha = \alpha' + \alpha''$. By Lemma 3.1, we have $c_1(\alpha') = c_1(\alpha'') = 0$. Then, by Lemma 3.2 (2), the images of $\bigcup_{a \neq a_0} \Sigma_a$ is contained entirely in the zero section D_0. So the image of Σ_{a_0} must be intersecting with L at p. Applying Lemma 3.2 (1) again, we see that α' must be zero. The result follows. □

Note that φ_0 is a nonconstant holomorphic map from $\mathbb{P}^1$ to X whose image contains p. Arguing as in the proof of Proposition 4.4, we see that the image of φ_0 is the fiber C_p of the $\mathbb{P}^1$-bundle $X \to Y$ which contains p, and $\varphi_0(\mathbb{P}^1) \cap L = S^1$. We can then split $\Sigma_0 \cong \mathbb{P}^1$ into two disks $\Sigma_0 = \Sigma_0' \cup \Sigma_0'' \cong D^2 \cup D^2$, and split φ_0 into two holomorphic maps $\varphi_0' : (\Sigma_0', \partial \Sigma_0') \to (X, L)$ and $\varphi_{m+1}' : (\Sigma_0'', \partial \Sigma_0'') \to (X, L)$ which represent the classes β_0 and β_{m+1} respectively. Now, let $\Sigma^L := \Sigma_0' \cup \Sigma_1$ and $\varphi' := \varphi|_{\Sigma^L}$. Then $\varphi' : (\Sigma^L, \partial \Sigma^L) \to (X, L)$ is a genus 0 bordered stable map such that $\varphi(\partial \Sigma^L)$ contains p, and $\sigma^L := ((\Sigma^L, z), \varphi')$ represents a point in $\mathcal{M}_p^L$. By our constructions, $j([\sigma^L]) = [\sigma^X]$. This defines a map $j^{-1} : \mathcal{M}_p^X \to \mathcal{M}_p^L$. Again, since any automorphism of $\sigma^X = ((\Sigma, z), \varphi)$ acts trivially on the component Σ_0, the map j^{-1} is well-defined. It is obvious that this is the inverse map of j. Hence, j is a bijective map.

Proposition 4.6. *Under the bijective map $j : \mathcal{M}_p^L \to \mathcal{M}_p^X$, the Kuranishi structures on $\mathcal{M}_p^L$ and $\mathcal{M}_p^X$ can be naturally identified.*

Proof. We shall first briefly recall the constructions of Kuranishi neighborhoods from [Fukaya and Ono 1999] and [Fukaya et al. 2009].

We begin with $\mathcal{M}_p^L$. Let $\sigma^L = ((\Sigma^L, z), \varphi)$ be representing a point in $\mathcal{M}_p^L$. By Proposition 4.3, we can decompose Σ^L into irreducible components $\Sigma^L = \Sigma_0 \cup \Sigma_1 \cup \ldots \cup \Sigma_k$, where $\Sigma_0 = D^2$ is a disk and $\Sigma_1, \ldots, \Sigma_k$ are copies of $\mathbb{P}^1$, such that the restrictions of φ to Σ_0 and $\bigcup_{a=1}^k \Sigma_a$ represent the classes β_0 and α respectively.

For each $a = 0, 1, \ldots, k$, let $W^{1,p}\big(\Sigma_a; \varphi^*(TX); L\big)$ be the space of sections v of $\varphi^*(TX)$ of $W^{1,p}$ class such that the restriction of v to $\partial\Sigma_a$ lies in $\varphi^*(TL)$, and $W^{0,p}\big(\Sigma_a; \varphi^*(TX) \otimes \bigwedge^{0,1}\big)$ be the space of sections of $\varphi^*(TX) \otimes \bigwedge^{0,1}$ of $W^{0,p}$ class. Note that L does not play a role in the definition of $W^{1,p}\big(\Sigma_a; \varphi^*(TX); L\big)$ for $a = 1, \ldots, k$. Then, let $W^{1,p}\big(\Sigma^L; \varphi^*(TX); L\big)$ be the subspace of

$$\bigoplus_{a=0}^{k} W^{1,p}\big(\Sigma_a; \varphi^*(TX); L\big)$$

consisting of elements

$$\{u = (u_a)\} \in \bigoplus_{a=0}^{k} W^{1,p}\big(\Sigma_a; \varphi^*(TX); L\big)$$

such that for any singular point $w \in \Sigma^L$ which is the intersection of two irreducible components Σ_a and Σ_b, we have $u_a(w) = u_b(w)$. Also let

$$W^{0,p}\big(\Sigma^L; \varphi^*(TX) \otimes \textstyle\bigwedge^{0,1}\big) = \displaystyle\bigoplus_{a=0}^{k} W^{0,p}\big(\Sigma_a; \varphi^*(TX) \otimes \textstyle\bigwedge^{0,1}\big).$$

Consider the linearization of the Cauchy–Riemann operator $\bar\partial$:

$$D_\varphi \bar\partial : W^{1,p}\big(\Sigma^L; \varphi^*(TX); L\big) \to W^{0,p}\big(\Sigma^L; \varphi^*(TX) \otimes \textstyle\bigwedge^{0,1}\big).$$

This is a Fredholm operator by ellipticity.

To construct the obstruction space, choose open subsets W_a of Σ_a whose closure is disjoint from the boundary of each of Σ_a and from the singular and marked points. Then, for each $a = 0, 1, \ldots, k$, by the unique continuation theorem, we can choose a finite dimensional subset E_a of $C_0^\infty(W_a; \varphi^*(TX))$ such that

$$\mathrm{Im}\ D_\varphi \bar\partial + \bigoplus_{a=0}^{k} E_a = W^{0,p}\big(\Sigma^L; \varphi^*(TX) \otimes \textstyle\bigwedge^{0,1}\big).$$

We also choose $\bigoplus_{a=0}^k E_a$ to be invariant under the group Γ_{σ^L} of automorphisms of σ^L. We set $E_{\sigma^L} = \bigoplus_{a=0}^k E_a$.

Let

$$\Pi : W^{0,p}\big(\Sigma^L; \varphi^*(TX) \otimes \textstyle\bigwedge^{0,1}\big) \to W^{0,p}\big(\Sigma^L; \varphi^*(TX) \otimes \textstyle\bigwedge^{0,1}\big)/E_{\sigma^L}$$

be the projection map. Let $V_{\mathrm{map},\sigma^L}$ be the kernel of the operator $\Pi \circ D_\varphi \bar\partial$. Now, consider the automorphism group $\mathrm{Aut}(\Sigma^L, z)$ of the marked bordered Riemann surface

(Σ^L, z). The group $\mathrm{Aut}(\Sigma^L, z)$ may not be finite since some components may be unstable. However, we can naturally embed the Lie algebra $\mathrm{Lie}(\mathrm{Aut}(\Sigma^L, z))$ into $V_{\mathrm{map},\sigma^L}$. Take its L^2 orthogonal complement (with respect to a certain metric). Then let $V'_{\mathrm{map},\sigma^L}$ be a small neighborhood of the zero of it.

On the other hand, let $V_{\mathrm{deform},\sigma^L}$ be a small neighborhood of the origin in the space of first order deformations of the stable components of (Σ^L, z). Also let $V_{\mathrm{resolve},\sigma^L}$ be a small neighborhood of the origin in the space $\bigoplus_w (T_w\Sigma_a \otimes T_w\Sigma_b)$, where the sum is over singular points $w \in \Sigma^L \setminus \Sigma_0$ and Σ_a, Σ_b are the two components such that $\Sigma_a \cap \Sigma_b = \{w\}$. There is a family of marked semistable bordered Riemann surfaces $\{(\Sigma^L_\zeta, z) : \zeta \in V_{\mathrm{deform},\sigma^L} \times V_{\mathrm{resolve},\sigma^L}\}$ over the product $V_{\mathrm{deform},\sigma^L} \times V_{\mathrm{resolve},\sigma^L}$. We remark that, since we do not deform the singular point in Σ_0, each Σ^L_ζ is singular and can be decomposed as $\Sigma^L_\zeta = \Sigma_0 \cup \Sigma'_\zeta$.

Let $V'_{\sigma^L} = V'_{\mathrm{map},\sigma^L} \times V_{\mathrm{deform},\sigma^L} \times V_{\mathrm{resolve},\sigma^L}$. By the proof of Proposition 12.23 in [Fukaya and Ono 1999], there exist a Γ_{σ^L}-equivariant smooth map

$$s_{\sigma^L} : V'_{\sigma^L} \to E_{\sigma^L}$$

and a family of smooth maps

$$\varphi_{u,\zeta} : (\Sigma^L_\zeta, \partial\Sigma^L_\zeta) \to (X, L)$$

for $(u, \zeta) \in V'_{\sigma^L}$ such that $\bar\partial\varphi_{u,\zeta} = s_{\sigma^L}(u, \zeta)$. Now we set $V_{\sigma^L} = \{(u, \zeta) \in V'_{\sigma^L} : \varphi_{u,\zeta}(z) = p\}$. By abuse of notation, denote the restriction of s_{σ^L} to V_{σ^L} also by s_{σ^L}. Then by [Fukaya and Ono 1999], there is a map ψ_{σ^L} mapping $s^{-1}_{\sigma^L}(0)/\Gamma_{\sigma^L}$ onto a neighborhood of $[\sigma^L]$ in $\mathcal{M}^L_p$. This finishes the review of the construction of a Kuranishi neighborhood $(V_{\sigma^L}, E_{\sigma^L}, \Gamma_{\sigma^L}, s_{\sigma^L}, \psi_{\sigma^L})$ of $[\sigma^L] \in \mathcal{M}^L_p$.

For a point in $\mathcal{M}^X_p$ represented by $\sigma^X = ((\Sigma, z), \varphi)$, using Proposition 4.5, we decompose Σ into irreducible components $\Sigma = \Sigma_0 \cup \Sigma_1 \cup \ldots \cup \Sigma_k$, where $\Sigma_0, \Sigma_1, \ldots, \Sigma_k$ are all copies of $\mathbb{P}^1$, such that the restrictions of φ to Σ_0 and $\bigcup_{a=1}^k \Sigma_a$ represent the classes h and α respectively. The construction of a Kuranishi neighborhood $(V_{\sigma^X}, E_{\sigma^X}, \Gamma_{\sigma^X}, s_{\sigma^X}, \psi_{\sigma^X})$ of $[\sigma^X] \in \mathcal{M}^X_p$ is more or less the same as above, except that $W^{1,p}(\Sigma_0; \varphi^*(TX); L)$ is replaced by the space $W^{1,p}(\Sigma_0; \varphi^*(TX))$ of sections v of $\varphi^*(TX)$ of class $W^{1,p}$.

We can now go back to the proof of the proposition.

Let $[\sigma^L] \in \mathcal{M}^L_p$, $[\sigma^X] \in \mathcal{M}^X_p$ be such that $j([\sigma^L]) = [\sigma^X]$. First of all, it is obvious that the automorphism groups Γ_{σ^L} and Γ_{σ^X} are the same. Next, since the moduli space of maps from $(D^2, \partial D^2)$ to (X, L) with class β_0 is unobstructed, we can choose $E_0 = 0$ for the obstruction space E_{σ^L}. Similarly, since the moduli space of maps from $\mathbb{P}^1$ to X with class h is unobstructed, we can also choose $E_0 = 0$ for the obstruction space E_{σ^X}. Hence, the obstruction spaces E_{σ^L} and E_{σ^X} are both of the form $0 \oplus E_1 \oplus \ldots \oplus E_k$ and can be identified naturally.

We can identify $V_{\text{deform},\sigma^L}$ with $V_{\text{deform},\sigma^X}$ since the component Σ_0 in Σ^L has no nontrivial deformations and the component Σ_0 in Σ is unstable. It is also clear that we can identify $V_{\text{resolve},\sigma^L}$ with $V_{\text{resolve},\sigma^X}$. Now, let $(u=(u_0, u_1, \dots, u_k), \zeta) \in V_{\sigma^L}$. Because $E_0 = 0$, we have $D_\varphi \bar{\partial} u_0 = 0$. From the construction of the family of smooth maps $\varphi_{u,\zeta} : (\Sigma^L_\zeta, \partial \Sigma^L_\zeta) \to (X, L)$, it follows that the restriction of $\varphi_{u,\zeta}$ to the component Σ_0 is a holomorphic map with class β_0. We also have $\varphi_{u,\zeta}(z) = p$. But there is a unique (up to automorphisms of the domain) holomorphic map from $(D^2, \partial D^2)$ to (X, L) with class β_0 whose boundary passes through p, which is given by $\varphi|_{\Sigma_0}$. So we must have $u_0 = 0$. By a similar argument, all $(u, \zeta) \in V_{\sigma^X}$ also have $u_0 = 0$. Therefore, we can naturally identify V_{σ^L} and V_{σ^X}.

Finally, we can identify the families of maps $\{\varphi_{u,\zeta} : (\Sigma^L_\zeta, \partial \Sigma^L_\zeta) \to (X, L) : (u, \zeta) \in V_{\sigma^L}\}$ with $\{\varphi_{u,\zeta} : \Sigma_\zeta \to X : (u, \zeta) \in V_{\sigma^X}\}$ by the gluing construction that we used in the definition of the map j. Hence, the maps s_{σ^L} and ψ_{σ^L} can also be naturally identified with the maps s_{σ^X} and ψ_{σ^X} respectively.

This completes the proof of the proposition. □

Theorem 1.1 now follows from Propositions 4.1 and 4.6.

5. Applications to mirror symmetry

In this section, we apply Theorem 1.1 to study mirror symmetry for the toric manifolds $X = \mathbb{P}(K_Y \oplus \mathbb{O}_Y)$. We shall first briefly review the constructions of the mirrors for toric manifolds, following [Cho and Oh 2006; Auroux 2007; 2009; Fukaya et al. 2010a; 2011; Chan and Leung 2010a; 2010b].

As usual, $N \cong \mathbb{Z}^n$ is a rank n lattice, $M = \text{Hom}(N, \mathbb{Z})$ is the dual lattice and $\langle \cdot, \cdot \rangle : M \times N \to \mathbb{Z}$ is the dual pairing. Also let $N_\mathbb{R} = N \otimes_\mathbb{Z} \mathbb{R}$, $M_\mathbb{R} = M \otimes_\mathbb{Z} \mathbb{R}$, and denote by T_N and T_M the real tori $N_\mathbb{R}/N$ and $M_\mathbb{R}/M$ respectively.

Let $X = X_\Delta$ be an n-dimensional smooth projective toric variety defined by a fan Δ in $N_\mathbb{R}$. Let $v_1, \dots, v_d$ be the primitive generators of the 1-dimensional cones in Δ. We equip X with a toric Kähler structure ω. Let P be the corresponding moment polytope and $\mu : X \to P$ be the moment map. P is defined by a set of inequalities

$$P = \{x \in M_\mathbb{R} \mid \langle x, v_i \rangle \geq \lambda_i \text{ for } i = 1, \dots, d\},$$

for some $\lambda_1, \dots, \lambda_d \in \mathbb{R}$. For $i = 1, \dots, d$, we let $l_i : M_\mathbb{R} \to \mathbb{R}$ be the affine linear function defined by $l_i(x) = \langle x, v_i \rangle - \lambda_i$.

We are interested in the mirror symmetry for the Kähler manifold X, equipped with the toric Kähler structure ω and the nowhere zero meromorphic n-form $\Omega = d \log w_1 \wedge \dots \wedge d \log w_n$, where $w_1, \dots, w_n$ are the standard complex coordinates on the open dense orbit $U = N \otimes_\mathbb{Z} \mathbb{C}^* \cong (\mathbb{C}^*)^n \subset X$. From the point of view of [Auroux 2007], we are looking at the mirror symmetry for X relative to the toric divisor $D_\infty = \bigcup_{i=1}^d D_i = X \setminus U$. As before, D_i is the toric prime divisor in X

corresponding to v_i. The mirror geometry is given by a Landau–Ginzburg model $(X^\vee, W)$ consisting of a bounded domain $X^\vee \subset (\mathbb{C}^*)^n$ and a holomorphic function $W : X^\vee \to \mathbb{C}$ called the mirror superpotential.

As discussed in [Auroux 2007; Chan and Leung 2010a; 2010b], the mirror manifold $X^\vee$ can be obtained by dualizing Lagrangian torus fibrations (so-called T-duality) as follows. Restricting the moment map $\mu : X \to P$ to the open dense orbit $U \subset X$ gives a torus bundle $\mu : U \to \mathrm{Int}(P)$, where $\mathrm{Int}(P)$ denotes the interior of the polytope P. In fact this bundle is trivial, so we have $U = \mathrm{Int}(P) \times \sqrt{-1}\, T_N$. The mirror manifold $X^\vee$ is given by the total space of the dual torus bundle, i.e.,

$$X^\vee = \mathrm{Int}(P) \times \sqrt{-1}\, T_N^\vee = \mathrm{Int}(P) \times \sqrt{-1}\, T_M.$$

$X^\vee$ comes with a natural Kähler structure. In particular, as a complex manifold, $X^\vee$ is biholomorphic to a bounded domain in $(\mathbb{C}^*)^n = M_\mathbb{R} \times \sqrt{-1}\, T_M$. If $y = (y_1, \dots, y_n) \in (\mathbb{R}/2\pi\mathbb{Z})^n$ are the fiber coordinates on T_M and the complex coordinates on $(\mathbb{C}^*)^n$ are given by $z_j = \exp(-x_j - \sqrt{-1}\, y_j)$, $j = 1, \dots, n$, where $x = (x_1, \dots, x_n) \in \mathrm{Int}(P)$, then $X^\vee \subset (\mathbb{C}^*)^n$ can be written as

$$X^\vee = \bigl\{(z_1, \dots, z_n) \in (\mathbb{C}^*)^n : |e^{\lambda_i} z_i| < 1,\ i = 1, \dots, d\bigr\}.$$

Geometrically, $X^\vee$ should be viewed as the moduli space of pairs (L, ∇) consisting of a (special) Lagrangian torus fiber of the moment map $\mu : X \to P$ together with a flat $U(1)$-connection ∇ on the trivial line bundle $\underline{\mathbb{C}}$ over L. More precisely, to a point

$$z = \bigl(z_1 = \exp(x_1 + \sqrt{-1}\, y_1), \dots, z_1 = \exp(x_1 + \sqrt{-1}\, y_1)\bigr) \in X^\vee,$$

we associate the flat $U(1)$-connection $\nabla_y = d + (\sqrt{-1}/2) \sum_{j=1}^n y_j\, du_j$ on the trivial line bundle $\underline{\mathbb{C}}$ over the Lagrangian torus $L_x = \mu^{-1}(x) \cong T_N$, where $u = (u_1, \dots, u_n) \in (\mathbb{R}/2\pi\mathbb{Z})^n$ are the fiber coordinates on T_N. This picture is motivated by the SYZ conjecture for mirror Calabi–Yau manifolds proposed by Strominger, Yau and Zaslow in [Strominger et al. 1996].

On the other hand, it turns out that the mirror superpotential $W : X^\vee \to \mathbb{C}$ acts as the mirror of the obstruction $\mathfrak{m}_0$ to the Floer homology of Lagrangian torus fibers in X.[5] As shown in [Fukaya et al. 2009], $\mathfrak{m}_0$ comes from the virtual counting of Maslov index two holomorphic disks in X with boundary in the Lagrangian torus fibers L. This leads to the following expression for W: For $\beta \in \pi_2(X, L)$, we define a holomorphic function Z_β on $X^\vee$ by

$$Z_\beta(L, \nabla) = \exp\Bigl(-\frac{1}{2\pi} \int_\beta \omega\Bigr) \mathrm{hol}_\nabla(\partial\beta).$$

[5] Fukaya, Oh, Ohta and Ono call W the potential function and they define it over the Novikov ring $\bigwedge_0$ instead of $\mathbb{C}$; see [Fukaya et al. 2009; 2010a; 2011].

Then the mirror superpotential $W : X^\vee \to \mathbb{C}$ is given by the following holomorphic function

$$W(L, \nabla) = \sum_{\beta\in\pi_2(X,L),\ \mu(\beta)=2} c_\beta\, Z_\beta(L, \nabla), \tag{5-1}$$

assuming that the sum converges. See [Cho and Oh 2006; Auroux 2007; 2009; Fukaya et al. 2010a; 2011] for more details.

For $i = 1, \ldots, d$, let $\beta_i \in \pi_2(X, L)$ be the relative homotopy class such that $D_j \cdot \beta_i = \delta_{ij}$. Then, by the symplectic area formula of Cho and Oh (Theorem 8.1 in [Cho and Oh 2006]), we have

$$\int_{\beta_i} \omega = 2\pi l_i(x) = 2\pi(\langle x, v_i\rangle - \lambda_i),$$

where $x \in \mathrm{Int}(P)$ is the image of L under the moment map (i.e., $L = \mu^{-1}(x)$). Hence, for the basic classes β_i, the function Z_{β_i} is given in local coordinates by

$$Z_{\beta_i}(L_x, \nabla_y) = \exp(-l_i(x))\ \exp\bigl(-\sqrt{-1}\langle y, v_i\rangle\bigr) = e^{\lambda_i} z^{v_i},$$

where z^v denotes the monomial $z_1^{v^1} \cdots z_n^{v^n}$.

By Theorem 4.2, we have $c_{\beta_i} = 1$ for $i = 1, \ldots, d$. In particular, when X is Fano (i.e., the anticanonical bundle K_X^{-1} is ample), $\beta_1, \ldots, \beta_d \in \pi_2(X, L)$ are the only Maslov index two classes. Hence, the mirror superpotential is given explicitly by

$$W = Z_{\beta_1} + \cdots + Z_{\beta_d} = e^{\lambda_1} z^{v_1} + \cdots + e^{\lambda_d} z^{v_d}.$$

However, in the non-Fano cases, the invariants c_β and hence W are in general very hard to compute. The only non-Fano examples whose mirror superpotentials are explicitly computed are the Hirzebruch surfaces F_2 and F_3, first computed in [Auroux 2009]. Later, Fukaya, Oh, Ohta and Ono gave a different proof for the F_2 case in [Fukaya et al. 2010b].

Let's go back to our toric manifolds $X = \mathbb{P}(K_Y \oplus \mathbb{O}_Y)$. We want to compute their mirror superpotentials using Theorem 1.1.

Lemma 5.1. *If $\beta \in \pi_2(X, L)$ is a Maslov index two class with $c_\beta \neq 0$, then β must either be one of $\beta_1, \ldots, \beta_m, \beta_{m+1}$ or of the form $\beta_0 + \alpha$ for some effective class $\alpha \in H_2(X, \mathbb{Z})$ with $c_1(\alpha) = 0$.*

Proof. First of all, since X is semi-Fano, $c_1(\alpha) \geq 0$ for any effective class $\alpha \in H_2(X, \mathbb{Z})$. Hence, if $\beta \in \pi_2(X, L)$ is a Maslov index two class, then it must be of the form $\beta_i + \alpha$ for some $i = 0, 1, \ldots, m, m+1$ and some effective class $\alpha \in H_2(X, \mathbb{Z})$ with $c_1(\alpha) = 0$. Let $\varphi : (\Sigma^L, \partial\Sigma^L) \to (X, L)$ be a stable map from a genus 0 bordered Riemann surface $(\Sigma^L, \partial\Sigma^L)$ to (X, L) representing the class $\beta_i + \alpha$. Suppose that $\alpha \neq 0$. Then, by the proof of Proposition 4.3, we can decompose

Σ^L into $\Sigma_0^L \cup \Sigma_1$, where $\Sigma_0^L = D^2$ is a disk and Σ_1 is a genus 0 nodal curve, such that the restrictions $\varphi_0 := \varphi|_{\Sigma_0^L}$ and $\varphi_1 := \varphi|_{\Sigma_1}$ represent the classes β_i and α respectively. However, by Lemma 3.2 (2), the image of φ_1 must be contained entirely in the toric prime divisor D_0. Since $\varphi(\Sigma_0^L) \cdot D_0 = \delta_{0i}$ and the domain of φ is connected, we must have $i = 0$. Hence $c_{\beta_i+\alpha} = 0$ unless $i = 0$ or $\alpha = 0$. □

Theorem 5.2. *For the $\mathbb{P}^1$-bundle $X = \mathbb{P}(K_Y \oplus \mathbb{O}_Y)$ over a toric Fano manifold Y, the mirror superpotential $W : X^\vee \to \mathbb{C}$ is given by*

$$W = C Z_{\beta_0} + Z_{\beta_1} + \cdots + Z_{\beta_m} + Z_{\beta_{m+1}},$$

where

$$C = 1 + \sum_{\substack{\alpha \in H_2^{\mathrm{eff}}(X,\mathbb{Z}),\\ \alpha \neq 0,\ c_1(\alpha)=0}} \mathrm{GW}_{0,1}^{X,h+\alpha}([\mathrm{pt}])\, q^\alpha,$$

and $q^\alpha = \exp(-\frac{1}{2\pi}\int_\alpha \omega)$.

Proof. This is a consequence of formula (5-1), Lemma 5.1 and Theorem 1.1. □

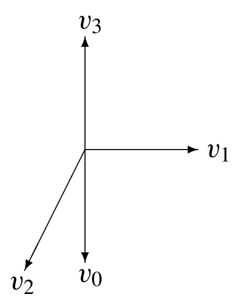
Example: The Hirzebruch surface F_2. Consider $X = F_2 = \mathbb{P}(K_{\mathbb{P}^1} \oplus \mathbb{O}_{\mathbb{P}^1})$. We choose the primitive generators of the 1-dimensional cones in the fan Δ defining F_2 to be[6]

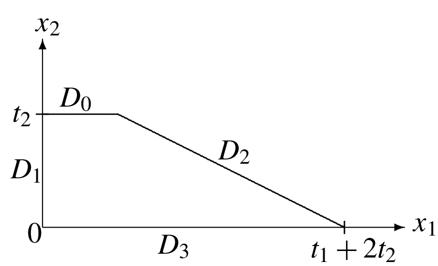
$$v_0 = (0, -1), \quad v_1 = (1, 0), \quad v_2 = (-1, -2), \quad v_3 = (0, 1)$$

in $N = \mathbb{Z}^2$. We equip F_2 with a toric Kähler structure so that moment polytope P is given by

$$P = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, x_2 \leq t_2, x_1 + 2x_2 \leq t_1 + 2t_2\},$$

where $t_1, t_2 > 0$. Here is a depiction of the fan Δ defining F_2 (left) and its moment polytope P (right).





[6]This choice of generators is different from the one in Section 3. This does not alter any of our results. We make this choice just to make our notation consistent with that of [Auroux 2009; Fukaya et al. 2010b].

The effective cone $H_2^{\text{eff}}(F_2, \mathbb{Z})$ is generated by two primitive relations

$$v_0 + v_3 = 0 \quad \text{and} \quad v_1 + v_2 - 2v_0 = 0.$$

Let $h := (1, 0, 0, 1), \alpha := (-2, 1, 1, 0) \in H_2^{\text{eff}}(F_2, \mathbb{Z})$ be the corresponding homology classes, which represent the fiber and the base of F_2 respectively. Then

$$t_1 = \int_\alpha \omega_X \quad \text{and} \quad t_2 = \int_h \omega_X.$$

Let $q_i = \exp(-t_i)$ for $i = 1, 2$. We also have $c_1(h) = 2$ and $c_1(\alpha) = 0$.

Now, the mirror manifold $X^\vee$ is a bounded domain in $(\mathbb{C}^*)^2$. By Theorem 5.2, the mirror superpotential $W : X^\vee \to \mathbb{C}$ is given by

$$W = C Z_{\beta_0} + Z_{\beta_1} + Z_{\beta_2} + Z_{\beta_3} = C\frac{q_2}{z_2} + z_1 + \frac{q_1 q_2^2}{z_1 z_2^2} + z_2,$$

where

$$C = \sum_{k=0}^{\infty} \mathrm{GW}_{0,1}^{F_2, h+k\alpha}\big(\mathrm{PD}[\mathrm{pt}]\big)\, q_1^k,$$

and z_1, z_2 are the standard coordinates on $(\mathbb{C}^*)^2$. F_2 is symplectomorphic to $F_0 = \mathbb{P}^1 \times \mathbb{P}^1$ with induced isomorphism on degree-2 homology given by

$$H_2(F_2, \mathbb{Z}) \to H_2(F_0, \mathbb{Z}) : \quad \alpha \mapsto l_1 - l_2 \quad \text{and} \quad h \mapsto l_2,$$

where $l_1, l_2 \in H_2(F_0, \mathbb{Z})$ are the line classes in the two $\mathbb{P}^1$ factors. Since Gromov–Witten invariants are symplectic invariants, the Gromov–Witten invariants of F_2 are all equal to those of F_0. So we have

$$\mathrm{GW}_{0,1}^{F_2, h+k\alpha}\big(\mathrm{PD}[\mathrm{pt}]\big) = \mathrm{GW}_{0,1}^{F_0, kl_1+(1-k)l_2}\big(\mathrm{PD}[\mathrm{pt}]\big) = \begin{cases} 1, & \text{if } k = 0 \text{ or } k = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Hence, $c_{\beta_0+k\alpha} = 0$ for $k \geq 2$ and $c_{\beta_0+\alpha} = c_{\beta_0} = 1$. We conclude that $C = 1 + q_1$ and the mirror superpotential is given by

$$W = z_1 + z_2 + \frac{q_1 q_2^2}{z_1 z_2^2} + \frac{q_2 + q_1 q_2}{z_2}.$$

This agrees with the formula in Proposition 3.1 in [Auroux 2007]. □

The formula in Theorem 1.1 has been applied to investigate mirror symmetry for various classes of toric manifolds. In [Lau et al. 2011], the formula was generalized and used to compute open Gromov–Witten invariants for toric Calabi–Yau 3-folds. In [Chan et al. 2010] and [Lau et al. 2010], the formula and its generalization in [Lau et al. 2011] were used to obtain an enumerative meaning for the (inverse) mirror maps for toric Calabi–Yau 2- and 3-folds. In particular, this explains why we always get integral coefficients for the Taylor expansions of these mirror maps.

In [Chan and Lau 2010], the formula was used to compute mirror superpotentials for all semi-Fano toric surfaces.

Acknowledgements. The ideas in this paper were motivated by our joint work with Siu-Cheong Lau and Prof. Conan Leung [Chan et al. 2010]. I would like to thank both of them for numerous very useful and inspiring discussions. Thanks are also due to Baosen Wu for generously sharing many of his ideas and insights. I am also grateful to Prof. Shing-Tung Yau and Mr. Sze-Lim Li for their continuous encouragement and support. I also thank IHÉS for hospitality and providing an excellent research environment during my stay in the fall of 2010 when the revision of this paper was done. This research was partially supported by Harvard University, a Croucher Foundation Fellowship and a William Hodge Fellowship.

References

[Auroux 2007] D. Auroux, "Mirror symmetry and T-duality in the complement of an anticanonical divisor", *J. Gökova Geom. Topol. GGT* **1** (2007), 51–91. MR 2009f:53141 Zbl 1181.53076

[Auroux 2009] D. Auroux, "Special Lagrangian fibrations, wall-crossing, and mirror symmetry", pp. 1–47 in *Geometry, analysis, and algebraic geometry: forty years of the Journal of Differential Geometry*, edited by H.-D. Cao and S.-T. Yau, Surv. Differ. Geom. **13**, Int. Press, Somerville, MA, 2009. MR 2010j:53181 Zbl 1184.53085

[Batyrev 1991] V. V. Batyrev, "On the classification of smooth projective toric varieties", *Tohoku Math. J.* (2) **43**:4 (1991), 569–585. MR 92j:14065 Zbl 0792.14026

[Chan and Lau 2010] K. Chan and S.-C. Lau, "Open Gromov–Witten invariants and superpotentials for semi-Fano toric surfaces", preprint, 2010. arXiv 1010.5287

[Chan and Leung 2010a] K. Chan and N. C. Leung, "Mirror symmetry for toric Fano manifolds via SYZ transformations", *Adv. Math.* **223**:3 (2010), 797–839. MR 2011k:14047 Zbl 1201.14029

[Chan and Leung 2010b] K. Chan and N. C. Leung, "On SYZ mirror transformations", pp. 1–30 in *New developments in algebraic geometry, integrable systems and mirror symmetry* (RIMS, Kyoto, 2008), edited by M.-H. Saito et al., Adv. Stud. Pure Math. **59**, Math. Soc. Japan, Tokyo, 2010. MR 2011g:53186 Zbl 1213.14073

[Chan et al. 2010] K. Chan, S.-C. Lau, and N. Leung, "SYZ mirror symmetry for toric Calabi–Yau manifolds", preprint, 2010. To appear in *J. Differential Geom.* arXiv 1006.3830

[Cho and Oh 2006] C.-H. Cho and Y.-G. Oh, "Floer cohomology and disc instantons of Lagrangian torus fibers in Fano toric manifolds", *Asian J. Math.* **10**:4 (2006), 773–814. MR 2007k:53150 Zbl 1130.53055

[Fukaya and Ono 1999] K. Fukaya and K. Ono, "Arnold conjecture and Gromov–Witten invariant", *Topology* **38**:5 (1999), 933–1048. MR 2000j:53116 Zbl 0946.53047

[Fukaya et al. 2009] K. Fukaya, Y.-G. Oh, H. Ohta, and K. Ono, *Lagrangian intersection Floer theory: anomaly and obstruction*, Part I, AMS/IP Studies in Advanced Mathematics **46**, American Mathematical Society, Providence, RI, 2009. MR 2011c:53217 Zbl 1181.53002

[Fukaya et al. 2010a] K. Fukaya, Y.-G. Oh, H. Ohta, and K. Ono, "Lagrangian Floer theory on compact toric manifolds. I", *Duke Math. J.* **151**:1 (2010), 23–174. MR 2011d:53220 Zbl 1190.53078

[Fukaya et al. 2010b] K. Fukaya, Y.-G. Oh, H. Ohta, and K. Ono, "Toric degeneration and non-displaceable Lagrangian tori in $S^2 \times S^2$", preprint, 2010. To appear in *Int. Math. Res. Not.* arXiv 1002.1660

[Fukaya et al. 2011] K. Fukaya, Y.-G. Oh, H. Ohta, and K. Ono, "Lagrangian Floer theory on compact toric manifolds, II: bulk deformations", *Selecta Math.* (*N.S.*) **17**:3 (2011), 609–711. MR 2827178 Zbl 05951105

[Lau et al. 2010] S.-C. Lau, N. Leung, and B. Wu, "Mirror maps equal SYZ maps for toric Calabi–Yau surfaces", preprint, 2010. To appear in *Bull. London Math. Soc.* arXiv 1008.4753

[Lau et al. 2011] S.-C. Lau, N. Leung, and B. Wu, "A relation for Gromov–Witten invariants of local Calabi–Yau threefolds", *Math. Res. Lett.* **18**:5 (2011), 943–956.

[Strominger et al. 1996] A. Strominger, S.-T. Yau, and E. Zaslow, "Mirror symmetry is T-duality", *Nuclear Phys. B* **479**:1-2 (1996), 243–259. MR 97j:32022 Zbl 0896.14024

Received October 21, 2010. Revised March 3, 2011.

KWOKWAI CHAN
DEPARTMENT OF MATHEMATICS
THE CHINESE UNIVERSITY OF HONG KONG
SHATIN
HONG KONG
kwchan@math.cuhk.edu.hk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

A NOTE ON p-HARMONIC l-FORMS ON COMPLETE MANIFOLDS

LIANG-CHU CHANG AND CHIUNG-JUE ANNA SUNG

Let (M^m, g) be an m-dimensional complete noncompact manifold. We show that for all $p > 1$ and $l > 1$, any bounded set of p-harmonic l-forms in $L^q(M)$, with $0 < q < \infty$, is relatively compact with respect to the uniform convergence topology if the curvature operator of M is asymptotically nonnegative.

1. Introduction

Let (M^m, g) be an m-dimensional complete oriented Riemannian manifold with associated Riemannian metric g. Let d be the exterior differential operator and let

$$\delta \equiv * d *$$

be the codifferential operator, where the linear operator $*$ is defined pointwise by

$$*(\omega_1 \wedge \cdots \wedge \omega_l) \equiv \omega_{l+1} \wedge \cdots \wedge \omega_m,$$

for a positively oriented orthonormal coframe $\{\omega_1, \omega_2, \ldots, \omega_m\}$ at the point. The Hodge–Laplace–Beltrami operator $\triangle$ acting on the space of smooth l-forms $\Lambda^l(M)$ is defined by

$$\triangle \equiv -(d\delta + \delta d).$$

Definition 1.1. An l-form ω on M is a p-harmonic l-form if ω satisfies $d\omega = 0$ and $\delta(|\omega|^{p-2}\omega) = 0$ for all $p > 1$.

When $p = 2$, the p-harmonic l-form $\omega \in \Lambda^l(M)$ is called a harmonic l-form on (M, g), that is,

$$\triangle_g \omega = 0.$$

When $l = 0$, let Ω be a compact domain on the Riemannian manifold (M, g), and let ω be a real smooth function on M. For $p > 1$, the p-energy of ω on Ω is

$$E_p(\Omega, \omega) \equiv \frac{1}{p} \int_\Omega |\nabla \omega|^p \, dV_g.$$

The second author was partially supported by NSC.
MSC2010: primary 58A10; secondary 53C21.
Keywords: curvature operator, p-harmonic l-forms, p-harmonic map.

The function ω is said to be p-harmonic on M if ω is a critical point of $E_p(\Omega, \cdot)$ for all $\Omega \subset M$, that is, if ω satisfies the Euler–Lagrange equation

$$\operatorname{div}(|\nabla\omega|^{p-2}\nabla\omega) = 0.$$

A curvature operator K_l on manifold M^m is defined as follows:

$$K_l = \begin{cases} \text{lower bound of the curvature operator on } M & \text{for } l > 1; \\ (m-1)^{-1} \times (\text{lower bound of the Ricci curvature}) & \text{for } l = 1. \end{cases}$$

We call this curvature operator K_l of M asymptotically nonnegative if $K_l \geq -K(r)$, where

$$K(r) : [0, \infty) \to [0, \infty)$$

is a nonnegative and nonincreasing continuous function of distance r to a fixed point $z \in M$, with

$$\int_0^\infty r K(r) < \infty.$$

Yau [1975] proved that any positive harmonic function on a manifold with nonnegative Ricci curvature must be constant. Much work has been done in the finite dimension of space of polynomial growth harmonic functions of growth order at most d [Li 1997; Colding and Minicozzi 1997; Li and Tam 1995; Li and Wang 1999]. Concerning general harmonic l-forms, Li [1980] established a dimension estimate of the space of polynomial growth harmonic forms. In this paper, we study general p-harmonic l-forms and p-harmonic maps on complete noncompact manifolds, for $p > 1$ and $l \neq 0$. For $p = 2$, Chen and Sung [2007] considered the space consisting of all harmonic l-forms of polynomial growth for all $l \geq 1$, and gave a dimension estimate of such a space when M has asymptotically nonnegative curvature. Since the set of p-harmonic l-forms is no longer linear, it is interesting to study the set of p-harmonic l-forms and to seek topological and geometrical links. Interestingly, Zhang [2001] proved that any $L^q(M)$ p-harmonic 1-forms must be zero on a manifold with nonnegative Ricci curvature for $p > 1$ and $0 < q < \infty$. Chang et al. [2010] generalized Zhang's result to a complete manifold M with asymptotically nonnegative curvature and finite first Betti number. They proved that a bounded set of $L^q(M)$ p-harmonic 1-forms on (M, g) has a uniformly convergent subsequence.

Next we introduce the Sobolev inequality. A geodesic ball $B_x(r)$ in a complete manifold M is said to admit a Sobolev inequality $S(C, \nu)$ if there exist constants $C > 0$ and $\nu > 2$ such that for all $f \in C_0^\infty(B_x(r))$, we have

$$\left(\int_{B_x(r)} |f|^{2\nu/(\nu-2)}\right)^{(\nu-2)/\nu} \leq C r^2 V_x^{-2/\nu}(r) \int_{B_x(r)} (|\nabla f|^2 + r^{-2} f^2),$$

where $V_x(r)$ is the volume of geodesic ball $B_x(r)$. Using the Bochner formula, the Moser iteration [1961] and the Sobolev inequality, Chang et al. [2010] showed that any bounded set of p-harmonic 1-forms in $L^q(M)$, with $0 < q < \infty$, is relatively compact with respect to the uniform convergence topology if M has asymptotically nonnegative Ricci curvature and finite first Betti number. However, the Bochner formula does not work for p-harmonic l-forms for $l > 1$. We derive a new type of Bochner formula to overcome this obstacle. We study the set of p-harmonic l-forms, for $l > 1$, on a complete noncompact manifold M, and then study the set of p-harmonic maps from a complete manifold M to a complete manifold N. In Section 2, we derive a different type of Bochner formula for p-harmonic l-forms and prove that any bounded set of p-harmonic l-forms in $L^q(M)$, with $0 < q < \infty$, must be relatively compact with respect to the uniform convergence topology if the curvature operator of M is asymptotically nonnegative. Of course, this implies that the linear space of harmonic l-forms must be finite-dimensional when $p = 2$ and $l \geq 0$. Also, there is no nonzero p-harmonic l-form on M in $L^q(M)$ if the curvature operator of M is nonnegative. In Section 3, we also derive a different type of Bochner formula for p-harmonic maps from M with asymptotically nonnegative Ricci curvature to N with nonpositive sectional curvature. We prove that the set of such p-harmonic maps with finite p-energy on M has a uniformly convergent subsequence. The p-harmonic map is constant if M is compact with nonnegative Ricci curvature, which is an extension of the fact in the harmonic map case ($p = 2$).

2. *p*-harmonic *l*-forms

Any smooth l-form on an m-dimensional manifold M satisfies the Kato inequality:

Lemma 2.1 [Wan and Xin 2004; Calderbank et al. 2000; Herzlich 2000]. *Let ω be a differentiable l-form on M. Then*

$$\big|\nabla|\omega|^2\big| \leq 2|\omega|\,|\nabla\omega|.$$

Lemma 2.2 [Bochner 1946]. *Let $\omega = \sum_I a_I\, \omega_I$ be an l-form on M. Then*

$$\Delta|\omega|^2 = 2\langle\Delta\omega, \omega\rangle + 2|\nabla\omega|^2 + 2K_l\langle\omega, \omega\rangle.$$

Let (M, g) be a complete noncompact manifold. We wish to study the set of L^q p-harmonic l-forms on M for $l > 1$ and $0 < q < \infty$. To prove the main theorem for all $l > 1$, we show a different type of Bochner formula for p-harmonic l-forms:

Lemma 2.3 (Bochner-type formula for p-harmonic forms). *Let ω be a p-harmonic l-form on an m-dimensional complete Riemannian M^m. Then*

$$|\omega|\,\Delta|\omega|^{p-1} = \big\langle\Delta(|\omega|^{p-2}\omega), \omega\big\rangle + |\omega|^{2-p}\Big(\big|\nabla(|\omega|^{p-2}\omega)\big|^2 - \big|\nabla|\omega|^{p-1}\big|^2\Big) + K_l|\omega|^p,$$

in the sense of distributions.

Proof. The Bochner–Weitzenböck formula for $|\omega|^{p-2}\omega$ asserts that

$$\text{(2-1)}\quad \tfrac{1}{2}\Delta\big||\omega|^{p-2}\omega\big|^2 = \big\langle\Delta(|\omega|^{p-2}\omega),\ |\omega|^{p-2}\omega\big\rangle + \big|\nabla(|\omega|^{p-2}\omega)\big|^2 + K_l\big||\omega|^{p-2}\omega\big|^2.$$

The left side of (2-1) is given by

$$\tfrac{1}{2}\Delta\big||\omega|^{p-2}\omega\big|^2 = \tfrac{1}{2}\Delta|\omega|^{2p-2} = \tfrac{1}{2}\Delta\big(|\omega|^{p-1}\big)^2 = |\omega|^{p-1}\,\Delta|\omega|^{p-1} + \big|\nabla|\omega|^{p-1}\big|^2.$$

Hence,

$$|\omega|^{p-1}\,\Delta|\omega|^{p-1} + \big|\nabla|\omega|^{p-1}\big|^2 = \big\langle\Delta(|\omega|^{p-2}\omega),\ |\omega|^{p-2}\omega\big\rangle + \big|\nabla(|\omega|^{p-2}\omega)\big|^2 + K_l\,|\omega|^{2p-4}\,|\omega|^2.$$

It follows that

$$|\omega|^{p-1}\,\Delta|\omega|^{p-1} = |\omega|^{p-2}\big\langle\Delta(|\omega|^{p-2}\omega),\ \omega\big\rangle + \Big(\big|\nabla(|\omega|^{p-2}\omega)\big|^2 - \big|\nabla|\omega|^{p-1}\big|^2\Big) + K_l\,|\omega|^{2p-2}. \qquad \square$$

For l-forms with $l > 1$, the volume comparison property holds on M with asymptotically nonnegative curvature operator [Li and Tam 1995]. Therefore, inside geodesic ball $B_x(R)$ with $r(x) = 2R$, the volume doubling property holds [Li and Tam 1995]. Also, by [Saloff-Coste 1992], a local weak Poincaré inequality holds on geodesic ball $B_x(R)$, and hence we have the Sobolev inequality $S(C, \nu)$ on $B_x(R)$ [Hajłasz and Koskela 1995]; that is, there exists a real number $\nu > 2$ such that

$$\left(\int_{B_x(R)} |f|^{2\nu/(\nu-2)}\,dV\right)^{(\nu-2)/\nu} \le C\cdot r^2\cdot V^{-2/\nu}(B)\int_{B_x(R)} |\nabla f|^2\,dV,$$

for all $f \in C_0^\infty(B_x(r))$, where $r \le R$.

Theorem 2.4 (main theorem). *Let M^m be an m-dimensional complete Riemannian manifold with asymptotically nonnegative curvature operator K_l, for $l > 1$. Then a bounded set of $L^q(M)$ p-harmonic l-forms on (M^m, g) has a uniformly convergent subsequence, for $1 < p < \infty$ and $0 < q < \infty$.*

Proof. Let ω be a p-harmonic l-form on M^m. Lemma 2.3 asserts that

$$|\omega|^{p-1}\,\Delta|\omega|^{p-1} = |\omega|^{p-2}\big\langle\Delta(|\omega|^{p-2}\omega),\ \omega\big\rangle + \Big(\big|\nabla(|\omega|^{p-2}\omega)\big|^2 - \big|\nabla|\omega|^{p-1}\big|^2\Big) + K_l|\omega|^{2p-2}.$$

By the Kato inequality, we have

$$\big|\nabla|\omega|^{p-1}\big| = \big|\nabla\big||\omega|^{p-2}\omega\big|\big| \le \big|\nabla(|\omega|^{p-2}\omega)\big|.$$

Therefore,

$$|\omega|^{p-1}\Delta|\omega|^{p-1} \geq |\omega|^{p-2}\big\langle\Delta(|\omega|^{p-2}\omega),\, \omega\big\rangle - K(R)|\omega|^{2p-2},$$

where $-K(R)$ is the pointwise lower bound of the curvature operator. Let η be a compactly supported nonnegative smooth function on M.

$$\begin{aligned}\int_M \eta^2|\omega|^{p-1}\Delta|\omega|^{p-1} &\geq \int_M \eta^2\,|\omega|^{p-2}\big\langle\Delta(|\omega|^{p-2}\omega),\, \omega\big\rangle - K(R)\int_M \eta^2|\omega|^{2p-2}\\ &= \int_M \eta^2|\omega|^{p-2}\big\langle\delta d(|\omega|^{p-2}\omega),\, \omega\big\rangle - K(R)\int_M \eta^2|\omega|^{2p-2}\\ &= -K(R)\int_M \eta^2|\omega|^{2p-2}.\end{aligned}$$

Integration by parts yields

$$\begin{aligned}K(R)&\int_M \eta^2|\omega|^{2p-2}\\ &\geq \int_M \nabla(\eta^2|\omega|^{p-1})\cdot\nabla|\omega|^{p-1}\\ &\geq \frac{(p-1)^2}{4}\int_M \eta^2|\omega|^{2p-6}\big|\nabla|\omega|^2\big|^2 - (p-1)\int_M \eta|\nabla\eta||\omega|^{2p-4}\big|\nabla|\omega|^2\big|.\end{aligned}$$

It follows that

$$\begin{aligned}\text{(2-2)}\quad \frac{(p-1)^2}{4}&\int_M \eta^2|\omega|^{2p-6}\big|\nabla|\omega|^2\big|^2\\ &\leq (p-1)\int_M \eta|\nabla\eta||\omega|^{2p-4}\big|\nabla|\omega|^2\big| + K(R)\int_M \eta^2|\omega|^{2p-2},\end{aligned}$$

for all $p > 1$.

By Young's inequality, we have

$$(p-1)\eta\,|\nabla\eta|\,|\omega|^{2p-4}\big|\nabla|\omega|^2\big| \leq \frac{(p-1)^2}{8}\eta^2|\omega|^{2p-6}\big|\nabla|\omega|^2\big|^2 + 2|\nabla\eta|^2\,|\omega|^{2p-2}.$$

Since

$$|\omega|^{2p-6}\big|\nabla|\omega|^2\big|^2 = \frac{4}{(p-1)^2}\big|\nabla|\omega|^{p-1}\big|^2,$$

then (2-2) can be written as

$$\text{(2-3)}\qquad \int_M \eta^2\big|\nabla|\omega|^{p-1}\big|^2 \leq 4\int_M |\nabla\eta|^2|\omega|^{2p-2} + 2K(R)\int_M \eta^2|\omega|^{2p-2},$$

for all $p > 1$.

For $R > 0$ and $x \in \partial B_z(2R)$, let $\eta \in \mathscr{C}_0^\infty(B_x(R))$ be a cut-off function satisfying

$$\eta(y) = \begin{cases}1 & \text{if } y \in B_x(\rho R),\\ 0 & \text{if } y \in M\setminus B_x(\gamma R).\end{cases}$$

Note that $\eta \in [0, 1]$ on M and $|\nabla\eta| \le 2/((\gamma-\rho)R)$, for $0 < \rho < \gamma \le 1$.

By the Sobolev inequality and (2-3),

$$\Big(\int_{B_x(\rho R)} (|\omega|^{p-1})^{2\alpha}\Big)^{1/\alpha} \le \Big(\int_{B_x(\gamma R)} (\eta|\omega|^{p-1})^{2\alpha}\Big)^{1/\alpha}$$
$$\le c_s(\nu)V_x(R)^{-2/\nu}R^2\,16\Big(\frac{1}{(\gamma-\rho)^2R^2}+K(R)\Big)\int_{B_x(\gamma R)}|\omega|^{2p-2},$$

where $\alpha = \nu/(\nu-2)$, and $c_s(\nu)$ is the Sobolev constant.

By the assumption on function $K(R)$, it is easy to see that

$$K(R) \le \frac{c}{R^2}$$

on ball $B_x(R)$. Therefore,

(2-4)
$$\Big(\int_{B_x(\rho R)} |\omega|^{2(p-1)\cdot\alpha}\Big)^{1/\alpha} \le c_s(\nu)V_x(R)^{-2/\nu}4^2\Big(\frac{1}{(\gamma-\rho)^2}\Big)\int_{B_x(\gamma R)}|\omega|^{2(p-1)},$$

where $\alpha = \nu/(\nu-2)$.

Define

$$p = q_0\,\alpha^i + 1 \quad \text{and} \quad R_i = (\rho + 2^{-i}(\gamma-\rho))R,$$

for $i = 0, 1, 2, 3, \ldots$. Observe that $\lim_{i\to\infty} R_i = \rho R$. Let $\rho R = R_{i+1}$ and $\gamma R = R_i$ in inequality (2-4) and iterate the inequality; then

$$\sup_{B_x(\rho R)} |\omega|^{2q_0} \le C V_x(R)^{-1}\Big(\frac{1}{\gamma-\rho}\Big)^{\nu}\int_{B_x(\gamma R)}|\omega|^{2q_0}. \tag{2-5}$$

When $q \ge 2q_0$, by (2-5), we have

$$|\omega|(x) \le C\Big(V_x(R)^{-1}\int_{B_x(R)}|\omega|^q\Big)^{1/q},$$

for some constant C.

When $0 < q < 2q_0$, let $h_i = \sum_{j=1}^{i+1} 2^{-j}$, $\rho = h_i$, and $\gamma = h_{i+1}$, for all $i = 0, 1, 2, 3\ldots$. By (2-5), we have

$$\sup_{B_x(h_i R)} |\omega|^{2q_0} \le C V_x(R)^{-1}2^{(i+2)\nu}\int_{B_x(h_{i+1}R)}|\omega|^q \cdot \sup_{B_x(h_{i+1}R)}|\omega|^{2q_0-q}. \tag{2-6}$$

Write $M(i) = \sup_{B_x(h_i R)} |\omega|^{2q_0}$. Inequality (2-6) becomes

$$M(i) \le C V_x(R)^{-1}2^{(i+2)\nu}\int_{B_x(R)}|\omega|^q M(i+1)^{(2q_0-q)/2q_0}. \tag{2-7}$$

Let $\lambda = 1 - q/2q_0 \in (0, 1)$; iterating inequality (2-7), we have

$$M(0) \le \prod_{i=0}^{j-1} \tilde{c}^{\lambda^i} M^{\lambda^j}(j) = \prod_{i=0}^{j-1} \left(C V_x(R)^{-1} 2^{\nu(i+1)} \int_{B_x(R)} |\omega|^q \right)^{\lambda^i} M^{\lambda^j}(j).$$

Let $j \to \infty$; we have

$$M(0) \le (C)^{2q_0/q} V_x(R)^{-2q_0/q} \left(\int_{B_x(R)} |\omega|^q \right)^{2q_0/q}.$$

Hence,

$$|\omega|(x) \le (C)^{1/q} V_x(R)^{-1/q} \left(\int_{B_x(R)} |\omega|^q \right)^{1/q} \le C V_x(R)^{-1/q} \left(\int_{B_x(R)} |\omega|^q \right)^{1/q},$$

for some constant C.

For ω a p-harmonic l-form on M, and $x \in \partial B_z(2R)$, we have

$$|\omega|(x) \le C \left(V_x(R)^{-1} \int_{B_x(R)} |\omega|^q \right)^{1/q}.$$

When the $L^q(M)$ norm of ω is assumed to be bounded by a fixed constant, since we also have $V_x(R) \ge cR$, we conclude that for any given $\epsilon > 0$, by taking R to be sufficiently large, $|\omega| < \epsilon$ on $M \setminus B_z(R)$. On the other hand, using the standard elliptic PDE theory, on ball $B_z(R)$, the length of ω and all its covariant derivatives can be bounded by the $L^q(M)$ norm of ω. In particular, we conclude that any bounded sequence of such ω admits a uniformly convergent subsequence on M. This finishes the proof of the theorem. □

An immediate corollary is obtained from the proof of Theorem 2.4.

Corollary 2.5. *Let (M^m, g) be a complete noncompact manifold with nonnegative curvature operator. Then any bounded $L^q(M)$ p-harmonic l-forms on (M, g) must be zero.*

3. p-Harmonic maps

Here we derive a different type of Bochner formula for p-harmonic maps and study the set of p-harmonic maps with finite p-energy. Let (M^m, g) be a complete Riemannian manifold (without boundary) of dimension m with metric g, and let (N^n, g') be a complete manifold of dimension n with metric g'. For any smooth map $f : M \to N$ and compact domain $\Omega \subset M$, we define the p-energy of f on Ω:

$$E_p(\Omega, f) \equiv \frac{1}{p} \int_\Omega |df(x)|^p \, dV_g,$$

where $|df(x)|$ is the norm of the differential $df(x)$ of f at $x \in \Omega$, dV_g is the volume element of M, and $1 < p < \infty$ is a fixed number. Let $f^{-1}TN$ be the induced vector bundle by f over M. Then df can be viewed as a section of the bundle $\Lambda^1(f^{-1}TN) = T^*M \otimes f^{-1}TN$. We denote by $|df(x)|$ its norm at a point x of M, induced by the metrics g and g'.

A map f is called p-harmonic if it is a critical point of p-energy functional $E_p(\Omega, \cdot)$ for any compact domain $\Omega \subset M$. That is, f is a p-harmonic map if and only if

$$\frac{dE_p(f_s)}{ds} = 0$$

at $s = 0$ for any one-parameter family of maps $f_s : M \to N$ with $f_0 = f$ and $f_s(x) = f(x)$ if $x \in M \setminus \Omega$. We define the p-tension field $\tau_p(f)$ of f by

$$\tau_p(f) = -\delta(|df|^{p-2}df),$$

where $\delta : \Lambda^1(f^{-1}TN) \to \Lambda^0(f^{-1}TN)$ is the codifferential operator. Equivalently, a smooth map $f : M \to N$ is p-harmonic if and only if $\tau_p(f) = 0$.

Assume that (M, g) is a complete noncompact manifold with asymptotically nonnegative Ricci curvature, and that (N, g') is a complete manifold with non-positive sectional curvature. We denote the Ricci tensor of (M, g) by Ricci_M, and the curvature tensor of (N, g') by R_N. Let $\{e_1, \dots, e_m\}$ be a local orthonormal frame on M; by the Weitzenböck formula [Eells and Lemaire 1983], we have

$$\begin{aligned}(3\text{-}1)\quad \tfrac{1}{2}\Delta|df|^2 &= \langle \Delta df, df\rangle + |\nabla df|^2 + \sum_{i=1}^{m}\big\langle df(\mathrm{Ricci}_M(e_i))\cdot df(e_i)\big\rangle \\ &\qquad - \sum_{i,j=1}^{m}\big\langle R_N\big(df(e_j), df(e_i)\big)df(e_i), df(e_j)\big\rangle \\ &\geq \langle \Delta df, df\rangle + |\nabla df|^2 - K|df|^2.\end{aligned}$$

Lemma 3.1 (Bochner-type formula for p-harmonic maps). *Let $u : M \to N$ be a smooth p-harmonic map and $\{e_i\}_{i=1}^m$ be an orthonormal basis of the tangent space of M. Then*

$$\begin{aligned}(3\text{-}2)\quad |du|^{p-1}\Delta|du|^{p-1} &= |du|^{p-2}\big\langle \Delta(|du|^{p-2}du), du\big\rangle \\ &\quad + \Big(\big|\nabla(|du|^{p-2}du)\big|^2 - \big|\nabla|du|^{p-1}\big|^2\Big) \\ &\quad + |du|^{2p-4}\sum_{i}^{m}\big\langle \mathrm{Ricci}_M(du(e_i)), du(e_i)\big\rangle \\ &\quad - |du|^{2p-4}\sum_{i,j=1}^{n}\big\langle R_N(du(e_i), du(e_j))du(e_i), du(e_j)\big\rangle,\end{aligned}$$

in the sense of distributions. Also, if $\mathrm{Ricci}_M \geq 0$ *and* $K_N \leq 0$, *then*

$$
\begin{aligned}
&|du|^{p-1}\,\Delta|du|^{p-1}\\
&\qquad\geq |du|^{p-2}\big\langle\Delta(|du|^{p-2}du),\, du\big\rangle+\Big(\big|\nabla(|du|^{p-2}du)\big|^2-\big|\nabla|du|^{p-1}\big|^2\Big).
\end{aligned}
$$

Proof. The Bochner–Weitzenböck formula for $|du|^{p-1}$ asserts that

$$
\begin{aligned}
\tfrac{1}{2}\Delta|du|^{2p-2} &= \tfrac{1}{2}\Delta\big||du|^{p-2}du\big|^2\\
&=\big\langle\Delta(|du|^{p-2}du),\, |du|^{p-2}du\big\rangle+\big|\nabla(|du|^{p-2}du)\big|^2\\
&\qquad+\sum_i^m\big\langle|du|^{p-2}(\mathrm{Ricci}_M(du(e_i)),\, |du|^{p-2}du(e_i)\big\rangle\\
&\qquad-\sum_{i,j=1}^n\big\langle|du|^{p-2}R_N(du(e_i),du(e_j))du(e_i),\, |du|^{p-2}du(e_j)\big\rangle\\
&=\big\langle\Delta(|du|^{p-2}du),\, |du|^{p-2}du\big\rangle+\big|\nabla(|du|^{p-2}\, du)\big|^2\\
&\qquad+|du|^{2p-4}\sum_i^m\big\langle\mathrm{Ricci}_M(du(e_i)),\, du(e_i)\big\rangle\\
&\qquad-|du|^{2p-4}\sum_{i,j=1}^n\big\langle R_N(du(e_i),du(e_j))du(e_i),\, du(e_j)\big\rangle.
\end{aligned}
$$

On the other hand,

$$
\tfrac{1}{2}\Delta|du|^{2p-2}=\tfrac{1}{2}\Delta(|du|^{p-1})^2=|du|^{p-1}\Delta|du|^{p-1}+\big|\nabla|du|^{p-1}\big|^2.
$$

Hence,

$$
\begin{aligned}
&|du|^{p-1}\Delta|du|^{p-1}+\big|\nabla|du|^{p-1}\big|^2\\
&\qquad=\big\langle\Delta(|du|^{p-2}du),\, |du|^{p-2}du\big\rangle+\big|\nabla(|du|^{p-2}du)\big|^2\\
&\qquad\qquad+|du|^{2p-4}\sum_i^m\big\langle(\mathrm{Ricci}_M(du(e_i)),\, du(e_i)\big\rangle\\
&\qquad\qquad-|du|^{2p-4}\sum_{i,j=1}^n\big\langle R_N(du(e_i),du(e_j))du(e_i),\, du(e_j)\big\rangle.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&|du|^{p-1}\Delta|du|^{p-1}\\
&\quad=|du|^{p-2}\big\langle\Delta(|du|^{p-2}du),du\big\rangle+\Big(\big|\nabla(|du|^{p-2}du)\big|^2-\big|\nabla|du|^{p-1}\big|^2\Big)\\
&\qquad\qquad+|du|^{2p-4}\sum_i^m\big\langle(\mathrm{Ricci}_M(du(e_i)),\, du(e_i)\big\rangle\\
&\qquad\qquad-|du|^{2p-4}\sum_{i,j=1}^n\big\langle R_N(du(e_i),du(e_j))du(e_i),\, du(e_j)\big\rangle.
\end{aligned}
$$

If $\mathrm{Ricci}_M \geq 0$ and $K_N \leq 0$, then

$$|du|^{p-1}\Delta|du|^{p-1} \\ \geq |du|^{p-2}\big\langle\Delta(|du|^{p-2}du), du\big\rangle + \Big(\big|\nabla(|du|^{p-2}du)\big|^2 - \big|\nabla|du|^{p-1}\big|^2\Big). \qquad \square$$

Theorem 3.2. *Let (M, g) be a complete noncompact manifold with asymptotically nonnegative Ricci curvature, and let (N, g') be a complete Riemannian manifold with nonpositive sectional curvature. Then the set of p-harmonic maps u from M to N with $\int_M |du|^p dV_g \leq C$, for some $C > 0$ and $1 < p < \infty$, has a uniformly convergent subsequence.*

Proof. Let u be a p-harmonic map; if $K_N < 0$, the Bochner type formula (3-2) asserts that

$$|du|^{p-1}\Delta|du|^{p-1} \geq du|^{p-2}\big\langle\Delta(|du|^{p-2}du), du\big\rangle \\ + \Big(\big|\nabla(|du|^{p-2}du)\big|^2 - \big|\nabla|du|^{p-1}\big|^2\Big) - |du|^{2p-2}K(R).$$

By the Kato inequality, we have

$$\big|\nabla|du|^{p-1}\big| = \big|\nabla\big||du|^{p-2}du\big|\big| \leq \big|\nabla(|du|^{p-2}du)\big|.$$

Thus,

$$|du|^{p-1}\Delta|du|^{p-1} \geq |du|^{p-2}\big\langle\Delta(|du|^{p-2}du), du\big\rangle - |du|^{2p-2}K(R). \tag{3-3}$$

Dividing both sides of (3-3) by $|du|^{p-2}$, we get

$$|du|\Delta|du|^{p-1} \geq \big\langle\Delta(|du|^{p-2}du), du\big\rangle - |du|^p K(R).$$

Let η be a compactly supported nonnegative smooth function on M; then

$$\begin{aligned}\int_M \eta^2|du|\Delta|du|^{p-1} &\geq \int_M \eta^2\langle(d\delta+\delta d)|du|^{p-2}du, du\rangle - \int_M \eta^2|du|^p K(R) \\ &= \int_M \eta^2\langle d|du|^{p-2}du, d(du)\rangle - \int_M \eta^2|du|^p K(R) \\ &= -\int_M \eta^2|du|^p K(R).\end{aligned}$$

On the other hand, by integration by parts,

(3-4)

$$\begin{aligned}
-\int_M \eta^2|du|^p K(R) &\le \int_M \eta^2|du|\Delta|du|^{p-1}\\
&= -\int_M \nabla(\eta^2|du|)\cdot\nabla|du|^{p-1}\\
&= -\int_M (\eta^2\nabla|du| + |du|2\eta\cdot\nabla\eta)\cdot\big((p-1)|du|^{p-2}\nabla|du|\big)\\
&= -(p-1)\int_M \eta^2|du|^{p-2}\big|\nabla|du|\big|^2\\
&\quad -2(p-1)\int_M \eta\cdot\nabla\eta|du|^{p-1}\cdot\nabla|du|.
\end{aligned}$$

Since

$$\frac{4}{p^2}\big|\nabla|du|^{p/2}\big|^2 = \frac{4}{p^2}\Big|\frac{p}{2}|du|^{(p/2)-1}\nabla|du|\Big|^2 = \big|du|^{p-2}|\nabla|du|\big|^2$$

and

$$\frac{2}{p}|du|^{p/2}\nabla|du|^{p/2} = \frac{2}{p}|du|^{p/2}\frac{p}{2}|du|^{(p/2)-1}\nabla|du| = |du|^{p-1}\nabla|du|,$$

inequality (3-4) can be rewritten as

$$\begin{aligned}
-\int_M &\eta^2|du|^p K(R)\\
&\le -\frac{4(p-1)}{p^2}\int_M \eta^2|\nabla|du|^{p/2}|^2 - \frac{4(p-1)}{p}\int_M |du|^{p/2}\cdot\nabla\eta\cdot\eta\cdot\nabla|du|^{p/2}.
\end{aligned}$$

By Young's inequality,

$$\begin{aligned}
-\int_M &\eta^2|du|^p K(R)\\
&\le -\frac{4(p-1)}{p^2}\int_M \eta^2\big|\nabla|du|^{p/2}\big|^2 + \Big(\zeta\int_M \eta^2\big|\nabla|du|^{p/2}\big|^2 + \frac{c_1}{\zeta}\int_M |\nabla\eta|^2|du|^p\Big),
\end{aligned}$$

for some positive constants c_1 and $0<\zeta<1$. Therefore,

$$\begin{aligned}
\text{(3-5)}\quad \Big(\frac{4(p-1)}{p^2} - 2\zeta\Big)&\int_M \eta^2\big|\nabla|du|^{p/2}\big|^2\\
&\le \frac{c_2}{\zeta}\Big(\int_M |\nabla\eta|^2|du|^p + \int_M \eta^2|du|^p K(R)\Big).
\end{aligned}$$

For $R>0$ and $x\in\partial B_z(2R)$, let $\eta\in C_0^\infty(B_x(R))$ be a cut-off function such that

$$\eta(y) = \begin{cases} 1 & \text{if } y\in B_x(\rho R),\\ 0 & \text{if } y\in M\setminus B_x(\gamma R).\end{cases}$$

Note that $\eta \in [0, 1]$ on M and $|\nabla\eta| \le c_3/R$, for $0 < \rho < \gamma \le 1$ and some positive constant c_3.

By the curvature assumption on function $K(R)$, we have

$$K(R) \le \frac{c_4}{R^2},$$

for some constant c_4. Let $\zeta = (p-1)/p^2$; then inequality (3-5) becomes

$$\int_{B_x(R)} \big|\nabla|du|^{p/2}\big|^2 \le \frac{c_5}{R^2}\int_{B_x(R)} |du|^p + \int_{B_x(R)} \frac{c_6}{R^2}|du|^p \le \frac{C}{R^2}\int_{B_x(R)} |du|^p.$$

Therefore, for u a p-harmonic map from M to N and $x \in \partial B_z(2R)$, we have

$$\int_{B_x(R)} \big|\nabla|du|^{p/2}\big|^2 \le \frac{C}{R^2}\int_M |du|^p.$$

When $\int_M |du|^p$ is assumed to be bounded by a fixed constant, by taking R to be sufficiently large, for any $\epsilon > 0$, we have $|\nabla|du|^{p/2}| < \epsilon$ on $M \backslash B_z(R)$. On the other hand, $\big|\nabla|du|^{p/2}\big|$ can be bounded by the finite energy of u on ball $B_z(R)$. We conclude that the set of such p-harmonic maps admits a uniformly convergent subsequence. If M is a compact manifold with nonnegative Ricci curvature, then the p-harmonic map is constant, which is an extension of the fact in the harmonic map case ($p = 2$) [Eells and Sampson 1964]. □

Acknowledgement

Part of this work was done when the second author was visiting the Department of Mathematics at Princeton University. She thanks the members of the department for their hospitality during her visit.

References

[Bochner 1946] S. Bochner, "Vector fields and Ricci curvature", *Bull. Amer. Math. Soc.* **52** (1946), 776–797. MR 8,230a Zbl 0060.38301

[Calderbank et al. 2000] D. M. J. Calderbank, P. Gauduchon, and M. Herzlich, "Refined Kato inequalities and conformal weights in Riemannian geometry", *J. Funct. Anal.* **173**:1 (2000), 214–255. MR 2001f:58046 Zbl 0960.58010

[Chang et al. 2010] L.-C. Chang, C.-L. Guo, and C.-J. A. Sung, "p-harmonic 1-forms on complete manifolds", *Arch. Math.* (*Basel*) **94**:2 (2010), 183–192. MR 2011c:53064 Zbl 1210.58001

[Chen and Sung 2007] J.-T. R. Chen and C.-J. A. Sung, "Dimension estimate of harmonic forms on complete manifolds", *Pacific J. Math.* **232**:1 (2007), 91–109. MR 2009c:53049 Zbl 1151.53058

[Colding and Minicozzi 1997] T. H. Colding and W. P. Minicozzi, II, "Harmonic functions on manifolds", *Ann. of Math.* (2) **146**:3 (1997), 725–747. MR 98m:53052 Zbl 0928.53030

[Eells and Lemaire 1983] J. Eells and L. Lemaire, *Selected topics in harmonic maps*, CBMS Regional Conference Series in Mathematics **50**, American Mathematical Society, Providence, RI, 1983. MR 85g:58030 Zbl 0515.58011

[Eells and Sampson 1964] J. Eells, Jr. and J. H. Sampson, "Harmonic mappings of Riemannian manifolds", *Amer. J. Math.* **86** (1964), 109–160. MR 29 #1603 Zbl 0122.40102

[Hajłasz and Koskela 1995] P. Hajłasz and P. Koskela, "Sobolev meets Poincaré", *C. R. Acad. Sci. Paris Sér. I Math.* **320**:10 (1995), 1211–1215. MR 96f:46062 Zbl 0837.46024

[Herzlich 2000] M. Herzlich, "Refined Kato inequalities in Riemannian geometry", exposé VI in *Journées "Équations aux Dérivées Partielles"*, Université de Nantes, 2000. MR 2001i:58068 Zbl 1028.58025

[Li 1980] P. Li, "On the Sobolev constant and the p-spectrum of a compact Riemannian manifold", *Ann. Sci. École Norm. Sup.* (4) **13**:4 (1980), 451–468. MR 82h:58054 Zbl 0466.53023

[Li 1997] P. Li, "Harmonic sections of polynomial growth", *Math. Res. Lett.* **4**:1 (1997), 35–44. MR 98i:53054 Zbl 0880.53039

[Li and Tam 1995] P. Li and L.-F. Tam, "Green's functions, harmonic functions, and volume comparison", *J. Differential Geom.* **41**:2 (1995), 277–318. MR 96f:53054 Zbl 0827.53033

[Li and Wang 1999] P. Li and J. Wang, "Counting massive sets and dimensions of harmonic functions", *J. Differential Geom.* **53**:2 (1999), 237–278. MR 2001k:53063 Zbl 1039.53036

[Moser 1961] J. Moser, "On Harnack's theorem for elliptic differential equations", *Comm. Pure Appl. Math.* **14** (1961), 577–591. MR 28 #2356 Zbl 0111.09302

[Saloff-Coste 1992] L. Saloff-Coste, "Uniformly elliptic operators on Riemannian manifolds", *J. Differential Geom.* **36**:2 (1992), 417–450. MR 93m:58122 Zbl 0735.58032

[Wan and Xin 2004] T. Y. H. Wan and Y. L. Xin, "Vanishing theorems for conformally compact manifolds", *Comm. Partial Differential Equations* **29**:7-8 (2004), 1267–1279. MR 2005h:53058 Zbl 1086.53060

[Yau 1975] S. T. Yau, "Harmonic functions on complete Riemannian manifolds", *Comm. Pure Appl. Math.* **28** (1975), 201–228. MR 55 #4042 Zbl 0291.31002

[Zhang 2001] X. Zhang, "A note on p-harmonic 1-forms on complete manifolds", *Canad. Math. Bull.* **44**:3 (2001), 376–384. MR 2002c:58027 Zbl 0989.58006

Received January 12, 2011.

LIANG-CHU CHANG
CENTER FOR GENERAL EDUCATION
NATIONAL FORMOSA UNIVERSITY
HUWEI, YUNLIN COUNTY 632
TAIWAN

lachu4@gmail.com

CHIUNG-JUE ANNA SUNG
DEPARTMENT OF MATHEMATICS
NATIONAL TSING HUA UNIVERSITY
HSINCHU 30013
TAIWAN

cjsung@math.nthu.edu.tw

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

THE CHEEGER CONSTANT OF CURVED STRIPS

DAVID KREJČIŘÍK AND ALDO PRATELLI

We study the Cheeger constant and Cheeger set for domains obtained as strip-like neighborhoods of curves in the plane. If the reference curve is complete and finite (a "curved annulus"), then the strip itself is a Cheeger set and the Cheeger constant equals the inverse of the half-width of the strip. The latter holds true for unbounded strips as well, but there is no Cheeger set. Finally, for strips about noncomplete finite curves, we derive lower and upper bounds to the Cheeger set, which become sharp for infinite curves. The paper is concluded by numerical results for circular sectors.

1. Introduction

Let Ω be an open connected set in the plane $\mathbb{R}^2$. The *Cheeger constant* of Ω is defined as

$$h(\Omega) := \inf_{S\subseteq\Omega} \frac{P(S)}{|S|}, \qquad (1)$$

where the infimum is taken over all sets $S \subseteq \Omega$ of finite perimeter. We use $P(S)$ and $|S|$ to denote the perimeter and the area of S, respectively. Any minimizer of (1), if it exists, is called a *Cheeger set* of Ω and is denoted by $\mathscr{C}_\Omega$.

The problems of existence, uniqueness, and regularity of Cheeger sets have been widely studied in recent years; see, for example, [Kawohl and Fridman 2003; Hebey and Saintier 2006; Saintier 2007; Caselles et al. 2007]. We briefly list and discuss here some of the known general properties.

Theorem 1.1 (general facts). (i) *While for a general Ω neither existence nor uniqueness is guaranteed, there is always some Cheeger set if Ω is a bounded open set.*

(ii) *If $\Omega_1 \subseteq \Omega_2$, then $h(\Omega_1) \geq h(\Omega_2)$, but the strict inclusion does not imply the strict inequality.*

This work was partially supported by the Czech Ministry of Education, Youth and Sports within the project LC06002, by the Spanish Ministry of Education and Science within the project MTM2008-03541, and by the ERC within the advanced grant number 226234 and the starting grant number 258685. Krejčiřík acknowledges support from Ikerbasque (Basque Foundation for Science), Spain.
MSC2010: 28A75, 49Q20, 35P15, 51M16.
Keywords: Cheeger sets, Cheeger constant, curved strips.

(iii) *The boundary of any Cheeger set* $\mathscr{C}_\Omega$ *intersects the boundary of the set* Ω.

(iv) *The part of* $\partial\mathscr{C}_\Omega$ *which is inside* Ω *is made by arcs of circle, each of which starts and ends touching the boundary of* Ω *and is of radius* $1/h(\Omega)$.

(v) *A Cheeger set cannot have corners (that is, discontinuities in the tangent vector to the boundary giving rise to an angle smaller than* π*). In particular, the arcs of circle of* $\partial\mathscr{C} \cap \Omega$ *must intersect the boundary of* Ω *tangentially or in "open corners" (that is, angles bigger than* π*).*

(vi) *If there is a Cheeger set, there is a connected Cheeger set.*

Concerning property (i), examples of nonexistence or nonuniqueness can be found in [Kawohl and Lachand-Robert 2006], while the existence is immediate by the compactness results for BV functions; see, for example, [Evans and Gariepy 1992; Ambrosio et al. 2000]. Property (ii) is immediate by the definition (1), and examples for the nonstrict inequality can be found in [Kawohl and Lachand-Robert 2006]. Property (iii) comes immediately by a rescaling of $\mathscr{C}$ with a factor bigger than 1, since this lowers the ratio in (1). Property (iv) comes from a standard variational argument; see, for example, [Kawohl and Fridman 2003, Remark 9]. Property (v) comes directly by noticing that "cutting a corner" of a small length ε decreases $|\mathscr{C}_\Omega|$ by at most $C\varepsilon^2$ and the perimeter by at least $c\varepsilon$. By "corner" we mean a point of the boundary where the tangent vector is discontinuous and makes an angle smaller than π (with respect to the internal part of Ω). In the case of angles bigger than π, we talk about "open corners", and they cannot be excluded from $\partial\mathscr{C}$, since, as pointed out in [Kawohl and Lachand-Robert 2006], there are open corners (or "reentrant corners" in their terminology) in an L-shaped set. Finally, property (vi) is immediate because if a Cheeger set has different connected components, each of these components must also be a Cheeger set thanks to the characterization (1).

Apart from the above-mentioned general properties, it is usually a difficult task to find the Cheeger constant or the Cheeger set of a given domain Ω. The situation is simplified when Ω is a bounded convex set, which is a well-studied situation. In fact, in this case it is known that there is a unique open Cheeger set, which is also convex; see [Alter et al. 2005; Kawohl and Lachand-Robert 2006; Caselles et al. 2007]. Moreover, it is possible to give the following characterization.

Theorem 1.2 [Kawohl and Lachand-Robert 2006]. *Let* Ω *be a bounded convex subset of* $\mathbb{R}^2$. *For* $r \geq 0$, *define*

$$\Omega^r := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > r\}.$$

There exists a unique value $r = r^* > 0$ *such that*

$$|\Omega^r| = \pi r^2. \tag{2}$$

Then $h(\Omega) = 1/r^*$ and the Cheeger set of Ω is the Minkowski sum $\mathscr{C}_\Omega = \Omega^{r^*} + B_{r^*}$, with B_{r^*} denoting the disc of radius r^*.

This theorem can be used to find explicitly $h(\Omega)$ and $\mathscr{C}_\Omega$ in some cases (for example for discs, rectangles and triangles). In particular, the Cheeger sets of rectangles and triangles are obtained by suitably "cutting the corners". Furthermore, it provides a constructive algorithm for the determination of the Cheeger constant and Cheeger set for general convex domains; in particular for convex polygons.

Unfortunately, there is no such constructive method for nonconvex domains. Only one particular case seems to be explicitly known in the literature, namely the annulus, for which it is known that $\mathscr{C}_\Omega = \Omega$. In general, while a trivial strategy to find upper estimates for $h(\Omega)$ is to choose a suitable "test domain" S in (1), it is less clear how to obtain lower estimates. One possibility is given by the following result concerning "test vector fields".

Theorem 1.3 [Grieser 2006]. *Let $V : \Omega \to \mathbb{R}^2$ be a smooth vector field on Ω, $h \in \mathbb{R}$, and assume that the pointwise inequalities $|V| \leq 1$ and $\operatorname{div} V \geq h$ hold in Ω. Then $h(\Omega) \geq h$.*

An example of the applicability of this criterion is the above-mentioned result for the annulus, which can be obtained by employing the vector field of [Bellettini et al. 2002, Section 11, Example 4] (see also Remark 2.6 below, where the corresponding vector field can be found explicitly). However, for a general set Ω it is not easy at all to find a vector field producing nontrivial lower bounds by this criterion.

The purpose of this paper is to introduce a class of nonconvex planar domains for which the Cheeger constant and the Cheeger set can be determined explicitly, namely the curved strips. This class of sets has been intensively studied in the last two decades as an effective configuration space for curved quantum waveguides (see [Duclos and Exner 1995; Krejčiřík and Kříž 2005] and the references therein).

More precisely, we call a tubular neighborhood of a curve without boundary in the plane a "curved strip". There are then few possibilities: a "curved annulus", a "finite curved strip", an "infinite curved strip", or a "semi-infinite curved strip"; see Figure 1 (we leave the formal definitions to Section 1).

Our main results, Theorems 3.1 and 3.2, describe the situation in all of these cases. In particular, for a curved annulus the situation is analogous to the standard annulus, that is, the strip itself is the unique Cheeger set and the Cheeger constant only depends on the width of the strip, irrespectively of the curvature of the curve. More precisely, the Cheeger constant is the inverse of the half-width (Theorem 3.1, part (i)). For an infinite or a semi-infinite curved strip, again the Cheeger constant equals the inverse of the half-width of the strip, but there is no Cheeger set (Theorem 3.1, part (ii)). Finally, for a finite curved strip, the situation is analogous to the standard rectangle, that is, there exists a Cheeger set, which

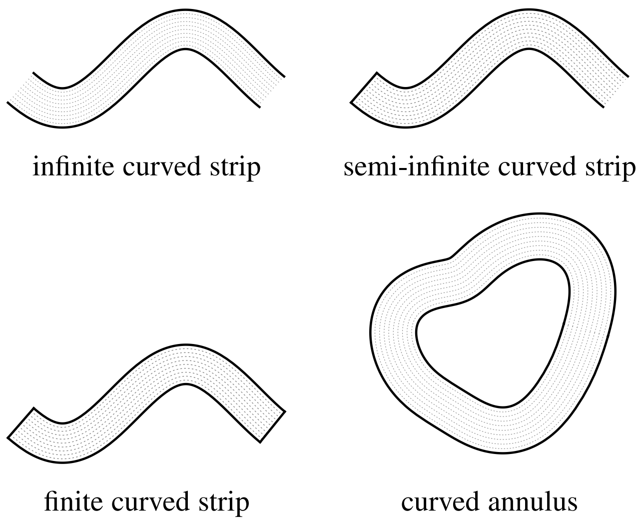


Figure 1. The four possible types of strips.

is not the whole strip because of the corners, and the Cheeger constant is strictly bigger than the inverse of the half-width. Moreover, in this last case we can also give a (sharp) upper and a lower bound, which depend only on the width and length of the strip (Theorem 3.2).

We conclude this introductory section with a couple of comments. First of all, it should mentioned that in the study of the Cheeger problem an important role is played by those sets Ω which are Cheeger sets of themselves. This is what happens in many situations, such as the discs, the annuli, and, as we show in the present paper, the "curved annuli". Those sets are called *calibrable* and are intensively studied in the image processing literature; see, for instance, [Bellettini et al. 2002].

A second remark has to be made on the connection between the Cheeger constant and the eigenvalue problems. In fact, the *Cheeger inequality* tells us that

$$\lambda_p(\Omega) \geq \left(\frac{h(\Omega)}{p}\right)^p \tag{3}$$

for any $p \in (1, \infty)$, where $\lambda_p(\Omega)$ is the first eigenvalue of the p-Laplacian. Moreover, as shown in [Kawohl and Fridman 2003], $h(\Omega) = \lim_{p \searrow 1} \lambda_p(\Omega)$. In this regard, it is interesting to notice one property of the curved strips. It is well known that the first eigenvalue of the Dirichlet Laplacian (or, more generally, the infimum of the Rayleigh quotient, in the case of unbounded strips for which there might be no eigenvalues) for a curved strip strongly depends on its curvature; see, for instance, [Duclos and Exner 1995; Exner et al. 2004; Krejčiřík and Kříž 2005]. On the other hand, the Cheeger constant is much less sensitive, because, as we will show, for infinite and semi-infinite curve strips, as well as for curved annuli, the Cheeger constant depends not on the curvature of the strip, but on its width.

The geometrical setting. In this section we set the notations for the geometrical situation that we will consider throughout the paper. Let Γ be a C^2, connected curve in $\mathbb{R}^2$ (that is, the homeomorphic image of (0, 1) or $\mathbb{S}^1$ under a C^2 function), and let us denote by $|\Gamma| = \int_\Gamma dq$ its length, dq being the arclength element of Γ. Additionally, let $N : \Gamma \to \mathbb{R}^2$ be a C^1 vector field giving the normal vector in the points of Γ, and let $\kappa : \Gamma \to \mathbb{R}$ be the associated curvature (notice that the sign of κ depends on the choice of the orientation of N). We recall that to define κ it is enough to take a unit-speed parametrization γ of Γ, and hence it is

$$\kappa(q) = \ddot{\gamma}(\gamma^{-1}(q)) \cdot N(q), \tag{4}$$

where the dot denotes the standard scalar product in $\mathbb{R}^2$. We now introduce a mapping $\mathscr{L}$ from $\Gamma \times \mathbb{R}$ to $\mathbb{R}^2$,

$$\mathscr{L}(q, t) := q + tN(q),$$

and we introduce the set

$$\Omega_{\Gamma,a} := \mathscr{L}(\Gamma \times (-a, a)),$$

for any positive a. We are interested in sets $\Omega_{\Gamma,a}$ that are non-self-intersecting tubular neighborhoods of Γ. More precisely, we will always assume that

$$\mathscr{L} \text{ is injective in } \Gamma \times [-a, a], \tag{5}$$

hence the set is as in Figure 2. Using the expression for the bilinear form

$$d\mathscr{L}^2 = (1 - \kappa(q)t)^2\, dq^2 + dt^2 \tag{6}$$

that follows from (4), we can easily see that, by the inverse function theorem, the assumption (5) forces a to be small compared to the curvature. More precisely, (5) implies that $|\kappa(q)|a \le 1$ for any $q \in \Gamma$, the boundary of $\Omega_{\Gamma,a}$ is $C^{1,1}$, and $\mathscr{L}$ is in fact a C^1 diffeomorphism between $\Gamma \times (-a, a)$ and $\Omega_{\Gamma,a}$.

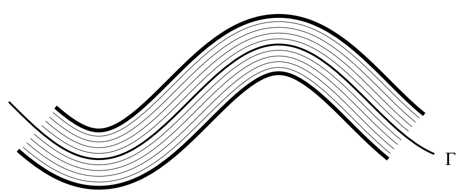


Figure 2. The geometry of a curved strip $\Omega_{\Gamma,a}$ and the corresponding curve Γ; the parallel lines correspond to the curves $s \mapsto \mathscr{L}(s, t)$ with fixed $t \in (-a, a)$.

Summing up, under the hypothesis (5) $\Omega_{\Gamma,a}$ has the geometrical meaning of an open non-self-intersecting strip contained between the parallel curves

$$q \mapsto q \pm a N(q),$$

with $q \in \Gamma$, and it can be identified with the Riemannian manifold $\Gamma \times (-a, a)$ equipped with the metric (6).

In this paper, we will call any set $\Omega_{\Gamma,a}$ satisfying the assumption (5) a curved strip. Notice that when Γ is contained in a line, Ω reduces to a rectangle. But the most interesting situation is when Γ has a more complicated geometry, since then the associated set is not convex, hence not covered by the preceding known results for the Cheeger problem. It is easy to characterize the four possible situations occurring for a curved strip, to each of which we will associate a name to fix the ideas. The four kinds of strips are shown in Figure 1. First of all, if the curve Γ is not finite, it may be either infinite or semi-infinite (that is, not finite but complete, or not finite and not complete, respectively). We will call the corresponding sets $\Omega_{\Gamma,a}$ an *infinite curved strip* and a *semi-infinite curved strip*. On the other hand, if the curve is finite, then it can be either compact or not compact (homeomorphic to a circle or to an open segment, respectively). In the first case, we will speak about a *curved annulus*, the annulus corresponding to the case when Γ is exactly a circle, and in the other case about a *finite curved strip*.

2. The main geometrical results

In this section we give some general technical properties, which will be used later to show our main results. We can easily obtain an upper bound for the curved strips. In the next result, for a curve Γ which is not finite, we consider a unit-speed parametrization $\gamma : (0, +\infty) \to \mathbb{R}^2$ (respectively, $\gamma : (-\infty, +\infty) \to \mathbb{R}^2$) if the strip is semi-infinite (respectively, infinite). We will denote by Γ_L the subset of Γ given by $\gamma(0, L)$ or $\gamma(-L, L)$ for the semi-infinite or infinite case, respectively.

Lemma 2.1 (upper bound). *Let Γ be infinite or compact (that is, $\Omega_{\Gamma,a}$ is a semi-infinite or infinite curved strip, or a curved annulus). Then*

$$h(\Omega_{\Gamma,a}) \leq \frac{1}{a}.$$

In particular, if $\Omega_{\Gamma,a}$ is a curved annulus,

$$\frac{P(\Omega_{\Gamma,a})}{|\Omega_{\Gamma,a}|} = \frac{1}{a},$$

while if $\Omega_{\Gamma,a}$ is a semi-infinite or infinite curved strip,

$$\frac{P(\Omega_{\Gamma_L,a})}{|\Omega_{\Gamma_L,a}|} \xrightarrow[L\to\infty]{} \frac{1}{a}.$$

Proof. If $\Omega_{\Gamma,a}$ is a curved annulus, then we take the whole $S = \Omega_{\Gamma,a}$ as a test domain in (1). Recalling (6), we then have

$$\frac{P(S)}{|S|} = \frac{\int_\Gamma (1+\kappa(q)a)\,dq + \int_\Gamma (1-\kappa(q)a)\,dq}{\int_\Gamma \int_{-a}^{a} (1-\kappa(q)t)\,dt\,dq} = \frac{2|\Gamma|}{2a|\Gamma|} = \frac{1}{a}.$$

Notice that, by the symmetry of the set S, the curvature term cancels both in the numerator and in the denominator.

On the other hand, if Γ is not finite, then the whole strip is not admissible because it has both infinite area and perimeter. However, for any $L > 0$, we can consider the finite curved strip $S = \Omega_{\Gamma_L,a}$, which is of course contained in $\Omega_{\Gamma,a}$. Therefore, one can easily evaluate

$$\begin{aligned} \frac{P(S)}{|S|} &= \frac{4a + \int_{\Gamma_L} \big(1+\kappa(q)a\big)\,dq + \int_{\Gamma_L} \big(1-\kappa(q)a\big)\,dq}{\int_{\Gamma_L} \int_{-a}^{a} \big(1-\kappa(q)t\big)\,dt\,dq} \\ &= \frac{4a + 2|\Gamma_L|}{2a|\Gamma_L|} \xrightarrow[L\to\infty]{} \frac{1}{a}. \end{aligned} \tag{7}$$

In the formula for the perimeter, notice the term $4a$ corresponding to the two "vertical" parts of ∂S at the start and at the end. Thanks to the definition (1), the two above estimates give the thesis. □

The lower bound is much more complicated to obtain. To find it, we introduce an operation that, in a sense, fills in the "holes" and the "bays" in the test domains S. More precisely, let us take an open set $S \subseteq \Omega_{\Gamma,a}$, and define the set Γ_S as

$$\Gamma_S := \{q \in \Gamma : \mathscr{L}(\{q\} \times (-a, a)) \cap S \neq \varnothing\},$$

and the functions $f_\pm : \Gamma_S \to [-a, a]$ as

$$f_-(q) := \inf\{t \in (-a, a) : (q, t) \in S\}, \quad f_+(q) := \sup\{t \in (-a, a) : (q, t) \in S\}.$$

Therefore, S is contained between the two graphs of f_+ and f_-. Notice now that, if S is connected, then of course so is Γ_S. In particular, there are two possibilities: either Γ_S is a subinterval of Γ and, in this case, we call q_l and q_r its extremes, or Γ_S is a closed curve. Observe that if Γ is not compact (that is, always except when $\Omega_{\Gamma,a}$ is a curved annulus), then Γ_S must necessarily be a subinterval of Γ; on the other hand, if $\Omega_{\Gamma,a}$ is a curved annulus, then both the situations — that Γ_S is a subinterval of Γ, and that Γ_S is a closed curve — are possible, and, in particular, Γ_S is a closed curve if and only if $\Gamma_S = \Gamma$.

Definition 2.2. Let S be an open subset of $\Omega_{\Gamma,a}$ with finite perimeter, and let Γ_S and $f_\pm$ be defined as above. We define

$$S^* := \{\mathscr{L}(q, t) \in \Omega_{\Gamma,a} : q \in \Gamma_S,\ f_-(q) < t < f_+(q)\}.$$

We can now show the main property of the set S^*, which will be fundamental for our purposes.

Lemma 2.3 (area and perimeter of S^*). *Let S be an open, bounded, and connected subset of Ω of finite perimeter. Then*

$$|S^*| \geq |S|, \quad P(S^*) \leq P(S),$$

and $f_\pm \in BV(\Gamma_S)$. Moreover, calling $f'_\pm \, dq$ the absolute continuous part of $Df_\pm$, and $D_s f_\pm$ its singular part, we have the validity of the formula

$$\begin{aligned}(8) \quad P(S^*) = &\int_{\Gamma_S} \sqrt{(1-\kappa(q) f_+(q))^2 + f'_+(q)^2}\, dq \\ &+ \int_{\Gamma_S} \sqrt{(1-\kappa(q) f_-(q))^2 + f'_-(q)^2}\, dq + |D_s f_+|(\Gamma_S) \\ &+ |D_s f_-|(\Gamma_S) + (f_+(q_l) - f_-(q_l) + (f_+(q_r) - f_-(q_r)),\end{aligned}$$

where if Γ_S is a subinterval of Γ, we denote by q_l and q_r its extremes, and if Γ_S is compact, the term $(f_+(q_l) - f_-(q_l)) + (f_+(q_r) - f_-(q_r))$ has to be intended as 0.

Proof. First of all, the fact that $|S^*| \geq |S|$ is obvious, since by definition $S^* \supseteq S$. Concerning the inequality for the perimeter, we start by noticing that, by standard arguments, it is admissible to assume that S is smooth. In fact, by the compactness theorem for BV functions [Ambrosio et al. 2000], we can take a sequence S_j of smooth sets converging in the L^1 sense to S in such a way that $P(S_j) \to P(S)$. By definition, the corresponding sets S_j^* converge to S^*, and by the lower semicontinuity of the perimeter this yields $P(S^*) \leq \liminf P(S_j^*)$. As an immediate consequence, once we establish the validity of this lemma for smooth sets, it will directly follow also in full generality.

The inequality $P(S^*) \leq P(S)$ for smooth sets is very easy to guess, but a bit boring to prove. For simplicity, we will divide the proof into several steps.

Step I. Nonintersecting curves cannot pass "from above to below". In this first step, we underline the following very easy topological fact.

Claim. *Let $\gamma_1, \gamma_2 \subseteq \mathbb{R}^2$ be continuous, nonintersecting plane curves such that $\min \pi_1\gamma_1 = \min \pi_1\gamma_2 = q_0 \in \mathbb{R}$, where $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$ denotes the first projection. If $\max\{t : (q_0, t) \in \gamma_1\} = t_1$ and $\max\{t : (q_0, t) \in \gamma_2\} = t_2$, with $t_1 > t_2$, then $\max\{t : (q, t) \in \gamma_1\} > \max\{t : (q, t) \in \gamma_2\}$ for all $q \in \pi_1\gamma_1 \cap \pi_1\gamma_2$.*

The meaning of this claim is very simple: if one has two continuous and nonintersecting curves in the plane, and the least abscissa of points in the two curves coincide (otherwise, it is obvious that the claim is false), then the curve which starts above always remains above.

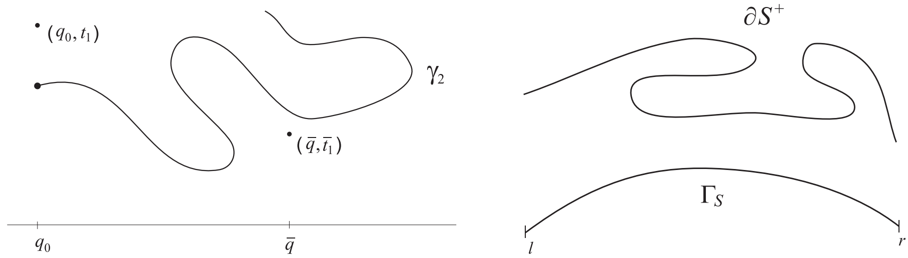


Figure 3. Left: the situation of Step I in the proof of Lemma 2.3. Right: a possible ∂S^+ in Step III.

To show the validity of the claim, suppose it is not true, and let $\bar{q} \in \pi_1\gamma_1 \cap \pi_1\gamma_2$ be a point for which

$$\bar{t}_1 := \max\{t : (\bar{q}, t) \in \gamma_1\} < \max\{t : (\bar{q}, t) \in \gamma_2\} =: \bar{t}_2.$$

Because the curves do not intersect the equality cannot hold true. Figure 3, left, shows the situation.
The curve γ_1, then, is contained by definition in

$$A := \{(q, t) \in \mathbb{R}^2 \setminus \gamma_2 : q \geq q_0, \} \setminus \{(\bar{q}, t) \in \mathbb{R}^2 : t > \bar{t}_1\}. \tag{9}$$

This is a contradiction with the continuity of the curve γ_1, since the points (q_0, t_1) and $(\bar{q}, \bar{t}_1)$ are in γ_1 but belong to two distinct connected components of A. Therefore, the Claim is proved.

Step II. First properties and some definitions. We can immediately observe some simple properties of ∂S and give some related definitions. First of all, since S is smooth, ∂S is the union of finitely many closed curves γ_i, $1 \leq i \leq N$. Exactly one of them, say γ_1, encloses all of S.

Let us then consider the two cases (Γ_S is either compact or not) separately. If Γ_S is not compact and is thus an interval (q_l, q_r), then it is immediate to observe that for every $q \in \Gamma_S$ the points $\mathscr{L}(q, f_\pm(q))$ belong to γ_1. Let us then call ∂S^+ the part of γ_1 starting from $\mathscr{L}(l, f_+(l))$, ending at $\mathscr{L}(r, f_+(r))$ and containing $\mathscr{L}(q, f_+(q))$ for every $q \in \Gamma_S$; similarly, we denote by ∂S^- the part of γ_1 starting from $\mathscr{L}(r, f_-(r))$, ending at $\mathscr{L}(l, f_-(l))$, and containing $\mathscr{L}(q, f_-(q))$ for every $q \in \Gamma_S$. An easy geometric argument ensures that ∂S^+ and ∂S^- are well defined and do not intersect each other. We then obtain

$$\partial S \supseteq \gamma_1 = \partial S^+ \cup \partial S^r \cup \partial S^- \cup \partial S^l, \tag{10}$$

being ∂S^r (respectively ∂S^l) the part of $\gamma \setminus (\partial S^+ \cup \partial S^-)$ connecting $\mathscr{L}(r, f_+(r))$ and $\mathscr{L}(r, f_-(r))$ (respectively $\mathscr{L}(l, f_-(l))$ and $\mathscr{L}(l, f_+(l))$).

Consider now the case when Γ_S is compact ($\Gamma_S = \Gamma$). In this case, we can directly call $\partial S^+ = \gamma_1$, and again it is easy to observe that for every $q \in \Gamma$ one has $\mathscr{L}(q, f_+(q)) \in \partial S^+$. On the other hand, all the points $\mathscr{L}(q, f_-(q))$ belong to the same connected component of ∂S different from γ_1, say γ_2. We then call $\partial S^- = \gamma_2$ and $\partial S^r = \partial S^l = \emptyset$, so that also in this case (10) holds true.

We conclude this step noticing that, for the set S^*, the inclusion (10) is in fact an equality by construction.

Step III. The "upper boundary" is well-ordered. We show that the curve ∂S^+ reaches all the points $\mathscr{L}(q, f_+(q))$ in the "correct order". This means that, if we parametrize ∂S^+ as $\gamma([0, 1])$ with $\gamma(0) = \mathscr{L}(l, f_+(l))$ and $\gamma(1) = \mathscr{L}(r, f_+(r))$, then

(11) If $\gamma(\sigma_1) = \mathscr{L}(q_1, f_+(q_1))$, $\gamma(\sigma_2) = \mathscr{L}(q_2, f_+(q_2))$, then $\sigma_1 < \sigma_2 \Longleftrightarrow q_1 < q_2$.

Notice that this fact is not trivial, since the curve ∂S^+ does not have to be a graph on Γ_S, and, therefore, it can sometimes move to the left, as in Figure 3, right. However, the figure itself suggests that the points $(q, f_+(q))$ are in any case reached "from left to right". Let us now show (11). To do so, suppose by contradiction that it is not true. Hence, there exist σ_1, σ_2, q_1, and q_2 such that $\gamma(\sigma_i) = \mathscr{L}(q_i, f_+(q_i))$ for $i = 1, 2$, but one has $\sigma_1 > \sigma_2$ and $q_1 < q_2$. We can then give the following definitions, π being the projection from Ω to Γ.

$$\begin{aligned}
\sigma_3 &= \min\{\sigma \in (\sigma_1, 1) : \pi(\gamma(\sigma)) = q_2\} \\
q^* &= \min\{\pi(\gamma(\sigma)) : \sigma \in (\sigma_1, \sigma_3)\}, \\
\sigma_0 &= \max\{\sigma \in (0, \sigma_2) : \pi(\gamma(\sigma)) = q^*\}.
\end{aligned}$$

By construction one has $0 < \sigma_0 < \sigma_2 < \sigma_1 < \sigma_3 < 1$, as well as $q^* \le q_1 < q_2$. Now consider the two curves $\gamma_1 = \mathscr{L}^{-1}(\gamma_{|[\sigma_0,\sigma_2]})$ and $\gamma_2 = \mathscr{L}^{-1}(\gamma_{|[\sigma_1,\sigma_3]})$, which are continuous and nonintersecting. Moreover, $\min \pi_1\gamma_1 = \min \pi_1\gamma_2 = q^*$, so we can apply Step I to derive that γ_1 is either "always above" or "always below" γ_2, in the sense of the Claim. By checking $q = q_1$, we observe that γ_1 is below γ_2, since $\max\{\sigma : (q_1, \sigma) \in \gamma_2\} = f_+(q_1)$ is greater than $\max\{\sigma : (q_1, \sigma) \in \gamma_1\}$, by definition of f_+. On the other hand, by checking $q = q_2$, the very same reasoning shows that γ_1 is above γ_2, being $\max\{\sigma : (q_2, \sigma) \in \gamma_1\} = f_+(q_2)$. The contradiction shows the validity of (11).

Step IV. The functions $f_\pm$ are in $BV(\Gamma_S)$. Let us fix an arbitrary $N \in \mathbb{N}$, and an arbitrary sequence $l = q_0 < q_1 < \cdots < q_N < q_{N+1} = r$ in Γ_S. We claim that

$$\sum_{i=0}^{N} |f_+(q_i) - f_+(q_{i+1})| \le \mathscr{H}^1(\partial S^+), \tag{12}$$

$\mathcal{H}^1$ being Hausdorff measure of dimension one (length). Notice that this inequality would show that $f_+ \in BV(\Gamma_S)$, since S is of finite perimeter.

To show the estimate, let us call γ_i the part of the curve ∂S^+ which connects $\mathcal{L}(q_i, f_+(q_i))$ with $\mathcal{L}(q_{i+1}, f_+(q_{i+1}))$. By the preceding steps, we know that ∂S^+ consists of the disjoint union of the curves γ_i, so that

$$\mathcal{H}^1(\partial S^+) = \sum_{i=0}^{N} \mathcal{H}^1(\gamma_i).$$

Hence (12) will follow because for any $i = 0, \ldots, N$, one has

$$\mathcal{H}^1(\gamma_i) \geq \left|\mathcal{L}(q_i, f_+(q_i)) - \mathcal{L}(q_{i+1}, f_+(q_{i+1}))\right| > |f_+(q_i) - f_+(q_{i+1})|. \tag{13}$$

The first inequality is trivial, since it just says that the length of the curve γ_i is greater than the distance of its extreme points. Concerning the strict inequality let us instead use the following notation for brevity.

$$P := \mathcal{L}(q_i, f_+(q_i)), \quad Q := \mathcal{L}(q_{i+1}, f_+(q_{i+1})), \quad Q' := \mathcal{L}(q_i, f_+(q_{i+1})),$$
$$S' := \mathcal{L}(q_i, 0), \quad S := \mathcal{L}(q_{i+1}, 0).$$

Hence, assuming that $f_+(q_i) \geq f_+(q_{i+1}) \geq 0$ (it is then trivial to modify the argument to cover the other cases), one has

$$\overline{PQ'} + \overline{Q'S'} = \overline{PS'} < \overline{PS} < \overline{PQ} + \overline{QS} = \overline{PQ} + \overline{Q'S'},$$

where the first inequality is due to the fact that, by definition, S' is the closest point to P inside Γ. The inequality above says that $\overline{PQ'} < \overline{PQ}$, which is precisely the missing inequality in (13). As explained above, this implies the validity of (12), hence the fact that $f_+ \in BV(\Gamma_S)$.

Of course, the very same argument shows that $f_- \in BV(\Gamma_S)$.

Step V. One has $\mathcal{H}^1(\partial S^+) \geq \mathcal{H}^1(\partial S^{*+})$. Let us define $\{q_i, i \in \mathbb{N}\} \subseteq \Gamma_S$ the jump points of f_+, which are countably many since $f_+ \in BV(\Gamma_S)$. For any i, set

$$f_+^l(q_i) = \lim_{q \uparrow q_i} f_+(q), \quad f_+^r(q_i) = \lim_{q \downarrow q_i} f_+(q).$$

Since $f_+ \in BV(\Gamma_S)$, these two limits exist and correspond to the lim inf and the lim sup of f_+ for $q \to q_i$. In particular, one has that

$$\partial(S^{*+}) = \{\mathcal{L}(q, f_+(q)) : q \in \Gamma_S\} \cup \bigcup_{i \in \mathbb{N}} J_i,$$

where J_i is the segment joining $\mathcal{L}(q_i, f_+^l(q_i))$ and $\mathcal{L}(q_i, f_+^r(q_i))$. Let us now fix $\varepsilon > 0$, so that there exists $N \in \mathbb{N}$ such that

$$\sum_{i>N} |J_i| < \varepsilon.$$

We can assume that the points q_i are ordered so that $l < q_1 < \cdots < q_n < r$. We can now pick, for any $1 \le i \le N$, two points $q_i^l < q_i < q_i^r$ in Γ_S so that

- the different intervals (q_i^l, q_i^r) are disjoint,
- $|f_+(q_i^l) - f_+^l(q_i)| + |f_+(q_i^r) - f_+^r(q_i)| \le \dfrac{\varepsilon}{N}$ for any i, and
- $\mathcal{H}^1\big((\partial S^{*+}) \cap \mathcal{L}((q_i^l, q_i^r) \times (-a, a))\big) \le |J_i| + \dfrac{\varepsilon}{N} = |f_+^l(q_i) - f_+^r(q_i)| + \dfrac{\varepsilon}{N}$.

We now consider the "bad" intervals $B_i = (q_i^l, q_i^r)$, where there are high jumps, and the "good" intervals $G_i = (q_i^r, q_{i+1}^l)$, where there are not. Define also $G_0 = (l, q_1^l)$, while $G_N = (q_N^r, r)$. Therefore, we have decomposed $\Gamma_S = \cup_{i \le N} B_i \cup G_i$. For any good interval G_i, one has

$$\partial S^{*+} \cap \mathcal{L}(G_i \times (-a, a)) = \{\mathcal{L}(q, f_+(q)) : q \in G_i\} \cup \bigcup_{j \in \mathbb{N}} \widetilde{J}_{i,j},$$

where $\widetilde{J}_{i,j}$ are the jumps of f_+ contained in the interval G_i. Of course all the jumps $\widetilde{J}_{i,j}$, varying $0 \le i \le N$ and $j \in \mathbb{N}$, correspond to different jumps J_i for $i > N$. For any bad interval B_i, moreover, call γ_i the part of the curve ∂S^+ from $\mathcal{L}(q_i^l, f_+(q_i^l))$ to $\mathcal{L}(q_i^r, f_+(q_i^r))$. Thanks to Step III, all the curves γ_i are disjoint, and, in particular, $\mathcal{L}(q, f_+(q))$ belongs to γ_i if and only if $q \in B_i$. Since we know that $\mathcal{L}(q, f_+(q)) \in \partial S^+$ for all $q \in \Gamma_S$, this implies that

$$\mathcal{H}^1(\partial S^+) \ge \mathcal{H}^1\Big(\Big\{\mathcal{L}(q, f_+(q)) : q \in \bigcup\nolimits_{i=0}^N G_i\Big\}\Big) + \sum_{i=1}^N \mathcal{H}^1(\gamma_i).$$

Notice also that, as shown with (13) in Step IV, we have that for each $1 \le i \le N$

$$\mathcal{H}^1(\gamma_i) > |f_+(q_i^l) - f_+(q_i^r)|.$$

Finally, using all the properties listed above, we conclude that

$$\begin{aligned}
&\mathcal{H}^1(\partial S^{*+}) \\
&\quad = \sum_{i=0}^N \mathcal{H}^1\Big(\partial S^{*+} \cap \mathcal{L}(G_i \times (-a, a))\Big) + \sum_{i=1}^N \mathcal{H}^1\big(\partial S^{*+} \cap \mathcal{L}(B_i \times (-a, a))\big) \\
&\quad \le \sum_{i=0}^N \bigg(\mathcal{H}^1\big(\{\mathcal{L}(q, f_+(q)) : q \in G_i\}\big) + \sum_{j \in \mathbb{N}} \big|\widetilde{J}_{i,j}\big|\bigg) + \sum_{i=1}^N \Big(|f_+^l(q_i) - f_+^r(q_i)| + \frac{\varepsilon}{N}\Big) \\
&\quad \le \mathcal{H}^1\Big(\Big\{\mathcal{L}(q, f_+(q)) : q \in \bigcup_{i=0}^N G_i\Big\}\Big) + \sum_{i > N} |J_i| + \sum_{i=1}^N \Big(|f_+(q_i^l) - f_+(q_i^r)| + 2\frac{\varepsilon}{N}\Big) \\
&\quad \le \mathcal{H}^1\Big(\Big\{\mathcal{L}\big(q, f_+(q)\big) : q \in \bigcup\nolimits_{i=0}^N G_i\Big\}\Big) + \varepsilon + \sum_{i=1}^N \Big(\mathcal{H}^1(\gamma_i) + 2\frac{\varepsilon}{N}\Big) \\
&\quad \le \mathcal{H}^1(\partial S^+) + 3\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this step is concluded.

Step VI. Conclusion. By (10), we know that

$$\partial S \supseteq \partial S^+ \cup \partial S^- \cup \partial S^l \cup \partial S^r,$$

and the union is disjoint. But, as noticed at the end of Step II, we have

$$\partial S^* = \partial S^{*+} \cup \partial S^{*-} \cup \partial S^{*l} \cup \partial S^{*r}.$$

By Step V we know that $\mathcal{H}^1(\partial S^+) \geq \mathcal{H}^1(\partial S^{*+})$, and, in the same way, we know that $\mathcal{H}^1(\partial S^-) \geq \mathcal{H}^1(\partial S^{*-})$. Let us then focus for a moment on ∂S^l and on ∂S^{*l}. If Γ_S is compact, they are both empty. Otherwise, ∂S^l is a curve between $\mathcal{L}(l, f_-(l))$ and $\mathcal{L}(l, f_+(l))$, while ∂S^{*l} is the segment joining the same points. As a consequence, one has $\mathcal{H}^1(\partial S^l) \geq \mathcal{H}^1(\partial S^{*l})$, and, similarly, $\mathcal{H}^1(\partial S^r) \geq \mathcal{H}^1(\partial S^{*r})$. Adding up the four inequalities, we finally get that $\mathcal{H}^1(\partial S) \geq \mathcal{H}^1(\partial S^*)$.

Concerning formula (8), it is immediate to obtain for smooth functions f_- and f_+, while the generalization for BV functions is standard. □

With the above result at hand, it will be quite easy to obtain the lower bound.

Lemma 2.4 (lower bound). *For a curved strip $\Omega_{\Gamma,a}$ of any kind, one has*

$$h(\Omega_{\Gamma,a}) \geq \frac{1}{a}.$$

Moreover, if the inequality above is an equality and there is a Cheeger set, then this Cheeger set must be $\Omega_{\Gamma,a}$ itself.

Proof. Let S be any open connected set of finite perimeter in $\Omega_{\Gamma,a}$, and let S^* be as in Definition 2.2. Setting

$$t_- := \inf\{f_-(q) : q \in \Gamma_S\},$$
$$t_+ := \sup\{f_+(q) : q \in \Gamma_S\},$$

we can easily estimate

$$\begin{aligned} |S^*| &= \int_\Gamma \int_{f_-(q)}^{f_+(q)} \big(1 - \kappa(q)t\big)\, dt\, dq \\ &= \int_\Gamma \big(f_+(q) - f_-(q)\big)\left(1 - \kappa(q)\frac{f_+(q) + f_-(q)}{2}\right) dq \\ &\leq (t_+ - t_-) \int_\Gamma \left(1 - \kappa(q)\frac{f_+(q) + f_-(q)}{2}\right) dq. \end{aligned} \tag{14}$$

On the other hand, by (8) it is easy to estimate the perimeter of S^* as

$$\begin{aligned} P(S^*) &= \int_{\Gamma_S} \sqrt{(1-\kappa(q) f_+(q))^2 + f_+'(q)^2}\, dq \\ &\quad + \int_{\Gamma_S} \sqrt{(1-\kappa(q) f_-(q))^2 + f_-'(q)^2}\, dq + |D_s f_+|\,(\Gamma_S) \\ &\quad + |D_s f_-|\,(\Gamma_S) + (f_+(q_0) - f_-(q_0)) + (f_+(q_1) - f_-(q_1)) \\ &\geq 2 \int_{\Gamma_S} \left(1 - \kappa(q) \frac{f_+(q) + f_-(q)}{2}\right) dq \end{aligned}$$

simply by neglecting both the absolutely continuous and the singular part of Df. Hence, thanks to Lemma 2.3, we can readily deduce that

$$\frac{P(S)}{|S|} \geq \frac{P(S^*)}{|S^*|} \geq \frac{2}{t_+ - t_-} \geq \frac{1}{a},$$

where the last inequality is due to the trivial bounds $-a \leq t_- < t_+ \leq a$. Finally, if $h(\Omega_{\Gamma,a}) = 1/a$ and there is some Cheeger set $\mathscr{C} = \mathscr{C}_{\Omega_{\Gamma,a}}$, then all the preceding inequalities must be equalities for $S = \mathscr{C}$, from which it immediately follows that f_+ and f_- are constant, and that $t_\pm = \pm a$. Thus $\mathscr{C} = \Omega_{\Gamma,a}$. □

Remark 2.5. As a consequence of (3) for $p = 2$, from the above result we get the lower bound

$$\lambda_2(\Omega_{\Gamma,a}) \geq \frac{1}{4a^2},$$

which is in fact weaker than the bound

$$\lambda_2(\Omega_{\Gamma,a}) \geq \frac{j_{0,1}^2}{4a^2}$$

known from [Exner et al. 2004]. Here, $j_{0,1} \approx 2.4$ denotes the first positive zero of the Bessel function J_0. In fact, an even better bound, reflecting the local geometry of Γ and valid in arbitrary dimensions, is established in [Exner et al. 2004].

Remark 2.6. It is possible to establish the lower bound of Lemma 2.4 directly from Theorem 1.3 without using the "stripization" procedure S^* of Definition 2.2 and its properties stated in Lemma 2.3. Inspired by the formula of [Bellettini et al. 2002, Sec. 11, Ex. 4] for the annulus, we introduce the function $V_t : \Gamma \times (-a, a) \to \mathbb{R}$,

$$V_t(q,t) := \begin{cases} \dfrac{(1-\kappa(q)a)(1+\kappa(q)a) - (1-\kappa(q)t)^2}{2a\kappa(q)(1-\kappa(q)t)} & \text{if } \kappa(q) \neq 0, \\ \dfrac{t}{a} & \text{if } \kappa(q) = 0. \end{cases}$$

The value for vanishing curvature corresponds to taking the limit $\kappa(q) \to 0$ in the formula for positive curvatures. We check that, where the components are considered with respect to the coordinates (q, t), the vector field $V(q,t) := (0, V_t(q,t))$

satisfies $\|V\|_{L^\infty(\Gamma\times(-a,a))} = 1$ and

$$(\operatorname{div} V)(q,t) = \frac{1}{1-\kappa(q)t}\partial_t[(1-\kappa(q)t)V_t(q,t)] = \frac{1}{a}$$

for every $(q,t) \in \Gamma \times (-a,a)$. Hence the searched lower bound is a consequence of Theorem 1.3. However, Lemma 2.3 is needed to establish some finer properties of the Cheeger constant and Cheeger set.

3. The main results

This section is devoted to show our two main results, namely Theorem 3.1, which deals with the case of curved annuli or non-finite curved strips, and Theorem 3.2, which deals with finite curved strips.

The case of a curved annulus and that of a non-finite curved strip.

Theorem 3.1. *Let Γ be compact, infinite, or semi-infinite. Then*

$$h(\Omega_{\Gamma,a}) = \frac{1}{a}. \tag{15}$$

In particular:

(i) *If Γ is compact (that is, $\Omega_{\Gamma,a}$ is a curved annulus), then the infimum of* (1) *is attained and the unique Cheeger set is $\mathscr{C}_{\Omega_{\Gamma,a}} = \Omega_{\Gamma,a}$.*

(ii) *If Γ is infinite or semi-infinite (that is, $\Omega_{\Gamma,a}$ is an infinite or semi-infinite curved strip), then the infimum of* (1) *is not attained, but the sequence $\Omega_{\Gamma_L,a}$ of Lemma 2.1 is an optimizing sequence for $L \to \infty$.*

Proof. The equality (15) follows directly from the upper estimate of Lemma 2.1 and the lower estimate of Lemma 2.4.

From the characterization of Lemma 2.4, moreover, we know that the unique possible Cheeger set is the whole $\Omega_{\Gamma,a}$. Since this set has an infinite area and perimeter in the case of an infinite or semi-infinite curved strip, we get the nonexistence result of a minimizer for the case (ii), while the fact that $\Omega_{\Gamma_L,a}$ is a minimizing sequence for $L \to \infty$ follows by Lemma 2.1. On the other hand, in case (i) we know by compactness that some Cheeger set must exist, hence the existence and uniqueness of the whole $\Omega_{\Gamma,a}$ as a Cheeger set again comes by Lemma 2.4. □

The case of a finite curved strip.

Theorem 3.2. *Let Γ be noncomplete and finite (hence, $\Omega_{\Gamma,a}$ is a finite curved strip). Then there exists a positive universal constant c such that*

$$\frac{1}{a} + \frac{c}{|\Gamma|} \leq h(\Omega_{\Gamma,a}) \leq \frac{1}{a} + \frac{2}{|\Gamma|}. \tag{16}$$

For instance, one may take $c = \frac{1}{400}$. *Moreover, the infimum in* (1) *is attained for some connected set* $\mathscr{C}_{\Omega_{\Gamma,a}} \subsetneq \Omega_{\Gamma,a}$.

Proof. Concerning the existence of a Cheeger set $\mathscr{C} = \mathscr{C}_{\Omega_{\Gamma,a}}$, and in particular of a connected one, this follows by Theorem 1.1. From the same Theorem, we know also that $\partial\mathscr{C} \cap \Omega_{\Gamma,a}$ is made by arcs of circle of radius $h(\Omega_{\Gamma,a})^{-1}$, and, again by Theorem 1.1, it cannot coincide with the whole set $\Omega_{\Gamma,a}$, since $\mathscr{C}$ may not have corners. As a consequence, by the characterization of Lemma 2.4 we deduce that $h(\Omega_{\Gamma,a}) > 1/a$. To conclude, we have then only to give a proof of the bounds (16), which will be done in several steps.

Step I. The upper bound. Obtaining the upper bound is very easy: it is enough to recall that

$$P(\Omega_{\Gamma,a}) = 2|\Gamma| + 4a, \quad |\Omega_{\Gamma,a}| = 2a|\Gamma|,$$

which was already checked, for instance in (7), and then

$$h(\Omega_{\Gamma,a}) \leq \frac{P(\Omega_{\Gamma,a})}{|\Omega_{\Gamma,a}|} = \frac{2|\Gamma|+4a}{2a|\Gamma|} = \frac{1}{a} + \frac{2}{|\Gamma|}.$$

Step II. The lower bound: the behavior of the arcs of $\partial\mathscr{C} \cap \Omega_{\Gamma,a}$. Thanks to Theorem 1.1, we know that $\partial\mathscr{C}$ cannot have corners. Hence $\partial\mathscr{C} \cap \Omega_{\Gamma,a}$ is not empty, and it is done by some arcs of circle, all of radius $1/h(\Omega_{\Gamma,a})$ and hence strictly smaller than a as noticed above, such that all four corners of $\Omega_{\Gamma,a}$ are ruled out from $\mathscr{C}$. Denoting by q_0 and q_1 the extreme points of Γ, let us call for simplicity "up", "down", "left", and "right" the four parts of $\partial\Omega_{\Gamma,a}$ given by the points of the form $\mathscr{L}(q, a)$, $\mathscr{L}(q, -a)$, $\mathscr{L}(q_0, t)$, and $\mathscr{L}(q_1, t)$ for $q \in \Gamma$ and $t \in (-a, a)$ respectively. We aim to show this:

Claim. *All the arcs of circle of* $\partial\mathscr{C} \cap \Omega_{\Gamma,a}$ *connect two points of* $\Omega_{\Gamma,a}$, *at least one of which is either in the left or in the right part.*

To show this claim, we have to exclude the case of an arc of circle starting and ending in the upper part, and the case of an arc connecting the up and the down (the case of an arc starting and ending in the bottom part is exactly the same as the first one).

Suppose first that there is an arc of circle connecting the points P and Q, both in the upper part. Thus $P = \mathscr{L}(q', a)$ and $Q = \mathscr{L}(q'', a)$. By Theorem 1.1, we know that the circle is tangent to $\partial\Omega_{\Gamma,a}$ at P and Q, hence its center O is the intersection between the two lines which are normal to $\partial\Omega_{\Gamma,a}$ at P and Q, which are $t \mapsto \mathscr{L}(q', t)$ and $t \mapsto \mathscr{L}(q'', t)$. Since the radius r of these circles is at most a, the two lines must intersect at the point

$$\mathscr{L}(q', a - r) \equiv \mathscr{L}(q'', a - r),$$

while this is impossible for any $r \leq 2a$ because $\mathscr{L}$ is one-to-one.

A very similar argument works assuming that an arc of circle connects the point $P = \mathscr{L}(q', a)$ in the upper part with the point $Q = \mathscr{L}(q'', -a)$ in the lower part. Indeed, again the circle would be tangent to $\partial\Omega_{\Gamma,a}$ at both P and Q, so that its center would be at the intersection of the segments $t \mapsto \mathscr{L}(q', t)$ and $t \mapsto \mathscr{L}(q'', t)$. This is impossible if the circle has radius smaller than $2a$ for $q' \neq q''$, but it is also impossible for a radius strictly smaller than a in the case $q' = q''$. This shows the Claim.

Notice now that by definition the left and the right parts of $\partial\Omega_{\Gamma,a}$ are segments, so the case of a circle starting and ending on the left is impossible, as is an arc starting and ending on the right. In conclusion, we now know that there can be either 2, 3, or 4 arcs of circle in $\partial\mathscr{C}$. The simplest case is when there are four arcs, each of which make a "rounded corner". This happens, for instance, for a rectangle (that is, if Γ is a segment), and more generally if a is sufficiently small with respect to $|\Gamma|$. However, it is also possible that there are only three arcs, one of which connects the left and the right part of the boundary. This happens, for instance, whenever the upper or lower part of the boundary is very short due to a big (but still admissible) curvature of Γ. An example of this situation is a sector of an annulus with very small inner radius, which then is very similar to a triangle: in this case the boundary of $\mathscr{C}$ does not touch the inner circle (some examples of this kind are shown in the next section). As to the last possibility (only two arcs of circle both connecting left and right), we have no example in mind and it may be impossible, but we do not need to exclude this case within this proof. Indeed, in Steps III and IV we will show the theorem in the case of four rounded corners, while in Step V we will show how it is always possible to reduce to this case.

Step III. The lower bound: the case when $\mathscr{C}$ has four rounded corners; statement of the Claim (18). To show the lower bound, we start from the case when $\mathscr{C}$ has four rounded corners. Let us recall that, as shown by (8) in Lemma 2.3, one has

$$
\begin{aligned}
(17) \qquad P(\mathscr{C}) = &\int_\Gamma \sqrt{(1-\kappa(q) f_+(q))^2 + f_+'(q)^2}\, dq \\
&+ \int_\Gamma \sqrt{(1-\kappa(q) f_-(q))^2 + f_-'(q)^2}\, dq + |D_s f_+|(\Gamma) \\
&+ |D_s f_-|(\Gamma) + (f_+(q_0) - f_-(q_0)) + (f_+(q_1) - f_-(q_1)),
\end{aligned}
$$

where $f_\pm'\, dq$ is the absolute continuous part of $Df_\pm$ and $D_s f_\pm$ its singular part (notice that, in the language of Definition 2.2, we have $\Gamma_{\mathscr{C}} = \Gamma$ thanks to Step II). In particular,

$$
P(\mathscr{C}) \geq \int_\Gamma \big(2 - \kappa(q)(f_+(q) + f_-(q))\big)\, dq.
$$

We claim that, at least in the case when $\mathscr{C}$ has four corners,

$$P(\mathscr{C}) \geq \int_\Gamma \big(2-\kappa(q)(f_+(q)+f_-(q))\big)\,dq + \tfrac{1}{50}a. \tag{18}$$

We will prove this estimate in the next step. Now we show how this implies the thesis. We can easily estimate the area of $\mathscr{C}$ as in (14):

$$\begin{aligned} |\mathscr{C}| &= \int_\Gamma \int_{f_-(q)}^{f_+(q)} (1-\kappa(q)t)\,dt\,dq \\ &= \int_\Gamma \big(f_+(q)-f_-(q)\big)\bigg(1-\kappa(q)\frac{f_+(q)+f_-(q)}{2}\bigg)\,dq \\ &\leq a\int_\Gamma \big(2-\kappa(q)(f_+(q)+f_-(q))\big)\,dq. \end{aligned}$$

Hence, using (18), we get (16) because, recalling that $a\|\kappa\|_{L^\infty(\Gamma)} \leq 1$ (as pointed out in Section 1),

$$\begin{aligned} h(\Omega_{\Gamma,a}) = \frac{P(\mathscr{C})}{|\mathscr{C}|} &\geq \frac{\int_\Gamma \big(2-\kappa(q)(f_+(q)+f_-(q))\big)\,dq + \frac{1}{50}a}{a\int_\Gamma \big(2-\kappa(q)(f_+(q)+f_-(q))\big)\,dq} \\ &= \frac{1}{a} + \frac{1}{50\int_\Gamma \big(2-\kappa(q)(f_+(q)+f_-(q))\big)\,dq} \\ &\geq \frac{1}{a} + \frac{1}{50|\Gamma|\big(2+2a\|\kappa\|_{L^\infty(\Gamma)}\big)} \geq \frac{1}{a} + \frac{1}{200|\Gamma|}. \end{aligned} \tag{19}$$

Step IV. The lower bound: the case when $\mathscr{C}$ has four rounded corners; proof of Claim (18). We show that, assuming that $\partial\mathscr{C} \cap \Omega_{\Gamma,a}$ consists of four arcs of circle, (18) holds. This will be done by considering a single arc. To choose it, we start by noticing that (17) already trivially implies (18) if

$$f_+(q_1) - f_-(q_1) \geq \tfrac{1}{50}a.$$

As a consequence, we can assume that

$$f_+(q_1) \leq \tfrac{1}{100}a, \tag{20}$$

and we concentrate on the arc of circle corresponding to the "upper right corner". Of course, if (20) were not true, one could assume $f_-(q_1) \geq -a/100$ and then make the completely symmetric considerations on the "lower right corner". As shown in Figure 4, we call the arc of circle that we are considering γ, and we can also look at γ in the reference rectangle, where of course it is no longer part of a circle. We define the length η as in Figure 4.

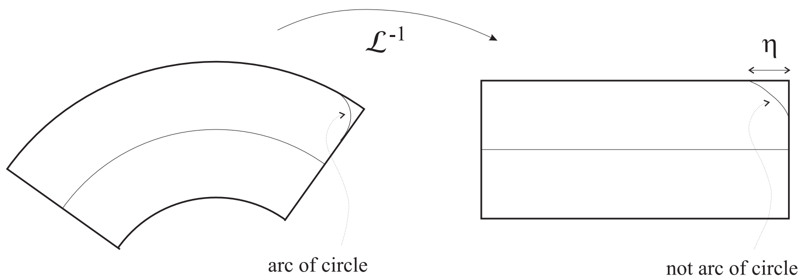


Figure 4. The situation (both in Ω and in the reference configuration) of Step IV.

We denote by Γ_γ the part of Γ related to the curve γ. Γ_γ is therefore the subset of Γ such that

$$\gamma = \{(q, f_+(q)) : q \in \Gamma_\gamma\},$$

we can subdivide Γ_γ into two parts, namely

$$\Gamma_1 := \left\{q \in \Gamma_\gamma : |f_+'(q)| < \tfrac{1}{5}\right\}, \quad \Gamma_2 := \left\{q \in \Gamma_\gamma : |f_+'(q)| \geq \tfrac{1}{5}\right\}.$$

Notice that the above subdivision makes sense because f_+ has no singular part inside Γ_γ (since the image of its graph under $\mathscr{L}$ is an arc of circle). By definition,

$$\int_{\Gamma_1} |f_+'(q)| dq \leq \tfrac{1}{5}|\Gamma_1| \leq \tfrac{1}{5}\eta \leq \tfrac{2}{5}a. \tag{21}$$

In the last inequality we used the fact that $\eta \leq 2a$, which is true because γ is an arc of circle or radius smaller than a (keep in mind that by Lemma 2.4 we already know that $h(\Omega_{\Gamma,a}) \geq 1/a$) and the lengths in the reference rectangle are at most the double of the true lengths (recall also that by Theorem 1.1 the arcs of circle touch $\partial\Omega_{\Gamma,a}$ tangentially). But then, thanks to (20), one has

$$\int_{\Gamma_\gamma} \left|f_+'(q)\right| dq \geq \tfrac{99}{100}a,$$

so that by (21) we get

$$\int_{\Gamma_2} \left|f_+'(q)\right| dq \geq \left(\tfrac{99}{100} - \tfrac{2}{5}\right)a = \tfrac{59}{100}a. \tag{22}$$

Recalling again that $0 < 1 - \kappa(q) f_+(q) < 2$, a trivial calculation ensures that for any $q \in \Gamma_2$

$$\sqrt{(1 - \kappa(q) f_+(q))^2 + f_+'(q)^2} \geq (1 - \kappa(q) f_+(q)) + \tfrac{1}{25}|f_+'(q)|. \tag{23}$$

Hence, thanks to (22) and (23), we can estimate the length of γ as

$$\begin{aligned}|\gamma| &= \int_{\Gamma_\gamma} \sqrt{\big(1-\kappa(q) f_+(q)\big)^2 + f_+'(q)^2}\, dq \\ &\geq \int_{\Gamma_\gamma} \big(1-\kappa(q) f_+(q)\big)\, dq + \tfrac{1}{25} \int_{\Gamma_2} \big|f_+'(q)\big|\, dq \\ &\geq \int_{\Gamma_\gamma} \big(1-\kappa(q) f_+(q)\big)\, dq + \frac{59}{25\cdot 100} a.\end{aligned}$$

Recalling formula (17) for the perimeter of $\mathscr{C}$, we finally conclude that

$$P(\mathscr{C}) \geq \int_{\Gamma} \big(2-\kappa(q)(f_+(q)+f_-(q))\big)\, dq + \frac{59}{25\cdot 100} a,$$

thus finally getting (18).

Step V. The lower bound: general case. In this last step we conclude the proof of the theorem. Thanks to the above steps, we already know that the result holds in the case of four rounded corners, so we can now assume that $\partial\mathscr{C}$ has only two or three arcs. In this case, there exist two maximal numbers $a^\pm \leq a$ such that $\mathscr{C} \subseteq \mathscr{L}(\Gamma \times (-a^-, a^+))$. Let us now introduce a new strip $\Omega_{\widetilde{\Gamma},\tilde{a}}$:

$$\bar{t} := \frac{a^+ - a^-}{2}, \quad \widetilde{\Gamma} := \mathscr{L}(\Gamma \times \{\bar{t}\}), \quad \tilde{a} := \frac{a^+ + a^-}{2}.$$

Notice that there is a bijective map $\varphi : \Gamma \to \widetilde{\Gamma}$ given by $\varphi(q) = \mathscr{L}(q, \bar{t})$, and that since $\widetilde{\Gamma}$ is parallel to Γ by construction, the normal vector $N(q)$ to Γ at q coincides with the normal vector $\widetilde{N}(\varphi(q))$ to $\widetilde{\Gamma}$ at $\varphi(q)$. Thus, $\Omega_{\widetilde{\Gamma},\tilde{a}}$ being a subset of $\Omega_{\Gamma,a}$, the injectivity condition (5) trivially holds for $\widetilde{\Gamma}$ and $\tilde{a}$, and we can conclude that the strip $\Omega_{\widetilde{\Gamma},\tilde{a}}$ is admissible for our purposes.

By construction, we have $\mathscr{C} \subseteq \Omega_{\widetilde{\Gamma},\tilde{a}}$, so $\mathscr{C}$ is also the Cheeger set of $\Omega_{\widetilde{\Gamma},\tilde{a}}$. Moreover, by maximality of $a^\pm$ we know that $\mathscr{C}$ touches all four parts of the boundary of $\Omega_{\widetilde{\Gamma},\tilde{a}}$, so the preceding steps, and in particular (19), allow us to deduce that

$$h(\Omega_{\Gamma,a}) = \frac{P(\mathscr{C})}{|\mathscr{C}|} = h(\Omega_{\widetilde{\Gamma},\tilde{a}}) \geq \frac{1}{\tilde{a}} + \frac{1}{200|\widetilde{\Gamma}|}.$$

Finally, by definition $\tilde{a} \leq a$, while

$$|\widetilde{\Gamma}| = \int_{\Gamma} \big(1-\bar{t}\kappa(q)\big)\, dq \leq 2|\Gamma|.$$

Thus, we get (16) with the constant $c = \frac{1}{400}$. □

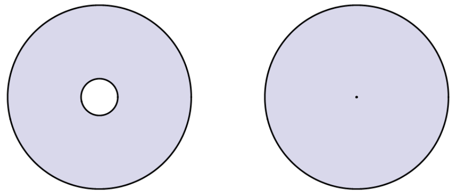


Figure 5. The annulus and the disc considered as its limit case.

4. Solvable models

In this section we discuss our results on the basis of several examples of curved strips about circles and circular arcs. They are referred to as *solvable models* since the determination of the Cheeger constant and the Cheeger set is reduced to solving an explicit algebraic equation. Where the exact solution is not available, we have solved the problem with the help of standard numerical tools.

Annuli. Probably the simplest example is given by annuli, that is, strips built about (full) circles; see Figure 5. Then the Cheeger set is the strip itself and the Cheeger constant equals half of the distance between the boundary curves. It follows from Theorem 3.1 that exactly the same situation holds for general curved annuli. Let us remark that discs can also be thought of as examples of curved strips. Indeed, a disc with its central point removed has the same Cheeger set (up to the point) and Cheeger constant as the disc, and the former set can be considered as the limit case of the annulus built about the circle of radius $a+\varepsilon$ when $\varepsilon \to 0+$.

Rectangles. The rectangle $\mathcal{R}_{a,b} := (-b, b) \times (-a, a)$, with $a, b > 0$, can be considered as a strip built about the segment $\Gamma := (-b, b) \times \{0\}$. Using Theorem 1.2, it is easy to find its Cheeger constant explicitly:

$$h(\mathcal{R}_{a,b}) = \frac{a+b+\sqrt{(a-b)^2+\pi ab}}{2ab}.$$

Notice the scaling $h(\mathcal{R}_{a,b}) = h(\mathcal{R}_{1,b/a})$. The procedure also determines the Cheeger set of $\mathcal{R}_{a,b}$ as the rectangle with its corners rounded off by circular arcs of radius $h(\mathcal{R})^{-1}$; see Figure 6.

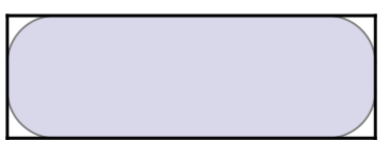


Figure 6. The rectangle and its Cheeger set (light gray) for $b/a = 3$.

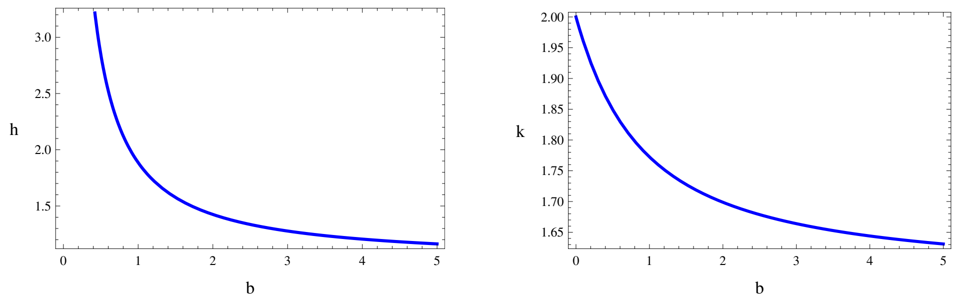


Figure 7. The Cheeger constant h and the quantity k for rectangles with $a = 1$.

The Cheeger constant can be written as

$$h(\mathcal{R}_{a,b}) = \frac{1}{a} + \frac{k(a,b)}{|\Gamma|}, \quad \text{where } k(a,b) := \frac{a - b + \sqrt{(a-b)^2 + \pi ab}}{a}$$

and $|\Gamma| = 2b$. Notice the scaling $k(a,b) = k(1, b/a)$. It is straightforward to check that $b/a \mapsto k(a,b)$ is a decreasing function with the limits $k(a,b) \to 2$ as $b/a \to 0$ and $k(a,b) \to \pi/2$ as $b/a \to \infty$. Hence the upper bound of Theorem 3.2 becomes sharp in the limit of very narrow rectangles. The dependence of the Cheeger constant h and of the quantity k on rectangle parameters is shown in Figure 7.

Sectors. Let Γ_a be the circle of curvature $\kappa = a^{-1}$ and consider its part Γ_a^α of length $|\Gamma_a^\alpha| = \alpha a$, with any $\alpha \in (0, 2\pi)$; see Figure 8. The corresponding strip $\Omega_a^\alpha := \Omega_{\Gamma_a^\alpha, a}$ does not satisfy the assumption (5). However, since $\mathcal{L}$ is in fact injective in $\Gamma_a \times (-a, a)$, it can be considered as a limit case of admissible strips along corresponding parts of the circle of radius $a + \varepsilon$ when $\varepsilon \to 0+$.

The Cheeger constant and the Cheeger set of Ω_a^α can be found as follows. First, we construct a family of domains S_r, with $r \in (0, a)$, defined by rounding off the corners in Ω_a^α of angle smaller than π by circular arcs of radius r. This can be done by a straightforward usage of elementary geometric rules. Secondly, we

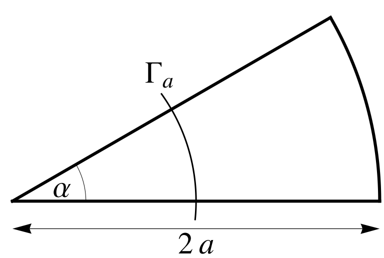


Figure 8. The sector of a disc considered as a strip built about the $(\frac{\alpha}{2\pi})$-th part of a circle.

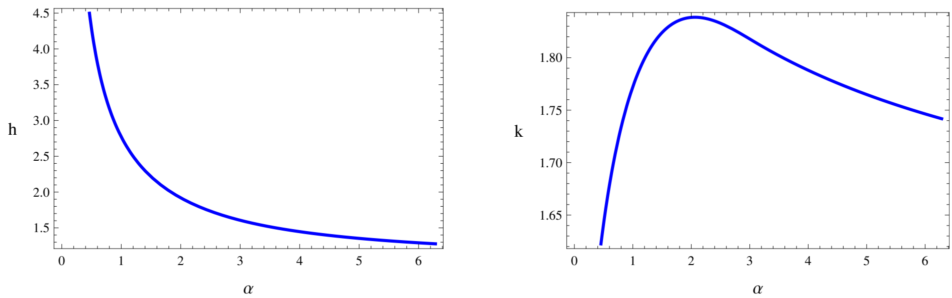


Figure 9. The Cheeger constant h and the constant k for sectors with $a = 1$.

minimize the quotient $P(S_r)/|S_r|$ with respect to r, which is done with the help of a numerical optimization. The minimum of the quotient corresponds to the Cheeger constant, and the minimizer is the Cheeger set. The procedure is equivalent to using Theorem 1.2, which seems to remain valid also for $\alpha > \pi$, corresponding to nonconvex sectors.

In view of the obvious scaling $h(\Omega_a^\alpha) = h(\Omega_1^\alpha)/a$, one can set $a = 1$, without loss of generality. The dependence of the Cheeger constant on α is shown in Figure 9. Table 1 contains numerical values for some specific angles.

Writing the Cheeger constant as

$$h(\Omega_a^\alpha) = \frac{1}{a} + \frac{k(\alpha)}{|\Gamma_a^\alpha|},$$

we also study the dependence of the constant $k(\alpha)$ on α; see Figure 9 and Table 1. The third value of α in Table 1 corresponds to the maximal point of the curve $\alpha \mapsto k(\alpha)$ from Figure 9. In any case, we see that the upper bound of Theorem 3.2 is quite good for all the sectors. Finally, Figure 10 shows a numerical approximation of the Cheeger sets for some annuli.

α	$\pi/10$	$\pi/2$	$0.656749\,\pi$	$3\pi/4$	π	$3\pi/2$	2π
$h(\Omega_1^\alpha)$	5.92687	2.16358	1.89111	1.77915	1.57714	1.37582	1.27722
$k(\Omega_1^\alpha)$	1.54782	1.82774	1.83856	1.83583	1.81315	1.77101	1.74184

Table 1. The Cheeger constant h and the constant k for sectors with $a = 1$.

Acknowledgements

The authors acknowledge the hospitality of the Nuclear Physics Institute ASCR in Řež and the University of Pavia.

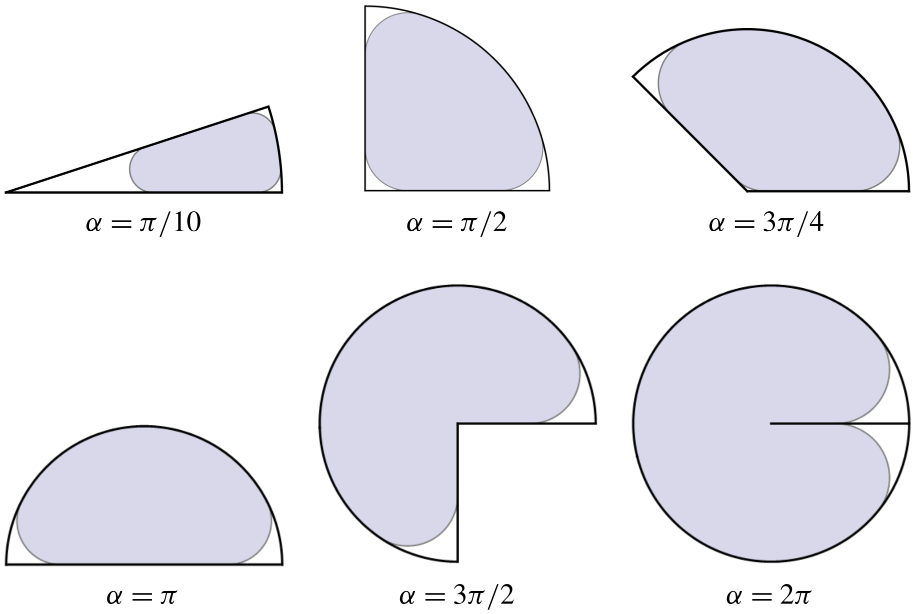


Figure 10. The sectors and their Cheeger sets (light gray).

References

[Alter et al. 2005] F. Alter, V. Caselles, and A. Chambolle, "Evolution of characteristic functions of convex sets in the plane by the minimizing total variation flow", *Interfaces Free Bound.* **7**:1 (2005), 29–53. MR 2006b:35154 Zbl 1084.49003

[Ambrosio et al. 2000] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, New York, 2000. MR 2003a:49002 Zbl 0957.49001

[Bellettini et al. 2002] G. Bellettini, V. Caselles, and M. Novaga, "The total variation flow in $\mathbb{R}^N$", *J. Differential Equations* **184**:2 (2002), 475–525. MR 2003g:35105 Zbl 1036.35099

[Caselles et al. 2007] V. Caselles, A. Chambolle, and M. Novaga, "Uniqueness of the Cheeger set of a convex body", *Pacific J. Math.* **232**:1 (2007), 77–90. MR 2008j:49012 Zbl 1221.35171

[Duclos and Exner 1995] P. Duclos and P. Exner, "Curvature-induced bound states in quantum waveguides in two and three dimensions", *Rev. Math. Phys.* **7**:1 (1995), 73–102. MR 95m:81239 Zbl 0837.35037

[Evans and Gariepy 1992] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992. MR 93f:28001 Zbl 0804.28001

[Exner et al. 2004] P. Exner, P. Freitas, and D. Krejčiřík, "A lower bound to the spectral threshold in curved tubes", *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460**:2052 (2004), 3457–3467. MR 2006d:81100 Zbl 02185441

[Grieser 2006] D. Grieser, "The first eigenvalue of the Laplacian, isoperimetric constants, and the max flow min cut theorem", *Arch. Math.* (*Basel*) **87**:1 (2006), 75–85. MR 2007e:35213 Zbl 1105.35062

[Hebey and Saintier 2006] E. Hebey and N. Saintier, "Stability and perturbations of the domain for the first eigenvalue of the 1-Laplacian", *Arch. Math.* (*Basel*) **86**:4 (2006), 340–351. MR 2007b:35245 Zbl 1099.58012

[Kawohl and Fridman 2003] B. Kawohl and V. Fridman, "Isoperimetric estimates for the first eigenvalue of the p-Laplace operator and the Cheeger constant", *Comment. Math. Univ. Carolin.* **44**:4 (2003), 659–667. MR 2005g:35053 Zbl 1105.35029

[Kawohl and Lachand-Robert 2006] B. Kawohl and T. Lachand-Robert, "Characterization of Cheeger sets for convex subsets of the plane", *Pacific J. Math.* **225**:1 (2006), 103–118. MR 2007e:52002 Zbl 1133.52002

[Krejčiřík and Kříž 2005] D. Krejčiřík and J. Kříž, "On the spectrum of curved planar waveguides", *Publ. Res. Inst. Math. Sci.* **41**:3 (2005), 757–791. MR 2006k:35244 Zbl 1113.35143

[Saintier 2007] N. Saintier, "Shape derivative of the first eigenvalue of the 1-Laplacian", preprint, 2007. arXiv 0706.0873v1

Received November 15, 2010. Revised October 2, 2011.

DAVID KREJČIŘÍK
DEPARTMENT OF THEORETICAL PHYSICS
NUCLEAR PHYSICS INSTITUTE ASCR
25068 ŘEŽ
CZECH REPUBLIC
krjcirik@ujf.cas.cz

ALDO PRATELLI
DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DI PAVIA
VIA FERRATA, 1
I-27100 PAVIA
ITALY
aldo.pratelli@unipv.it

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

STRUCTURE OF SOLUTIONS OF 3D AXISYMMETRIC NAVIER–STOKES EQUATIONS NEAR MAXIMAL POINTS

ZHEN LEI AND QI S. ZHANG

Let v be a solution of the axially symmetric Navier–Stokes equation. We determine the structure of a certain (possible) maximal singularity of v in the following sense. Let (x_0, t_0) be a point where the flow speed $Q_0 = |v(x_0, t_0)|$ is comparable with the maximum flow speed at and before time t_0. We show, after a space-time scaling with the factor Q_0 and the center (x_0, t_0), that the solution is arbitrarily close in $C^{2,1,\alpha}_{\text{local}}$ norm to a nonzero constant vector in a fixed parabolic cube, provided that $r_0 Q_0$ is sufficiently large. Here r_0 is the distance from x_0 to the z axis. Similar results are also shown to be valid if $|r_0 v(x_0, t_0)|$ is comparable with the maximum of $|r v(x, t)|$ at and before time t_0. This mirrors a numerical result of Hou for the Euler equation: there exists a certain "calm spot" or depletion of vortex stretching in a region of high flow speed.

1. Introduction

We study the structure, in a space-time region with maximum flow speed, of solutions to the three-dimensional incompressible Navier–Stokes equations

$$\begin{cases} \partial_t v + v \cdot \nabla v + \nabla p = \mu \Delta v, \\ \nabla \cdot v = 0, \end{cases} \qquad t \geq 0, \quad x \in \mathbb{R}^3, \tag{1-1}$$

with the axially symmetric initial data

$$v_0(x) = a^r(r, z, t)e_r + a^\theta(r, z, t)e_\theta + a^z(r, z, t)e_z. \tag{1-2}$$

In cylindrical coordinates, the solution $v = v(x, t)$ is of the form

$$v(x, t) = v^r(r, z, t)e_r + v^\theta(r, z, t)e_\theta + v^z(r, z, t)e_z. \tag{1-3}$$

Here $x = (x_1, x_2, z)$ and $r = \sqrt{x_1^2 + x_2^2}$, while

$$e_r = \begin{pmatrix} x_1/r \\ x_2/r \\ 0 \end{pmatrix}, \quad e_\theta = \begin{pmatrix} -x_2/r \\ x_1/r \\ 0 \end{pmatrix}, \quad e_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{1-4}$$

MSC2010: 35Q30, 76D05.
Keywords: Axisymmetric Navier–Stokes equations, structure of singularities.

are the three orthogonal unit vectors along the radial, angular, and axial directions. Also, the angular, swirl and axial components v^r, v^θ and v^z of the velocity field are solutions of the axially symmetric Navier–Stokes equations (or ASNS)

$$
\begin{cases}
\partial_t v^r + b\cdot\nabla v^r - \dfrac{(v^\theta)^2}{r} + \partial_r p = \Big(\Delta - \dfrac{1}{r^2}\Big)v^r,\\
\partial_t v^\theta + b\cdot\nabla v^\theta + \dfrac{v^r v^\theta}{r} = \Big(\Delta - \dfrac{1}{r^2}\Big)v^\theta,\\
\partial_t v^z + b\cdot\nabla v^z + \partial_z p = \Delta v^z,\\
b = v^r e_r + v^z e_z, \quad \nabla\cdot b = \partial_r v^r + \dfrac{v^r}{r} + \partial_z v^z = 0.
\end{cases}
\tag{1-5}
$$

Here, without loss of generality, we set the viscosity constant μ equal to 1.

Although the axially symmetric case is a special instance of the full Navier–Stokes equations, the main regularity problem is just as wide open. Let us briefly discuss some interesting results on the axially symmetric Navier–Stokes equations. When $v^\theta = 0$, that is, in the no swirl case, Ladyzhenskaya [1968] and Ukhovskii and Iudovich [1968] proved that weak solutions are regular for all time. See also [Leonardi et al. 1999]. More recent activity, in the presence of swirl, includes [Chen et al. 2008; 2009], where it is proven that suitable axially symmetric solutions bounded by $Cr^{-\alpha}\sqrt{|t|}^{-1+\alpha}$ $(0 \le \alpha \le 1)$ are smooth. Here, r is the distance from a point to the z axis, and t is time. See also [Koch et al. 2009] and its local version using different methods, [Seregin and Šverák 2009]. Also in the presence of swirl, there is [Neustupa and Pokorný 2000], proving that regularity of one component (either v^r or v^θ) implies regularity of the other components of the solution. Also proving regularity, under an assumption of sufficiently small zero-dimension scaled norms, is [Jiu and Xin 2003].

We also wish to mention the regularity results of Chae and Lee [2002], who prove regularity results assuming finiteness of another zero-dimensional integral. On the other hand, Tian and Xin [1998] constructed a family of singular axis symmetric solutions with singular initial data, and Hou and Li [2008] found a special class of global smooth solutions. See also the recent extension [Hou et al. 2008].

In this paper, we take another approach to ASNS, seeking to understand the local structure of solutions when the flow velocity is very high. This is akin to the approach taken by Hamilton and Perelman in the study of Ricci flow. We can reach understanding when the flow speed $|v(x_0, t_0)|$ at a space-time point (x_0, t_0) is comparable with the maximum flow speed, or $r_0|v(x_0, t_0)|$ at a space-time point (x_0, t_0) is comparable with the maximum of $r|v(x, t)|$, at and before time t_0.

In order to present the result, we introduce some notations. Let $v = v(x, t)$ be a solution to ASNS. Here (x, t) is a point in space-time. Given a number $a > 0$ and a point in space-time (x_0, t_0), we define the parabolic cube

$$P(x_0, t_0, a) \equiv \{(x, t) \ : \ |x_0 - x| < a, \ t_0 - a^2 \le t \le t_0\}.$$

Unless stated otherwise, we use r, r_0, r_k to denote the distance between points x, x_0, x_k in space and the z-axis, respectively.

Now we are ready to state the main result of the paper.

Theorem 1.1. *Let $v = v(x, t)$, with $(x, t) \in \mathbb{R}^3 \times [0, T_0)$ and $T_0 > 0$, be a smooth solution to the three-dimensional ASNS, with initial condition v_0 satisfying*

$$\|v_0\|_{L^\infty(\mathbb{R}^3)} \le N_0, \quad \|v_0\|_{L^2(\mathbb{R}^3)} \le N_0, \quad |r v_0| \le N_0. \tag{1-6}$$

Here N_0 is any positive number. For any sufficiently small constant $\epsilon > 0$ and two other constants $\sigma_0 > 0$ and $0 < \alpha < 1$, there exists some $\rho_0 = \rho_0(\epsilon, N_0, \sigma_0, \alpha) > 0$ with the following properties.

(a) *Suppose*

$$r_0 |v(x_0, t_0)| \ge \rho_0^{-2}$$

at some point (x_0, t_0), where $x_0 \in \mathbb{R}^3$ and $t_0 \in (0, T_0)$. Suppose also that (x_0, t_0) is an almost maximal point in the sense that

$$|v(x_0, t_0)| \ge \tfrac{1}{4} \sup_{x \in \mathbb{R}^3, \ t \le t_0} |v(x, t)|.$$

Then the velocity v in the cube

$$P(x_0, t_0, (\sigma_0 \epsilon Q)^{-1}), \qquad Q \equiv |v(x_0, t_0)|,$$

after scaling by the factor Q, that is, $Q^{-1} v(Q^{-1}x + x_0, Q^{-2}t + t_0)$, is ϵ-close in $C^{2,1,\alpha}_{\text{local}}$ norm to a nonzero constant vector.

(b) *The conclusion in* (*a*) *still holds if*

$$r_0 |v(x_0, t_0)| \ge \rho_0^{-2}$$

at (x_0, t_0) and

$$r_0 |v(x_0, t_0)| \ge \tfrac{1}{4} \sup_{x \in \mathbb{R}^3, \ t \le t_0} r |v(x, t)|.$$

Remark 1.2. According to [Seregin and Šverák 2009] and [Chen et al. 2009], if a smooth solution blows up in finite time, then the scaling invariant quantity $r|v(x, t)|$ must also blow up in finite time near singularity. So the condition in (b) can always be satisfied if the solution develops finite time singularity.

Remark 1.3. The factor $\frac{1}{4}$ in the statement of the theorem can be replaced by any fixed positive number smaller than or equal to 1. In particular, the statement is true if (x_0, t_0) is a point such that $r_0 |v(x_0, t_0)| = \sup_{x \in \mathbb{R}^3, \, t \le t_0} r |v(x, t)|$.

An important open question is to generalize the current result in (a) to the case when $|v(x_0, t_0)|$ is very large but still much smaller than maximum.

Another question is: what happens when $r_0|v(x_0, t_0)|$ is not large, but $|v(x_0, t_0)|$ is large at almost maximal point (x_0, t_0)?

Remark 1.4. The result and parameters in the theorem depend only on the norms of the initial value in (1-6). They do not depend on individual solutions.

We end the introduction by stating the main result in a more intuitive manner.

Definition 1.5 (calm spot). Let v be a solution of (1-5), and $\epsilon > 0$. We say the ball $B(x, 1/s)$ is an ϵ-calm spot of speed s if

$$\sup_{B(x,1/s)} |v| = s \quad \text{and} \quad \sup_{B(x,1/s)} |\nabla v| \le \epsilon s^2.$$

When ϵ is small, the gradient of the velocity is much smaller than the speed in an ϵ-calm spot, after scaling by s.

Corollary 1.6. *Let v be a solution of* (1-5) *whose initial value satisfies* (1-6). *If the flow becomes turbulent*, *that is*, *the speed becomes arbitrarily large*, *then there exist ϵ-calm spots of arbitrarily high speed. Here ϵ is any given positive number.*

By axial symmetry, there is a ring of very small vorticity. In [Hou 2009], one can find a related numerical result for the Euler equation, which is called deletion of vortex stretching. As an application, the method in this paper has helped to prove regularity of solutions in the BMO^{-1} class for axially symmetric Navier–Stokes equations. See [Lei and Zhang 2011].

2. Proof of Theorem 1.1

Let us prove part (a) first, after which the proof of (b) follows easily.

Proof. From the condition

$$\|v_0\|_{L^\infty(\mathbb{R}^3)} \le N_0, \quad \|v_0\|_{L^2(\mathbb{R}^3)} \le N_0, \quad |r v_0| \le N_0,$$

by standard theory (see [Koch et al. 2009, Proposition 4.1], for example), there exists a time h_0 such that

$$\|v(\cdot, t)\|_{L^\infty(\mathbb{R}^3)} \le 2N_0, \quad t \le h_0. \tag{2-1}$$

The proof is divided into several steps and uses the method of contradiction.

Step 1 (setting up a limit solution). Suppose part (a) of the theorem is false. Then for some $\epsilon > 0$ and $\sigma_0 > 0$, there exists a sequence of solutions v_k, with associated pressure $p_k = (-\Delta)^{-1}\nabla\cdot(v_k \cdot \nabla v_k)$ and initial condition satisfying (1-6), defined on the time interval $[0, T_k)$ for some $T_k > h_0$, which satisfies the following conditions:

(i) There exist sequences of positive numbers $\rho_k \to 0$, points $x_k \in \mathbb{R}^3$, and times $t_k \in [0, T_k)$ such that

$$r_k |v_k(x_k, t_k)| \geq \rho_k^{-2}.$$

(ii) For each k, the solution v_k in the parabolic region

$$P(x_k, t_k, [cQ_k]^{-1}) \equiv \{(x, t) \in [0, T_k) \;:\; |x_k - x| < (cQ_k)^{-1},\; t_k - (cQ_k)^{-2} \leq t \leq t_k\}$$

is not, after scaling by the factor Q_k, ϵ close, in $C^{2,1,\alpha}$ norm, to a nonzero constant vector. Here $c = \sigma_0 \epsilon$ and also

$$Q_k = |v_k(x_k, t_k)| \geq \tfrac{1}{4} \sup_{t \in [0, t_k],\; x \in R^3} |v_k(x, t)|.$$

Write $\alpha_k = r_k Q_k = r_k |v_k(x_k, t_k)|$. We consider v_k in the space-time cube

$$P\left(x_k, t_k, \frac{r_k}{\sqrt{\alpha_k}}\right) \equiv B\left(x_k, \frac{r_k}{\sqrt{\alpha_k}}\right) \times \left[t_k - \left(\frac{r_k}{\sqrt{\alpha_k}}\right)^2, t_k\right].$$

Note that

$$\beta_k \equiv \frac{r_k}{\sqrt{\alpha_k}} = \frac{r_k}{\sqrt{r_k Q_k}} = o(r_k), \tag{2-2}$$

$$Q_k \beta_k = \sqrt{r_k Q_k} \to \infty, \quad k \to \infty.$$

Define the scaled function

$$\tilde{v}_k = Q_k^{-1} v_k(Q_k^{-1} \tilde{x} + x_k, Q_k^{-2} \tilde{t} + t_k). \tag{2-3}$$

Then $\tilde{v}_k$ is a solution of the Navier–Stokes equation in the slab $\mathbb{R}^3 \times [-(Q_k \beta_k)^2, 0]$. By the assumption on Q_k, we know that $|\tilde{v}_k| \leq 4$ whenever it is defined. Since $\tilde{v}_k$ is a bounded mild solution, [Koch et al. 2009, Proposition 4.1] gives, for example, that the $C^{2,1,\alpha}$ norm of $\tilde{v}_k$ are uniformly bounded in $\mathbb{R}^3 \times [-(Q_k \beta_k)^2 + 1, 0]$. The pressure

$$P_k = Q_k^{-2} p_k(Q_k^{-1} \tilde{x} + x_k, Q_k^{-2} \tilde{t} + t_k),$$

satisfying $\Delta P_k = \operatorname{div}(\tilde{v}_k \cdot \nabla \tilde{v}_k)$, also has uniformly bounded $C^{2,1,\alpha}_{\text{local}}$ norm, by virtue of standard Schauder theory. Indeed, all $C^{p,p/2}$ norms are bounded for $p \geq 1$, though we do not need this fact here.

Let us restrict the solution $\tilde{v}_k$ to the cube

$$P(0, 0, Q_k \beta_k) = \{(\tilde{x}, \tilde{t}) \;:\; |\tilde{x}| \leq Q_k \beta_k, -(Q_k \beta_k)^2 \leq \tilde{t} \leq 0\}.$$

By the uniform bounds on the $C^{2,1,\alpha}_{\text{local}}$ norm and the fact that $Q_k \beta_k \to \infty$, we know there exists a subsequence, still called $\{\tilde{v}_k\}$, that converges to an ancient solution of the Navier–Stokes equation in $C^{2,1,\alpha}_{\text{local}}$ sense. Let us call this ancient solution $\tilde{v}$. Note that $\tilde{v}$ has length 1 at (0, 0) and is hence nontrivial. In the next step, we show that it is a spatial 2-dimensional solution, one dimension being the z-dimension.

Step 2 (proving $\tilde{v}$ is a 2D solution). Denote by v_k^θ the angular component of v_k. For the given initial value, it is known that

$$|v_k^\theta(x,t)| \le \frac{N_0}{r}.$$

For $x \in B(x_k, \beta_k)$, we have, by (2-2),

$$|v_k^\theta(x,t)| \le \frac{2N_0}{r_k}$$

when k is sufficiently large. Therefore,

$$Q_k^{-1}|v_k^\theta(x,t)| \le \frac{2N_0}{Q_k r_k} \to 0, \quad k \to \infty. \tag{2-4}$$

In the standard basis for $\mathbb{R}^3$, put $x_k = (x_{k,1}, x_{k,2}, x_{k,3})$, with the third component being the z-component, and let $\xi_k = (0, 0, x_{k,3})$. Since $(x_k - \xi_k)/|x_k - \xi_k|$ are unit vectors, there exists a subsequence, still labeled by k, that converges to a unit vector $\zeta = (\zeta_1, \zeta_2, 0)$. We use

$$\zeta, \zeta' = (-\zeta_2, \zeta_1, 0), (0, 0, 1)$$

as the basis of a new coordinate. Since this basis is obtained by a rotation around the z-axis, we know v_k is invariant. From now on, when we mention the coordinates of a point, we use the new basis with the same origin. We still use (θ, r, z) to denote the variables for the cylindrical system corresponding to this new basis.

For $x \in B(x_k, \beta_k)$, we recall that θ is the angle between x and ζ. Then

$$\cos\theta = \frac{(x - \xi_k)\cdot\zeta}{|x - (0,0,x_3)^T|} = \frac{(x_k - \xi_k)\cdot\zeta}{|x_k - \xi_k|} + \frac{O(\beta_k)}{r_k} \to 1, \quad k \to \infty. \tag{2-5}$$

For $v_k = v_k(x,t)$ in $B(x_k, \beta_k) \times [t_k - \beta_k^2, t_k]$, we have defined

$$\tilde{v}_k = \tilde{v}_k(\tilde{x}, \tilde{t}) = Q_k^{-1} v_k(Q_k^{-1}\tilde{x} + x_k, Q_k^{-2}\tilde{t} + t_k),$$

where $x = Q_k^{-1}\tilde{x} + x_k$ and $t = Q_k^{-2}\tilde{t} + t_k$. Then for $x = (x^{(1)}, x^{(2)}, x^{(3)})$ and $\tilde{x} = (\tilde{x}^{(1)}, \tilde{x}^{(2)}, \tilde{x}^{(3)})$, we have

$$\begin{cases} \partial_r v_k(x,t) = \partial_{x^{(1)}} v_k(x,t)\cos\theta + \partial_{x^{(2)}} v_k(x,t)\sin\theta \\ \qquad\qquad = Q_k^2 \partial_{\tilde{x}^{(1)}} \tilde{v}_k(\tilde{x},\tilde{t})\cos\theta + Q_k^2 \partial_{\tilde{x}^{(2)}} \tilde{v}_k(\tilde{x},\tilde{t})\sin\theta, \\ \partial_r^2 v_k(x,t) = Q_k^3 \partial_{\tilde{x}^{(1)}}^2 \tilde{v}_k(\tilde{x},\tilde{t})\cos^2\theta + 2Q_k^3 \partial_{\tilde{x}^{(1)}\tilde{x}^{(2)}}^2 \tilde{v}_k(\tilde{x},\tilde{t})\sin\theta\cos\theta \\ \qquad\qquad\qquad + Q_k^3 \partial_{\tilde{x}^{(2)}}^2 \tilde{v}_k(\tilde{x},\tilde{t})\sin^2\theta, \\ \partial_z^2 v_k(x,t) = Q_k^3 \partial_{\tilde{x}^{(3)}}^2 \tilde{v}_k(\tilde{x},\tilde{t}), \\ \partial_t^2 v_k(x,t) = Q_k^3 \partial_{\tilde{t}}^2 \tilde{v}_k(\tilde{x},\tilde{t}). \end{cases} \tag{2-6}$$

For the pressure $p_k = p_k(x,t)$, recall that

$$\tilde{p}_k = \tilde{p}_k(\tilde{x},\tilde{t}) = Q_k^{-2} p_k(Q_k^{-1}\tilde{x} + x_k,\, Q_k^{-2}\tilde{t} + t_k).$$

Therefore,

$$\partial_r p_k(x,t) = Q_k^3 \partial_{\tilde{x}^{(1)}} \tilde{p}_k(\tilde{x},\tilde{t}) \cos\theta + Q_k^3 \partial_{\tilde{x}^{(2)}} \tilde{p}_k(\tilde{x},\tilde{t}) \sin\theta. \tag{2-7}$$

Writing $v_k = v_k^r e_r + v_k^\theta e_\theta + v_k^z e_z$, we get

$$\begin{aligned} v_k^r \partial_r v_k^r + v_k^z \partial_z v_k^r = Q_k^3 \big(& v_k^r(\tilde{x},\tilde{t}) \partial_{\tilde{x}^{(1)}} \tilde{v}_k^r(\tilde{x},\tilde{t}) \cos\theta \\ & + v_k^r(\tilde{x},\tilde{t}) \partial_{\tilde{x}^{(2)}} \tilde{v}_k^r(\tilde{x},\tilde{t}) \sin\theta + \tilde{v}_k^z \partial_{\tilde{x}^{(3)}} \tilde{v}_k^r(\tilde{x},\tilde{t}) \big). \end{aligned} \tag{2-8}$$

Substituting these identities into the equation for v_k^r,

$$\Big(\partial_r^2 + \frac{1}{r}\partial_r - \frac{1}{r^2}\Big) v_k^r - (b\cdot\nabla) v_k^r + \frac{(v_k^\theta)^2}{r} - \frac{\partial p_k}{\partial r} - \frac{\partial v_k^r}{\partial t} = 0,$$

we arrive at

$$\begin{aligned} & (\partial_{\tilde{x}^{(1)}}^2 + \partial_{\tilde{x}^{(3)}}^2) \tilde{v}_k^r - (\tilde{v}_k^r \partial_{\tilde{x}^{(1)}} + \tilde{v}_k^z \partial_{\tilde{x}^{(3)}}) \tilde{v}_k^r - \partial_{\tilde{x}^{(1)}} \tilde{p}_k - \partial_{\tilde{t}} v_k^r \\ & + \frac{1}{Q_k r}\big(\partial_{\tilde{x}^{(1)}} \tilde{v}_k(\tilde{x},\tilde{t}) \cos\theta + \partial_{\tilde{x}^{(2)}} \tilde{v}_k(\tilde{x},\tilde{t}) \sin\theta\big) - \frac{1}{(Q_k r)^2} \tilde{v}_k^r + \frac{(r v_k^\theta)^2}{(Q_k r)^3} + O(\theta) = 0. \end{aligned}$$

Here $O(\theta)$ represents all the terms that vanish when $\theta \to 0$ as $k \to \infty$. In particular, all terms involving the derivative with respect to $\tilde{x}^{(2)}$ are included in $O(\theta)$.

Recall that $Q_k r$ is comparable to $Q_k r_k$, which goes to ∞. Letting $k \to \infty$ and noting that v_k and derivatives are uniformly bounded, we know that $\tilde{v}^1$, the limit of $\tilde{v}_k^r$, satisfies

$$(\partial_{\tilde{x}^{(1)}}^2 + \partial_{\tilde{x}^{(3)}}^2) \tilde{v}^{(1)} - (\tilde{v}^{(1)} \partial_{\tilde{x}^{(1)}} + \tilde{v}^{(3)} \partial_{\tilde{x}^{(3)}}) \tilde{v}^{(1)} - \partial_{\tilde{x}^{(1)}} \tilde{p} - \partial_{\tilde{t}} v^{(1)} = 0.$$

Here $\tilde{v}^{(3)}$ is the limit of v_k^z, for which we have, in a similar manner,

$$(\partial_{\tilde{x}^{(1)}}^2 + \partial_{\tilde{x}^{(3)}}^2) \tilde{v}^{(3)} - (\tilde{v}^{(1)} \partial_{\tilde{x}^{(1)}} + \tilde{v}^{(3)} \partial_{\tilde{x}^{(3)}}) \tilde{v}^{(3)} - \partial_{\tilde{x}^{(3)}} \tilde{p} - \partial_{\tilde{t}} v^{(3)} = 0.$$

Note that $\tilde{v}_k$ and its derivatives are uniformly bounded in the region of concern. When $k \to \infty$, then $\theta \to 0$ in the region of concern. Hence $\tilde{v}_k^\theta$ and derivatives all vanish when $k \to \infty$.

Finally, we need to show that $\tilde{v}^{(1)}$ and $\tilde{v}^{(3)}$ are independent of the variable $\tilde{x}^{(2)}$. To prove this, let us recall that $\partial_\theta v_k^r = \partial_\theta v_k^z = 0$. Hence

$$-\partial_{x^{(1)}} v_k^r \sin\theta + \partial_{x^{(2)}} v_k^r \cos\theta = -\partial_{x^{(1)}} v_k^z \sin\theta + \partial_{x^{(2)}} v_k^z \cos\theta = 0.$$

This implies

$$\partial_{\tilde{x}^{(2)}} \tilde{v}_k = \partial_{\tilde{x}^{(1)}} \tilde{v}_k \tan\theta.$$

Taking $k \to \infty$ (and therefore $\theta \to 0$), we see the desired result.

Step 3. Here we use a regularity result already cited [Koch et al. 2009, Proposition 4.1] and the fact that 2-dimensional ancient (mild) solutions are constants [Koch et al. 2009] to conclude that $\tilde{v}_k$, with k large, is ϵ-close to a nonzero constant vector in $C^{2,1,\alpha}_{\text{local}}$ sense. This contradiction with the condition (ii) completes the proof.

Now we prove part (b). Suppose it is false. Then for some $\epsilon > 0$, there exists a sequence of solutions v_k with normalized initial condition as above, defined on the time interval $[0, T_k)$ for some $T_k \in [h_0, T_0]$, that satisfies the following conditions.

(i) There exist sequences of positive numbers $\rho_k \to 0$, points $x_k \in \mathbb{R}^3$, and times $t_k \in [0, T_k)$ such that

$$r_k |v_k(x_k, t_k)| \geq \rho_k^{-2}.$$

(ii) For each k, the solution v_k in the parabolic region

$$P(x_k, t_k, [cQ_k]^{-1}) \equiv \{(x, t) \in [0, T_k) : |x - x_k| < (cQ_k)^{-1},\ t_k - (cQ_k)^{-2} \leq t \leq t_k\}$$

is not, after scaling by the factor Q_k, ϵ-close, in $C^{2,1,\alpha}_{\text{local}}$ norm, to a nonzero constant vector. Here $c = \sigma_0 \epsilon$ and also

$$r_k |v_k(x_k, t_k)| \geq \tfrac{1}{4} \sup_{t \in [0, t_k],\ x \in R^3} r |v_k(x, t)|.$$

As before, define $Q_k = |v(x_k, t_k)|$. Suppose k is large. Then for $x \in B(x_k, \beta_k)$ with $\beta_k = r_k / \sqrt{r_k Q_k} = o(r_k)$, there holds, for $t \leq t_k$,

$$r |v(x, t)| \leq r_k |v(x_k, t_k)| = r_k Q_k$$

and $\frac{1}{2} r_k \leq r \leq 2 r_k$, when k is large. This shows, in the ball $B(x_k, \beta_k)$ and for $t \leq t_k$, that

$$|v(x, t)| \leq 2 Q_k.$$

Now we can scale by Q_k^{-1} in the above ball again, as in the proof of part (a). By [Seregin and Šverák 2009, Theorem 2.8], the limit of scaled solutions is again a bounded, mild, ancient solution. Similar arguments as in part (a) lead to a contradiction, proving part (b). □

Acknowledgment

Zhang thanks the mathematics department of Fudan university for its hospitality during his visits.

This work was supported in part by NSFC (grant No. 11171072), the Foundation for Innovative Research Groups of NSFC (grant No. 11121101), FANEDD (grant No. 200920), Shanghai Rising Star Program (10QA1400300), Innovation Program of Shanghai Municipal Education Commission (grant No. 12ZZ012) and SGST 09DZ2272900.

References

[Chae and Lee 2002] D. Chae and J. Lee, "On the regularity of the axisymmetric solutions of the Navier–Stokes equations", *Math. Z.* **239**:4 (2002), 645–671. MR 2003d:35205 Zbl 0992.35068

[Chen et al. 2008] C.-C. Chen, R. M. Strain, H.-T. Yau, and T.-P. Tsai, "Lower bound on the blow-up rate of the axisymmetric Navier–Stokes equations", *Int. Math. Res. Not.* **2008**:9 (2008), Art. ID rnn016, 31. MR 2009i:35233 Zbl 1154.35068

[Chen et al. 2009] C.-C. Chen, R. M. Strain, T.-P. Tsai, and H.-T. Yau, "Lower bounds on the blow-up rate of the axisymmetric Navier–Stokes equations, II", *Comm. Partial Differential Equations* **34**:1-3 (2009), 203–232. MR 2010f:35279 Zbl 1173.35095

[Hou 2009] T. Y. Hou, "Blow-up or no blow-up? A unified computational and analytic approach to 3D incompressible Euler and Navier–Stokes equations", *Acta Numer.* **18** (2009), 277–346. MR 2010g:35224 Zbl 05622314

[Hou and Li 2008] T. Y. Hou and C. Li, "Dynamic stability of the three-dimensional axisymmetric Navier–Stokes equations with swirl", *Comm. Pure Appl. Math.* **61**:5 (2008), 661–697. MR 2009j:35253 Zbl 1138.35077

[Hou et al. 2008] T. Y. Hou, Z. Lei, and C. Li, "Global regularity of the 3D axi-symmetric Navier–Stokes equations with anisotropic data", *Comm. Partial Differential Equations* **33**:7-9 (2008), 1622–1637. MR 2009g:35219 Zbl 1146.76010

[Jiu and Xin 2003] Q. Jiu and Z. Xin, "Some regularity criteria on suitable weak solutions of the 3-D incompressible axisymmetric Navier–Stokes equations", pp. 119–139 in *Lectures on partial differential equations*, edited by S.-Y. A. Chang et al., New Stud. Adv. Math. **2**, Int. Press, Somerville, MA, 2003. MR 2005i:35210 Zbl 1054.35053

[Koch et al. 2009] G. Koch, N. Nadirashvili, G. A. Seregin, and V. Šverák, "Liouville theorems for the Navier–Stokes equations and applications", *Acta Math.* **203**:1 (2009), 83–105. MR 2010i:35281 Zbl 1208.35104

[Ladyzhenskaya 1968] O. A. Ladyzhenskaya, "Unique global solvability of the three-dimensional Cauchy problem for the Navier–Stokes equations in the presence of axial symmetry", *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* (*LOMI*) **7** (1968), 155–177. MR 39 #3170

[Lei and Zhang 2011] Z. Lei and Q. S. Zhang, "A Liouville theorem for the axially-symmetric Navier–Stokes equations", *J. Funct. Anal.* **261**:8 (2011), 2323–2345. MR 2824580 Zbl 05956041

[Leonardi et al. 1999] S. Leonardi, J. Málek, J. Nečas, and M. Pokorný, "On axially symmetric flows in $\mathbb{R}^3$", *Z. Anal. Anwendungen* **18**:3 (1999), 639–649. MR 2000h:76038 Zbl 0943.35066

[Neustupa and Pokorný 2000] J. Neustupa and M. Pokorný, "An interior regularity criterion for an axially symmetric suitable weak solution to the Navier–Stokes equations", *J. Math. Fluid Mech.* **2**:4 (2000), 381–399. MR 2001m:76034 Zbl 0980.35129

[Seregin and Šverák 2009] G. Seregin and V. Šverák, "On type I singularities of the local axi-symmetric solutions of the Navier–Stokes equations", *Comm. Partial Differential Equations* **34**:1-3 (2009), 171–201. MR 2010k:35356 Zbl 1180.35002

[Tian and Xin 1998] G. Tian and Z. Xin, "One-point singular solutions to the Navier–Stokes equations", *Topol. Methods Nonlinear Anal.* **11**:1 (1998), 135–145. MR 99j:35170 Zbl 0923.35121

[Ukhovskii and Iudovich 1968] M. R. Ukhovskii and V. I. Iudovich, "Axially symmetric flows of ideal and viscous fluids filling the whole space", *J. Appl. Math. Mech.* **32** (1968), 52–61. MR 39 #650 Zbl 0172.53405

Received February 16, 2011. Revised September 16, 2011.

ZHEN LEI
SCHOOL OF MATHEMATICAL SCIENCES
LMNS AND SHANGHAI KEY LABORATORY FOR CONTEMPORARY APPLIED MATHEMATICS
FUDAN UNIVERSITY
SHANGHAI 200433
CHINA
zlei@fudan.edu.cn

QI S. ZHANG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
RIVERSIDE, CA 92521
UNITED STATES
qizhang@math.ucr.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

LOCAL COMPARISON THEOREMS FOR KÄHLER MANIFOLDS

GANG LIU

We establish a sharp relative volume comparison theorem for small balls on Kähler manifolds with lower bound on Ricci curvature, assuming real analyticity of the metric. The model spaces being compared to are complex space forms, that is, Kähler manifolds with constant holomorphic sectional curvature. Moreover, we give an example showing that on Kähler manifolds, the pointwise Laplacian comparison theorem does not hold when the Ricci curvature is bounded from below.

1. Introduction

Comparison theorems are fundamental tools in geometric analysis. They are vital in estimates of spectra, heat kernels and the Sobolev constants. The classical Bishop–Gromov relative volume comparison theorem [Bishop and Crittenden 1964; Gromov 1981; Li 1993] in Riemannian geometry is this:

Theorem 1.1. *Let M^n be a complete Riemannian manifold of dimension n such that* $\operatorname{Ric} \geq (n-1)K$. *For any $p \in M$ and $0 < a < b$, the volume of geodesic balls satisfies*

$$\frac{\operatorname{Vol} B_p(b)}{\operatorname{Vol} B_p(a)} \leq \frac{\operatorname{Vol} B_{M_K}(b)}{\operatorname{Vol} B_{M_K}(a)},$$

where M_K is the simply connected real space form with sectional curvature K and $\operatorname{Vol} B_{M_K}(r)$ *is the volume of the geodesic ball in M_K with radius r. Equality holds if and only if $B_p(b)$ is isometric to $B_{M_K}(b)$.*

The key ingredient in Theorem 1.1 is the Laplacian comparison theorem [Cheeger and Ebin 2008; Schoen and Yau 1994]:

Theorem 1.2. *Let M^n be a complete Riemannian manifold with* $\operatorname{Ric} \geq (n-1)K$. *Let M_k be the simply connected real space form with sectional curvature K. Denote by $r_M(x)$ the distance function from p to x in M. Let r_{M_k} be the distance function on M_k. Then for any $x \in M$ and $y \in M_k$ with $r_M(x) = r_{M_k}(y)$,*

$$\Delta r_M(x) \leq \Delta r_{M_k}(y).$$

MSC2010: primary 53BXX; secondary 53B35, 53B20.
Keywords: volume comparison, Kähler manifolds.

The model spaces in the theorems above are real space forms. In the Kähler category, a natural question is whether we can replace the model spaces by Kähler models, that is, complex space forms which are Kähler manifolds with constant holomorphic sectional curvature. Li and Wang [2005] showed that when the bisectional curvature has a lower bound, both of the theorems above hold with Kähler models. So the question left is: what can we get if we only assume the lower bound of the Ricci curvature? This note addresses the local case. Our main theorem is:

Theorem 1.3. *Let M^n be a Kähler manifold of complex dimension n with a real analytic metric. Assume* $\mathrm{Ric} \geq K$, *where K is any real number. Given any point $p \in M$, there exists $r = r(p, M) > 0$ such that for any $0 < a < b < r$, the volume of geodesic balls satisfies*

$$\frac{\mathrm{Vol}\, B_{M^n}(p, b)}{\mathrm{Vol}\, B_{M^n}(p, a)} \leq \frac{\mathrm{Vol}\, B_{N_K}(b)}{\mathrm{Vol}\, B_{N_K}(a)},$$

where N_K denotes the rescaled complex space form with $\mathrm{Ric} = K$ *and $\Delta_{N_K} r$ is the Laplacian of distance function on N_K. Equality holds if and only if M is locally isometric to N_K.*

Remark. Theorem 1.3 is a local version of the Bishop–Gromov relative volume comparison theorem on Kähler manifolds. However, one cannot directly extend the result to any radius. A simple counterexample is a product of several $\mathbb{P}^1$ with the standard product metric: the diameter is greater than that of the complex space form. Thus, when r is large, the inequality in Theorem 1.3 does not hold.

We will prove a slightly stronger result:

Theorem 1.4. *Under the hypotheses of Theorem 1.3, there exists $r_0 = r_0(p, M) > 0$ such that for any $r < r_0$, the average Laplacian comparison holds,*

$$\frac{\int_{\partial B_p(r)} \Delta r}{A(\partial B_p(r))} \leq \Delta_{N_K} r(r),$$

where $\Delta_{N_K} r$ is the Laplacian of distance function on N_K. Moreover, the equality holds if and only if M is locally isometric to N_K.

Remark. Theorem 1.4 is a local version of Theorem 1.2 in the average sense. However, on Kähler manifolds with lower bound on Ricci curvature, the pointwise Laplacian comparison does not even hold locally (see Section 6).

The idea of the proof of Theorem 1.4 is very simple. We shall expand the area of the geodesic sphere $A(\partial B_p(r))$ as a power series, then compare the coefficients with those of the rescaled complex space form. The computation is complicated since it involves the covariant derivatives of the curvature tensor of arbitrary order.

This note is organized as follows:

In Section 2, we state two propositions which demonstrate the relation between the derivatives of $A(\partial B_p(r))$ and the covariant derivatives of the curvature tensor at p. Section 3 is the first part of the proof of Proposition 2.1. We shall estimate the derivatives of $A(\partial B_p(r))$ up to order 4. In the estimate of the 4-th derivative, the Kähler condition is employed. The most important part is Section 4. We use induction to prove Proposition 2.1. Besides the routine computation, there are two technical lemmas (Lemma 4.4 and Lemma 4.6) which simplify the computation of higher order covariant derivatives of the curvature tensor significantly. One should note that the Kähler condition is essential in these two lemmas. We complete the proof of Proposition 2.2 and Theorem 1.4 in Section 5. The last section is devoted to giving an example showing that the pointwise Laplacian comparison with the complex space form does not necessarily hold if the complex dimension is greater or equal to 2.

2. Basic set up

Throughout this note, we implicitly evaluate derivatives of functions of r at $r = 0$. Given a point p on a Kähler manifold M^n, fix a unit vector $e_0 \in T_pM$. Along the geodesic l from p with initial direction e_0, consider the Jacobian equation $J'' = R(e_0, J)\, e_0$. Set up an orthonormal frame $\{e_k\}$ at p such that

$$Je_{2i} = e_{2i+1} \quad \text{and} \quad Je_{2i+1} = -e_{2i}$$

for $0 \le i \le n-1$. Parallel transport the frame along the geodesic l. Consider the Jacobian field J_u with initial value $J_u(0) = 0$, $J'_u(0) = e_u$.

We may write

$$J_u = J_u(r, e_0) = \sum_{i=1}^{\infty} \sum_{v=0}^{2n-1} r^i C^v_{u,i}\, e_v \tag{2-1}$$

where $C^v_{u,i}$ are constants independent of r. Denote $R_{e_0 e_u e_0 e_v}$ by R_{uv} when e_0 is fixed. Plugging (2-1) into the Jacobian equation, we get

$$\sum_i \sum_v i(i-1) r^{i-2} C^v_{u,i}\, e_v = \sum_k \sum_w r^k C^w_{u,k} R(e_0, e_w)\, e_0. \tag{2-2}$$

Along the geodesic l,

$$R(e_0, e_w)\, e_0 = \sum_{s=0}^{2n-1} \sum_{j=0}^{\infty} \frac{R^{(j)}_{sw}}{j!} e_s r^j$$

where $R^{(j)}_{sw}$ denotes the j-th covariant derivative of R_{sw} along e_0 at p.

Inserting this into (2-2), we get

$$\sum_{i,v} i(i-1)r^{i-2}C^v_{u,i}\, e_v = \sum_{k,j,w,s} r^{k+j} C^w_{u,k} \frac{R^{(j)}_{sw}}{j!} e_s.$$

Comparing coefficients, we obtain

$$C^v_{u,i} = \sum_{k+j=i-2,\ w} C^w_{u,k} \frac{R^{(j)}_{vw}}{j!\, i(i-1)}. \tag{2-3}$$

A simple iteration now gives the constants $C^v_{u,i}$. First we have $C^v_{u,1} = \delta^v_u$ and $C^w_{u,2} = 0$. Then we get

$$C^v_{u,3} = \sum_w C^w_{u,1} \frac{R_{vw}}{6} = \frac{R_{uv}}{6}, \quad C^v_{u,4} = \sum_w C^w_{u,1} \frac{R'_{vw}}{12} = \frac{R'_{vu}}{12},$$
$$C^v_{u,5} = \sum_w \Bigl(C^w_{u,1} \frac{R''_{vw}}{40} + C^w_{u,3} \frac{R_{vw}}{20} \Bigr) = \frac{1}{120} \Bigl(\sum_s R_{us} R_{sv} + 3R''_{uv} \Bigr).$$

Plugging these values into (2-1), we have

$$J_u = re_u + \frac{r^3}{6} R_{uv} e_v + \frac{r^4}{12} R'_{uv} e_v + \frac{r^5}{120} \Bigl(\sum_s R_{us} R_{sv} + 3R''_{uv} \Bigr) e_v + O(r^6). \tag{2-4}$$

We write dA for the standard measure of the unit tangent bundle $UT_p(M)$ at p, and we write $\int_{\partial B(p,r)} dA$ as $\int$. We define

$$W = \frac{\int \sqrt{\det\langle J_u, J_v\rangle}}{r^{2n-1}},$$

and introduce two propositions:

Proposition 2.1. *Assume the hypotheses of Theorem 1.3. Let the derivatives of W of order from 1 to $2m-1$ for $m \geq 1$ be the same as that of the complex space form.*

(1) *If $m = 1, 2$, then* $\mathrm{Ric} = K$ *at p.*
If $m \geq 3$, then
$$R_{i\bar{j}k\bar{l}} = \frac{K}{n+1} (\delta_{ij}\delta_{kl} + \delta_{il}\delta_{jk})$$
at p. Moreover, if $u, v, e_0 \in UT_p(M)$ are any unit vectors, then $R^{(\lambda)}_{uv} = 0$ for $1 \leq \lambda \leq m-3$ and $\mathrm{Ric}^{(l)}(e_0, e_0) = 0$ for $1 \leq l \leq 2m-4$. The superscripts are the orders of covariant derivatives along the direction e_0.

(2) *In either case, $W^{(2m)}$ is less than or equal to that of the complex space form.*

Proposition 2.2. *Under the same conditions as in Theorem 1.3, if the derivatives of W of order 1 to $2m$ for $m \geq 1$ are the same as the complex space form, then $W^{(2m+1)} = 0$.*

We divide the proof of Proposition 2.1 into two parts: $m = 1, 2$ and $m \geq 3$.

3. The proof of Proposition 2.1, case $m = 1, 2$

By (2-1), we have

$$\frac{\langle J_u, J_v\rangle}{r^2} = \sum_{i,j,w} r^{i+j-2} C^w_{u,i} C^w_{v,j}. \tag{3-1}$$

By (2-4),

$$\frac{\langle J_u, J_u\rangle}{r^2} = 1 + \tfrac{1}{3} R_{uu} r^2 + \tfrac{1}{6} R'_{uu} r^3 + \Big(\tfrac{2}{45}\sum_s R^2_{us} + \tfrac{1}{20} R''_{uu}\Big) r^4 + O(r^5).$$

If $u \neq v$,

$$\frac{\langle J_u, J_v\rangle}{r^2} = \tfrac{1}{3} R_{uv} r^2 + \tfrac{1}{6} R'_{uv} r^3 + \Big(\tfrac{2}{45}\sum_s R_{us} R_{vs} + \tfrac{1}{20} R''_{uv}\Big) r^4 + O(r^5).$$

Now use the above two expressions to see that

$$\begin{aligned}\frac{\det\langle J_u, J_v\rangle}{r^{4n-2}} &= 1 + \tfrac{1}{3}\sum_u R_{uu} r^2 + \tfrac{1}{6}\sum_u R'_{uu} r^3 \\ &\quad + \Big(\tfrac{2}{45}\sum_{u,s} R^2_{us} + \tfrac{1}{20}\sum_u R''_{uu} + \tfrac{1}{9}\sum_{u<v}\big(R_{uu}R_{vv} - R^2_{uv}\big)\Big) r^4 + O(r^5).\end{aligned} \tag{3-2}$$

Considering the identity $\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + O(x^3)$, we get

$$\begin{aligned}\frac{\sqrt{\det\langle J_u, J_v\rangle}}{r^{2n-1}} &= 1 + \tfrac{1}{6}\sum_u R_{uu} r^2 + \tfrac{1}{12}\sum_u R'_{uu} r^3 \\ &\quad + \Big(\tfrac{1}{45}\sum_{u,s} R^2_{us} + \tfrac{1}{40}\sum_u R''_{uu} + \tfrac{1}{18}\sum_{u<v}\big(R_{uu}R_{vv} - R^2_{uv}\big) - \tfrac{1}{72}\Big(\sum_u R_{uu}\Big)^2\Big) r^4 + O(r^5).\end{aligned} \tag{3-3}$$

Since $W = \frac{1}{r^{2n-1}} \int \sqrt{\det\langle J_u, J_v\rangle}$, we find

$$W'(0) = 0 \quad \text{and} \quad W''(0) = -c\, s,$$

where c is a positive constant depending only on n, and s is the scalar curvature at p. Therefore $W''(0)$ is less than or equal to that of the complex space form. This proves Proposition 2.1 for $m = 1$.

Now we consider $m = 2$. According to the assumption of Proposition 2.1, W'' is the same as that of the complex space form. Therefore $s = nK$ at p. Since the Ricci curvature is bounded from below by K, $\mathrm{Ric} = Kg$ at p. By (3-3), it is simple to see that the r^3 coefficient of W is 0 by symmetry. Thus to complete the proof for $m = 2$, we just need to show that the 4th derivative of W is less than or equal to that of the complex space form.

We keep in mind that $\mathrm{Ric} = Kg$ at p. The r^4 coefficient of W is

$$
\begin{aligned}
c_4 &= \int \Big(\tfrac{1}{45} \sum_{u,s} R_{us}^2 + \tfrac{1}{40} \sum_u R''_{uu} + \tfrac{1}{18} \sum_{u<v} \big(R_{uu}R_{vv} - R_{uv}^2\big) - \tfrac{1}{72}\Big(\sum_u R_{uu}\Big)^2 \Big) \\
&= \frac{1}{360} \int \Big(8 \sum_u R_{uu}^2 + 16 \sum_{u<v} R_{uv}^2 + 9 \sum_u R''_{uu} \\
&\qquad\qquad + 20 \sum_{u<v} R_{uu}R_{vv} - 20 \sum_{u<v} R_{uv}^2 - 5\Big(\sum_u R_{uu}\Big)^2 \Big) \\
&= \frac{1}{360} \int \Big(-2 \sum_u R_{uu}^2 + 10(\sum_u R_{uu})^2 - 4 \sum_{u<v} R_{uv}^2 + 9 \sum_u R''_{uu} - 5\Big(\sum_u R_{uu}\Big)^2 \Big) \\
&= \frac{1}{360} \int \Big(9 \sum_u R''_{uu} - 4 \sum_{u<v} R_{uv}^2 - 2 \sum_u R_{uu}^2 + 5\Big(\sum_u R_{uu}\Big)^2 \Big).
\end{aligned}
$$

Note that the Ricci curvature attains the minimum K at p, so

$$
\sum_u R''_{uu} = -\mathrm{Ric}''(e_0, e_0) \le 0.
$$

Therefore we have

$$
\begin{aligned}
\text{(3-4)} \qquad c_4 &= \frac{1}{360} \int \Big(9 \sum_u R''_{uu} - 4 \sum_{u<v} R_{uv}^2 - 2 \sum_u R_{uu}^2 + 5K^2 \Big) \\
&\le -\frac{1}{360} \int \Big(2 \sum_u R_{uu}^2 - 5K^2 \Big) \\
&= -\frac{1}{360} \int \Big(2 \sum_{u\neq 1} R_{uu}^2 + 2R_{11}^2 - 5K^2 \Big) \\
&\le -\frac{1}{360} \int \Big(\frac{1}{n-1}\Big(\sum_{u\neq 1} R_{uu}\Big)^2 + 2R_{11}^2 - 5K^2 \Big) \\
&= -\frac{1}{360} \int \Big(\frac{1}{n-1}(\mathrm{Ric}(e_0, e_0) + R_{11})^2 + 2R_{11}^2 - 5K^2 \Big) \\
&= -\frac{1}{360} \int \Big(\frac{1}{n-1}K^2 + \frac{2}{n-1}KR_{11} + \Big(\frac{1}{n-1} + 2\Big)R_{11}^2 - 5K^2 \Big) \\
&\le -\frac{1}{360} \Big(\int \frac{1}{n-1}K^2 + \frac{2}{n-1}K \int R_{11} + C_1\Big(\int R_{11}\Big)^2 - \int 5K^2 \Big) \\
&= C_2 K^2,
\end{aligned}
$$

where C_1, C_2 are constants depending only on n.

We explain the inequalities above. In the first inequality, we drop the two terms $\sum_{u<v} R_{uv}^2$ and $\sum_u R''_{uu}$. In the second inequality, we apply the Schwartz inequality for directions e_u that are perpendicular to e_1, e_0. In the third inequality we use the Schwartz inequality $\int R_{11}^2 \ge C\big(\int R_{11}\big)^2$. We make use of the Kähler condition to

obtain $\int R_{11} = C_3\, s = nC_3K$, where C_3 is a constant depending only on n. This explains the last equality.

The right hand side of (3-4) is exactly the case of the complex space form. Therefore when W' and W'' are the same as the complex space form, $W^{(3)} = 0$ and $W^{(4)}$ is less than or equal to that of the complex space form. Equation (3-4) becomes an equality if and only if the holomorphic sectional curvature is constant at p and $\mathrm{Ric}''(e_0, e_0) = 0$ for any $e_0 \in UT_pM$. This completes the proof for $m = 2$.

4. The proof of Proposition 2.1, case $m \geq 3$

Denote $\mathrm{Ric}^{(l)}(e_0, e_0)$ by $\mathrm{Ric}^{(l)}$. According to the assumption of Proposition 2.1, the derivatives of W of order 1 to $2m-1$ are the same as the complex space form. From results in the last section, the holomorphic sectional curvature is constant at p and $\mathrm{Ric}'' = 0$ for any e_0. That is to say,

$$R_{i\bar{j}k\bar{l}} = \frac{K}{n+1}(\delta_{ij}\delta_{kl} + \delta_{il}\delta_{jk}) \quad \text{and} \quad \mathrm{Ric}'' = 0$$

at p. Therefore, we have proved part (1) of Proposition 2.1 for $m = 3$.

Now we use induction. Assuming that part (1) of Proposition 2.1 holds for $k = m$, we shall prove that it holds for $k = m + 1$.

Claim 4.1. *Let $C^v_{u,i}$ be the coefficients defined in* (2-1) *for $i \leq m$. Under the hypothesis of the induction above, $C^v_{u,i}$ are constants independent of the direction e_0. In fact, they are the same as that of the complex space form.*

Proof. The claim follows if we insert the induction hypothesis into (2-3). □

Let us write

$$\frac{\det\langle J_u, J_v\rangle}{r^{4n-2}} = 1 + \sum_{i=1}^{m-1} a_i r^i + \sum_{j=m}^{2m} b_j r^j + O(r^{2m+1}). \tag{4-1}$$

Combining Claim 4.1 with (3-1), we find that a_i are constants independent of the direction e_0. Equation (3-1) also yields $C^v_{u,m+1} = C^u_{v,m+1}$ for all u, v. Direct expansion of the determinant via (3-1) gives

$$\begin{aligned} b_{2m} = \sum_{u,v}\big(C^v_{u,m+1}\big)^2 + 4\sum_{u<v} C^u_{u,m+1}C^v_{v.m+1} + 2\sum_u C^u_{u,2m+1} \\ -4\sum_{u<v} C^v_{u,m+1}C^u_{v,m+1} + \sum_{i=1}^{m} C^v_{u,m+i}C_{i,m,u,v} + C_{0,m} \end{aligned} \tag{4-2}$$

where $C_{i,m,u,v}$ and $C_{0,m}$ are all constants independent of the direction e_0.

Note also

$$b_m = 2\sum_u C^u_{u,m+1} + \text{Constant}. \tag{4-3}$$

Let us set $\gamma_m = \sum_{i=1}^{m-1} a_i r^i + \sum_{j=m}^{2m} b_j r^j$ for $m \geq 1$. Applying the Taylor series expansion

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \sum_{k=3}^{\infty} \lambda_k x^k$$

for $|x| < 1$, we obtain

$$\text{(4-4)} \qquad \frac{\sqrt{\det\langle J_u, J_v\rangle}}{r^{2n-1}} = 1 + \tfrac{1}{2}\gamma_m - \tfrac{1}{8}\gamma_m^2 + \sum_{k=3}^{\infty} \lambda_k \gamma_m^k + O(r^{2m+1}).$$

Lemma 4.2. *The* $2m$*-th order coefficient of the expansion of* W *is*

$$\text{(4-5)} \quad c_{2m} = \int \Big(\tfrac{1}{2} \sum_{u,v} \big(C^v_{u,m+1}\big)^2 + 2 \sum_{u<v} C^u_{u,m+1} C^v_{v,m+1} + \sum_u C^u_{u,2m+1} \\ -2 \sum_{u<v} C^v_{u,m+1} C^u_{v,m+1} - \tfrac{1}{2} \Big(\sum_u C^u_{u,m+1} \Big)^2 + \sum_{i=1}^{m} C^v_{u,m+i} \widetilde{C}_{i,m,u,v} \Big) + \widetilde{C}_{0,m}$$

where $\widetilde{C}_{i,m,u,v}$ *and* $\widetilde{C}_{0,m}$ *are constants independent of the direction* e_0.

Proof. It suffices to find out the contribution of each term in (4-4) to c_{2m}. We keep in mind that coefficients a_i in (4-1) are independent of e_0.

By (4-2), the contribution of the term $1 + \frac{1}{2}\gamma_m$ to c_{2m} is

$$\text{(4-6)} \quad \int \tfrac{1}{2} \sum_{u,v} \big(C^v_{u,m+1}\big)^2 + 2 \sum_{u<v} C^u_{u,m+1} C^v_{v,m+1} + \sum_u C^u_{u,2m+1} \\ -2 \sum_{u<v} C^v_{u,m+1} C^u_{v,m+1} + \tfrac{1}{2} \Big(\sum_{i=1}^{m} C^v_{u,m+i} C_{i,m,u,v} + C_{0,m} \Big).$$

The contribution of the term $-\frac{1}{8}\gamma_m^2$ to c_{2m} is

$$\text{(4-7)} \qquad -\int \Big(\tfrac{1}{8} b_m^2 + \sum_{i=1}^{m} C^v_{u,m+i} p_{i,m,u,v} \Big) + p_{0,m}.$$

By (4-3), it could be written as

$$\text{(4-8)} \qquad -\int \Big(\tfrac{1}{2} \Big(\sum_u C^u_{u,m+1} \Big)^2 + \sum_{i=1}^{m} C^v_{u,m+i} p_{i,m,u,v} \Big) + p_{0,m}.$$

The contribution of $\sum_{k=3}^{\infty} \lambda_k \gamma_m^k$ to c_{2m} is

$$\text{(4-9)} \qquad \int \sum_{i=1}^{m} C^v_{u,m+i} q_{i,m,u,v} + q_{0,m}.$$

In (4-7), (4-8), (4-9), $p_{i,m,u,v}$, $q_{i,m,u,v}$, $p_{0,m}$ and $q_{0,m}$ are all constants independent of the direction e_0. Lemma 4.2 follows if we combine (4-6), (4-7), (4-8) and (4-9). □

Lemma 4.3. *There is a negative definite quadratic form Q, constants $h_{m,i}$ and C and a negative constant C_m such that*

$$c_{2m} = \int Q\big(R_{uv}^{(m-2)}\big) + \sum_{i=-2}^{m-4} h_{m,i} \int R_{11}^{(m+i)} + C_m \int \mathrm{Ric}^{(2m-2)} + C. \tag{4-10}$$

Proof. By the induction hypothesis and (2-3), we have

$$\begin{aligned} C^u_{u,2m+1} &= \sum_{k+j=2m-1,w} \frac{C^w_{u,k} R^{(j)}_{uw}}{j!\,(2m+1)2m} \\ &= \frac{1}{(2m+1)2m}\bigg(\sum_w\bigg(\frac{R^{(m-2)}_{uw} C^w_{u,m+1}}{(m-2)!} + \sum_{j=m-1}^{2m-2} B_{j,m,w,u} R^{(j)}_{uw}\bigg) + R_{uu} C^u_{u,2m-1}\bigg) \end{aligned} \tag{4-11}$$

where $B_{j,m,w,u}$ are constants. For $i \le m$, we have

$$C^v_{u,m+i} = \sum_{j=m-2}^{m+i-3} d_{m,i,j,w,u} R^{(j)}_{uw} + C \tag{4-12}$$

where C and $d_{m,i,j,w,u}$ are constants. In particular, we have

$$C^v_{u,m+1} = \sum_{k+j=m-1,w} C^w_{u,k} \frac{R^{(j)}_{vw}}{j!\,m(m+1)} = \frac{1}{m(m+1)}\bigg(\frac{R^{(m-2)}_{vu}}{(m-2)!} + C^v_{u,m-1} R_{vv}\bigg). \tag{4-13}$$

By the induction hypothesis,

$$\sum_u R^{(m-2)}_{uu} = -\mathrm{Ric}^{(m-2)} = 0. \tag{4-14}$$

Therefore

$$\begin{aligned} \sum_u \big(R^{(m-2)}_{uu}\big)^2 &= \Big(\sum_u R^{(m-2)}_{uu}\Big)^2 - 2\sum_{u<v} R^{(m-2)}_{uu} R^{(m-2)}_{vv} \\ &= -2\sum_{u<v} R^{(m-2)}_{uu} R^{(m-2)}_{vv}. \end{aligned} \tag{4-15}$$

Inserting (4-11), (4-12), (4-13) into (4-5), we find

$$c_{2m} = \int Q\big(R^{(m-2)}_{uv}\big) + \sum_{i=-2}^{m-2} \int \sum_{u,v} h_{m,i,u,v} R^{(m+i)}_{uv} + C. \tag{4-16}$$

Now we prove that Q is negative definite. Let us check each term in (4-5). By (4-13), the term $\frac{1}{2}\sum_{u,v}\big(C^v_{u,m+1}\big)^2$ in (4-5) contributes to the quadratic term

$$\sum_{u,v} \frac{1}{2m^2(m+1)^2((m-2)!)^2} \big(R^{(m-2)}_{uv}\big)^2. \tag{4-17}$$

The term $2\sum_{u<v} C^u_{u,m+1}C^v_{v,m+1}$ contributes to the quadratic term

$$\sum_{u<v} \frac{2}{m^2(m+1)^2((m-2)!)^2} R^{(m-2)}_{uu} R^{(m-2)}_{vv}. \tag{4-18}$$

By (4-15), it could be written as

$$-\frac{1}{m^2(m+1)^2((m-2)!)^2} \sum_u \big(R^{(m-2)}_{uu}\big)^2. \tag{4-19}$$

By (4-11) and (4-13), the term $\sum_u C^u_{u,2m+1}$ contributes to the quadratic term

$$\sum_{u,v} \frac{1}{2m^2(m+1)(2m+1)((m-2)!)^2} \big(R^{(m-2)}_{uv}\big)^2. \tag{4-20}$$

The term $-2\sum_{u<v} C^v_{u,m+1}C^u_{v,m+1}$ contributes to the quadratic term

$$-\sum_{u<v} \frac{2}{m^2(m+1)^2((m-2)!)^2} \big(R^{(m-2)}_{uv}\big)^2. \tag{4-21}$$

The term $-\frac{1}{2}\big(\sum_u C^u_{u,m+1}\big)^2$ is obviously negative semidefinite.

By combining (4-17), (4-18), (4-19), (4-20) and (4-21), it follows that the quadratic form in (4-10) is negative definite.

Consider the linear terms in (4-16). By the induction hypothesis, the coefficients $h_{m,i,u,v}$ are unchanged if we take a unitary transformation keeping the direction e_0 fixed. Comparing the coefficients of the linear order terms, we see that $h_{m,i,u,v} = 0$ if $u \neq v$, and $h_{m,i,u,u} = h_{m,i,v,v}$ if $u \neq e_1$ and $v \neq e_1$. Therefore, the linear terms $h_{m,i,u,u} R^{(m+i)}_{uu}$ could be absorbed into $\mathrm{Ric}^{(m+i)}$ with the terms $-h_{m,i} R^{(m+i)}_{11}$ left. Also note that by induction hypothesis, $\mathrm{Ric}^{(l)} = 0$ for $0 < l \leq 2m-3$ (the term $\mathrm{Ric}^{(2m-3)}$ vanishes as the Ricci curvature attains its minimum at p). Finally, one verifies that $\sum_u C^u_{u,2m+1}$ is the only term in (4-5) that has contribution to $R^{(2m-2)}_{uv}$. Therefore the linear terms in (4-16) could be written as

$$\sum_{i=-2}^{m-4} h_{m,i} \int R^{(m+i)}_{11} + C_m \int \mathrm{Ric}^{(2m-2)}.$$

From (4-11), it is simple to check that C_m is negative. □

By the induction hypothesis and that the Ricci curvature attains its minimum at p, we have $\mathrm{Ric}^{(2m-2)} \geq 0$. It follows from Lemma 4.3 that

$$c_{2m} \leq \sum_{i=-2}^{m-4} h_{m,i} \int R^{(m+i)}_{11} + \text{Constant}. \tag{4-22}$$

We would like to prove that the linear terms $\int R^{(m+i)}_{11}$ vanish for $-2 \leq i \leq m-4$. Note that by symmetry, if $m+i$ is odd, the integral equals 0. Let us deal with case when $m+i$ is even. We shall check when $i=m-4$. The other cases are similar. Let

$$A=-\frac{1}{4}\int R_{11}^{(2m-4)}. \tag{4-23}$$

Set up an orthonormal frame $\{f_i\}$ at p such that $Jf_{2j}=f_{2j+1}$ and $Jf_{2j+1}=-f_{2j}$ for $0\le j\le n-1$. Letting $\beta_j=\frac{1}{2}(f_{2j}-\sqrt{-1}f_{2j+1})$, in a small neighborhood of p, we parallel transport the frame along each geodesic through p. Suppose that

$$e_0=\sum_{j=0}^{n-1}\bigl(z_j\beta_j+\overline{z_j\beta_j}\bigr). \tag{4-24}$$

Lemma 4.4. *Under the assumption of the induction in Proposition 2.1*, $Rm^{(\lambda)}=0$ *at* p *for* $1\le\lambda\le m-3$, *where* $Rm^{(\lambda)}$ *denotes any covariant derivative of the curvature tensor of order* λ *at* p.

Proof. We use induction. If $\lambda=0$, the result automatically holds since there is nothing to prove. Suppose the result holds for $k<\lambda$. For $k=\lambda$, we plug (4-24) in $R_{uv}^{(\lambda)}$.

Claim 4.5. *We can commute the covariant derivatives of* $R_{uv}^{(\lambda)}$.

Proof. To prove the claim, we only need to consider the case $\lambda\ge 2$. By the induction hypothesis of Lemma 4.4, the covariant derivatives of the curvature tensor vanish up to order $\lambda-1$ at p. If $\lambda>3$, the claim follows from the Ricci identity. Now suppose $\lambda=2$. By the Ricci identity, the difference of commuting the covariant derivatives is a function of the curvature tensor. Note that the curvature tensor at p is the same as of the complex space form. This completes the proof for $\lambda=2$. □

We insert (4-24) into $R_{Je_0Je_0}^{(\lambda)}$. By Claim 4.5 and the Bianchi identities, $R_{Je_0Je_0}^{(\lambda)}$ becomes a polynomial in the variables $z_j,\bar{z}_j$. The coefficients of the polynomial are exactly all the covariant derivatives of Rm at p of order λ. According to the assumption of Lemma 4.4, $R_{Je_0Je_0}^{(\lambda)}$ is identically 0 for all e_0. Therefore, the coefficients of the polynomial are all 0. This completes the induction of Lemma 4.4. □

Lemma 4.6. *Under the assumption of the induction in Proposition 2.1*, A *could be written as* $\sum_{i=1}^{m-2}g_{i,m}\Delta^i s$, *where* s *denotes the scalar curvature*, *and* $g_{i,m}$ *are constants depending only on* n, m *and* i.

Proof. Define $X=\frac{1}{2}(e_0-\sqrt{-1}Je_0)$, then $A=\int R_{X\overline{X}X\overline{X},e_0e_0\dots e_0}$, where the number of e_0 is $2m-4$. Integrating and plugging (4-24) into the result, we find

$$A=\sum_{\alpha_1\alpha_2\dots\alpha_{2m}}\left(\int\alpha_1\alpha_2\dots\alpha_{2m}\right)R_{\alpha_1\alpha_2\dots\alpha_{2m}} \tag{4-25}$$

where each α_i is either z_j or $\overline{z}_k$ for $0 \le j, k \le n-1$, with the further condition that $\alpha_1, \alpha_3 \in \{z_j\}$, and $\alpha_2, \alpha_4 \in \{\overline{z_k}\}$. Under the subscript of R, z_j stands for β_j, and $\overline{z}_k$ stands for $\overline{\beta}_k$.

From the expression of (4-25), we see that $z_i, \overline{z}_i$ must all go in pairs in the sequence $\alpha_1\alpha_2 \dots \alpha_{2m}$, otherwise the integral $\int \alpha_1\alpha_2 \dots \alpha_{2m}$ would equal 0. Using the Kähler identities, we can switch the covariant derivatives in (4-25) and rearrange it as

$$A = \sum_{I_1, I_2, \dots I_n} C_{I_1 I_2 \dots I_n} R_{I_1 I_2 \dots I_n} + B. \tag{4-26}$$

Here the symbol I_j denotes $z_j\overline{z}_j \dots z_j\overline{z}_j$; we have $\sum_j |I_j| = 2m$; subscripts after the fourth subscript of R denote covariant derivatives; $C_{I_1 I_2 \dots I_n}$ are the coefficients in (4-25); and B is a combination of covariant derivatives of Rm of lower order.

From (4-23), we see that the coefficients $C_{I_1 I_2 \dots I_n}$ in (4-26) are unitary invariants. For fixed $I_3, I_4, \dots, I_n$, let $d = |I_1| + |I_2|$. Denote the coefficient $C_{I_1 I_2 \dots I_n}$ by C_p, where $0 \le |I_1| = p \le d$. We want to find relations between the different C_p. Define a unitary transformation by setting $\widetilde{\beta_i} = \beta_i$ for $i \neq 1, 2$ and let

$$\beta_1 = \cos\theta \widetilde{\beta_1} + \sin\theta \widetilde{\beta_2} \quad \text{and} \quad \beta_2 = -\sin\theta \widetilde{\beta_1} + \cos\theta \widetilde{\beta_2}.$$

Insert the unitary transformation above in (4-26). Then the new coefficient $\tilde{C_d}$ becomes $\sum_{p=0}^{d} C_p \cos^{2p}\theta \sin^{2(d-p)}\theta$. Therefore we have:

$$\sum_{p=0}^{d} C_p \cos^{2p}\theta \sin^{2(d-p)}\theta = C_d = C_d(\cos^2\theta + \sin^2\theta)^d. \tag{4-27}$$

Claim 4.7. $C_p = C_d\binom{d}{p}$.

Proof. Divide by $cos^{2d}\theta$ on both sides, then (4-27) becomes

$$\sum_{p=0}^{d} C_p \tan^{2(d-p)}\theta = C_d = C_d(1 + \tan^2\theta)^d.$$

Since θ is arbitrary, the claim follows. □

By Claim 4.7, $C_p/C_d = \binom{d}{p}$. Since we can substitute any index u, v for 1, 2, the ratios of all coefficients in (4-26) are determined. Note that to get the relations between C_p, we only use the condition that the form (4-23) is unitary invariant. Since $\Delta^{m-2}s$ is also unitary invariant with respect to the frame, we can write it in the same form as (4-26). By the same argument, the ratios of coefficients of $\Delta^{m-2}s$ are the same as of coefficients in (4-26). It follows that the term $\sum_{I_1, I_2, \dots I_n} C_{I_1 I_2 \dots I_n} R_{I_1 I_2 \dots I_n}$ in (4-26) equals $C(m, n)\Delta^{(m-2)}s$ modulo lower order covariant derivatives, where $C(m, n)$ is a constant depending only on m, n.

Now we make an important observation. From the Ricci identity,

$$R_{i_1\bar{i}_2\dots i_p\alpha\beta i_{p+3}\dots i_{2m}} - R_{i_1\bar{i}_2\dots i_p\beta\alpha i_{p+3}\dots i_{2m}}$$

is the sum of $(Rm\,Rm^{(p-4)})_{,i_{p+3}\dots i_{2m}}$. By Lemma 4.4, $Rm^{(\lambda)} = 0$ for $1 \le \lambda \le m-3$. It follows that $(Rm\,Rm_{,i_5\dots ip})_{,i_{p+3}\dots i_{2m}}$ can be expanded as a linear combination of the covariant derivatives of curvature tensor. Therefore $A - C(m,n)\Delta^{(m-2)}s$ can be written as a linear combination of the covariant derivatives of the curvature tensor with the highest order $2m-6$. Furthermore it is unitary invariant since the curvature tensor is unitary invariant at p. By induction, we have completed the proof of Lemma 4.6. □

From the induction in Proposition 2.1, $\mathrm{Ric}^{(l)} = 0$ for $1 \le l \le 2m-4$. Integrating over the unit sphere in T_pM we find, by similar arguments as in the proof of Lemma 4.6, that for l even

$$0 = \int \mathrm{Ric}_{e_0e_0,e_0e_0\dots e_0} = \sum_{k=1}^{l/2} C_{l,k}\Delta^k s \tag{4-28}$$

where the order of the covariant derivative above is l. It is straightforward to check that the highest order coefficient $C_{l,l/2}$ is not equal to 0. Then, by induction, $\Delta^k s = 0$ at p for $1 \le k \le m-2$. Combining this with Lemma 4.6, it follows that $A = 0$. Similarly all linear terms in (4-10) vanish. Therefore, under the induction hypothesis in Proposition 2.1, in order that c_{2m} in (4-10) achieves the maximum, we must have $\mathrm{Ric}^{(2m-2)} = 0$ and $R^{(\lambda)}_{uv} = 0$ for $1 \le \lambda \le m-2$. This is exactly the case of the complex space form. Therefore we have completed the induction step for part (1) in Proposition 2.1 and, as a byproduct, we have proved part (2) as well. The proof is thus complete. □

5. The proof of Theorem 1.4

Proof of Proposition 2.2. Using the same argument as in the last section, we find that $W^{(2m+1)}$ is a linear combination of $\int R_{11}^{(m+i)}$ for $1 \le i \le m-3$ (the terms of order greater than $2m-3$ can be absorbed into $\mathrm{Ric}^{(m+i)}$). Then $W^{(2m+1)}$ is equal to 0 by similar arguments as in the proof of Lemma 4.6. □

Proof of Theorem 1.4. Consider the two cases below:

1. All coefficients of the power series of W are equal to that of the complex space form. From Proposition 2.1, all covariant derivatives of the curvature tensor at p are the same as the complex space form. Since the metric is real analytic, we conclude that near p, the manifold is isometric to the complex space form.

2. There is a $i_0 \ge 1$ such that for all $i < i_0$, the coefficients of the power series of W are equal to that of the complex space form, but the i_0-th coefficient is less than that of the complex space form. Checking the power series of W'/W at p,

we find that for sufficiently small r, W'/W is less than that of the complex space form. From the definition of W we have, for small r,

$$\frac{\int_{\partial B_p(r)} \Delta r}{A(\partial B_p(r))} = \frac{\int \sqrt{\det\langle J_u, J_v\rangle}\,'}{\int \sqrt{\det\langle J_u, J_v\rangle}} < \Delta_{N_K} r(r). \qquad \square$$

6. An example

In this section we give an example showing that the analogous Laplacian comparison theorem is not true on Kähler manifolds when the Ricci curvature is bounded from below by a nonzero constant. The example is in dimension 2. For higher dimensions, the construction is similar.

Identify $\mathbb{R}^4$ with $\mathbb{C}^2$ in the usual way. The corresponding almost complex structure J is given by

$$J\frac{\partial}{\partial x_1} = \frac{\partial}{\partial x_2}, \quad J\frac{\partial}{\partial x_2} = -\frac{\partial}{\partial x_1}, \quad J\frac{\partial}{\partial x_3} = \frac{\partial}{\partial x_4} \quad \text{and} \quad J\frac{\partial}{\partial x_4} = -\frac{\partial}{\partial x_3}.$$

Given a small ball near the origin of $\mathbb{C}^2$, define a function f to be

$$\begin{aligned} f = {} & |z_1|^2 + |z_2|^2 + a|z_1|^4 + 8a|z_1|^2|z_2|^2 + a|z_2|^4 + \tfrac{8}{3}a^2|z_1|^6 \\ & +28a^2|z_1|^4|z_2|^2 + 28a^2|z_1|^2|z_2|^4 + \tfrac{8}{3}a^2|z_2|^6 + p(|z_1|, |z_2|), \end{aligned}$$

where a is a nonzero constant and p is a homogeneous polynomial of degree 8 to be determined later. We define

$$\omega = \tfrac{\sqrt{-1}}{2}\partial\bar\partial f = \tfrac{\sqrt{-1}}{2}\sum_{i,j} g_{i\bar j}\, dz_i \wedge d\bar z_j.$$

It is straightforward to check that ω defines a Kähler metric g if the ball is sufficiently small (note that the metric is not complete).

Direct computation gives

$$\begin{aligned} g_{1\bar 1} = {} & 1 + 4a|z_1|^2 + 8a|z_2|^2 + 24a^2|z_1|^4 \\ & +112a^2|z_1|^2|z_2|^2 + 28a^2|z_2|^4 + O((|z_1| + |z_2|)^6), \end{aligned}$$

$$g_{1\bar 2} = 8a\bar z_1 z_2 + 56a^2 z_1\bar z_1^2 z_2 + 56a^2\bar z_1 z_2^2\bar z_2 + O((|z_1| + |z_2|)^6),$$

and $g_{2\bar 2} = g_{1\bar 1}$. Therefore

$$\begin{aligned} \det(g_{i\bar j}) &= g_{1\bar 1}g_{2\bar 2} - |g_{1\bar 2}|^2 \\ &= \big(1 + 4a|z_1|^2 + 8a|z_2|^2 + 24a^2|z_1|^4 + 112a^2|z_1|^2|z_2|^2 + 28a^2|z_2|^4\big)^2 \\ &\quad - \big|8a\bar z_1 z_2 + 56a^2 z_1\bar z_1^2 z_2 + 56a^2\bar z_1 z_2^2\bar z_2\big|^2 + O((|z_1| + |z_2|)^6) \\ &= 1 + 12a(|z_1|^2 + |z_2|^2) + 84a^2(|z_1|^4 + |z_2|^4) \\ &\quad +240a^2|z_1|^2|z_2|^2 + O((|z_1| + |z_2|)^6). \end{aligned}$$

Using $\log(1+x) = x - \frac{1}{2}x^2 + O(x^3)$, we have

$$\mathrm{Ric} + 12ag = \partial\bar{\partial}(-\log(\det g_{i\bar{j}}) + 12af) = \partial\bar{\partial}\, O((|z_1| + |z_2|)^6).$$

Therefore $\mathrm{Ric}+12ag$ vanishes up to order 3 at the origin. If we choose the function p to be $-\lambda(|z_1|^8 + |z_2|^8 + 8(|z_1|^6|z_2|^2 + |z_1|^2|z_2|^6))$, a direct computation gives

$$\mathrm{Ric} + 12ag = \partial\bar{\partial}(24\lambda(|z_1|^2 + |z_2|^2)^3 + O((|z_1| + |z_2|)^6)$$

where the term $O((|z_1| + |z_2|)^6)$ does not depend on λ. If λ is sufficiently large, $\mathrm{Ric} + 12ag \geq 0$ near the origin. Set $K = -12a$. Thus, near the origin, $\mathrm{Ric} \geq K$. By direct computation $R_{1212} = R_{1313} = R_{1414} = 4a$ and $R_{1u1v} = 0$ at the origin if $u \neq v$. Combining this with the fact that the second derivatives of the Ricci tensor vanish at the origin we find, after a slight computation, that the fourth order term of (3-3) is greater than that of the complex space form if $e_0 = \partial/\partial x_1$. So when r is very small, $\sqrt{\det\langle J_u, J_v\rangle}$ is greater than that of the complex space form along the geodesic with initial direction $\partial/\partial x_1$ at the origin. Since

$$\Delta r = \frac{\partial \log \sqrt{\det\langle J_u, J_v\rangle}}{\partial r},$$

it follows that the pointwise Laplacian comparison with the complex space forms is not true for Kähler manifolds.

Acknowledgments

The author is grateful to his advisor Professor Jiaping Wang for continuous encouragement and helpful discussions during the work.

References

[Bishop and Crittenden 1964] R. L. Bishop and R. J. Crittenden, *Geometry of manifolds*, Pure and Applied Mathematics **15**, Academic Press, New York, 1964. MR 29 #6401 Zbl 0132.16003

[Cheeger and Ebin 2008] J. Cheeger and D. G. Ebin, *Comparison theorems in Riemannian geometry*, AMS Chelsea Publishing, Providence, RI, 2008. Revised reprint of the 1975 original. MR 2009c:53043 Zbl 1142.53003

[Gromov 1981] M. Gromov, *Structures métriques pour les variétés riemanniennes*, Textes Mathématiques **1**, CEDIC, Paris, 1981. MR 85e:53051 Zbl 0509.53034

[Li 1993] P. Li, *Lecture notes on geometric analysis*, Lecture Notes Series **6**, Research Institute of Mathematics, Seoul, 1993. MR 96m:58269 Zbl 0822.58001

[Li and Wang 2005] P. Li and J. Wang, "Comparison theorem for Kähler manifolds and positivity of spectrum", *J. Differential Geom.* **69**:1 (2005), 43–74. MR 2006g:53111 Zbl 1087.53067

[Schoen and Yau 1994] R. Schoen and S.-T. Yau, *Lectures on differential geometry*, Conference Proceedings and Lecture Notes in Geometry and Topology **1**, International Press, Cambridge, MA, 1994. MR 97d:53001 Zbl 0830.53001

Received March 14, 2011.

GANG LIU
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MINNESOTA
127 VINCENT HALL, 206 CHURCH ST. SE
MINNEAPOLIS, MN 55455
UNITED STATES
liuxx895@umn.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

STRUCTURABLE ALGEBRAS OF SKEW-RANK 1 OVER THE AFFINE PLANE

SUSANNE PUMPLÜN

Let k be a field of characteristic not 2 or 3. Infinitely many mutually non-isomorphic structurable algebras of rank 20 over $k[X, Y]$ are constructed whose fiber is a given structurable algebra over k of skew-rank 1.

Introduction

Let R be a ring such that $\frac{1}{6} \in R$ and k a field of characteristic not 2 or 3. Let A be a unital nonassociative algebra over R with an involution $^{-}$. The pair $(A, ^{-})$ is called a *structurable algebra* if

$$\{x, y, \{z, w, q\}\} - \{z, w, \{x, y, q\}\} = \{\{x, y, z\}, w, q\} - \{z, \{y, x, w\}, q\}$$

for $x, y, z, w, q \in A$, where

$$\{x, y, z\} = (x\bar{y})z + (z\bar{y})x - (z\bar{x})y.$$

Structurable algebras were introduced in [Allison 1978]: an analogue of the Koecher–Kantor–Tits functor gives a correspondence between a structurable algebra and a Lie algebra. Using this functor all classical simple isotropic Lie algebras can be obtained [Allison 1979].

In [Parimala et al. 1999], nontrivial Albert algebra bundles over the affine plane were constructed whose associated principal F_4 bundle admits no reduction of the structure group to any proper connected reductive subgroup. (For an analogous result with the associated principal G bundle being of type G_2, see [Parimala et al. 1997; 1999].) Over a field, every Albert algebra arises from the first or second Tits construction and the associated F_4 bundle admits reduction of the structure group to $\mathrm{SL}_1(B)$ for a central simple algebra B either over k or to $\mathrm{SU}(B, \sigma)$ for a central simple algebra B over a quadratic field extension of k, σ an involution of the second type. Hence the patched Albert algebras over the affine plane arise neither from a first nor a second Tits construction (and correspondingly, there are

MSC2010: 17C40.
Keywords: structurable algebra, Jordan algebra, affine plane, principal bundle.

patched octonion algebras over the affine plane which do not arise from a Cayley-Dickson doubling process or are constructed involving a ternary hermitian form and a two-dimensional subalgebra).

In the present paper we employ the patching arguments from [Parimala et al. 1999] to obtain infinitely many structurable algebras M_i of rank 20 over the affine plane $\mathbb{A}_k^2$, which are not extended from k and mutually nonisomorphic and whose fiber is a given matrix algebra over k (Theorem 4). In order to achieve this, we show that the matrix algebra $M(T, N, N^\vee)$ over $k[X, Y]$ admits a unique extension to a matrix algebra over $\mathbb{P}_k^2$ in Section 2. In Section 3, we look at forms of these matrix algebras. For a nonfree projective left $D[X, Y]$-module P of rank one, the structurable algebra $S(D, \sigma, P, N)$ over $k[X, Y]$ admits a unique extension to a structurable algebra $S(\mathcal{D}, \sigma, \widetilde{P}, N)$ over $\mathbb{P}_k^2$, where $\widetilde{P}$ is an indecomposable vector bundle. We use this result to construct infinitely many mutually nonisomorphic structurable algebras A^i over $\mathbb{A}_k^2$ such that

$$A^i \otimes_k K \cong M_i,$$

where K is a separable quadratic field extension of k (Theorem 9). In Section 4, some general results on extending structurable algebras from affine to projective space are obtained.

If a structurable algebra over $\mathbb{A}_k^2$ has rank 56, it corresponds to the structure group E_7. Such bundles were constructed in [Raghunathan 1989] for a connected reductive absolutely almost simple k-group G, which is k-anisotropic and is not of type F_4 or G_2 (for the type G_2 and F_4, see [Knus et al. 1994; Parimala et al. 1997; 1999]). The results show that although $GL(r)$-bundles over the affine plane $\mathbb{A}_k^n$ are trivial, this is not the case for a general reducible structure group.

It is also known that if G is a k-anisotropic reductive absolutely almost simple algebraic k-group, there are infinite families of mutually nonisomorphic, nontrivial (sometimes indecomposable) principal G-bundles over $\mathbb{A}_k^2$, which do not admit a reduction of its structure group to any proper connected reductive subgroup of G.

The author is not able to say whether the new principal G-bundle constructed in this paper admit a reduction of their structure group to a proper reductive subgroup or not.

We use the results and terminology from [Achhammer 1995] (see also [Pumplün 2008; 2010a; 2010b) and [Parimala et al. 1999]. The approach in this last work is mostly functorial and formulated for base rings R which are domains with $\frac{1}{6} \in R$, the one in [Achhammer 1995] works instead for arbitrary base rings. Both were originally developed to generalize the first and second Tits construction for Jordan algebras over rings.

For the standard terminology on Jordan algebras, see [McCrimmon 2004; Jacobson 1968; Schafer 1966].

1. Preliminaries

1.1. ***Algebras over R.*** For $P \in \operatorname{Spec} R$, let R_P be the local ring of R at P and m_P the maximal ideal of R_P. The corresponding residue class field is denoted by $k(P) = R_P/m_P$. For an R-module F the localization of F at P is denoted by F_P. The *rank* of F is defined to be $\sup\{\operatorname{rank}_{R_P} F_P \mid P \in \operatorname{Spec} R\}$. The term "$R$-algebra" always refers to nonassociative R-algebras which are unital and finitely generated projective of finite constant rank as R-modules.

An antiautomorphism $\sigma : A \to A$ of order 2 is called an *involution* on A. Define $H(A, \sigma) = \{a \in A \mid \sigma(a) = a\}$ and $\mathrm{S}(A, \sigma) = \{a \in A \mid \sigma(a) = -a\}$. Then $A = H(A, \sigma) \oplus \mathrm{S}(A, \sigma)$.

1.2. ***Structurable algebras.*** An *algebra with involution* is a pair $(A, ^-)$ consisting of an R-algebra A and an involution $^- : A \to A$. A *structurable algebra* is an algebra with involution $(A, ^-)$ satisfying

$$\{x, y, \{z, w, q\}\} - \{z, w, \{x, y, q\}\} = \{\{x, y, z\}, w, q\} - \{z, \{y, x, w\}, q\}$$

for all elements $x, y, z, w, q \in A$, where

$$\{x, y, z\} = (x\bar{y})z + (z\bar{y})x - (z\bar{x})y$$

[Allison 1978, (3) and Corollary 5]. If B is an R-submodule if A closed under multiplication, we call B a subalgebra of A. If, additionally, $\overline{B} = B$ we call $(B, ^-)$ a subalgebra of $(A, ^-)$.

An *isotopy* from $(A, ^-) \to (A', ^{-\prime})$ is an R-linear bijective map $a : A \to A'$ such that

$$a\{x, y, z\} = \{ax, \hat{a}y, az\}$$

for all $x, y, z \in A$ and some R-linear map $\hat{a} : A \to A'$. Two structurable algebras (A^-) and $(A, ^-)$ are *isotopic* if there exists an isotopy from A to A'. This is equivalent to $(A', ^{-\prime}) \cong (A, ^-)^{\langle u \rangle}$ for some invertible $u \in A$. Every isomorphism between structurable algebras is an isotopy.

In the following, we will only deal with structurable algebras $(A, ^-)$ over R whose residue class algebras $A(P) = A_P \otimes_{R_P} k(P)$ are central simple structurable algebras of skew-dimension 1.

1.3. Let W and W' be two finitely generated projective R-modules of constant rank with cubic forms $N : W \to R$ and $N' : W' \to R$, paired by a nondegenerate bilinear form $T : W \times W' \to R$. That is, T induces R-module isomorphisms

$$T : W \to \operatorname{Hom}_R(W', R), \quad x \mapsto T(x, \cdot)$$

and

$$T : W' \to \operatorname{Hom}_R(W, R), \quad y' \mapsto T(\cdot, y').$$

We say that the triple (T, N, N') *is defined on* (W, W'). Let $N(x, y, z)$ denote the trilinear form associated with N and $N'(x', y', z')$ the trilinear form associated with N'. Let $x \in W$, $x' \in W'$ and define quadratic maps

$$\sharp : W \to W' \quad \text{and} \quad \sharp' : W' \to W$$

via

$$D_y N(x) = T(y, x^{\sharp}) \quad \text{and} \quad D_{y'} N'(x') = T(x'^{\sharp'}, y')$$

for all elements $x, y \in W$, $x', y' \in W'$; i.e.,

$$3N(x, x, y) = T(y, x^{\sharp}) \quad \text{and} \quad 3N'(x', x', y') = T(x'^{\sharp'}, y')$$

for all elements $x, y \in W$, $x', y' \in W'$. The triple (T, N, N') *satisfies the adjoint identities* if

$$(x^{\sharp})^{\sharp'} = N(x)x \quad \text{and} \quad (x'^{\sharp'})^{\sharp} = N'(x')x'.$$

If $N = 0$ and $N' = 0$ these identities are trivially satisfied. If $N \neq 0$ or $N' \neq 0$ then both N and N' are nonzero and (T, N, N') is called *nontrivial*.

Let (T, N, N') be a triple defined on (W, W'). Define symmetric bilinear maps $\times : W \times W \to W'$ and $\times' : W' \times W' \to W$ via

$$x \times y = (x + y)^{\sharp} - x^{\sharp} - y^{\sharp}, \quad x' \times' y' = (x' + y')^{\sharp'} - x'^{\sharp'} - y'^{\sharp'}.$$

Then

$$x^{\sharp} = \tfrac{1}{2} x \times x, \quad x'^{\sharp'} = \tfrac{1}{2} x' \times' x',$$
$$N(x, y, z) = T(x, y \times z), \quad N'(x', y', z') = T(x' \times' y', z').$$

If the triple (T, N, N') satisfies the adjoint identities then the matrix algebra

$$A = M(T, N, N') = \begin{bmatrix} R & W \\ W' & R \end{bmatrix}$$

with multiplication

$$\begin{bmatrix} a & x \\ x' & b \end{bmatrix} \begin{bmatrix} c & y \\ y' & d \end{bmatrix} = \begin{bmatrix} ac + T(x, y') & ay + dx + x' \times' y' \\ cx' + by' + x \times y & bd + T(y, x') \end{bmatrix}$$

and involution

$$\overline{\begin{bmatrix} a & x \\ x' & b \end{bmatrix}} = \begin{bmatrix} b & x \\ x' & a \end{bmatrix}$$

is a structurable algebra [Allison and Faulkner 1984, p. 194; [Pumplün 2010b, Theorem 1]. We have $S(A, ^-) = s_0 R$ with

$$s_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

invertible and $s_0^2 = 1 \in R^\times$ and the residue class algebras $A(P) = A_P \otimes k(P)$ are central simple structurable algebras of skew-dimension 1 over $k(P)$ [Allison and Faulkner 1984; Pumplün 2010b]. Let

$$u = \begin{bmatrix} a & x \\ x' & b \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} c & y \\ y' & d \end{bmatrix}$$

with $a, b, c, d \in R$ and $x, y \in W$, $x', y' \in W'$. The (*conjugate*) norm

$$\nu : M(T, N, N') \to R$$

is given by

$$\nu(u) = 4aN(x) + 4bN'(x') - 4T(x'^{\sharp'}, x^\sharp) + (ab - T(x, x'))^2$$

and is isotropic since $\nu(u) = 0$ for

$$u = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The *trace* $\chi : M(T, N, N') \times M(T, N, N') \to R$ is defined by

$$\chi(u, v) = 2(ad + bc + T(x, y') + T(y, x')).$$

Note that $\chi(u, u) = 0$.

1.4. Let B be an Azumaya algebra over R of degree 3, $B^+ = (N_B, \sharp_B, 1)$ with $(N_B, \sharp_B, 1)$ a cubic form with adjoint and base point (see for instance [Pumplün 2010b, 1.4]). Let $\mathrm{Pic}_l B$ denote the set of isomorphism classes of locally free left B-modules of rank 1. Let $P \in \mathrm{Pic}_l B$ such that $N_B(P) \cong R$ and let $N : P \to R$ be a norm on P. Let $N^\vee : P^\vee \to R$ be the uniquely determined norm and $\sharp : P \to P^\vee$, $\check{\sharp} : P^\vee \to P$ be the uniquely determined adjoints satisfying

(1) $\langle w, w^\sharp \rangle = N(w)1$,

(2) $\langle \check{w}^{\check{\sharp}}, \check{w} \rangle = N^\vee(\check{w})1$, and

(3) $w^{\sharp\check{\sharp}} = N(w)w$

for all $w \in P$, $\check{w} \in P^\vee$ (these are identities (7), (8), (9) in [Pumplün 2010b]). Let $\times : P^\vee \times P^\vee \to P$ denote the bilinear map associated to the quadratic map $\sharp$ and $\check{\times} : P^\vee \times P^\vee \to P$ the bilinear map associated to the quadratic map $\check{\sharp}$ (see for instance [ibid., 3.2]). Define $T : P \times P^\vee \to R$ via

$$T(w, \check{w}) = T_B(\langle w, \check{w} \rangle).$$

For any $\mu \in R^\times$, the triple $(\mu T, \mu N, \mu^2 N^\vee)$ satisfies the adjoint identities [Pumplün 2010b, Theorem 6], hence

$$M = M(\mu T, \mu N, \mu^2 N^\vee) = \begin{bmatrix} R & P \\ P^\vee & R \end{bmatrix}$$

is a structurable algebra over R with automorphism group isomorphic to the semidirect product of $\mathbb{Z}/2$ and the group of bijective norm isometries of P; see [ibid., Corollary 7 and Theorem 18].

The group $\mathrm{Inv}(M)$ defined in Section 4 is an absolutely almost simple linear algebraic group, which is connected except in the case that M has rank 9. In that case its connected component is a subgroup of index 2 in $\mathrm{Inv}(M)$ [Krutelevich 2007, p. 941 ff.].

1.5. Let R' be a ring and B a unital separable associative algebra over R'. Let $* : R' \to R'$ be an involution on R' and $*_B$ an involution on B such that $*_B|_{R'} = *$. Let $(N_B, \sharp_B, 1)$ be a cubic form with adjoint and base point on B such that $B^+ = J(N_B, \sharp_B, 1)$, with 1 the unit element in B, and that the conditions

$$\begin{gathered} xyx = T_B(x, y)x - x^{\sharp_B} \times_B y, \\ N_B(xy) = N_B(x)N_B(y), \\ N(x^{*_B}) = N(x)^{*_B} \end{gathered}$$

are satisfied for all $x, y \in B$ (these are identities (1), (2), (3) in [Pumplün 2010b]). Let $(H(B, *_B), H(R', *_B))$ be a B-ample pair, and define $R = H(R', *_B)$. Let $P \in \mathrm{Pic}_l\, B$ be such that $N_B(P) \cong R'$ and such that there is a nondegenerate *hermitian* form $h : P \times P \longrightarrow B$ satisfying

$$h(w, w) \in H(B, *_B) \quad \text{and} \quad N_B(h(w, w)) = N(w)N(w)^{*_B}$$

for $w \in P$. Denote by $*$ the $H(B, *_B)$-admissible involution $j_h : P \to \overline{P}^\vee$ on P induced by h. Let $N : P \to R'$ be a norm on P. Let $N^\vee : P^\vee \to R'$ be the uniquely determined norm and $\sharp : P \to P^\vee$, $\check{\sharp} : P^\vee \to P$ be the uniquely determined adjoints satisfying equations (1), (2), (3). We can also write

$$\langle u, v^* \rangle = h(u, v), \quad v^* = j_h(v) \quad \text{and} \quad \check{v}^{\check{*}} = j_h^{-1}(\check{v})$$

for $j_h : P \to \overline{P}^\vee$ induced by h. The R-module $S(B, *_B, P, N, h) = R' \oplus P$ together with the multiplication

$$(a, u)(b, v) = (ab + T_B(\langle u, v^* \rangle), b^{*_B} u + av + (u \times v)^{\check{*}})$$

and the involution

$$\overline{(a, u)} = (\bar{a}, u)$$

for $a, b \in R'$, $u, v \in P$ is a structurable algebra over R, which is a form of the structurable algebra $M(T, N, N^\vee)$ [Pumplün 2010b, Theorem 20]. We define the (*conjugate*) *norm* $\nu : S(B, *_B, P, N, h) \to R$ of $S(B, *_B, P, N, h)$ via

$$\nu((\lambda, w)) = N_B(\lambda\lambda^* - h(w, w)).$$

If R' is a field this definition coincides with [Allison and Faulkner 1992, Theorem 6.1]. ν is a quartic form. Even if B is a division algebra and R' is a field, the norm is isotropic: then $\nu((\lambda, w)) = 0$ if and only if (λ, w) is an admissible scalar; i.e., $\mu \in R'^\times$, $w \in H(B, *_B)^\times$ and $N_B(w) = \mu\mu^*$.

If R' is a quadratic étale ring extension of the ring R then $R' = \mathrm{Cay}(R, P, N)$ with $L \in \mathrm{Pic}\, R$ of order 2, since $2 \in R^\times$. For $A = S(B, *_B, P, N, h)$ this means $S(A, {}^-) = \{(r, 0) \mid r \in S(R', *)\} = L$. If R is a domain and $R' = \mathrm{Cay}(R, c) = R(\sqrt{c})$ then $S(A, {}^-) = (\sqrt{c}, 0)R$ and $s_0 = (\sqrt{c}, 0)$ satisfies $s_0^2 = (c, 0) = c1_A$ with $c \in R^\times$. This means we can define the (conjugate) norm $\nu : A \to R$ also by

$$\nu(x) = \frac{1}{12c}\chi(s_0 x, \{x, s_0 x, x\}),$$

and also a *trace* $\chi : A \times A \to R$ on A by

$$\chi(x, y) = \frac{2}{c}\psi(s_0 x, y)s_0 = \frac{2}{c}(V^\delta_{y,x} s_0)s_0,$$

analogously as in [Allison and Faulkner 1984; 1992], where $\psi(x, y) = x\bar{y} - y\bar{x}$ [Allison and Faulkner 1992, 5.4]. χ is a nondegenerate symmetric bilinear form independent of the choice of s_0 and $\chi(1, 1) = 4$. (Nondegeneracy follows from [Allison and Faulkner 1984, Proposition 2.5] applied to the residue class forms.)

2. Nontrivial structurable algebras over the affine plane which locally are matrix algebras

2.1. We mostly use the results and notation of [Parimala et al. 1999, Section 4]. Occasionally, we also use the notation of [Pumplün 2008]: in the notation of [Parimala et al. 1999], the map $\times$ in [Pumplün 2008] or Section 1.4 is denoted by ϕ and the map $\check{\times}$ in Section 1.4 by ϕ_*. There is the obvious notion of a structurable algebra over a locally ringed space; see [Pumplün 2010b]. We identify structurable algebras over $k[X, Y]$ and over $\mathbb{A}^2_k$ using the canonical equivalence described in [ibid., 6.2]. Let $X = \mathbb{P}^2_k$.

Remark 1. Let D be a central simple algebra over k of degree 3. Once we have picked a locally free left $D[X, Y]$-module of rank 1 with $N_{D[X,Y]}(P) \cong k[X, Y]$, the choice of a norm $N : P \to k[X, Y]$ automatically determines $N^\vee$ and the adjoints $\sharp$ and $\check{\sharp}$; see [ibid., 3.2]. This fact is expressed in [Parimala et al. 1999] by explicitly choosing a trivialization $\widetilde{\mu} : N_{D[X,Y]}(P) \to k[X, Y]$ which in turn determines uniquely the choice of N, hence of $N^\vee$, $\sharp$ and $\check{\sharp}$. Recall that the norm N is uniquely determined up to a scalar $\mu \in k^\times$. For any $\mu \in k^\times$, the adjoint belonging to μN is $\mu\,\sharp$ and $(\mu N)^\vee = \mu^2 N^\vee$, $(\mu\,\sharp)^\vee = \mu^2\,\sharp$.

2.2. Let D be a central division algebra over k of degree 3. Let De be a free module of rank 1 over D with e as a basis element such that $N_D(De) \cong k$ and let $\mu_0 : N_D(De) \to k$ be such an isomorphism. Let $\{g_i\}$ be an infinite family of mutually coprime polynomials in $k[X]$. Then there exist nonfree projective left modules P_i of rank 1 over $D[X,Y]$ and polynomials $f_i \in k[X]$ with $(f_i, f_j) = 1$ for $i \neq j$, $(f_i, g_j) = 1$ for all i, j, such that $P_i \otimes k[X]_{f_i}[Y]$ is free for each i. Further, there exists

$$\tilde{\mu}_i : N_{D[X,Y]}(P_i) \to k[X,Y]$$

such that $(P_i, \tilde{\mu}_i)$ modulo Y is $(De, \mu_0) \otimes_k k[X]$; see [Parimala et al. 1999, 4.1]. The P_i are mutually nonisomorphic $D[X,Y]$-modules [ibid., 4.2].

2.3. Let P be a nonfree projective $D[X,Y]$-module such that $N_{D[X,Y]}(P) \cong k[X,Y]$, the isomorphism given by the trivialization $\tilde{\mu} : N_{D[X,Y]}(P) \to k[X,Y]$ of the reduced norm. Then the pair $(P, \tilde{\mu})$ is a principal $\mathrm{SL}_1(D)$-bundle over $\mathbb{A}^2_k$ which admits an extension $(\tilde{P}, \tilde{\mu})$ to $\mathbb{P}^2_k$; the bundle $\tilde{P}$ is simply an extension of the $D[X,Y]$-module P [ibid., p. 31] (by abuse of notation, we denote both $\tilde{\mu}$ and its extension by the same name). Let $N : P \to k[X,Y]$ be the norm on P determined by the choice of the trivialization $\tilde{\mu}$. The choice of $\tilde{\mu}$ also determines the maps $\times : P \times P \to P^\vee$, $\check{\times} : P^\vee \times P^\vee \to P$, and $N^\vee$, and therefore also $\sharp$ and $\check{\sharp}$; see [ibid., p. 16]. Take $T(u, \check{v}) = T_{D[X,Y]}(\langle u, \check{v}\rangle)$. The adjoints satisfy the adjoint identities [ibid., 1.2].

Analogously, $\tilde{\mu}$ determines extensions

$$(\dagger) \qquad \tilde{N} : \tilde{P} \to \mathbb{O}_X, \quad \tilde{N}^\vee : \tilde{P}^\vee \to \mathbb{O}_X, \quad \tilde{\sharp}, \quad \text{and} \quad \tilde{\check{\sharp}}$$

of N, $N^\vee$, $\sharp$ and $\check{\sharp}$, respectively, which satisfy the adjoint identities. Let $\tilde{T}(u, \check{v}) = T_{D\otimes\mathbb{O}_X}(\langle u, \check{v}\rangle)$.

Proposition 2. *The matrix algebra*

$$M(T, N, N^\vee) = \begin{bmatrix} k[X,Y] & P \\ P^\vee & k[X,Y] \end{bmatrix}$$

over $k[X,Y]$ admits a unique extension to a matrix algebra

$$M(\tilde{T}, \tilde{N}, \tilde{N}^\vee) = \begin{bmatrix} \mathbb{O}_X & \tilde{P} \\ \tilde{P}^\vee & \mathbb{O}_X \end{bmatrix}$$

over $\mathbb{P}^2_k$. The vector bundles $\tilde{P}$ and $\tilde{P}^\vee$ are indecomposable and $\tilde{P}$ and $\tilde{P}^\vee$ are not isomorphic as vector bundles on X.

Proof. There is a unique extension $\widetilde{P}$ over $X = \mathbb{P}^2_k$ of P of norm one that is a locally free right $D \otimes \mathcal{O}_X$-module: by [Parimala et al. 1999, p. 29], P extends to a vector bundle $\widetilde{P}$, which is unique up to a line bundle $\mathcal{L}$. Since we require $\widetilde{P}$ to be of norm one this implies $\mathcal{L}^3 \cong \mathcal{O}_X$, hence $\mathcal{L} = \mathcal{O}_X$ and the extension is unique. Let $N : P \to k[X, Y]$ be the norm on P determined by the choice of the trivialization $\widetilde{\mu}$. Two extensions $\widetilde{N} : \widetilde{P} \to \mathcal{O}_X$ and, say $\widetilde{N}' : \widetilde{P} \to \mathcal{O}_X$ of N, can only differ by a scalar $\lambda \in k^\times$. Being its extension, the algebra $M(\widetilde{T}, \widetilde{N}, \widetilde{N}^\vee)$ restricts to the structurable matrix algebra

$$M(T, N, N^\vee) = \begin{bmatrix} k[X,Y] & P \\ P^\vee & k[X,Y] \end{bmatrix}$$

over $\mathbb{A}^2_k$. Therefore $\widetilde{N}|_{\mathbb{A}^2_k} = N = \widetilde{N}'|_{\mathbb{A}^2_k}$ implies that $\lambda = 1$. Thus the maps listed in (†), which are the extensions of N, $N^\vee$, $\sharp$ and $\check{\sharp}$ from $\mathbb{A}^2_k$ to $\mathbb{P}^2_k$ determined by the trivializations $\widetilde{\mu}$ and μ, are uniquely determined as well.[-2pt]

The proof of the second statement follows from [Parimala et al. 1999, 3.2]. □

More precisely, by [ibid., Remark] and [Arason et al. 1992], $\widetilde{P} \cong tr_{l/k}(\mathcal{P}_0)$ for some cubic field extension l/k and a suitable vector bundle $\mathcal{P}_0$ over $\mathbb{P}^2_l$ that is absolutely indecomposable and of rank 3.

2.4. Let J be an Albert algebra over k that is a first Tits construction and a division algebra. Choose two cyclic division algebras D_1, D_2 of degree 3 over k such that the Jordan algebras D_1^+ and D_2^+ are subalgebras of J with $D_1^+ \cap D_2^+ = k$. By [Parimala et al. 1999, 4.3], these can be even chosen such that $D_2^+ = \Phi(D_1^+)$ for a suitable automorphism Φ of J; that is, we can and will assume that additionally we have $D_1^+ \cong D_2^+$. Then $J = J(D_1, e_1, \mu_1) = J(D_2, e_2, \mu_2)$ for some $e_i \in J$ and isomorphisms $\mu_i : N(D_i e_i) \to k$. Again, the choice of μ_i determines a norm $N_i : D_i \to k$, (a scalar multiple of N_{D_i}) and an adjoint $\sharp_i : D_i \to D_i$ (a scalar multiple of $\sharp_{D_i}$), so with $T_i(a, b) = T_{D_i}(ab)$ we obtain the structurable algebra

$$M = M(T_1, N_1, N_1) \cong M(T_2, N_2, N_2)$$

over k. By 2.2, for every $i \geq 1$ there exists a pair $(P_i^1, \widetilde{\mu}_i^1)$, where P_i^1 is a nonfree projective $D_1[X, Y]$-module of rank 1 and $\widetilde{\mu}_i^1$ a trivialization of its reduced norm and a polynomial $f_i \in k[X]$ such that:

(4) The polynomials f_i and f_j are coprime for $i \neq j$ and $(P_i^1)_{f_i}$ is free.

(5) The reduction of $(P_i^1, \widetilde{\mu}_i^1)$ modulo Y is $(D_1 e_1, \mu_1) \otimes k[X]$.

Similarly, for every $i \geq 1$, there is a pair $(P_i^2, \widetilde{\mu}_i^2)$, where P_i^2 is a nonfree projective $D_2[X, Y]$-module of rank 1 and $\widetilde{\mu}_i^2$ a trivialization of its reduced norm and a polynomial $g_i \in k[X]$ such that:

(6) The polynomials g_i and g_j are coprime for $i \neq j$, the polynomials f_i and g_j are coprime for all i, j, and $(P_i^2)_{g_i}$ is free.

(7) The reduction of $(P_i^2, \widetilde{\mu}_i^2)$ modulo Y is $(D_2 e_2, \mu_2) \otimes k[X]$.

For each pair $(P_i^j, \widetilde{\mu}_i^j)$, $j = 1, 2$, let

$$N_i^j : P_i^j \to k[X, Y]$$

be the norm on P_i^j induced by $\widetilde{\mu}_i^j$, let

$$T_i^j : P_i^j \times (P_i^j)^\vee \to k[X, Y]$$

be the usual trace, given by $T_i^j(u, \check{v}) = T_{D_j}(\langle u, \check{v}\rangle)$, and let $\sharp_i^j$ be the induced adjoint.

Define matrix algebras

$$M_i^1 = M(T_i^1, N_i^1, N_i^{1\vee}) = \begin{bmatrix} k[X,Y] & P_i^1 \\ (P_i^1)^\vee & k[X,Y] \end{bmatrix}$$

and

$$M_i^2 = M(T_i^2, N_i^2, N_i^{2\vee}) = \begin{bmatrix} k[X,Y] & P_i^2 \\ (P_i^2)^\vee & k[X,Y] \end{bmatrix}$$

of rank 20. Then $\{M_i^j \mid j = 1, 2, i \geq 1\}$ is a family of structurable algebras over $k[X, Y]$ such that $M_i^j = M \otimes k[X]$ modulo Y and

$$M_i^1 \otimes k[X]_{f_i}[Y] \cong M \otimes k[X]_{f_i}[Y], \quad M_i^2 \otimes k[X]_{g_i}[Y] \cong M \otimes k[X]_{g_i}[Y],$$

with $(f_i, f_j) = 1 = (g_i, g_j)$ for $i \neq j$, $(f_i, g_j) = 1$ for all i, j. As in [Parimala et al. 1997, 4.5] we can then conclude:

Proposition 3. *The matrix algebras M_i^1, respectively M_i^2, over $k[X, Y]$ are mutually nonisomorphic.*

Proof. Suppose there are $i \neq j$ such that $M_i^1 \cong M_j^1$. Since M_i^1 and M_j^1 are extended after inverting f_i and f_j, respectively, and since $(f_i, f_j) = 1$, M_i^1 is extended from $M \otimes k[X]$. Let $\tau : X \to k$ be the structure morphism. Since the extension $\widetilde{M}_i^1$ of M_i^1 to $\mathbb{P}_k^2$ is unique, it must be thus isomorphic to $\tau^*(M)$. Therefore, the underlying vector bundles must be isomorphic; i.e.,

$$\mathbb{O}_X^2 \oplus \widetilde{P}_i^1 \oplus (\widetilde{P}_i^1)^\vee \cong \mathbb{O}_X^{20}.$$

This is a contradiction, since $\widetilde{P}_i^1$ is an indecomposable vector bundle by [Parimala et al. 1999, 3.2]. □

2.5. Let

$$\pi_i^1 : (P_i^1, \widetilde{\mu}_i^1) \otimes k[X]_{f_i}[Y] \to (D_1 e_1, \mu_1) \otimes k[X]_{f_i}[Y]$$

and

$$\pi_i^2 : (P_i^2, \widetilde{\mu}_i^2) \otimes k[X]_{g_i}[Y] \to (D_1 e_2, \mu_2) \otimes k[X]_{g_i}[Y]$$

be isomorphisms such that $\overline{\pi}_i^j = \text{id}$, $j = 1, 2$ (we may assume this by [Parimala et al. 1997, 6.1]). These canonically induce isomorphisms

$$M(\pi_i^1) : M_i^1 \otimes k[X]_{f_i}[Y] \to M \otimes k[X]_{f_i}[Y]$$

and

$$M(\pi_i^2) : M_i^2 \otimes k[X]_{g_i}[Y] \to M \otimes k[X]_{g_i}[Y]$$

with $\overline{M(\pi_i^j)} = \text{id}$, $j = 1, 2$. Let M_i be the structurable algebra obtained by patching M_i^1 on $k[X]_{g_i}[Y]$ and M_i^2 on $k[X]_{f_i}[Y]$ over $k[X]_{f_i g_i}[Y]$ by $\phi_i = M(\pi_i^2)^{-1} M(\pi_i^1)$.

We obtain an involution $^- : M_i \to M_i$ by analogously patching the involutions of M_i^1 on $k[X]_{g_i}[Y]$ and of M_i^2 on $k[X]_{f_i}[Y]$ over $k[X]_{f_i g_i}[Y]$ by $\phi_i = M(\pi_i^2)^{-1} M(\pi_i^1)$.

Since $\overline{M}_i^j = M$ modulo Y and $\overline{M(\pi_i^j)} = \text{id}$, we get $\overline{\phi}_i = \text{id}$ and $\overline{M}_i = M \otimes k[X]$ modulo Y. By construction,

$$M_i \otimes k[X]_{f_i g_i}[Y] \cong M \otimes k[X]_{f_i g_i}[Y]$$

and the polynomials $r_i := f_i g_i$ are mutually coprime. The algebras M_i are mutually nonisomorphic by the same argument as given in [Parimala et al. 1999, p. 33], and thus we can conclude:

Theorem 4. *The structurable algebras M_i on $\mathbb{A}_k^2$ have the following properties*:

(i) $\overline{M}_i = M \otimes k[X]$ *modulo* Y.

(ii) *There are mutually coprime polynomials $r_i \in k[X]$ such that $M_i \otimes k[X]_{r_i}[Y] \cong M \otimes k[X]_{r_i}[Y]$.*

(iii) *The algebras M_i are nonextended and mutually nonisomorphic.*

Proof. By construction, we have

$$M_i \otimes k[X]_{f_i g_i}[Y] \cong M \otimes k[X]_{f_i g_i}[Y]$$

and the polynomials $r_i = f_i g_i$ are mutually coprime. To show that the algebras M_i are mutually nonisomorphic, suppose that there are $i \neq j$ such that $M_i \cong M_j$. Then both $(M_i)_{r_i}$ and $(M_i)_{r_j}$ are extended from M. Since $(r_i, r_j) = 1$, $M_i \cong M \otimes k[X, Y]$. Restrict M_i to $k[X]_{g_i}[Y]$. This yields that $M_i^1 \otimes k[X]_{g_i}[Y]$ and $M_i^1 \otimes k[X]_{f_i}[Y]$ are extended. Since $(f_i, g_i) = 1$, M_i^1 is extended from M. This contradicts Proposition 3. □

Note that all the ingredients for these proofs have been provided in [Parimala et al. 1999, Section 4].

It is not clear that these structurable algebras are again matrix algebras. We are not able to say if the corresponding principal G-bundle P_{M_i} admits reduction of the structure group to a proper reductive subgroup of G or not. They are subalgebras of a 56-dimensional matrix algebra:

2.6. Let J_i^1 and J_i^2 be the infinitely many mutually nonisomorphic Albert algebras over $k[X,Y]$ used in [Parimala et al. 1999, Proposition 4.5]. They give rise to infinitely many matrix algebras

$$M(J_i^1)=\begin{bmatrix} k[X,Y] & J_i^1 \\ J_i^1 & k[X,Y]\end{bmatrix} \quad\text{and}\quad M(J_i^2)=\begin{bmatrix} k[X,Y] & J_i^2 \\ J_i^2 & k[X,Y]\end{bmatrix}$$

over $k[X,Y]$ of rank 56 which contain the mutually nonisomorphic subalgebras

$$M_i^1=M(T_i^1,N_i^1,N_i^{1\vee})=\begin{bmatrix} k[X,Y] & P_i^1 \\ (P_i^1)^\vee & k[X,Y]\end{bmatrix}$$

and

$$M_i^2=M(T_i^2,N_i^2,N_i^{2\vee})=\begin{bmatrix} k[X,Y] & P_i^2 \\ (P_i^2)^\vee & k[X,Y]\end{bmatrix}$$

of rank 20, which are stable under the involution $^-$. They also contain the subalgebras

$$M(D_1)=M(T_{D_1},N_{D_1},N_{D_1})=\begin{bmatrix} k[X,Y] & D_1 \\ D_1 & k[X,Y]\end{bmatrix}$$

and

$$M(D_2)=M(T_{D_2},N_{D_2},N_{D_2})=\begin{bmatrix} k[X,Y] & D_2 \\ D_2 & k[X,Y]\end{bmatrix}$$

of rank 20, which are again stable under the involution $^-$ [Pumplün 2010b, Theorem 10].

Let J_i be the Jordan algebra we get if we patch J_i^1 on $k[X]_{g_i}[Y]$ and J_i^2 on $k[X]_{f_i}[Y]$ over $k[X]_{f_i g_i}[Y]$ using the isomorphisms $J(\pi_i^1)$ and $J(\pi_i^2)$ respectively, that are canonically induced by the π_i^j, $j=1,2$, as described in [Parimala et al. 1999, p. 32]. The algebras J_i are nonextended, mutually nonisomorphic and no longer a first Tits construction starting with some Azumaya algebra of degree 3 [Parimala et al. 1999, 6.3]. The matrix algebra

$$M(J_i)=\begin{bmatrix} k[X,Y] & J_i \\ J_i & k[X,Y]\end{bmatrix}$$

can then be also viewed as obtained from the matrix algebras

$$M(J_i^1)=\begin{bmatrix} k[X,Y] & J_i^1 \\ J_i^1 & k[X,Y]\end{bmatrix} \quad\text{and}\quad M(J_i^2)=\begin{bmatrix} k[X,Y] & J_i^2 \\ J_i^2 & k[X,Y]\end{bmatrix}$$

by patching them using the obvious induced isomorphisms. Call them $S(\pi_i^j)$, $j=1,2$.

By construction, M_i is then clearly a subalgebra of the matrix algebra $M(J_i)$ (the isomorphisms used to patch it are restrictions of the $S(\pi_i^j)$) and there are mutually coprime polynomials $r_i \in k[X]$ with $M(J_i)\otimes k[X]_{r_i}[Y] \cong M(J)\otimes k[X]_{r_i}[Y]$ and $M_i \otimes k[X]_{r_i}[Y] \cong M \otimes k[X]_{r_i}[Y]$, where $M \cong M(D_1^+) \cong M(D_2^+) \subset M(J)$.

Remark 5. We observe independently of this that be the infinitely many mutually nonisomorphic reduced Albert algebras A_i over $k[X,Y]$ constructed in [Parimala et al. 1997, Step I and 6.2], also give rise to matrix algebras

$$H_i = \begin{bmatrix} k[X,Y] & A_i \\ A_i & k[X,Y] \end{bmatrix}$$

over $k[X,Y]$ of rank 56 which are mutually nonisomorphic, which is proved analogously to [Parimala et al. 1997, 6.2].

3. Structurable algebras over $\mathbb{A}_k^2$ which are forms of matrix algebras

Remark 6. Let T be a quadratic étale algebra over $k[X,Y]$ with anisotropic norm. As in [Parimala et al. 1997, 4.6], one can see that T extends uniquely to a quadratic étale algebra $\mathcal{T} = \mathrm{Cay}(\mathcal{O}_X, \mathcal{L}, N)$ over $X = \mathbb{P}_k^2$. Since $\mathrm{Pic}\, X = \mathbb{Z}$, $\mathcal{L} \cong \mathcal{O}_X$ and $\mathcal{T}$ is defined over k, thus so is T. We conclude that every quadratic étale algebra over $k[X,Y]$ with anisotropic norm is of the kind $K \otimes_k k[X,Y] \cong K[X,Y]$ with $K = k(\sqrt{c})$ a separable quadratic field extension. As a consequence, every quadratic étale ring extension R' of $k[X,Y]$ satisfies $R' = k(\sqrt{c})[X,Y]$ and every form of a matrix algebra of the type $S(B,*,P,N,h)$, B a central simple algebra over R' has $S(A,^-) = (\sqrt{c},0)R$.

3.1. Let K be a separable quadratic field extension of k. Let D be a central division algebra over K of degree 3 with an involution σ of the second kind over K/k. Let $X = \mathbb{P}_k^2$, $X' = X \otimes_k K = \mathbb{P}_K^2$ and $\mathcal{D} = D \otimes_K \mathcal{O}_{X'}$.

Proposition 7. *Let P be a nonfree projective left $D[X,Y]$-module of rank one. The structurable algebra $S(D,\sigma,P,N) = K[X,Y] \oplus P$ over $k[X,Y]$ admits a unique extension to a structurable algebra $S(\mathcal{D},\sigma,\widetilde{P},N) = \mathcal{O}_{X'} \oplus \widetilde{P}$ over $X = \mathbb{P}_k^2$. The vector bundle $\widetilde{P}$ over X' is indecomposable.*

Proof. There is a unique extension of the quadratic étale algebra $K[X,Y]$ over $k[X,Y]$ to a quadratic étale algebra $\mathcal{O}_{X'} = K \otimes_k \mathcal{O}_X$ over X. There is a unique extension $\widetilde{P}$ over $X' = \mathbb{P}_K^2$ of P of norm one that is a locally free left $\mathcal{D}$-module: by [Parimala et al. 1999, p. 29], P extends to a vector bundle $\widetilde{P}$ over X' that is unique up to a line bundle $\mathcal{L} \in \mathrm{Pic}\, X'$. Since we require $\widetilde{P}$ to be of norm one this implies $\mathcal{L}^3 \cong \mathcal{O}_{X'}$, hence $\mathcal{L} = \mathcal{O}_{X'}$ and the extension is unique. More precisely, by [ibid., Remark] and [Arason et al. 1992], $\widetilde{P} \cong tr_{L'/K'}(\mathcal{P}_0)$ for some cubic field extension L'/K' and a suitable vector bundle $\mathcal{P}$ over $\mathbb{P}^2_{L'}$ that is absolutely indecomposable and must have rank 3. In particular, N and h can be extended as well.

The algebra $S(\mathfrak{D}, \sigma, \widetilde{P}, N) = \mathbb{O}_{X'} \oplus \widetilde{P}$ restricts to the structurable algebra

$$S(D, \sigma, P, N) = K[X, Y] \oplus P$$

over $\mathbb{A}^2_k$. The second statement follows from [Parimala et al. 1999, 3.2]. □

3.2. Let K be a separable quadratic field extension of k. Let D be a central division algebra over K of degree 3 with an involution σ of the second kind over K/k. Let (u, μ) be an admissible scalar; i.e., $\mu \in K^\times$, $c \in H(B, *_B)^\times$ and $N_B(c) = \mu\mu^*$. By [ibid., p. 33], there exists a projective left $D[X, Y]$-module P of rank 1 together with a nondegenerate hermitian form $h : P \times P \to D[X, Y]$ and a trivialization $\widetilde{\mu} : \mathrm{disc}(h) \to (K[X, Y], \langle 1\rangle)$ such that:

(8) The reduction of $(P, h, \widetilde{\mu})$ modulo Y is isomorphic to $(D, \langle u\rangle, \mu)$, where $\langle u\rangle$ denotes the hermitian form $a \to au\sigma(a)$ and μ is treated as a trivialization of the discriminant of $\langle u\rangle$. Moreover, $(De, u_e, \mu_e) \otimes k[X] = (P, h, \widetilde{\mu})$ modulo Y, where De is the free module of rank one over D with e a basis element, u_e the hermitian form on De given by $u_e(xe, ye) = xu\sigma(y)$ and $\mu_e N_D(e) = \mu$.

(9) There exists $f \in k[X]$, $f(0) \neq 0$, such that

$$(P, h, \widetilde{\mu}) \otimes k[X]_f[Y] \cong (D, \langle u\rangle, \mu) \otimes k[X]_f[Y].$$

(10) The principal $SU(D, \sigma)$-bundle on $\mathbb{A}^2_k$ associated to $(P, h, \widetilde{\mu})$ admits no reduction of the structure group to any proper connected reductive subgroup of $SU(D, \sigma)$. In particular, $(P, h, \widetilde{\mu})$ is not extended from $(D, \langle u\rangle, \mu)$.

Now let J be an Albert division algebra over k that is a second Tits construction but not a first one. We may write

$$J = J(D^1 e_1, u_{e_1}, \mu_{e_1}) = J(D^2 e_2, u_{e_2}, \mu_{e_2})$$

where D^1, D^2 are two isomorphic central simple algebras of degree 3 over a quadratic extension F/k with involution σ^1, σ^2 of the second kind and norms N_1 and N_2, such that $H(D^1, \sigma^1) \cap H(D^2, \sigma^2) = k$; see [ibid., 5.2].

Define the structurable algebra

$$A = S(D^1, \sigma^1, D^1, N_1, u_{e_1}) \cong S(D^2, \sigma^2, D^2, N_2, u_{e_2}).$$

By [Parimala et al. 1999, p. 35], there exist nontrivial hermitian spaces $(P^i_1, h^i_1, \widetilde{\mu}^i_1)$ over $(D^1[X, Y], \sigma^1)$ and $(P^i_2, h^i_2, \widetilde{\mu}^i_2)$ over $(D^2[X, Y], \sigma^2)$ of rank 1, and $f_i, g_i \in k[X]$ such that:

(11) $(P_1^i, h_1^i, \widetilde{\mu}_1^i)$ modulo Y reduces to $(D^1 e_1, u_{e_1}, \mu_{e_1})$, $(P_2^i, h_2^i, \widetilde{\mu}_2^i)$ modulo Y reduces to $(D^2 e_2, u_{e_2}, \mu_{e_2})$.

(12) $(P_1^i, h_1^i, \widetilde{\mu}_1^i) \otimes k[X]_{f_i}[Y]$ is isomorphic to $(D^1 e_1, u_{e_1}, \mu_{e_1}) \otimes k[X]_{f_i}[Y]$ and $(P_2^i, h_2^i, \widetilde{\mu}_2^i) \otimes k[X]_{g_i}[Y]$ is isomorphic to $(D^2 e_2, u_{e_2}, \mu_{e_2}) \otimes k[X]_{g_i}[Y]$ with $(f_i, f_j) = 1 = (g_i, g_j)$ for all $i \neq j$ and $(f_i, g_j) = 1$ for all i, j.

(13) The vector bundles (P_1^i, h_1^i) and (P_2^i, h_2^i) are not extended from D^1 and D^2, respectively.

Let $N_j^i : P_j^i \to D_j[X, Y]$ denote the norm on P_j^i determined by the choice of $\widetilde{\mu}_j^i$, $j = 1, 2$. We define two families of structurable algebras

$$A_1^i = S(D^1, \sigma^1, P_1^i, N_1^i, h_1^i) \quad \text{and} \quad A_2^i = S(D^2, \sigma^2, P_2^i, N_2^i, h_2^i)$$

over $k[X, Y]$ with underlying modules structures

$$A_1^i \cong K[X, Y] \oplus P_1^i \quad \text{and} \quad A_2^i \cong K[X, Y] \oplus P_2^i.$$

Let

$$\pi_1^i : (P_1^i, h_1^i, \widetilde{\mu}_1^i)_{f_i} \to (D^1 e_1, u_{e_1}, \mu_{e_1}) \otimes k[X]_{f_i}[Y],$$
$$\pi_2^i : (P_2^i, h_2^i, \widetilde{\mu}_2^i)_{g_i} \to (D^2 e_2, u_{e_2}, \mu_{e_2}) \otimes k[X]_{g_i}[Y]$$

be isometries such that $\overline{\pi}_j^i = \mathrm{id}$ for $j = 1, 2$. These isometries induce isomorphisms

$$A(\pi_1^i) : A_1^i \otimes k[X]_{f_i}[Y] \to A \otimes k[X]_{f_i}[Y],$$
$$A(\pi_1^i) : A_2^i \otimes k[X]_{g_i}[Y] \to A \otimes k[X]_{g_i}[Y],$$

which reduce to the identity map modulo Y.

Proposition 8. *The structurable algebras A_1^i and A_2^i over $k[X, Y]$ have the following properties*:

(i) *A_1^i and A_2^i modulo Y reduce to A.*

(ii) *$A_1^i \otimes k[X]_{f_i}[Y]$ is extended from $A \otimes k[X]_{f_i}[Y]$ and $A_2^i \otimes k[X]_{g_i}[Y]$ is extended from $A \otimes k[X]_{g_i}[Y]$, with $(f_i, f_j) = 1 = (g_i, g_j)$ for $i \neq j$, $(f_i, g_j) = 1$ for all i, j.*

(iii) *$A_j^i \otimes_k K \cong M(T_j^i, N_j^i, N_j^{i\,\vee}) = M_j^i$ for $j = 1, 2$, where the matrix algebras M_j^i are the ones constructed in Section 2.4.*

(iv) *All the A_1^i are mutually nonisomorphic and all the A_2^i are mutually nonisomorphic.*

Proof. Properties (i) and (ii) are immediate consequences of the properties of $(P^i_j, h^i_j, \widetilde{\mu}^i_j)$. Property (iii) follows from the construction of the algebras. We use the identification from the proof of [Pumplün 2010b, Theorem 20].

(iv) Since $A^i_1 \otimes_k K \cong M(T^i_j, N^i_j, N^i_j)$ is not extended from $M = M(T_{D^1}, N^1, N^1)$ and P^i_1 is not free, it follows that A^i_1 is not extended from A by [Bass et al. 1977]. Thus the algebras A^i_1 are mutually nonisomorphic. The same argument holds for the A^i_2. □

We now patch the structurable algebras $(A^i_1)_{g_i}$ over $k[X]_{g_i}[Y]$ and $(A^i_2)_{f_i}$ over $k[X]_{f_i}[Y]$ over $k[X]_{f_i g_i}[Y]$ and their involutions using the isomorphism

$$\psi_i : A^i_1 \otimes k[X]_{f_i g_i}[Y] \to A^i_2 \otimes k[X]_{f_i g_i}[Y], \quad \psi_i = A(\pi^i_2)^{-1} A(\pi^i_1).$$

This way we obtain a structurable algebra A^i over $k[X, Y]$.

Theorem 9. *The structurable algebras A^i over $\mathbb{A}^2_k$ have the following properties*:

(1) $\overline{A}^i = A \otimes k[X]$ *modulo* Y.

(2) *There exists* $\pi^i : A^i \otimes k[X]_{s_i}[Y] \to A \otimes k[X]_{s_i}[Y]$ *such that* $\overline{\pi}^i = \mathrm{id}$, *for some* $s_i \in k[X]$ *with* $(s_i, s_j) = 1$ *for* $i \neq j$.

(3) *The* A^i *are mutually nonisomorphic.*

(4) $A^i \otimes_k K \cong M_i$ *with the* M_i *as constructed in Section 2.5.*

Proof. Since A^i_j reduces modulo Y to A and $\overline{\psi}_i = \mathrm{id}$, A^i reduces modulo Y to A. By construction,

$$A^i \otimes k[X]_{f_i g_i}[Y] \cong A \otimes k[X]_{f_i g_i}[Y]$$

and the polynomials $s_i := f_i g_i$ satisfy $(s_i, s_j) = 1$ for $i \neq j$. As in the proof of Theorem 4, it follows that the A^i are mutually nonisomorphic. □

Again, the ingredients for the results were provided in [Parimala et al. 1999, Section 5].

4. On extending structurable algebras from the affine to the projective plane

We conclude with some general results about extending structurable algebras from the affine to the projective plane, imitating the techniques used in [Parimala et al. 1997, 4.1, 4.2, 4.3]. Let R be a domain with $\frac{1}{6} \in R$.

4.1. For a structurable algebra $(A, ^-)$, an *isotopy* from $(A, ^-)$ to $(A, ^-)$ is an element $\alpha \in \mathrm{GL}(A)$ such that

$$\alpha\{x, y, z\} = \{\alpha(x), \widehat{\alpha}(y), \alpha(z)\}$$

for all $x, y, z \in A$ and some $\widehat{\alpha} \in \mathrm{GL}(A)$. $\widehat{\alpha}$ is uniquely determined by α. The *structure group* $\Gamma(A, ^-)$ of $(A, ^-)$ is the subgroup of $\mathrm{GL}(A)$ which consists of all isotopies of $(A, ^-)$ onto itself.

Let $(A, ^-)$ be a structurable algebra of skew-rank one such that $S(A, ^-) = s_0 R$ for some $s_0 \in S(A, ^-)$ that is conjugate invertible, which means that left multiplication L_{s_0} with s_0 is invertible. Since $\widehat{s}_0 \in S(A, ^-)$ for its conjugate inverse $\widehat{s}_0$, there is $\beta \in R$, $\beta \neq 0$, such that $\widehat{s}_0 = \beta s_0$ and since $s_0 \widehat{s}_0 = -1_A$ we obtain $\beta s_0^2 = -1_A$. Assume that $\beta \in R^\times$ and denote $c = \beta^- 1$. Then $s_0^2 = c 1_A$ with $c \in R^\times$. Suppose in addition that the invertible elements in $(A, ^-)$ are Zariski dense in A. Then we can define a (*conjugate*) norm $\nu : A \to R$ on A via

$$\nu(x) = \frac{1}{12c} \chi(s_0 x, \{x, s_0 x, x\}),$$

a *trace* $\chi : A \times A \to R$ on A by

$$\chi(x, y) = \frac{2}{c} \psi(s_0 x, y) s_0 = \frac{2}{c} (V^\delta_{y,x} s_0) s_0,$$

and a nondegenerate skew-symmetric bilinear form on A

$$\langle x, y \rangle = \psi(x, y) s_0 = \tfrac{1}{2} \chi(s_0 x, y)$$

analogously as in [Allison and Faulkner 1984; 1992], where $\psi(x, y) = x\bar{y} - y\bar{x}$ [Allison and Faulkner 1992, 5.4]. (The nondegeneracy of $\langle\ ,\ \rangle$ follows from [Allison and Faulkner 1984, p. 192], applied to the residue class forms.) ν is a quartic form such that $\nu(1_A) = 1$. χ is a nondegenerate symmetric bilinear form independent of the choice of s_0 and $\chi(1_A, 1_A) = 4$. (Nondegeneracy follows from [ibid., Proposition 2.5], applied to the residue class forms.) Note that if desired, A can be viewed as a Freudenthal triple system as explained in [ibid., 2.18], in this setting. An element $x \in A$ is conjugate invertible if and only if $\nu(x) \neq 0$ [ibid., 4.4]. So if the norm is anisotropic, every nonzero element of A is conjugate invertible and, if R is a field, $(A, ^-)$ a conjugate division algebra [Allison and Faulkner 1984, 2.11]. The norm ν is a semi-invariant for the structure group $\Gamma(A, ^-)$, which is proved analogously as in [Allison and Faulkner 1992, 4.7]. Denote the group of all invertible linear transformations on A that preserve the norm and the skew-symmetric bilinear form $\langle\ ,\ \rangle$ by $\mathrm{Inv}(A)$.

Theorem 10. *Let $(A_1, ^-)$ and $(A_2, ^-)$ be structurable algebras of skew-rank one over R. Suppose that $(A_2, ^-)$ satisfies all of the criteria in Section 4.1 (i.e., it carries a conjugate norm), and that the conjugate norm of $(A_2 \otimes R/(p), ^-)$ is anisotropic. Let*

$$\alpha : (A_1 \otimes R[1/p], ^-) \to (A_2 \otimes R[1/p], ^-)$$

be an isotopy of structurable algebras. Then α extends uniquely to an isotopy

$$\tilde{\alpha} : (A_1, {}^-) \to (A_2, {}^-).$$

In particular, every isomorphism $\alpha : (A_1 \otimes R[1/p], {}^-) \to (A_2 \otimes R[1/p], {}^-)$ of the structurable algebras $(A_1, {}^-)$ and $(A_2, {}^-)$ extends uniquely to an isomorphism $\tilde{\alpha} : (A_1, {}^-) \to (A_2, {}^-)$.

Proof. We show that $\alpha(A_1) = A_2$, which is sufficient: let $x \in A_1$ and assume that $\alpha(x) \notin A_2$. Let n be the least integer such that $y = p^n \alpha(x) \in A_2$ and $p^{n-1}\alpha(x) \notin A_2$. Then $n \geq 1$. ν is a semi-invariant for the structure group of $(A, {}^-)$, i.e., there is $0 \neq r \in R$ such that $\nu(\alpha(x)) = r\nu(x)$ for all $x \in A_1$. Thus we obtain $\nu(y) = rp^{4n}\nu(x)$. Hence $\nu(y) = 0$ modulo p and $y \neq 0$ modulo p. This contradicts the assumption that the norm $\nu \otimes R/(p)$ of $(A_2 \otimes R/(p), {}^-)$ is anisotropic. □

4.2. There is an obvious notion of a structurable algebra over a locally ringed space [Pumplün 2010b, Section 6]. Let $(A, {}^-)$ be a structurable algebra of skew-rank one over $X = \mathbb{P}^n_k$ such that $S(A, {}^-) = s_0 \mathcal{O}_X$ for some $s_0 \in H^0(X, S(A, {}^-)) = k$ which is conjugate invertible, which means that left multiplication L_{s_0} with s_0 is invertible. Since $\widehat{s}_0 \in H^0(X, S(A, {}^-))$ for its conjugate inverse $\widehat{s}_0$, there is $c \in k^\times$, such that $\widehat{s}_0 = -c^{-1}s_0$ and since $s_0\widehat{s}_0 = -1_A$ we obtain $s_0^2 = c1_A$. Suppose in addition that the invertible elements in $H^0(U, (A, {}^-))$ are Zariski dense in $H^0(U, A)$ for every open subset $U \subset X$. Then we can define a (*conjugate*) norm $\nu : A \to \mathcal{O}_X$ via

$$\nu(x) = \frac{1}{12c}\chi(s_0 x, \{x, s_0 x, x\}),$$

a *trace* $\chi : A \times A \to \mathcal{O}_X$ on A by

$$\chi(x, y) = \frac{2}{c}\psi(s_0 x, y)s_0 = \frac{2}{c}(V^{\delta}_{y,x}s_0)s_0$$

and a nondegenerate skew-symmetric bilinear form $\nu : A \times A \to \mathcal{O}_X$

$$\langle x, y\rangle = \psi(x, y)s_0 = \frac{1}{2}\chi(s_0 x, y)$$

analogously as in 4.1, $\psi(x, y) = x\bar{y} - y\bar{x}$. ν is a quartic form such that $\nu(1_A) = 1$. χ is a nondegenerate symmetric bilinear form independent of the choice of s_0 and $\chi(1_A, 1_A) = 4$.

Theorem 10 now implies:

Corollary 11. *Let $(\mathcal{A}_1, {}^-{}_1)$, $(\mathcal{A}_2, {}^-{}_2)$ be two structurable algebras of skew-rank one over $\mathbb{P}^n_k$ which satisfy the assumptions of* 4.2. *Suppose that the restrictions $(\mathcal{A}_1)_\xi$ and $(\mathcal{A}_2)_\xi$ to the generic point ξ have anisotropic norms. Then every isotopy $\alpha : \mathcal{A}_1 \to \mathcal{A}_2$ over $\mathbb{A}^n_k$ extends uniquely to an isotopy $\widetilde{\alpha} : \mathcal{A}_1 \to \mathcal{A}_2$ over $\mathbb{P}^n_k$.*

In particular, every isomorphism $\alpha : \mathcal{A}_1 \to \mathcal{A}_2$ *over* $\mathbb{A}^n_k$ *extends uniquely to an isomorphism* $\tilde{\alpha} : \mathcal{A}_1 \to \mathcal{A}_2$ *over* $\mathbb{P}^n_k$.

The proof is verbatim to the proof of [Parimala et al. 1997, 4.3], substituting "isotopy" (and "isomorphism") for "isometry" throughout.

From Corollary 11 and [Parimala et al. 1997, 4.5], we obtain:

Corollary 12. *Let* k *have characteristic* 0. *Let* $(\mathcal{A}, ^-)$ *be a structurable algebra of skew-rank one over* $\mathbb{A}^2_k$ *satisfying the conditions of Section 4.1, such that its restriction* $\mathcal{A}_\xi$ *to the generic point* ξ *has an anisotropic norm. Then* $(\mathcal{A}, ^-)$ *extends uniquely to an algebra* $(\mathcal{A}, ^-)$ *over* $\mathbb{P}^2_k$.

If $H = \operatorname{Inv}(A)$ *is a connected reductive algebraic group defined over* k *then every* H*-bundle over* $\mathbb{A}^2_k$ *extends to* $\mathbb{P}^2_k$ *as an* H*-bundle.*

If the structurable algebra bundle has rank 56 *and admits a reduction of the structure group to a proper connected reductive subgroup of* E_7, *its corresponding extension to* $\mathbb{P}^2_k$ *has the same property.*

References

[Achhammer 1995] G. Achhammer, *Albert Algebren über lokal geringten Räumen*, Ph.D. thesis, FernUniversität in Hagen, 1995, Available at http://homepage.uibk.ac.at/~c70202/jordan/archive/achhammer/achhammer.pdf. Zbl 0864.17044

[Allison 1978] B. N. Allison, "A class of nonassociative algebras with involution containing the class of Jordan algebras", *Math. Ann.* **237**:2 (1978), 133–156. MR 81h:17003 Zbl 0368.17001

[Allison 1979] B. N. Allison, "Models of isotropic simple Lie algebras", *Comm. Algebra* **7**:17 (1979), 1835–1875. MR 81d:17005 Zbl 0422.17006

[Allison and Faulkner 1984] B. N. Allison and J. R. Faulkner, "A Cayley–Dickson process for a class of structurable algebras", *Trans. Amer. Math. Soc.* **283**:1 (1984), 185–210. MR 85i:17001 Zbl 0543.17015

[Allison and Faulkner 1992] B. N. Allison and J. R. Faulkner, "Norms on structurable algebras", *Comm. Algebra* **20**:1 (1992), 155–188. MR 93b:17005 Zbl 0753.17004

[Arason et al. 1992] J. K. Arason, R. Elman, and B. Jacob, "On indecomposable vector bundles", *Comm. Algebra* **20**:5 (1992), 1323–1351. MR 93e:14051 Zbl 0769.14004

[Bass et al. 1977] H. Bass, E. H. Connell, and D. L. Wright, "Locally polynomial algebras are symmetric algebras", *Invent. Math.* **38**:3 (1977), 279–299. MR 55 #5613 Zbl 0371.13007

[Jacobson 1968] N. Jacobson, *Structure and representations of Jordan algebras*, Colloquium Publications **39**, American Mathematical Society, Providence, RI, 1968. MR 40 #4330 Zbl 0218.17010

[Knus et al. 1994] M.-A. Knus, R. Parimala, and R. Sridharan, "On compositions and triality", *J. Reine Angew. Math.* **457** (1994), 45–70. MR 96a:11037 Zbl 0808.11028

[Krutelevich 2007] S. Krutelevich, "Jordan algebras, exceptional groups, and Bhargava composition", *J. Algebra* **314**:2 (2007), 924–977. MR 2008k:20103 Zbl 1163.17032

[McCrimmon 2004] K. McCrimmon, *A taste of Jordan algebras*, Springer, New York, 2004. MR 2004i:17001 Zbl 1044.17001

[Parimala et al. 1997] R. Parimala, V. Suresh, and M. L. Thakur, "Jordan algebras and F_4 bundles over the affine plane", *J. Algebra* **198**:2 (1997), 582–607. MR 98k:17035

[Parimala et al. 1999] R. Parimala, R. Sridharan, and M. L. Thakur, "Tits' constructions of Jordan algebras and F_4 bundles on the plane", *Compositio Math.* **119**:1 (1999), 13–40. MR 2000m:17035

[Pumplün 2008] S. Pumplün, "Albert algebras over curves of genus zero and one", *J. Algebra* **320**:12 (2008), 4178–4214. MR 2009k:17053 Zbl 1209.17025

[Pumplün 2010a] S. Pumplün, "Jordan algebras over algebraic varieties", *Comm. Algebra* **38**:2 (2010), 714–751. MR 2011d:17052 Zbl 05686617

[Pumplün 2010b] S. Pumplün, "Classes of structurable algebras of skew-rank 1", *Israel J. Math.* **180** (2010), 425–460. MR 2011j:17002 Zbl 05834351

[Raghunathan 1989] M. S. Raghunathan, "Principal bundles on affine space and bundles on the projective line", *Math. Ann.* **285**:2 (1989), 309–332. MR 91b:11045 Zbl 0661.14016

[Schafer 1966] R. D. Schafer, *An introduction to nonassociative algebras*, Pure and Applied Mathematics **22**, Academic Press, New York, 1966. Reprinted Dover, New York, 1995. MR 35 #1643 Zbl 0145.25601

Received December 15, 2010. Revised December 5, 2011.

SUSANNE PUMPLÜN
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF NOTTINGHAM
UNIVERSITY PARK
NOTTINGHAM NG7 2RD
UNITED KINGDOM

susanne.pumpluen@nottingham.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

AN ANALOGUE OF KREIN'S THEOREM FOR SEMISIMPLE LIE GROUPS

SANJOY PUSTI

We give an integral representation of K-positive definite functions on a real rank n connected, noncompact, semisimple Lie group with finite centre. Moreover, we characterize the λ's for which the τ-spherical function $\phi^{\tau}_{\sigma,\lambda}$ is positive definite for the group $G = \mathrm{Spin}_e(n, 1)$ and the complex spin representation τ.

1. Introduction

A continuous function f on $\mathbb{R}$ is said to be *positive definite* if for any real numbers $x_1, \ldots, x_m$ and complex numbers $\xi_1, \ldots, \xi_m$ the following holds:

$$\sum_{k,j=1}^{m} f(x_j - x_k)\xi_k\overline{\xi_j} \geq 0.$$

This definition is equivalent to

$$\int_{\mathbb{R}} f(x)(\phi * \phi^*)(-x)\,dx \geq 0 \quad \text{for all } \phi \in C_c^\infty(\mathbb{R}),$$

where $\phi^*(x) = \overline{\phi(-x)}$. Also, an even continuous function f on $\mathbb{R}$ is said to be *evenly positive definite* if

$$\int_{\mathbb{R}} f(x)(\phi * \phi^*)(-x)\,dx = \int_{\mathbb{R}} f(x)(\phi * \phi^*)(x)\,dx \geq 0 \quad \text{for all } \phi \in C_c^\infty(\mathbb{R})_e,$$

where $C_c^\infty(\mathbb{R})_e$ denotes the set of infinitely differentiable compactly supported even functions on $\mathbb{R}$. Then it is clear that the set of even positive definite functions is a subset of the set of evenly positive definite functions. Bochner's theorem and M. G. Krein's theorem respectively give integral representations of positive definite functions and evenly positive definite functions. Precisely, for a positive definite function f on $\mathbb{R}$, there exists a finite positive measure μ on $\mathbb{R}$ such that

$$f(x) = \int_{\mathbb{R}} e^{i\lambda x}\,d\mu(\lambda).$$

MSC2010: primary 43A85; secondary 22E30.
Keywords: positive definite functions, K-positive definite functions, τ-positive definite functions.

Also for an evenly positive definite function f on $\mathbb{R}$, there exists a finite positive even measure σ on $\mathbb{R} \cup i\mathbb{R}$ such that

$$f(x) = \int_{\mathbb{R}\cup i\mathbb{R}} e^{i\lambda x}\, d\sigma(\lambda).$$

From this integral representation it follows that a bounded evenly positive definite function is a positive definite function. We note that the measure σ in the integral representation of an evenly positive definite function is not unique, whereas the measure μ in the integral representation of a positive definite function is unique. However, if an evenly positive definite function satisfies a certain restriction on its growth for $|x| \to \infty$, then the integral representation becomes unique [Gelfand and Vilenkin 1964].

Let G be a connected, noncompact semisimple Lie group with finite centre, and let K be a fixed maximal compact subgroup of G. Integral representations of K-positive definite distributions and K-positive definite functions have been derived for real rank-one semisimple Lie groups with finite centre in [Sitaram 1978] and [Pusti 2011], respectively.

An analogue of Krein's theorem on $\mathbb{R}^n$ has been obtained by N. Bopp [1979]. In this case, instead of evenly positive definite functions, one considers functions which are positive definite relative to the action of a finite subgroup of $\mathrm{O}(n)$. Here too, if we impose a certain growth condition, then the integral representation of these functions is unique. In this paper, using Bopp's result, we derive an integral representation for the K-positive definite functions on a real rank n connected, noncompact, semisimple Lie group with finite centre. We observe that the set of positive definite functions is a proper subset of the set of K-positive definite functions. Next, we consider the τ-positive definite functions, $\tau \in \widehat{K}$. The K-positive definite functions are a special instance of the τ-positive definite functions (for τ equals the trivial representation). We give an example in which the set of τ-positive definite functions is same as the set of positive definite functions. That is, the same conclusion (as in K-positive definite function) is not true for τ-positive definite functions. Finally we characterize the λ's for which the τ-spherical function $\phi^{\tau}_{\sigma,\lambda}$ is a positive definite function for the group $G = \mathrm{Spin}_e(n, 1)$ and the complex spin representation τ. We note that G. van Dijk and A. Pasquale [1999] studied positive definiteness of $\phi^{\tau}_{\sigma,\lambda}$ for the group $G = \mathrm{Sp}(1, n)$.

2. Preliminaries

Most of our notations are standard and can be found in [Anker 1991]. Let G be a real rank n connected, noncompact, semisimple Lie group with finite centre with Lie algebra $\mathfrak{g}$, and let $\mathcal{U}(\mathfrak{g})$ be the universal enveloping algebra of $\mathfrak{g}$. Let $\mathfrak{g} = \mathfrak{k} + \mathfrak{s}$ be a Cartan decomposition of $\mathfrak{g}$, and let K be the maximal compact subgroup of G

with Lie algebra $\mathfrak{k}$. We fix a maximal abelian subspace $\mathfrak{a}$ of $\mathfrak{s}$. Since G is of real rank n, we have $\dim \mathfrak{a} = n$. We denote the real dual of $\mathfrak{a}$ by $\mathfrak{a}^*$ and its complex dual by $\mathfrak{a}^*_{\mathbb{C}}$. The Killing form of $\mathfrak{g}$ induces an Ad K-invariant scalar product on $\mathfrak{s}$ and hence a G-invariant Riemannian metric on G/K (or $K\backslash G$). With this structure, G/K is a Riemannian globally symmetric space of the noncompact type. Also, the Killing form of $\mathfrak{g}$ induces a scalar product on $\mathfrak{a}$ and hence on $\mathfrak{a}^*$. We denote by $\langle \cdot, \cdot \rangle$ its $\mathbb{C}$-bilinear extension to $\mathfrak{a}^*_{\mathbb{C}}$.

Let $\mathfrak{g} = \mathfrak{g}_0 \oplus (\oplus_{\alpha \in \Sigma} \mathfrak{g}_\alpha)$ be the root space decomposition of $\mathfrak{g}$. Here, $\mathfrak{g}_0 = \mathfrak{m} \oplus \mathfrak{a}$, where $\mathfrak{m}$ is the centralizer of $\mathfrak{a}$ in $\mathfrak{k}$ and $\Sigma \subseteq \mathfrak{a}^*$ is the root system of $(\mathfrak{g}, \mathfrak{a})$. Let W be the Weyl group associated to Σ. We choose a set Σ^+ of positive roots. Let $\mathfrak{a}^+ \subseteq \mathfrak{a}$ be the corresponding positive Weyl chamber and let $\overline{\mathfrak{a}^+}$ be its closure. We denote by $(\mathfrak{a}^*)^+$ and $\overline{(\mathfrak{a}^*)^+}$ the similar cones in $\mathfrak{a}^*$. Let $\mathfrak{n} = \bigoplus_{\alpha \in \Sigma^+} \mathfrak{g}_\alpha$. Then $\mathfrak{n}$ is a nilpotent subalgebra of $\mathfrak{g}$. The element $\rho \in \mathfrak{a}^*$ is defined by

$$\rho(H) = \tfrac{1}{2} \Sigma_{\alpha \in \Sigma^+} m_\alpha \alpha(H),$$

where $m_\alpha = \dim \mathfrak{g}_\alpha$. Let A be the analytic subgroup of G with Lie algebra $\mathfrak{a}$. Then A is a closed subgroup of G and the exponential map is an isomorphism from $\mathfrak{a}$ onto A. We set $A^+ = \exp \mathfrak{a}^+$. Its closure is $\overline{A^+} = \exp \overline{\mathfrak{a}^+}$. Let N be the analytic subgroup of G with Lie algebra $\mathfrak{n}$, and let M be the centralizer of A in K.

The group G can be decomposed as $G = K\overline{A^+}K$. It is called the Cartan decomposition of G and every element x of G can be decomposed as $x = k_1 a k_2$ with $k_1, k_2 \in K$ and $a \in \overline{A^+}$. We let x^+ be the $\overline{\mathfrak{a}^+}$-component of $x \in G$ in the decomposition $G = K(\exp \overline{\mathfrak{a}^+})K$ and let $|x| = \|x^+\|$. Viewed on G/K, $|\cdot|$ is the distance to the origin $0 = \{K\}$. Also, the group G has Iwasawa decomposition $G = KAN$. Let $k(x)$ and $H(x)$ be the components of $x \in G$ in K and $\mathfrak{a}$. Then any element $x \in G$ can be expressed as $x = k(x) \exp H(x) n$ for some $n \in N$.

For $\lambda \in \mathfrak{a}^*_{\mathbb{C}}$, the elementary spherical function ϕ_λ on G is given by

$$\phi_\lambda(x) = \int_K e^{-(i\lambda + \rho) H(x^{-1}k)}\, dk.$$

It satisfies the following properties:

(1) It is K-biinvariant, that is, $\phi_\lambda(k_1 x k_2) = \phi_\lambda(x)$ for all $k_1, k_2 \in K$ and $x \in G$. Also, it is W-invariant in $\lambda \in \mathfrak{a}^*_{\mathbb{C}}$, that is, $\phi_{w \cdot \lambda}(x) = \phi_\lambda(x)$ for all $w \in W$ and $x \in G$.

(2) The function $\phi_\lambda(x)$ is C^∞ in x and holomorphic in λ.

(3) It is a joint eigenfunction for all G-invariant differential operators on G/K; in particular for the Laplacian Δ on G/K,

$$\Delta \phi_\lambda = -(\langle \lambda, \lambda \rangle + \|\rho\|^2) \phi_\lambda.$$

A function f on G is called K-biinvariant if $f(k_1xk_2)=f(x)$ for all $k_1,k_2\in K$ and $x\in G$. For a K-biinvariant function f on G, its spherical Fourier transform is defined by

$$\hat{f}(\lambda)=\int_G f(x)\phi_\lambda(x^{-1})\,dx$$

for suitable $\lambda\in\mathfrak{a}_{\mathbb{C}}^*$.

The set of infinitely differentiable compactly supported K-biinvariant functions and infinitely differentiable K-biinvariant functions are denoted by $C_c^\infty(G/\!/K)$ and $C^\infty(G/\!/K)$, respectively. For $0<p\le 2$ the L^p-Schwartz space $\mathscr{C}^p(G/\!/K)$ is the set of all functions $f\in C^\infty(G/\!/K)$ such that

$$\sup_{x\in G}(1+|x|)^s\phi_0(x)^{-2/p}\,|f(D;x;E)|<\infty$$

for any $D,E\in\mathscr{U}(\mathfrak{g})$ and any integer $s\ge 0$. The Schwartz space $\mathscr{C}^p(G/\!/K)$ is topologized by the seminorms

$$\sigma^p_{D,E,s}(f)=\sup_{x\in G}(1+|x|)^s\phi_0(x)^{-2/p}\,|f(D;x;E)|\,.$$

Then it follows that $C_c^\infty(G/\!/K)$ is dense in $\mathscr{C}^p(G/\!/K)$ and $\mathscr{C}^p(G/\!/K)$ is dense in $L^p(G/\!/K)$.

Let $\mathscr{P}(\mathfrak{a}_{\mathbb{C}}^*)$ be the space of entire functions on $\mathfrak{a}_{\mathbb{C}}^*$, which are of exponential type and rapidly decreasing. The set of W-invariant elements in $\mathscr{P}(\mathfrak{a}_{\mathbb{C}}^*)$ is denoted by $\mathscr{P}(\mathfrak{a}_{\mathbb{C}}^*)^W$.

For fixed $\epsilon>0$, let $C^{\epsilon\rho}$ be the convex hull of the set $W\cdot\epsilon\rho$ in $\mathfrak{a}^*$, and let $\mathfrak{a}_\epsilon^*=\mathfrak{a}^*+iC^{\epsilon\rho}$ be the tube in $\mathfrak{a}_{\mathbb{C}}^*$ with base $C^{\epsilon\rho}$. For $\epsilon=0$, $\mathfrak{a}_\epsilon^*$ reduces to $\mathfrak{a}^*$. Let $\mathbf{S}(\mathfrak{a}^*)$ be the symmetric algebra over $\mathfrak{a}^*$. We define the Schwartz space $\mathscr{S}(\mathfrak{a}_\epsilon^*)$ as the space of all complex valued functions h such that the following hold true.

(1) h is holomorphic in the interior of $\mathfrak{a}_\epsilon^*$.

(2) h and all its derivatives extend continuously to $\mathfrak{a}_\epsilon^*$.

(3) for any polynomial $P\in\mathbf{S}(\mathfrak{a}^*)$ and any (integer) $t\ge 0$,

$$\sup_{\lambda\in\mathfrak{a}_\epsilon^*}(1+\|\lambda\|)^t\left|P\Big(\frac{\partial}{\partial\lambda}\Big)h(\lambda)\right|<\infty.$$

The space $\mathscr{S}(\mathfrak{a}_\epsilon^*)$ is topologized by the seminorms

$$\tau^\epsilon_{P,t}(h)=\sup_{\lambda\in\mathfrak{a}_\epsilon^*}(1+\|\lambda\|)^t\left|P\Big(\frac{\partial}{\partial\lambda}\Big)h(\lambda)\right|.$$

We denote by $\mathscr{S}(\mathfrak{a}_\epsilon^*)^W$ the subspace of W-invariant functions in $\mathscr{S}(\mathfrak{a}_\epsilon^*)$. For $\epsilon=0$, $\mathscr{S}(\mathfrak{a}_\epsilon^*)$ becomes the classical Schwartz space on $\mathfrak{a}^*$. Then for $\epsilon\ge 0$, $\mathscr{S}(\mathfrak{a}_\epsilon^*)^W$ is a Fréchet algebra (under pointwise multiplication) and $\mathscr{P}(\mathfrak{a}_{\mathbb{C}}^*)^W$ is a dense subalgebra of $\mathscr{S}(\mathfrak{a}_\epsilon^*)^W$.

We consider the function

$$\cosh_{\epsilon\rho}(H) = \frac{1}{|W|}\sum_{w\in W} e^{w\cdot\epsilon\rho(H)} \quad \text{on } \mathfrak{a}.$$

Then we define the space $\mathcal{S}_{\epsilon\rho}(\mathfrak{a})$ consisting of all functions $g \in C^\infty(\mathfrak{a})$ such that

$$\sup_{H\in\mathfrak{a}}(1+\|H\|)^s \cosh_{\epsilon\rho}(H)\left|P\Big(\frac{\partial}{\partial H}\Big)g(H)\right| < \infty$$

for any polynomial $P \in \mathbf{S}(\mathfrak{a})$ (the symmetric algebra over $\mathfrak{a}$) and any $s \geq 0$.

Theorem 2.1 [Anker 1991]. (1) *The spherical Fourier transform $f \mapsto \hat{f}$ is a topological isomorphism between $C_c^\infty(G/\!/K)$ and $\mathcal{P}(\mathfrak{a}_{\mathbb{C}}^*)^W$ and also between $\mathcal{C}^p(G/\!/K)$ and $\mathcal{S}(\mathfrak{a}_\epsilon^*)^W$, where $\epsilon = 2/p - 1$.*

(2) *The Euclidean Fourier transform $f \mapsto \tilde{f}$ is a topological isomorphism between $\mathcal{S}_{\epsilon\rho}(\mathfrak{a})^W$ and $\mathcal{S}(\mathfrak{a}_\epsilon^*)^W$, where $\tilde{f}(\lambda) = \int_{\mathfrak{a}} f(H)e^{-i\lambda(H)}dH$, $\lambda \in \mathfrak{a}^*$.*

For a suitable K-biinvariant function f on G, the Abel transform is defined by

$$\mathcal{A}f(H) = e^{\rho(H)}\int_N f(\exp Hn)\,dn.$$

It satisfies the relation $\hat{f}(\lambda) = \widetilde{\mathcal{A}f}(\lambda)$ for a suitable K-biinvariant function f on G. Therefore it follows from Theorem 2.1 that the Abel transform $f \mapsto \mathcal{A}f$ is a topological isomorphism between $\mathcal{C}^p(G/\!/K)$ and $\mathcal{S}_{\epsilon\rho}(\mathfrak{a})^W$ for $\epsilon = 2/p - 1$.

3. M. G. Krein's theorem

For $\alpha \geq 0$, we define

$$\mathcal{S}_\alpha(\mathbb{R}^n) = \{\phi \in C^\infty(\mathbb{R}^n) : \|\phi\|_p < \infty \text{for any nonnegative integer } p\},$$

where

$$\|\phi\|_p = \max_{|q|\leq p}\sup_{x\in\mathbb{R}^n}(1+|x|^2)^p e^{\alpha|x|}\left|D^q\phi(x)\right|.$$

Then $S_\alpha(\mathbb{R}^n)$ becomes a Fréchet space and $C_c^\infty(\mathbb{R}^n)$ is a dense subspace of $S_\alpha(\mathbb{R}^n)$. For a finite subgroup E of $\mathrm{O}(n)$, let $\mathcal{S}_\alpha(\mathbb{R}^n)^E$ be the subspace of E-invariant functions in $\mathcal{S}_\alpha(\mathbb{R}^n)$.

Theorem 3.1 [Bopp 1979]. *Let E be a finite subgroup of* $\mathrm{O}(n)$ *and let*

$$T : \mathcal{S}_\alpha(\mathbb{R}^n) \to \mathbb{C}$$

be a continuous, linear functional such that

(1) $T(\eta\cdot\phi) = T(\phi)$ *for all* $\eta \in E,\ \phi \in \mathcal{S}_\alpha(\mathbb{R}^n)$.

(2) $T(\phi * \phi^*) \geq 0$ *for all* $\phi \in \mathcal{S}_\alpha(\mathbb{R}^n)^E$.

Then there exists a unique positive tempered measure σ, invariant under the E-action, such that for all $\phi \in \mathcal{S}_\alpha(\mathbb{R}^n)$,

$$T(\phi) = \int_{M \cap T_\alpha} \widetilde{\phi}(\xi)\, d\sigma(\xi),$$

where $M = \{\xi \in \mathbb{C}^n : \text{there exists } \eta \in E \text{ such that } \eta.\xi = \overline{\xi}\}$ *and*

$$T_\alpha = \{\xi \in \mathbb{C}^n : |\operatorname{Im} \xi| \leq \alpha\}.$$

Since we have an isomorphism between $\mathcal{S}_{\epsilon\rho}(\mathbb{R}^n)$ and $\mathcal{S}_{\epsilon\rho}(\mathfrak{a})$ we can rewrite the theorem above in the following way:

Theorem 3.2. *Let $T : \mathcal{S}_{\epsilon\rho}(\mathfrak{a})^W \to \mathbb{C}$ be a continuous, linear functional such that*

$$T(\phi * \phi^*) \geq 0 \quad \text{for all } \phi \in \mathcal{S}_{\epsilon\rho}(\mathfrak{a})^W.$$

Then there exists a unique positive tempered measure σ, invariant under the W-action, such that for all $\phi \in \mathcal{S}_{\epsilon\rho}(\mathfrak{a})$,

$$T(\phi) = \int_{\mathcal{M} \cap \mathfrak{a}_\epsilon^*} \widetilde{\phi}(\lambda)\, d\sigma(\lambda),$$

where $\mathcal{M} = \{\lambda \in \mathfrak{a}_{\mathbb{C}}^* : \text{ there exists } w \in W \text{ such that } w.\lambda = \overline{\lambda}\}$.

We call a K-biinvariant continuous function f on G K-positive definite if for all $g \in C_c^\infty(G/\!/K)$,

$$\int_G f(x)(g * g^*)(x^{-1})\, dx \geq 0,$$

where $g^*(x) = \overline{g(x^{-1})}$ for all $x \in G$. If the equation above is true for every $g \in C_c^\infty(G)$ we say that f is a positive definite function. We prove the following analogue of M. G. Krein's theorem for K-positive definite functions on semisimple Lie groups.

Theorem 3.3. *For a K-positive definite function $f \in \mathscr{C}^p(G/\!/K)'$ $(0 < p \leq 2)$, there exists a unique finite positive measure σ on $\mathcal{M} \cap \mathfrak{a}_\epsilon^*$, invariant under the Weyl group action, such that for all $x \in G$*

$$f(x) = \int_{\mathcal{M} \cap \mathfrak{a}_\epsilon^*} \phi_\lambda(x)\, d\sigma(\lambda),$$

where $\epsilon = 2/p - 1$.

Proof. We define a linear functional $T_f : \mathcal{S}_{\epsilon\rho}(\mathfrak{a})^W \to \mathbb{C}$ by

$$T_f(h) = \int_G f(x)(\mathcal{A}^{-1}h)(x^{-1})\, dx.$$

The integral exists and is continuous by the given condition on f and the isomorphism of the Abel transform on $\mathscr{C}^p(G/\!/K)$. Since $\hat{f}(\lambda) = \widetilde{\mathscr{A}f}(\lambda)$ for all $f \in \mathscr{C}^p(G/\!/K)$, it follows that

$$\widehat{\mathscr{A}^{-1}h} = \widetilde{h} \quad \text{for all } h \in \mathscr{S}_{\epsilon\rho}(\mathfrak{a})^W.$$

Using this, we easily check that

$$\mathscr{A}^{-1}(h_1 * h_2) = \mathscr{A}^{-1}h_1 * \mathscr{A}^{-1}h_2 \quad \text{and} \quad \mathscr{A}^{-1}h_1^* = \big(\mathscr{A}^{-1}h_1\big)^*$$

for all $h_1, h_2 \in \mathscr{S}_{\epsilon\rho}(\mathfrak{a})^W$. Then

$$T_f(h * h^*) = \int_G f(x)\,\big(\mathscr{A}^{-1}h * (\mathscr{A}^{-1}h)^*\big)\,(x^{-1})\,dx \geq 0,$$

since f is K-positive definite. Therefore, by Theorem 3.2, there exists a unique positive tempered measure σ on $\mathscr{M} \cap \mathfrak{a}_\epsilon^*$, invariant under the Weyl group action, such that

$$T_f(h) = \int_{\mathscr{M}\cap\mathfrak{a}_\epsilon^*} \widetilde{h}(\lambda)\,d\sigma(\lambda) \quad \text{for all } h \in \mathscr{S}_{\epsilon\rho}(\mathfrak{a})^W.$$

This shows that

$$\text{(3-1)} \qquad \int_G f(x)g(x^{-1})\,dx = \int_{\mathscr{M}\cap\mathfrak{a}_\epsilon^*} \widehat{g}(\lambda)\,d\sigma(\lambda) \quad \text{for all } g \in \mathscr{C}^p(G/\!/K).$$

Now we show that the measure σ is finite. For this let $\{g_n\}$ be a Dirac-delta sequence in $\mathscr{C}^p(G/\!/K)$. Then $\{g_n * g_n^*\}$ is also a Dirac-delta sequence in $\mathscr{C}^p(G/\!/K)$. Applying this sequence to the previous equation we get

$$\int_G f(x)(g_n * g_n^*)(x^{-1})\,dx = \int_{\mathscr{M}\cap\mathfrak{a}_\epsilon^*} |\widehat{g_n}(\lambda)|^2\,d\sigma(\lambda).$$

Now we take the limit as $n \to \infty$ on both sides of the equation and apply Fatou's lemma to get $\sigma(\mathscr{M} \cap \mathfrak{a}_\epsilon^*) \leq f(e)$. Therefore the measure σ is finite. From (3-1) we get, using Fubini's theorem,

$$\int_G f(x)g(x^{-1})\,dx = \int_{\mathscr{M}\cap\mathfrak{a}_\epsilon^*} \widehat{g}(\lambda)\,d\sigma(\lambda) = \int_G g(x^{-1}) \int_{\mathscr{M}\cap\mathfrak{a}_\epsilon^*} \phi_\lambda(x)\,d\sigma(\lambda)\,dx.$$

This is true for every $g \in \mathscr{C}^p(G/\!/K)$. Hence

$$f(x) = \int_{\mathscr{M}\cap\mathfrak{a}_\epsilon^*} \phi_\lambda(x)\,d\sigma(\lambda) \quad \text{with } \epsilon = \frac{2}{p} - 1. \qquad \square$$

We can easily check that a function f on G which has an integral representation as in Theorem 3.3 is a K-positive definite function. That is, the converse of the Theorem 3.3 holds.

A K-biinvariant distribution T on G is called a K-positive definite distribution if $T(\phi * \phi^*) \geq 0$ for all $\phi \in C_c^\infty(G/\!/K)$. It is a positive definite distribution if the inequality above holds for all $\phi \in C_c^\infty(G)$. Barker [1975, p. 201] raised the question whether a K-positive definite distribution is a positive definite distribution. We shall see that the answer is negative, that is, the set of positive definite distributions is a proper subset of the set of K-positive definite distributions. For this let us consider $\lambda_0 \in \mathcal{M} \setminus \mathfrak{a}_1^*$. Our claim is that ϕ_{λ_0} is a K-positive definite distribution but not a positive definite distribution. By the Helgason–Johnson theorem ϕ_λ is bounded if and only if $\lambda \in \mathfrak{a}_1^*$. Since $\lambda_0 \notin \mathfrak{a}_1^*$, ϕ_{λ_0} is not bounded. Therefore, ϕ_{λ_0} is not a positive definite function. Hence ϕ_{λ_0} is not a positive definite distribution. Now $\lambda_0 \in \mathcal{M}$ implies that there exists $w \in W$ such that $w.\lambda_0 = \overline{\lambda}_0$. This shows that $\phi_{\overline{\lambda}_0} = \phi_{\lambda_0}$. Therefore, for a suitable K-biinvariant function f on G,

$$\int_G \phi_{\lambda_0}(x)(f * f^*)(x^{-1})\, dx = \hat{f}(\lambda_0)\widehat{f^*}(\lambda_0) = |\hat{f}(\lambda_0)|^2 \geq 0.$$

This proves our claim.

The same example also shows that the set of positive definite functions is a proper subset of the set of K-positive definite functions.

We now see in the real rank-one case that if we restrict our attention to certain classes of functions, then the set of positive definite functions is same as the set of K-positive definite functions. Any real rank-one connected, noncompact, semisimple Lie group G with finite centre can be classified (up to coverings) as

(1) $G = \mathrm{SO}_e(1, n)$, for which $m_\alpha = n - 1$ and $m_{2\alpha} = 0$,

(2) $G = \mathrm{SU}(1, n)$, for which $m_\alpha = 2n - 2$ and $m_{2\alpha} = 1$,

(3) $G = \mathrm{Sp}(1, n)$, for which $m_\alpha = 4n - 4$ and $m_{2\alpha} = 3$, or

(4) $G = \mathrm{F}_{4(-20)}$, for which $m_\alpha = 8$ and $m_{2\alpha} = 7$.

Let $\mathcal{P}_K$ and $\mathcal{P}$ be the set of K-positive definite functions and the set of positive definite functions on G respectively.

Proposition 3.4. (1) *For the groups* $G = \mathrm{SO}_e(1, n)$ *and* $G = \mathrm{SU}(1, n)$, *we have* $\mathcal{P}_K \cap L^\infty(G/\!/K) = \mathcal{P} \cap L^\infty(G/\!/K)$.

(2) *For the group* $G = \mathrm{Sp}(1, n)$, *we have* $\mathcal{P}_K \cap L^r(G/\!/K) = \mathcal{P} \cap L^r(G/\!/K)$, *for any* $2 < r \leq (2n + 1)$.

(3) *For the group* $G = \mathrm{F}_{4(-20)}$, *we have* $\mathcal{P}_K \cap L^r(G/\!/K) = \mathcal{P} \cap L^r(G/\!/K)$, *for any* $2 < r \leq \frac{11}{3}$.

Proof. It is known [Flensted-Jensen and Koornwinder 1979] that ϕ_λ is positive definite if and only if $\lambda \in \mathbb{R}$ or $\lambda = i\eta$ for $\eta \in [-s_0, s_0] \cup \{\pm\rho\}$, where $s_0 = \rho$ if $m_{2\alpha} = 0$, otherwise $s_0 = \frac{1}{2}m_\alpha + 1$. Therefore:

(a) For the groups $G = \mathrm{SO}_e(1, n)$ and $G = \mathrm{SU}(1, n)$, ϕ_λ is positive definite if and only if $\lambda \in \mathbb{R}$ or $\lambda = i\eta$ with $\eta \in [-\rho, \rho]$.

(b) For the group $G = \mathrm{Sp}(1, n)$, ϕ_λ is positive definite if and only if $\lambda \in \mathbb{R}$ or $\lambda = i\eta$ with $\eta \in [-(2n-1), (2n-1)] \cup \{\pm(2n+1)\}$.

(c) For the group $G = \mathrm{F}_{4(-20)}$, ϕ_λ is positive definite if and only if $\lambda \in \mathbb{R}$ or $\lambda = i\eta$ with $\eta \in [-5, 5] \cup \{\pm 11\}$.

We know that

$$\phi_\lambda \in L^\infty(G/\!/K) \text{ if and only if } \lambda \in S_1, \tag{3-2}$$

where $S_r = \{\lambda \in \mathbb{C} : |\mathrm{Im}\,\lambda| \le (2/r - 1)\rho\}$, $r > 0$. Also, for $r > 1$,

$$\phi_\lambda \in L^{r'}(G/\!/K) \text{ if and only if } \lambda \in S_r^\circ, \tag{3-3}$$

where $1/r + 1/r' = 1$ [Pusti et al. 2011]. Now, from Theorem 3.3, a K-positive definite function can be expressed as

$$f(x) = \int_{\mathbb{R}} \phi_\lambda(x)\,d\mu_1(\lambda) + \int_{i\mathbb{R}} \phi_\lambda(x)\,d\mu_2(\lambda),$$

where μ_1 is a finite positive measure and μ_2 is a positive measure such that the integral $\int_{\mathbb{R}} \phi_{i\lambda}(x)\,d\mu_2(\lambda)$ exists for every $x \in G$.

If the K-positive definite function f is in $L^\infty(G/\!/K)$, the measure μ_2 must be supported in $i[-\rho, \rho]$ by (3-2). Therefore a K-positive definite function f which is in $L^\infty(G/\!/K)$ has an integral form:

$$f(x) = \int_{\mathbb{R}} \phi_\lambda(x)\,d\mu_1(\lambda) + \int_{i[-\rho,\rho]} \phi_\lambda(x)\,d\mu_2(\lambda).$$

However, this is a positive definite function for the groups $G = \mathrm{SO}_e(1, n)$ and $G = \mathrm{SU}(1, n)$, by (a). This proves (1).

To prove (2), note that if the K-positive definite function f is in $L^r(G/\!/K)$, $2 < r \le (2n+1)$ by (3-3) it follows that the measure μ_2 must be supported in

$$i\mathbb{R} \cap S_{r'}^\circ \subseteq i(-(2n-1), (2n-1)).$$

This proves that the function f is positive definite.

The proof of (3) is similar. □

Remark 3.5. For the groups $G = \mathrm{SO}_e(1, n)$ and $G = \mathrm{SU}(1, n)$, it follows from the above that a K-positive definite function is a positive definite function if and only if $f \in L^\infty(G/\!/K)$. However, a similar statement is not true for the groups $G = \mathrm{Sp}(1, n)$ and $G = \mathrm{F}_{4(-20)}$. In fact, for the group $G = \mathrm{Sp}(1, n)$, the function $\phi_{(2n+1)i}$ is K-positive definite as well as positive definite, but it does not belong to any $L^r(G/\!/K)$, $2 < r \le (2n+1)$. Similarly, for the group $G = \mathrm{F}_{4(-20)}$, the function

ϕ_{11i} is K-positive definite as well as positive definite but it does not belong to any $L^r(G/\!/K)$, $2 < r \leq \frac{11}{3}$.

Let $f \in \mathscr{C}^2(G/\!/K)$ be a K-positive definite function. Then by Theorem 3.3 there exists a finite positive measure σ, invariant under the Weyl group action such that

$$f(x) = \int_{\mathcal{M}\cap\mathfrak{a}^*} \phi_\lambda(x)\, d\sigma(\lambda).$$

However, this is a positive definite function on G because ϕ_λ is positive definite for $\lambda \in \mathfrak{a}^*$. Hence $\mathscr{P}_K \cap \mathscr{C}^2(G/\!/K) = \mathscr{P} \cap \mathscr{C}^2(G/\!/K)$ (cf. [Bopp 1979] for distributions).

4. τ-positive definite functions

In this section we give an example in which the set of τ-positive definite functions is same as the set of positive definite functions (without imposing any decay condition on functions). For defining the τ-positive definite functions we recall some basic facts [Camporesi 1997; Camporesi and Pedon 2001].

Definition 4.1. For $\tau \in \widehat{K}$ a scalar valued function f on G is said to be τ-radial if $f(kxk^{-1}) = f(x)$ for all $k \in K$, $x \in G$ and if $d_\tau \overline{\chi}_\tau * f = f = f * d_\tau \overline{\chi}_\tau$, where χ_τ and d_τ are respectively the character and dimension of τ.

When τ is the trivial representation of K, a τ-radial function is a K-biinvariant function. We note that the τ-radial functions are radial sections of the homogeneous vector bundle over G/K associated with the representation $\tau \in \widehat{K}$. The set of all compactly supported τ-radial infinitely differentiable functions and infinitely differentiable τ-radial functions are denoted by $C^\infty_{c,\tau}(G)$ and $C^\infty_\tau(G)$, respectively.

Definition 4.2. A τ-radial continuous function f on G is called *τ-positive definite* if

$$\int_G f(x)(g * g^*)(x^{-1})\, dx \geq 0, \quad \text{for all } g \in C^\infty_{c,\tau}(G).$$

Let $G = \mathrm{Spin}_e(n, 1)$, the identity component of $\mathrm{Spin}(n, 1)$. Then, in the notation of the previous section, $K = \mathrm{Spin}(n)$ and $M = \mathrm{Spin}(n-1)$. In the rest of the section we fix these meanings for G, K, M.

Let τ_n be the complex spin representation of K. The following proposition gives information about the irreducibility of τ_n.

Proposition 4.3 [Camporesi and Pedon 2001]. (1) *If n is even, then τ_n splits into two irreducible components given by the positive and negative half-spin representations $\tau_n = \tau_n^+ \oplus \tau_n^-$ and $\tau_n^\pm|_M = \sigma_{n-1}$, where σ_{n-1} is the spin representation of M.*

(2) *If n is odd, then τ_n is irreducible and $\tau_n|_M = \sigma_{n-1}^+ \oplus \sigma_{n-1}^-$, where $\sigma_{n-1}^\pm$, are irreducible components of the spin representation σ_{n-1} of M.*

It is known that (G, K, τ) is a Gelfand triple, that is, the convolution algebra $C^\infty_{c,\tau}(G)$ is commutative when $\tau \in \widehat{K}$ is either τ_n^+ or τ_n^- if n is even and τ_n if n is odd. For n even the $\tau_n^\pm$-spherical function is given by

$$\phi_\lambda^{\tau_n^\pm}(x) = \int_K e^{-(i\lambda+\rho)H(xk)} \chi_{\tau_n^\pm}(kK(xk)^{-1})\,dk.$$

Also, it satisfies $\phi_{-\lambda}^{\tau_n^\pm}(x) = \phi_\lambda^{\tau_n^\pm}(x)$.

For n odd the τ_n-spherical functions are denoted by $\phi^{\tau_n}_{\sigma^+_{n-1},\lambda}$ and $\phi^{\tau_n}_{\sigma^-_{n-1},\lambda}$. They are given by the integral formula

$$\phi^{\tau_n}_{\sigma^\pm_{n-1},\lambda}(x) = 2d_{\sigma^\pm_{n-1}} \int_K \int_M e^{-(i\lambda+\rho)H(xk)} \chi_{\tau_n}(km^{-1}K(xk)^{-1}) \chi_{\sigma^\pm_{n-1}}(m)\,dm\,dk.$$

They satisfy $\phi^{\tau_n}_{\sigma^+_{n-1},-\lambda}(x) = \phi^{\tau_n}_{\sigma^-_{n-1},\lambda}(x)$.

From now on by $\tau \in \widehat{K}$ we will mean either $\tau = \tau_n^+$ or $\tau = \tau_n^-$ if n is even and $\tau = \tau_n$ if n is odd. For n even we shall write the $\tau_n^\pm$-spherical functions $\phi^\tau_{\sigma,\lambda}$ instead of $\phi^{\tau_n^\pm}_{\sigma_{n-1},\lambda}$. Also for n odd we write the τ_n-spherical functions

$$\phi^\tau_{\sigma^\pm,\lambda} \quad \text{instead of} \quad \phi^{\tau_n}_{\sigma^\pm_{n-1},\lambda}.$$

Henceforth while dealing with $G = \mathrm{Spin}_e(n, 1)$ and τ as above we shall simply say *when n is even* and *when n is odd* to distinguish between these two cases.

For a τ-radial function f its spherical Fourier transform is defined by

$$\hat{f}(\sigma, \lambda) = \int_G f(x)\phi^\tau_{\sigma,\lambda}(x^{-1})\,dx$$

when n is even. For n odd it is defined by

$$\hat{f}(\sigma^\pm, \lambda) = \int_G f(x)\phi^\tau_{\sigma^\pm,\lambda}(x^{-1})\,dx.$$

Theorem 4.4 [Gelfand and Vilenkin 1964, Theorem 3, p. 157, Theorem 5, p. 226].

(a) *Let T be a positive definite distribution on $\mathbb{R}$, that is,*

$$T(\phi * \phi^*) \geq 0 \quad \textit{for all } \phi \in C_c^\infty(\mathbb{R}).$$

Then there exists a positive tempered measure μ on $\mathbb{R}$ such that

$$T(\phi) = \int_{\mathbb{R}} \widetilde{\phi}(\lambda)\,d\mu(\lambda) \quad \textit{for all } \phi \in C_c^\infty(\mathbb{R}).$$

(b) *Let T be an evenly positive definite distribution on $\mathbb{R}$, that is, $T(\phi * \phi^*) \geq 0$ for all $\phi \in C_c^\infty(\mathbb{R})_e$. Then there exists positive even measures μ_1 and μ_2 such that*

$$T(\phi) = \int_{\mathbb{R}} \widetilde{\phi}(\lambda)\,d\mu_1(\lambda) + \int_{\mathbb{R}} \widetilde{\phi}(i\lambda)\,d\mu_2(\lambda) \quad \textit{for all } \phi \in C_c^\infty(\mathbb{R})_e,$$

where μ_1 is a tempered measure and μ_2 is such that

$$\int_{\mathbb{R}} e^{a|\lambda|}\, d\mu_2(\lambda) < \infty \quad \textit{for all } a > 0.$$

The next theorem gives integral representations of τ-positive definite functions on $G = \mathrm{Spin}_e(n, 1)$.

Theorem 4.5. *Let $G = \mathrm{Spin_e}(n, 1)$ and let τ denote one of $\{\tau_n^+, \tau_n^-\}$ when n is even and τ_n when n is odd.*

(a) *Let n be even and let f be a τ-positive definite function on G. Then there exists even positive measures μ_1 and μ_2 such that for all $x \in G$*

$$f(x) = \int_{\mathbb{R}} \phi^\tau_{\sigma,\lambda}(x)\, d\mu_1(\lambda) + \int_{\mathbb{R}} \phi^\tau_{\sigma,i\lambda}(x)\, d\mu_2(\lambda),$$

where μ_1 is finite measure and μ_2 is such that

$$\int_{\mathbb{R}} e^{a|\lambda|}\, d\mu_2(\lambda) < \infty \quad \textit{for all } a > 0.$$

(b) *Let n be odd and let f be a τ-positive definite function on G. Then there exists a finite positive measure μ such that for all $x \in G$*

$$f(x) = \int_{\mathbb{R}} \phi^\tau_{\sigma^+,\lambda}(x)\, d\mu(\lambda).$$

Proof. We shall prove (b). The proof of (a) is similar. Let n be odd and let f be a τ-positive definite function on G. We define the linear functional T_f on $C_c^\infty(\mathbb{R})$ as follows:

$$T_f(h) = \int_G f(x)(\mathscr{A}^{-1}h)(x^{-1})\, dx \quad \text{for all } h \in C_c^\infty(\mathbb{R}).$$

Here $\mathscr{A}$ is the Abel transform, which is a topological isomorphism between $C^\infty_{c,\tau}(G)$ and $C_c^\infty(\mathbb{R})$. We also have $\hat{f}(-\lambda) = \widetilde{\mathscr{A}f}(\lambda)$ for all $f \in C^\infty_{c,\tau}(G)$. Then it follows that $\widehat{\mathscr{A}^{-1}h}(-\lambda) = \widetilde{h}(\lambda)$ for all $h \in C_c^\infty(\mathbb{R})$. Using this, we easily check that

$$\mathscr{A}^{-1}(h_1 * h_2) = \mathscr{A}^{-1}h_1 * \mathscr{A}^{-1}h_2$$

and $\mathscr{A}^{-1}h_1^* = (\mathscr{A}^{-1}h_1)^*$ for all $h_1, h_2 \in C_c^\infty(\mathbb{R})$. Then

$$T_f(h * h^*) = \int_G f(x)(\mathscr{A}^{-1}h * (\mathscr{A}^{-1}h)^*)(x^{-1})\, dx \geq 0$$

as f is τ-positive definite. Therefore, by (a), there exists a positive tempered measure μ on $\mathbb{R}$ such that for all $h \in C_c^\infty(\mathbb{R})$

$$T_f(h) = \int_{\mathbb{R}} \widetilde{h}(\lambda)\, d\mu(\lambda).$$

This shows that for all $g \in C^\infty_{c,\tau}(G)$

$$\int_G f(x)g(x^{-1})\,dx = \int_{\mathbb{R}} \widehat{g}_+(\lambda)\,d\mu(\lambda). \tag{4-1}$$

Using approximate identity techniques we can easily prove that the measure μ is finite. Then from Equation (4-1), using Fubini's theorem we get

$$\begin{aligned}\int_G f(x)g(x^{-1})\,dx &= \int_{\mathbb{R}}\int_G g(x)\phi^\tau_{\sigma^+,\lambda}(x^{-1})\,dx\,d\lambda\\ &= \int_G g(x^{-1})\int_{\mathbb{R}} \phi^\tau_{\sigma^+,\lambda}(x)\,d\mu(\lambda)\,dx.\end{aligned}$$

Since this is true for every $g \in C^\infty_{c,\tau}(G)$, it follows that

$$f(x) = \int_{\mathbb{R}} \phi^\tau_{\sigma^+,\lambda}(x)\,d\mu(\lambda).$$

□

It is easy to check that the converse of Theorem 4.5 holds true. We get the following corollary from Theorem 4.5(b):

Corollary 4.6. *The set of τ-positive definite functions is same as the set of positive definite functions when $\tau = \tau_n$ and n is odd.*

Remark 4.7. We saw after Theorem 3.3 that the function ϕ_{λ_0}, $\lambda_0 \in \mathcal{M} \setminus \mathfrak{a}^*_1$ is K-positive definite but not positive definite. When $\tau = \tau_n$ and n is odd, we could try to find a similar example by considering the function $\phi^\tau_{\sigma^+,\lambda_0}$, $\lambda_0 \in i\mathbb{R} \setminus i[-1, 1]$. But $\phi^\tau_{\sigma^+,\lambda_0}$ is neither a τ-positive definite function nor a positive definite function. The argument used in the K-positive definite case does not work here. Indeed, unlike the case of the spherical functions ϕ_λ, which are W-invariant in $\lambda \in \mathfrak{a}^*_{\mathbb{C}}$, there is no relation between $\phi^\tau_{\sigma^+,\lambda}$ and $\phi^\tau_{\sigma^+,-\lambda}$ when $\tau = \tau_n$ and n is odd.

We Now characterize the λ's for which $\phi^\tau_{\sigma,\lambda}$ is positive definite for

$$G = \mathrm{Spin_e}(n, 1)$$

when τ is the irreducible component of the complex spin representation.

Theorem 4.8. *Let $G = \mathrm{Spin_e}(n, 1)$ and let τ denote one of $\{\tau_n^+, \tau_n^-\}$ when n is even and τ_n when n is odd. Then*

(a) *$\phi^\tau_{\sigma,\lambda}$ is positive definite if and only if $\lambda \in \mathbb{R}$ when $n(\geq 4)$ is even, and*

(b) *$\phi^\tau_{\sigma^\pm,\lambda}$ are positive definite if and only if $\lambda \in \mathbb{R}$ when n is odd.*

Proof. (a) Let n be even and $n \geq 4$. The τ-spherical function $\phi^\tau_{\sigma,\lambda}$ is positive definite if and only if τ is contained in the unitary principal, discrete or complementary series representations. It is well-known that there is no discrete series representation which contains τ. Also, by [Knapp and Stein 1971, Proposition 55] and the Frobenius reciprocity theorem there is no complementary series containing τ. Hence $\phi^{\tau}_{\sigma,\lambda}$ is positive definite if and only if $\lambda \in \mathbb{R}$.

(b) For the case n odd we prove the result without using representation theory. By Corollary 4.6 the τ-spherical function $\phi^{\tau}_{\sigma^+,\lambda}$ is positive definite if and only if it is a τ-positive definite function. That is equivalent to

$$\int_G (f * f^*)(x)\phi^{\tau}_{\sigma^+,\lambda}(x^{-1})\,dx \geq 0 \quad \text{for all } f \in \mathscr{C}^2_\tau(G),$$

where $\mathscr{C}^2_\tau(G)$ is the set of τ-radial L^2-Schwartz class functions on G. That is,

$$\hat{f}(\sigma^+, \lambda)\overline{\hat{f}(\sigma^+, \overline{\lambda})} \geq 0 \quad \text{for all } f \in \mathscr{C}^2_\tau(G), \tag{4-2}$$

since $\overline{\phi^{\tau}_{\sigma,\lambda}(x)} = \phi^{\tau}_{\sigma,\overline{\lambda}}(x^{-1})$. Let us consider a function

$$f \in \mathscr{C}^2_\tau(G)$$

such that $\hat{f}(\sigma^+, \lambda) = \lambda e^{-\lambda^2}$. Such a function exists by the Schwartz space isomorphism theorem [Camporesi and Pedon 2001, Theorem 6.3]. Then (4-2) is true if and only if $\lambda \in \mathbb{R}$. □

Acknowledgments

We are grateful to Prof. J. Faraut for giving us the problem and many important suggestions.

References

[Anker 1991] J.-P. Anker, "The spherical Fourier transform of rapidly decreasing functions. A simple proof of a characterization due to Harish-Chandra, Helgason, Trombi, and Varadarajan", *J. Funct. Anal.* **96**:2 (1991), 331–349. MR 92d:22008 Zbl 0732.43006

[Barker 1975] W. H. Barker, "The spherical Bochner theorem on semisimple Lie groups", *J. Functional Analysis* **20**:3 (1975), 179–207. MR 53 #3197 Zbl 0308.43002

[Bopp 1979] N. Bopp, "Les distributions de type positif relativement à un groupe fini d'isométries", pp. 15–50 in *Analyse harmonique sur les groupes de Lie, II* (Sém.,Nancy-Strasbourg 1976–1978), edited by P. Eymard et al., Lecture Notes in Math. **739**, Springer, Berlin, 1979. MR 81e:22018 Zbl 0422.46034

[Camporesi 1997] R. Camporesi, "The spherical transform for homogeneous vector bundles over Riemannian symmetric spaces", *J. Lie Theory* **7**:1 (1997), 29–60. MR 98j:58114 Zbl 0871.43005

[Camporesi and Pedon 2001] R. Camporesi and E. Pedon, "Harmonic analysis for spinors on real hyperbolic spaces", *Colloq. Math.* **87**:2 (2001), 245–286. MR 2002c:43012 Zbl 0968.22009

[van Dijk and Pasquale 1999] G. van Dijk and A. Pasquale, "Harmonic analysis on vector bundles over $\mathrm{Sp}(1,n)/\mathrm{Sp}(1) \times \mathrm{Sp}(n)$", *Enseign. Math.* (2) **45**:3-4 (1999), 219–252. MR 2001b:43011 Zbl 0967.43005

[Flensted-Jensen and Koornwinder 1979] M. Flensted-Jensen and T. H. Koornwinder, "Positive definite spherical functions on a noncompact, rank one symmetric space", pp. 249–282 in *Analyse harmonique sur les groupes de Lie, II* (Sém., Nancy and Strasbourg 1976–1978), edited by P. Eymard et al., Lecture Notes in Math. **739**, Springer, Berlin, 1979. MR 81j:43015 Zbl 0433.43014

[Gelfand and Vilenkin 1964] I. M. Gelfand and N. Y. Vilenkin, *Applications of harmonic analysis*, Generalized functions **4**, Academic Press, New York, 1964. MR 30 #4152 Zbl 0136.11201

[Knapp and Stein 1971] A. W. Knapp and E. M. Stein, "Intertwining operators for semisimple groups", *Ann. of Math.* (2) **93** (1971), 489–578. MR 57 #536 Zbl 0257.22015

[Pusti 2011] S. Pusti, "An analogue of Bochner's theorem for Damek–Ricci spaces", preprint, 2011. arXiv 1111.3157

[Pusti et al. 2011] S. Pusti, S. K. Ray, and R. P. Sarkar, "Wiener–Tauberian type theorems for radial sections of homogenous vector bundles on certain rank one Riemannian symmetric spaces of noncompact type.", *Math. Z.* **269**:1-2 (2011), 555–586. Zbl 05963446

[Sitaram 1978] A. Sitaram, "Positive definite distributions on $K|G/K$. II", *J. Indian Math. Soc.* (*N.S.*) **42**:1-4 (1978), 95–104 (1979). MR 81e:43024 Zbl 0499.43010

Received July 10, 2011. Revised July 19, 2011.

SANJOY PUSTI
MATHEMATICS RESEARCH UNIT
UNIVERSITY OF LUXEMBOURG, CAMPUS KIRCHBERG
6, RUE RICHARD COUDENHOVE-KALERGI
1359 LUXEMBOURG
LUXEMBOURG

sanjoy.pusti@uni.lu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

UNE REMARQUE DE DYNAMIQUE SUR LES VARIÉTÉS SEMI-ABÉLIENNES

GAËL RÉMOND

Soit φ l'endomorphisme de multiplication par un entier sur une variété semi-abélienne A définie sur $\overline{\mathbb{Q}}$ et soit X une sous-variété algébrique de A. Il existe (pour des raisons évidentes) un entier N avec la propriété suivante : si les N premiers itérés de φ envoient un point x dans X alors ceci vaut pour tous les itérés. Nous montrons que N peut être choisi indépendamment de φ. Nous montrons aussi qu'un tel N peut être calculé explicitement si A est une variété abélienne ou un tore. La preuve repose sur un résultat d'effectivité dans la solution de la conjecture de Mordell–Lang et sur un résultat combinatoire de Crittenden et Vanden Eynden sur les progressions arithmétiques.

Let φ be the endomorphism of multiplication by an integer on a semi-abelian variety A defined over $\overline{\mathbb{Q}}$ and let X be an algebraic subvariety of A. There exists (for obvious reasons) an integer N with the property that if the first N iterates of φ map a point x into X then this is true of all iterates. We prove that N can be chosen independently of φ. Moreover we show that such an N can be explicitly computed if A is either an abelian variety or a torus. The proof relies on an effectivity result in the solution of the Mordell–Lang conjecture together with a combinatorial result of Crittenden and Vanden Eynden on arithmetic progressions.

1. Introduction

Lorsque l'on dispose d'une application $\varphi\colon A \to A$ d'un ensemble A dans lui-même, on peut étudier la dynamique de cette application, c'est-à-dire s'intéresser à l'action des itérés de φ. Dans cette note, nous considérons en particulier le problème suivant : une partie X de A étant fixée, quels sont les points dont toute l'orbite est contenue dans X ? Autrement dit, nous introduisons la *partie stable* de X

$$Y = \{x \in X \mid \varphi^n(x) \in X \text{ pour tout } n \in \mathbb{N}\} = \bigcap_{n\in\mathbb{N}} (\varphi^n)^{-1}X.$$

MSC2010: 11G10, 11G35.
Mots-clefs: semi-abelian varieties, algebraic dynamics, Mordell–Lang problem, variétés semi-abéliennes, problème de Mordell–Lang effectif, dynamique algébrique.

Dans la suite, A et X seront des variétés algébriques et φ un morphisme de variétés. Dans ce cadre, on constate immédiatement que Y est un fermé de X et de plus (par nœthérianité) qu'il existe un entier N tel que

$$Y = \bigcap_{n=0}^{N} (\varphi^n)^{-1} X.$$

Nous étudions plus précisément la valeur de cet entier N lorsque A est une variété semi-abélienne définie sur le corps $\overline{\mathbb{Q}}$ et φ un endomorphisme de A laissant stable toute sous-variété semi-abélienne de A. Bien entendu cette condition sur φ (nécessaire dans la preuve actuelle, voir aussi la remarque à la fin de la partie 3) s'avère assez restrictive. Elle comporte toutefois, comme exemple principal, le cas de $\varphi = [m]$ la multiplication par un entier m (au sens de la loi de groupe de A notée additivement). En outre, on peut construire des exemples différents de φ convenables comme $[m_1] \times \cdots \times [m_s]$ sur un produit $\prod_{i=1}^{s} A_i$ de variétés abéliennes A_i vérifiant $\mathrm{Hom}(A_i, A_j) = 0$ si $i \neq j$. Dans le cas d'une variété abélienne à multiplication complexe, on peut également penser à des relevés d'endomorphismes de Frobenius. Dans le cas du tore $A = \mathbb{G}_{\mathrm{m}}^n$, en revanche, la condition de stabilité force $\varphi = [m]$ pour un entier m.

Dans ce cadre, notre résultat principal affirme que l'entier N peut être choisi indépendamment de φ.

Théorème 1.1. *Pour toute variété semi-abélienne A définie sur $\overline{\mathbb{Q}}$ et tout fermé X de A, il existe un entier N tel que pour tout endomorphisme φ de A laissant stable toute sous-variété semi-abélienne on ait*

$$\bigcap_{n\in\mathbb{N}} (\varphi^n)^{-1} X = \bigcap_{n=0}^{N} (\varphi^n)^{-1} X.$$

Précisons que, dans cet énoncé, on ne suppose pas X définie sur $\overline{\mathbb{Q}}$. De plus, dans certains cas, nous pouvons même calculer une valeur pour N.

Théorème 1.2. *Soient X un fermé de $\mathbb{G}_{\mathrm{m},\mathbb{C}}^g$ et m un entier. Alors*

$$\bigcap_{n\in\mathbb{N}} [m^n]^{-1} X = \bigcap_{n=0}^{2^S-1} [m^n]^{-1} X,$$

où $S = (\deg X)^{2g^2(\dim X+1)^{3(\dim X+1)^2}}$, le degré étant calculé dans le plongement de Segre $\mathbb{G}_{\mathrm{m}}^g \hookrightarrow (\mathbb{P}^1)^g \hookrightarrow \mathbb{P}^{2^g-1}$.

Pour une variété abélienne principalement polarisée A définie sur $\overline{\mathbb{Q}}$ nous disposons aussi d'une valeur effective de la même forme en remplaçant S par la partie entière de

$$(2^{34} h_0(A) \deg X)^{3g^{5(\dim X+1)^2+2}},$$

où $h_0(A)$ est une constante ne dépendant que de A (comme variété polarisée ; il s'agit essentiellement d'un terme de hauteur, voir partie suivante).

Le théorème 1.2 améliore un résultat d'Aliev et Smyth [2008] qui fournissait une borne dépendant de m et beaucoup plus grande (définie de manière récursive, elle comportait au moins une tour de puissance à n étages).

La démonstration consiste à remarquer que les itérés $\varphi^n(x)$ se trouvent dans un groupe de rang fini Γ de A. Les résultats sur la conjecture de Mordell–Lang nous disent alors que les points de $X \cap \Gamma$ se répartissent sur un nombre fini S de translatés de sous-variétés semi-abéliennes tracées sur X. Ensuite un calcul élémentaire montre que les entiers n tels que $\varphi^n(x)$ appartient à un translaté donné sont en progression arithmétique. Par ailleurs un théorème de Crittenden et Vanden Eynden [1970] affirme que si S progressions arithmétiques couvrent les entiers de 0 à $2^S - 1$ alors elles couvrent tout $\mathbb{N}$. On en déduit immédiatement le théorème avec $N = 2^S - 1$ à condition bien sûr que l'entier S intervenant ne dépende ni de x ni de φ. Pour cela nous avons besoin d'une précision quantitative dans le problème de Mordell–Lang : il nous faut savoir que S ne dépend de Γ que par son rang. Nous examinons ceci dans la partie suivante (ainsi que les valeurs explicites dans certains cas) avant de passer à la démonstration du théorème principal dans la troisième partie. Enfin, nous donnons les arguments de spécialisation permettant d'étendre dans notre situation des résultats de $\overline{\mathbb{Q}}$ à un corps de caractéristique zéro quelconque.

2. Mordell–Lang

Les travaux de Laurent (pour les tores), Faltings, Hindry (pour les variétés abéliennes), Vojta et McQuillan (cas général) permettent de donner l'énoncé qualitatif suivant.

Théorème 2.1. *Soit A une variété semi-abélienne sur un corps de caractéristique zéro. Pour tout sous-groupe Γ de rang fini de A et tout fermé X de A, on peut écrire*

$$X \cap \Gamma = \bigcup_{i=1}^{S} (x_i + B_i) \cap \Gamma$$

pour un entier S, des sous-variétés semi-abéliennes B_i et des points x_i avec la condition $x_i + B_i \subset X$.

Pour donner des versions quantitatives de ce résultat, nous introduisons quelques notations. Nous écrirons $g = \dim A$ et $m = \dim X + 1$. Nous désignons en outre par r le rang de Γ (on rappelle qu'il s'agit de la dimension du $\mathbb{Q}$-espace vectoriel $\Gamma \otimes \mathbb{Q}$).

Comme dans le théorème 1.2, si A est un tore (c'est-à-dire $\mathbb{G}_{\mathrm{m}}^{g}$), nous le plongeons dans $\mathbb{P}^{2^g-1}$ à la Segre et utilisons ce plongement pour parler du degré de X.

Théorème 2.2. *Si A est un tore le théorème 2.1 vaut avec*

$$S = (\deg X)^{(r+1)g^2m^{3m^2}}.$$

Si $K = \overline{\mathbb{Q}}$ ceci est contenu dans le théorème 1.1 de [Rémond 2002]. Pour passer au cas général, on met en œuvre un argument de spécialisation. La démarche est assez classique mais l'argument ne semble pas être écrit sous la forme où nous en avons besoin (on trouvera dans la partie 3 de [Evertse et al. 2002] le cas où X est un hyperplan et où l'on se limite à des points non dégénérés). Nous le redonnons donc en détails dans la dernière partie (voir théorème 4.1).

Dans le cas d'une variété abélienne sur $\overline{\mathbb{Q}}$ nous avons un résultat de même nature. Il provient essentiellement de [Rémond 2000] précisé par une estimation de David et Philippon pour le problème de Bogomolov effectif. Nous le citons sous la forme du théorème 1.3 de [Rémond 2002] qui impose que A soit principalement polarisée par un faisceau inversible symétrique $\mathscr{L}$. Dans ce cas, on définit une hauteur thêta h_θ de A associée à $\mathscr{L}^{\otimes 16}$ (la hauteur de l'origine dans le plongement thêta) et on pose $h_0(A) = [L : \mathbb{Q}] \max(1, h_\theta)$ où L est un corps de définition de $(A, \mathscr{L})$.

Théorème 2.3. *Si $(A, \mathscr{L})$ est une variété abélienne principalement polarisée définie sur $\overline{\mathbb{Q}}$, le théorème 2.1 vaut avec*

$$S = \left[(2^{34} h_0(A) \deg_{\mathscr{L}} X)^{(r+1)g^{5m^2}}\right].$$

Comme plus haut, le résultat de [Rémond 2002] donne ceci avec la restriction supplémentaire $K = \overline{\mathbb{Q}}$ (c'est-à-dire X et Γ définis sur $\overline{\mathbb{Q}}$) et le théorème 4.1 ci-dessous permet de s'en affranchir.

En général, pour une variété semi-abélienne, nous ne disposons pas à l'heure actuelle d'un énoncé effectif dans le problème de Bogomolov. Ceci nous empêche de donner une formule explicite dans ce cas. Avec les méthodes existantes, nous pouvons tout de même montrer un résultat suffisamment uniforme.

Théorème 2.4. *Si A est une variété semi-abélienne définie sur $\overline{\mathbb{Q}}$, le théorème 2.1 vaut avec*

$$S = f(A, \deg X)^{r+1}$$

pour une certaine fonction f (le degré étant calculé dans n'importe quel plongement projectif de A).

Démonstration. Nous reprenons la stratégie de [Rémond 2000] lorsque $K = \overline{\mathbb{Q}}$ (là encore, on passe au cas général grâce au théorème 4.1). La preuve peut se scinder en trois étapes :

1. Les grands points ne sont pas trop espacés (inégalité de Vojta).
2. Les grands points sont assez espacés (inégalité de Mumford).
3. Les petits points sont assez espacés (propriété de Bogomolov).

Dans le cas semi-abélien, l'étape 1 est entièrement décrite dans [Rémond 2003]. La deuxième étape n'est écrite que dans les cas abélien [Rémond 2000] et torique [Rémond 2002]. Toutefois la comparaison de ces deux cas montre que la preuve s'étend à l'identique dans le cas semi-abélien. Dans la première étape, on obtient des constantes qui ne dépendent que de A et du degré de X. Dans la seconde intervient en plus une borne en dimension inférieure. En raisonnant par récurrence, nous pouvons supposer que celle-ci est de la forme $f_1(A, \deg X)^{r+1}$.

Pour la dernière étape, nous n'avons pas de constante explicite mais la proposition 3 de [Rémond 2005] (avec $\Gamma = 0$) montre qu'elle ne dépend aussi que de A et de $\deg X$.

Nous pouvons alors combiner les trois étapes pour l'estimation de S. Les deux premières permettent de compter les grands points en les répartissant en un nombre fini de cônes dans $\Gamma \otimes \mathbb{R}$ d'angle au sommet fixé (indépendant de Γ). Ceci conduit donc à une borne de la forme $f_2(A, \deg X)^{r+1}$ pour le nombre de grands points.

De manière analogue, l'étape 3 permet de répartir les petits points dans un nombre fini de boules de rayon fixé (indépendant de Γ) et ceci conduit à une borne de la même forme pour le nombre de petits points. Dans les deux cas, on trouvera des énoncés de décompte précis pour ces cônes ou boules au sens de la hauteur dans le lemme 5.1 de [Rémond 2003]. □

Nous nous sommes limités ici au cas arithmétique où A est définie sur $\overline{\mathbb{Q}}$. Cependant la preuve du théorème principal serait tout aussi valable dans d'autres cas si l'on connaît l'existence d'une borne ne dépendant de Γ que par son rang. Par exemple, dans un cas diagonalement opposé au nôtre, Buium [1993] donne une telle borne si A est une variété abélienne dont aucune sous-variété abélienne non nulle n'est définie sur $\overline{\mathbb{Q}}$ et X est lisse. Sa borne est explicite (mais beaucoup plus grande que les nôtres dans la mesure où elle comporte une factorielle itérée au moins $3 \max(r, \deg X)^2$ fois).

3. Démonstration du théorème principal

Nous nous plaçons donc sous les hypothèses du théorème 1.1. Nous associons à tout point x le groupe

$$\Gamma_x = \mathrm{End}(A) \cdot x = \{\psi(x) \mid \psi \in \mathrm{End}(A)\}$$

de rang fini au plus égal au rang r_0 de $\mathrm{End}(A)$, que l'on sait être fini. En vue d'appliquer le théorème 2.4, nous posons $S = f(A, \deg X)^{r_0+1}$ puis $N = 2^S - 1$.

Nous écrivons comme dans le théorème 2.1

$$X \cap \Gamma_x = \bigcup_{i=1}^{S} (x_i + B_i) \cap \Gamma_x$$

avec la condition $x_i + B_i \subset X$ (noter que comme l'on a augmenté S il peut être nécessaire de répéter certains translatés).

Supposons à présent $x \in \bigcap_{n=0}^{N} (\varphi^n)^{-1} X$. Bien entendu nous avons $\varphi^n(x) \in X \cap \Gamma_x$ pour $0 \le n \le N$. Définissons ensuite pour $1 \le i \le S$ la partie

$$P_i = \{n \in \mathbb{N} \mid \varphi^n(x) \in x_i + B_i\}.$$

La remarque précédente se traduit par $\{0, \ldots, N\} \subset \bigcup_{i=1}^{S} P_i$. Notre but est de montrer $\bigcup_{i=1}^{S} P_i = \mathbb{N}$. Pour l'atteindre, nous allons établir que chaque P_i est une progression arithmétique.

Soient donc $n_1, n_2, n_3 \in P_i$ avec $n_1 \le n_2$. Nous écrivons

$$\varphi^{n_3+n_2-n_1}(x) - \varphi^{n_2}(x) = \varphi^{n_2-n_1}\big(\varphi^{n_3}(x) - \varphi^{n_1}(x)\big).$$

Dans cette expression, $\varphi^{n_3}(x) - \varphi^{n_1}(x)$ est la différence de deux éléments de $x_i + B_i$ donc appartient à B_i. D'après l'hypothèse sur φ, on a $\varphi(B_i) \subset B_i$ donc le second membre de notre égalité est également un point de B_i. Par suite $\varphi^{n_2}(x) \in x_i + B_i$ entraîne $\varphi^{n_3+n_2-n_1}(x) \in x_i + B_i$ et donc $n_3 + n_2 - n_1 \in P_i$.

Cette propriété assure que si P_i est non vide alors il est de la forme $a_i + b_i\mathbb{N}$ avec $a_i, b_i \in \mathbb{N}$: en effet, il suffit de définir a_i comme le plus petit élément de P_i et $Q_i = P_i - a_i \subset \mathbb{N}$; alors Q_i jouit de la même propriété et contient 0 ; il est donc stable par somme et par différence positive donc de la forme $b_i\mathbb{N}$.

Il reste simplement à appliquer le théorème-titre de [Crittenden et Vanden Eynden 1970] avec $n = S$ (on vérifie que la nullité éventuelle de certains b_i ne pose pas de problème). Nous avons donc comme prévu $\bigcup_{i=1}^{s} P_i = \mathbb{N}$ ce qui signifie que pour tout $n \in \mathbb{N}$ il existe i avec $\varphi^n(x) \in x_i + B_i \subset X$. Par conséquent, nous avons bien $x \in \bigcap_{n \in \mathbb{N}} (\varphi^n)^{-1} X$ et le théorème 1.1 est établi.

Nous en déduisons immédiatement la formule citée dans l'introduction pour les variétés abéliennes à l'aide du théorème 2.3 et de la majoration du rang $r_0 \le 2g^2$.

Pour les tores, on déduit le théorème 1.2 du théorème 2.2 en remarquant que cette fois le rang r_0 peut être remplacé par 1 : en effet, comme $\varphi = [m]$ l'argument s'applique au groupe $\mathbb{Z}x$ de rang au plus 1.

Remarque. Si l'on ne suppose pas que φ laisse stables les sous-variétés semi-abéliennes de A alors la partie P_i ci-dessus n'est plus en général une progression arithmétique. Toutefois Ghioca et Tucker [2009] ont montré que P_i est toujours une union finie de progressions arithmétiques. Si l'on savait majorer le nombre de progressions nécessaires par une borne T indépendante de x, de i et de φ alors la démonstration précédente permettrait de conclure de même avec $N = 2^{ST} - 1$.

4. Spécialisation

Dans cette partie, on fixe une variété semi-abélienne A_0 sur $\overline{\mathbb{Q}}$ ainsi qu'un plongement de A_0 dans $\mathbb{P}^n_{\overline{\mathbb{Q}}}$ qui nous permet de calculer le degré des fermés de A_0. Pour toute extension K de $\overline{\mathbb{Q}}$, tout fermé X de $A = A_0 \times \operatorname{Spec} K$ et toute partie F de $A(K)$, on note $\lambda_{K,X}(F) \in \overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ l'infimum des entiers naturels S pour lesquels il existe des sous-variétés semi-abéliennes B_i de A et des points $a_i \in A(K)$ tels que

$$X \cap F \subset \bigcup_{i=1}^{S} a_i + B_i \subset X.$$

(Si aucun tel S n'existe, on a $\lambda_{K,X}(F) = \infty$.) En outre on appelle rang de F le rang du sous-groupe de $A(K)$ engendré par F.

Nous montrons alors que toute borne suffisamment uniforme sur $\overline{\mathbb{Q}}$ s'étend automatiquement à K quelconque.

Théorème 4.1. *S'il existe une fonction $f : \overline{\mathbb{N}}^3 \to \overline{\mathbb{N}}$ telle que pour tous X et F on ait $\lambda_{\overline{\mathbb{Q}},X}(F) \leq f(\dim X, \deg X, \operatorname{rang} F)$ alors on a*

$$\lambda_{K,X}(F) \leq f(\dim X, \deg X, \operatorname{rang} F)$$

pour tous K, X et F.

Démonstration. Nous commençons par rappeler que l'on peut choisir les B_i dans un ensemble fini indépendant de X. On note $\Delta(X) = (\deg X)^{(\dim X+1)^2/4}$.

Lemme 1. *On ne change pas la définition de $\lambda_{K,X}(F)$ si l'on y impose de plus $a_i \in X \cap F$ et* $\deg B_i \leq \Delta(X)$.

Démonstration. Si $(a_i + B_i) \cap F = \varnothing$ on peut supprimer ce translaté en diminuant S. Si au contraire $x \in (a_i + B_i) \cap F$ alors $a_i + B_i = x + B_i$ et l'on peut remplacer a_i par $x \in X \cap F$. Pour le degré, on pose

$$Z = \bigcap_{b \in B_i} X - b \qquad \text{et} \qquad B = \operatorname{Stab}(Z) = \bigcap_{z \in Z} Z - z$$

de sorte que $\deg Z \leq (\deg X)^{\dim X - \dim Z + 1}$ et $\deg B \leq (\deg Z)^{\dim Z - \dim B + 1}$. Ces deux relations entraînent $\deg B \leq (\deg X)^{(\dim X + 2 - \dim B)^2/4}$ et donc, si l'on note B^0 la composante neutre de B, on a $\deg B^0 \leq \Delta(X)$ (car si $\dim B = 0$ alors $B^0 = 0$). Maintenant $a_i + B_i \subset X$ donne $a_i \in Z$ d'où $a_i + B \subset Z \subset X$. De plus $B_i \subset \operatorname{Stab}(Z)$ donne $B_i \subset B^0$ donc $a_i + B_i \subset a_i + B^0 \subset X$ ce qui montre que l'on peut remplacer B_i par B^0. □

Nous nous ramenons ensuite à une partie F finie.

Lemme 2. *Pour tout F on a*

$$\lambda_{K,X}(F) = \sup\{\lambda_{K,X}(F') \mid F' \subset F,\ F' \text{ finie}\}.$$

Démonstration. Il suffit de montrer que si le membre de droite est un entier $S \in \mathbb{N}$ alors $\lambda_{K,X}(F) \le S$. Fixons une partie finie F' de F telle que $\lambda_{K,X}(F') = S$ et considérons tous les choix de $a_i \in F'$ et de B_i de degré au plus $\Delta(X)$ tels que

$$X \cap F' \subset \bigcup_{i=1}^{S} a_i + B_i \subset X.$$

Ces choix sont en nombre fini. Si pour l'un d'entre eux on a $X \cap F \subset \bigcup_{i=1}^{S} a_i + B_i$ le résultat est acquis. Sinon on pourrait choisir à chaque fois un élément $y \in X \cap F$ avec $y \notin \bigcup_{i=1}^{S} a_i + B_i$. Nous formons alors la partie finie F'' constituée de F' et de tous les points y obtenus. Par hypothèse $\lambda_{K,X}(F'') \le S$ ce qui permet d'écrire

$$X \cap F'' \subset \bigcup_{i=1}^{S} a_i + B_i \subset X$$

où $a_i \in F''$ et $\deg B_i \le \Delta(X)$. Si $(a_i + B_i) \cap F' \ne \varnothing$ on peut supposer $a_i \in F'$. Si $(a_i + B_i) \cap F' = \varnothing$ on obtiendrait un recouvrement de $X \cap F'$ (inclus dans $X \cap F''$) par $S - 1$ translatés, ce qui contredit $\lambda_{K,X}(F') = S$. Par suite, le choix obtenu de a_i et B_i est l'un de ceux considérés ci-dessus donc il existe $y \in X \cap F''$ tel que $y \notin \bigcup_{i=1}^{S} a_i + B_i$, ce qui est contradictoire. □

Ce lemme implique que, pour montrer l'inégalité

$$\lambda_{K,X}(F) \le f(\dim X, \deg X, \operatorname{rang} F)$$

que nous avons en vue, nous pouvons supposer F finie. En outre, comme X et F sont maintenant définis sur un corps de type fini sur $\overline{\mathbb{Q}}$, nous pouvons supposer que K est un tel corps. Ceci permet de regarder X comme une famille algébrique de fermés de A_0 et de mettre en place un argument de spécialisation.

Plus précisément, considérons un anneau $R \subset K$ de type fini sur $\overline{\mathbb{Q}}$ dont le corps des fractions est K. Il correspond à un schéma intègre $V = \operatorname{Spec} R$ de type fini sur $\overline{\mathbb{Q}}$. Nous allons passer du point générique (de corps résiduel K) aux points fermés (de corps résiduel $\overline{\mathbb{Q}}$) dont on rappelle qu'ils forment une partie dense.

Dans la suite, nous aurons besoin à plusieurs reprises de remplacer V par un ouvert non vide. Nous le choisirons toujours de la forme $D(u) = \operatorname{Spec} R_u$ pour $u \in R \setminus \{0\}$: ceci revient à remplacer R par le localisé R_u sans changer aucune des propriétés de l'anneau ; nous conserverons la notation R pour notre anneau à chaque étape. Tout l'argument repose bien sûr sur la possibilité d'atteindre la situation désirée en un nombre *fini* de telles étapes.

Notons à présent $\mathscr{A} = A_0 \times V$ (de fibre générique A_0) et $\mathscr{X}$ l'adhérence de X dans $\mathscr{A}$ (munie de sa structure de sous-schéma fermé réduit). Le morphisme $\mathscr{X} \to V$ (de type fini) nous donne la famille algébrique cherchée. Quitte à restreindre V, nous pouvons la supposer plate de sorte que, pour tout $t \in V$, on ait $\dim \mathscr{X}_t = \dim X$ et $\deg \mathscr{X}_t = \deg X$ (pour cette dernière égalité, nous voyons $\mathscr{A}$ comme plongée dans $\mathbb{P}^n_R$ et donc $\mathscr{X}_t$ dans $\mathbb{P}^n_{k(t)}$).

Intéressons-nous ensuite aux points de l'ensemble F. Ils s'identifient à des sections $\operatorname{Spec} K \to A$ que nous souhaitons étendre en $\operatorname{Spec} R \to \mathscr{A}$. C'est possible après restriction à un ouvert : par exemple un point $\operatorname{Spec} K \to \mathbb{P}^n$ donné par des coordonnées $(x_0 : \cdots : x_n)$ choisies dans R s'étend en $\operatorname{Spec} R_{x_i} \to \mathbb{P}^n$ pour tout i tel que $x_i \neq 0$; ensuite la section $\operatorname{Spec} R \to \mathbb{P}^n$ se factorise en $s \colon \operatorname{Spec} R \to \overline{\mathscr{A}}$ (immersion fermée) ; enfin on retire à V, par une nouvelle restriction, le fermé strict $s^{-1}(\overline{\mathscr{A}} \setminus \mathscr{A})$ pour obtenir $V \to \mathscr{A}$. Ceci fait, tout élément $x \in F$ donne donc naissance à des spécialisations

$$x_t \in \mathscr{A}_t \simeq A_0 \times \operatorname{Spec} k(t)$$

pour $t \in V$.

Finalement, nous restreignons V pour avoir les propriétés suivantes pour tous $x, y \in F$, toute sous-variété semi-abélienne B de A_0 de degré au plus $\Delta(X)$ et tout $t \in V$:

(1) Si $x - y \notin B$ alors $x_t - y_t \notin B$.

(2) Si $x + B \not\subset X$ alors $x_t + B \not\subset \mathscr{X}_t$.

(On note ici B pour $B \times \operatorname{Spec} k(t') \subset \mathscr{A}_{t'}$ tant pour $t' = t$ que pour le point générique de V.) Ceci est possible car il y a un nombre fini de conditions et chacune revient à demander comme ci-dessus à une section $\operatorname{Spec} R \to \mathscr{A}$ d'éviter un fermé : c'est direct pour (1) et on choisit pour (2) un point de torsion ξ de B (défini sur $\overline{\mathbb{Q}}$) tel que $x + \xi \notin X$ de sorte qu'il suffit bien sûr d'assurer $x_t + \xi \notin \mathscr{X}_t$.

Nous pouvons alors fixer définitivement un point fermé t de V. En particulier nous avons $\mathscr{A}_t = A_0$. En outre, en notant $F_t = \{x_t \mid x \in F\}$, il vient clairement $\operatorname{rang} F_t \le \operatorname{rang} F$ (si $\sum n_i x_i = 0$ avec $n_i \in \mathbb{Z}$ et $x_i \in F$ alors $\sum n_i (x_i)_t = 0$). Choisissons un sous-groupe G de $A_0(\overline{\mathbb{Q}})$ contenant F_t tel que $\operatorname{rang} G = \operatorname{rang} F$. Nous savons par hypothèse que

$$\begin{aligned} S = \lambda_{\overline{\mathbb{Q}},\mathscr{X}_t}(F_t) \le \lambda_{\overline{\mathbb{Q}},\mathscr{X}_t}(G) &\le f(\dim \mathscr{X}_t, \deg \mathscr{X}_t, \operatorname{rang} G) \\ &= f(\dim X, \deg X, \operatorname{rang} F). \end{aligned}$$

Nous pouvons donc écrire

$$X_t \cap F_t \subset \bigcup_{i=1}^{S} (x_i)_t + B_i \subset \mathscr{X}_t$$

avec $(x_i)_t \in F_t$ et $\deg B_i \leq \Delta(\mathscr{X}_t) = \Delta(X)$. Nos réductions permettent alors d'affirmer

$$X \cap F \subset \bigcup_{i=1}^{S} x_i + B_i \subset X.$$

En effet : si $y \in X \cap F$ alors $y_t \in \mathscr{X}_t \cap F_t$ donc il existe i tel que $y_t - (x_i)_t \in B_i$; par (1) il est impossible que $y - x_i \notin B_i$; donc $y \in x_i + B_i$, ce qui donne la première inclusion ; pour la seconde si $x_i + B_i \not\subset X$ on aurait $(x_i)_t + B_i \not\subset \mathscr{X}_t$ par (2) ce qui est absurde. Nous venons de montrer $\lambda_{K,X}(F) \leq f(\dim X, \deg X, \operatorname{rang} F)$ et ceci termine la démonstration. □

Bibliographie

[Aliev et Smyth 2008] I. Aliev et C. Smyth, "Power maps and subvarieties of the complex algebraic n-torus", prépublication, 2008. arXiv 0802.2938

[Buium 1993] A. Buium, "Effective bound for the geometric Lang conjecture", *Duke Math. J.* **71**:2 (1993), 475–499. MR 95c:14055 Zbl 0812.14029

[Crittenden et Vanden Eynden 1970] R. B. Crittenden et C. L. Vanden Eynden, "Any n arithmetic progressions covering the first 2^n integers cover all integers", *Proc. Amer. Math. Soc.* **24**:3 (1970), 475–481. MR 41 #3365 Zbl 0192.39001

[Evertse et al. 2002] J.-H. Evertse, H. P. Schlickewei et W. M. Schmidt, "Linear equations in variables which lie in a multiplicative group", *Ann. of Math.* (2) **155**:3 (2002), 807–836. MR 2003f:11037 Zbl 1026.11038

[Ghioca et Tucker 2009] D. Ghioca et T. J. Tucker, "Periodic points, linearizing maps, and the dynamical Mordell–Lang problem", *J. Number Theory* **129**:6 (2009), 1392–1403. MR 2010i:37219 Zbl 1186.14047

[Rémond 2000] G. Rémond, "Décompte dans une conjecture de Lang", *Invent. Math.* **142**:3 (2000), 513–545. MR 2002f:14058 Zbl 0972.11054

[Rémond 2002] G. Rémond, "Sur les sous-variétés des tores", *Compositio Math.* **134**:3 (2002), 337–366. MR 2003m:11093 Zbl 1101.14030

[Rémond 2003] G. Rémond, "Approximation Diophantienne sur les variétés semi-abéliennes", *Ann. Sci. École Norm. Sup.* (4) **36**:2 (2003), 191–212. MR 2004d:11064 Zbl 1081.11053

[Rémond 2005] G. Rémond, "Borne générique pour le problème de Mordell–Lang", *Manuscripta Math.* **118**:1 (2005), 85–97. MR 2006f:14023 Zbl 1136.11039

Received January 11, 2011.

GAËL RÉMOND
INSTITUT FOURIER, UMR 5582
BOÎTE POSTALE 74
38402 SAINT-MARTIN-D'HÈRES CEDEX
FRANCE

Gael.Remond@ujf-grenoble.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

FOURIER TRANSFORMS OF SEMISIMPLE ORBITAL INTEGRALS ON THE LIE ALGEBRA OF SL_2

LOREN SPICE

The Harish-Chandra–Howe local character expansion expresses the characters of reductive, p-adic groups in terms of Fourier transforms of nilpotent orbital integrals on their Lie algebras, and Murnaghan–Kirillov theory expresses many characters of reductive, p-adic groups in terms of Fourier transforms of semisimple orbital integrals (also on their Lie algebras). In many cases, the evaluation of these Fourier transforms seems intractable, but for SL_2, the nilpotent orbital integrals have already been computed. We compute Fourier transforms of semisimple orbital integrals using a variant of Huntsinger's integral formula and the theory of p-adic special functions.

1. Introduction

1A. ***History.*** Harish-Chandra's p-adic Lefschetz principle suggests that results in real harmonic analysis should have analogues in p-adic harmonic analysis. This principle has had too many successes to list, but it is interesting that the paths to

The author was partially supported by NSF award DMS-0854897.
MSC2010: 22E35, 22E50.
Keywords: p-adic, orbital integral, special function.

results in the Archimedean and non-Archimedean settings are often different. One striking manifestation of this is that the characters for the discrete series of real groups were found *before* the representations to which they were associated were constructed (see [Harish-Chandra 1966, Theorem 16; Schmid 1968, Theorem 4]), whereas, in the p-adic setting, we now have explicit constructions of many representations (see [Howe 1971; Corwin 1989; Moy 1986; Morris 1992; Bushnell and Kutzko 1993a; 1993b; 1994; Moy and Prasad 1994; Adler 1998; Yu 2001; Stevens 2008], among many others), but explicit character tables are still very rare.

This scarcity is of particular concern because, as suggested by Sally, it should be the case that "characters tell all" [Sally and Spice 2009, p. 104]. Note, for example, the recent work of Langlands [2011], which uses in a crucial way (see [ibid., Section 1.d]) the character formulae of [Sally and Shalika 1968] to show the existence of a transfer map dual to the transfer of stable characters, but only for SL_2. It seems likely that one of the main obstacles to extending the results of [Langlands 2011] to other groups is the absence of explicit character formulae for them.

The good news here is that much *is* known about the behaviour of characters in general. For example, the Harish-Chandra–Howe local character expansion [Howe 1973; Harish-Chandra 1999; DeBacker 2002] and Murnaghan–Kirillov theory [Murnaghan 1995a; 1995b; 1996a; 1996b; 2000; Kim and Murnaghan 2003; 2006] give information about the asymptotics (near the identity element) of characters of p-adic groups in terms of Fourier transforms of orbital integrals (nilpotent or semisimple) on the Lie algebra, and many existing character formulae are stated in terms of such orbital integrals; see, for example, [DeBacker 1997, Theorem 5.3.2; Spice 2005, Theorems 6.6 and 7.18; Adler and Spice 2009, Theorem 7.1; DeBacker and Reeder 2009, Lemma 10.0.4]. See also [Adler and Spice 2009, Section 0.1] for a more exhaustive description of what is known in the supercuspidal case.

The bad news is that many applications require completely explicit character tables — in particular, the evaluation of Fourier transforms of orbital integrals when they appear — but Hales [1994] has shown that the orbital integrals may themselves be "nonelementary". This term has a technical meaning, but, for our purposes, it suffices to regard it informally as meaning "difficult to evaluate". (Note, though, that the asymptotic behaviour of orbital integrals "near ∞" is understood in all cases; see [Waldspurger 1995, Proposition VIII.1].) Since SL_2 is both simple enough for many explicit computations to be tractable (for example, the Fourier transforms of nilpotent orbital integrals have already been computed [DeBacker and Sally 2000, Appendix A.3–A.4]), and complicated enough for interesting phenomena to be apparent (for example, unlike GL_2 and PGL_2, it admits nonstable characters), it is a natural focus for our investigations.

Another perspective on the behaviour of characters in the range where Murnaghan–Kirillov theory holds is offered in [Corwin et al. 1995, Theorem 4.2(d); Takahashi

2003, Proposition 2.9(2); 2005, Theorem 2.5], where explicit mention of orbital integrals is replaced (on the "bad shell" — see Section 10B) by arithmetically interesting sums, identified in [Takahashi 2003; 2005] as Kloosterman sums. In fact, exponential sums — specifically, Gauss sums — have long been observed in p-adic harmonic analysis; see, for example, [Corwin et al. 1995, Proposition 3.7; Waldspurger 1995, Section VIII.1; DeBacker 1997, p. 55; Shalika 2004, Section 1.3; Adler and Spice 2009, Section 5.2].

The work recorded here was carried out while preparing [Adler et al. 2011], which provides a proof of the aforementioned SL_2 character formulae [Sally and Shalika 1968] by specialising the results of [Adler and Spice 2009; DeBacker and Reeder 2009]. As discussed above, these general results are stated in terms of Fourier transforms of orbital integrals (see Definition 5.5); so, in order to obtain completely explicit formulae, it was necessary to evaluate those Fourier transforms. The author of the present paper was surprised to discover that this latter evaluation reduced to the computation of *Bessel functions* (see Section 7 and Proposition 8.11). In retrospect, by the p-adic Lefschetz principle mentioned on the first page, it seems natural that the "special functions" described in [Sally and Taibleson 1966] will play some important role in p-adic harmonic analysis, since their classical analogues are so integral to real harmonic analysis (see, for just one example, [Gindikin and Karpelevič 1962, Theorem 2], where Harish-Chandra's $\boldsymbol{c}$-function is calculated in terms of Γ-functions). Relationships between a different sort of Bessel function and a different sort of orbital integral (adapted to the Jacquet–Ye relative trace formula) have already been demonstrated by Baruch [1997; 2001; 2003; 2004; 2005]. We will investigate further applications of complex-valued p-adic special functions in future work.

See also [Cunningham and Gordon 2009, Section 4] for a motivic approach to the calculation of Fourier transforms of semisimple orbital integrals.

1B. ***Outline of the paper.*** We need a lot of notation in order to be completely explicit; we describe it in Sections 2–7. Specifically, Sections 2–4 describe the basic notation for working with groups over p-adic fields, adapted to the particular setting of the group SL_2. Since our formulae will be written "torus-by-torus" (à la [Harish-Chandra 1970, Theorem 12]), we need to describe the tori in SL_2. This can be done very concretely; see Definition 4.1.

In Section 5, we define the functions $\hat{\mu}^G_{X^*}$ (Fourier transforms of orbital integrals) that we want to compute as representing functions for certain invariant distributions on $\mathfrak{sl}_2$ (see Definition 5.5 and Notation 5.7). Since these functions are defined only up to scalar multiples, it is important to be aware of the normalisations involved in their construction. We specify the (Haar) measures that we are using in Definition 2.1 and Proposition 11.2.

As mentioned in Section 1A, p-adic harmonic analysis tends to involve Gauss sums and other fourth roots of unity, and our calculations are no exception; we define and compare some of the relevant constants in Section 6. Finally, with these ingredients in place, we can follow [Sally and Taibleson 1966] in defining the Bessel functions that we will use to evaluate $\hat{\mu}^G_{X^*}$. Already, [Sally and Taibleson 1966] offers considerable information about the values of these functions, but we need to carry the calculations further, especially far from the identity (see Proposition 7.5) and on the "bad shell" (see Proposition 7.7), where (twisted) Kloosterman sums make an appearance.

In Section 8, we define a function $M^G_{X^*}$ (see Definition 8.4), which we will spend most of the rest of the paper computing. This is a reasonable focus because, once the computations are completed, Proposition 11.2 will show that we have actually been computing $\hat{\mu}^G_{X^*}$. The definition of $M^G_{X^*}$ involves a rather remarkable function φ_θ (see Definition 8.2 and Lemma 8.3); it seems likely that generalising our techniques will require understanding the proper replacement for φ_θ.

Proposition 8.11 describes $M^G_{X^*}$ in terms of Bessel functions, and Proposition 8.13 uses Theorem 7.4 to describe their behaviour near 0.

We now proceed according to the "type" of X^* (as in Definition 4.4). The calculations when X^* is split, and when it is unramified, are quite similar; we combine them in Section 9. We split into cases depending on whether the argument to $M^G_{X^*}$ is far from (as in Section 9A) or close to (as in Section 9B) zero; there are qualitative differences in the behaviour, as can be seen by comparing, for example, Theorems 9.5 and 9.7. When X^* is ramified, it turns out that, in addition to the behaviour far from (as in Section 10A) and close to (as in Section 10C) zero, there is a third range of interest in the middle. This is the so called "bad shell" (see Section 10B), and it seems likely that the particularly complicated nature of the formulae here is a reflection of the "nonelementary" behaviour of orbital integrals (hence, by Murnaghan–Kirillov theory, also of characters) described in [Hales 1994].

Finally, we show in Section 11 that the function we have been evaluating actually does represent the desired distribution, that is, equals $\hat{\mu}^G_{X^*}$. (See Proposition 11.2.) We close with some observations (see Theorem 11.3) about the qualitative behaviour of orbital integrals that does not depend (much) on the "type" of X^*.

2. Notation

Suppose that k is a nondiscrete, non-Archimedean local field. We do not make any assumptions on its characteristic, but we assume that its residual characteristic p is not 2. (We occasionally cite [Shalika 2004], which works only with characteristic-0 fields, but we shall not use any results from there that require this restriction.) Let

R denote the ring of integers in k, $\wp$ the prime ideal of R, and ord the valuation on k with value group $\mathbb{Z}$.

Let $\mathfrak{f}$ denote the residue field $R/\wp$ of k. Write $q = |\mathfrak{f}|$ for the number of elements in $\mathfrak{f}$, and put $|x| = q^{-\operatorname{ord}(x)}$ for $x \in k$. If $\alpha \in \mathbb{C}$, then write ν^α for the (multiplicative) character $x \mapsto |x|^\alpha$ of $k^\times$.

Put $\boldsymbol{G} = \mathrm{SL}_2$ and $G = \boldsymbol{G}(k)$, and let $\mathfrak{g}$ and $\mathfrak{g}^*$ denote the Lie algebra and dual Lie algebra of G, respectively.

It is important for our calculations to be quite specific about the Haar measures that we are using. For convenience, we fix the ones used in [Sally and Taibleson 1966, p. 280].

Definition 2.1. Throughout, we shall use the (additive) Haar measure $\mathrm{d}x$ on k that assigns measure 1 to R, and the associated (multiplicative) Haar measure $\mathrm{d}^\times x = |x|^{-1}\,\mathrm{d}x$ on $k^\times$ that assigns measure $1 - q^{-1}$ to $R^\times$. When convenient, we shall write $\mathrm{d}t$ instead of $\mathrm{d}x$.

Definition 2.2. If Φ is an (additive) character of k, then define $\Phi_b : x \mapsto \Phi(bx)$ for $b \in k$. The *depth* of Φ is

$$\mathrm{d}(\Phi) := \begin{cases} \min\{i \in \mathbb{Z} : \Phi \text{ is trivial on } \wp^{i+1}\} & \text{if } \Phi \text{ is nontrivial,} \\ -\infty & \text{otherwise.} \end{cases}$$

The depth of a character is related to what is often called its *conductor* by $\mathrm{d}(\Phi) = \omega(\Phi) - 1$ (in the notation of [Shalika 2004, Section 1.3]). We have that

$$\mathrm{d}(\Phi_b) = \mathrm{d}(\Phi) - \operatorname{ord}(b). \tag{2.3}$$

The notion of depth and the symbol d will be used in multiple contexts (see Definition 4.9); we rely on the context to disambiguate them.

Notation 2.4. Φ is a nontrivial (additive) character of k.

One of the crucial tools of Harish-Chandra's approach to harmonic analysis is the reduction, whenever possible, of questions about a group to questions about its Lie algebra. The exponential map often allows one to effect this reduction, but, since it might converge only in a very small neighbourhood of 0, we replace it with a "mock-exponential map" (see [Adler 1998, Section 1.5]) which has many of the same properties (see Lemma 2.6).

Definition 2.5. The *Cayley map* $\mathrm{c} : k \setminus \{1\} \to k \setminus \{-1\}$ is defined by

$$\mathrm{c}(X) = (1 + X)(1 - X)^{-1} \quad \text{for } X \in k \setminus \{1\}.$$

The Cayley function is available in many settings; we are using it only as a function defined almost everywhere on k.

Lemma 2.6.

- *The map* c *is a bijection.*
- $\mathrm{c}(-X) = \mathrm{c}(X)^{-1} = \mathrm{c}^{-1}(X)$ *for* $X \in k \setminus \{\pm 1\}$.
- *The map* c *carries* $\wp^i$ *to* $1 + \wp^i$ *for all* $i \in \mathbb{Z}_{>0}$.
- *In the notation of Definition 2.1, the pullback along* c *of the measure* $\mathrm{d}^\times x$ *on* $1 + \wp$ *is the measure* $\mathrm{d}x$ *on* $\wp$.
- *If* $X \in \wp^i$ *and* $Y \in \wp^j$, *with* $i, j \in \mathbb{Z}_{>0}$, *then*

$$\mathrm{c}(X + Y) \equiv \mathrm{c}(X) + 2Y \pmod{1 + \wp^n},$$

where $n = j + \min\{2i, j\}$.

Proof. It is easy to check that $x \mapsto (1 - x)(1 + x)^{-1}$ is inverse to c and satisfies the desired equalities and check that $\mathrm{c}(\wp^i) \subseteq 1 + \wp^i$ and $\mathrm{c}^{-1}(1 + \wp^i) \subseteq \wp^i$. If $f \in C^\infty(1 + \wp)$, then there is some $i \in \mathbb{Z}_{>0}$ such that $f \in C(1 + \wp/1 + \wp^i)$. Since $\mathrm{meas}_{\mathrm{d}x}(\wp^i) = q^{-i} = \mathrm{meas}_{\mathrm{d}^\times x}(1 + \wp^i)$, we see that

$$\begin{aligned}\int_{1+\wp} f(x)\, \mathrm{d}^\times x &= \sum_{x \in 1+\wp/1+\wp^i} f(x)\, \mathrm{meas}_{\mathrm{d}^\times x}(1 + \wp^i) \\ &= \sum_{x \in \wp/\wp^i} (f \circ \mathrm{c})(x) q^{-i}\, \mathrm{meas}_{\mathrm{d}x}(\wp^i) = \int_{\wp} (f \circ \mathrm{c})(x)\, \mathrm{d}x.\end{aligned}$$

Finally, under the stated conditions on X and Y,

$$\begin{aligned}\big(\mathrm{c}(X) + 2Y\big)\big(1 - (X + Y)\big) &= \mathrm{c}(X) \cdot (1 - X) + Y\big(2(1 - (X + Y)) - \mathrm{c}(X)\big) \\ &= (1 + X + Y) + Y\big((1 - 2X - \mathrm{c}(X)) - 2Y\big).\end{aligned}$$

Since $\mathrm{c}(X) = 1 + 2X(1 - X)^{-1}$, we have that $1 - 2X - \mathrm{c}(X) \in \wp^{2i}$. The result follows. □

3. Fields and algebras

Definition 3.1. For $\theta \in k^\times$, write k_θ for the k-algebra that is $k \oplus k$ (as a vector space), equipped with multiplication $(a, b) \cdot (c, d) = (ac + bd\theta, ad + bc)$. Write $\sqrt{\theta}$ for the element $(0, 1) \in k_\theta$, so that $(a, b) = a + b\sqrt{\theta}$.

We also use the notation $\sqrt{\theta}$ for a matrix (see Definition 4.1); we shall rely on context to make the meaning clear.

If $\theta \notin (k^\times)^2$, then k_θ is isomorphic to $k(\sqrt{\theta})$ (as k-algebras) via the map $(a, b) \mapsto a + b\sqrt{\theta}$, and we shall not distinguish between them.

If $\theta = x^2$, with $x \in k$, then k_θ is isomorphic to $k \oplus k$ (as k-algebras) via the map $(a, b) \mapsto (a + bx, a - bx)$.

Definition 3.2. Define

$$\begin{aligned} N_\theta(a+b\sqrt{\theta}) &= a^2 - b^2\theta, \quad \mathrm{tr}_\theta(a+b\sqrt{\theta}) = 2a, \\ \mathrm{Re}_\theta(a+b\sqrt{\theta}) &= a, \qquad\qquad \mathrm{Im}_\theta(a+b\sqrt{\theta}) = b, \\ \mathrm{ord}_\theta(a+b\sqrt{\theta}) &= \tfrac{1}{2}\,\mathrm{ord}\big(N_\theta(a+b\sqrt{\theta})\big) \end{aligned}$$

for $a+b\sqrt{\theta} \in k_\theta$. Write $C_\theta = \ker N_\theta$ and $V_\theta = \ker \mathrm{tr}_\theta$, and let sgn_θ be the unique (multiplicative) character of $k^\times$ with kernel precisely $N_\theta(k_\theta^\times)$.

If $\theta \notin (k^\times)^2$, then N_θ and tr_θ are the usual norm and trace maps associated to the quadratic extension of fields k_θ/k, and ord_θ is the valuation on k_θ extending ord. In any case, $k_\theta^\times = \{z \in k_\theta : N_\theta(z) \neq 0\}$.

We can describe the signum character explicitly by

$$\mathrm{sgn}_\theta(x) = \begin{cases} 1 & \theta \text{ split}, \\ (-1)^{\mathrm{ord}(x)} & \theta \text{ unramified}, \end{cases} \tag{3.3}$$

$$\begin{cases} \mathrm{sgn}_\theta(\theta) = \mathrm{sgn}_{\mathfrak{f}}(-1) \\ \mathrm{sgn}_\theta(x) = \mathrm{sgn}_{\mathfrak{f}}(\bar{x}) \quad \text{for } x \in R^\times, \end{cases} \tag{3.4}$$

where $\mathrm{sgn}_{\mathfrak{f}}$ is the quadratic character of $\mathfrak{f}^\times$ and $x \mapsto \bar{x}$ the reduction map $R \to \mathfrak{f}$.

4. Tori and filtrations

We begin by defining a few model tori.

Definition 4.1. For $\theta \in k$, put

$$\boldsymbol{T}_\theta = \left\{ \begin{pmatrix} a & b \\ b\theta & a \end{pmatrix} : a^2 - b^2\theta = 1 \right\}.$$

Then

$$\mathfrak{t}_\theta := \mathrm{Lie}(\boldsymbol{T}_\theta) = \left\{ \begin{pmatrix} 0 & b \\ b\theta & 0 \end{pmatrix} \right\}.$$

Write $\sqrt{\theta}$ for the element

$$\begin{pmatrix} 0 & 1 \\ \theta & 0 \end{pmatrix}$$

so that $\mathfrak{t}_\theta = \mathrm{Span}_k \sqrt{\theta}$. Call a maximal k-torus in $\boldsymbol{G}$ *standard* exactly when it is of the form $\boldsymbol{T}_\theta$ for some $\theta \in k$.

We also denote by $\sqrt{\theta}$ a specific element of an extension of k (see Definition 4.1); we shall rely on context to make the meaning clear.

Remark 4.2. The group T_θ is isomorphic to $C_\theta = \ker N_\theta$ and the Lie algebra $\mathfrak{t}_\theta$ to $V_\theta = \ker \mathrm{tr}_\theta$ in each case via the map

$$\begin{pmatrix} a & b \\ b\theta & a \end{pmatrix} \mapsto (a, b).$$

We shall use the terms "split", "unramified", and "ramified" in many different contexts.

Remark 4.3. If $\boldsymbol{T}$ is a maximal k-torus in $\boldsymbol{G}$ and $\mathfrak{t} = \mathrm{Lie}(T)$, then we shall identify $\mathfrak{t}$ (respectively, $\mathfrak{t}^*$) with the spaces of fixed points for the adjoint (respectively, coadjoint) action on $\mathfrak{g}$ (respectively, $\mathfrak{g}^*$). By abuse of language, we shall sometimes say that $X^* \in \mathfrak{g}^*$ or $Y \in \mathfrak{g}$ lies in, or belongs to, the torus $\boldsymbol{T}$ to mean that $X^* \in \mathfrak{t}^*$ and $Y \in \mathfrak{t}$; equivalently, that $C_{\boldsymbol{G}}(X^*) = \boldsymbol{T} = C_{\boldsymbol{G}}(Y)$. In particular, "$X^*$ and Y belong to a common torus" is shorthand for "$C_{\boldsymbol{G}}(X^*) = C_{\boldsymbol{G}}(Y)$".

Definition 4.4. A maximal k-torus in $\boldsymbol{G}$ is called (un)ramified if it is elliptic and splits over an (un)ramified extension of k. An element $\theta \in k$ is called split, unramified, or ramified if $\boldsymbol{T}_\theta$ has that property. A regular, semisimple element of $\mathfrak{g}$ or $\mathfrak{g}^*$ is called split, unramified, or ramified if the torus to which it belongs has that property.

Remark 4.5. To be explicit, squares in $k^\times$ are split, and a nonsquare $\theta \in k$ is unramified or ramified if $\max\{\mathrm{ord}(x^2\theta) : x \in k\}$ is even or odd, respectively.

Notation 4.6. If $\boldsymbol{T}$ is a maximal k-torus in $\boldsymbol{G}$ with $T = \boldsymbol{T}(k)$, then write $W(\boldsymbol{G}, \boldsymbol{T}) = N_{\boldsymbol{G}}(\boldsymbol{T})/\boldsymbol{T}$ for the absolute and $W(G, T) = N_G(T)/T$ for the relative Weyl group of $\boldsymbol{T}$ in $\boldsymbol{G}$.

Every maximal k-torus in $\boldsymbol{G}$ is G-conjugate to some $\boldsymbol{T}_\theta$. (See, for example, [DeBacker and Sally 2000, Section A.2].) In particular,

$$\mathrm{Int}\begin{pmatrix} 1 & 1 \\ -1/2 & 1/2 \end{pmatrix}\left\{\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} : ad = 1\right\} = \boldsymbol{T}_1.$$

Remark 4.7. For all $\theta \in k$, the group $W(\boldsymbol{G}, \boldsymbol{T}_\theta)$ has order 2, with the nontrivial element acting on $\boldsymbol{T}_\theta$ by inversion. If $\mathrm{sgn}_\theta(-1) = 1$ (in particular, if θ is split or unramified), say, with $N_\theta(a + b\sqrt{\theta}) = -1$, then $W(G, T_\theta)$ also has order 2, with the nontrivial element represented by

$$\begin{pmatrix} a & b \\ -b\theta & -a \end{pmatrix}.$$

If $\theta = 1$, then we may take $(a, b) = (0, 1)$ to recover the familiar Weyl group element. Otherwise (that is, if $\mathrm{sgn}_\theta(-1) = -1$), $W(G, T_\theta)$ is trivial.

The concept of *stable conjugacy* was introduced by Langlands [1979, pp. 2–3] as part of the foundation of the Langlands conjectures.

Definition 4.8. Two

- maximal k-tori $\boldsymbol{T}_i$ in $\boldsymbol{G}$,
- regular semisimple elements $X_i^* \in \mathfrak{g}^*$, or
- regular semisimple elements $Y_i \in \mathfrak{g}$,

with $i = 1, 2$, are called *stably conjugate* exactly when there are a field extension E/k and an element $g \in \boldsymbol{G}(E)$ such that

- $\mathrm{Int}(g)T_1 = T_2$,
- $\mathrm{Ad}^*(g)X_1^* = X_2^*$, or
- $\mathrm{Ad}(g)X_1 = X_2$,

where $T_i = \boldsymbol{T}_i(k)$ for $i = 1, 2$. If the conjugacy can be carried out without passing to an extension field (that is, if we may take $g \in G$), then we will sometimes emphasise this by saying that the tori or elements are *rationally conjugate*.

The Zariski-density of T_i in $\boldsymbol{T}_i$ implies that $\mathrm{Int}(g)\boldsymbol{T}_1 = \boldsymbol{T}_2$, but that this is a strictly weaker condition; indeed, given *any* two maximal tori, there is an element g, defined over some extension field of k, satisfying this condition. In our special case (of $\boldsymbol{G} = \mathrm{SL}_2$), two tori or elements are stably conjugate if and only if they are conjugate in $\mathrm{GL}_2(k)$.

More concretely, two tori $\boldsymbol{T}_\theta$ and $\boldsymbol{T}_{\theta'}$ are stably conjugate if and only if $\theta \equiv \theta' \pmod{(k^\times)^2}$. The stable conjugacy class of the split torus $\boldsymbol{T}_1$ is also a rational conjugacy class.

Suppose that ϵ is an unramified and ϖ a ramified, nonsquare. Then the stable conjugacy class of $\boldsymbol{T}_\epsilon$ splits into 2 rational conjugacy classes, represented by $\boldsymbol{T}_\epsilon$ and $\boldsymbol{T}_{\varpi^2\epsilon}$. The stable conjugacy class of $\boldsymbol{T}_\varpi$ is also a rational conjugacy class if $\mathrm{sgn}_\varpi(-1) = -1$, but it splits into 2 rational conjugacy classes, represented by $\boldsymbol{T}_\varpi$ and $\boldsymbol{T}_{\epsilon^2\varpi}$, if $\mathrm{sgn}_\varpi(-1) = 1$.

We also need filtrations on the Lie algebra and dual Lie algebra of a torus. These definitions are standard (see, for example, [Adler 1998, Section 1.4]) and can be made in far more generality (see [Moy and Prasad 1994, Section 3; 1996, Section 3.3]); we give only simple definitions adapted to $\boldsymbol{G} = \mathrm{SL}_2$.

Definition 4.9. Let $\boldsymbol{T}$ be a maximal k-torus in $\boldsymbol{G}$. Recall that $\boldsymbol{T}$ is G-conjugate to $\boldsymbol{T}_\theta$ for some $\theta \in k$, so that $\mathfrak{t} = \mathrm{Lie}(T)$ is isomorphic to $V_\theta = \ker \mathrm{tr}_\theta \subseteq k_\theta$. For $r \in \mathbb{R}$, write $\mathfrak{t}_r$ for the preimage of $\{Y \in V_\theta : \mathrm{ord}_\theta(Y) \geq r\}$ and $\mathfrak{t}_{r+}$ for the preimage of $\{Y \in V_\theta : \mathrm{ord}_\theta(Y) > r\}$; then write $\mathfrak{t}_r^* = \{X^* \in \mathfrak{t}^* : \Phi(\langle X^*, Y\rangle) = 1 \text{ for all } Y \in \mathfrak{t}_{(-r)+}\}$ (where Φ is the additive character of Notation 2.4).

If $X^* \in \mathfrak{t}^*$ and $Y \in \mathfrak{t}$, then define $\mathrm{d}(X^*) = \max\{r \in \mathbb{R} : X^* \in \mathfrak{t}_r^*\}$ and $\mathrm{d}(Y) = \max\{r \in \mathbb{R} : Y \in \mathfrak{t}_r\}$.

One can define a notion of depth in more generality (see, for example, [Adler and DeBacker 2002, Section 3.3 and Example 3.4.6; Kim and Murnaghan 2003, Section 2.1 and Lemma 2.1.5]), but we only need the special case above. (The only remaining case to consider for $\mathfrak{g} = \mathfrak{sl}_2(k)$ is the depth of a nilpotent element, which is ∞.)

5. Orbital integrals

Our goal in this paper is to compute Fourier transforms of regular, semisimple orbital integrals on $\mathfrak{g}$ (see Definition 5.5 below). Since the Fourier transforms of nilpotent orbital integrals were computed in [DeBacker and Sally 2000, Appendix A], this covers all Fourier transforms of orbital integrals on $\mathfrak{g}$ (for our particular case $\boldsymbol{G} = \mathrm{SL}_2$). The case of orbital integrals on G was discussed in [Sally and Shalika 1984], as the culmination of the series of papers that began with [Sally and Shalika 1968; 1969].

We begin by choosing a representative for the regular, semisimple orbit of interest. By Section 4, we may choose this representative in a standard torus (in the sense of Definition 4.1).

Notation 5.1. $\beta, \theta \in k^\times$, and $X^* = \beta \cdot \sqrt{\theta} \in \mathfrak{t}_\theta^*$.

Here, we are implicitly using the identification of $\mathfrak{t}_\theta$ with $\mathfrak{t}_\theta^*$ via the trace form; what we really mean is that $\langle X^*, Y\rangle = \operatorname{tr} \beta \cdot \sqrt{\theta} \cdot Y$ for $Y \in \mathfrak{t}_\theta$, where $\langle\,\cdot\,,\cdot\,\rangle$ is the usual pairing between $\mathfrak{t}_\theta^*$ and $\mathfrak{t}_\theta$.

As in Definition 2.2, we may define a new character Φ_β of k, which we use often in our calculations.

Notation 5.2. $-r = \mathrm{d}(X^*)$, $\Phi' = \Phi_\beta$, and $r' = \mathrm{d}(\Phi')$.

By Definition 4.9, $Y \mapsto \Phi(\langle X^*, Y\rangle)$ is trivial on $(\mathfrak{t}_\theta)_{r+}$, but not on $(\mathfrak{t}_\theta)_r$. Therefore, $r' = r + \frac{1}{2}\operatorname{ord}(\theta)$.

Since $C_G(X^*) = T_\theta$ is Abelian, it is unimodular, so there exists a measure on $G/C_G(X^*)$ invariant under the action of G by left translation.

Notation 5.3. Let $\mathrm{d}\dot{g}$ be a translation-invariant measure on $G/C_G(X^*)$.

Since the orbit, $\mathcal{O}_{X^*}^G$, of X^* under the coadjoint action of G is isomorphic as a G-set to $G/C_G(X^*)$, we could transport to it the measure on the latter space; but we do not find it convenient to do so.

Since X^* is semisimple, $\mathcal{O}_{X^*}^G$ is closed in $\mathfrak{g}^*$; see, for example, [Tauvel and Yu 2005, Proposition 34.3.2]. Therefore, the restriction to $\mathcal{O}_{X^*}^G$ of a locally constant, compactly supported function on $\mathfrak{g}^*$ remains locally constant and compactly supported, so that the following definition makes sense.

Definition 5.4. The orbital integral of X^* is the distribution $\mu_{X^*}^G$ on $\mathfrak{g}^*$ defined by

$$\mu_{X^*}^G(f^*) = \int_{G/C_G(X^*)} f^*(\mathrm{Ad}^*(g)X^*)\,\mathrm{d}\dot{g} \quad \text{for all } f^* \in C_c^\infty(\mathfrak{g}^*).$$

We are interested in the Fourier transform of $\mu_{X^*}^G$. The definition of the Fourier transform (of distributions or of functions) requires, in addition to a choice of additive character (see Notation 2.4), also a choice of Haar measure $\mathrm{d}Y$ on $\mathfrak{g}^*$; but we shall build this choice into our representing function (see Notation 5.7), so that it will not show up in our final answer.

Definition 5.5. The Fourier transform of the orbital integral of X^* is the distribution $\hat{\mu}^G_{X^*}$ on $\mathfrak{g}$ defined for all $f \in C_c^\infty(\mathfrak{g})$ by

$$\hat{\mu}^G_{X^*}(f) = \mu^G_{X^*}(\hat{f}),$$

where

$$\hat{f}(Y^*) = \int_{\mathfrak{g}} f(Y)\Phi(\langle Y^*, Y\rangle)\,\mathrm{d}Y \quad \text{for all } Y^* \in \mathfrak{g}^*.$$

It is a result of Harish-Chandra [1999, Theorem 1.1] that $\hat{\mu}^G_{X^*}$ is *representable* on $\mathfrak{g}$, that is, there exists a locally integrable function F on $\mathfrak{g}$ such that

$$\hat{\mu}^G_{X^*}(f) = \int_G f(Y)F(Y)\,\mathrm{d}Y \quad \text{for all } f \in C_c^\infty(\mathfrak{g}).$$

One can say more about the behaviour and asymptotics of the function F. For example, it blows up as Y approaches 0, but its blow-up is controlled by a power of a discriminant function.

Definition 5.6. The *Weyl discriminant* on $\mathfrak{g}$ is the function $D_{\mathfrak{g}} : \mathfrak{g} \to \mathbb{C}$ such that, for all $Y \in \mathfrak{g}$, $D_{\mathfrak{g}}(Y)$ is the coefficient of the degree-1 term in the characteristic polynomial of $\operatorname{ad}(Y)$. Concretely,

$$D_{\mathfrak{g}}\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = 4(a^2 + bc).$$

Our main interest, however, is in the restriction of the function F above to the set $\mathfrak{g}^{\text{rss}}$ of regular, semisimple elements, where it is locally constant.

Notation 5.7. By abuse of notation, write again $\hat{\mu}^G_{X^*}$ for the function that represents the restriction to $\mathfrak{g}^{\text{rss}}$ of $\hat{\mu}^G_{X^*}$.

When we refer to the computation of the Fourier transform of an orbital integral, it is actually the (scalar) function of Notation 5.7 that we are trying to compute. The main tool in this direction is a general integral formula of Huntsinger (see [Adler and DeBacker 2004, Theorem A.1.2]), but we find it easier to evaluate an integral adapted to our current setting (see Definition 8.4). The computation of this integral will occupy most of the paper; we finally prove it actually represents the distribution $\hat{\mu}^G_{X^*}$ in Proposition 11.2.

Finally, we fix an element at which to evaluate the functions of interest. Since $\hat{\mu}^G_{X^*}$, as just defined, and $M^G_{X^*}$ in Definition 8.4 are G-invariant functions on $\mathfrak{g}^{\text{rss}}$, we may again consider only elements of standard tori.

Notation 5.8. $s, \theta' \in k^\times$, and $Y = s \cdot \sqrt{\theta'} \in \mathfrak{t}_{\theta'}$.

We phrase our computations in terms of the values of two "basic" functions at Y.

Lemma 5.9. $\mathrm{d}(Y) = \frac{1}{2}\operatorname{ord}(s^2\theta')$ *and* $D_{\mathfrak{g}}(Y) = 4s^2\theta'$.

Proof. This is a straightforward consequence of Definitions 4.9 and 5.6. □

6. Roots of unity and other constants

The computation of Fourier transforms of orbital integrals on $\mathfrak{g}$ via Murnaghan–Kirillov theory [Murnaghan 1995a; Kim and Murnaghan 2003; 2006; Adler and DeBacker 2004; Adler and Spice 2009] and also of the values near the identity of characters of G (see [Sally and Shalika 1968; Adler et al. 2011]) involves a somewhat bewildering array of 4-th roots of unity, for each of which there is a variety of notation available. All of these can be expressed in terms of a single "basic" quantity, the Gauss sum, denoted by $G(\Phi)$ in [Shalika 2004, Lemma 1.3.2]. The definition there implicitly depends on a choice of uniformiser, denoted there by π. Although the choice is arbitrary, for later convenience we denote it by $-\varpi$. Recall from Notation 2.4 that Φ is a nontrivial (additive) character of k.

Definition 6.1. If ϖ is a uniformiser of k, then

$$G_{\varpi}(\Phi) := q^{-1/2} \sum_{X \in R/\wp} \Phi_{(-\varpi)^{\mathrm{d}(\Phi)}}(X^2).$$

It is possible to compute these values exactly (see, for example, [Lidl and Niederreiter 1997, Theorem 5.15]), but we only require a few transformation laws.

Lemma 6.2. *If ϖ is a uniformiser of k, then*

$$\begin{aligned} G_{b\varpi}(\Phi) &= \operatorname{sgn}_{\varpi}(b)^{\mathrm{d}(\Phi)} G_{\varpi}(\Phi) \quad \textit{for } b \in R^\times, \\ G_{\varpi}(\Phi_b) &= \operatorname{sgn}_{\varpi}(b) G_{\varpi}(\Phi) \qquad \textit{for } b \in k^\times, \\ G_{\varpi}(\Phi)^2 &= \operatorname{sgn}_{\varpi}(-1), \\ G_{\varpi}(\Phi) &= q^{-1/2} \operatorname{sgn}_{\varpi}(-1)^{\mathrm{d}(\Phi)} \textstyle\sum_{X \in \mathfrak{f}^\times} \overline{\Phi}(X) \operatorname{sgn}_{\mathfrak{f}}(X), \end{aligned}$$

where $\operatorname{sgn}_{\mathfrak{f}}$ is the quadratic character of $\mathfrak{f}^\times$, and $\overline{\Phi}$ the (additive) character of $\mathfrak{f} = R/\wp$ arising from the restriction to R of the depth-0 character $\Phi_{\varpi^{\mathrm{d}(\Phi)}}$ of k.

Proof. Since $\sum_{X \in \mathfrak{f}} \overline{\Phi}(X) = 0$, we have that

$$\begin{aligned} \sum_{X \in \mathfrak{f}^\times} \overline{\Phi}(X) \operatorname{sgn}_{\mathfrak{f}}(X) &= \overline{\Phi}(0) + \sum_{X \in \mathfrak{f}^\times} \overline{\Phi}(X)\bigl(1 + \operatorname{sgn}_{\mathfrak{f}}(X)\bigr) \\ &= \overline{\Phi}(0) + 2 \sum_{X \in (\mathfrak{f}^\times)^2} \overline{\Phi}(X) = \sum_{X \in \mathfrak{f}} \overline{\Phi}(X^2) = q^{1/2} G_{\varpi}(\Phi_{(-1)^{\mathrm{d}(\Phi)}}). \end{aligned}$$

In other words,

$$G_{\varpi}(\Phi_{(-1)^{\mathrm{d}(\Phi)}}) = q^{-1/2} G(\operatorname{sgn}_{\mathfrak{f}}, \overline{\Phi}), \tag{$*$}$$

where the notation on the right is as in [Lidl and Niederreiter 1997, Section 5.2] (except that their ψ is our $\mathrm{sgn}_{\mathfrak{f}}$, the quadratic character of $\mathfrak{f}^\times$, and their χ is our $\overline{\Phi}$). The third equality, and the second equality for $b \in R^\times$, now follow from [ibid., Theorem 5.12]. The first equality follows from the second since $G_{b\varpi}(\Phi) = G_\varpi(\Phi_{b^{\mathrm{d}(\Phi)}})$. Taking $b = (-1)^{\mathrm{d}(\Phi)}$ and combining with $(*)$ gives the fourth equality. Finally, by definition, $G_\varpi(\Phi_{(-\varpi)^n}) = G_\varpi(\Phi) = \mathrm{sgn}_\varpi(-\varpi)^n G_\varpi(\Phi)$ for all $n \in \mathbb{Z}$. □

By Proposition 8.11 and Theorem 7.4, our calculations will involve the Γ-factors defined in [Sally and Taibleson 1966, Section 3]. The factor $\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi)$ is of particular interest. By [ibid., Theorem 3.1(iii)], $\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi)^2 = \mathrm{sgn}_\varpi(-1)$, so by Lemma 6.2, $\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi) = \pm G_\varpi(\Phi)$. It will be useful to identify the sign.

Lemma 6.3. *If ϖ is a uniformiser of k, then*

$$\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi) = \mathrm{sgn}_\varpi(-1)^{\mathrm{d}(\Phi)+1} G_\varpi(\Phi).$$

Proof. Write $\overline{\Phi} = \Phi_{\varpi^{\mathrm{d}(\Phi)}}$; this is a depth-0 character of k. The definitions of [Sally and Taibleson 1966] depend on a depth-(-1) additive character χ; we take it to be $\overline{\Phi}_\varpi$. The definition of $\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi)$ involves a principal-value integral (see Definition 8.4), but, as pointed out in the proof of [Sally and Taibleson 1966, Theorem 3.1], by [ibid., Lemma 3.1] and (3.4) it simplifies to

$$\begin{aligned}\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi) &= \int_{\mathrm{ord}(x)=-1} \overline{\Phi}_\varpi(x)|x|^{1/2}\,\mathrm{sgn}_\varpi(x)\,\mathrm{d}^\times x\\ &= \int_{R^\times} \overline{\Phi}_\varpi(\varpi^{-1}x)|\varpi^{-1}x|^{1/2}\,\mathrm{sgn}_\varpi(\varpi^{-1}x)\,\mathrm{d}^\times x\\ &= q^{1/2}\,\mathrm{sgn}_\varpi(-1)\,\mathrm{meas}_{\mathrm{d}^\times x}(1+\wp) \sum_{x\in R^\times/1+\wp} \overline{\Phi}(x)\,\mathrm{sgn}_{\mathfrak{f}}(x),\end{aligned}$$

where $\mathrm{d}^\times x$ is the Haar measure on $k^\times$ giving $R^\times$ measure $1-q^{-1}$ (see Definition 2.1). Since $\mathrm{meas}_{\mathrm{d}^\times x}(1+\wp) = q^{-1}$, the result now follows from Lemma 6.2. □

We will also need some constants associated to specific elements.

Waldspurger [1995, Proposition VIII.1] describes the "behaviour at ∞" of Fourier transforms of semisimple orbital integrals on general reductive, p-adic Lie algebras. His description involves a 4-th root of unity $\gamma_\psi(X^*, Y)$ (see [ibid., p. 79]); since his ψ is our Φ (Notation 2.4), we denote it by $\gamma_\Phi(X^*, Y)$. See Theorem 11.3 for our quantitative analogues (for the special case of $\mathfrak{sl}_2$) of his result.

We would like to avoid choosing "standard" representatives for $k^\times/(k^\times)^2$ (see Remark 6.9), but doing this is notationally unwieldy. Although our proofs will make use of these choices, none of the statements of the main results (except Theorems 10.8 and 10.9, via Remark 10.7) relies on them.

Notation 6.4. Let ϵ be a lift to $R^\times$ of a nonsquare in $\mathfrak{f}^\times$ and ϖ a uniformiser of k.

Definition 6.5. Recall Notations 5.1 and 5.8. If X^* and Y lie in stably conjugate tori, so that $\theta \equiv \theta' \pmod{(k^\times)^2}$, then

$$\gamma_\Phi(X^*, Y) = \begin{cases} 1 & \theta \equiv 1, \\ \gamma_{\text{un}}(s) & \theta \equiv \epsilon, \\ \gamma_{\text{ram}}(s) & \theta \equiv \varpi, \\ -\gamma_{\text{un}}(s)\gamma_{\text{ram}}(s) & \theta \equiv \epsilon\varpi, \end{cases}$$

where all congruences are taken modulo $(k^\times)^2$ and where

$$\gamma_{\text{un}}(s) := (-1)^{r'+1}\operatorname{sgn}_\epsilon(s) \quad \text{and} \quad \gamma_{\text{ram}}(s) := \operatorname{sgn}_\varpi(-s)G_\varpi(\Phi')$$

(with notation as in Notation 5.2 and Definition 6.1). It simplifies our notation considerably also to put $\gamma_\Phi(X^*, Y) = 1$ if X^* is elliptic and Y is split, and otherwise put $\gamma_\Phi(X^*, Y) = 0$ if X^* and Y do not lie in stably conjugate tori.

Remark 6.6. The dependence of $\gamma_\Phi(X^*, Y)$ on X^* is via r' and Φ' (see Notation 5.2). Expanding these definitions shows that $\gamma_\Phi(X^*, Y) = c_{\theta,\phi} \cdot \operatorname{sgn}_\theta(\beta s)$ when X^* and Y lie in stably conjugate tori, using Notations 5.1 and 5.8.

We have defined $\gamma_\Phi(X^*, Y)$ only when X^* and Y belong to (possibly different) standard tori, in the sense of Definition 4.1. A direct computation shows that, if we replace X^* or Y by a rational conjugate, or replace the pair (X^*, Y) by a stable conjugate, such that X^* and Y still lie in standard tori, then the constant $\gamma_\Phi(X^*, Y)$ does not change. (In the notation of Definition 8.2, $\operatorname{Ad}^*(g)X^*$ lies in a standard torus if and only if $\varphi_\theta(g) = (\alpha, 0)$, in which case $\operatorname{Ad}^*(g)X^* = \beta N_\theta(\alpha) \cdot \sqrt{N_\theta(\alpha)^{-2}\theta}$; and similarly for Y.) This allows us to define $\gamma_\Phi(X^*, Y)$ for all pairs of regular, semisimple elements, if desired.

By Lemma 6.2,

$$\gamma_{\text{ram}}(s)^2 = \operatorname{sgn}_\varpi(-1). \tag{6.7}$$

To make use of Propositions 7.5 and 7.7 below, we need the computation

$$\begin{aligned} \operatorname{sgn}_\varpi(v)G_\varpi(\Phi'_{\varpi^{r'+1}}) &= \operatorname{sgn}_\varpi(\varpi^{-(r'+1)}s\theta) \cdot \operatorname{sgn}_\varpi(\varpi^{r'+1})G_\varpi(\Phi') \\ &= \operatorname{sgn}_\varpi(-\theta)\gamma_{\text{ram}}(s). \end{aligned} \tag{6.8}$$

Remark 6.9. We will be interested exclusively in the case when $\theta \in \{1, \epsilon, \varpi\}$. This means we seem to be omitting the cases when $\theta \in \{\varpi^2\epsilon, \epsilon^2\varpi, \epsilon^{\pm 1}\varpi\}$, but, actually, this problem is not serious. Indeed, for $b \in k$, write

$$g_b := \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix} \in \mathrm{GL}_2(k).$$

Then

$$\operatorname{Ad}^*(g_b)X^* = \operatorname{Ad}^*(g_b)(\beta \cdot \sqrt{\theta}) = \beta b^{-1} \cdot \sqrt{b^2\theta}$$

(where we identify $\mathfrak{t}_\theta^*$ with $\mathfrak{t}_\theta$ via the trace pairing, as in Notation 5.1); and $\hat{\mu}^G_{X^*} = \hat{\mu}^G_{\mathrm{Ad}^*(g_b)X^*} \circ \mathrm{Ad}(g_b)$. This covers $\theta = \varpi^2\epsilon$ (by taking $b = \varpi^{-1}$) and $\theta = \epsilon^2\varpi$ (by taking $b = \epsilon^{-1}$). Handling $\theta \in \{\epsilon^{\pm 1}\varpi\}$ requires a different observation: since our choice of uniformiser was arbitrary, it could as well have been $\epsilon^{\pm 1}\varpi$ (or, for that matter, $\epsilon^2\varpi$) as ϖ itself. Thus, the formulae for the cases $\theta = \epsilon^n\varpi$ can be obtained by simple substitution.

The definition of $\gamma_\Phi(X^*, Y)$ when $\theta \equiv \epsilon\varpi \pmod{(k^\times)^2}$ is an instance of this; namely, by Lemma 6.2,

$$\begin{aligned}-\gamma_{\mathrm{un}}(s)\gamma_{\mathrm{ram}}(s) &= (-1)^{r'} \operatorname{sgn}_\epsilon(s) \cdot \operatorname{sgn}_\varpi(-s) G_\varpi(\Phi') \\ &= \operatorname{sgn}_{\epsilon\varpi}(-s) \cdot \operatorname{sgn}_\varpi(\epsilon)^{r'} G_\varpi(\Phi') = \operatorname{sgn}_{\epsilon\varpi}(-s) G_{\epsilon\varpi}(\Phi'),\end{aligned}$$

where we have used that $\operatorname{sgn}_\epsilon(-1) = 1$ and $\operatorname{sgn}_\varpi(\epsilon) = -1$.

We next define a constant $c_0(X^*)$ for use in Theorems 9.7 and 10.10. Those theorems (and Proposition 11.2) show that, as the notation suggests, it is the coefficient of the trivial orbit in the expansion of the germ of $\hat{\mu}^G_{X^*}$ in terms of Fourier transforms of nilpotent orbital integrals (see [Harish-Chandra 1999, Theorem 5.11]).

Definition 6.10.

$$c_0(X^*) = \begin{cases} -2q^{-1} & X^* \text{ split}, \\ -q^{-1} & X^* \text{ unramified}, \\ -\frac{1}{2}q^{-2}(q+1) & X^* \text{ ramified}. \end{cases}$$

Recall that $\hat{\mu}^G_{X^*}$ is defined in terms of the measure $\mathrm{d}\dot{g}$ of Proposition 11.2; in the notation of that proposition, whenever X^* is elliptic,

$$c_0(X^*) = (q-1)^{-1} \operatorname{meas}_{\mathrm{d}\dot{g}}(\dot{K}).$$

7. Bessel functions

Our strategy for computing Fourier transforms of orbital integrals is to reduce them to p-adic Bessel functions (see Proposition 8.11, (9.3), and (10.2)). In this context, we are referring to the complex-valued Bessel functions defined in [Sally and Taibleson 1966, Section 4], not the p-adic-valued ones defined in [Dwork 1974].

The definition of these functions depends on an additive character, denoted by χ in [Sally and Taibleson 1966], and a multiplicative character π of k. For internal consistency, we will instead denote the additive character by Φ and the multiplicative character by χ; but, for consistency with their work, we require throughout this section that $\mathrm{d}(\Phi) = -1$, that is, Φ is trivial on R but not on $\wp^{-1}$.

Definition 7.1 [Sally and Taibleson 1966, (4.1)]. For $\chi \in \widehat{k^\times}$, the *p-adic Bessel function of order* χ is given by

$$J_\chi(u, v) = \fint_{k^\times} \Phi(ux + vx^{-1})\chi(x)\,\mathrm{d}^\times x \quad \text{for } u, v \in k^\times,$$

where $\mathrm{d}^\times x$ is the Haar measure on $k^\times$ fixed in Definition 2.1. Also put $J_\chi^\theta = \frac{1}{2}(J_\chi + J_{\chi\,\mathrm{sgn}_\theta})$, with notation as in Definition 3.2.

The locally constant K-Bessel function $K(z|\chi)$ of [Trimble 1994, Definition 3.2] is $J_\chi(\varpi^t, \varpi^t)$ (in the notation of that definition), where ϖ is a uniformiser.

For $\chi \neq 1$, it is natural to extend the Bessel function by putting $J_\chi(u, 0) = \chi(u)^{-1}\Gamma(\chi)$ and $J_\chi(0, v) = \chi(v)\Gamma(\chi^{-1})$, where the Γ-factors are as in [Sally and Taibleson 1966, Section 3]. Under some conditions on χ, we can even define $J_\chi(0, 0)$ (either as 0 or the sum of a geometric series), but we do not need to do this.

The notation J_χ^θ arises naturally in our computations; see Proposition 8.11.

Definition 7.2. We say that a character $\chi \in \widehat{k^\times}$ is *mildly ramified* if χ is trivial on $1 + \wp$, but nontrivial on $k^\times$.

Since our orbital-integral calculations require information about J_χ only for χ mildly ramified, and since more precise information in that case is available in general, we focus our attention there.

Notation 7.3. Fix the following notation for the remainder of the section.

- $u, v \in k^\times$,
- $m = -\operatorname{ord}(uv)$, and
- $\chi \in \widehat{k^\times}$.

This is consistent with Notation 8.6. After Proposition 7.5, we will assume that χ is mildly ramified.

Of particular interest to us later will be the cases where χ is an unramified twist of one of the characters $\mathrm{sgn}_{\theta'}$ of Definition 3.2 (that is, is of the form $\nu^\alpha \,\mathrm{sgn}_\theta$ for some $\alpha \in \mathbb{C}$). Note that $\mathrm{sgn}_\epsilon = \nu^{\pi i/\ln(q)}$.

Theorem 7.4 [Sally and Taibleson 1966, Theorems 4.8 and 4.9].

$$J_\chi(u, v) = \begin{cases} \chi(v)\Gamma(\chi^{-1}) + \chi(u)^{-1}\Gamma(\chi) & m \le 1, \\ \chi(u)^{-1}F_\chi(m/2, uv) & m \ge 2 \textit{ and } m \textit{ even}, \\ 0 & m > 2 \textit{ and } m \textit{ odd}, \end{cases}$$

where the Γ*-factors are as in* [Sally and Taibleson 1966, Section 3], *and*

$$F_\chi(m/2, uv) := \int_{\operatorname{ord}(x) = -m/2} \Phi(x + uvx^{-1})\chi(x)\,\mathrm{d}^\times x.$$

The Γ-factor tables of [Sally and Taibleson 1966, Theorem 3.1], together with Lemma 6.3, mean that we understand $J_\chi(u, v)$ completely when $m < 2$, but further calculation is necessary in the remaining cases.

Proposition 7.5. *If*

- $h \in \mathbb{Z}_{>0}$,
- χ *is trivial on* $1 + \wp^h$, *and*
- $m \geq 4h - 1$,

then $J_\chi(uv) = 0$ *if* $uv \notin (k^\times)^2$; *and, if* $w \in k^\times$ *satisfies* $uv = w^2$, *then*

$$J_\chi(u, v) = q^{-m/4}\chi(u^{-1}w) \times \begin{cases} \Phi(2w) + \chi(-1)\Phi(-2w) & 4 \mid m, \\ \operatorname{sgn}_\varpi(w) G_\varpi(\Phi)\bigl(\Phi(2w) + (\chi\operatorname{sgn}_\varpi)(-1)\Phi(-2w)\bigr) & 4 \nmid m. \end{cases}$$

Proof. If m is odd, then the vanishing result follows from Theorem 7.4, so we assume that m is even. In this case, $m \geq 4h$, and, by Theorem 7.4, $J_\chi(u, v) = \chi(u)^{-1} F_\chi(m/2, uv)$.

We evaluate the integral defining $F_\chi(m/2, uv)$ by splitting it into pieces. Write

$$S_{uv} = \bigl\{x \in k : \operatorname{ord}(x) = -m/2 \text{ and } \operatorname{ord}(x - uvx^{-1}) < -m/2 + h\bigr\},$$
$$T_{uv} = \bigl\{x \in k : \operatorname{ord}(x) = -m/2 \text{ and } \operatorname{ord}(x - uvx^{-1}) \geq -m/2 + h\bigr\}.$$

Both S_{uv} and T_{uv} are invariant under multiplication by $1 + \wp$, and if $x \in T_{uv}$, then $uv \in x^2(1 + \wp^h) \subseteq (k^\times)^2$. We claim that the relevant integral may be taken over only T_{uv}.

If $X \in \wp^{m/2-h}$, then by Lemma 2.6 and the fact that $2(m/2 - h) \geq m/2$, we have

$$\mathrm{c}(X) \equiv 1 + 2X \pmod{\wp^{m/2}} \quad \text{and} \quad \mathrm{c}(X)^{-1} \equiv 1 - 2X \pmod{\wp^{m/2}},$$

so

$$\int_{S_{uv}} \Phi(x + uvx^{-1})\chi(x)\,\mathrm{d}^\times x$$
$$= (\star)\int_{\wp^{m/2-h}}\int_{S_{uv}} \Phi\bigl(x\cdot\mathrm{c}(X) + uvx^{-1}\cdot\mathrm{c}(X)^{-1}\bigr)\chi\bigl(x\cdot\mathrm{c}(X)\bigr)\,\mathrm{d}^\times x\,\mathrm{d}X$$
$$= (\star)\int_{S_{uv}} \Phi(x + uvx^{-1})\chi(x)\int_{\wp^{m/2-h}} \Phi_{2(x-uvx^{-1})}(X)\,\mathrm{d}X\,\mathrm{d}^\times x,$$

where $(\star) = \operatorname{meas}_{\mathrm{d}X}(\wp^{m/2-h})^{-1}$ is a constant and we used that Φ is trivial on $x\wp^{m/2} \cup uvx^{-1}\wp^{m/2} \subseteq R$ and χ is trivial on $\mathrm{c}(\wp^{m/2-h}) = 1 + \wp^{m/2-h} \subseteq 1 + \wp^h$. By (2.3), we have that $\mathrm{d}(\Phi_{2(x-uvx^{-1})}) > m/2 - h + 1$ (that is, $\Phi_{2(x-uvx^{-1})}$ is a nontrivial character on $\wp^{m/2-h}$) whenever $x \in S_{uv}$, so the inner integral is 0. This shows that, as desired, the integral defining $F_\chi(m/2, uv)$ may be taken over only T_{uv}.

If $uv \notin (k^\times)^2$, then $T_{uv} = \varnothing$, so $J_\chi(u, v) = \chi(u)^{-1} F_\chi(m/2, uv) = 0$; whereas, if $w \in k^\times$ satisfies $w^2 = uv$, then $T_{uv} = w(1 + \wp^h) \sqcup -w(1 + \wp^h)$, so

$$(*) \quad J_\chi(u, v) = \chi(u)^{-1}\biggl(\int_{w(1+\wp^h)} \Phi(x + uvx^{-1})\chi(x)\,\mathrm{d}^\times x + \int_{-w(1+\wp^h)} \Phi(x + uvx^{-1})\chi(x)\,\mathrm{d}^\times x\biggr).$$

Note that $\operatorname{ord}(w) = -m/2$.

We show a detailed calculation of the first integral; of course, that of the second is identical. Note that the integral no longer involves χ. By Lemma 2.6 again, $X \mapsto w \cdot \mathrm{c}(X)$ is a measure-preserving bijection from $\wp^h$ to $w(1 + \wp^h)$, so

$$\int_{w(1+\wp^h)} \Phi(x + uvx^{-1})\chi(x)\,\mathrm{d}^\times x = \chi(w) \int_{\wp^h} \Phi_w\bigl(\mathrm{c}(X) + \mathrm{c}(X)^{-1}\bigr)\,\mathrm{d}X,$$

where we used $uvw^{-1} = w$ and again that χ is trivial on $\mathrm{c}(\wp^h) = 1 + \wp^h$. We will evaluate the latter integral by breaking it into "shells" on which $\operatorname{ord}(X)$ is constant, using the following facts. By direct computation (and Definition 2.5),

$$\mathrm{c}(X) + \mathrm{c}(X)^{-1} = 2\mathrm{c}(X^2)$$

for $X \in k \setminus \{1\}$. If $\operatorname{ord}(X) = i$ and $\operatorname{ord}(Y) = j$, then Lemma 2.6 implies

$$\mathrm{c}\bigl((X + Y)^2\bigr) \equiv \mathrm{c}(X^2 + 2XY) \pmod{\wp^{2j}},$$
$$\mathrm{c}(X^2 + 2XY) \equiv \mathrm{c}(X^2) + 4XY \pmod{\wp^{2j}}.$$

(The second congruence could be made much finer, but we do not need this.)

In particular, fix $i \ge h$ with $2i < m/2 - 1$, so that $\mathrm{d}(\Phi) = m/2 - 1 < 2(m/2 - 1 - i)$ (that is, Φ is trivial on $\wp^{2(m/2-1-i)}$). Then

$$\begin{aligned}\int_{\operatorname{ord}(X)=i} &\Phi_w\bigl(\mathrm{c}(X) + \mathrm{c}(X)^{-1}\bigr)\,\mathrm{d}X \\ &= (\star) \int_{\wp^{m/2-1-i}} \int_{\operatorname{ord}(X)=i} (\Phi_{2w} \circ \mathrm{c})\bigl((X + Y)^2\bigr)\,\mathrm{d}X\,\mathrm{d}Y \\ &= (\star) \int_{\operatorname{ord}(X)=i} (\Phi_{2w} \circ \mathrm{c})(X^2) \int_{\wp^{m/2-1-i}} \Phi_{8wX}(Y)\,\mathrm{d}Y\,\mathrm{d}X,\end{aligned}$$

where $(\star) = \operatorname{meas}(\wp^{m/2-1-i})^{-1}$ is a constant. Since $\mathrm{d}(\Phi_{8wX}) = \mathrm{d}(\Phi_w) - \operatorname{ord}(8X) = m/2 - 1 - i$, the inner integral is 0.

Note that $\lceil (m/2 - 1)/2 \rceil \ge h$. Thus

$$J_\chi(u, v) = \int_{\wp^{\lceil (m/2-1)/2 \rceil}} (\Phi_{2w} \circ \mathrm{c})(X^2)\,\mathrm{d}X.$$

If $m/2$ is even, then the integral is over $\wp^{m/4}$, and $c(X^2) \equiv 1 \pmod{\wp^{m/2} \subseteq \ker \Phi_{2w}}$ for all $X \in \wp^{m/4}$. Thus, in that case,

$$J_\chi(u, v) = \mathrm{meas}_{\mathrm{d}X}(\wp^{m/4})\Phi_{2w}(1) = q^{-m/4}\Phi(2w).$$

If $m/2$ is odd, then the integral is over $\wp^{m/4-1/2}$, and $c(X^2) \equiv 1+2X^2 \pmod{\wp^{m/2}}$ for all $X \in \wp^{m/4-1/2}$. So, in that case,

$$\begin{aligned} J_\chi(u, v) &= \mathrm{meas}_{\mathrm{d}X}(\wp^{m/4+1/2})\Phi_{2w}(1) \sum_{X \in \wp^{m/4-1/2}/\wp^{m/4+1/2}} \Phi_{4w}(X^2) \\ &= q^{-m/4}\Phi(2w)q^{-1/2} \sum_{X \in R/\wp} \Phi_{4w\varpi^{m/2-1}}(X^2). \end{aligned}$$

By Lemma 6.2, and the fact that $m/2$ is odd, this can be rewritten as

$$q^{-m/4}\Phi(2w)\,\mathrm{sgn}_\varpi(-1)^{m/2-1}G_\varpi(\Phi_{4w}) = q^{-m/4}\Phi(2w)\,\mathrm{sgn}_\varpi(w)G_\varpi(\Phi).$$

The result now follows from $(*)$. □

From now on, we assume that χ is mildly ramified. In particular, we may take $h = 1$, so that Proposition 7.5 holds whenever $m > 2$.

Definition 7.6. For

- $\xi \in \mathfrak{f}^\times$,
- $\overline{\Phi}$ an (additive) character of $\mathfrak{f}$, and
- $\overline{\chi}$ a (multiplicative) character of $\mathfrak{f}^\times$,

define the corresponding *twisted Kloosterman sum* by

$$K(\overline{\chi}, \overline{\Phi}; \xi) := \sum_{x \in \mathfrak{f}^\times} \overline{\Phi}(x + \xi x^{-1})\overline{\chi}(x).$$

Proposition 7.7. *If $m = 2$, then*

$$J_\chi(u, v) = q^{-1}\chi(u\varpi)^{-1}K(\overline{\chi}, \overline{\Phi}; \xi).$$

Here,

- *ξ is the image in $\mathfrak{f}^\times$ of $\varpi^2 uv \in R^\times$,*
- *$\overline{\Phi}$ is the (additive) character of $\mathfrak{f} = R/\wp$ arising from the restriction to R of the depth-0 character $\Phi_{\varpi^{-1}}$ of k, and*
- *$\overline{\chi}$ is the (multiplicative) character of $\mathfrak{f}^\times \cong R^\times/1+\wp$ arising from the restriction to $R^\times$ of χ.*

Proof. By Theorem 7.4,

$$\begin{aligned}\chi(u\varpi)J_\chi(u,v) &= \chi(\varpi)\int_{\operatorname{ord}(x)=-1}\Phi(x+uvx^{-1})\chi(x)\,\mathrm{d}^\times x\\ &= \int_{R^\times}\Phi(\varpi^{-1}x+uv\cdot\varpi x^{-1})\chi(x)\,\mathrm{d}^\times x\\ &= \operatorname{meas}_{\mathrm{d}^\times x}(1+\wp)\sum_{x\in R^\times/1+\wp}\Phi_{\varpi^{-1}}(x+\varpi^2uvx^{-1})\chi(x)\,\mathrm{d}^\times x.\end{aligned}$$

Since $\operatorname{meas}_{\mathrm{d}^\times x}(1+\wp)=q^{-1}$, the result follows. □

Corollary 7.8. *Suppose $m=2$. Then for $\alpha\in\mathbb{C}$,*

$$J_{\nu^\alpha}(u,v)=q^{\alpha-1}|u|^{-\alpha}\sum_{\substack{c\in\wp^{-1}/R\\ c^2\neq uv}}\Phi(2c)\operatorname{sgn}_\varpi(c^2-uv),$$

$$J_{\nu^\alpha\operatorname{sgn}_\varpi}(u,v)=q^{\alpha-1/2}|u|^{-\alpha}\operatorname{sgn}_\varpi(v)G_\varpi(\Phi)\sum_{\substack{c\in\wp^{-1}/R\\ c^2=uv}}\Phi(2c).$$

Proof. If $\chi=\nu^\alpha$, then $\overline{\chi}=1$, so [Lidl and Niederreiter 1997, Theorem 5.47] gives

$$\begin{aligned}K(\overline{\chi},\overline{\Phi};\xi) &= \sum_{\substack{c\in\mathfrak{f}\\ c^2\neq\xi}}\overline{\Phi}(2c)\operatorname{sgn}_\mathfrak{f}(c^2-\alpha)\\ &= \sum_{\substack{c\in R/\wp\\ c^2\neq\varpi^2uv}}\overline{\Phi}(2c)\operatorname{sgn}_\varpi(c^2-\varpi^2uv)\\ &= \sum_{\substack{c\in\wp^{-1}/R\\ c^2\neq uv}}\Phi(2c)\operatorname{sgn}_\varpi(c^2-uv).\end{aligned}$$

(Note that our $\overline{\Phi}$ is their χ, and they write $K(\chi;a,b)$ where we write $K(\overline{\Phi},1;ab)$.)

If $\chi=\nu^\alpha\operatorname{sgn}_\varpi$, then $\overline{\chi}=\operatorname{sgn}_\mathfrak{f}$, therefore [Lidl and Niederreiter 1997, Exercises 5.84–85] gives that

$$K(\overline{\chi},\overline{\Phi};\xi)=\operatorname{sgn}_\mathfrak{f}(\xi)G(\operatorname{sgn}_\mathfrak{f},\overline{\Phi})\sum_{\substack{c\in\mathfrak{f}\\ c^2=\xi}}\overline{\Phi}(2c)=\operatorname{sgn}_\varpi(uv)G(\operatorname{sgn}_\mathfrak{f},\overline{\Phi})\sum_{\substack{c\in\wp^{-1}/R\\ c^2=uv}}\Phi(2c),$$

where $G(\operatorname{sgn}_\mathfrak{f},\overline{\Phi})=\sum_{X\in\mathfrak{f}^\times}\overline{\Phi}(X)\operatorname{sgn}_\mathfrak{f}(X)$. (Note that our $\overline{\Phi}$ is their χ and our $\overline{\chi}$ their η, and our $K(\overline{\chi},\overline{\Phi};\xi)$ is their $K(\eta,\chi;1,\xi)$.) Since $\mathrm{d}(\Phi)=-1$, Lemma 6.2 gives that $G(\operatorname{sgn}_\mathfrak{f},\overline{\Phi})=q^{1/2}\operatorname{sgn}_\varpi(-1)G_\varpi(\Phi)$.

The result now follows from Proposition 7.7. □

The following apparently specialised corollary allows simplification of many of our “shallow” computations (see Section 9A and Section 10A).

Corollary 7.9. *If* $m \geq 2$ *and* $\operatorname{ord}(u) = \operatorname{ord}(v)$, *then* $J_{\nu^\alpha \chi}(u, v)$ *is independent of* $\alpha \in \mathbb{C}$; *in particular*,

$$J_\chi^\epsilon(u, v) = J_\chi(u, v) \quad \textit{and} \quad J_\chi^\varpi(u, v) = J_{\chi\,\mathrm{sgn}_\epsilon}^\varpi(u, v).$$

If $m \geq 2$ *and* $\operatorname{ord}(u) = \operatorname{ord}(v) + 2$, *then* $J_{\nu^\alpha \chi}(u, v) = q^\alpha J_\chi(u, v)$; *in particular*,

$$J_\chi^\epsilon(u, v) = 0 \quad \textit{and} \quad J_\chi^\varpi(u, v) = -J_{\chi\,\mathrm{sgn}_\epsilon}^\varpi(u, v).$$

Proof. Suppose that $m > 2$. If $uv \notin (k^\times)^2$, then $J_{\nu^\alpha \chi}(u, v) = 0$ for all $\alpha \in \mathbb{C}$. If $uv = w^2$, then the only dependence on α in Proposition 7.5 is via the factor $\chi(u^{-1}w)$. If $\operatorname{ord}(u) = \operatorname{ord}(v)$, then also $\operatorname{ord}(w) = \operatorname{ord}(u)$, so $\nu^\alpha(u^{-1}w) = 1$. If $\operatorname{ord}(u) = \operatorname{ord}(v) + 2$, then $\operatorname{ord}(w) = \operatorname{ord}(u) - 1$, so $\nu^\alpha(u^{-1}w) = q^\alpha$.

Now suppose that $m = 2$, that is, that $\operatorname{ord}(uv) = -2$. Since $\overline{\nu^\alpha \chi} = \overline{\chi}$, the only dependence on α in Proposition 7.7 is via the factor $\chi(u\varpi)^{-1}$. If $\operatorname{ord}(u) = \operatorname{ord}(v)$, then $\operatorname{ord}(u) = -1$, so $\nu^\alpha(u\varpi) = 1$. If $\operatorname{ord}(u) = \operatorname{ord}(v) + 2$, then $\operatorname{ord}(u) = 0$, so $\nu^\alpha(u\varpi) = q^{-\alpha}$. □

8. A mock Fourier transform

We introduce a function $M_{X^*}^G$ specified by an integral formula (see Definition 8.4) reminiscent of the usual one for (the function representing) $\hat{\mu}_{X^*}^G$ (see [Adler and DeBacker 2004, Theorem A.1.2]). We will eventually show (see Proposition 11.2) that it is actually *equal* to $\hat{\mu}_{X^*}^G$, but first we spend some time computing it.

In the notation of Definition 4.1, we have

$$\operatorname{tr} g \cdot \sqrt{\theta} \cdot g^{-1} \cdot \sqrt{\theta'} = N_\theta(\alpha) \cdot \theta' + N_\theta(\gamma), \tag{8.1}$$

where

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

$\alpha = a + b\sqrt{\theta}$, and $\gamma = c + d\sqrt{\theta}$. Since $1 = ad - bc = \operatorname{Im}_\theta(\overline{\alpha} \cdot \gamma)$, we have that $\gamma = \overline{\alpha}^{-1} \cdot (t + \sqrt{\theta})$ for some $t \in k$; specifically, $t = \operatorname{Re}_\theta(\overline{\alpha} \cdot \gamma) = ac - bd\theta$. This calculation motivates the definition of the following map.

Definition 8.2. Define $\varphi_\theta : G \to k_\theta^\times \times k$ by

$$\varphi_\theta(g) = (a + b\sqrt{\theta}, ac - bd\theta)$$

for

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G.$$

Note that φ_θ is a bianalytic map (of k-manifolds), with inverse

$$(\alpha, t) \mapsto \begin{pmatrix} \mathrm{Re}_\theta(\alpha) & \mathrm{Im}_\theta(\alpha) \\ N_\theta(\alpha)^{-1}\bigl(t\cdot\mathrm{Re}_\theta(\alpha)+\theta\cdot\mathrm{Im}_\theta(\alpha)\bigr) & N_\theta(\alpha)^{-1}\bigl(\mathrm{Re}_\theta(\alpha)+t\cdot\mathrm{Im}_\theta(\alpha)\bigr) \end{pmatrix}.$$

It is *not* an isomorphism, but its restrictions to T_θ, A, and

$$\left\{\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} : b \in k\right\}$$

are isomorphisms onto $C_\theta \times \{0\}$, $k^\times \times \{0\}$, and $\{1\} \times k$, respectively. In fact, the next lemma says a bit more.

Lemma 8.3. *If $g \in G$ satisfies $\varphi_\theta(g) = (\alpha, t)$, and*

- *$h \in T_\theta$ is identified with $\eta \in C_\theta$,*
- *$a = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ with $\lambda \in k^\times$, and*
- *$\bar{u} = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$ with $b \in k$,*

then

$$\begin{aligned} \varphi_\theta(gh) &= (\alpha\eta, t), \\ \varphi_\theta(ag) &= (\lambda\alpha, t), \\ \varphi_\theta(\bar{u}g) &= (\alpha, t + N_\theta(\alpha)b). \end{aligned}$$

Proof. This is a straightforward computation. □

We can now define our "mock orbital integral". Again, Proposition 11.2 will eventually show that it is actually equal to the function in which we are interested.

Definition 8.4. For $\alpha \in k_\theta^\times$ and $t \in k$, put

$$\langle X^*, Y\rangle_{\alpha,t} := \beta s\bigl(N_\theta(\alpha)\cdot\theta' + N_\theta(\alpha)^{-1}\cdot\theta - N_\theta(\alpha)^{-1}\cdot t^2\bigr).$$

The dependence on α is only via $N_\theta(\alpha)$. Thus, we may define

$$M_{X^*}^G(Y) := \operatorname{P\!\!\!\int}_{k_\theta^\times/C_\theta} \operatorname{P\!\!\!\int}_k \Phi(\langle X^*, Y\rangle_{\alpha,t})\,\mathrm{d}t\,\mathrm{d}^\times\dot{\alpha},$$

where

$$\begin{aligned} \operatorname{P\!\!\!\int}_k f(x)\,\mathrm{d}t &:= \sum_{n\in\mathbb{Z}} \int_{\mathrm{ord}(x)=n} f(x)\,\mathrm{d}t, \\ \operatorname{P\!\!\!\int}_{k^\times} f(x)\,\mathrm{d}^\times x &:= \sum_{n\in\mathbb{Z}} \int_{\mathrm{ord}(x)=n} f(x)\,\mathrm{d}^\times x, \\ \operatorname{P\!\!\!\int}_{k_\theta^\times/C_\theta} (f\circ N_\theta)(\alpha)\,\mathrm{d}^\times\dot{\alpha} &:= \operatorname{P\!\!\!\int}_{k^\times} [N_\theta(k^\times)](x) f(x)\,\mathrm{d}^\times x \end{aligned}$$

(for those $f \in C^\infty(k)$ for which the sum converges) are "principal-value" integrals, as in [Sally and Taibleson 1966, p. 282]. Here, $\mathrm{d}t$ and $\mathrm{d}^\times x$ are the measures of Definition 2.1, and $[S]$ denotes the characteristic function of S.

By (8.1) (and Notations 5.1 and 5.8), we have that

$$\textbf{(8.5)} \qquad \langle X^*, Y\rangle_{\alpha,t} = \langle \mathrm{Ad}^*(g)X^*, Y\rangle \quad \text{when } \varphi_\theta(g) = (\alpha, t),$$

where the pairing $\langle\,\cdot\,,\cdot\,\rangle$ on the right is the usual pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$.

Notation 8.6. $u = \varpi^{-(r'+1)}s\theta'$, $v = \varpi^{-(r'+1)}s\theta$, and $m = -\operatorname{ord}(uv)$.

This is a special case of Notation 7.3. These particular values of u and v will be fixed for the remainder of the paper. It follows that

$$\textbf{(8.7)} \qquad uv = (\varpi^{-(r'+1)}s)^2 \cdot \theta\theta',$$

therefore

$$\textbf{(8.8)} \qquad uv \in (k^\times)^2 \quad \Leftrightarrow \quad \theta\theta' \in (k^\times)^2.$$

We use Lemma 5.9 to compute

$$\textbf{(8.9)} \qquad \operatorname{ord}(u) = -(r'+1) + \operatorname{ord}(s\theta') = -\bigl(r' + 1 + \tfrac{1}{2}\operatorname{ord}(\theta')\bigr) + \mathrm{d}(Y),$$

$$\textbf{(8.10)} \qquad m = 2(r'+1) - \operatorname{ord}(s^2\theta') - \operatorname{ord}(\theta) = 2(r'+1-\mathrm{d}(Y)) - \operatorname{ord}(\theta).$$

8A. ***Mock Fourier transforms and Bessel functions.*** We can now evaluate the integral occurring in Definition 8.4 in terms of Bessel functions — or, rather, the sums J_χ^θ of Definition 7.1.

Proposition 8.11. *Let J_χ^θ be as in Definition 7.1 and $\gamma_{\mathrm{un}}(s)$ and $\gamma_{\mathrm{ram}}(s)$ be as in Definition 6.5. Then*

$$M_{X^*}^G(Y) = \tfrac{1}{2}|s|^{-1/2}q^{-(r'+1)/2} \times \Bigl(\bigl(J_{\nu^{1/2}}^\theta(u,v) + \gamma_{\mathrm{un}}(s)J_{\nu^{1/2}\,\mathrm{sgn}_\epsilon}^\theta(u,v)\bigr) + \gamma_{\mathrm{ram}}(s)\bigl(J_{\nu^{1/2}\,\mathrm{sgn}_\varpi}^\theta(u,v) - \gamma_{\mathrm{un}}(s)J_{\nu^{1/2}\,\mathrm{sgn}_{\epsilon\varpi}}^\theta(u,v)\bigr)\Bigr),$$

Proof. Recall the notation $\Phi' = \Phi_\beta$ from Notation 5.2. By Definition 8.4,

$$(*) \quad M_{X^*}^G(Y) = \fint_{k_\theta^\times/C_\theta} \Phi'_s\bigl(N_\theta(\alpha)\cdot\theta' + N_\theta(\alpha)^{-1}\cdot\theta\bigr)\cdot \fint_k \Phi'(-sN_\theta(\alpha)^{-1}t^2)\,\mathrm{d}t\,\mathrm{d}^\times\dot\alpha$$
$$= q^{-(r'+1)/2}\fint_{k^\times}[N_\theta(k_\theta^\times)](x)j(\theta',\theta;x)\mathcal{H}(\Phi',-sx^{-1})\,\mathrm{d}^\times x,$$

where

- $j(\theta',\theta;x) := \Phi'_s(\theta' x + \theta x^{-1}) = \Phi\bigl(\beta s(\theta' x + \theta x^{-1})\bigr)$ for $x \in k^\times$, and
- $\mathcal{H}(\Phi',b) = \fint_k \Phi'(bt^2)\,\mathrm{d}_{\Phi'}t$ for $b \in k^\times$ is as in [Shalika 2004, p. 6].

In particular, $\mathrm{d}_{\Phi'}t$ is the Φ'-self-dual Haar measure on k; by [Shalika 2004, p. 5], it satisfies $\mathrm{d}t = q^{-(r'+1)/2}\,\mathrm{d}_{\Phi'}t$. This is the reason for the appearance of $q^{-(r'+1)/2}$ on the last line of the computation.

The significance of j is that integrating it against a (multiplicative) character χ of $k^\times$ corresponds to evaluating a Bessel function of order χ, in the sense of Definition 7.1. To be precise, our character Φ' has depth r', not -1, so we must work instead with $\Phi'_{\varpi^{r'+1}}$. Then

$$j(\theta',\theta;x) = \Phi'_{\varpi^{r'+1}}\bigl((\varpi^{-(r'+1)}s\theta')x + (\varpi^{-(r'+1)}s\theta)x^{-1}\bigr) = \Phi'_{\varpi^{r'+1}}(ux + vx^{-1}),$$

where (u, v) is as in Notation 8.6, so

$$(\dagger)\qquad \int\!\!\!\!\!\!-_{k^\times} j(\theta',\theta;x)\chi(x)\,\mathrm{d}^\times x = J_\chi(u,v)$$

for $\chi \in \widehat{k^\times}$.

Now $\frac{1}{2}(1+\mathrm{sgn}_\theta)$ is the characteristic function of $N_\theta(k_\theta^\times)$, so we may rewrite $(*)$:

$$(**)\qquad q^{-(r'+1)/2}\int\!\!\!\!\!\!-_{k^\times} \tfrac{1}{2}(1+\mathrm{sgn}_\theta(x))\cdot j(\theta',\theta;x)\mathscr{K}(\Phi',-sx^{-1})\,\mathrm{d}^\times x.$$

By [Shalika 2004, Lemma 1.3.2] and Lemma 6.2, we have

$$\mathscr{K}(\Phi',b) = |b|^{-1/2}\begin{cases}\mathrm{sgn}_\varpi(b)G_\varpi(\Phi') & r'-\mathrm{ord}(b)\text{ even},\\ 1 & r'-\mathrm{ord}(b)\text{ odd}.\end{cases}$$

We find it useful to describe $\mathscr{K}(\Phi',b)$ without explicit use of cases. As above, $\frac{1}{2}(1+(-1)^n\,\mathrm{sgn}_\epsilon)$ is the characteristic function of $\{b\in k^\times : \mathrm{ord}(b)\equiv n \pmod 2\}$, so we may rewrite

$$\mathscr{K}(\Phi',b) = \tfrac{1}{2}(1+(-1)^{r'}\,\mathrm{sgn}_\epsilon(b))\,\mathrm{sgn}_\varpi(b)G_\varpi(\Phi') + \tfrac{1}{2}(1-(-1)^{r'}\,\mathrm{sgn}_\epsilon(b)).$$

Plugging this into $(**)$ with $b=-st^{-1}$ gives

$$\begin{aligned}M^G_{X^*}(Y) = \tfrac{1}{2}|s|^{-1/2}q^{-(r'+1)/2}&\\ \times \int\!\!\!\!\!\!-_{k^\times}\tfrac{1}{2}(1+\mathrm{sgn}_\theta(x))&\\ \times\bigl((1-\gamma_{\mathrm{un}}(s)\,\mathrm{sgn}_\epsilon(x))\gamma_{\mathrm{ram}}(s)\,\mathrm{sgn}_\varpi(x) &+ (1+\gamma_{\mathrm{un}}(s)\,\mathrm{sgn}_\epsilon(x))\bigr)\\ &\times |x|^{1/2}j(\theta',\theta;x)\,\mathrm{d}^\times x.\end{aligned}$$

Expanding the product and applying $(\dagger)$ gives the desired formula. □

8B. *"Deep" Bessel functions.* By Proposition 8.11, one approach to computing $M^G_{X^*}(Y)$ (hence $\hat\mu^G_{X^*}(Y)$, by Proposition 11.2) is to evaluate many Bessel functions, and this is exactly what we do. As Theorem 7.4 makes clear, the behaviour of Bessel functions is more predictable when $m<2$ than otherwise. We introduce a convenient shorthand for referring to Bessel functions in this range; we will only use it in this section, and Sections 9B and 10C.

Notation 8.12. Define

$$[A;\,B]_{\theta,r'}(\theta') := |\theta|^{1/2}A + q^{-(r'+1)}|D_{\mathfrak{g}}(Y)|^{-1/2}B(\theta').$$

We usually suppress the subscript on $[A;\,B]$ and sometimes write

$$[A;\,B(1),\,B(\epsilon),\,B(\varpi),\,B(\epsilon\varpi)](\theta')$$

for the same quantity.

Proposition 8.13. *With Notations 5.2, 5.8, 8.6, and the notation of Definition 6.5, if $m < 2$, then*

$$|s|^{-1/2}q^{-(r'+1)/2}J_{\nu^{1/2}\chi}(u,v) = \begin{cases} [Q_3(q^{-1/2});\,1](\theta') & \chi = 1, \\ \gamma_{\mathrm{un}}(s)[\mathrm{sgn}_\epsilon(\theta)Q_3(-q^{-1/2});\,\mathrm{sgn}_\epsilon](\theta') & \chi = \mathrm{sgn}_\epsilon, \\ \gamma_{\mathrm{ram}}(s)^{-1}[\mathrm{sgn}_\varpi(\theta)q^{-1};\,\mathrm{sgn}_\varpi](\theta') & \chi = \mathrm{sgn}_\varpi, \\ -\gamma_{\mathrm{un}}(s)\gamma_{\mathrm{ram}}(s)^{-1}[\mathrm{sgn}_{\epsilon\varpi}(\theta)q^{-1};\,\mathrm{sgn}_{\epsilon\varpi}](\theta') & \chi = \mathrm{sgn}_{\epsilon\varpi}, \end{cases}$$

where

$$Q_3(T) = -T(T^2+T+1).$$

The unexpected factor $|s|^{-1/2}q^{-(r'+1)/2}$ on the left-hand side crops up repeatedly in calculations (see, for example, Proposition 8.11), so it simplifies matters to include it in this calculation.

Proof. By Theorem 7.4 and Lemma 5.9,

$$\begin{aligned} J_{\nu^{1/2}\chi}(u,v) &= (\nu^{1/2}\chi)(v)\Gamma(\nu^{-1/2}\chi) + (\nu^{-1/2}\chi)(u)\Gamma(\nu^{1/2}\chi) \\ &= (\nu^{1/2}\chi)(v\theta^{-1}) \\ &\qquad \times \bigl((\nu^{1/2}\chi)(\theta)\Gamma(\nu^{-1/2}\chi) + (\nu^{-1/2}\chi)(uv\theta^{-1})\Gamma(\nu^{1/2}\chi)\bigr) \\ &= |s|^{1/2}q^{(r'+1)/2}\chi(\varpi^{r'+1}s)\bigl[\chi(\theta)\Gamma(\nu^{-1/2}\chi);\,\Gamma(\nu^{1/2}\chi)\cdot\chi\bigr](\theta') \end{aligned}$$

whenever $\chi^2 = 1$.

In particular, using [Sally and Taibleson 1966, Theorem 3.1(i, ii)] to compute the Γ-factors, we see that $|s|^{-1/2}q^{-(r'+1)/2}J_{\nu^{1/2}\chi}(u,v)$ is given by

$$(*)\quad \begin{cases} [Q_3(q^{-1/2});\,1](\theta') & \chi = 1, \\ \gamma_{\mathrm{un}}(s)[\mathrm{sgn}_\epsilon(\theta)Q_3(-q^{-1/2});\,\mathrm{sgn}_\epsilon](\theta') & \chi = \mathrm{sgn}_\epsilon, \\ \mathrm{sgn}_\varpi(\varpi^{r'+1}s)\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi)[\mathrm{sgn}_\varpi(\theta)q^{-1};\,\mathrm{sgn}_\varpi](\theta') & \chi = \mathrm{sgn}_\varpi, \\ \gamma_{\mathrm{un}}(s)\,\mathrm{sgn}_\varpi(\varpi^{r'+1}s)\Gamma(\nu^{1/2}\,\mathrm{sgn}_{\epsilon\varpi})[\mathrm{sgn}_{\epsilon\varpi}(\theta)q^{-1};\,\mathrm{sgn}_{\epsilon\varpi}](\theta') & \chi = \mathrm{sgn}_{\epsilon\varpi}. \end{cases}$$

By [ibid., Theorem 3.1(ii)] again and the fact that $\mathrm{sgn}_{\epsilon\varpi} = \nu^{i\pi/\ln(q)}\,\mathrm{sgn}_\varpi$, we have $\Gamma(\nu^{1/2}\,\mathrm{sgn}_{\epsilon\varpi}) = -\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi)$; and, by Lemma 6.3, Definition 6.5, and (6.7),

$$\begin{aligned}\mathrm{sgn}_\varpi(\varpi^{r'+1}s)\Gamma(\nu^{1/2}\,\mathrm{sgn}_\varpi) &= \mathrm{sgn}_\varpi(-1)^{r'+1}\,\mathrm{sgn}_\varpi(s)\cdot\mathrm{sgn}_\varpi(-1)^{r'+1}G_\varpi(\Phi') \\ &= \mathrm{sgn}_\varpi(s)G_\varpi(\Phi') \\ &= \gamma_{\mathrm{ram}}(s)^{-1}.\end{aligned}$$

This shows that $(*)$ reduces to the table in the statement. □

9. Split and unramified orbital integrals

Throughout this section, we have

$$\theta = 1 \text{ or } \theta = \epsilon, \quad \text{so that} \quad r' = r. \tag{9.1}$$

In the split case, $J^1_\chi = J_\chi$ for $\chi \in \widehat{k^\times}$, so Proposition 8.11 gives

$$\begin{aligned} M^G_{X^*}(Y) = \tfrac{1}{2}|s|^{-1/2}q^{-(r+1)/2} \\ \times\Big(\big(J_{\nu^{1/2}}(u,v) + \gamma_{\mathrm{un}}(s)J_{\nu^{1/2}\,\mathrm{sgn}_\epsilon}(u,v)\big) \\ + \gamma_{\mathrm{ram}}(s)\big(J_{\nu^{1/2}\,\mathrm{sgn}_\varpi}(u,v) - \gamma_{\mathrm{un}}(s)J_{\nu^{1/2}\,\mathrm{sgn}_{\epsilon\varpi}}(u,v)\big)\Big).\end{aligned} \tag{9.2}$$

In the unramified case, $J^\epsilon_\chi = J^\epsilon_{\chi\,\mathrm{sgn}_\epsilon}$ for $\chi \in \widehat{k^\times}$, so Proposition 8.11 gives

$$\begin{aligned} M^G_{X^*}(Y) = \tfrac{1}{2}|s|^{-1/2}q^{-(r+1)/2} \\ \times\big((1+\gamma_{\mathrm{un}}(s))J^\epsilon_{\nu^{1/2}}(u,v) + \gamma_{\mathrm{ram}}(s)(1-\gamma_{\mathrm{un}}(s))J^\epsilon_{\nu^{1/2}\,\mathrm{sgn}_\varpi}(u,v)\big).\end{aligned} \tag{9.3}$$

By (6.8) and (6.7),

$$\mathrm{sgn}_\varpi(v)G_\varpi(\Phi'_{\varpi^{r+1}}) = \begin{cases}\mathrm{sgn}_\varpi(-1)\gamma_{\mathrm{ram}}(s) = \gamma_{\mathrm{ram}}(s)^{-1} & \theta = 1,\\ \mathrm{sgn}_\varpi(-\epsilon)\gamma_{\mathrm{ram}}(s) = -\gamma_{\mathrm{ram}}(s)^{-1} & \theta = \epsilon.\end{cases} \tag{9.4}$$

9A. ***Far from zero.*** The results of this section are special cases for split and unramified orbital integrals of results of Waldspurger [1995, Proposition VIII.1]. We prove analogues of these results for ramified orbital integrals in Section 10A.

The qualitative behaviour of unramified orbital integrals does not change as we pass from elements of depth less than r to those of depth exactly r; this is unlike the situation for ramified orbital integrals. See Section 10B.

Theorem 9.5. *If* $\mathrm{d}(X^*) + \mathrm{d}(Y) \le 0$ *and* X^* *is split or unramified*, *then* $M^G_{X^*}(Y) = 0$ *unless* X^* *and* Y *lie in* G*-conjugate tori.*

Proof. Recall that $\theta = 1$ if X^* is split, and $\theta = \epsilon$ if X^* is unramified.

By (8.10), $m \ge 2$; in fact, $m > 2$ (indeed, m is odd) unless $\mathrm{ord}(\theta')$ is even.

If $m > 2$, then Proposition 7.5 and (8.8) show that $M^G_{X^*}(Y) = 0$ unless $\theta\theta' \in (k^\times)^2$. By Section 4, it therefore suffices to consider the cases when $\theta = \epsilon$ and $\theta' = \varpi^2\epsilon$, that is, X^* and Y lie in stably, but not rationally, conjugate tori; and when $m = 2$ and $\{\theta, \theta'\} = \{1, \epsilon\}$, that is, one of X^* or Y is split, and the other unramified.

Suppose first that $\theta = \epsilon$ and $\theta' = \varpi^2\epsilon$, so that $\operatorname{ord}(u) = \operatorname{ord}(v) + 2$. By Corollary 7.9, (9.3) becomes $M^G_{X^*}(Y) = 0$.

Now suppose $\{\theta, \theta'\} = \{1, \epsilon\}$ and $m = 2$. By Corollary 7.9, since $\operatorname{ord}(u) = \operatorname{ord}(v)$,

$$J_{\nu^{1/2}}(u, v) = J_{\nu^{1/2}\operatorname{sgn}_\epsilon}(u, v) \quad \text{and} \quad J_{\nu^{1/2}\operatorname{sgn}_\varpi}(u, v) = J_{\nu^{1/2}\operatorname{sgn}_{\epsilon\varpi}}(u, v),$$

so (9.2) agrees with (9.3). We shall work with (9.3), since it is simpler.

By Corollary 7.8 and (8.8), $J_{\nu^\alpha \operatorname{sgn}_\varpi}(u, v) = 0$ for all $\alpha \in \mathbb{C}$, in particular, for $\alpha = 1/2$ and $\alpha = 1/2 + i\pi/\ln(q)$. By (8.10), $\operatorname{ord}(s) = r$, so, by Definition 6.5, $\gamma_{\mathrm{un}}(s) = -1$, and (9.3) (hence also (9.2)) becomes

$$M^G_{X^*}(Y) = \tfrac{1}{2}|s|^{-1/2} J^\epsilon_{\nu^{1/2}\operatorname{sgn}_\varpi}(u, v) = 0. \qquad \square$$

Theorem 9.6. *If* $\mathrm{d}(X^*) + \mathrm{d}(Y) \le 0$ *and* X^* *and* Y *lie in a common split or unramified torus* $\boldsymbol{T}$ *(with* $T = \boldsymbol{T}(k)$*), then*

$$M^G_{X^*}(Y) = q^{-(r+1)}|D_{\mathfrak{g}}(Y)|^{-1/2}\gamma_\Phi(X^*, Y) \sum_{\sigma \in W(G,T)} \Phi(\langle \mathrm{Ad}^*(\sigma)X^*, Y\rangle),$$

where $\gamma_\Phi(X^*, Y)$ *is as in Definition 6.5.*

Proof. The hypothesis implies that $\theta = \theta'$, so $u = v$. By Corollary 7.9,

$$J_{\nu^{1/2}}(u, v) = J_{\nu^{1/2}\operatorname{sgn}_\epsilon}(u, v) \quad \text{and} \quad J_{\nu^{1/2}\operatorname{sgn}_\varpi}(u, v) = J_{\nu^{1/2}\operatorname{sgn}_{\epsilon\varpi}}(u, v),$$

so (9.2) agrees with (9.3). We again work with (9.3), since it is simpler.

By Remark 4.7, $W(G, T_\theta) = \{1, \sigma_\theta\}$, where $\mathrm{Ad}^*(\sigma_\theta)X^* = -X^*$.

We may take the square root w of uv in Proposition 7.5 to be just u. By (8.10),

$$(*) \qquad q^{-m/4} = q^{-(r+1)/2} q^{\operatorname{ord}(s)/2} = q^{-(r+1)/2}|s|^{-1/2}.$$

By Notations 5.2 and 8.6,

$$(**) \qquad \Phi'_{\varpi^{r+1}}(\pm 2w) = \Phi'(\pm 2s\theta) = \Phi(\pm 2\beta s\theta) = \Phi(\pm\langle X^*, Y\rangle)$$

(the last equality following, for example, from (8.5)).

Suppose $\operatorname{ord}(s) \not\equiv r \pmod 2$, so $\gamma_{\mathrm{un}}(s) = 1$ and $\gamma_\Phi(X^*, Y) = 1$. By Corollary 7.9, since $u = v$, (9.3) (hence also (9.2)) becomes

$$(\dagger) \qquad \begin{aligned} M^G_{X^*}(Y) &= \tfrac{1}{2}|s|^{-1/2} q^{-(r+1)/2} \cdot 2 \cdot J^\epsilon_{\nu^{1/2}}(u, v) \\ &= |s|^{-1/2} q^{-(r+1)/2} J_{\nu^{1/2}}(u, v). \end{aligned}$$

Since $m > 2$ and $4 \mid m$ by (8.10), combining Proposition 7.5, $(*)$, and $(**)$ gives

$$(\dagger\dagger)\quad J_{\nu^{1/2}}(u,v) = q^{-(r+1)/2}|s|^{-1/2}\bigl(\Phi(\langle X^*, Y\rangle) + \Phi(-\langle X^*, Y\rangle)\bigr)$$
$$= q^{-(r+1)/2}|s|^{-1/2} \sum_{\sigma \in W(G,T_\theta)} \Phi(\langle \mathrm{Ad}^*(\sigma)X^*, Y\rangle)$$
$$= q^{-(r+1)/2}|s\theta'|^{-1/2}\gamma_\Phi(X^*, Y) \sum_{\sigma \in W(G,T_\theta)} \Phi(\langle \mathrm{Ad}^*(\sigma)X^*, Y\rangle).$$

The result (in this case) now follows from Lemma 5.9 by combining $(\dagger)$ and $(\dagger\dagger)$.

Suppose now that $\mathrm{ord}(s) \equiv r \pmod 2$, so that $\gamma_{\mathrm{un}}(s) = -1$ and

$$\gamma_\Phi(X^*, Y) = \begin{cases} 1 & \theta = 1, \\ -1 & \theta = \epsilon. \end{cases}$$

Again by Corollary 7.9, since $u = v$, (9.3) (hence also (9.2)) becomes (as in $(\dagger)$)

$$(\dagger')\qquad M^G_{X^*}(Y) = |s|^{-1/2} q^{-(r+1)/2}\gamma_{\mathrm{ram}}(s) J_{\nu^{1/2}\,\mathrm{sgn}_\varpi}(u,v).$$

Since $4 \nmid m$ by (8.10), if $m > 2$, then combining Proposition 7.5, $(*)$, (9.4), and $(**)$ gives (as in $(\dagger\dagger)$)

$$(\dagger\dagger'_{<r})\quad J_{\nu^{1/2}\,\mathrm{sgn}_\varpi}(u,v)$$
$$= q^{-(r+1)/2}|s\theta'|^{-1/2}\gamma_{\mathrm{ram}}(s)^{-1}\gamma_\Phi(X^*, Y) \sum_{\sigma \in W(G,T_\theta)} \Phi(\langle \mathrm{Ad}^*(\sigma)X^*, Y\rangle).$$

If $m = 2$, then $|s| = q^{-r}$ and $\mathrm{ord}(u) = -1$ by Lemma 5.9, (8.9), and (8.10). Thus, combining Corollary 7.8, (9.4), and $(**)$ gives

$$(\dagger\dagger'_{=r})\quad J_{\nu^{1/2}\,\mathrm{sgn}_\varpi}(u,v)$$
$$= q^{-1/2}\gamma_{\mathrm{ram}}(s)^{-1}\gamma_\Phi(X^*, Y) \sum_{\sigma \in W(G,T_\theta)} \Phi(\langle \mathrm{Ad}^*(\sigma)X^*, Y\rangle)$$
$$= q^{-(r+1)/2}|s\theta'|^{-1/2}\gamma_{\mathrm{ram}}(s)^{-1}\gamma_\Phi(X^*, Y) \sum_{\sigma \in W(G,T_\theta)} \Phi(\langle \mathrm{Ad}^*(\sigma)X^*, Y\rangle).$$

The result follows by combining $(\dagger')$ and $(\dagger\dagger'_{<r})$ *or* $(\dagger\dagger'_{=r})$ with Lemma 5.9. □

9B. *Close to zero.*

Theorem 9.7. *If* $\mathrm{d}(X^*)+\mathrm{d}(Y) > 0$, *and* X^* *is split or unramified, then let* $\gamma_\Phi(X^*, Y)$ *and* $c_0(X^*)$ *be as in Definitions 6.5 and 6.10, respectively. Then*

$$M^G_{X^*}(Y) = c_0(X^*) + \frac{2}{n(X^*)} q^{-(r+1)}|D_{\mathfrak{g}}(Y)|^{-1/2}\gamma_\Phi(X^*, Y),$$

where

$$n(X^*) = \begin{cases} 1 & \text{for } X^* \text{ split}, \\ 2 & \text{for } X^* \text{ elliptic}. \end{cases}$$

Proof. By (8.10), $m < 2$.

By Proposition 8.13, using Notation 8.12, (9.2) becomes

$$M^G_{X^*}(Y) = \tfrac{1}{2}\big[Q_3(q^{-1/2}) + Q_3(q^{-1/2}) - q^{-1} - q^{-1};\ 1 + \mathrm{sgn}_\epsilon + \mathrm{sgn}_\varpi + \mathrm{sgn}_{\epsilon\varpi}\big](\theta').$$

Since

$$Q_3(q^{-1/2}) + Q_3(-q^{-1/2}) = -2T^2\big|_{T=q^{-1/2}} = -2q^{-1}, \tag{9.8}$$

this simplifies (by the Plancherel formula on $k^\times/(k^\times)^2$!) to

$$M^G_{X^*}(Y) = [-2q^{-1}; 2, 0, 0, 0].$$

Similarly, (9.3) becomes

$$M^G_{X^*}(Y) = \tfrac{1}{2}\Big(\tfrac{1}{2}(1 + \gamma_{\mathrm{un}}(s))\underbrace{\big[Q_3(q^{-1/2}) + \gamma_{\mathrm{un}}(s)Q_3(-q^{-1/2});\ 1 + \gamma_{\mathrm{un}}(s)\,\mathrm{sgn}_\epsilon\big]}_{\text{(I)}}$$
$$+ \tfrac{1}{2}(1 - \gamma_{\mathrm{un}}(s))\underbrace{\big[-(1 - \gamma_{\mathrm{un}}(s))q^{-1};\ (1 - \gamma_{\mathrm{un}}(s)\,\mathrm{sgn}_\epsilon)\,\mathrm{sgn}_\varpi\big]}_{\text{(II)}}\Big)(\theta').$$

Since $\gamma_{\mathrm{un}}(s) = \pm 1$ (see Definition 6.5), we may replace $\gamma_{\mathrm{un}}(s)$ by 1 in (I) and by -1 in (II), then use (9.8) and check case-by-case to see that the formula simplifies to

$$M^G_{X^*}(Y) = [-q^{-1}; 1, \gamma_{\mathrm{un}}(s), 0, 0](\theta'). \qquad \square$$

10. Ramified orbital integrals

Throughout this section, we have

$$\theta = \varpi, \quad \text{so that} \quad r' = r + \tfrac{1}{2} =: h. \tag{10.1}$$

Then $J^\varpi_\chi = J^\varpi_{\chi\,\mathrm{sgn}_\varpi}$ for $\chi \in \widehat{k^\times}$, so Proposition 8.11 gives

$$\begin{aligned} M^G_{X^*}(Y) &\\ = \tfrac{1}{2}|s|^{-1/2}\big((1 + \gamma_{\mathrm{ram}}(s))J^\varpi_{\nu^{1/2}}(u, v) &+ \gamma_{\mathrm{un}}(s)(1 - \gamma_{\mathrm{ram}}(s))J^\varpi_{\nu^{1/2}\,\mathrm{sgn}_\epsilon}(u, v)\big). \end{aligned} \tag{10.2}$$

By (6.8),

$$\mathrm{sgn}_\varpi(v)G_\varpi(\Phi'_{\varpi^{h+1}}) = \mathrm{sgn}_\varpi(-\varpi)\gamma_{\mathrm{ram}}(s) = \gamma_{\mathrm{ram}}(s). \tag{10.3}$$

10A. ***Far from zero.*** As in Section 9A, the results of this section are special cases of [Waldspurger 1995, Proposition VIII.1].

Theorem 10.4. *If* $\mathrm{d}(X^*) + \mathrm{d}(Y) < 0$ *and* X^* *is ramified, then* $M^G_{X^*}(Y) = 0$ *unless* X^* *and* Y *lie in* G*-conjugate tori.*

Proof. By (8.10), $m > 2$, so Proposition 7.5 and (8.8) show that $M^G_{X^*}(Y) = 0$ unless $\varpi\theta' \in (k^\times)^2$. By Section 4, it therefore suffices to consider the case when $-1 \in (\mathfrak{f}^\times)^2$ (so $\operatorname{sgn}_\varpi(-1) = 1$) and $\theta' = \epsilon^2\varpi$, that is, X^* and Y lie in stably, but not rationally, conjugate tori.

By (8.7), we may take the square root w of uv to be $w = \varpi^{-h}s\epsilon = \epsilon^{-1}u$. Then $u^{-1}w = \epsilon^{-1}$, so Proposition 7.5 shows (whether or not 4 divides m) that, if χ is mildly ramified *and* trivial at -1, then

$$J_{\chi\operatorname{sgn}_\varpi}(u, v) = \operatorname{sgn}_\varpi(u^{-1}\varpi)J_\chi(u, v) = -J_\chi(u, v),$$

hence $J^\varpi_\chi(u, v) = 0$. In particular, this equality holds for $\chi = \nu^{1/2}$ and $\chi = \nu^{1/2}\operatorname{sgn}_\epsilon$. It follows from (10.2) that $M^G_{X^*}(Y) = 0$. □

Theorem 10.5. *If* $\mathrm{d}(X^*) + \mathrm{d}(Y) < 0$, *and* X^* *and* Y *lie in a common ramified torus* $\boldsymbol{T}$ (*with* $T = \boldsymbol{T}(k)$), *then*

$$M^G_{X^*}(Y) = q^{-(h+1)}|D_{\mathfrak{g}}(Y)|^{-1/2}\gamma_\Phi(X^*, Y) \sum_{\sigma\in W(G,T)} \Phi(\langle \mathrm{Ad}^*(\sigma)(X^*), Y\rangle),$$

where $\gamma_\Phi(X^*, Y)$ *is as in Definition 6.5.*

Proof. Since we have fixed $\theta = \varpi$, the hypothesis implies that $\theta' = \varpi$. In particular, $u = v$. Write σ_ϖ for the nontrivial element of $W(\boldsymbol{G}, \boldsymbol{T}_\varpi)(k_\varpi)$, so that $\mathrm{Ad}^*(\sigma_\varpi)X^* = -X^*$. It is possible that σ_ϖ is not k-rational. More precisely, by Section 4, we have that

$$W(G, T_\varpi) = \begin{cases} \{1, \sigma_\varpi\} & \operatorname{sgn}_\varpi(-1) = 1, \\ \{1\} & \operatorname{sgn}_\varpi(-1) = -1. \end{cases}$$

By (8.10),

$$(*) \qquad q^{-m/4} = q^{-(h-\operatorname{ord}(s))/2} = q^{-h/2}|s|^{-1/2}.$$

By Corollary 7.9, since $u = v$,

$$J^\varpi_{\nu^{1/2}}(u, v) = J^\varpi_{\nu^{1/2}\operatorname{sgn}_\epsilon}(u, v),$$

so (10.2) becomes

$$(\dagger) \quad M^G_{X^*}(Y) = \tfrac{1}{2}|s|^{-1/2}q^{-(h+1)/2}\big((1+\gamma_{\mathrm{ram}}(s)) + \gamma_{\mathrm{un}}(s)(1-\gamma_{\mathrm{ram}}(s))\big)J^\varpi_{\nu^{1/2}}(u, v).$$

It remains to compute $J^\varpi_{\nu^{1/2}}(u, v)$.

We will use Proposition 7.5, but, for simplicity, we want to avoid splitting into cases depending on whether or not $4 \mid m$. By (8.10), the restrictions to $k \setminus \wp^{h-1}$ of $\frac{1}{2}(1 + (-1)^h \operatorname{sgn}_\epsilon) = \frac{1}{2}(1 - \gamma_{\mathrm{un}})$ and $\frac{1}{2}(1 + \gamma_{\mathrm{un}})$ are characteristic functions that indicate whether $4 \mid m$ or $4 \nmid m$, respectively. (We omit $\wp^{h-1}$ because we are concerned with the case where $\mathrm{d}(Y) < r$, so that $\operatorname{ord}(s) < r - \frac{1}{2} = h - 1$.)

Thus, if $\operatorname{sgn}_\varpi(-1) = -1$, then combining Proposition 7.5, $(*)$, and (10.3) gives

$$(*_{\rm ns})\qquad J_{\nu^\alpha}(u, v) = q^{-h/2}|s|^{-1/2} \times \Big(\tfrac{1}{2}\big((1-\gamma_{\rm un}(s)) + (1+\gamma_{\rm un}(s))\gamma_{\rm ram}(s)\big) \times \Phi'_{\varpi^{h+1}}(2\varpi^{-h}s) \overset{(\S)}{+} \tfrac{1}{2}\big((1-\gamma_{\rm un}(s)) \overset{(\P)}{-} (1+\gamma_{\rm un}(s))\gamma_{\rm ram}(s)\big) \times \Phi'_{\varpi^{h+1}}(-2\varpi^{-h}s)\Big)$$
$$= \tfrac{1}{2}q^{-(h+1)/2}|s\theta'|^{-1/2} \times \Big(\big((1+\gamma_{\rm ram}(s)) - \gamma_{\rm un}(s)(1-\gamma_{\rm ram}(s))\big)\Phi(\langle X^*, Y\rangle) + \big((1-\gamma_{\rm ram}(s)) - \gamma_{\rm un}(s)(1+\gamma_{\rm ram}(s))\big)\Phi(\langle \operatorname{Ad}^*(\sigma_\varpi)X^*, Y\rangle)$$

and (changing the sign at (§), but not at (¶)) that

$$J_{\nu^\alpha \operatorname{sgn}_\varpi}(u, v) = \tfrac{1}{2}q^{-(h+1)/2}|s\theta'|^{-1/2} \times \Big(\big((1+\gamma_{\rm ram}(s)) - \gamma_{\rm un}(s)(1-\gamma_{\rm ram}(s))\big)\Phi(\langle X^*, Y\rangle) - \big((1-\gamma_{\rm ram}(s)) - \gamma_{\rm un}(s)(1+\gamma_{\rm ram}(s))\big)\Phi(\langle \operatorname{Ad}^*(\sigma_\varpi)X^*, Y\rangle)\Big),$$

so that

$$(\ddagger_{\rm ns})\qquad J^\varpi_{\nu^\alpha}(u, v) = \tfrac{1}{2}q^{-(h+1)/2}|s\theta'|^{-1/2}\big((1+\gamma_{\rm ram}(s)) - \gamma_{\rm un}(s)(1-\gamma_{\rm ram}(s))\big)\Phi(\langle X^*, Y\rangle).$$

Similarly, if $\operatorname{sgn}_\varpi(-1) = 1$, then (changing the sign at (¶), but not at (§), in $(*_{\rm ns})$) we obtain

$$(*_{\rm s})\qquad J_{\nu^\alpha}(u, v) = J_{\nu^\alpha \operatorname{sgn}_\varpi}(u, v) = \tfrac{1}{2}q^{-(h+1)/2}|s\theta'|^{-1/2}\big((1+\gamma_{\rm ram}(s)) - \gamma_{\rm un}(s)(1-\gamma_{\rm ram}(s))\big) \times \big(\Phi(\langle X^*, Y\rangle) + \Phi(\langle \operatorname{Ad}^*(\sigma_\varpi)X^*, Y\rangle)\big),$$

so that

$$(\ddagger_{\rm s})\qquad J^\varpi_{\nu^\alpha}(u, v) = J_{\nu^\alpha}(u, v) = \tfrac{1}{2}q^{-(h+1)/2}|s\theta'|^{-1/2}\big((1+\gamma_{\rm ram}(s)) - \gamma_{\rm un}(s)(1-\gamma_{\rm ram}(s))\big) \times \big(\Phi(\langle X^*, Y\rangle) + \Phi(\langle \operatorname{Ad}^*(\sigma_\varpi)X^*, Y\rangle)\big).$$

We may write $(\ddagger_{\rm ns})$ and $(\ddagger_{\rm s})$ uniformly as

$$(\ddagger)\qquad J^\varpi_{\nu^\alpha}(u, v) = \tfrac{1}{2}q^{-(h+1)/2}|s\theta'|^{-1/2}\big((1+\gamma_{\rm ram}(s)) - \gamma_{\rm un}(s)(1-\gamma_{\rm ram}(s))\big) \times \sum_{\sigma \in N_G(T_\varpi)/T_\varpi} \Phi(\langle \operatorname{Ad}^*(\sigma)X^*, Y\rangle).$$

Upon combining (†), (‡), and Lemma 5.9, we obtain the desired formula from

$$\big((1+\gamma_{\mathrm{ram}}(s))+\gamma_{\mathrm{un}}(s)(1-\gamma_{\mathrm{ram}}(s))\big)\cdot\big((1+\gamma_{\mathrm{ram}}(s))-\gamma_{\mathrm{un}}(s)(1-\gamma_{\mathrm{ram}}(s))\big)$$
$$= (1+\gamma_{\mathrm{ram}}(s))^2-\gamma_{\mathrm{un}}(s)^2(1-\gamma_{\mathrm{ram}}(s))^2 = 4\gamma_{\mathrm{ram}}(s) = 4\gamma_\Phi(X^*, Y)$$

(since $\gamma_{\mathrm{un}}(s)^2 = 1$). □

10B. *The bad shell.* We shall be concerned in this section with the behaviour of $M^G_{X^*}$ (hence $\hat\mu^G_{X^*}$, by Proposition 11.2) at the "bad shell", that is, on those regular, semisimple elements Y such that $\mathrm{d}(Y) = r$. We assume this is the case throughout the section. By (8.10), this implies that $m = 2$ and that $\mathrm{ord}(\theta')$ is odd, that is, Y belongs to a ramified torus. By Section 4, we can in fact assume $\mathrm{ord}(\theta') = 1$. Then, by Lemma 5.9,

$$\textbf{(10.6)}\qquad \mathrm{ord}(s) = h-1 \quad\Rightarrow\quad \mathrm{sgn}_\epsilon(s) = (-1)^{h-1} \text{ and } |s\theta'| = q^{-h}.$$

By Definition 6.5, the formula that holds in the situation of Theorem 10.9 holds also, suitably understood, in the situation of Theorem 10.8. We find it useful to separate them anyway.

Remark 10.7. In this section only, we need to name the specific ramified torus in which we are interested. We therefore assume in Theorems 10.8 and 10.9 that $X^* \in \mathfrak{t}^*_\varpi$. See Remark 6.9 for a discussion of how to handle other ramified tori.

Theorem 10.8. *If* $\mathrm{d}(X^*)+\mathrm{d}(Y) = 0$, *and* Y *lies in a ramified torus that is* not *stably conjugate to* $\boldsymbol{T}_\varpi$, *then*

$$M^G_{X^*}(Y) = \tfrac12 q^{-(h+1)}\cdot q^{-1/2}|D_{\mathfrak g}(Y)|^{-1/2}\sum_{Z\in(\mathfrak{t}_\varpi)_{r:r+}}\Phi(\langle X^*, Z\rangle)\,\mathrm{sgn}_\varpi(Y^2-Z^2),$$

where we identify the scalar matrices Y^2 *and* Z^2 *with elements of* k *in the natural way.*

Proof. By Section 4, it suffices to consider the case where $\theta' = \epsilon\varpi$.

By Corollary 7.9, since $\mathrm{ord}(u) = \mathrm{ord}(v)$,

$$J^\varpi_{\nu^{1/2}}(u, v) = J^\varpi_{\nu^{1/2}\,\mathrm{sgn}_\epsilon}(u, v),$$

and, by Corollary 7.8 and (8.8), $J_{\nu^{1/2}\,\mathrm{sgn}_\varpi}(u, v) = 0$, so

$$J^\varpi_{\nu^{1/2}}(u, v) = \tfrac12 J_{\nu^{1/2}}(u, v).$$

Hence, by (10.2) and (10.6),

$$(*)\qquad \begin{aligned} M^G_{X^*}(Y) &= \tfrac{1}{4}|s|^{-1/2}q^{-(h+1)/2} \\ &\qquad \times\bigl((1+\gamma_{\mathrm{ram}}(s)) - (-1)^h \operatorname{sgn}_\epsilon(s)(1-\gamma_{\mathrm{ram}}(s))\bigr) J_{\nu^{1/2}}(u, v) \\ &= \tfrac{1}{4}|s|^{-1/2}q^{-(h+1)/2}\cdot 2\cdot J_{\nu^{1/2}}(u, v) \\ &= \tfrac{1}{2}|s|^{-1/2}J_{\nu^{1/2}}(u, v). \end{aligned}$$

Finally, another application of Corollary 7.8, together with (8.9), gives that

$$J_{\nu^{1/2}}(u, v) = q^{-1}\sum_{c\in\wp^{-1}/R} \Phi'_{\varpi^{h+1}}(2c)\operatorname{sgn}_\varpi(c^2 - (\varpi^{-h}s)^2\epsilon).$$

Replacing c by $\varpi^{-h}c$ and using (10.6) again allows us to rewrite

$$(**)\qquad J_{\nu^{1/2}}(u, v) = q^{-(h+2)/2}|s\theta'|^{-1/2}\sum_{c\in\wp^{h-1}/\wp^h} \Phi(2\beta\varpi c)\operatorname{sgn}_\varpi(c^2 - s^2\epsilon).$$

By Definition 4.9, the isomorphism $c\mapsto c\cdot\sqrt{\varpi}$ of k with $\mathfrak{t}_\varpi$ identifies $\wp^{h-1}/\wp^h$ with $(\mathfrak{t}_\varpi)_{(h-1/2):(h+1/2)} = (\mathfrak{t}_\varpi)_{r:r+}$. If c is mapped to Z, then (by (8.5), for example) $2\beta\varpi c = \langle X^*, Z\rangle$, and

$$\operatorname{sgn}_\varpi(c^2 - s^2\epsilon) = \operatorname{sgn}_\varpi(s^2\epsilon\varpi - c^2\varpi) = \operatorname{sgn}_\varpi(Y^2 - Z^2).$$

Combining this with $(*)$, $(**)$, and Lemma 5.9 yields the desired formula. □

Theorem 10.9. *If* $\mathrm{d}(X^*) + \mathrm{d}(Y) = 0$, *and* $\widetilde{Y}$ *is a stable conjugate of* Y *that lies in a torus with* X^*, *then*

$$\begin{aligned} M^G_{X^*}(Y) &= \tfrac{1}{2}q^{-(h+1)}|D_{\mathfrak{g}}(Y)|^{-1/2} \\ &\quad\times\Biggl(\gamma_\Phi(X^*, Y)\sum_{\sigma\in W(\boldsymbol{G},\boldsymbol{T}_\varpi)} \Phi(\langle \mathrm{Ad}^*(\sigma)X^*, \widetilde{Y}\rangle) \\ &\qquad + q^{-1/2}\sum_{\substack{Z\in(\mathfrak{t}_\varpi)_{r:r+}\\ Z\neq\pm\widetilde{Y}}} \Phi(\langle X^*, Z\rangle)\operatorname{sgn}_\varpi(Y^2 - Z^2)\Biggr), \end{aligned}$$

where $\gamma_\Phi(X^*, Y)$ *is as in Definition 6.5.*

Proof. Implicit in the statement is the hypothesis that $\mathfrak{t} = \mathfrak{t}_{\theta'}$ is stably conjugate to $\mathfrak{t}_\varpi$, so that, by Section 4, we have $\theta' = x^2\varpi$ for some $x\in R^\times$. The proof proceeds much as in Theorem 10.8.

By (10.6) and Corollary 7.9, since $\operatorname{ord}(u) = \operatorname{ord}(v)$, (10.2) becomes

$$(*)\qquad M^G_{X^*}(Y) = |s|^{-1/2}q^{-(h+1)/2}J^{\varpi}_{\nu^{1/2}}(u, v).$$

By (8.7), we may take the square root w of uv to be $w = \varpi^{-h}xs$.

Combining Corollary 7.8 with (8.7), (8.9), and (10.6) gives

$$\begin{aligned}J_{\nu^\alpha}(u,v) &= q^{-1}\sum_{\substack{c\in\wp^{-1}/R\\ c\neq\pm\varpi^{-h}xs}} \Phi'_{\varpi^{h+1}}(2c)\,\operatorname{sgn}_\varpi(c^2-(\varpi^{-h}xs)^2)\\ &= q^{-1}\sum_{\substack{c\in\wp^{h-1}/\wp^h\\ c\neq\pm xs}} \Phi(2\beta\varpi c)\,\operatorname{sgn}_\varpi(c^2-x^2s^2)\\ &= q^{-(h+2)/2}|s\theta'|^{-1/2}\sum_{\substack{c\in\wp^{h-1}/\wp^h\\ c\neq\pm xs}} \Phi(2\beta\varpi c)\,\operatorname{sgn}_\varpi(c^2-x^2s^2).\end{aligned}$$

Note that $Y^2 = s^2\theta' = x^2s^2\varpi$, and that

$$\widetilde{Y} := xs\sqrt{\varpi} = \operatorname{Ad}\begin{pmatrix}\sqrt{x} & 0\\ 0 & \sqrt{x}^{-1}\end{pmatrix} Y$$

is a stable conjugate of Y that lies in $\mathfrak{t}_\varpi$. (Here, $\sqrt{\varpi}$ is an element of $\mathfrak{g}$, but $\sqrt{x}$ is an element of an extension field of k.) As in Theorem 10.8, if $Z = c\cdot\sqrt{\varpi}$, then $\langle X^*, Z\rangle = 2\beta\varpi c$ and $\operatorname{sgn}_\varpi(c^2-x^2s^2) = \operatorname{sgn}_\varpi(Y^2-Z^2)$. That is, upon using again the bijection $\wp^{h-1}/\wp^h \to (\mathfrak{t}_\varpi)_{r:r+}$ given by $c\mapsto c\cdot\sqrt{\varpi}$, we obtain

$$(**_1)\qquad J_{\nu^{1/2}}(u,v) = q^{-(h+2)/2}|s\theta'|^{-1/2}\sum_{\substack{Z\in(\mathfrak{t}_\varpi)_{r:r+}\\ Z\neq 0,\pm\widetilde{Y}}} \Phi(\langle X^*, Z\rangle)\,\operatorname{sgn}_\varpi(Y^2-Z^2).$$

Similarly, combining Corollary 7.8 with (8.9), Lemma 5.9, and (10.3) gives

$$\begin{aligned}(**_\varpi)\quad J_{\nu^{1/2}\operatorname{sgn}_\varpi}(u,v) &= q^{-1/2}\gamma_{\mathrm{ram}}(s)\big(\Phi(2\beta\varpi xs)+\Phi(-2\beta\varpi xs)\big)\\ &= q^{-(h+1)/2}|s\theta'|^{-1/2}\gamma_{\mathrm{ram}}(s)\sum_{\sigma\in W(\boldsymbol{G},\boldsymbol{T})}\Phi(\langle \operatorname{Ad}^*(\sigma)X^*, \widetilde{Y}\rangle).\end{aligned}$$

Combining $(*)$, $(**_1)$, $(**_\varpi)$, and Lemma 5.9 gives the desired formula. □

10C. *Close to zero.*

Theorem 10.10. *If* $\mathrm{d}(X^*)+\mathrm{d}(Y) > 0$, *and* X^* *is ramified, then let* $\gamma_\Phi(X^*, Y)$ *and* $c_0(X^*)$ *be as in Definitions 6.5 and 6.10, respectively. Then*

$$M^G_{X^*}(Y) = c_0(X^*) + q^{-(h+1)}|D_{\mathfrak{g}}(Y)|^{-1/2}\gamma_\Phi(X^*, Y).$$

Proof. By (8.10), $m < 2$.

By Proposition 8.13 and (6.7), using Notation 8.12 changes (10.2) into

$$\begin{aligned}&M^G_{X^*}(Y)\\ &= \tfrac12\Big(\tfrac12(1+\gamma_{\mathrm{ram}}(s))\big[Q_3(q^{-1/2})+\gamma_{\mathrm{ram}}(s)q^{-1};\ 1+\gamma_{\mathrm{ram}}(s)^{-1}\operatorname{sgn}_\varpi\big]\\ &\quad+\tfrac12(1-\gamma_{\mathrm{ram}}(s))\times\big[-Q_3(-q^{-1/2})+\gamma_{\mathrm{ram}}(s)q^{-1};\ (1-\gamma_{\mathrm{ram}}(s)^{-1}\operatorname{sgn}_\varpi)\operatorname{sgn}_\epsilon\big]\Big)(\theta').\end{aligned}$$

By (9.8) and the fact that

$$Q_3(q^{-1/2}) - Q_3(-q^{-1/2}) = -2T(T^2+1)\big|_{T=q^{-1/2}} = -2q^{-3/2}(q+1),$$

we may check case-by-case to see that this simplifies to

$$M^G_{X^*}(Y) = \big[-\tfrac{1}{2}q^{-3/2}(q+1);\ 1,\ 0,\ \gamma_{\mathrm{ram}}(s),\ 0\big](\theta'). \qquad \square$$

11. An integral formula

Our efforts so far have focused on computing the function $M^G_{X^*}$ of Definition 8.4, whereas we are really interested in the function $\hat{\mu}^G_{X^*}$ of Notation 5.7. We can now show that they are actually equal.

Lemma 11.1. *If $f \in L^1(G)$, then*

$$\int_G f(g)dg = \int_{k_\theta^\times} \int_k f\big(\varphi_\theta^{-1}(\alpha, t)\big)\,\mathrm{d}t\,\mathrm{d}^\times\alpha.$$

In Lemma 11.1, $\mathrm{d}g$, $\mathrm{d}t$, and $\mathrm{d}^\times\alpha$ are Haar measures on the obvious groups. Given any two of them, the third can be chosen so that the identity is satisfied. Since Definition 5.4 requires a measure on $G/C_G(X^*)$, not on G, we do not worry much here about normalisations (although a specific one is used in the proof).

Proof. With respect to the coordinate charts

$$(a, b, c) \mapsto \begin{pmatrix} a & b \\ c & (1+bc)/a \end{pmatrix}$$

(for $a \neq 0$) on G and

$$(a, b, t) \mapsto (a + b\sqrt{\theta}, t)$$

on $k_\theta^\times \times k$, the Jacobian of φ_θ at

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

(with $a \neq 0$) is $a^{-1}N_\theta(\alpha)$, where $\varphi_\theta(g) = (\alpha, t)$.

In particular, the Haar measure

$$|a|^{-1}\,\mathrm{d}a\,\mathrm{d}b\,\mathrm{d}c$$

on G is carried to the measure

$$|N_\theta(a + b\sqrt{\theta})|^{-1}\,\mathrm{d}a\,\mathrm{d}b\,\mathrm{d}t = |N_\theta(\alpha)|^{-1}\,\mathrm{d}\alpha\,\mathrm{d}t = \mathrm{d}^\times\alpha\,\mathrm{d}t$$

on $k_\theta^\times \times k$, as desired. $\square$

Proposition 11.2. *If $X^* \in \mathfrak{g}^*$ and $Y \in \mathfrak{g}$ are regular and semisimple*, *then*

$$\hat{\mu}^G_{X^*}(Y) = M^G_{X^*}(Y),$$

where $M^G_{X^*}$ is as in Definition 8.4, *and the Haar measure $\mathrm{d}\dot{g}$ on $G/C_G(X^*)$ of Notation 5.3 is normalised so that*

$$\mathrm{meas}_{\mathrm{d}\dot{g}}(\dot{K}) = \begin{cases} q^{-1}(q+1) & \text{for } X^* \text{ split}, \\ q^{-1}(q-1) & \text{for } X^* \text{ unramified}, \\ \frac{1}{2}q^{-2}(q^2-1) & \text{for } X^* \text{ ramified}, \end{cases}$$

where $\dot{K}$ is the image in $G/C_G(X^*)$ of $\mathrm{SL}_2(R)$.

Proof. We will maintain Notation 5.1. In particular, $X^* \in \mathfrak{t}^*_\theta$.

By the explicit formulae of the previous sections (specifically, Theorems 9.5, 9.6, 9.7, 10.4, 10.5, 10.8, 10.9, and 10.10), $M^G_{X^*} \in C^\infty(\mathfrak{g}^{\mathrm{rss}})$. This result plays the role of [Adler and DeBacker 2004, Corollary A.3.4]; we now imitate the proof of [ibid., Theorem A.1.2].

If $f \in C_c(\mathfrak{g}^{\mathrm{rss}})$, then there is a lattice $\mathscr{L} \subseteq \mathfrak{g}$ such that $f \cdot M^G_{X^*}$ is invariant under translation by $\mathscr{L}$. Then

$$\int_{\mathfrak{g}} f(Y) M^G_{X^*}(Y)\,\mathrm{d}Y = \mathrm{meas}_{\mathrm{d}Y}(\mathscr{L}) \sum_{Y \in \mathfrak{g}/\mathscr{L}} f(Y) \cdot \fint_{k^\times_\theta/C_\theta} \fint_k \Phi(\langle X^*, Y\rangle_{\alpha,t})\,\mathrm{d}t\,\mathrm{d}^\times\dot{\alpha}.$$

Since the sum is finitely supported, we may bring it inside the integral. By (8.5) and Definition 5.5,

$$(*) \qquad \begin{aligned} &\int_{\mathfrak{g}} f(Y) M^G_{X^*}(Y)\,\mathrm{d}Y \\ &= \fint_{k^\times_\theta/C_\theta} \fint_k \mathrm{meas}_{\mathrm{d}Y}(\mathscr{L}) \sum_{Y \in \mathfrak{g}/\mathscr{L}} f(Y) \Phi\big(\langle \mathrm{Ad}^*(\varphi_\theta^{-1}(\alpha,t))X^*, Y\rangle\big)\,\mathrm{d}t\,\mathrm{d}^\times\dot{\alpha} \\ &= \fint_{k^\times_\theta/C_\theta} \fint_k \int_{\mathfrak{g}} f(Y) \Phi\big(\langle \mathrm{Ad}^*(\varphi_\theta^{-1}(\alpha,t))X^*, Y\rangle\big)\,\mathrm{d}Y\,\mathrm{d}t\,\mathrm{d}^\times\dot{\alpha} \\ &= \fint_{k^\times_\theta/C_\theta} \fint_k \hat{f}\big(\mathrm{Ad}^*(\varphi_\theta^{-1}(\alpha,t))X^*\big)\,\mathrm{d}t\,\mathrm{d}^\times\dot{\alpha}, \end{aligned}$$

where φ_θ is as in Definition 8.2.

On the other hand, again by Definition 5.5,

$$\hat{\mu}^G_{X^*}(f) := \mu^G_{X^*}(\hat{f}) = \int_{G/T_\theta} \hat{f}\big(\mathrm{Ad}^*(g)X^*\big)\,\mathrm{d}\dot{g} = \int_{\overline{U}\backslash G/T_\theta} \int_{\overline{U}} \hat{f}\big(\mathrm{Ad}^*(\bar{u}g)X^*\big)\,\mathrm{d}\bar{u}\,\mathrm{d}\ddot{g},$$

where

$$\overline{U} = \left\{ \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} : b \in k \right\}.$$

By Lemmata 11.1 and 8.3, and (∗), if $\mathrm{d}\dot{g}$ is properly normalised, then

$$\hat{\mu}^G_{X^*}(f) = \int_{k_\theta^\times/C_\theta} \int_k \hat{f}\big(\mathrm{Ad}^*(\varphi_\theta^{-1}(\alpha, t))X^*\big)\,\mathrm{d}t\,\mathrm{d}^\times\dot{\alpha} = \int_{\mathfrak{g}} f(Y) M^G_{X^*}(Y)\,\mathrm{d}Y.$$

It remains only to compute the normalisation of $\mathrm{d}\dot{g}$. We do so case-by-case. If X^* is split, so that we may take $\theta = 1$, then the image under φ_1 of

$$(1 + \wp_1) \times \wp \subseteq k_1^\times \times k$$

is precisely the kernel K_+ of the (component-wise) reduction map $\mathrm{SL}_2(R) \to \mathrm{SL}_2(\mathfrak{f})$. Here, we have written $1 + \wp_1 = \{(a, b) \in k_1 : a \in 1 + \wp, b \in \wp\}$. Thus,

$$(1 + \wp_1)C_1/C_1 \times \wp \xrightarrow{\sim} K_+T_1/T_1.$$

Now $N_1 : 1 + \wp_1 \to 1 + \wp$ is surjective, so by Definitions 2.1 and 8.4, the measure (in $k_1/C_1 \times k$) of the domain is

$$\mathrm{meas}_{\mathrm{d}^\times x}(1 + \wp) \cdot \mathrm{meas}_{\mathrm{d}x}(\wp) = q^{-2}.$$

Since $\dot{K} = \mathrm{SL}_2(R)T_1/T_1$ is tiled by

$$\big[\mathrm{SL}_2(R)T_1 \colon K_+T_1\big] = \big[\mathrm{SL}_2(R) \colon K_+(T_1 \cap \mathrm{SL}_2(R))\big] = \big[\mathrm{SL}_2(\mathfrak{f}) \colon \mathsf{T}_1(\mathfrak{f})\big] = q(q+1)$$

copies of K_+T_1/T_1, where

$$\mathsf{T}_1 := \left\{\begin{pmatrix} a & b \\ b & a\end{pmatrix} : a^2 - b^2 = 1\right\},$$

we see that, in this case, $\mathrm{d}\dot{g}$ assigns $\dot{K}$ measure $q^{-2} \cdot q(q+1) = q^{-1}(q+1)$.

The remaining cases are easier, since C_θ is contained in the ring R_θ of integers in k_θ, and (for our choices of θ) T_θ is contained in $\mathrm{SL}_2(R)$. If X^* is unramified, so that we may take $\theta = \epsilon$, then the image under φ_ϵ of $R_\epsilon^\times \times R$ is precisely $\mathrm{SL}_2(R)$. Since $N_\epsilon : R_\epsilon^\times \to R^\times$ is surjective, we see that, in this case, $\mathrm{d}\dot{g}$ assigns $\dot{K}$ measure $\mathrm{meas}_{\mathrm{d}^\times x}(R_\epsilon^\times) \cdot \mathrm{meas}_{\mathrm{d}x}(R) = q^{-1}(q-1)$.

If X^* is ramified, so that we may take $\theta = \varpi$, then the image under φ_ϖ of $R_\varpi^\times \times \wp$ is precisely the Iwahori subgroup $\mathcal{I}$, that is, the preimage in $\mathrm{SL}_2(R)$ of

$$\mathsf{B}(\mathfrak{f}) := \left\{\begin{pmatrix} a & b \\ 0 & a\end{pmatrix} : a \in \mathfrak{f}^\times, b \in \mathfrak{f}\right\}$$

under the reduction map $\mathrm{SL}_2(R) \to \mathrm{SL}_2(\mathfrak{f})$. Since $N_\varpi : R_\varpi^\times \to R^\times$ has cokernel of order 2, we see that, in this case, $\mathrm{d}\dot{g}$ assigns $\dot{K}$ measure

$$\tfrac{1}{2}\,\mathrm{meas}_{\mathrm{d}^\times x}(R^\times) \cdot \mathrm{meas}_{\mathrm{d}x}(\wp) \cdot \big[\mathrm{SL}_2(\mathfrak{f}) \colon \mathsf{B}\big] = \tfrac{1}{2}q^{-2}(q^2 - 1). \qquad \square$$

Thus, all the results we have proven for $M^G_{X^*}$ are actually results about $\hat{\mu}^G_{X^*}$. We close by summarising some results that can be stated in a fairly uniform fashion (that is, mostly independent of the "type" of X^*, in the sense of Definition 4.4). This theorem does *not* cover everything we have shown about Fourier transforms of semisimple orbital integrals (in particular, it says nothing about the behaviour of ramified orbital integrals on the "bad shell", as in Section 10B); for that, the reader should refer to the detailed results of Sections 9–10.

Theorem 11.3. *If* $\mathrm{d}(X^*) + \mathrm{d}(Y) < 0$ (*or* $\mathrm{d}(X^*) + \mathrm{d}(Y) \le 0$ *and* X^* *is split or unramified*), *then*

$$\hat{\mu}^G_{X^*}(Y) = q^{-(r'+1)} |D_{\mathfrak{g}}(Y)|^{-1/2} \gamma_\Phi(X^*, Y) \sum_{\sigma \in W(G,T)} \Phi(\langle \mathrm{Ad}^*(\sigma) X^*, Y \rangle)$$

if X^* *and* Y *lie in a common torus* $\boldsymbol{T}$ (*with* $T = \boldsymbol{T}(k)$), *and*

$$\hat{\mu}^G_{X^*}(Y) = 0$$

if X^* *and* Y *do not lie in* G*-conjugate tori. Here*, r' *is as in Notation 5.2*, *and* $\gamma_\Phi(X^*, Y)$ *is as in Definition 6.5.*

If $\mathrm{d}(X^*) + \mathrm{d}(Y) > 0$, *then*

$$\hat{\mu}^G_{X^*}(Y) = c_0(X^*) + q^{-(r'+1)} |D_{\mathfrak{g}}(Y)|^{-1/2} \gamma_\Phi(X^*, Y).$$

Here, $\gamma_\Phi(X^*, Y)$ *and* $c_0(X^*)$ *are as in Definitions 6.5 and 6.10*, *respectively.*

Proof. This is an amalgamation of parts of Theorems 9.5, 9.6, 9.7, 10.4, 10.5, and 10.10 and Proposition 11.2. □

Acknowledgements

This paper, and the paper [Adler et al. 2011] that follows it, would not have been possible without the advice and guidance of Paul J. Sally, Jr. It is a pleasure to thank him, as well as Stephen DeBacker and Jeffrey D. Adler, both of whom offered useful suggestions regarding this paper.

References

[Adler 1998] J. D. Adler, "Refined anisotropic K-types and supercuspidal representations", *Pacific J. Math.* **185**:1 (1998), 1–32. MR 2000f:22019 Zbl 0924.22015

[Adler and DeBacker 2002] J. D. Adler and S. M. DeBacker, "Some applications of Bruhat–Tits theory to harmonic analysis on the Lie algebra of a reductive p-adic group", *Michigan Math. J.* **50**:2 (2002), 263–286. MR 2003g:22016 Zbl 1018.22013

[Adler and DeBacker 2004] J. D. Adler and S. M. DeBacker, "Murnaghan–Kirillov theory for supercuspidal representations of tame general linear groups", *J. Reine Angew. Math.* **575** (2004), 1–35. MR 2005j:22008 Zbl 1054.22014

[Adler and Spice 2009] J. D. Adler and L. R. Spice, "Supercuspidal characters of reductive p-adic groups", *Amer. J. Math.* **131**:4 (2009), 1137–1210. MR 2011a:22018 Zbl 1173.22012 arXiv 0707.3313

[Adler et al. 2011] J. D. Adler, S. M. DeBacker, P. J. Sally, Jr., and L. R. Spice, "Supercuspidal characters of SL_2 over a p-adic field", pp. 19–69 in *Harmonic analysis on reductive, p-adic groups* (San Francisco, CA, 2010), edited by R. S. Doran et al., Contemp. Math. **543**, Amer. Math. Soc., Providence, RI, 2011. MR 2798422 Zbl 1222.22007

[Baruch 1997] E. M. Baruch, "On Bessel distributions for GL_2 over a p-adic field", *J. Number Theory* **67**:2 (1997), 190–202. MR 98k:22042 Zbl 0887.22023

[Baruch 2001] E. M. Baruch, "On Bessel distributions for quasi-split groups", *Trans. Amer. Math. Soc.* **353**:7 (2001), 2601–2614. MR 2002b:22029 Zbl 0976.22015

[Baruch 2003] E. M. Baruch, "Bessel functions for GL(3) over a p-adic field", *Pacific J. Math.* **211**:1 (2003), 1–34. MR 2004k:22022 Zbl 1055.22013

[Baruch 2004] E. M. Baruch, "Bessel distributions for GL(3) over the p-adics", *Pacific J. Math.* **217**:1 (2004), 11–27. MR 2005h:22025 Zbl 1064.22008

[Baruch 2005] E. M. Baruch, "Bessel functions for $\mathrm{GL}(n)$ over a p-adic field", pp. 1–40 in *Automorphic representations, L-functions and applications: progress and prospects* (Columbus, OH, 2003), edited by J. W. Cogdell et al., Ohio State Univ. Math. Res. Inst. Publ. **11**, de Gruyter, Berlin, 2005. MR 2007d:22011 Zbl 1103.22006

[Bushnell and Kutzko 1993a] C. J. Bushnell and P. C. Kutzko, *The admissible dual of* $\mathrm{GL}(N)$ *via compact open subgroups*, Annals of Mathematics Studies **129**, Princeton University Press, Princeton, NJ, 1993. MR 94h:22007 Zbl 0787.22016

[Bushnell and Kutzko 1993b] C. J. Bushnell and P. C. Kutzko, "The admissible dual of $\mathrm{SL}(N)$, I", *Ann. Sci. École Norm. Sup.* (4) **26**:2 (1993), 261–280. MR 94a:22033 Zbl 0787.22017

[Bushnell and Kutzko 1994] C. J. Bushnell and P. C. Kutzko, "The admissible dual of $\mathrm{SL}(N)$, II", *Proc. London Math. Soc.* (3) **68**:2 (1994), 317–379. MR 94k:22035 Zbl 0801.22011

[Corwin 1989] L. Corwin, "The unitary dual for the multiplicative group of arbitrary division algebras over local fields", *J. Amer. Math. Soc.* **2**:3 (1989), 565–598. MR 90d:22021 Zbl 0706.22013

[Corwin et al. 1995] L. Corwin, A. Moy, and P. J. Sally, Jr., "Supercuspidal character formulas for GL_ℓ", pp. 1–11 in *Representation theory and harmonic analysis: conference in honor of Ray A. Kunze* (Cincinnati, OH, 1994), edited by T. Ton-That et al., Contemp. Math. **191**, Amer. Math. Soc., Providence, RI, 1995. MR 96m:22037 Zbl 0839.22022

[Cunningham and Gordon 2009] C. Cunningham and J. Gordon, "Motivic proof of a character formula for SL(2)", *Experiment. Math.* **18**:1 (2009), 11–44. MR 2010i:22014 Zbl 1179.22018

[DeBacker 1997] S. M. DeBacker, *On supercuspidal characters of* GL_ℓ*, ℓ a prime*, thesis, The University of Chicago, 1997.

[DeBacker 2002] S. M. DeBacker, "Homogeneity results for invariant distributions of a reductive p-adic group", *Ann. Sci. École Norm. Sup.* (4) **35**:3 (2002), 391–422. MR 2003i:22019 Zbl 0999.22013

[DeBacker and Reeder 2009] S. M. DeBacker and M. Reeder, "Depth-zero supercuspidal L-packets and their stability", *Ann. of Math.* (2) **169**:3 (2009), 795–901. MR 2010d:22023 Zbl 1193.11111

[DeBacker and Sally 2000] S. M. DeBacker and P. J. Sally, Jr., "Germs, characters, and the Fourier transforms of nilpotent orbits", pp. 191–221 in *The mathematical legacy of Harish-Chandra* (Baltimore, MD, 1998), edited by R. S. Doran and V. S. Varadarajan, Proc. Sympos. Pure Math. **68**, Amer. Math. Soc., Providence, RI, 2000. MR 2001i:22022 Zbl 0960.22017

[Dwork 1974] B. Dwork, "Bessel functions as p-adic functions of the argument", *Duke Math. J.* **41**:4 (1974), 711–738. MR 52 #8124 Zbl 0302.14008

[Gindikin and Karpelevič 1962] S. G. Gindikin and F. I. Karpelevič, "Plancherel measure for symmetric Riemannian spaces of non-positive curvature", *Dokl. Akad. Nauk SSSR* **145** (1962), 252–255. In Russian; translated in *Sov. Math* **3** (1962), 962–965. MR 27 #240 Zbl 0168.27503

[Hales 1994] T. C. Hales, "Hyperelliptic curves and harmonic analysis: why harmonic analysis on reductive p-adic groups is not elementary", pp. 137–169 in *Representation theory and analysis on homogeneous spaces* (New Brunswick, NJ, 1993), edited by S. Gindikin et al., Contemp. Math. **177**, Amer. Math. Soc., Providence, RI, 1994. MR 96d:22024 Zbl 0834.22017

[Harish-Chandra 1966] Harish-Chandra, "Discrete series for semisimple Lie groups, II: Explicit determination of the characters", *Acta Math.* **116**:1 (1966), 1–111. MR 36 #2745 Zbl 0199.20102

[Harish-Chandra 1970] Harish-Chandra, *Harmonic analysis on reductive p-adic groups*, Lecture Notes in Mathematics **162**, Springer, Berlin, 1970. MR 54 #2889 Zbl 0202.41101

[Harish-Chandra 1999] Harish-Chandra, *Admissible invariant distributions on reductive p-adic groups*, University Lecture Series **16**, Amer. Math. Soc., Providence, RI, 1999. MR 2001b:22015 Zbl 0928.22017

[Howe 1971] R. E. Howe, "Representation theory for division algebras over local fields: tamely ramified case", *Bull. Amer. Math. Soc.* **77**:6 (1971), 1063–1066. MR 44 #4146 Zbl 0223.22009

[Howe 1973] R. E. Howe, "Two conjectures about reductive p-adic groups", pp. 377–380 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 #3044 Zbl 0284.22004

[Kim and Murnaghan 2003] J.-L. Kim and F. Murnaghan, "Character expansions and unrefined minimal K-types", *Amer. J. Math.* **125**:6 (2003), 1199–1234. MR 2004k:22024 Zbl 1037.22035

[Kim and Murnaghan 2006] J.-L. Kim and F. Murnaghan, "K-types and Γ-asymptotic expansions", *J. Reine Angew. Math.* **592** (2006), 189–236. MR 2007m:22016 Zbl 1089.22011

[Langlands 1979] R. P. Langlands, "Stable conjugacy: definitions and lemmas", *Canad. J. Math.* **31**:4 (1979), 700–725. MR 82j:10054 Zbl 0421.12013

[Langlands 2011] R. P. Langlands, "Singularités et transfert", 2011, Available at http://publications.ias.edu/node/139. To appear in *Annales des Sciences Mathématiques du Québec*.

[Lidl and Niederreiter 1997] R. Lidl and H. Niederreiter, *Finite fields*, 2nd ed., Encyclopedia of Mathematics and its Applications **20**, Cambridge University Press, Cambridge, 1997. MR 97i:11115 Zbl 0866.11069

[Morris 1992] L. Morris, "Tamely ramified supercuspidal representations of classical groups, II: Representation theory", *Ann. Sci. École Norm. Sup.* (4) **25**:3 (1992), 233–274. MR 93h:22032 Zbl 0782.22012

[Moy 1986] A. Moy, "Local constants and the tame Langlands correspondence", *Amer. J. Math.* **108**:4 (1986), 863–930. MR 88b:11081 Zbl 0597.12019

[Moy and Prasad 1994] A. Moy and G. Prasad, "Unrefined minimal K-types for p-adic groups", *Invent. Math.* **116**:1 (1994), 393–408. MR 95f:22023 Zbl 0804.22008

[Moy and Prasad 1996] A. Moy and G. Prasad, "Jacquet functors and unrefined minimal K-types", *Comment. Math. Helv.* **71**:1 (1996), 98–121. MR 97c:22021 Zbl 0860.22006

[Murnaghan 1995a] F. Murnaghan, "Characters of supercuspidal representations of $\mathrm{SL}(n)$", *Pacific J. Math.* **170**:1 (1995), 217–235. MR 96k:22030 Zbl 0848.22022

[Murnaghan 1995b] F. Murnaghan, "Local character expansions for supercuspidal representations of $\mathrm{U}(3)$", *Canad. J. Math.* **47**:3 (1995), 606–640. MR 96i:22026 Zbl 0909.22030

[Murnaghan 1996a] F. Murnaghan, "Characters of supercuspidal representations of classical groups", *Ann. Sci. École Norm. Sup.* (4) **29**:1 (1996), 49–105. MR 98c:22016 Zbl 0856.22025

[Murnaghan 1996b] F. Murnaghan, "Local character expansions and Shalika germs for $\mathrm{GL}(n)$", *Math. Ann.* **304**:1 (1996), 423–455. MR 98b:22020 Zbl 0841.22008

[Murnaghan 2000] F. Murnaghan, "Germs of characters of admissible representations", pp. 501–515 in *The mathematical legacy of Harish-Chandra* (Baltimore, MD, 1998), edited by R. S. Doran and V. S. Varadarajan, Proc. Sympos. Pure Math. **68**, Amer. Math. Soc., Providence, RI, 2000. MR 2001i:22023 Zbl 0957.22020

[Sally and Shalika 1968] P. J. Sally, Jr. and J. A. Shalika, "Characters of the discrete series of representations of SL(2) over a local field", *Proc. Nat. Acad. Sci. USA* **61**:4 (1968), 1231–1237. MR 38 #5994 Zbl 0198.18203

[Sally and Shalika 1969] P. J. Sally, Jr. and J. A. Shalika, "The Plancherel formula for SL(2) over a local field", *Proc. Nat. Acad. Sci. USA* **63**:3 (1969), 661–667. MR 51 #813 Zbl 0201.17703

[Sally and Shalika 1984] P. J. Sally, Jr. and J. A. Shalika, "The Fourier transform of orbital integrals on SL_2 over a p-adic field", pp. 303–340 in *Lie group representations, II* (College Park, MD, 1982/1983), edited by R. Herb et al., Lecture Notes in Math. **1041**, Springer, Berlin, 1984. MR 86a:22017 Zbl 0542.22017

[Sally and Spice 2009] P. J. Sally, Jr. and L. R. Spice, "Character theory or reductive p-adic groups", pp. 103–111 in *Ottawa lectures on admissible representations of reductive p-adic groups*, edited by C. Cunningham and M. Nevins, Fields Inst. Monogr. **26**, Amer. Math. Soc., Providence, RI, 2009. MR 2010b:22012 Zbl 1176.22016

[Sally and Taibleson 1966] P. J. Sally, Jr. and M. H. Taibleson, "Special functions on locally compact fields", *Acta Math.* **116**:1 (1966), 279–309. MR 34 #6168 Zbl 0173.07005

[Schmid 1968] W. Schmid, "Homogeneous complex manifolds and representations of semisimple Lie groups", *Proc. Nat. Acad. Sci. USA* **59**:1 (1968), 56–59. MR 37 #1520 Zbl 0164.15803

[Shalika 2004] J. A. Shalika, "Representation of the two by two unimodular group over local fields", pp. 1–38 in *Contributions to automorphic forms, geometry, and number theory* (Baltimore, MD, 2002), edited by H. Hida et al., Johns Hopkins University Press, Baltimore, MD, 2004. MR 2005f: 22028 Zbl 1076.11032

[Spice 2005] L. R. Spice, "Supercuspidal characters of SL_l over a p-adic field, l a prime", *Amer. J. Math.* **127**:1 (2005), 51–100. MR 2005k:22028 Zbl 1059.22020

[Stevens 2008] S. Stevens, "The supercuspidal representations of p-adic classical groups", *Invent. Math.* **172**:2 (2008), 289–352. MR 2010e:22008 Zbl 1140.22016 arXiv math.RT/0607622

[Takahashi 2003] T. Takahashi, "Formulas for tamely ramified supercuspidal characters of GL_3", *Trans. Amer. Math. Soc.* **355**:2 (2003), 567–591. MR 2003i:22020 Zbl 1033.22005

[Takahashi 2005] T. Takahashi, "On some constants in the supercuspidal characters of GL_l, l a prime $\neq p$", *Trans. Amer. Math. Soc.* **357**:6 (2005), 2509–2526. MR 2006d:22027 Zbl 1065.22014

[Tauvel and Yu 2005] P. Tauvel and R. W. T. Yu, *Lie algebras and algebraic groups*, Springer, Berlin, 2005. MR 2006c:17001 Zbl 1068.17001

[Trimble 1994] C. E. Trimble, "Finite ring sums from p-adic K-Bessel functions", pp. 369–378 in *Finite fields: theory, applications, and algorithms: Proceedings of the Second International Conference* (Las Vegas, NV, 1993), edited by G. L. Mullen and P. J.-S. Shiue, Contemp. Math. **168**, Amer. Math. Soc., Providence, RI, 1994. MR 96a:11134 Zbl 0835.11044

[Waldspurger 1995] J.-L. Waldspurger, "Une formule des traces locale pour les algèbres de Lie p-adiques", *J. Reine Angew. Math.* **465** (1995), 41–99. MR 96i:22039 Zbl 0829.11030

[Yu 2001] J.-K. Yu, "Construction of tame supercuspidal representations", *J. Amer. Math. Soc.* **14**:3 (2001), 579–622. MR 2002f:22033 Zbl 0971.22012

Received October 7, 2010.

LOREN SPICE
DEPARTMENT OF MATHEMATICS
TEXAS CHRISTIAN UNIVERSITY
TCU BOX 298900
2840 W. BOWIE ST
FORT WORTH, TX 76109
UNITED STATES

l.spice@tcu.edu
http://faculty.tcu.edu/lspice

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

ON NONCOMPACT τ-QUASI-EINSTEIN METRICS

LIN FENG WANG

In this paper, we will study the τ-quasi-Einstein metrics on complete noncompact Riemannian manifolds and get a rigid property. We will also obtain lower and upper estimates for scalar curvatures on these metrics by using the maximum principle.

1. Introduction

For a given smooth potential function f, the τ-Bakry–Émery Ricci curvature tensor

$$\mathrm{Ric}_{f,\tau} = \mathrm{Ric} + \mathrm{Hess}\, f - \frac{\nabla f \otimes \nabla f}{\tau}$$

is always used to replace the Ricci curvature tensor when one tries to study the weighted measure $d\mu = e^{-f} dx$, where $0 < \tau \leq +\infty$ and dx is the Riemann–Lebesgue measure determined by the metric. There has been an active interest in the study of the weighted measure under some conditions about the τ-Bakry–Émery Ricci curvature tensor; see [Li 2005; Wang 2010] and the references therein.

According to [Kim and Kim 2003; Case 2010; Case et al. 2011; Wang 2011], we call a metric g τ-quasi-Einstein with potential function f, if for some constant λ,

$$\mathrm{Ric} + \mathrm{Hess}\, f - \frac{\nabla f \otimes \nabla f}{\tau} = \lambda g, \tag{1-1}$$

where $0 < \tau \leq +\infty$. A τ-quasi-Einstein metric becomes an Einstein metric when the potential function f is constant. We note that an ∞-quasi-Einstein metric indicates a gradient Ricci soliton. As in [Hamilton 1995; Perelman 2002; Cao and Zhu 2006], a gradient Ricci soliton is shrinking, steady or expanding when $\lambda > 0$, $\lambda = 0$ or $\lambda < 0$, respectively.

For a positive integer τ, the τ-quasi-Einstein metric is closely relative to the existence of warped product Einstein manifolds [Besse 1987; Case 2010; Case et al. 2011]. Let (M, g) and (N^τ, h) be two Riemannian manifolds. Then, for

The author was supported in part by the NSF of China(10871070, 10971066, 11171254).
MSC2010: primary 53C21; secondary 53C25.
Keywords: τ-quasi-Einstein metric, rigidity properties, maximum principle, potential function.

some potential function f on M, the warped product manifold $(M \times N, \tilde{g})$ with product metric

$$\tilde{g} = g \oplus \exp\left(-\frac{2f}{\tau}\right)h$$

is Einstein if and only if (N^τ, h) is Einstein and the Ricci curvature tensor of M satisfies the quasi-Einstein equation (1-1) for some constant λ.

It was proved in [Qian 1997; Wei and Wylie 2007] that a manifold with a τ-quasi-Einstein metric (τ is finite) is automatically compact when $\lambda > 0$. It was also proved in [Ivey 1993] that any expanding or steady gradient Ricci solitons on closed manifolds should be trivial. The same rigid properties for the τ-quasi-Einstein metrics on closed manifolds were proved in [Kim and Kim 2003; Wang 2011]. But for the τ-quasi-Einstein metrics on closed manifolds with $\lambda > 0$, the rigid properties rely on the constant μ which appears in the following identity:

$$R + \frac{\tau - 1}{\tau}|\nabla f|^2 + (\tau - n)\lambda = \mu e^{2f/\tau}, \tag{1-2}$$

where R is the scalar curvature. This identity was proved in [Kim and Kim 2003]. See also [Wang 2011], where the author proved that the quasi-Einstein metrics with $\lambda > 0$ should be trivial when $\mu \leq 0$. In fact, the authors of [Lü et al. 2004] constructed nontrivial τ-quasi-Einstein metrics with $\lambda > 0$ and $\tau > 1$, which also satisfy $\mu > 0$.

In this paper, we will study the τ-quasi-Einstein metrics on complete noncompact Riemannian manifolds with $\lambda \leq 0$. Our first result is Theorem 1.1, which is about the rigidity.

Theorem 1.1. *Let M be a complete noncompact Riemannian manifold and g a τ-quasi-Einstein metric on M with potential function f and $\lambda \leq 0$ a constant. If*

$$R_0^{-2} \int_{B_{2R_0} \setminus B_{R_0}} |\nabla f|^2 \exp\left(-\frac{\tau + 2}{\tau} f\right) dx \to 0 \tag{1-3}$$

as $R_0 \to \infty$, where B_{R_0} denotes the geodesic ball centered at a fixed point $O \in M$ with radius R_0, then e^f is a harmonic function on M, that is, $\Delta e^f = 0$. Moreover, if $\lambda < 0$, then g is trivial in the sense that f is constant.

The following theorem for gradient Ricci solitons was proved in [Zhang 2009]. In fact, part 1 is a consequence of [Chen 2009, Corollary 2.5].

Theorem 1.2. *Let (M^n, g) be a complete noncompact gradient Ricci soliton with potential function f and soliton constant λ.*

(1) *If the gradient Ricci soliton is shrinking or steady, then $R \geq 0$.*

(2) *If the gradient Ricci soliton is expanding, then there exists a positive constant $C(n)$ such that $R \geq C(n)\lambda$.*

Zhang [2011] pointed out that $R \geq n\lambda$ is right in Theorem 1.2(2). The lower bound estimates for scalar curvatures play important roles in the study of geometric properties of gradient Ricci solitons. Based on these estimates, compactness theorems for gradient Ricci solitons were proved in [Zhang 2006] and some results about the volume growth for noncompact gradient Ricci solitons were deduced in [Cao 2009; Cao and Zhou 2010; Munteanu 2009; Zhang 2011].

In [Case et al. 2011], the authors got estimates for R on closed τ-quasi-Einstein metrics. Later, Wang [2011] studied the lower bound estimate for scalar curvature R on complete noncompact τ-quasi-Einstein metrics with $\lambda \leq 0$. We state this result as follows.

Theorem 1.3. *Let M be an n-dimensional complete noncompact Riemannian manifold, metric g is τ-quasi-Einstein with potential function f and constant $\lambda \leq 0$, where $\tau \geq 1$. If $\mu \leq 0$ or $\mu > 0$ and f is bounded from above by a constant C, then*

$$R(y) \geq n\lambda \tag{1-4}$$

for any $y \in M$.

The proof of this theorem in [Wang 2011] relies on a gradient estimate of f, this gradient estimate shows that $|\nabla f|^2$ is bounded from above if $\mu \leq 0$ or $\mu > 0$ and f is bounded from above by a constant C. We will give a nontrivial τ-quasi-Einstein metric with $\lambda < 0$, but f is not bounded from above; see Example 2.1. The second main result of this paper is to improve Theorem 1.3. That is to say, we will show that the lower estimate (1-4) is always right for τ-quasi-Einstein metrics with $\lambda \leq 0$.

Theorem 1.4. *Let M be an n-dimensional complete noncompact Riemannian manifold, g be a τ-quasi-Einstein metric with potential function f and $\lambda \leq 0$ be a constant, where $\tau > 0$. Then* (1-4) *holds for any $y \in M$.*

Remark 1.5. If $\tau = \infty$, we recover the lower bound estimate for R on a complete noncompact steady or expanding gradient Ricci soliton given in [Zhang 2011].

It remains interesting to find out whether R is bounded from above by a constant for noncompact quasi-Einstein metrics. The following theorem states that the scalar curvature of a quasi-Einstein metric with $\lambda \leq 0$ is bounded from above if $\mu \leq 0$.

Theorem 1.6. *Let g be a τ-quasi-Einstein metric with $\lambda \leq 0$ and $\mu \leq 0$. Then*

$$R(y) \leq (n - \max\{\tau, 1\})\lambda \tag{1-5}$$

for any $y \in M$.

2. Examples of quasi-Einstein metrics

In this section, we assume that $M = \mathbb{R} \times N^{n-1}$ is a warped product manifold with the product metric given by

$$ds_M^2 = dt^2 + \varphi^2(t)\, ds_N^2,$$

where ds_N^2 is a fixed metric on N and φ is a positive function on $\mathbb{R}$. Consider the orthonormal coframe $\{\theta_\alpha : 2 \le \alpha \le n\}$ on N^{n-1}; then

$$\{\omega_1 = dt,\, \omega_\alpha = \varphi(t)\theta_\alpha : 2 \le \alpha \le n\}$$

is an orthonormal coframe on M^n. We use $R_{M,ijkl}$ and $R_{N,\alpha\beta\gamma\delta}$ to denote the Riemannian curvature tensors of M and N respectively. After the same calculation as in [O'Neill 1983; Wang 2011], we conclude that

$$R_{M,1\alpha ij} = \begin{cases} -(\log\varphi(t))'' - ((\log\varphi(t))')^2 & \text{if } i = 1,\ j = \alpha, \\ (\log\varphi(t))'' + ((\log\varphi(t))')^2 & \text{if } i = \alpha,\ j = 1, \\ 0 & \text{otherwise} \end{cases} \tag{2-1}$$

and

$$R_{M,\alpha\beta ij} = \begin{cases} \varphi^{-2}(t)R_{N,\alpha\beta\gamma\theta} + ((\log\varphi(t))')^2(\delta_{\alpha\theta}\delta_{\beta\gamma} - \delta_{\alpha\gamma}\delta_{\beta\theta}) & \text{if } i = \gamma,\ j = \theta, \\ 0 & \text{otherwise.} \end{cases} \tag{2-2}$$

If we use $R_{N,\alpha\beta}$ to denote the Ricci curvature tensor on N, by (2-1) and (2-2), the Ricci curvature tensor of M can be expressed as

$$R_{M,1i} = -(n-1)\big((\log\varphi(t))'' + ((\log\varphi(t))')^2\big)\delta_{1i}, \tag{2-3}$$

$$R_{M,\alpha\beta} = \varphi^{-2}(t)R_{N,\alpha\beta} - \big((\log\varphi(t))'' + (n-1)((\log\varphi(t))')^2\big)\delta_{\alpha\beta}. \tag{2-4}$$

Example 2.1. For $\tau > 0$, we assume that N is a flat manifold with

$$R_{N,\alpha\beta} = 0.$$

Let

$$f(t,x) = f(t) = \tau t, \quad \varphi(t) = e^{-t}.$$

It is easy to testify that

$$R_{M,ij} + f_{ij} - \frac{f_i f_j}{\tau} = \lambda g_{ij} \tag{2-5}$$

for $\lambda = -(n + \tau - 1)$. Hence M is τ-quasi-Einstein with potential function $f = \tau t$ and $\lambda = -(n + \tau - 1)$. Moreover, by (2-3) and (2-4), the scalar curvature of M is

$$R_M = -n(n-1),$$

which means that (1-2) follows with $\mu = 0$. It is easy to see that the potential function f is not bounded from above.

Example 2.2. For $\tau > 0$, we assume that N is an Einstein manifold with

$$R_{N,\alpha\beta} = -(n+\tau-2)\delta_{\alpha\beta}.$$

Choose

$$f(t,x) = f(t) = -\tau \log\cosh t, \quad \varphi(t) = \cosh t.$$

It is easy to testify that (2-5) holds for $\lambda = -(n+\tau-1)$. Hence M is τ-quasi-Einstein with potential function $f = -\tau \log\cosh t$ and $\lambda = -(n+\tau-1)$. Moreover, by (2-3) and (2-4), the scalar curvature of M is

$$R_M = -n(n-1) - \frac{(n-1)\tau}{\cosh^2 t},$$

which means that (1-2) follows with

$$\mu = -\tau(\tau+n-2).$$

It is easy to see that $\mu < 0$ and R_M is bounded from above.

3. Basic formulas

In this section, we will first give some basic formulas for quasi-Einstein metrics in Lemma 3.1. These formulas are well-established in [Case et al. 2011; Kim and Kim 2003; Wang 2011].

Lemma 3.1. *If g is a τ-quasi-Einstein metric with potential function f and λ is a constant, then one can get*

$$\tag{3-1} \tfrac{1}{2}\Delta R - \frac{\tau+2}{2\tau}\nabla f\cdot\nabla R = -\frac{\tau-1}{\tau}\Big|Ric - \frac{1}{n}Rg\Big|^2 - \frac{n+\tau-1}{n\tau}(R-n\lambda)\Big(R - \frac{n(n-1)}{n+\tau-1}\lambda\Big).$$

Moreover, there exists a constant μ such that

$$\tag{3-2} R + \frac{\tau-1}{\tau}|\nabla f|^2 + (\tau-n)\lambda = \mu e^{2f/\tau}.$$

And also one can get

$$\tag{3-3} \nabla\Delta f\cdot\nabla f = \frac{2}{\tau}\Delta f|\nabla f|^2 - 2\mathrm{Ric}(\nabla f,\nabla f)$$

$$\tag{3-4} \Delta f - |\nabla f|^2 - \tau\lambda + \mu e^{2f/\tau} = 0.$$

In the following, we will calculate the weighted Laplacian of $\varphi(R + 2xe^{2f/\tau})$ by using Lemma 3.1, where $x > 0$ is a constant and φ is a smooth cutoff function.

Lemma 3.2. *Let*

$$\tag{3-5} Q = \varphi(R + 2xe^{2f/\tau}),$$

where $x > 0$ is a constant and φ is a smooth cutoff function. If $\tau > 1$ and $\mu > 0$, then for $\epsilon > 0$,

$$\tag{3-6} \begin{aligned} \tfrac{1}{2}\Delta_f Q \le{}& \frac{\Delta_f\varphi}{2\varphi}Q + \frac{\nabla\varphi\cdot\nabla Q}{\varphi} - \frac{|\nabla\varphi|^2}{\varphi^2}Q + \frac{\varphi^2}{4\epsilon\tau}\left|\frac{\nabla Q}{\varphi} - \frac{Q\nabla\varphi}{\varphi^2}\right|^2 \\ &+ \frac{4x(n+\tau-1)}{n\tau}Qe^{2f/\tau} - \frac{n+\tau-1}{n\tau\varphi}Q^2 \\ &+ \Big(\frac{2n-2+\tau}{\tau}\lambda - \frac{\epsilon}{(\tau-1)\varphi}\Big)Q + \varphi A - \frac{n(n-1)\varphi}{\tau}\lambda^2 + \frac{n-\tau}{\tau-1}\epsilon\lambda \end{aligned}$$

holds at $y \in M$ with $\varphi(y) \neq 0$, where

$$\tag{3-7} \Delta_f = \Delta - \nabla f\cdot\nabla$$

and A, depending on $x, n, \tau, \mu, \lambda, \epsilon, \varphi$, is defined in (3-14).

Proof. Let

$$\tag{3-8} G = R + 2xe^{2f/\tau}.$$

It is easy to see that

$$\tag{3-9} \Delta_f e^{2f/\tau} = \Delta e^{2f/\tau} - \nabla e^{2f/\tau}\cdot\nabla f = \Big(\frac{4-2\tau}{\tau^2}|\nabla f|^2 + \frac{2}{\tau}\Delta f\Big)e^{2f/\tau},$$

which, together with (3-1), shows that, for $\epsilon > 0$,

$$\begin{aligned} \tfrac{1}{2}\Delta_f G &\le \frac{1}{\tau}\nabla R\cdot\nabla f + x\Big(\frac{4-2\tau}{\tau^2}|\nabla f|^2 + \frac{2}{\tau}\Delta f\Big)e^{2f/\tau} - \frac{n+\tau-1}{n\tau}(R-n\lambda)\Big(R - \frac{n(n-1)}{n+\tau-1}\lambda\Big) \\ &= \frac{1}{\tau}\nabla G\cdot\nabla f + x\Big(-\frac{2}{\tau}|\nabla f|^2 + \frac{2}{\tau}\Delta f\Big)e^{2f/\tau} - \frac{n+\tau-1}{n\tau}(R-n\lambda)\Big(R - \frac{n(n-1)}{n+\tau-1}\lambda\Big) \\ &\le \frac{\varphi}{4\epsilon\tau}|\nabla G|^2 + \frac{\epsilon}{\tau\varphi}|\nabla f|^2 + x\Big(-\frac{2}{\tau}|\nabla f|^2 + \frac{2}{\tau}\Delta f\Big)e^{2f/\tau} \\ &\qquad - \frac{n+\tau-1}{n\tau}(R-n\lambda)\Big(R - \frac{n(n-1)}{n+\tau-1}\lambda\Big) \end{aligned}$$

holds at $y \in M$ when $\varphi(y) \neq 0$. By (3-8) and (3-2), we get

$$\tag{3-10} R = G - 2xe^{2f/\tau}$$

and

$$\tag{3-11} |\nabla f|^2 = -\frac{\tau}{\tau-1}G + \frac{\tau(2x+\mu)}{\tau-1}e^{2f/\tau} + \frac{\tau(n-\tau)}{\tau-1}\lambda.$$

Plugging (3-8), (3-10), (3-11) and (3-4) into 3 yields

$$\begin{aligned}\frac{1}{2}\Delta_f G \le{}& \frac{\varphi}{4\epsilon\tau}|\nabla G|^2 + \frac{4x(n+\tau-1)}{n\tau}Ge^{2f/\tau} - \frac{n+\tau-1}{n\tau}G^2 \\ &+ \Big(\frac{2n-2+\tau}{\tau}\lambda - \frac{\epsilon}{(\tau-1)\varphi}\Big)G - \frac{4x^2(n+\tau-1)+2xn\mu}{n\tau}e^{4f/\tau} \\ &- \left(\frac{4x(n-1)}{\tau}\lambda - \frac{\epsilon(2x+\mu)}{(\tau-1)\varphi}\right)e^{2f/\tau} - \frac{n(n-1)}{\tau}\lambda^2 + \frac{n-\tau}{(\tau-1)\varphi}\epsilon\lambda.\end{aligned} \tag{3-12}$$

Since for all $a > 0$,

$$-ax^2 + bx \le \frac{b^2}{4a},$$

we conclude that

$$-\frac{4x^2(n+\tau-1)+2xn\mu}{n\tau}e^{4f/\tau} - \left(\frac{4x(n-1)}{\tau}\lambda - \frac{\epsilon(2x+\mu)}{(\tau-1)\varphi}\right)e^{2f/\tau} \le A \tag{3-13}$$

with

$$A = \frac{n\tau}{16x^2(n+\tau-1)+8nx\mu}\left(\frac{4x(n-1)}{\tau}\lambda - \frac{\epsilon(2x+\mu)}{(\tau-1)\varphi}\right)^2. \tag{3-14}$$

It is easy to see that

$$\nabla G = \frac{\nabla Q}{\varphi} - \frac{Q\nabla\varphi}{\varphi^2}$$

and

$$\Delta_f Q = \frac{Q}{\varphi}\Delta_f\varphi + 2\nabla\varphi\cdot\nabla G + \varphi\Delta_f G.$$

Hence

$$\Delta_f Q = \frac{\Delta_f\varphi}{\varphi}Q + \frac{2\nabla\varphi\cdot\nabla Q}{\varphi} - \frac{2|\nabla\varphi|^2}{\varphi^2}Q + \varphi\Delta_f G. \tag{3-15}$$

Plugging (3-5), (3-12) and (3-13) into (3-15) yields (3-6). □

4. A rigid property

In this section, we will prove Theorem 1.1, a rigid property of τ-quasi-Einstein metrics with $\lambda \le 0$ on complete noncompact Riemannian manifolds.

Proof. Consider a smooth function $\theta(t) : [0, +\infty) \to [0, 1]$:

$$\theta(t) = \begin{cases} 1 & \text{if } 0 \le t \le 1, \\ 0 & \text{if } t \ge 2, \end{cases} \tag{4-1}$$

so that

$$-10\sqrt{\theta} \le \theta' \le 0. \tag{4-2}$$

For $R_0 > 0$, let

$$\varphi(x) = \theta\Big(\frac{r(x)}{R_0}\Big)$$

be a cutoff function, where $r(x)$ is the distance function determined by $O \in M$. Then

$$0 \le \varphi \le 1, \quad |\nabla\varphi|(x) \le \frac{C}{R_0}$$

and $\varphi(x) = 1$ on B_{R_0}, $\varphi(x) = 0$ outside of B_{2R_0}. Let

$$\alpha = -\frac{\tau+2}{\tau}.$$

Plugging (1-1) into (3-3) yields

$$-\nabla\Delta f \cdot \nabla f + \nabla|\nabla f|^2 \cdot \nabla f = 2\lambda|\nabla f|^2 + \frac{2}{\tau}|\nabla f|^4 - \frac{2}{\tau}\Delta f|\nabla f|^2. \tag{4-3}$$

Integrating (4-3) on M and using the fact that $\lambda \le 0$, we obtain

$$\begin{aligned}-\int_M \nabla\Delta f \cdot \nabla f \varphi e^{\alpha f}\,dx &+ \int_M \nabla|\nabla f|^2 \cdot \nabla f \varphi e^{\alpha f}\,dx \\ &\le \frac{2}{\tau}\int_M |\nabla f|^4 \varphi e^{\alpha f}\,dx - \frac{2}{\tau}\int_M \Delta f|\nabla f|^2 \varphi e^{\alpha f}\,dx.\end{aligned} \tag{4-4}$$

Integrating by parts yields

$$\int_M \nabla|\nabla f|^2 \cdot \nabla f \varphi e^{\alpha f}\,dx = -\int_M |\nabla f|^2(\Delta f\varphi + \alpha|\nabla f|^2\varphi + \nabla f \cdot \nabla\varphi)e^{\alpha f}\,dx \tag{4-5}$$

and

$$\int_M \nabla\Delta f \cdot \nabla f \varphi e^{\alpha f}\,dx = -\int_M \big((\Delta f)^2\varphi + \alpha\Delta f|\nabla f|^2\varphi + \Delta f\nabla f \cdot \nabla\varphi\big)e^{\alpha f}\,dx. \tag{4-6}$$

Taking (4-5) and (4-6) into (4-4) yields

$$\begin{aligned}\int_M \big((\Delta f)^2 &- 2\Delta f|\nabla f|^2 + |\nabla f|^4\big)\varphi e^{\alpha f}\,dx \\ &\le -\int_M \big(\Delta f\nabla f \cdot \nabla\varphi - |\nabla f|^2\nabla f \cdot \nabla\varphi\big)e^{\alpha f}\,dx \\ &\le \Big(\int_M (\Delta f - |\nabla f|^2)^2\varphi e^{\alpha f}\,dx\Big)^{1/2}\Big(\int_{B_{2R_0}\setminus B_{R_0}} \frac{|\nabla f \cdot \nabla\varphi|^2}{\varphi}e^{\alpha f}\,dx\Big)^{1/2}.\end{aligned}$$

Observing that

$$|\nabla f \cdot \nabla\varphi| \le |\nabla f|\,|\nabla\varphi| \le \frac{C}{R_0}|\nabla f|,$$

we get

$$\int_{B_{R_0}} (\Delta f - |\nabla f|^2)^2 e^{\alpha f\ dx} \le \int_M (\Delta f - |\nabla f|^2)^2 \varphi e^{\alpha f} dx$$
$$\le C R_0^{-2} \int_{B_{2R_0} \setminus B_{R_0}} |\nabla f|^2 e^{\alpha f} dx.$$

Letting $R_0 \to \infty$, by (1-3), we conclude that

$$\int_M (\Delta f - |\nabla f|^2)^2 e^{\alpha f} dx = 0.$$

Hence $\Delta e^f = 0$.

When $\lambda < 0$, we deduce from $\Delta e^f = 0$ that $\Delta f = |\nabla f|^2$. Equation (4-3) is then equivalent to $2\lambda |\nabla f|^2 = 0$, which means that f is constant. □

5. Lower bound of the scalar curvature

In this section, we will prove Theorem 1.4 and Theorem 1.6 by using the weighted Laplacian comparison theorem and the maximum principle. We first introduce the weighted Laplacian comparison theorem, which can be found in [Lott 2003; Wang 2010].

Lemma 5.1. *Let (M, g) be an n-dimensional complete Riemannian manifold, f a real value smooth function on M and $\Delta_f = \Delta - \nabla f \cdot \nabla$ the weighted Laplacian. Assume that the τ-Bakry–Émery Ricci curvature on M is bounded by*

$$\mathrm{Ric}_{f,\tau} \ge \lambda$$

with constant λ and $r(x) = \mathrm{dist}(O, x)$ is the distance function determined by a fixed point O. If a_λ is a solution to the Riccati equation

$$\frac{\partial a_\lambda}{\partial r} = \lambda - \frac{a_\lambda^2}{n + \tau - 1}, \quad \lim_{r \searrow 0} r a_\lambda = n + \tau - 1,$$

then at $y \notin \mathrm{Cut}(O)$,

$$\Delta_f r \le a_\lambda(r).$$

In particular, if $\lambda \le 0$,

$$\Delta_f r \le \frac{n+\tau-1}{r}\left(1 + \sqrt{-\frac{\lambda}{n+\tau-1}}\, r\right).$$

We need the following estimate, which can be proved by using the maximum principle [Pigola et al. 2005; Schoen and Yau 1994; Yau 1975; Cheng and Yau 1975].

Theorem 5.2. *Let M be an n-dimensional complete noncompact Riemannian manifold, g a τ-quasi-Einstein metric with potential function f and $\lambda \le 0$ a constant. We also assume that $\tau > 1$ and $\mu > 0$. Then for $x > 0$,*

$$R(y)+2xe^{2f/\tau(y)} \ge \frac{n(2n-2+\tau)+n\sqrt{\Delta}}{2(n+\tau-1)}\lambda \tag{5-1}$$

holds for any $y \in M$, where

$$\Delta = \tau^2 + \frac{8(n+\tau-1)(n-1)^2 x}{2x(n+\tau-1)+n\mu}. \tag{5-2}$$

Proof. Consider a smooth function $\theta(t) : [0, +\infty) \to [0, 1]$,

$$\theta(t) = \begin{cases} 1 & \text{if } 0 \le t \le 1, \\ 0 & \text{if } t \ge 2, \end{cases}$$

so that

$$-10\theta^{1/2} \le \theta' \le 0, \theta'' \ge -10. \tag{5-3}$$

For a large enough constant $R_0 > 0$, define the smooth cutoff function $\varphi : M \to \mathbb{R}$ by

$$\varphi(x, t) = \theta\Big(\frac{r(x)}{R_0}\Big).$$

Then

$$\nabla\varphi = \frac{\theta' \nabla r}{R_0}. \tag{5-4}$$

By Lemma 5.1, we have that for $y \in B_{2R_0}$,

$$\begin{aligned} \Delta_f \varphi(y) = \Delta\varphi - \nabla\varphi \cdot \nabla f &= \frac{\theta''}{R_0^2} + \frac{\theta' \Delta_f r}{R_0} \\ &\ge \frac{\theta''}{R_0^2} + \frac{(n+\tau-1)\theta'(1+\sqrt{K}R_0)}{R_0^2}, \end{aligned} \tag{5-5}$$

where

$$K = -\frac{\lambda}{n+\tau-1} \ge 0.$$

Let

$$Q = \varphi G = \varphi(R + 2xe^{2f/\tau}).$$

If for any $R_0 > 0$ the minimal value of G on B_{R_0} is not smaller than zero, then Theorem 5.2 holds. Hence we can assume that for some large enough value of $R_0 > 0$, the minimal value of G on B_{R_0} is negative. If we assume that Q achieves its minimal value at x_0 on B_{2R_0}, then

$$Q(x_0) \le \min_{x \in B_{R_0}} Q(x) = \min_{x \in B_{R_0}} G(x) < 0,$$

which means that x_0 is not on the boundary of B_{2R_0}. Hence $\varphi(x_0) > 0$ and

$$\nabla Q = 0, \tag{5-6}$$

$$\Delta_f Q \geq 0 \tag{5-7}$$

hold at x_0. By (3-6), (5-6) and (5-7), we get that, at x_0,

$$\begin{aligned} 0 \leq \frac{\Delta_f \varphi}{2} Q - \frac{|\nabla\varphi|^2}{\varphi} Q + \frac{|\nabla\varphi|^2}{4\epsilon\tau\varphi} Q^2 + \frac{4x(n+\tau-1)\varphi}{n\tau} Q e^{2f/\tau} - \frac{n+\tau-1}{n\tau} Q^2 \\ + \Big(\frac{2n-2+\tau}{\tau}\varphi\lambda - \frac{\epsilon}{\tau-1}\Big) Q + \varphi^2 A - \frac{n(n-1)\varphi^2}{\tau}\lambda^2 + \frac{n-\tau}{\tau-1}\varphi\epsilon\lambda. \end{aligned} \tag{5-8}$$

Noticing that, at x_0,

$$\frac{4x(n+\tau-1)\varphi}{n\tau} Q e^{2f/\tau} \leq 0. \tag{5-9}$$

Taking (5-4), (5-5) and (5-9) into (5-8), and using (5-3), we get that, at x_0,

$$\begin{aligned} 0 \leq -\frac{105Q}{R_0^2} - \frac{5(n+\tau-1)(1+\sqrt{K}R_0)Q}{R_0^2} + \frac{25}{\epsilon\tau R_0^2} Q^2 - \frac{n+\tau-1}{n\tau} Q^2 \\ + \Big(\frac{2n-2+\tau}{\tau}\varphi\lambda - \frac{\epsilon}{\tau-1}\Big) Q + \varphi^2 A - \frac{n(n-1)\varphi^2}{\tau}\lambda^2 + \frac{n-\tau}{\tau-1}\varphi\epsilon\lambda. \end{aligned} \tag{5-10}$$

By (3-14) and the fact that $\lambda \leq 0$, we have that, at x_0,

$$\varphi^2 A \leq B, \tag{5-11}$$

where

$$B = \frac{n\tau}{16x^2(n+\tau-1)+8nx\mu}\Big(\frac{4x(n-1)}{\tau}\lambda - \frac{\epsilon(2x+\mu)}{\tau-1}\Big)^2. \tag{5-12}$$

For a large enough value of $R_0 > 0$, define

$$\sigma(R_0) = \frac{\inf\{G(x) : x \in B(O, R_0)\}}{\inf\{G(x) : x \in B(O, 2R_0)\}}.$$

It is easy to see that

$$Q(x_0) = \varphi(x_0)G(x_0) \leq \inf\{G(x) : x \in B(O, R_0)\}$$

and

$$Q(x_0) = \varphi(x_0)G(x_0) \geq \varphi(x_0)\inf\{G(x) : x \in B(O, 2R_0)\}.$$

Using the assumption that $\inf\{G(x) : x \in B(O, R_0)\} < 0$, we get that

$$\sigma(R_0) \leq \varphi(x_0) \leq 1. \tag{5-13}$$

From (5-10), (5-11) and (5-13), it follows that, at x_0,

$$0 \le -\frac{105Q}{R_0^2} - \frac{5(n+\tau-1)(1+\sqrt{K}R_0)Q}{R_0^2} + \frac{25}{\epsilon\tau R_0^2}Q^2 - \frac{n+\tau-1}{n\tau}Q^2 + \Big(\frac{2n-2+\tau}{\tau}\sigma(R_0)\lambda - \frac{\epsilon}{\tau-1}\Big)Q + B - \frac{n(n-1)\sigma^2(R_0)}{\tau}\lambda^2 - \frac{|n-\tau|}{\tau-1}\epsilon\lambda. \tag{5-14}$$

Now, assume that R_0 is large enough so that

$$\frac{25}{\epsilon\tau R_0^2} < \frac{n+\tau-1}{n\tau}.$$

Then (5-14) gives us that, at x_0,

$$Q \ge \frac{-E(R_0) - \sqrt{E^2(R_0) + 4D(R_0)F(R_0)}}{2D(R_0)}, \tag{5-15}$$

where

$$D(R_0) = \frac{n+\tau-1}{n\tau} - \frac{25}{\epsilon\tau R_0^2},$$
$$E(R_0) = \frac{105 + 5(n+\tau-1)(1+\sqrt{K}R_0)}{R_0^2} - \frac{(2n-2+\tau)\sigma(R_0)\lambda}{\tau} + \frac{\epsilon}{\tau-1},$$
$$F(R_0) = B - \frac{n(n-1)\sigma^2(R_0)\lambda^2}{\tau} - \frac{|n-\tau|}{\tau-1}\epsilon\lambda.$$

Hence, for all $y \in B_{R_0}$,

$$G(y) = \varphi(y)G(y) \ge Q(x_0) \ge \frac{-E(R_0) - \sqrt{E^2(R_0) + 4D(R_0)F(R_0)}}{2D(R_0)}. \tag{5-16}$$

Noting that $0 \le \sigma(R_0) \le 1$ and B is a constant independent of R_0, we deduce from (5-16) that for a large enough value of R_0, G is bounded from below by a constant independent of R_0. Hence

$$\lim_{R_0\to\infty} \sigma(R_0) = 1,$$

which means that

$$D = \lim_{R_0\to\infty} D(R_0) = \frac{n+\tau-1}{n\tau},$$
$$E = \lim_{R_0\to\infty} E(R_0) = -\frac{(2n-2+\tau)\lambda}{\tau} + \frac{\epsilon}{\tau-1},$$
$$F = \lim_{R_0\to\infty} F(R_0) = B - \frac{n(n-1)\lambda^2}{\tau} - \frac{|n-\tau|}{\tau-1}\epsilon\lambda.$$

By (5-16), we obtain that, for all $y \in M$, $x > 0$ and $\epsilon > 0$,

$$G(y) \geq \frac{-E - \sqrt{E^2 + 4DF}}{2D}. \tag{5-17}$$

Note that

$$\lim_{\epsilon \searrow 0} E = -\frac{(2n-2+\tau)\lambda}{\tau}, \quad \lim_{\epsilon \searrow 0} F = \frac{2xn(n-1)^2\lambda^2}{(2x(n+\tau-1)+n\mu)\tau} - \frac{n(n-1)\lambda^2}{\tau}.$$

Letting $\epsilon \searrow 0$ in (5-17) leads to (5-1). □

The following result is useful.

Theorem 5.3. *If $\lambda < 0$ and $\mu < 0$, then for all $y \in M$,*

$$f(y) \leq \frac{\tau}{2} \ln \frac{\tau\lambda}{\mu}. \tag{5-18}$$

Proof. Let φ be the cutoff function defined in the proof of Theorem 5.2 and

$$H = \varphi e^{2f/\tau}.$$

Noting that

$$\nabla H = \nabla\varphi e^{2f/\tau} + \varphi \nabla e^{2f/\tau},$$

by (3-4) and (3-9), we have that

$$\Delta_f e^{2f/\tau} \geq \frac{2}{\tau} \Delta_f f e^{2f/\tau} = -\frac{2\mu}{\tau} e^{4f/\tau} + 2\lambda e^{2f/\tau}.$$

Hence

$$\begin{aligned} \Delta_f H &= \varphi \Delta_f e^{2f/\tau} + e^{2f/\tau} \Delta_f \varphi + 2\nabla\varphi \cdot \nabla e^{2f/\tau} \\ &\geq -\frac{2\mu\varphi}{\tau} e^{4f/\tau} + 2\lambda\varphi e^{2f/\tau} + \frac{\Delta_f \varphi}{\varphi} H + 2\frac{\nabla\varphi \cdot \nabla H}{\varphi} - 2\frac{|\nabla\varphi|^2}{\varphi^2} H \end{aligned} \tag{5-19}$$

holds at $y \in M$ when $\varphi(y) > 0$. We assume that H achieves its maximum at x_0 on B_{2R_0}. If $\varphi(x_0) = 0$, then for all $x \in B_{R_0}$,

$$e^{2f/\tau(x)} = \varphi(x) e^{2f/\tau(x)} = H(x) \leq H(x_0) = 0,$$

which is impossible, so $\varphi(x_0) > 0$. Noting that

$$\Delta_f H \leq 0, \quad \nabla H = 0$$

hold at x_0, by (5-3), (5-4), (5-5) and (5-19), we get that, at x_0,

$$
\begin{aligned}
0 &\geq -\frac{2\mu\varphi}{\tau}e^{4f/\tau} + 2\lambda\varphi e^{2f/\tau} + \frac{\Delta_f\varphi}{\varphi}H - \frac{2|\nabla\varphi|^2}{\varphi^2}H \\
&\geq -\frac{2\mu}{\tau\varphi}H^2 + 2\lambda H + \frac{\theta'' + (n+\tau-1)(1+\sqrt{K}R_0)\theta'}{R_0^2\varphi}H - \frac{2|\theta'|^2}{R_0^2\varphi^2}H \\
&\geq -\frac{2\mu}{\tau\varphi}H^2 + 2\lambda H - \frac{210 + 10(n+\tau-1)(1+\sqrt{K}R_0)\sqrt{\varphi}}{R_0^2\varphi}H.
\end{aligned}
$$

By using the fact that $0 < \varphi(x_0) \leq 1$, we have that, at x_0,

$$
\begin{aligned}
-\frac{2\mu}{\tau}H^2 &\leq -2\lambda\varphi H + \frac{210 + 10(n+\tau-1)(1+\sqrt{K}R_0)\sqrt{\varphi}}{R_0^2}H \\
&\leq -2\lambda H + \frac{210 + 10(n+\tau-1)(1+\sqrt{K}R_0)}{R_0^2}H
\end{aligned}
\tag{5-20}
$$

or

$$
H(x_0) \leq \frac{\tau\lambda}{\mu} - \frac{105\tau + 5\tau(n+\tau-1)(1+\sqrt{K}R_0)}{\mu R_0^2}.
$$

Hence, for all $y \in B_{R_0}$,

$$
\begin{aligned}
e^{2f/\tau(y)} = \varphi(y)e^{2f/\tau(y)} = H(y) &\leq H(x_0)m \\
&\leq \frac{\tau\lambda}{\mu} - \frac{105\tau + 5\tau(n+\tau-1)(1+\sqrt{K}R_0)}{\mu R_0^2}.
\end{aligned}
$$

Letting $R_0 \to \infty$ yields (5-18). □

Remark 5.4. Equation (1-1) still holds if we shift the function f by a constant. However, from Equation (3-2), the constant μ will change after this shift.

Proof of Theorem 1.4. When $\tau \geq 1$ and $\mu \leq 0$, (1-4) follows from Theorem 1.3. When $\tau > 1$ and $\mu > 0$, (1-4) follows by letting $x \searrow 0$ in (5-1). We only need to consider the case that $0 < \tau \leq 1$. Now if $\mu \geq 0$, (3-2) tells us that

$$
R = \mu e^{2f/\tau} + (n-\tau)\lambda + \frac{1-\tau}{\tau}|\nabla f|^2 \geq (n-\tau)\lambda \geq n\lambda.
$$

If $\mu < 0$ and $\lambda < 0$, by Theorem 5.3, we have

$$
R \geq \mu e^{2f/\tau} + (n-\tau)\lambda \geq n\lambda.
$$

Hence (1-4) follows. If $\mu < 0$ and $\lambda = 0$, [Wang 2011, Theorem 3.2] tells us that f is constant, and Theorem 1.4 follows. □

Proof of Theorem 1.6. When $\mu \leq 0$, (3-2) tells us that

$$R \leq (n-\tau)\lambda + \frac{1-\tau}{\tau}|\nabla f|^2.$$

If $\tau \geq 1$, then $R \leq (n-\tau)\lambda$. If $0 < \tau < 1$, by the gradient estimate in [Wang 2011], we have $|\nabla f|^2 \leq -\tau\lambda$. Hence $R \leq (n-1)\lambda$. □

Acknowledgments

The author thanks the referee for valuable comments and suggestions. The author is also grateful to Professor Shengliang Pan for his constant encouragement.

References

[Besse 1987] A. L. Besse, *Einstein manifolds*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)] **10**, Springer, Berlin, 1987. MR 88f:53087 Zbl 0613.53001

[Cao 2009] H. Cao, "Geometry of complete gradient shrinking solitons", preprint, 2009. arXiv 0903.3927

[Cao and Zhou 2010] H.-D. Cao and D. Zhou, "On complete gradient shrinking Ricci solitons", *J. Differential Geom.* **85**:2 (2010), 175–185. MR 2011k:53040 Zbl 05835154

[Cao and Zhu 2006] H.-D. Cao and X.-P. Zhu, "A complete proof of the Poincaré and geometrization conjectures—application of the Hamilton-Perelman theory of the Ricci flow", *Asian J. Math.* **10**:2 (2006), 165–492. MR 2008d:53090 Zbl 1200.53057

[Case 2010] J. S. Case, "The nonexistence of quasi-Einstein metrics", *Pacific J. Math.* **248**:2 (2010), 277–284. MR 2011m:53053 Zbl 1204.53032

[Case et al. 2011] J. Case, Y.-J. Shu, and G. Wei, "Rigidity of quasi-Einstein metrics", *Differential Geom. Appl.* **29**:1 (2011), 93–100. MR 2784291 Zbl 1215.53033

[Chen 2009] B.-L. Chen, "Strong uniqueness of the Ricci flow", *J. Differential Geom.* **82**:2 (2009), 363–382. MR 2010h:53095 Zbl 1177.53036

[Cheng and Yau 1975] S. Y. Cheng and S. T. Yau, "Differential equations on Riemannian manifolds and their geometric applications", *Comm. Pure Appl. Math.* **28**:3 (1975), 333–354. MR 52 #6608 Zbl 0312.53031

[Hamilton 1995] R. S. Hamilton, "The formation of singularities in the Ricci flow", pp. 7–136 in *Surveys in differential geometry, Vol. II* (Cambridge, MA, 1993), edited by S. Yau and C. Hsiung, Int. Press, 1995. MR 97e:53075 Zbl 0867.53030

[Ivey 1993] T. Ivey, "Ricci solitons on compact three-manifolds", *Differential Geom. Appl.* **3**:4 (1993), 301–307. MR 94j:53048 Zbl 0788.53034

[Kim and Kim 2003] D.-S. Kim and Y. H. Kim, "Compact Einstein warped product spaces with nonpositive scalar curvature", *Proc. Amer. Math. Soc.* **131**:8 (2003), 2573–2576. MR 2004b:53063 Zbl 1029.53027

[Li 2005] X.-D. Li, "Liouville theorems for symmetric diffusion operators on complete Riemannian manifolds", *J. Math. Pures Appl.* (9) **84**:10 (2005), 1295–1361. MR 2006f:58046 Zbl 1082.58036

[Lott 2003] J. Lott, "Some geometric properties of the Bakry–Émery–Ricci tensor", *Comment. Math. Helv.* **78**:4 (2003), 865–883. MR 2004i:53044 Zbl 1038.53041

[Lü et al. 2004] H. Lü, D. N. Page, and C. N. Pope, "New inhomogeneous Einstein metrics on sphere bundles over Einstein–Kähler manifolds", *Phys. Lett. B* **593**:1-4 (2004), 218–226. MR 2005f:53063

[Munteanu 2009] O. Munteanu, "The volume growth of complete gradient shrinking Ricci solitons", preprint, 2009. arXiv 0904.0798

[O'Neill 1983] B. O'Neill, *Semi-Riemannian geometry*, Pure and Applied Mathematics **103**, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1983. With applications to relativity. MR 85f:53002 Zbl 0531.53051

[Perelman 2002] G. Perelman, "The entropy formula for the Ricci flow and its geometric applications", preprint, 2002. Zbl 1130.53001 arXiv math/0211159

[Pigola et al. 2005] S. Pigola, M. Rigoli, and A. G. Setti, "Maximum principles on Riemannian manifolds and applications", *Mem. Amer. Math. Soc.* **174**:822 (2005), x+99. MR 2006b:53048

[Qian 1997] Z. Qian, "Estimates for weighted volumes and applications", *Quart. J. Math. Oxford Ser.* (2) **48**:190 (1997), 235–242. MR 98e:53058 Zbl 0902.53032

[Schoen and Yau 1994] R. Schoen and S.-T. Yau, *Lectures on differential geometry*, Conference Proceedings and Lecture Notes in Geometry and Topology **1**, International Press, Cambridge, MA, 1994. MR 97d:53001 Zbl 0830.53001

[Wang 2010] L.-F. Wang, "The upper bound of the L^2_μ spectrum", *Ann. Global Anal. Geom.* **37**:4 (2010), 393–402. MR 2011d:53067 Zbl 1190.53034

[Wang 2011] L. F. Wang, "Rigid properties of quasi-Einstein metrics", *Proc. Amer. Math. Soc.* **139**:10 (2011), 3679–3689. MR 2813397 Zbl 05959506

[Wei and Wylie 2007] G. Wei and W. Wylie, "Comparison geometry for the smooth metric measure spaces", pp. 191–202 in *Proceedings of the fourth ICCM, Vol. 2* (Hangzhou, China), edited by S. Yau, 2007.

[Yau 1975] S. T. Yau, "Harmonic functions on complete Riemannian manifolds", *Comm. Pure Appl. Math.* **28** (1975), 201–228. MR 55 #4042 Zbl 0291.31002

[Zhang 2006] X. Zhang, "Compactness theorems for gradient Ricci solitons", *J. Geom. Phys.* **56**:12 (2006), 2481–2499. MR 2008b:53049 Zbl 1107.53045

[Zhang 2009] Z.-H. Zhang, "On the completeness of gradient Ricci solitons", *Proc. Amer. Math. Soc.* **137**:8 (2009), 2755–2759. MR 2010a:53057 Zbl 1176.53046

[Zhang 2011] S. J. Zhang, "On a sharp volume estimate for gradient Ricci solitons with scalar curvature bounded below", *Acta Math. Sin.* (*Engl. Ser.*) **27**:5 (2011), 871–882. MR 2786449 Zbl 1219.53039

Received January 10, 2011. Revised November 7, 2011.

LIN FENG WANG
SCHOOL OF SCIENCE
NANTONG UNIVERSITY
NANTONG 226007, JIANGSU
CHINA
wlf711178@126.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

DECOMPOSITION OF DE RHAM COMPLEXES WITH SMOOTH HORIZONTAL COEFFICIENTS FOR SEMISTABLE REDUCTIONS

QIHONG XIE

We generalize Illusie's result to prove the decomposition of the de Rham complex with smooth horizontal coefficients for a semistable S-morphism $f : X \to Y$ which is liftable over $\mathbb{Z}/p^2\mathbb{Z}$. As an application, we prove the Kollár vanishing theorem in positive characteristic for a semistable S-morphism $f : X \to Y$ which is liftable over $\mathbb{Z}/p^2\mathbb{Z}$, where all concerned horizontal divisors are smooth over Y.

1. Introduction

The decomposition of de Rham complexes is one of the most important results in algebraic geometry of positive characteristic, which has been discovered by Deligne and Illusie [1987] and successfully used to give a purely algebraic proof of the Kodaira vanishing theorem. More precisely, let k be a perfect field of characteristic $p > 0$, and $W_2(k)$ the ring of Witt vectors of length two of k. Let S be a k-scheme, $\widetilde{S}$ a lifting of S over $W_2(k)$, X a smooth S-scheme, and $F : X \to X_1$

This paper was partially supported by the National Natural Science Foundation of China (Grant No. 10901037), the Ph.D. Programs Foundation of the Ministry of Education of China (Grant No. 20090071120004) and the Scientific Research Foundation for Returned Overseas Chinese Scholars, State Education Ministry.
MSC2010: 14F17.
Keywords: de Rham complex, semistable reduction, positive characteristic.

the relative Frobenius morphism of X over S. If X has a lifting $\widetilde{X}$ over $\widetilde{S}$ and $\dim(X/S) < p$, then we have a decomposition in $D(X_1)$:

$$\bigoplus_i \Omega^i_{X_1/S}[-i] \xrightarrow{\sim} F_*\Omega^\bullet_{X/S}.$$

Illusie [1990] generalized the result above to the relative case for a semistable S-morphism $f : X \to Y$ to obtain the decomposition of de Rham complexes with coefficients in the Gauss–Manin systems. Roughly speaking, let E be the branch divisor of f, $D = X \times_Y E$, and $\mathbb{H} = \bigoplus_i \boldsymbol{R}^i f_*\Omega^\bullet_{X/Y}(\log D/E)$ the Gauss–Manin system. If $f : X \to Y$ has a lifting $\widetilde{f} : \widetilde{X} \to \widetilde{Y}$ over $\widetilde{S}$ and $\dim(X/S) < p$, then we have a decomposition in $D(Y_1)$:

$$\bigoplus_i \mathrm{gr}^i\, \Omega^\bullet_{Y_1/S}(\log E_1)(\mathbb{H}_1) \xrightarrow{\sim} F_*\Omega^\bullet_{Y/S}(\log E)(\mathbb{H}).$$

In this paper, we generalize Illusie's result to the case where smooth horizontal coefficients are taken into account. Roughly speaking, let D be an adapted divisor on X, i.e., D consists of three parts: all singular fibers of f, some smooth fibers of f, and some smooth horizontal divisors with respect to f (see Definition 2.2 for more details). Let

$$\mathbb{H} = \textstyle\bigoplus_i \boldsymbol{R}^i f_*\Omega^\bullet_{X/Y}(\log D/E_a).$$

Then we prove that if $f : (X, D) \to (Y, E_a)$ has a lifting $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ and $\dim(X/S) < p$, then there is a decomposition in $D(Y_1)$ (see Theorem 5.9 for more details):

$$\textstyle\bigoplus_i \mathrm{gr}^i\, \Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}_1) \xrightarrow{\sim} F_*\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}),$$

from which follows the Kollár vanishing theorem in positive characteristic, saying that

$$H^i\big(Y, \mathscr{L} \otimes R^j f_*\omega_{X/S}(D)\big) = 0$$

holds for any $i > 0$, $j \geq 0$, and any ample invertible sheaf $\mathscr{L}$ on Y (see Theorem 6.3 for more details). It should be mentioned that the proofs of all of the results in this paper follow Illusie's arguments very closely.

In general, we may put forward the following conjecture, called logarithmic Kollár vanishing for semistable reductions in positive characteristic (see [Kollár 1995, Theorem 10.19] for the logarithmic Kollár vanishing theorem in characteristic zero):

Conjecture 1.1. *Let X and Y be proper and smooth S-schemes, $f : X \to Y$ an E-semistable S-morphism, and D a simple normal crossing divisor on X containing the divisor $X \times_Y E$. Let H be a $\mathbb{Q}$-divisor on X such that the support of the* *fractional part of H is contained in D and $H \sim_{\mathbb{Q}} f^*L$, where L is an ample $\mathbb{Q}$-divisor on Y. Assume that $f:(X,D)\to(Y,E)$ has a lifting $\widetilde{f}:(\widetilde{X},\widetilde{D})\to(\widetilde{Y},\widetilde{E})$ over $\widetilde{S}$ and* $\dim(X/S)<p$. *Then $H^i\big(Y, R^jf_*\mathcal{O}_X(K_{X/S}+\lceil H\rceil)\big)=0$ holds for any $i>0$ and $j\geq 0$, where $\lceil H\rceil$ denotes the round-up of H.*

There are several difficulties in dealing with this conjecture. First, we need some technique to change the $\mathbb{Q}$-divisor argument into the integral divisor argument. Second, the situation of the horizontal divisors contained in D or H is complicated. Third, the decomposition of de Rham complexes with horizontal coefficients is completely unknown. In this sense, all of the results obtained in this paper may be regarded as the first step to resolving Conjecture 1.1.

Notation. We denote the support of a divisor D by $\mathrm{Supp}(D)$, the relative dualizing sheaf of $f:X\to Y$ by $\omega_{X/Y}$, and the divisor defined by $x=0$ by $\mathrm{div}_0(x)$.

2. Definitions and preliminaries

This section is parallel to [Illusie 1990, §1], and all proofs follow Illusie's proofs very closely.

Definition 2.1. Let S be a scheme, X and Y smooth S-schemes, and $f:X\to Y$ an S-morphism. Let $E\subset Y$ be a divisor relatively simple normal crossing over S (RSNC for short), and $E_X=X\times_Y E$. We say that $f:X\to Y$ is E-semistable, or that f has a semistable reduction along E if, locally for the étale topology, f is the product of S-morphisms of one of the following types:

(i) $\mathrm{pr}_1:\mathbb{A}^n_S\to\mathbb{A}^1_S$, $E=\varnothing$;

(ii) $h:\mathbb{A}^n_S\to\mathbb{A}^1_S$, $h^*y=x_1\cdots x_n$, where $\mathbb{A}^n_S=\mathrm{Spec}\,\mathcal{O}_S[x_1,\dots,x_n]$, $\mathbb{A}^1_S=\mathrm{Spec}\,\mathcal{O}_S[y]$, and $E=\mathrm{div}_0(y)$.

Definition 2.2. Let $f:X\to Y$ be an E-semistable morphism as in Definition 2.1. A divisor $D\subset X$ is said to be adapted to f if the following conditions hold:

(i) D admits a decomposition $D=E_X+D_a+D_h$ of irreducible components, where D_a is the sum of the irreducible components of D whose images under f are divisors not contained in E and D_h is the sum of those whose images under f are the whole Y.

(ii) D is RSNC over S, D_h is RSNC over Y, and the union of the divisor $A:=f(D_a)$ and E is RSNC over S.

Remark 2.3. (1) The divisor E_X is adapted to f.

(2) In Definition 2.2, for any irreducible component D_{h1} of D_h, the restriction morphism $f|_{D_{h1}}:D_{h1}\to Y$ is smooth.

Definition 2.4. Let $f : X \to Y$ be an E-semistable S-morphism with an adapted divisor D as in Definition 2.2. For simplicity, denote $E + A$ by E_a. Let $\Omega^\bullet_{X/S}(\log D)$ and $\Omega^\bullet_{Y/S}(\log E_a)$ be the de Rham complexes of X and Y with logarithmic poles along D and E_a, respectively. We define $\Omega^\bullet_{X/Y}(\log D/E_a)$, the de Rham complex of X over Y with relative logarithmic poles along D over E_a, in the following way.

By Lemma 2.5, the quotient

$$\Omega^1_{X/Y}(\log D/E_a) := \frac{\Omega^1_{X/S}(\log D)}{\operatorname{Im}\big(f^*\Omega^1_{Y/S}(\log E_a)\big)}$$

is a locally free sheaf on X of rank $d = n - e$, where $n = \dim(X/S)$ and $e = \dim(Y/S)$. Let $\Omega^i_{X/Y}(\log D/E_a) = \bigwedge^i \Omega^1_{X/Y}(\log D/E_a)$, and define the differential d by passage to the quotient of that of the complex $\Omega^\bullet_{X/S}(\log D)$.

It is easy to see that if f is smooth, then $\Omega^\bullet_{X/Y}(\log D/E_a) = \Omega^\bullet_{X/Y}(\log D_h)$ in the usual sense.

Lemma 2.5. *With notation as in Definition 2.4, there is an exact sequence of locally free sheaves of finite type*:

$$0 \to f^*\Omega^1_{Y/S}(\log E_a) \to \Omega^1_{X/S}(\log D) \to \Omega^1_{X/Y}(\log D/E_a) \to 0. \tag{2-1}$$

Proof. We only have to prove the statement locally for the étale topology, so it suffices to check for the following three types, where $\mathbb{A}^n_S = \operatorname{Spec} \mathcal{O}_S[x_1, \dots, x_n]$ and $\mathbb{A}^1_S = \operatorname{Spec} \mathcal{O}_S[y]$:

(i) $\mathrm{pr}_1 : \mathbb{A}^n_S \to \mathbb{A}^1_S$, where $E = \varnothing$, $A = \varnothing$, $\mathrm{pr}_1^*(y) = x_1$, and $D_h = \operatorname{div}_0(x_2 \cdots x_r)$ for $1 \le r \le n$;

(ii) $\mathrm{pr}_1 : \mathbb{A}^n_S \to \mathbb{A}^1_S$, where $E = \varnothing$, $A = \operatorname{div}_0(y)$, $\mathrm{pr}_1^*(y) = x_1$, $D_a = \operatorname{div}_0(x_1)$, and $D_h = \operatorname{div}_0(x_2 \cdots x_r)$ for $1 \le r \le n$;

(iii) $h : \mathbb{A}^n_S \to \mathbb{A}^1_S$, where $E = \operatorname{div}_0(y)$, $h^*y = x_1 \cdots x_s$, $A = \varnothing$, and $D_h = \operatorname{div}_0(x_{s+1} \cdots x_n)$ for $2 \le s \le n$.

(i) $f^*\Omega^1_{Y/S}(\log E_a)$ is generated by $f^*(dy) = dx_1$, $\Omega^1_{X/S}(\log D)$ is generated by $dx_1, dx_2/x_2, \dots, dx_r/x_r, dx_{r+1}, \dots, dx_n$; hence, $\Omega^1_{X/Y}(\log D/E_a)$ is generated by $dx_2/x_2, \dots, dx_r/x_r, dx_{r+1}, \dots, dx_n$, so the conclusion is clear.

(ii) $f^*\Omega^1_{Y/S}(\log E_a)$ is generated by $f^*(dy/y) = dx_1/x_1$, and $\Omega^1_{X/S}(\log D)$ is generated by $dx_1/x_1, \dots, dx_r/x_r, dx_{r+1}, \dots, dx_n$; hence, $\Omega^1_{X/Y}(\log D/E_a)$ is generated by $dx_2/x_2, \dots, dx_r/x_r, dx_{r+1}, \dots, dx_n$, so the conclusion is clear.

(iii) $f^*\Omega^1_{Y/S}(\log E_a)$ is generated by $f^*(dy/y) = \sum_{i=1}^s dx_i/x_i$, $\Omega^1_{X/S}(\log D)$ is generated by $dx_1/x_1, \dots, dx_n/x_n$; hence

$$\Omega^1_{X/Y}(\log D/E_a) = \mathbb{O}_X\left\langle \frac{dx_1}{x_1}, \ldots, \frac{dx_s}{x_s}\right\rangle \Big/ \left(\sum_{i=1}^{s} \frac{dx_i}{x_i}\right) \bigoplus \mathbb{O}_X\left\langle \frac{dx_{s+1}}{x_{s+1}}, \ldots, \frac{dx_n}{x_n}\right\rangle,$$

so the conclusion is clear. □

Remarks 2.6. 1. Let $f_i : X_i \to Y_i$ be E_i-semistable S-morphisms with adapted divisors D_i as in Definition 2.2. Let X, Y, and $f : X \to Y$ be the external products over S of X_i, Y_i, and f_i, respectively. Then $\Omega^\bullet_{X/Y}(\log D/E_a)$ is the external tensor product of $\Omega^\bullet_{X_i/Y_i}(\log D_i/E_{ai})$ over S.

2. In the exact sequence (2-1), taking the top exterior tensor product gives rise to the canonical isomorphism

$$f^*\Omega^e_{Y/S}(\log E_a) \otimes \Omega^d_{X/Y}(\log D/E_a) \xrightarrow{\sim} \Omega^n_{X/S}(\log D).$$

Since $\Omega^e_{Y/S}(\log E_a) = \omega_{Y/S}(E_a)$, $\Omega^n_{X/S}(\log D) = \omega_{X/S}(D)$, we have

$$\Omega^d_{X/Y}(\log D/E_a) \cong \omega_{X/Y}(D_h).$$

3. Let $f' : X' \to Y'$ be deduced from $f : X \to Y$ by a base change $Y' \to Y$. Put

$$\Omega^\bullet_{X'/Y'}(\log D'/E'_a) = \Omega^\bullet_{X/Y}(\log D/E_a) \otimes_{\mathbb{O}_X} \mathbb{O}_{X'}. \tag{2-2}$$

Note that, in general, X' is no longer smooth over S, and that it is no longer possible to interpret $\Omega^\bullet_{X'/Y'}(\log D'/E'_a)$ as a de Rham complex with relative logarithmic poles.

4. Let $j : U \hookrightarrow X$ be the open subset over which f is smooth. Then we have a canonical isomorphism:

$$\Omega^\bullet_{X/Y}(\log D/E_a) \xrightarrow{\sim} j_*\Omega^\bullet_{U/Y}(\log D|_U/E_a). \tag{2-3}$$

In fact, for any point $s \in S$, $X_s - U_s$ is of codimension at least 2 in X_s; therefore, $\Omega^\bullet_{X/Y}(\log D/E_a)$ is the unique prolongation of $\Omega^\bullet_{U/Y}(\log D|_U/E_a)$ with components being locally free of finite type.

From now on, let S be a scheme of characteristic $p > 0$, and $f : X \to Y$ an E-semistable S-morphism with an adapted divisor D as in Definition 2.2. Let F_X and F_Y be the absolute Frobenius morphisms of X and Y, which fit into the commutative diagram

$$\begin{array}{ccccc} X & \xrightarrow{F} & X' & \longrightarrow & X \\ & \searrow^{f} & \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ & & Y & \xrightarrow{F_Y} & Y \end{array} \tag{2-4}$$

where the square is cartesian and the composition of the upper horizontal morphisms is equal to F_X.

The differential d of the complex $F_*\Omega^\bullet_{X/Y}(\log D/E_a)$ is $\mathcal{O}_{X'}$-linear, so we would like to calculate its cohomology $\mathcal{O}_{X'}$-modules by a Cartier-type isomorphism. Consider the following commutative diagram with cartesian square:

$$
\begin{array}{ccccc}
X & \xrightarrow{F_{X/S}} & X_1 & \longrightarrow & X \\
& \searrow & \downarrow & & \downarrow \\
& & S & \xrightarrow{F_S} & S,
\end{array}
\tag{2-5}
$$

where $F_{X/S} : X \to X_1$ is the relative Frobenius morphism of X over S.

By [Katz 1970, 7.2], we have the Cartier isomorphism

$$C^{-1} : \Omega^1_{X_1/S}(\log D_1) \xrightarrow{\sim} \mathcal{H}^1\big(F_{X/S*}\Omega^\bullet_{X/S}(\log D)\big),$$

where D_1 is the pullback of D by F_S. By adjunction of (F_S^*, F_{S*}) and abuse of notation, we have the homomorphism

$$C^{-1} : \Omega^1_{X/S}(\log D) \to \mathcal{H}^1\big(F_{X*}\Omega^\bullet_{X/S}(\log D)\big),$$

which sends dx and dx/x to the cohomology class of, respectively, $x^{p-1}dx$ and dx/x in the $\mathcal{O}_X$-module $\mathcal{H}^1\big(F_{X*}\Omega^\bullet_{X/S}(\log D)\big)$. The natural surjective morphism of complexes of $\mathcal{O}_X$-modules $F_{X*}\Omega^\bullet_{X/S}(\log D) \to F_{X*}\Omega^\bullet_{X/Y}(\log D/E_a)$ induces a natural homomorphism

$$\pi : \mathcal{H}^1\big(F_{X*}\Omega^\bullet_{X/S}(\log D)\big) \to \mathcal{H}^1\big(F_{X*}\Omega^\bullet_{X/Y}(\log D/E_a)\big),$$

which kills all cohomology classes of $y^{p-1}dy$ (respectively, dy/y), where dy (respectively, dy/y) are local sections of $f^*\Omega^1_{Y/S}(\log E_a)$. The composition

$$\pi \circ C^{-1} : \Omega^1_{X/S}(\log D) \to \mathcal{H}^1\big(F_{X*}\Omega^\bullet_{X/Y}(\log D/E_a)\big)$$

vanishes on $f^*\Omega^1_{Y/S}(\log E_a)$, which defines the homomorphism

$$C^{-1} : \Omega^1_{X/Y}(\log D/E_a) \to \mathcal{H}^1\big(F_{X*}\Omega^\bullet_{X/Y}(\log D/E_a)\big).$$

By adjunction of (F_Y^*, F_{Y*}), we have the Cartier homomorphism

$$C^{-1} : \Omega^1_{X'/Y}(\log D'/E_a) \to \mathcal{H}^1\big(F_*\Omega^\bullet_{X/Y}(\log D/E_a)\big).$$

The exterior product gives rise to a homomorphism of graded $\mathcal{O}_{X'}$-algebras:

$$C^{-1} : \Omega^*_{X'/Y}(\log D'/E_a) \to \mathcal{H}^*\big(F_*\Omega^\bullet_{X/Y}(\log D/E_a)\big). \tag{2-6}$$

Proposition 2.7. *The homomorphism* (2-6) *is an isomorphism.*

Proof. Since (2-6) is compatible with étale topology and external tensor products over S, it suffices to prove the statement for those three types described in the proof of Lemma 2.5.

(i) and (ii) In these cases, f is smooth; hence, $\Omega^\bullet_{X/Y}(\log D/E_a) = \Omega^\bullet_{X/Y}(\log D_h)$. Thus (2-6) is just the usual Cartier isomorphism [Katz 1970, 7.2].

(iii) In this case, we can further assume $S = \operatorname{Spec} \mathbb{F}_p$. Diagram (2-4) corresponds to the following diagram of rings:

$$\begin{array}{ccccc} B & \longrightarrow & B' & \xrightarrow{F} & B \\ {\scriptstyle f^*}\uparrow & & {\scriptstyle f'^*}\uparrow & \nearrow{\scriptstyle f^*} & \\ A & \xrightarrow{F_A} & A, & & \end{array}$$

where $A = \mathbb{F}_p[y]$, $B = \mathbb{F}_p[x_1, \dots, x_n]$, $f^*(y) = x_1 \cdots x_s$, $F_A(y) = y^p$, and $F(x_i) = x_i^p$. If we identify B with the A-algebra $\mathbb{F}_p[x_1, \dots, x_n, y]/(y - x_1 \cdots x_s)$, then B' can be identified with the A-algebra $\mathbb{F}_p[x_1, \dots, x_n, y]/(y^p - x_1 \cdots x_s)$ since $y \in A$ is sent to y^p. Thus B' can also be identified with the A-algebra $\mathbb{F}_p[x_1^p, \dots, x_s^p, x_{s+1}, \dots, x_n, x_1 \cdots x_s]$. Define

$$\begin{aligned} B_1 &= \mathbb{F}_p[x_1, \dots, x_s], & B_1' &= \mathbb{F}_p[x_1^p, \dots, x_s^p, x_1 \cdots x_s], \\ B_2 &= \mathbb{F}_p[x_{s+1}, \dots, x_n], & B_2' &= \mathbb{F}_p[x_{s+1}, \dots, x_n]. \end{aligned}$$

Then $B = B_1 \otimes B_2$, $B' = B_1' \otimes B_2'$, $F : B' \to B$ factorizes into the external tensor product of $F_1 : B_1' \to B_1$ defined by the inclusion and $F_2 : B_2' \to B_2$ defined by the p-th power map, and $B \to B'$ factorizes into the external tensor product of $B_j \to B_j'$ for $j = 1, 2$, where $B_1 \to B_1'$ is defined by $x_i \mapsto x_i^p$ $(1 \le i \le s)$ and $B_2 \to B_2'$ is defined by $x_i \mapsto x_i$ $(s + 1 \le i \le n)$. To prove that (2-6) is an isomorphism, it suffices to prove that

$$C^{-1} : \Omega^*_{B_j'/A}(\log D'/E_a) \to \mathcal{H}^*\big(F_* \Omega^\bullet_{B_j/A}(\log D/E_a)\big) \tag{2-7}$$

is an isomorphism for $j = 1, 2$. When $j = 1$, it was proved in [Illusie 1990, Proposition 1.5]. When $j = 2$, (2-7) is just the usual Cartier isomorphism [Katz 1970, 7.2]. □

Remark 2.8. Note that in case (iii), f is no longer smooth, X' is no longer smooth over S, and $F : X \to X'$ is no longer flat.

3. Decomposition of de Rham complex with relative logarithmic poles

This section is parallel to [Illusie 1990, §2], and all proofs follow Illusie's proofs very closely.

Definition 3.1. Let S be a scheme of characteristic $p > 0$. A *lifting* of S over $\mathbb{Z}/p^2\mathbb{Z}$ is a scheme $\widetilde{S}$, defined and flat over $\mathbb{Z}/p^2\mathbb{Z}$ such that

$$\widetilde{S} \times_{\operatorname{Spec} \mathbb{Z}/p^2\mathbb{Z}} \operatorname{Spec} \mathbb{F}_p = S.$$

A lifting of the absolute Frobenius morphism $F_S : S \to S$ over $\widetilde{S}$ is an endomorphism $F_{\widetilde{S}} : \widetilde{S} \to \widetilde{S}$ of $\widetilde{S}$ such that $F_{\widetilde{S}}|_S = F_S$. A lifting of an E-semistable S-morphism $f : X \to Y$ with an adapted divisor D over $\widetilde{S}$ is an $\widetilde{E}$-semistable $\widetilde{S}$-morphism $\widetilde{f} : \widetilde{X} \to \widetilde{Y}$ with an adapted divisor $\widetilde{D}$ as in Definition 2.2, such that $\widetilde{X} \times_{\widetilde{S}} S = X$, $\widetilde{Y} \times_{\widetilde{S}} S = Y$, $\widetilde{D} \times_{\widetilde{S}} S = D$, $\widetilde{E} \times_{\widetilde{S}} S = E$, and $\widetilde{f}|_X = f$. We say that $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ is a lifting of $f : (X, D) \to (Y, E_a)$ over $\widetilde{S}$ if no confusion is likely.

In this section, let S be a scheme of characteristic $p > 0$, $\widetilde{S}$ a lifting of S over $\mathbb{Z}/p^2\mathbb{Z}$, and $F_{\widetilde{S}} : \widetilde{S} \to \widetilde{S}$ a lifting of the absolute Frobenius morphism $F_S : S \to S$ over $\widetilde{S}$. Let $f : X \to Y$ be an E-semistable S-morphism with an adapted divisor D as in Definition 2.2, and $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ a lifting of $f : (X, D) \to (Y, E_a)$ over $\widetilde{S}$ as in Definition 3.1. Let $\widetilde{D}_1 \subset \widetilde{X}_1$ be the $\widetilde{S}$-schemes deduced from $\widetilde{D} \subset \widetilde{X}$ by the base change $F_{\widetilde{S}}$, and $\widetilde{F} : \widetilde{X} \to \widetilde{X}_1$ an $\widetilde{S}$-morphism lifting the relative Frobenius morphism $F : X \to X_1$ of X over S:

$$\begin{array}{ccccc} \widetilde{X} & \xrightarrow{\widetilde{F}} & \widetilde{X}_1 & \longrightarrow & \widetilde{X} \\ & \searrow & \downarrow & & \downarrow \\ & & \widetilde{S} & \xrightarrow{F_{\widetilde{S}}} & \widetilde{S} \end{array}$$

We say that $\widetilde{F}$ is compatible with $\widetilde{D}$ if

$$\widetilde{F}^* \mathcal{O}_{\widetilde{X}_1}(-\widetilde{D}_1) = \mathcal{O}_{\widetilde{X}}(-p\widetilde{D}).$$

Locally for the étale topology on X, there exists a lifting $\widetilde{F} : \widetilde{X} \to \widetilde{X}_1$ compatible with $\widetilde{D}$. Indeed, if $\widetilde{X}$ is étale over $\mathbb{A}^n_{\widetilde{S}}$ via coordinates $\{\widetilde{x}_1, \dots, \widetilde{x}_n\}$ and $\widetilde{D} = \operatorname{div}_0(\widetilde{x}_1 \cdots \widetilde{x}_r)$, then there exists a unique lifting $\widetilde{F}$ such that $\widetilde{F}^*(\widetilde{x}_i \otimes 1) = \widetilde{x}_i^p$ for $1 \le i \le n$.[-3pt]

We recall the following results from [Deligne and Illusie 1987, 4.2.3]. Two compatible liftings $\widetilde{F}_1$, $\widetilde{F}_2$ differ by a derivation

$$h_{12} = (\widetilde{F}_2^* - \widetilde{F}_1^*)/p : \Omega^1_{X_1/S}(\log D_1) \to F_* \mathcal{O}_X.$$

In fact, if $\widetilde{X}$ is étale over $\mathbb{A}^n_{\widetilde{S}}$ via coordinates $\{\widetilde{x}_1, \dots, \widetilde{x}_n\}$ and $\widetilde{D} = \operatorname{div}_0(\widetilde{x}_1 \cdots \widetilde{x}_r)$, then we can write $\widetilde{F}_j^*(\widetilde{x}_i \otimes 1) = (1 + pa_{ij})\widetilde{x}_i^p$ for $1 \le i \le r$ and $j = 1, 2$. By an easy calculation, we have $h_{12}(dx_i/x_i \otimes 1) = a_{i2} - a_{i1}$ for $1 \le i \le r$. Furthermore, any lifting $\widetilde{F}$ compatible with $\widetilde{D}$ gives rise to a quasi-isomorphism of complexes:

$$\phi_{\widetilde{F}} : \bigoplus_{i<p} \Omega^i_{X_1/S}(\log D_1)[-i] \to \tau_{<p} F_* \Omega^\bullet_{X/S}(\log D), \tag{3-1}$$

which is given in degree 1 by $\phi^1_{\widetilde{F}} = \widetilde{F}^*/p$ and prolonged canonically through the exterior powers.

Theorem 3.2. *Let $f : X \to Y$ be an E-semistable S-morphism with an adapted divisor D, and $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ a lifting of $f : (X, D) \to (Y, E_a)$ over $\widetilde{S}$. Let $\widetilde{F}_{Y/S} : \widetilde{Y} \to \widetilde{Y}_1$ be a lifting over $\widetilde{S}$ of the relative Frobenius morphism $F_{Y/S} : Y \to Y_1$ of Y over S, which is compatible with the divisor $\widetilde{E}_a$. Then there is a canonical isomorphism in $D(X', \mathcal{O}_{X'})$:*

$$\phi_{(\widetilde{f}, \widetilde{F}_{Y/S})} : \bigoplus_{i<p} \Omega^i_{X'/Y}(\log D'/E_a)[-i] \to \tau_{<p} F_{X/Y*} \Omega^\bullet_{X/Y}(\log D/E_a), \tag{3-2}$$

which induces the Cartier isomorphism C^{-1} (2-6) *on $\mathcal{H}^i$.*

Proof. The proof is analogous to that of [Illusie 1990, Theorem 2.2]. It suffices to define, for any $i < p$, $\phi^i : \Omega^i_{X'/Y}(\log D'/E_a)[-i] \to F_{X/Y*}\Omega^\bullet_{X/Y}(\log D/E_a)$ inducing C^{-1} on $\mathcal{H}^i$. Since ϕ^i can be deduced from ϕ^1 by an argument similar to that of [Deligne and Illusie 1987, 2.1(a)], we only have to define ϕ^1. The definition of ϕ^1 is given in three steps.

Step 1: local case. To define ϕ^1, we first suppose that there is a lifting $\widetilde{F} : \widetilde{X} \to \widetilde{X}_1$ of the relative Frobenius morphism $F : X \to X_1$ of X over S, which is compatible with $\widetilde{D}$ and compatible with $\widetilde{F}_{Y/S}$ in the sense that the square is commutative:

$$\begin{array}{ccc} \widetilde{X} & \xrightarrow{\widetilde{F}} & \widetilde{X}_1 \\ {\scriptstyle \widetilde{f}}\downarrow & & \downarrow{\scriptstyle \widetilde{f}_1} \\ \widetilde{Y} & \xrightarrow{\widetilde{F}_{Y/S}} & \widetilde{Y}_1. \end{array} \tag{3-3}$$

The morphism $\phi^1_{\widetilde{F}} : \Omega^1_{X_1/S}(\log D_1)[-1] \to F_*\Omega^\bullet_{X/S}(\log D)$ in (3-1), composed with the projection of $F_*\Omega^\bullet_{X/S}(\log D)$ onto $F_*\Omega^\bullet_{X/Y}(\log D/E_a)$, vanishes on the subsheaf $f_1^*\Omega^1_{Y_1/S}(\log E_{a1})[-1]$; therefore, by passage to the quotient, it defines a morphism $\Omega^1_{X_1/Y_1}(\log D_1/E_{a1})[-1] \to F_*\Omega^\bullet_{X/Y}(\log D/E_a)$, and by adjunction, it defines a morphism

$$\phi^1 : \Omega^1_{X'/Y}(\log D'/E_a)[-1] \to F_{X/Y*}\Omega^\bullet_{X/Y}(\log D/E_a), \tag{3-4}$$

which induces the Cartier isomorphism C^{-1} on $\mathcal{H}^1$.

Step 2: from local to global. Assume that $\widetilde{F}_j : \widetilde{X} \to \widetilde{X}_1$ are liftings of the relative Frobenius of X over S for $j = 1, 2$, which are compatible with $\widetilde{D}$ and compatible with $\widetilde{F}_{Y/S}$. Then the derivation

$$(\widetilde{F}_2^* - \widetilde{F}_1^*)/p : \Omega^1_{X_1/S}(\log D_1) \to F_* \mathcal{O}_X$$

vanishes on the subsheaf $f_1^* \Omega^1_{Y_1/S}(\log E_{a1})$ by the commutativity of the square (3-3). Therefore, by passage to the quotient and by adjunction, it defines a homomorphism

$$h_{12} : \Omega^1_{X'/Y}(\log D'/E_a) \to F_{X/Y*} \mathcal{O}_X.$$

A calculation analogous to that of [Deligne and Illusie 1987, 2.1(c)] shows that $\phi_2^1 - \phi_1^1 = dh_{12}$ holds, where ϕ_j^1 are the morphisms (3-4) associated with $\widetilde{F}_j$ for $j = 1, 2$. By an argument similar to that of [Deligne and Illusie 1987, 2.1(c)], we have a relation of transitivity: $h_{12} + h_{23} = h_{13}$ for three liftings $\widetilde{F}_1$, $\widetilde{F}_2$, $\widetilde{F}_3$ of the relative Frobenius of X over S. Working on the étale topology instead of the Zariski topology on X, we can construct a global morphism ϕ^1 by the procedure of the Čech globalization described as in [Deligne and Illusie 1987, 2.1(d)].

Step 3: local existence of compatible liftings. We shall prove that locally for the étale topology on X, there exists a lifting $\widetilde{F} : \widetilde{X} \to \widetilde{X}_1$ compatible with $\widetilde{D}$ and compatible with $\widetilde{F}_{Y/S}$. Keeping in mind the types (i), (ii), and (iii) as in the proof of Lemma 2.5, we divide the argument into four cases.

(I) Assume $\operatorname{div}_0(y) \subset \widetilde{A}$ and $\widetilde{F}^*_{Y/S}(y \otimes 1) = (1 + pa)y^p$. Define $\widetilde{F}^*(x_1 \otimes 1) = (1 + pa)x_1^p$ and $\widetilde{F}^*(x_i \otimes 1) = x_i^p$ $(i \geq 2)$, where x_1 is the coordinate for the fiber over $\operatorname{div}_0(y)$, and x_i $(i \geq 2)$ are the coordinates for the divisor $\widetilde{D}_h$.

(II) Assume $\operatorname{div}_0(y) \subset \widetilde{E}$ and $\widetilde{F}^*_{Y/S}(y \otimes 1) = (1 + pa)y^p$. Define $\widetilde{F}^*(x_1 \otimes 1) = (1 + pa)x_1^p$ and $\widetilde{F}^*(x_i \otimes 1) = x_i^p$ $(i \geq 2)$, where x_1 is a prechosen coordinate for the fiber over $\operatorname{div}_0(y)$, and x_i $(i \geq 2)$ are the other coordinates for the fiber over $\operatorname{div}_0(y)$ or the coordinates for $\widetilde{D}_h$.

(III) Assume $\operatorname{div}_0(y) \not\subset \widetilde{E}_a$ and $\widetilde{F}^*_{Y/S}(y \otimes 1) = y^p + pb$. Define $\widetilde{F}^*(x_1 \otimes 1) = x_1^p + pb$ and $\widetilde{F}^*(x_i \otimes 1) = x_i^p$ $(i \geq 2)$, where x_1 is the coordinate for the fiber over $\operatorname{div}_0(y)$, and x_i $(i \geq 2)$ are the coordinates for $\widetilde{D}_h$.

(IV) Assume that all x_i are not the coordinates for the fiber over $\operatorname{div}_0(y)$. Define $\widetilde{F}^*(x_i \otimes 1) = x_i^p$.

It is easy to check that $\widetilde{F} : \widetilde{X} \to \widetilde{X}_1$ constructed above is a lifting of the relative Frobenius of X over S, which is compatible with $\widetilde{D}$ and compatible with $\widetilde{F}_{Y/S}$. □

Remarks 3.3. (1) If f is smooth, then the existence of a lifting of (X', D'_h) over $\widetilde{Y}$ such that $\widetilde{X}'$ is smooth over $\widetilde{Y}$ and $\widetilde{D}'_h$ is RSNC over $\widetilde{Y}$, gives rise to the decomposition of $\tau_{<p} F_{X/Y*} \Omega^\bullet_{X/Y}(\log D_h)$. Moreover, the gerbe of splittings

of $\tau_{\leq 1} F_{X/Y*}\Omega^\bullet_{X/Y}(\log D_h)$ is canonically isomorphic to the gerbe of liftings of (X', D'_h) over $\tilde{Y}$ (see [Deligne and Illusie 1987, 4.2.3]).

(2) Under the hypotheses of Theorem 3.2, suppose that f is of relative dimension $\leq p$ and $H^{p+1}\Big(X', \big(\Omega^p_{X'/Y}(\log D'/E_a)\big)^\vee\Big) = 0$ (this is the case, for example, if Y is affine and f is proper), then $F_{X/Y*}\Omega^\bullet_{X/Y}(\log D/E_a)$ is decomposable, i.e., there is an isomorphism in $D(X', \mathcal{O}_{X'})$

$$\bigoplus_i \Omega^i_{X'/Y}(\log D'/E_a)[-i] \xrightarrow{\sim} F_{X/Y*}\Omega^\bullet_{X/Y}(\log D/E_a),$$

which induces the Cartier isomorphism C^{-1} on $\mathcal{H}^i$. The proof of the decomposition is analogous to that of [Deligne and Illusie 1987, 3.7(b) and 4.2.3].

We shall state some corollaries for $\Omega^\bullet_{X/Y}(\log D/E_a)$ and omit their proofs, which are analogous to those in [Illusie 1990, §2].

Corollary 3.4. *Under the hypotheses of Theorem 3.2, suppose further that f is proper. Then*:

(i) *For any $i + j < p$, the $\mathcal{O}_Y$-modules $R^j f_*\Omega^i_{X/Y}(\log D/E_a)$ are locally free of finite type, and of formation compatible with any base change $Z \to Y$.*

(ii) *The Hodge spectral sequence*

$$E_1^{ij} = R^j f_*\Omega^i_{X/Y}(\log D/E_a) \Rightarrow \boldsymbol{R}^{i+j} f_*\Omega^\bullet_{X/Y}(\log D/E_a)$$

satisfies $E_1^{ij} = E_\infty^{ij}$ for any $i + j < p$.

(iii) *If f is of relative dimension $\leq p$, then* (i) *and* (ii) *are valid for any i, j.*

Corollary 3.5. *Let K be a field of characteristic zero, $S = \operatorname{Spec} K$, X, Y smooth S-schemes, and $f : X \to Y$ a proper E-semistable S-morphism with an adapted divisor D as in Definition 2.2. Then*:

(i) *The $\mathcal{O}_Y$-modules $R^j f_*\Omega^i_{X/Y}(\log D/E_a)$ are locally free of finite type, and of formation compatible with any base change $T \to Y$.*

(ii) *The Hodge spectral sequence*

$$E_1^{ij} = R^j f_*\Omega^i_{X/Y}(\log D/E_a) \Rightarrow \boldsymbol{R}^{i+j} f_*\Omega^\bullet_{X/Y}(\log D/E_a)$$

degenerates in E_1.

Corollary 3.6. *Under the hypotheses of Corollary 3.4, suppose further that f is of purely relative dimension $d \leq p$ and S is locally Noetherian and regular. Let $\mathcal{L}$ be an f-ample invertible $\mathcal{O}_X$-module. Then*

$$R^j f_*\big(\mathcal{L}^{-1} \otimes \Omega^i_{X/Y}(\log D/E_a)\big) = 0 \quad \textit{for all } i + j < d,$$
$$R^j f_*\big(\mathcal{L}(-D_h) \otimes \Omega^i_{X/Y}(\log D/E_a)\big) = 0 \quad \textit{for all } i + j > d.$$

Corollary 3.7. *Under the hypotheses of Corollary 3.5, let $\mathscr{L}$ be an f-ample invertible $\mathscr{O}_X$-module. Then*

$$R^j f_*\big(\mathscr{L}^{-1} \otimes \Omega^i_{X/Y}(\log D/E_a)\big) = 0 \quad \text{for all } i + j < d,$$
$$R^j f_*\big(\mathscr{L}(-D_h) \otimes \Omega^i_{X/Y}(\log D/E_a)\big) = 0 \quad \text{for all } i + j > d.$$

4. Variant with support

In this section, let S be a scheme of characteristic $p > 0$, and $f : X \to Y$ an E-semistable S-morphism with an adapted divisor D as in Definition 2.2. For simplicity, denote $E_X + D_a$ by D_v. Tensoring (2-1) with $f^*\mathscr{O}_Y(-E_a) = \mathscr{O}_X(-D_v)$, we obtain an exact sequence of locally free $\mathscr{O}_X$-modules:

$$0 \to f^*\Omega^1_{Y/S}(E_a, 0) \to \Omega^1_{X/S}(D_v, D_h) \to \Omega^1_{X/Y}(D_v, D_h) \to 0, \tag{4-1}$$

where

$$\Omega^1_{Y/S}(E_a, 0) := \Omega^1_{Y/S}(\log E_a) \otimes \mathscr{O}_Y(-E_a),$$
$$\Omega^1_{X/S}(D_v, D_h) := \Omega^1_{X/S}(\log D) \otimes \mathscr{O}_X(-D_v),$$
$$\Omega^1_{X/Y}(D_v, D_h) := \Omega^1_{X/Y}(\log D/E_a) \otimes \mathscr{O}_X(-D_v).$$

For any $i \geq 0$, define

$$\Omega^i_{X/Y}(D_v, D_h) = \Omega^i_{X/Y}(\log D/E_a) \otimes \mathscr{O}_X(-D_v).$$

Then it is easy to check that $\big(\Omega^\bullet_{X/Y}(D_v, D_h), d\big)$ is a well-defined complex.

Let F_Y be the absolute Frobenius of Y, and $F = F_{X/Y}$ the relative Frobenius of X over Y. We have the following commutative diagram with a cartesian square:

$$\begin{array}{ccccc} X & \xrightarrow{F} & X' & \longrightarrow & X \\ & \searrow^{f} & \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ & & Y & \xrightarrow{F_Y} & Y \end{array}$$

The differential d of the complex $F_*\Omega^\bullet_{X/Y}(D_v, D_h)$ is $\mathscr{O}_{X'}$-linear, so we would like to calculate its cohomology $\mathscr{O}_{X'}$-modules by a Cartier-type isomorphism. Consider the Cartier isomorphism (2-6) for $\Omega^\bullet_{X/Y}(\log D/E_a)$:

$$C^{-1} : \Omega^*_{X'/Y}(\log D'/E_a) \to \mathscr{H}^*\big(F_*\Omega^\bullet_{X/Y}(\log D/E_a)\big).$$

Since $F_*\Omega^\bullet_{X/Y}(D_v, D_h) = F_*\Omega^\bullet_{X/Y}(\log D/E_a) \otimes f'^*\mathscr{O}_Y(-E_a)$, tensoring this isomorphism with $f'^*\mathscr{O}_Y(-E_a)$ leads to this proposition:

Proposition 4.1. *There is an isomorphism of graded $\mathscr{O}_{X'}$-algebras*:

$$C^{-1} : \Omega^*_{X'/Y}(D'_v, D'_h) \to \mathscr{H}^*\big(F_*\Omega^\bullet_{X/Y}(D_v, D_h)\big). \tag{4-2}$$

We call (4-2) the *Cartier isomorphism* of $\Omega^\bullet_{X/Y}(D_v, D_h)$.

Tensoring 3.2 with $f'^*\mathcal{O}_Y(-E_a)$, we have the following theorem.

Theorem 4.2. *Let $f : X \to Y$ be an E-semistable S-morphism with an adapted divisor D, and $\tilde{f} : (\tilde{X}, \tilde{D}) \to (\tilde{Y}, \tilde{E}_a)$ a lifting of $f : (X, D) \to (Y, E_a)$ over $\tilde{S}$. Let $\tilde{F}_{Y/S} : \tilde{Y} \to \tilde{Y}_1$ be a lifting over $\tilde{S}$ of the relative Frobenius morphism $F_{Y/S} : Y \to Y_1$ of Y over S, which is compatible with the divisor $\tilde{E}_a$. Then there is a canonical isomorphism in $D(X', \mathcal{O}_{X'})$:*

$$\phi_{(\tilde{f}, \tilde{F}_{Y/S})} : \bigoplus_{i<p} \Omega^i_{X'/Y}(D'_v, D'_h)[-i] \xrightarrow{\sim} \tau_{<p} F_* \Omega^\bullet_{X/Y}(D_v, D_h), \tag{4-3}$$

which induces the Cartier isomorphism C^{-1} (4-2) *on $\mathcal{H}^i$.*

Remark 4.3. Under the hypotheses of Theorem 4.2, suppose that f is of relative dimension $\le p$ and $H^{p+1}\big(X', (\Omega^p_{X'/Y}(D'_v, D'_h))^\vee\big) = 0$ (this is the case, for example, if Y is affine and f is proper), then $F_*\Omega^\bullet_{X/Y}(D_v, D_h)$ is decomposable, i.e., there is an isomorphism in $D(X', \mathcal{O}_{X'})$:

$$\bigoplus_i \Omega^i_{X'/Y}(D'_v, D'_h)[-i] \xrightarrow{\sim} F_*\Omega^\bullet_{X/Y}(D_v, D_h),$$

which induces the Cartier isomorphism C^{-1} on $\mathcal{H}^i$. The proof of the decomposition is analogous to that of [Deligne and Illusie 1987, 3.7(b) and 4.2.3].

We shall state some corollaries for $\Omega^\bullet_{X/Y}(D_v, D_h)$ and omit their proofs, which are analogous to those in [Illusie 1990, §2].

Corollary 4.4. *Under the hypotheses of Theorem 4.2, suppose further that f is proper. Then*:

(i) *For any $i + j < p$, the $\mathcal{O}_Y$-modules $R^j f_* \Omega^i_{X/Y}(D_v, D_h)$ are locally free of finite type, and of formation compatible with any base change $Z \to Y$.*

(ii) *The Hodge spectral sequence*

$$E_1^{ij} = R^j f_* \Omega^i_{X/Y}(D_v, D_h) \Rightarrow \boldsymbol{R}^{i+j} f_* \Omega^\bullet_{X/Y}(D_v, D_h)$$

satisfies $E_1^{ij} = E_\infty^{ij}$ for any $i + j < p$.

(iii) *If f is of relative dimension $\le p$, then (i) and (ii) are valid for any i, j.*

Corollary 4.5. *Let K be a field of characteristic zero, $S = \operatorname{Spec} K$, X, Y smooth S-schemes, and $f : X \to Y$ a proper E-semistable S-morphism with an adapted divisor D as in Definition 2.2. Then*:

(i) *The $\mathcal{O}_Y$-modules $R^j f_* \Omega^i_{X/Y}(D_v, D_h)$ are locally free of finite type, and of formation compatible with any base change $T \to Y$.*

(ii) *The Hodge spectral sequence*

$$E_1^{ij} = R^j f_* \Omega^i_{X/Y}(D_v, D_h) \Rightarrow \boldsymbol{R}^{i+j} f_* \Omega^\bullet_{X/Y}(D_v, D_h)$$

degenerates in E_1.

Corollary 4.6. *Under the hypotheses of Corollary 4.4, suppose further that* f *is of purely relative dimension* $d \le p$ *and* S *is locally Noetherian and regular. Let* $\mathcal{L}$ *be an* f*-ample invertible* $\mathcal{O}_X$*-module. Then*

$$R^j f_*\big(\mathcal{L}^{-1} \otimes \Omega^i_{X/Y}(D_v, D_h)\big) = 0 \quad \textit{for all } i + j < d,$$
$$R^j f_*\big(\mathcal{L}(2D_v - D_h) \otimes \Omega^i_{X/Y}(D_v, D_h)\big) = 0 \quad \textit{for all } i + j > d.$$

Corollary 4.7. *Under the hypotheses of Corollary 4.5, let* $\mathcal{L}$ *be an* f*-ample invertible* $\mathcal{O}_X$*-module. Then*

$$R^j f_*\big(\mathcal{L}^{-1} \otimes \Omega^i_{X/Y}(D_v, D_h)\big) = 0 \quad \textit{for all } i + j < d,$$
$$R^j f_*\big(\mathcal{L}(2D_v - D_h) \otimes \Omega^i_{X/Y}(D_v, D_h)\big) = 0 \quad \textit{for all } i + j > d.$$

5. Decomposition of de Rham complex with smooth horizontal coefficients

This section is parallel to [Illusie 1990, §3], and all proofs follow Illusie's proofs very closely.

In this section, let S be a scheme of characteristic $p > 0$, and $f : X \to Y$ an E-semistable S-morphism with an adapted divisor D as in Definition 2.2. Then we have the following exact sequence of locally free $\mathcal{O}_X$-modules:

$$0 \to f^*\Omega^1_{Y/S}(\log E_a) \to \Omega^1_{X/S}(\log D) \to \Omega^1_{X/Y}(\log D/E_a) \to 0. \tag{5-1}$$

By definition, $\Omega^i_{X/Y}(\log D/E_a) = \bigwedge^i \Omega^1_{X/Y}(\log D/E_a)$ for any $i \ge 0$. Then $\Omega^d_{X/Y}(\log D/E_a) = \omega_{X/Y}(D_h)$, and the de Rham complex $(\Omega^\bullet_{X/Y}(\log D/E_a), d)$, where

$$d : \Omega^i_{X/Y}(\log D/E_a) \to \Omega^{i+1}_{X/Y}(\log D/E_a)$$

is the ordinary differential map.

Definition 5.1. Define $\mathbb{H} = \bigoplus_i \boldsymbol{R}^i f_* \Omega^\bullet_{X/Y}(\log D/E_a)$ to be a graded $\mathcal{O}_Y$-module. The Koszul filtration of $\Omega^\bullet_{X/S}(\log D)$ associated with (5-1) is defined as follows:

$$K^i \Omega^\bullet_{X/S}(\log D) = \operatorname{Im}\big(f^*\Omega^i_{Y/S}(\log E_a) \otimes \Omega^{\bullet - i}_{X/S}(\log D) \to \Omega^\bullet_{X/S}(\log D)\big).$$

Then $K^i \Omega^\bullet_{X/S}(\log D)$ are subcomplexes of $\Omega^\bullet_{X/S}(\log D)$ and induce a decreasing filtration of $\Omega^\bullet_{X/S}(\log D)$:

$$\cdots \supseteq K^i \Omega^\bullet_{X/S}(\log D) \supseteq K^{i+1} \Omega^\bullet_{X/S}(\log D) \supseteq \cdots .$$

It is easy to show that $K^i\Omega^j_{X/S}(\log D)$ is locally free for any i, j, and the associated graded complex

$$\mathrm{gr}^i_K\,\Omega^\bullet_{X/S}(\log D) = K^i/K^{i+1} = f^*\Omega^i_{Y/S}(\log E_a)\otimes\Omega^{\bullet-i}_{X/Y}(\log D/E_a).$$

The exact sequence

$$0\to K^1/K^2\to K^0/K^2\to K^0/K^1\to 0$$

is a short exact sequence of complexes:

$$0\to f^*\Omega^1_{Y/S}(\log E_a)\otimes\Omega^{\bullet-1}_{X/Y}(\log D/E_a)\to\Omega^\bullet_{X/S}(\log D)/K^2 \to\Omega^\bullet_{X/Y}(\log D/E_a)\to 0,$$

which induces a morphism in $D(X)$:

(5-2) $$\Omega^\bullet_{X/Y}(\log D/E_a)\to f^*\Omega^1_{Y/S}(\log E_a)\otimes\Omega^\bullet_{X/Y}(\log D/E_a).$$

Applying $\bigoplus_i \boldsymbol{R}^i f_*$ to (5-2), we obtain the Gauss–Manin connection

(5-3) $$\nabla:\mathbb{H}\to\Omega^1_{Y/S}(\log E_a)\otimes\mathbb{H},$$

and we can show that ∇ is an integrable connection with logarithmic poles along E_a. The complex

(5-4) $$\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}) = \bigl(\mathbb{H}\xrightarrow{\nabla}\Omega^1_{Y/S}(\log E_a)\otimes\mathbb{H}\xrightarrow{\nabla}\cdots\xrightarrow{\nabla}\Omega^i_{Y/S}(\log E_a)\otimes\mathbb{H}\xrightarrow{\nabla}\cdots\bigr)$$

is called the de Rham complex of Y over S with logarithmic poles along E_a and coefficients in the Gauss–Manin system $\mathbb{H}$. In fact, the Koszul filtration of $\Omega^\bullet_{X/S}(\log D)$ and the derived functor $\boldsymbol{R}f_*$ give rise to a spectral sequence

(5-5) $$E^{ij}_1=\boldsymbol{R}^{i+j}f_*\bigl(\mathrm{gr}^i_K\,\Omega^\bullet_{X/S}(\log D)\bigr) =\Omega^i_{Y/S}(\log E_a)\otimes\boldsymbol{R}^j f_*\Omega^\bullet_{X/Y}(\log D/E_a)\Rightarrow\boldsymbol{R}^{i+j}f_*\Omega^\bullet_{X/S}(\log D).$$

Then the de Rham complex $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H})$ is just [-3pt]the direct sum of the horizontal lines of E^{ij}_1, and the Gauss–Manin connection ∇ is just the direct sum of the differential operators $d_1: E^{ij}_1\to E^{i+1,j}_1$.

Variant. By definition, for any $i\geq 0$, we have

$$\Omega^i_{X/S}(D_v,D_h)=\Omega^i_{X/S}(\log D)\otimes\mathcal{O}_X(-D_v),$$
$$\Omega^i_{X/Y}(D_v,D_h)=\Omega^i_{X/Y}(\log D/E_a)\otimes\mathcal{O}_X(-D_v).$$

Define the graded $\mathcal{O}_Y$-module

$$\mathbb{H}^\dagger=\bigoplus_i\boldsymbol{R}^i f_*\Omega^\bullet_{X/Y}(D_v,D_h).$$

The Koszul filtration of $\Omega^\bullet_{X/S}(D_v, D_h)$ associated with (5-1) is defined as follows:

$$K^i\Omega^\bullet_{X/S}(D_v, D_h) = \operatorname{Im}\big(f^*\Omega^i_{Y/S}(\log E_a)\otimes\Omega^{\bullet-i}_{X/S}(D_v, D_h)\to\Omega^\bullet_{X/S}(D_v, D_h)\big).$$

Then $K^i\Omega^\bullet_{X/S}(D_v, D_h)$ are subcomplexes of $\Omega^\bullet_{X/S}(D_v, D_h)$ and induce a decreasing filtration of $\Omega^\bullet_{X/S}(D_v, D_h)$:

$$\cdots\supseteq K^i\Omega^\bullet_{X/S}(D_v, D_h)\supseteq K^{i+1}\Omega^\bullet_{X/S}(D_v, D_h)\supseteq\cdots.$$

It is easy to show that $K^i\Omega^j_{X/S}(D_v, D_h)$ is locally free for any i, j, and the associated graded complex $\operatorname{gr}^i_K\Omega^\bullet_{X/S}(D_v, D_h) = K^i/K^{i+1} = f^*\Omega^i_{Y/S}(\log E_a)\otimes\Omega^{\bullet-i}_{X/Y}(D_v, D_h)$. The exact sequence

$$0\to K^1/K^2\to K^0/K^2\to K^0/K^1\to 0$$

is a short exact sequence of complexes:

$$0\to f^*\Omega^1_{Y/S}(\log E_a)\otimes\Omega^{\bullet-1}_{X/Y}(D_v, D_h)\to\Omega^\bullet_{X/S}(D_v, D_h)/K^2 \to\Omega^\bullet_{X/Y}(D_v, D_h)\to 0,$$

which induces a morphism in $D(X)$:

(5-6) $$\Omega^\bullet_{X/Y}(D_v, D_h)\to f^*\Omega^1_{Y/S}(\log E_a)\otimes\Omega^\bullet_{X/Y}(D_v, D_h).$$

Applying $\bigoplus_i \boldsymbol{R}^i f_*$ to (5-6), we obtain the Gauss–Manin connection

(5-7) $$\nabla : \mathbb{H}^\dagger\to\Omega^1_{Y/S}(\log E_a)\otimes\mathbb{H}^\dagger,$$

and we can show that ∇ is an integrable connection with logarithmic poles along E_a. The complex

(5-8) $$\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^\dagger) = \big(\mathbb{H}^\dagger\xrightarrow{\nabla}\Omega^1_{Y/S}(\log E_a)\otimes\mathbb{H}^\dagger\xrightarrow{\nabla}\cdots\xrightarrow{\nabla}\Omega^i_{Y/S}(\log E_a)\otimes\mathbb{H}^\dagger\xrightarrow{\nabla}\cdots\big)$$

is called the de Rham complex of Y over S with logarithmic poles along E_a and coefficients in the Gauss–Manin system $\mathbb{H}^\dagger$. In fact, the Koszul filtration of $\Omega^\bullet_{X/S}(D_v, D_h)$ and the derived functor $\boldsymbol{R}f_*$ give rise to a spectral sequence

(5-9) $$E_1^{ij} = \boldsymbol{R}^{i+j}f_*\big(\operatorname{gr}^i_K\Omega^\bullet_{X/S}(D_v, D_h)\big) = \Omega^i_{Y/S}(\log E_a)\otimes\boldsymbol{R}^j f_*\Omega^\bullet_{X/Y}(D_v, D_h)\Rightarrow\boldsymbol{R}^{i+j}f_*\Omega^\bullet_{X/S}(D_v, D_h).$$

[-4pt]The de Rham complex $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^\dagger)$ is just the direct sum of the horizontal lines of E_1^{ij}, and the Gauss–Manin connection ∇ is just the direct sum of the differential operators $d_1 : E_1^{ij}\to E_1^{i+1,j}$.

Definition 5.2. The Hodge filtrations of $\mathbb{H}$ and $\mathbb{H}^\dagger$ are decreasing filtrations respectively defined by

$$\mathrm{Fil}^i\mathbb{H} = \mathrm{Im}\big(\textstyle\bigoplus_j \boldsymbol{R}^j f_*\Omega^{\geq i}_{X/Y}(\log D/E_a) \to \bigoplus_j \boldsymbol{R}^j f_*\Omega^\bullet_{X/Y}(\log D/E_a)\big),$$
$$\mathrm{Fil}^i\mathbb{H}^\dagger = \mathrm{Im}\big(\textstyle\bigoplus_j \boldsymbol{R}^j f_*\Omega^{\geq i}_{X/Y}(D_v, D_h) \to \bigoplus_j \boldsymbol{R}^j f_*\Omega^\bullet_{X/Y}(D_v, D_h)\big),$$

which induce the Hodge spectral sequences

$$E_1^{ij} = R^j f_*\Omega^i_{X/Y}(\log D/E_a) \Rightarrow \boldsymbol{R}^{i+j} f_*\Omega^\bullet_{X/Y}(\log D/E_a), \tag{5-10}$$
$$E_1^{ij} = R^j f_*\Omega^i_{X/Y}(D_v, D_h) \Rightarrow \boldsymbol{R}^{i+j} f_*\Omega^\bullet_{X/Y}(D_v, D_h). \tag{5-11}$$

Note that the Gauss–Manin connection satisfies Griffiths transversality:

$$\nabla(\mathrm{Fil}^i\mathbb{H}^!) \subset \Omega^1_{Y/S}(\log E_a) \otimes \mathrm{Fil}^{i-1}\mathbb{H}^!,$$

where ! stands for † or nothing. Hence, the Hodge filtration of $\mathbb{H}^!$ induces a decreasing filtration of the de Rham complex $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ by subcomplexes:

$$\mathrm{Fil}^i\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) = \big(\mathrm{Fil}^i\mathbb{H}^! \xrightarrow{\nabla} \Omega^1_{Y/S}(\log E_a) \otimes \mathrm{Fil}^{i-1}\mathbb{H}^! \xrightarrow{\nabla} \cdots \xrightarrow{\nabla} \Omega^j_{Y/S}(\log E_a) \otimes \mathrm{Fil}^{i-j}\mathbb{H}^! \xrightarrow{\nabla} \cdots\big). \tag{5-12}$$

Assume the Hodge spectral sequences (5-10) and (5-11) degenerate in E_1. Then

$$\textstyle\bigoplus_j R^{j-i} f_*\Omega^i_{X/Y}(\log D/E_a) \xrightarrow{\sim} \mathrm{gr}^i\,\mathbb{H}, \tag{5-13}$$
$$\textstyle\bigoplus_j R^{j-i} f_*\Omega^i_{X/Y}(D_v, D_h) \xrightarrow{\sim} \mathrm{gr}^i\,\mathbb{H}^\dagger. \tag{5-14}$$

An argument similar to that of [Katz 1970] shows that the Gauss–Manin connection $\nabla : \mathrm{gr}^i\,\mathbb{H}^! \to \Omega^1_{Y/S}(\log E_a) \otimes \mathrm{gr}^{i-1}\,\mathbb{H}^!$ can be identified with the cup product by the Kodaira–Spencer class $c \in \mathrm{Ext}^1_{\mathbb{O}_X}\big(\Omega^1_{X/Y}(\log D/E_a), f^*\Omega^1_{Y/S}(\log E_a)\big)$ defined by (5-1). For this reason, the graded complex of $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ associated with the Hodge filtration (5-12) is called the Kodaira–Spencer complex:

$$\mathrm{gr}^i\,\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) = \big(\mathrm{gr}^i\,\mathbb{H}^! \xrightarrow{\nabla} \Omega^1_{Y/S}(\log E_a) \otimes \mathrm{gr}^{i-1}\,\mathbb{H}^! \xrightarrow{\nabla} \cdots \xrightarrow{\nabla} \Omega^j_{Y/S}(\log E_a) \otimes \mathrm{gr}^{i-j}\,\mathbb{H}^! \xrightarrow{\nabla} \cdots\big), \tag{5-15}$$

where ! stands for † or nothing.

Definition 5.3. The conjugate filtrations of $\mathbb{H}$ and $\mathbb{H}^\dagger$ are increasing filtrations respectively defined by

$$\mathrm{Fil}_i\mathbb{H} = \mathrm{Im}\big(\textstyle\bigoplus_j \boldsymbol{R}^j f_*(\tau_{\leq i}\Omega^\bullet_{X/Y}(\log D/E_a)) \to \bigoplus_j \boldsymbol{R}^j f_*\Omega^\bullet_{X/Y}(\log D/E_a)\big),$$
$$\mathrm{Fil}_i\mathbb{H}^\dagger = \mathrm{Im}\big(\textstyle\bigoplus_j \boldsymbol{R}^j f_*(\tau_{\leq i}\Omega^\bullet_{X/Y}(D_v, D_h)) \to \bigoplus_j \boldsymbol{R}^j f_*\Omega^\bullet_{X/Y}(D_v, D_h)\big),$$

which induce the conjugate spectral sequences

$$ {}_cE_2^{ij} = R^i f_*\mathcal{H}^j\big(\Omega^\bullet_{X/Y}(\log D/E_a)\big) \Rightarrow \boldsymbol{R}^{i+j} f_*\Omega^\bullet_{X/Y}(\log D/E_a), \tag{5-16} $$

$$ {}_cE_2^{ij} = R^i f_*\mathcal{H}^j\big(\Omega^\bullet_{X/Y}(D_v, D_h)\big) \Rightarrow \boldsymbol{R}^{i+j} f_*\Omega^\bullet_{X/Y}(D_v, D_h). \tag{5-17} $$

The conjugate filtration is stable under the Gauss–Manin connection, i.e.,

$$ \nabla(\mathrm{Fil}_i\mathbb{H}^!) \subseteq \Omega^1_{Y/S}(\log E_a)\otimes \mathrm{Fil}_i\mathbb{H}^!; $$

hence, the conjugate filtration of $\mathbb{H}^!$ induces an increasing filtration of the de Rham complex $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ by subcomplexes:

$$ \mathrm{Fil}_i\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) = \big(\mathrm{Fil}_i\mathbb{H}^! \xrightarrow{\nabla} \Omega^1_{Y/S}(\log E_a)\otimes \mathrm{Fil}_i\mathbb{H}^! \xrightarrow{\nabla} \cdots \xrightarrow{\nabla} \Omega^j_{Y/S}(\log E_a)\otimes \mathrm{Fil}_i\mathbb{H}^! \xrightarrow{\nabla} \cdots\big). \tag{5-18} $$

[-4pt]From the increasing filtration Fil_i of $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$, we obtain a decreasing filtration Fil_{-i} of $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$, which gives rise to a spectral sequence

$$ E_1^{ij} = \mathcal{H}^{i+j}\big(\mathrm{gr}_{-i}\,\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big) \Rightarrow \mathcal{H}^{i+j}\big(\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big), \tag{5-19} $$

where ! stands for † or nothing.

From now on, let $\widetilde{S}$ be a lifting of S over $\mathbb{Z}/p^2\mathbb{Z}$, and $F_{\widetilde{S}} : \widetilde{S} \to \widetilde{S}$ a lifting of the absolute Frobenius morphism $F_S : S \to S$ over $\widetilde{S}$. We need the following assumptions:

Assumption 5.4. (i) $f : X \to Y$ is proper and of relative dimension $\leq p$;

(ii) $f : (X, D) \to (Y, E_a)$ has a lifting $\tilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ over $\widetilde{S}$; and

(iii) $F_{Y/S} : Y \to Y_1$ has a lifting $\widetilde{F}_{Y/S} : \widetilde{Y} \to \widetilde{Y}_1$ over $\widetilde{S}$, compatible with $\widetilde{E}_a$.

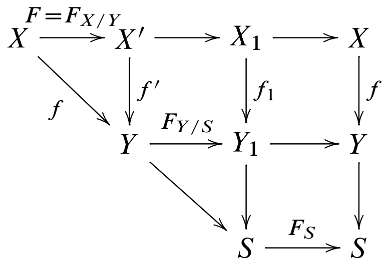
Under Assumption 5.4, by Corollaries 3.4 and 4.4, we have that for any i, j, $R^j f_*\Omega^i_{X/Y}(\log D/E_a)$ and $R^j f_*\Omega^i_{X/Y}(D_v, D_h)$ are locally free of finite type, and of formation compatible with any base change, and that the Hodge spectral sequences (5-10) and (5-11) degenerate in E_1. Furthermore:

Lemma 5.5. *Under Assumption 5.4, the conjugate spectral sequences* (5-16) *and* (5-17) *degenerate in* E_2.

Proof. It is a direct consequence of the Cartier isomorphism. For the degeneracy of (5-16), we use Proposition 2.7 and the degeneracy of (5-10). For the degeneracy of (5-17), we use Proposition 4.1 and the degeneracy of (5-11). □

For the reader's convenience, we recall the following commutative diagram with cartesian squares:

(5-20)



In the rest of this section, we assume that ! in $\mathbb{H}^!$ stands for † or nothing, unless otherwise stated. The degeneracy of the conjugate spectral sequences (5-16) and (5-17) in E_2 gives rise to the isomorphisms

$$\begin{aligned} \operatorname{gr}_i \mathbb{H} &\xrightarrow{\sim} \textstyle\bigoplus_j R^{j-i} f_* \mathcal{H}^i\big(\Omega^\bullet_{X/Y}(\log D/E_a)\big) \\ &\xrightarrow{\sim} \textstyle\bigoplus_j R^{j-i} f'_* \mathcal{H}^i\big(F_* \Omega^\bullet_{X/Y}(\log D/E_a)\big), \\ \operatorname{gr}_i \mathbb{H}^\dagger &\xrightarrow{\sim} \textstyle\bigoplus_j R^{j-i} f_* \mathcal{H}^i\big(\Omega^\bullet_{X/Y}(D_v, D_h)\big) \\ &\xrightarrow{\sim} \textstyle\bigoplus_j R^{j-i} f'_* \mathcal{H}^i\big(F_* \Omega^\bullet_{X/Y}(D_v, D_h)\big). \end{aligned}$$

By the Cartier isomorphisms and the base changes in (5-20), we have

$$\begin{aligned} \operatorname{gr}_i \mathbb{H} &\xrightarrow{\sim} \textstyle\bigoplus_j R^{j-i} f'_* \Omega^i_{X'/Y}(\log D'/E_a) \\ &\xrightarrow{\sim} \textstyle\bigoplus_j F^*_{Y/S} R^{j-i} f_{1*} \Omega^i_{X_1/Y_1}(\log D_1/E_{a1}) = F^*_{Y/S} \operatorname{gr}^i \mathbb{H}_1 = \operatorname{gr}^i \mathbb{H}_1 \otimes \mathcal{O}_Y, \\ \operatorname{gr}_i \mathbb{H}^\dagger &\xrightarrow{\sim} \textstyle\bigoplus_j R^{j-i} f'_* \Omega^i_{X'/Y}(D'_v, D'_h) \\ &\xrightarrow{\sim} \textstyle\bigoplus_j F^*_{Y/S} R^{j-i} f_{1*} \Omega^i_{X_1/Y_1}(D_{v1}, D_{h1}) = F^*_{Y/S} \operatorname{gr}^i \mathbb{H}^\dagger_1 = \operatorname{gr}^i \mathbb{H}^\dagger_1 \otimes \mathcal{O}_Y. \end{aligned}$$

By [Katz 1970, 2.3.1.3], the Gauss–Manin connection satisfies $\nabla_{\operatorname{gr}_i} = 1 \otimes d$ under these isomorphisms; hence, we obtain the following isomorphism of complexes, where the left one is the graded complex associated with (5-18), and the differential of the right one is $1 \otimes d$:

$$F_{Y/S*} \operatorname{gr}_i \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) \xrightarrow{\sim} \operatorname{gr}^i \mathbb{H}^!_1 \otimes F_{Y/S*} \Omega^\bullet_{Y/S}(\log E_a). \tag{5-21}$$

Since $\operatorname{gr}^i \mathbb{H}^!_1$ is locally free, we have the isomorphism for E_1 terms in (5-19):

$$\begin{aligned} E_1^{-i+j,i}\big(F_{Y/S*} \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!), \operatorname{Fil}_\bullet\big) &= \mathcal{H}^j\big(F_{Y/S*} \operatorname{gr}_{i-j} \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big) \\ &\xrightarrow{\sim} \operatorname{gr}^{i-j} \mathbb{H}^!_1 \otimes \mathcal{H}^j\big(F_{Y/S*} \Omega^\bullet_{Y/S}(\log E_a)\big) \overset{C}{\xrightarrow{\sim}} \operatorname{gr}^{i-j} \mathbb{H}^!_1 \otimes \Omega^j_{Y_1/S}(\log E_{a1}), \end{aligned}$$

whose inverse is called the Cartier isomorphism for $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$:

$$(5\text{-}22)\quad C^{-1} : \mathrm{gr}^{i-j}\,\mathbb{H}^!_1 \otimes \Omega^j_{Y_1/S}(\log E_{a1}) \xrightarrow{\sim} E_1^{-i+j,i}\big(F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!), \mathrm{Fil}_\bullet\big).$$

The left-hand side of (5-22) is the j-term in the Kodaira–Spencer complex (5-15) of $\mathbb{H}^!_1$ on Y_1. It follows from [Katz 1970, 3.2] that the right-hand side of (5-22) with the differential d_1 up to sign corresponds to the Kodaira–Spencer complex.

By definition (see [Deligne 1971, 1.3.3]), the delayed filtration $G_\bullet = \mathrm{Dec}(\mathrm{Fil}_\bullet)$ associated with the conjugate filtration $\mathrm{Fil}_\bullet$ is an increasing filtration of the complex $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$, which is defined by

$$G_i\,\Omega^j_{Y/S}(\log E_a)(\mathbb{H}^!) = \big\{x \in \Omega^j_{Y/S}(\log E_a) \otimes \mathrm{Fil}_{i-j}\mathbb{H}^! \mid \nabla(x) \in \Omega^{j+1}_{Y/S}(\log E_a) \otimes \mathrm{Fil}_{i-j-1}\mathbb{H}^!\big\}.$$

Similarly, we also have an increasing filtration of $F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ by subcomplexes of $\mathcal{O}_{Y_1}$-modules. There is a natural surjective homomorphism

$$\mathrm{gr}^G_i\,\Omega^j_{Y/S}(\log E_a)(\mathbb{H}^!) \to E_1^{-i+j,i}\big(\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!), \mathrm{Fil}_\bullet\big),$$

which is indeed an isomorphism and induces isomorphisms for all $r \geq 1$:

$$E_r\big(\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!), G_\bullet\big) \xrightarrow{\sim} E_{r+1}\big(\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!), \mathrm{Fil}_\bullet\big).$$

Objective. Under Assumption 5.4, we shall construct a decomposition in $D(Y_1)$:

$$G_{p-1}F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) \xrightarrow{\sim} \bigoplus_{i<p} \mathrm{gr}^i\,\Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1).$$

Fix $i < p$. For any $j \geq 0$, the decompositions in Theorems 3.2 and 4.2 give rise to morphisms in $D(X')$:

$$(5\text{-}23)\quad \phi^{i-j}_{(\tilde{f},\tilde{F}_{Y/S})} : \Omega^{i-j}_{X'/Y}(\log D'/E_a)[-i+j] \to \tau_{\leq i-j}F_*\Omega^\bullet_{X/Y}(\log D/E_a),$$

$$(5\text{-}24)\quad \phi^{i-j}_{(\tilde{f},\tilde{F}_{Y/S})} : \Omega^{i-j}_{X'/Y}(D'_v, D'_h)[-i+j] \to \tau_{\leq i-j}F_*\Omega^\bullet_{X/Y}(D_v, D_h).$$

Applying $\bigoplus_k \boldsymbol{R}^k f'_*$ to (5-23) and (5-24), we obtain these homomorphisms of $\mathcal{O}_Y$-modules:

$$(5\text{-}25)\quad u^{i-j} : \mathrm{gr}^{i-j}\,\mathbb{H}' = \textstyle\bigoplus_k R^{k-i+j}f'_*\Omega^{i-j}_{X'/Y}(\log D'/E_a) \to \mathrm{Fil}_{i-j}\big(\bigoplus_k \boldsymbol{R}^k f_*\Omega^\bullet_{X/Y}(\log D/E_a)\big) = \mathrm{Fil}_{i-j}\mathbb{H},$$

$$(5\text{-}26)\quad u^{i-j} : \mathrm{gr}^{i-j}\,\mathbb{H}^{\dagger\prime} = \textstyle\bigoplus_k R^{k-i+j}f'_*\Omega^{i-j}_{X'/Y}(D'_v, D'_h) \to \mathrm{Fil}_{i-j}\big(\bigoplus_k \boldsymbol{R}^k f_*\Omega^\bullet_{X/Y}(D_v, D_h)\big) = \mathrm{Fil}_{i-j}\mathbb{H}^\dagger.$$

On the other hand, $\widetilde{F}_{Y/S}$ gives rise to the homomorphism

$$v^1 = \widetilde{F}^*_{Y/S}/p : \Omega^1_{Y_1/S}(\log E_{a1}) \to \mathscr{Z}^1\big(F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)\big),$$

which, by exterior product, induces the homomorphism

$$v^j : \Omega^j_{Y_1/S}(\log E_{a1}) \to \mathscr{Z}^j\big(F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)\big) \subset F_{Y/S*}\Omega^j_{Y/S}(\log E_a). \tag{5-27}$$

By adjunction of $(F^*_{Y/S}, F_{Y/S*})$ and abuse of notation, (5-25) and (5-26) yield the homomorphism

$$u^{i-j} : \mathrm{gr}^{i-j}\, \mathbb{H}^!_1 \to F_{Y/S*}\mathrm{Fil}_{i-j}\mathbb{H}^!. \tag{5-28}$$

Combining (5-27) and (5-28), we obtain the homomorphism of $\mathbb{O}_{Y_1}$-modules:

$$v^j \otimes u^{i-j} : \Omega^j_{Y_1/S}(\log E_{a1}) \otimes \mathrm{gr}^{i-j}\, \mathbb{H}^!_1 \to F_{Y/S*}\big(\Omega^j_{Y/S}(\log E_a) \otimes \mathrm{Fil}_{i-j}\mathbb{H}^!\big). \tag{5-29}$$

Proposition 5.6. *Under Assumption 5.4, we have*:

(i) *The image of* $v^j \otimes u^{i-j}$ *is contained in* $G_i\, F_{Y/S*}\Omega^j_{Y/S}(\log E_a)(\mathbb{H}^!)$, *where* $G_\bullet$ *is the delayed filtration.*

(ii) *For any* $i < p$ *and any* $j \geq 0$, *the following square is commutative*:

$$\begin{array}{ccc} \Omega^j_{Y_1/S}(\log E_{a1}) \otimes \mathrm{gr}^{i-j}\, \mathbb{H}^!_1 & \xrightarrow{\nabla} & \Omega^{j+1}_{Y_1/S}(\log E_{a1}) \otimes \mathrm{gr}^{i-j-1}\, \mathbb{H}^!_1 \\ \downarrow{\scriptstyle v^j \otimes u^{i-j}} & & \downarrow{\scriptstyle v^{j+1} \otimes u^{i-j-1}} \\ F_{Y/S*}\big(\Omega^j_{Y/S}(\log E_a)(\mathbb{H}^!)\big) & \xrightarrow{F_{Y/S*}\nabla} & F_{Y/S*}\big(\Omega^{j+1}_{Y/S}(\log E_a)(\mathbb{H}^!)\big), \end{array} \tag{5-30}$$

where the upper horizontal morphism is the differential map of the Kodaira–Spencer complex. Parts (*i*) *and* (*ii*) *give rise to the morphism of complexes*

$$(v \otimes u)^i : \mathrm{gr}^i\, \Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1) \to G_i\, F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!). \tag{5-31}$$

(iii) *The composition of morphisms of complexes*

$$\mathrm{gr}^i\, \Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1) \overset{(5\text{-}31)}{\to} G_i\, F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) \to \mathrm{gr}^G_i\, F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) \to E_1^{-i+\bullet,i}\big(F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!), \mathrm{Fil}_\bullet\big) \tag{5-32}$$

induces the Cartier isomorphism (5-22) *for* $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$; *hence, it is a quasi-isomorphism. Hence, the following morphism is a quasi-isomorphism*:

$$\sum_{i<p} (v \otimes u)^i : \bigoplus_{i<p} \mathrm{gr}^i\, \Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1) \to G_{p-1}\, F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!).$$

The essential point in the proof of Proposition 5.6 is the compatibility (ii), which is deduced from a more general compatibility in the level of derived category between the morphism u and the Gauss–Manin connection ∇.

Lemma 5.7. *The Koszul filtrations $K^\bullet$ of the complexes*

$$\Omega^\bullet_{X/S}(\log D) \quad \textit{and} \quad \Omega^\bullet_{X/S}(D_v, D_h)$$

give rise to short exact sequences of complexes:

$$0 \to \mathrm{gr}_K^{i+1}\,\Omega^\bullet_{X/S}(\log D) \to K^i/K^{i+2}\big(\Omega^\bullet_{X/S}(\log D)\big) \to \mathrm{gr}_K^i\,\Omega^\bullet_{X/S}(\log D) \to 0,$$

$$0 \to \mathrm{gr}_K^{i+1}\Omega^\bullet_{X/S}(D_v,D_h) \to K^i/K^{i+2}\big(\Omega^\bullet_{X/S}(D_v,D_h)\big) \to \mathrm{gr}_K^i\Omega^\bullet_{X/S}(D_v,D_h) \to 0,$$

which yield the connecting morphisms $\partial : \Gamma(i) \to \Gamma'(i)$ in $D(X)$, where

$$\Gamma(i) = \begin{cases} f^*\Omega^i_{Y/S}(\log E_a) \otimes \Omega^{\bullet-i}_{X/Y}(\log D/E_a) & \textit{for } \Omega^\bullet_{X/S}(\log D), \\ f^*\Omega^i_{Y/S}(\log E_a) \otimes \Omega^{\bullet-i}_{X/Y}(D_v, D_h) & \textit{for } \Omega^\bullet_{X/S}(D_v, D_h), \end{cases}$$

$$\Gamma'(i) = \begin{cases} f^*\Omega^{i+1}_{Y/S}(\log E_a) \otimes \Omega^{\bullet-i}_{X/Y}(\log D/E_a) & \textit{for } \Omega^\bullet_{X/S}(\log D), \\ f^*\Omega^{i+1}_{Y/S}(\log E_a) \otimes \Omega^{\bullet-i}_{X/Y}(D_v, D_h) & \textit{for } \Omega^\bullet_{X/S}(D_v, D_h). \end{cases}$$

Then for any i, j, the following square is commutative:

$$\begin{array}{ccc} \Gamma(i) \otimes \Gamma(j) & \longrightarrow & \big(\Gamma'(i) \otimes \Gamma(j)\big) \oplus \big(\Gamma(i) \otimes \Gamma'(j)\big) \\ \big\downarrow \pi & & \big\downarrow \pi+\pi \\ \Gamma(i+j) & \xrightarrow{\quad\partial\quad} & \Gamma'(i+j), \end{array} \tag{5-33}$$

where the upper horizontal morphism is $\partial \otimes 1 + 1 \otimes \partial$, and π is the product morphism composed possibly with an isomorphism of commutativity.

Proof. It suffices to prove that the product morphism

$$\pi : \Omega^\bullet_{X/S}(\log D) \otimes \Omega^\bullet_{X/S}(\log D) \to \Omega^\bullet_{X/S}(\log D)$$

is compatible with the Koszul filtration. Thus we can use the morphisms of the corresponding short exact sequences of $K^n/K^{n+2}\big(\Omega^\bullet_{X/S}(\log D) \otimes \Omega^\bullet_{X/S}(\log D)\big) \to K^n/K^{n+2}\big(\Omega^\bullet_{X/S}(\log D)\big)$ to obtain the conclusion. The proof for $\Omega^\bullet_{X/S}(D_v, D_h)$ is similar. ☐

Applying $\bigoplus_k \boldsymbol{R}^k f_*$ to (5-33), we obtain the commutative square

$$\begin{array}{ccc} \Theta(i) \otimes \Theta(j) & \longrightarrow & \big(\Theta(i+1) \otimes \Theta(j)\big) \oplus \big(\Theta(i) \otimes \Theta(j+1)\big) \\ \big\downarrow \pi & & \big\downarrow \pi+\pi \\ \Theta(i+j) & \xrightarrow{\quad\nabla\quad} & \Theta(i+j+1), \end{array} \tag{5-34}$$

where $\Theta(i)=\Omega^i_{Y/S}(\log E_a)(\mathbb{H}^!)$, the upper horizontal morphism is $\nabla\otimes 1+1\otimes\nabla$, and π is the product morphism composed possibly with an isomorphism of commutativity. The diagram (5-34) implies that the complex $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ is a differential graded module over $\Omega^\bullet_{Y/S}(\log E_a)$.

Lemma 5.8. *For any $i<p$ and any $j\geq 0$, the following squares are commutative*:

$$\begin{array}{ccc}
f_1^*\Omega^j_{Y_1/S}(\log E_{a1})\otimes\Omega^{i-j}_{X_1/Y_1}(\log D_1/E_{a1})[-i+j] & \xrightarrow{\partial} & f_1^*\Omega^{j+1}_{Y_1/S}(\log E_{a1})\otimes\Omega^{i-j-1}_{X_1/Y_1}(\log D_1/E_{a1})[-i+j+1] \\
\Big\downarrow{\scriptstyle v^j\otimes\phi^{i-j}} & & \Big\downarrow{\scriptstyle v^{j+1}\otimes\phi^{i-j-1}} \\
F_{X/S*}\big(f^*\Omega^j_{Y/S}(\log E_a)\otimes\Omega^\bullet_{X/Y}(\log D/E_a)\big) & \xrightarrow{F_{X/S*}\partial} & F_{X/S*}\big(f^*\Omega^{j+1}_{Y/S}(\log E_a)\otimes\Omega^\bullet_{X/Y}(\log D/E_a)\big),
\end{array}$$

$$\begin{array}{ccc}
f_1^*\Omega^j_{Y_1/S}(\log E_{a1})\otimes\Omega^{i-j}_{X_1/Y_1}(D_{v1},D_{h1})[-i+j] & \xrightarrow{\partial} & f_1^*\Omega^{j+1}_{Y_1/S}(\log E_{a1})\otimes\Omega^{i-j-1}_{X_1/Y_1}(D_{v1},D_{h1})[-i+j+1] \\
\Big\downarrow{\scriptstyle v^j\otimes\phi^{i-j}} & & \Big\downarrow{\scriptstyle v^{j+1}\otimes\phi^{i-j-1}} \\
F_{X/S*}\big(f^*\Omega^j_{Y/S}(\log E_a)\otimes\Omega^\bullet_{X/Y}(D_v,D_h)\big) & \xrightarrow{F_{X/S*}\partial} & F_{X/S*}\big(f^*\Omega^{j+1}_{Y/S}(\log E_a)\otimes\Omega^\bullet_{X/Y}(D_v,D_h)\big),
\end{array}$$

where ϕ^{i-j} are deduced from (5-23) *and* (5-24) *by adjunction of* $(F^*_{Y/S},F_{Y/S*})$, *and the upper and lower horizontal morphisms are deduced from the short exact sequences of $K^j/K^{j+2}\big(\Omega^\bullet_{X/S}(\log D)\big)$ and $K^j/K^{j+2}\big(\Omega^\bullet_{X/S}(D_v,D_h)\big)$.*

Proof. We only deal with the case for $\Omega^\bullet_{X/S}(\log D)$. The proof divides into three steps.

Step 1: $i=1$, $j=0$. Recall the definition of $\phi^1:\Omega^1_{X'/Y}(\log D'/E_a)[-1]\to F_*\Omega^\bullet_{X/Y}(\log D/E_a)$ given in Theorem 3.2. We choose an étale covering $\mathcal{U}=(U_i)_{i\in I}$ of X, and a lifting $\widetilde{F}_i:\widetilde{U}_i\to\widetilde{U}'_i$ of F compatible with $\widetilde{D}$ for each $i\in I$. On U'_i, we take

$$f_i=\widetilde{F}_i^*/p:\Omega^1_{X'/Y}(\log D'/E_a)|_{U'_i}\to F_*\Omega^1_{X/Y}(\log D/E_a)|_{U'_i}.$$

On $U'_{ij}=U'_i\cap U'_j$, we take

$$h_{ij}=(\widetilde{F}_j^*-\widetilde{F}_i^*)/p:\Omega^1_{X'/Y}(\log D'/E_a)|_{U'_{ij}}\to F_*\mathcal{O}_X|_{U'_{ij}}.$$

We have $df_i = 0,\ f_j - f_i = dh_{ij},\ h_{ij} + h_{jk} = h_{ik}$. The morphism ϕ^1 is the composition of

$$u = (h_{ij}, f_i) : \Omega^1_{X'/Y}(\log D'/E_a)[-1] \to F_* \check{\mathscr{C}}\big(\mathfrak{U}, \Omega^\bullet_{X/Y}(\log D/E_a)\big)$$

and the inverse of the quasi-isomorphism

$$F_*\Omega^\bullet_{X/Y}(\log D/E_a) \to F_*\check{\mathscr{C}}\big(\mathfrak{U}, \Omega^\bullet_{X/Y}(\log D/E_a)\big).$$

By adjunction of $(F^*_{Y/S}, F_{Y/S*})$ and abuse of notation, we have a morphism

$$u : \Omega^1_{X_1/Y_1}(\log D_1/E_{a1})[-1] \to F_{X/S*}\check{\mathscr{C}}\big(\mathfrak{U}, \Omega^\bullet_{X/Y}(\log D/E_a)\big).$$

Similarly, the liftings $\widetilde{U}_i \xrightarrow{\widetilde{F}_i} \widetilde{U}'_i \to (\widetilde{U}_i)_1$ of $F_{X/S}$ provide a morphism

$$u_1 : \Omega^1_{X_1/S}(\log D_1)[-1] \to F_{X/S*}\check{\mathscr{C}}\big(\mathfrak{U}, \Omega^\bullet_{X/S}(\log D)\big).$$

Since $\check{\mathscr{C}}(\mathfrak{U}, \Omega^\bullet_{X/S}(\log D))$ coincides with $\check{\mathscr{C}}(\mathfrak{U}, \Omega^\bullet_{X/S}(\log D)/K^2)$ in degree at most 1, we can consider u_1 with values in $F_{X/S*}\check{\mathscr{C}}(\mathfrak{U}, \Omega^\bullet_{X/S}(\log D)/K^2)$ to obtain a morphism

$$u_1 : \Omega^1_{X_1/S}(\log D_1)[-1] \to F_{X/S*}\check{\mathscr{C}}\big(\mathfrak{U}, \Omega^\bullet_{X/S}(\log D)/K^2\big).$$

Finally, we take $v = \widetilde{F}^*_{Y/S}/p : \Omega^1_{Y_1/S}(\log E_{a1}) \to F_{Y/S*}\Omega^1_{Y/S}(\log E_a)$. By adjunction of $(F^*_{Y/S}, F_{Y/S*})$ and abuse of notation, we have a homomorphism

$$v : F^*_{Y/S}\Omega^1_{Y_1/S}(\log E_{a1}) \to \Omega^1_{Y/S}(\log E_a).$$

Applying f^* to the above homomorphism and using the commutativity of (5-20), we have a homomorphism

$$v : F^*_{X/S} f_1^*\Omega^1_{Y_1/S}(\log E_{a1}) \to f^*\Omega^1_{Y/S}(\log E_a).$$

By adjunction of $(F^*_{X/S}, F_{X/S*})$ and the composition with a natural morphism, we have a morphism

$$\begin{aligned} v : f_1^*\Omega^1_{Y_1/S}(\log E_{a1})[-1] \\ &\to F_{X/S*}\big(f^*\Omega^1_{Y/S}(\log E_a) \otimes \check{\mathscr{C}}(\mathfrak{U}, \Omega^{\bullet-1}_{X/Y}(\log D/E_a))\big) \\ &= F_{X/S*}\check{\mathscr{C}}\big(\mathfrak{U}, f^*\Omega^1_{Y/S}(\log E_a) \otimes \Omega^{\bullet-1}_{X/Y}(\log D/E_a)\big). \end{aligned}$$

We shall prove that v, u_1, and u fit into the following commutative diagram with exact rows:

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & f_1^*\Omega^1_{Y_1/S}(\log E_{a1})[-1] & \longrightarrow & \Omega^1_{X_1/S}(\log D_1)[-1] & \longrightarrow & \Omega^1_{X_1/Y_1}(\log D_1/E_{a1})[-1] & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle v} & & \downarrow{\scriptstyle u_1} & & \downarrow{\scriptstyle u} & & \\
0 & \longrightarrow & F_{X/S*}\check{\mathcal{C}}(\mathcal{U}, f^*\Omega^1_{Y/S}(\log E_a)\otimes\Omega^{\bullet-1}_{X/Y}(\log D/E_a)) & \longrightarrow & F_{X/S*}\check{\mathcal{C}}(\mathcal{U}, \Omega^\bullet_{X/S}(\log D)/K^2) & \longrightarrow & F_{X/S*}\check{\mathcal{C}}(\mathcal{U}, \Omega^\bullet_{X/Y}(\log D/E_a)) & \longrightarrow & 0 \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & F_{X/S*}\big(f^*\Omega^1_{Y/S}(\log E_a)\otimes\Omega^{\bullet-1}_{X/Y}(\log D/E_a)\big)_{q.i.} & \longrightarrow & F_{X/S*}\big(\Omega^\bullet_{X/S}(\log D)/K^2\big)_{q.i.} & \longrightarrow & F_{X/S*}\Omega^\bullet_{X/Y}(\log D/E_a)_{q.i.} & \longrightarrow & 0
\end{array}
$$

Since $v = (0, f_i^Y)$, $u_1 = (h_{ij}^X, f_i^X)$, and $u = (h_{ij}, f_i)$, the upper diagram is commutative. The lower one is a quasi-isomorphism of short exact sequences of complexes. The morphism of distinguished triangles defined by this diagram gives the commutativity of the diagram in Lemma 5.8 for the case $i = 1$, $j = 0$.

Step 2: $j = 0$. Recall that ϕ^i $(1 \le i < p)$ is deduced from ϕ^1 by the composition

$$\Omega^i_{X'/Y}(\log D'/E_a)[-i] \xrightarrow{a} \Omega^1_{X'/Y}(\log D'/E_a)^{\otimes i}[-i] \xrightarrow{(\phi^1)^{\otimes i}} \big(F_*\Omega^\bullet_{X/Y}(\log D/E_a)\big)^{\otimes i} \xrightarrow{\pi} F_*\Omega^\bullet_{X/Y}(\log D/E_a),$$

where π is the product map and a is the antisymmetrization map

$$a(x_1 \wedge \cdots \wedge x_i) = \frac{1}{i!} \sum_{\sigma \in S_i} \mathrm{sgn}(\sigma) x_{\sigma(1)} \wedge \cdots \wedge x_{\sigma(i)}.$$

Consider the diagram

$$\begin{array}{ccc}
\Omega^i_{X_1/Y_1}(\log D_1/E_{a1})[-i] & \xrightarrow{\partial} & f_1^*\Omega^1_{Y_1/S}(\log E_{a1}) \otimes \Omega^{i-1}_{X_1/Y_1}(\log D_1/E_{a1})[-i+1] \\
\downarrow a & & \downarrow 1\otimes a \\
\Omega^1_{X_1/Y_1}(\log D_1/E_{a1})^{\otimes i}[-i] & \longrightarrow & f_1^*\Omega^1_{Y_1/S}(\log E_{a1}) \otimes \big(\Omega^1_{X_1/Y_1}(\log D_1/E_{a1})\big)^{\otimes i-1}[-i+1] \\
\downarrow (\phi^1)^{\otimes i} & & \downarrow 1\otimes(\phi^1)^{\otimes i-1} \\
\big(F_{X/S*}\Omega^\bullet_{X/Y}(\log D/E_a)\big)^{\otimes i} & \longrightarrow & F_{X/S*}\big(f^*\Omega^1_{Y/S}(\log E_a) \otimes (\Omega^\bullet_{X/Y}(\log D/E_a))^{\otimes i-1}\big) \\
\downarrow \pi & & \downarrow \pi \\
F_{X/S*}\Omega^\bullet_{X/Y}(\log D/E_a) & \xrightarrow{F_{X/S*}\partial} & F_{X/S*}\big(f^*\Omega^1_{Y/S}(\log E_a) \otimes \Omega^\bullet_{X/Y}(\log D/E_a)\big),
\end{array}$$

where the unmarked horizontal morphisms are $\sum(1\otimes\cdots\otimes\partial\otimes\cdots\otimes 1)$ on the second row and $\sum F_{X/S*}(1\otimes\cdots\otimes\partial\otimes\cdots\otimes 1)$ on the third. Since the map a is compatible with the Koszul filtration, we obtain the commutativity of the upper diagram. The commutativity of the middle one follows from Step 1, and the commutativity of the lower one follows from Lemma 5.7.

Step 3: general case. It follows from Step 2 since $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ is a graded differential module over $\Omega^\bullet_{Y/S}(\log E_a)$. □

Proof of Proposition 5.6. By applying $\bigoplus_k \boldsymbol{R}^k f_{1*}$ to the diagrams in Lemma 5.8, we obtain the commutativity of the diagram (5-30). By definition, the image of

$v^j \otimes u^{i-j}$ is already contained in $F_{Y/S*}\big(\Omega^j_{Y/S}(\log E_a)\otimes \mathrm{Fil}_{i-j}\mathbb{H}^!\big)$. Since (5-30) is commutative, we have

$$\nabla\big(\mathrm{Im}(v^j \otimes u^{i-j})\big) \subset \mathrm{Im}\,\big(v^{j+1}\otimes u^{i-j-1}\big) \subset F_{Y/S*}\big(\Omega^{j+1}_{Y/S}(\log E_a)\otimes \mathrm{Fil}_{i-j-1}\mathbb{H}^!\big);$$

hence, $\mathrm{Im}(v^j \otimes u^{i-j}) \subset G_i F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ by the definition of G_i. By construction, $\mathrm{gr}^i\, \Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1) \to \mathrm{gr}^G_i\, F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ is a quasi-isomorphism, which implies the last sentence of Equation (5-6). □

We can eliminate the hypothesis (iii) in Assumption 5.4 to obtain the main theorem in this paper:

Theorem 5.9. *Let ! stand for † or nothing. Let $f : X \to Y$ be an E-semistable S-morphism with an adapted divisor D, and $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ a lifting of $f : (X, D) \to (Y, E_a)$ over $\widetilde{S}$. Assume that f is proper and* $\dim(X/Y) \le p$. *Then for any $i < p$, we have a morphism in $D(Y_1)$:*

$$\phi^i = \phi^i_{\widetilde{f}} : \mathrm{gr}^i\, \Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1) \to G_i F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!), \tag{5-35}$$

such that the composition of (5-35) *with the projection onto* gr^G_i *is a quasi-isomorphism, which is the Cartier isomorphism* (5-22).

Furthermore, for any $i < p$, we have an isomorphism in $D(Y_1)$:

$$\phi = \sum_{j\le i} \phi^j : \bigoplus_{j\le i} \mathrm{gr}^j\, \Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1) \xrightarrow{\sim} G_i F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!).$$

For any $i > \dim(X/S)$, we have $\mathrm{gr}^i\, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) = 0$. *Consequently, if* $\dim(X/S) < p$, *the preceding isomorphism gives rise to a decomposition in $D(Y_1)$:*

$$\phi : \bigoplus_{i} \mathrm{gr}^i\, \Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1) \xrightarrow{\sim} F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!). \tag{5-36}$$

Proof. Since a lifting of the relative Frobenius morphism $F_{Y/S} : Y \to Y_1$ always exists locally, Proposition 5.6 is indeed a local version of Theorem 5.9. The idea of the proof is to use Proposition 5.6 to obtain a coherent system of local splittings for $G_{p-1} F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$.

Take an étale covering $\mathcal{U} = (U_i)_{i\in I}$ of Y. By Proposition 5.6, on U_i for any $i \in I$, there is a splitting (v_i, u_i) of $G_{p-1} F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)|_{U_i}$, where v_i, u_i are defined as in (5-27) and (5-28) (see [Illusie 1990, 4.19] for the notion of splitting). By an argument similar to that of [Illusie 1990, Proposition 4.3], on $U_{ij} = U_i \cap U_j$ for any pair (i, j), there is a homomorphism

$$h_{ij} : \Omega^1_{Y_1/S}(\log E_{a1})|_{U_{ij1}} \to F_{Y/S*}\mathcal{O}_Y|_{U_{ij}},$$

such that the following conditions hold:

$$\begin{aligned} v_j^1 - v_i^1 &= dh_{ij} && \text{on } U_{ij}, \\ h_{ij} + h_{jk} &= h_{ik} && \text{on } U_{ijk}, \\ u_j^n - u_i^n &= (u_i h_{ij}) \circ d && \text{on } U_{ij} \text{ for } n < p, \end{aligned}$$

where in the third equality,

$$d : \mathrm{gr}^n\, \mathbb{H}_1^! \to \bigoplus_{0<m\le n} \mathrm{gr}^{n-m}\, \mathbb{H}_1^! \otimes \Gamma^m \Omega^1_{Y_1/S}(\log E_{a1})$$

is the differential map of the complex $NC(\mathrm{gr}^\bullet\, \mathbb{H}_1^!)$ defined as in [Illusie 1990, (4.1.7)], $u_i h_{ij}$ is given by $(u_i h_{ij})(x \otimes a) = u_i(x) h_{ij}(a)$, and the map

$$h_{ij} : \Gamma^m \Omega^1_{Y_1/S}(\log E_{a1}) \to F_{Y/S*}\mathcal{O}_Y$$

on U_{ij} is defined by the polynomial map $x^{[m]} \mapsto h_{ij}(x)^m/m!$ for $x \in \Omega^1_{Y_1/S}(\log E_{a1})$ (note that $\Gamma\Omega^1_{Y_1/S}(\log E_{a1})$ is the divided power algebra of $\Omega^1_{Y_1/S}(\log E_{a1})$).

Thus $G_{p-1}F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ has a coherent system of local splittings $(\mathcal{U} = (U_i), (v_i), (u_i), (h_{ij}))$. It follows from [Illusie 1990, Theorem 4.20] that there exist morphisms ϕ^i (5-35) satisfying all of the required properties. □

Corollary 5.10. *Let $f : X \to Y$ be an E-semistable S-morphism with an adapted divisor D. Assume that f is proper and $g : Y \to S$ is proper. Assume that $f : (X, D) \to (Y, E_a)$ has a lifting $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ over $\widetilde{S}$ and* $\dim(X/S) < p$. *Then the Hodge spectral sequence for $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ and $\boldsymbol{R}g_*$ degenerates in E_1:*

$$E_1^{ij} = \boldsymbol{R}^{i+j} g_*\, \mathrm{gr}^i\, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!) \Rightarrow \boldsymbol{R}^{i+j} g_* \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!),$$

and each E_1^{ij} is locally free of finite type, and of formation compatible with any base change.

Proof. We can use the decomposition (5-36) and an argument analogous to that of [Deligne and Illusie 1987, 4.1.2] to complete the proof. □

Corollary 5.11. *Let K be a field of characteristic zero, $S =$ Spec K, X, Y proper and smooth S-schemes, and $f : X \to Y$ an E-semistable S-morphism with an adapted divisor D. Then the Hodge spectral sequence for $\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)$ degenerates in E_1:*

$$E_1^{ij} = \boldsymbol{H}^{i+j}\big(Y, \mathrm{gr}^i\, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big) \Rightarrow \boldsymbol{H}^{i+j}\big(Y, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big).$$

Proof. It follows from Corollary 5.10 and the standard argument using the reduction modulo p technique (see [Deligne and Illusie 1987, 2.7] and [Illusie 1996]). □

6. Applications to vanishing theorems

In this section, let k be a perfect field of characteristic $p>0$, and $W_2(k)$ the ring of Witt vectors of length two of k. There are some applications of the main theorem to vanishing theorems.

Theorem 6.1. *Let $S=\operatorname{Spec} k$, $\widetilde{S}=\operatorname{Spec} W_2(k)$, and X, Y be proper and smooth S-schemes. Let $f: X\to Y$ be an E-semistable S-morphism with an adapted divisor D, and $\mathscr{L}$ an ample invertible sheaf on Y. Assume that $f:(X,D)\to (Y,E_a)$ has a lifting $\widetilde{f}:(\widetilde{X},\widetilde{D})\to(\widetilde{Y},\widetilde{E}_a)$ over $\widetilde{S}$ and* $\dim(X/S)<p$. *Then*

(6-1) $\quad \boldsymbol{H}^{i+j}\big(Y,\mathscr{L}\otimes \operatorname{gr}^i\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big)=0\quad$ *for any* $i+j>\dim(Y/S)$,

(6-2) $\quad \boldsymbol{H}^{i+j}\big(Y,\mathscr{L}^{-1}\otimes \operatorname{gr}^i\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big)=0\quad$ *for any* $i+j<\dim(Y/S)$.

Proof. We use an argument analogous to those of [Deligne and Illusie 1987, 2.8] and [Illusie 1990, Corollary 4.16]. Let $\mathscr{M}$ be an invertible sheaf on Y. Define

$$h^{ij}(\mathscr{M})=\dim \boldsymbol{H}^{i+j}\big(Y,\mathscr{M}\otimes \operatorname{gr}^i\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big).$$

Then for all n, we have

$$\sum_{i+j=n} h^{ij}(\mathscr{M})\le \sum_{i+j=n} h^{ij}(\mathscr{M}^p). \tag{6-3}$$

Indeed, denote by $\mathscr{M}_1$ the inverse image of $\mathscr{M}$ on Y_1; then we have $\mathscr{M}^p=F^*_{Y/S}\mathscr{M}_1$. The Hodge spectral sequence

$$E_1^{ij}=\boldsymbol{H}^{i+j}\big(Y_1,\mathscr{M}_1\otimes F_{Y/S*}\operatorname{gr}^i\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big)\Rightarrow \boldsymbol{H}^{i+j}\big(Y_1,\mathscr{M}_1\otimes F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big)$$

gives rise to the inequality

$$\dim \boldsymbol{H}^{n}\big(Y_1,\mathscr{M}_1\otimes F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big)\le \sum_{i+j=n} h^{ij}(\mathscr{M}^p).$$

On the other hand, by the decomposition (5-36), we have

$$\dim \boldsymbol{H}^{n}\big(Y_1,\mathscr{M}_1\otimes F_{Y/S*}\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big) = \sum_{i+j=n}\dim \boldsymbol{H}^{i+j}\big(Y_1,\mathscr{M}_1\otimes \operatorname{gr}^i\Omega^\bullet_{Y_1/S}(\log E_{a1})(\mathbb{H}^!_1)\big)=\sum_{i+j=n} h^{ij}(\mathscr{M}),$$

which proves (6-3).

Next, we shall prove $h^{ij}(\mathscr{L}^{p^N})=0$ for N sufficiently large and for all $i+j>\dim(Y/S)$. The stupid filtration of the Kodaira–Spencer complex

$$\operatorname{gr}^i\Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)=\big(\operatorname{gr}^i\mathbb{H}^!\xrightarrow{\nabla}\Omega^1_{Y/S}(\log E_a)\otimes \operatorname{gr}^{i-1}\mathbb{H}^!\xrightarrow{\nabla}\cdots\big)$$

gives rise to the spectral sequence

$$E_1^{rs} = H^s\big(Y, \mathscr{L}^{p^N} \otimes \Omega^r_{Y/S}(\log E_a) \otimes \mathrm{gr}^{i-r}\, \mathbb{H}^!\big) \Rightarrow \boldsymbol{H}^{r+s}\big(Y, \mathscr{L}^{p^N} \otimes \mathrm{gr}^i\, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big).$$

We focus on terms with $r+s = i+j > \dim(Y/S)$. If $s=0$ then $\Omega^r_{Y/S}(\log E_a) = 0$. If $s > 0$ then the choices of r and s are finite. By the Serre vanishing theorem, we can choose N sufficiently large that $E_1^{rs} = 0$ for all r and s; hence, $h^{ij}(\mathscr{L}^{p^N}) = 0$ holds for all $i + j > \dim(Y/S)$. Thanks to (6-3), we obtain the vanishing (6-1).

By a similar argument, we can prove $h^{ij}(\mathscr{L}^{-p^N}) = 0$ for N sufficiently large and for all $i + j < \dim(Y/S)$. Thanks to (6-3), we obtain the vanishing (6-2). □

Corollary 6.2. *Let K be a field of characteristic zero with $S = \operatorname{Spec} K$, and let X, Y be proper and smooth S-schemes. Let $f : X \to Y$ be an E-semistable S-morphism with an adapted divisor D, and $\mathscr{L}$ an ample invertible sheaf on Y. Then*

$$\boldsymbol{H}^{i+j}\big(Y, \mathscr{L} \otimes \mathrm{gr}^i\, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big) = 0 \quad \textit{for any } i + j > \dim(Y/S),$$

$$\boldsymbol{H}^{i+j}\big(Y, \mathscr{L}^{-1} \otimes \mathrm{gr}^i\, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^!)\big) = 0 \quad \textit{for any } i + j < \dim(Y/S).$$

Proof. It follows from Theorem 6.1 and the reduction modulo p technique. □

Theorem 6.3. *Set $S = \operatorname{Spec} k$ and $\widetilde{S} = \operatorname{Spec} W_2(k)$, and let X, Y be proper and smooth S-schemes. Let $f : X \to Y$ be an E-semistable S-morphism with an adapted divisor D, and $\mathscr{L}$ an ample invertible sheaf on Y. Assume that $f : (X, D) \to (Y, E_a)$ has a lifting $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ over $\widetilde{S}$ and $\dim(X/S) < p$. Then*

$$H^i\big(Y, \mathscr{L} \otimes R^j f_* \omega_{X/S}(D)\big) = 0 \quad \textit{and} \quad H^i\big(Y, \mathscr{L} \otimes R^j f_* \omega_{X/S}(D_h)\big) = 0$$

for any $i > 0$ and $j \geq 0$.

Proof. Suppose $\dim(X/S) = n$, $\dim(X/Y) = d$, and $\dim(Y/S) = e$.

(1) Consider $\mathrm{gr}^n\, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H})$, whose k-th component is

$$\Omega^k_{Y/S}(\log E_a) \otimes \mathrm{gr}^{n-k}\, \mathbb{H} = \Omega^k_{Y/S}(\log E_a) \otimes \big(\textstyle\bigoplus_l R^{l-n+k} f_* \Omega^{n-k}_{X/Y}(\log D/E_a)\big).$$

Since $\Omega^k_{Y/S}(\log E_a) = 0$ for any $k > e$ and $\Omega^{n-k}_{X/Y}(\log D/E_a) = 0$ for any $k < e$, we have

$$\begin{aligned}\mathrm{gr}^n\, \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}) &= \omega_{Y/S}(E_a) \otimes \big(\textstyle\bigoplus_l R^{l-d} f_* \omega_{X/Y}(D_h)\big)[-e] \\ &= \textstyle\bigoplus_{k \geq 0} R^k f_* \omega_{X/S}(D)[-e].\end{aligned}$$

In Theorem 6.1, taking $i = n$ and $r = i + j - e > 0$, we have

$$\bigoplus_{k \geq 0} H^r\big(Y, \mathscr{L} \otimes R^k f_* \omega_{X/S}(D)\big) = 0,$$

that is, $H^r\big(Y, \mathscr{L}\otimes R^k f_*\omega_{X/S}(D)\big)=0$ for any $r>0$ and $k\geq 0$.

(2) Consider $\operatorname{gr}^n \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^\dagger)$, whose k-th component is

$$\Omega^k_{Y/S}(\log E_a)\otimes \operatorname{gr}^{n-k}\mathbb{H}^\dagger = \Omega^k_{Y/S}(\log E_a)\otimes \big(\textstyle\bigoplus_l R^{l-n+k} f_*\Omega^{n-k}_{X/Y}(D_v, D_h)\big).$$

Since $\Omega^k_{Y/S}(\log E_a)=0$ for any $k>e$ and $\Omega^{n-k}_{X/Y}(D_v, D_h)=0$ for any $k<e$, we have

$$\begin{aligned}\operatorname{gr}^n \Omega^\bullet_{Y/S}(\log E_a)(\mathbb{H}^\dagger) &= \omega_{Y/S}(E_a)\otimes\big(\textstyle\bigoplus_l R^{l-d} f_*\omega_{X/Y}(D_h-D_v)\big)[-e]\\ &= \textstyle\bigoplus_{k\geq 0} R^k f_*\omega_{X/S}(D_h)[-e].\end{aligned}$$

In Theorem 6.1, taking $i=n$ and $r=i+j-e>0$, we have

$$\bigoplus_{k\geq 0} H^r\big(Y, \mathscr{L}\otimes R^k f_*\omega_{X/S}(D_h)\big)=0,$$

that is, $H^r\big(Y, \mathscr{L}\otimes R^k f_*\omega_{X/S}(D_h)\big)=0$ for any $r>0$ and $k\geq 0$. □

In order to give further applications, we need the following:

Definition 6.4 [Xie 2010, Definition 2.3]. Let X be a smooth scheme over k. X is said to be *strongly liftable* over $W_2(k)$ if

(i) X is liftable over $W_2(k)$, and
(ii) there is a lifting $\widetilde{X}$ of X, such that for any prime divisor D on X, (X, D) has a lifting $(\widetilde{X}, \widetilde{D})$ over $W_2(k)$, where $\widetilde{X}$ is fixed for all D.

It was proved in [Xie 2010; 2011] that $\mathbb{A}^n_k$, $\mathbb{P}^n_k$, smooth projective curves, smooth projective rational surfaces, certain smooth complete intersections in $\mathbb{P}^n_k$, and smooth toric varieties are strongly liftable over $W_2(k)$. As a consequence of Theorem 6.3, we can obtain some vanishing results for certain strongly liftable varieties.

Corollary 6.5. *Let $X=X(\Delta,k)$ be a smooth projective toric variety associated with a fan Δ with* $\operatorname{char} k = p > \dim X$, *$Y$ a smooth projective variety over k, $f: X\to Y$ an E-semistable morphism with an adapted divisor D, and $\mathscr{L}$ an ample invertible sheaf on Y. Then $f:(X,D)\to(Y,E_a)$ has a lifting $\tilde{f}:(\widetilde{X},\widetilde{D})\to(\widetilde{Y},\widetilde{E}_a)$ over $W_2(k)$. Consequently, we have*

$$H^i\big(Y, \mathscr{L}\otimes R^j f_*\omega_X(D)\big)=0 \quad \text{and} \quad H^i\big(Y, \mathscr{L}\otimes R^j f_*\omega_X(D_h)\big)=0$$

for any $i>0$ and $j\geq 0$.

Proof. Let H be a general very ample effective divisor on Y and $F=f^{-1}(H)$ the divisor on X. By [Xie 2011, Theorem 3.1], X is strongly liftable; hence, there are a lifting $\widetilde{X}=X\big(\Delta, W_2(k)\big)$ of $X=X(\Delta,k)$ and a lifting $\widetilde{D}+\widetilde{F}\subset\widetilde{X}$ of $D+F\subset X$ over $W_2(k)$. More precisely, let G be a torus invariant divisor on X

determined by the data $\{u(\sigma)\} \in \varprojlim M/M(\sigma)$ such that G is linearly equivalent to F. Then we can construct a torus invariant divisor $\widetilde{G}$ on $\widetilde{X}$ determined by the same data $\{u(\sigma)\}$ and prove that the natural map $H^0(\widetilde{X}, \widetilde{G}) \to H^0(X, G)$ is surjective. Thus we can take a lifting $\widetilde{F}$ of F such that $\widetilde{F}$ is linearly equivalent to $\widetilde{G}$.

By definition, the linear system $|F|$ is basepoint-free; hence, so is $|G|$. By [Fulton 1993, p. 68, Proposition], the continuous piecewise linear function ψ_G on $|\Delta|$ defined in [Fulton 1993, p. 66] is upper convex. Since the functions $\psi_{\widetilde{G}}$ and ψ_G are the same, $\psi_{\widetilde{G}}$ is also upper convex. Thus the linear system $|\widetilde{G}|$ is basepoint-free; hence, so is $|\widetilde{F}|$. Thus the linear system $|\widetilde{F}|$ defines a $W_2(k)$-morphism $\widetilde{f} : \widetilde{X} \to \widetilde{Y}$.

It is easy to verify that $\widetilde{Y}$ is a lifting of Y and $\widetilde{f}$ is a lifting of f over $W_2(k)$. By [Esnault and Viehweg 1992, Lemmas 8.13, 8.14] or [Xie 2011, Lemma 2.2], $\widetilde{D}$ is relatively simple normal crossing over $W_2(k)$. Hence, we can verify that $\widetilde{f} : \widetilde{X} \to \widetilde{Y}$ is an $\widetilde{E}$-semistable morphism and $\widetilde{D}$ is adapted to $\widetilde{f}$, which imply that $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ is a lifting of $f : (X, D) \to (Y, E_a)$ over $W_2(k)$. By Theorem 6.3, we obtain the required vanishings. □

Corollary 6.6. *Let X be a smooth projective rational surface over k with* $\operatorname{char} k = p > 3$, *$f : X \to \mathbb{P}^1_k$ an E-semistable morphism with an adapted divisor D, and $\mathscr{L}$ an ample invertible sheaf on $\mathbb{P}^1_k$. Then $f : (X, D) \to (Y, E_a)$ has a lifting $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\widetilde{Y}, \widetilde{E}_a)$ over $W_2(k)$. Consequently, we have*

$$H^i\big(\mathbb{P}^1_k, \mathscr{L} \otimes R^j f_* \omega_X(D)\big) = 0 \text{ and } H^i\big(\mathbb{P}^1_k, \mathscr{L} \otimes R^j f_* \omega_X(D_h)\big) = 0$$

for any $i > 0$ and $j \geq 0$.

Proof. Let $P \in \mathbb{P}^1_k$ be a general point and $F = f^{-1}(P)$ the fiber of f. By [Xie 2010, Theorem 1.3], X is strongly liftable; hence, there are a lifting $\widetilde{X}$ of X and a lifting $\widetilde{D} + \widetilde{F} \subset \widetilde{X}$ of $D + F \subset X$ over $W_2(k)$. Since both X and $\widetilde{X}$ are birational to certain smooth projective toric surfaces through a sequence of blow-ups along some closed points, by an argument similar to the proof of Corollary 6.5, we can show that the linear system $|\widetilde{F}|$ is basepoint-free, which gives rise to a $W_2(k)$-morphism $\widetilde{f} : \widetilde{X} \to \mathbb{P}^1_{W_2(k)}$. By [Esnault and Viehweg 1992, Lemmas 8.13, 8.14] or [Xie 2011, Lemma 2.2], it is easy to verify that $\widetilde{f} : (\widetilde{X}, \widetilde{D}) \to (\mathbb{P}^1_{W_2(k)}, \widetilde{E}_a)$ is a lifting of $f : (X, D) \to (\mathbb{P}^1_k, E_a)$ over $W_2(k)$. By Theorem 6.3, we obtain the required vanishings. □

Acknowledgments

I would like to express my gratitude to Professor Luc Illusie for invaluable advice and warm encouragement. I am grateful to the referee for pointing out some inaccuracies and giving many useful suggestions that which make this paper more readable.

References

[Deligne 1971] P. Deligne, “Théorie de Hodge. II”, *Inst. Hautes Études Sci. Publ. Math.* **40** (1971), 5–57. MR 58 #16653a Zbl 0219.14007

[Deligne and Illusie 1987] P. Deligne and L. Illusie, “Relèvements modulo p^2 et décomposition du complexe de de Rham”, *Invent. Math.* **89**:2 (1987), 247–270. MR 88j:14029 Zbl 0632.14017

[Esnault and Viehweg 1992] H. Esnault and E. Viehweg, *Lectures on vanishing theorems*, DMV Seminar **20**, Birkhäuser, Basel, 1992. MR 94a:14017 Zbl 0779.14003

[Fulton 1993] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies **131**, Princeton Univ. Press, Princeton, NJ, 1993. MR 94g:14028 Zbl 0813.14039

[Illusie 1990] L. Illusie, “Réduction semi-stable et décomposition de complexes de de Rham à coefficients”, *Duke Math. J.* **60**:1 (1990), 139–185. MR 91d:14007

[Illusie 1996] L. Illusie, “Frobenius et dégénérescence de Hodge”, pp. 113–168 in *Introduction à la théorie de Hodge*, Panor. Synthèses **3**, Soc. Math. France, Paris, 1996. MR 1409820

[Katz 1970] N. M. Katz, “Nilpotent connections and the monodromy theorem: Applications of a result of Turrittin”, *Inst. Hautes Études Sci. Publ. Math.* 39 (1970), 175–232. MR 45 #271 Zbl 0221.14007

[Kollár 1995] J. Kollár, *Shafarevich maps and automorphic forms*, Princeton Univ. Press, Princeton, NJ, 1995. MR 96i:14016 Zbl 0871.14015

[Xie 2010] Q. Xie, “Strongly liftable schemes and the Kawamata–Viehweg vanishing in positive characteristic”, *Math. Res. Lett.* **17**:3 (2010), 563–572. MR 2011j:14053 Zbl 1223.14026

[Xie 2011] Q. Xie, “Strongly liftable schemes and the Kawamata–Viehweg vanishing in positive characteristic, II”, *Math. Res. Lett.* **18**:2 (2011), 315–328. MR 2784674

Received November 11, 2010. Revised August 18, 2011.

QIHONG XIE
SCHOOL OF MATHEMATICAL SCIENCES
FUDAN UNIVERSITY
SHANGHAI, 200433
CHINA

qhxie@fudan.edu.cn
http://homepage.fudan.edu.cn/qhxie/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 254, No. 2, 2011

A DIFFERENTIABLE SPHERE THEOREM INSPIRED BY RIGIDITY OF MINIMAL SUBMANIFOLDS

HONG-WEI XU AND LING TIAN

We prove a vanishing theorem for the fundamental group of a compact submanifold in a space form, and then present a refined version of Ejiri's rigidity theorem for minimal submanifolds in a sphere. Inspired by the refined Ejiri theorem, we verify a new differentiable sphere theorem for compact submanifolds in space forms. We also show that our differentiable sphere theorem is sharp. We emphasize that our method of Ricci flow in the proof of the sphere theorem seems useful in the study of curvature and topology. Also, we obtain a differentiable pinching theorem for compact submanifolds in a Riemannian manifold.

1. Introduction

The investigation of geometrical and topological structures of manifolds plays an important role in global differential geometry. Since the first sphere theorem for Riemannian manifolds was proved by Rauch in 1951, there has been much progress in this field; see [Berger 2000; Brendle 2010; Shiohama 2000; Xu 2012]. Brendle and Schoen [2008] proved a remarkable classification theorem for compact manifolds with weakly $\frac{1}{4}$-pinched curvatures in the pointwise sense, implying this:

Theorem A. *Suppose that M is an n-dimensional complete and simply connected Riemannian manifold such that $\frac{1}{4} \le K_M \le 1$. Then M is either diffeomorphic to S^n or isometric to a compact rank-one symmetric space* (*CROSS*).

Petersen and Tao [2009] improved Brendle and Schoen's pinching constant in Theorem A to $\frac{1}{4} - \varepsilon_n$, with ε_n being a positive constant depending only on n. Motivated by Shen's topological sphere theorem [1989] for compact manifolds with positive Ricci curvature and Yau's open problem on the pinching theorem [Yau 1993, Problem 12], Gu and Xu [2011] obtained a differentiable sphere theorem for compact manifolds with positive scalar curvature. In particular, they proved that

Research supported by the NSFC, Grant No. 11071211; the Trans-Century Training Programme Foundation for Talents by the Ministry of Education of China.
MSC2010: 53C20, 53C24, 53C40.
Keywords: Submanifolds, differentiable sphere theorem, Ricci curvature, Ricci flow, stable currents.

if M is an n-dimensional compact Riemannian manifold ($n \geq 3$) whose sectional curvature and scalar curvature satisfy $K_M \leq 1$ and $R > n(n-1)\delta_n$, then M is diffeomorphic to a spherical space form. Here

$$\delta_n = 1 - \frac{20 - 8\,\mathrm{sgn}(n-3)}{5n(n-1)}$$

and $\mathrm{sgn}(\,\cdot\,)$ is the standard sign function.

In addition, results on sphere theorems for Riemannian submanifolds have been obtained by various authors [Huisken 1987; Lawson and Simons 1973; Shiohama 2000; Shiohama and Xu 1997; Xu and Zhao 2009]. Let M^n be an n-dimensional submanifold in an $(n+p)$-dimensional Riemannian manifold N^{n+p}. Denote by H and S the mean curvature and the squared length of the second fundamental form of M, respectively. Motivated by Lawson, Simons, Shiohama, and Xu's topological sphere theorem [Lawson and Simons 1973; Shiohama and Xu 1997] for submanifolds, Xu and Zhao [2009] proved that for $n \geq 4$, if M is an n-dimensional complete submanifold in S^{n+p} and $S < 2\sqrt{2}$, then M is diffeomorphic to S^n. Let $F^{n+p}(c)$ be an $(n+p)$-dimensional simply connected space form with constant curvature c. Xu and Gu [2010] extended Huisken's sphere theorem [1987] for hypersurfaces in spheres to the case of submanifolds in space forms, and proved the following optimal differentiable sphere theorem:

Theorem B. *Suppose that M is an n-dimensional oriented complete submanifold in $F^{n+p}(c)$ with $c \geq 0$. If*

$$\lambda(M) := \sup_M \left(S - \frac{n^2 H^2}{n-1} - 2c \right) < 0,$$

then M is diffeomorphic to S^n.

After the pioneering rigidity theorem for closed minimal submanifolds in a sphere due to Simons [1968], several fundamental rigidity results for minimal submanifolds were proved by various authors [Chern et al. 1970; Ejiri 1979; Lawson 1969; An-Min and Jimin 1992; Yau 1974; 1975]. Ejiri [1979] obtained the following rigidity theorem for minimal submanifolds in spheres:

Theorem C. *Suppose that M is an n-dimensional simply connected compact orientable minimal submanifold in an $(n+p)$-dimensional unit sphere S^{n+p}, with $n \geq 4$. If the Ricci curvature of M satisfies*

$$\mathrm{Ric}_M \geq n-2,$$

then M is either the totally geodesic submanifold S^n, the Clifford torus $S^m(\sqrt{1/2}) \times S^m(\sqrt{1/2})$ in S^{n+1} with $n = 2m$, or $\mathbb{C}P^2(\frac{4}{3})$ in S^7. Here $\mathbb{C}P^2(\frac{4}{3})$ denotes the 2-dimensional complex projective space minimally immersed into S^7 with constant holomorphic sectional curvature $\frac{4}{3}$.

Shen [1992] proved that if M is a 3-dimensional compact minimal submanifold in a unit sphere S^{3+p}, and if the Ricci curvature of M is larger than 1, then M is the totally geodesic submanifold S^3.

Here we investigate differentiable sphere theorems for compact submanifolds of positive Ricci curvature. Using convergence results of Hamilton [1982] and Brendle [2008] for Ricci flow and the Lawson–Simons–Xin formula for the nonexistence of stable currents [Lawson and Simons 1973; Xin 1984], we prove such a theorem inspired by rigidity of minimal submanifolds. We first prove the following differentiable sphere theorem for submanifolds in a Riemannian manifold.

Theorem 1.1. *Suppose that M is an n-dimensional compact submanifold in an $(n+p)$-dimensional Riemannian manifold N^{n+p}, with $n \geq 4$. Denote by $\overline{K}(x,\pi)$ the sectional curvature of N for tangent 2-plane $\pi(\subset T_xN)$ at point $x \in N$. Set $\overline{K}_{\max}(x) := \max_{\pi\subset T_xN} \overline{K}(x,\pi)$ and $\overline{K}_{\min}(x) := \min_{\pi\subset T_xN} \overline{K}(x,\pi)$. If*

$$\mathrm{Ric}_M > (n-\tfrac{2}{3})\overline{K}_{\max} - \tfrac{4}{3}\overline{K}_{\min} + \tfrac{1}{8}n^2H^2,$$

then the normalized Ricci flow with initial metric g_0

$$\frac{\partial}{\partial t}g(t) = -2\,\mathrm{Ric}_{g(t)} + \frac{2}{n}r_{g(t)}g(t)$$

exists for all time and converges to a constant curvature metric as $t \to \infty$. Also, M is diffeomorphic to a spherical space form. In particular, if M is simply connected, then M is diffeomorphic to S^n.

We give a vanishing theorem for the fundamental group of a submanifold.

Theorem 1.2. *Suppose that M is an n-dimensional compact submanifold in an $(n+p)$-dimensional space form $F^{n+p}(c)$ with $c \geq 0$. If the Ricci curvature of M satisfies*

$$\mathrm{Ric}_M > \tfrac{1}{2}(n-1)c + \tfrac{1}{8}n^2H^2,$$

then $\pi_1(M) = 0$; that is, M is simply connected.

Applying Theorem 1.2 to Theorem C, we obtain a refined version of the Ejiri rigidity theorem for minimal submanifolds in spheres, without the assumption that M is simply connected. Finally, inspired by the refined Ejiri rigidity theorem for minimal submanifolds, we prove the following differentiable sphere theorem for submanifolds in space forms.

Theorem 1.3 (Main Theorem). *Suppose that M is an n-dimensional compact submanifold in an $(n+p)$-dimensional space form $F^{n+p}(c)$, with $c \geq 0$ and $n \geq 3$. If the Ricci curvature of M satisfies*

$$\mathrm{Ric}_M > (n-2)c + \tfrac{1}{8}n^2H^2,$$

then M is diffeomorphic to S^n.

We show in Example 3.4 that the pinching condition in Theorem 1.3 is sharp. Our method in the proof of the main theorem seems very useful in the study of curvature and topology.

2. Notation and fundamental tools

Throughout, let M^n (where $n \geq 2$) be an n-dimensional compact submanifold in an $(n+p)$-dimensional Riemannian manifold N^{n+p}. We use the following conventions on the range of indices:

$$1 \leq A, B, C, \ldots \leq n+p, \quad 1 \leq i, j, k, \ldots \leq n, \quad n+1 \leq \alpha, \beta, \gamma, \ldots \leq n+p.$$

For an arbitrary fixed point $x \in M \subset N$, we choose an orthonormal local frame field $\{e_A\}$ in N^{n+p} such that e_i's are tangent to M. Let $\{\omega_A\}$ be the dual frame field of $\{e_A\}$. Let Rm, h and ξ denote the Riemannian curvature tensor, second fundamental form and mean curvature vector of M, respectively, and let $\overline{Rm}$ be the Riemannian curvature tensor of N. Then

$$Rm = \sum_{i,j,k,l} R_{ijkl}\omega_i \otimes \omega_j \otimes \omega_k \otimes \omega_l,$$

$$\overline{Rm} = \sum_{A,B,C,D} \bar{R}_{ABCD}\omega_A \otimes \omega_B \otimes \omega_C \otimes \omega_D,$$

$$h = \sum_{\alpha,i,j} h^{\alpha}_{ij}\omega_i \otimes \omega_j \otimes e_\alpha, \quad \xi = \frac{1}{n}\sum_{\alpha,i} h^{\alpha}_{ii} e_\alpha,$$

$$R_{ijkl} = \bar{R}_{ijkl} + \sum_{\alpha} \left(h^{\alpha}_{ik}h^{\alpha}_{jl} - h^{\alpha}_{il}h^{\alpha}_{jk}\right), \tag{2-1}$$

and the mean curvature H of M is given by $H := \left|(1/n)\sum_{\alpha,i} h^{\alpha}_{ii} e_\alpha\right|$. Denote by $\mathrm{Ric}(u)$ the Ricci curvature of M in the direction of $u \in UM$. From the Gauss equation (2-1), we have

$$\mathrm{Ric}(e_i) = \sum_j \bar{R}_{ijij} + \sum_{\alpha,j} \left(h^{\alpha}_{ii}h^{\alpha}_{jj} - h^{\alpha}_{ij}h^{\alpha}_{ij}\right).$$

Denote by A_α the shape operator of M with respect to e_α. Choose $\{e_\alpha\}$ such that e_{n+1} is parallel to ξ and

$$\operatorname{tr} A_{n+1} = nH, \quad \operatorname{tr} A_\alpha = 0, \quad \alpha \neq n+1. \tag{2-2}$$

Thus the mean curvature vector ξ is equal to He_{n+1}, and

$$\sum_j h^{\alpha}_{jj} = \begin{cases} nH, & \alpha = n+1, \\ 0, & \alpha \neq n+1. \end{cases}$$

Hence

$$\mathrm{Ric}(e_i) = \sum_j \bar{R}_{ijij} + nHh_{ii}^{n+1} - \sum_{\alpha,j} h_{ij}^{\alpha}h_{ij}^{\alpha}. \tag{2-3}$$

Denote by $K(x, \pi)$ the sectional curvature of M for tangent 2-plane $\pi(\subset T_x M)$ at point $x \in M$, and by $\bar{K}(x, \pi)$ the sectional curvature of N for tangent 2-plane $\pi(\subset T_x N)$ at point $x \in N$. Set

$$\bar{K}_{\max}(x) := \max_{\pi \subset T_x N} \bar{K}(x, \pi) \quad \text{and} \quad \bar{K}_{\min}(x) := \min_{\pi \subset T_x N} \bar{K}(x, \pi).$$

By Berger's inequality [Brendle 2010, Proposition 1.9],

$$|\bar{R}_{ABCD}| \leq \tfrac{2}{3}(\bar{K}_{\max} - \bar{K}_{\min}) \tag{2-4}$$

for all distinct indices A, B, C, D.

Theorem 2.1 [Hamilton 1982]. *Let (M, g_0) be a compact Riemannian manifold of dimension* 3. *If $Ric_M > 0$, then the normalized Ricci flow with initial metric g_0*

$$\frac{\partial}{\partial t} g(t) = -2\,\mathrm{Ric}_{g(t)} + \tfrac{2}{3} r_{g(t)} g(t)$$

exists for all time and converges to a constant curvature metric as $t \to \infty$. Here $r_{g(t)}$ denotes the mean value of the scalar curvature of $g(t)$.

Next we quote an important convergence result for the Ricci flow in higher dimensions.

Theorem 2.2 [Brendle 2008; Brendle and Schoen 2009]. *Suppose (M, g_0) is a compact Riemannian manifold of dimension $n \geq 4$. Assume that*

$$R_{1313} + \lambda^2 R_{1414} + R_{2323} + \lambda^2 R_{2424} - 2\lambda R_{1234} > 0 \tag{2-5}$$

for all orthonormal four-frames $\{e_1, e_2, e_3, e_4\}$ and all $\lambda \in [-1, 1]$. Then the normalized Ricci flow with initial metric g_0

$$\frac{\partial}{\partial t} g(t) = -2\,\mathrm{Ric}_{g(t)} + \frac{2}{n} r_{g(t)} g(t)$$

exists for all time and converges to a constant curvature metric as $t \to \infty$. Here $r_{g(t)}$ denotes the mean value of the scalar curvature of $g(t)$.

The following nonexistence theorem for stable currents in a compact Riemannian manifold M isometrically immersed into $F^{n+p}(c)$ was proved for $c > 0$ in [Lawson and Simons 1973] and was extended to $c = 0$ in [Xin 1984]. It is used to eliminate the homology groups $H_q(M; \mathbb{Z})$ for $1 \leq q \leq n-1$ and the fundamental group $\pi_1(M)$.

Theorem 2.3. *Let M be an n-dimensional compact submanifold in $F^{n+p}(c)$ with $c \geq 0$. Assume that*

$$\sum_{k=q+1}^{n} \sum_{i=1}^{q} \left[2|h(e_i, e_k)|^2 - \langle h(e_i, e_i), h(e_k, e_k)\rangle\right] < q(n-q)c \tag{2-6}$$

for any orthonormal basis $\{e_i\}$ of T_xM at any point $x \in M$, where q is an integer satisfying $1 \leq q \leq n-1$. Then there do not exist any stable q-currents. Also,

$$H_q(M; \mathbb{Z}) = H_{n-q}(M; \mathbb{Z}) = 0,$$

where $H_i(M; \mathbb{Z})$ is the i-th homology group of M with integer coefficients.

3. Proofs of the theorems

To complete the proof of the main theorem (Theorem 1.3), we need to prove the differentiable pinching theorem for submanifolds (Theorem 1.1) and the vanishing theorem for the fundamental group of a submanifold (Theorem 1.2).

Proof of Theorem 1.1. It is sufficient to show that inequality (2-5) in Theorem 2.2 holds for all $\lambda \in \mathbb{R}$. Suppose that $\{e_1, e_2, e_3, e_4\}$ is an orthonormal four-frame and that $\lambda \in \mathbb{R}$. From the Gauss equation, we have

$$\begin{aligned}
&R_{1313} + \lambda^2 R_{1414} + R_{2323} + \lambda^2 R_{2424} - 2\lambda R_{1234} \\
&= \bar{R}_{1313} + \lambda^2 \bar{R}_{1414} + \bar{R}_{2323} + \lambda^2 \bar{R}_{2424} - 2\lambda \bar{R}_{1234} \\
&\quad + \textstyle\sum_\alpha \left(h^\alpha_{11}h^\alpha_{33} - (h^\alpha_{13})^2 + h^\alpha_{22}h^\alpha_{33} - (h^\alpha_{23})^2\right) \\
&\quad + \lambda^2 \textstyle\sum_\alpha \left(h^\alpha_{11}h^\alpha_{44} - (h^\alpha_{14})^2 + h^\alpha_{22}h^\alpha_{44} - (h^\alpha_{24})^2\right) \\
&\quad - 2\lambda \textstyle\sum_\alpha \left(h^\alpha_{13}h^\alpha_{24} - h^\alpha_{14}h^\alpha_{23}\right) \\
&= \bar{R}_{1313} + \lambda^2 \bar{R}_{1414} + \bar{R}_{2323} + \lambda^2 \bar{R}_{2424} - 2\lambda \bar{R}_{1234} \\
&\quad + \textstyle\sum_\alpha \big(-(h^\alpha_{13})^2 - (h^\alpha_{23})^2 - \lambda^2 (h^\alpha_{14})^2 - \lambda^2 (h^\alpha_{24})^2 \\
&\qquad - 2\lambda h^\alpha_{13}h^\alpha_{24} + 2\lambda h^\alpha_{14}h^\alpha_{23} + h^\alpha_{11}h^\alpha_{33} + h^\alpha_{22}h^\alpha_{33} + \lambda^2 h^\alpha_{11}h^\alpha_{44} + \lambda^2 h^\alpha_{22}h^\alpha_{44}\big).
\end{aligned} \tag{3-1}$$

It follows from Berger's inequality that $\bar{R}_{1234} \leq \frac{2}{3}(\bar{K}_{\max} - \bar{K}_{\min})$. This together with (3-1) implies

$$\begin{aligned}
&R_{1313} + \lambda^2 R_{1414} + R_{2323} + \lambda^2 R_{2424} - 2\lambda R_{1234} \\
&\geq 2(1+\lambda^2)\bar{K}_{\min} - \tfrac{4}{3}|\lambda|(\bar{K}_{\max} - \bar{K}_{\min}) \\
&\quad + \textstyle\sum_\alpha \big(-(h^\alpha_{13})^2 - (h^\alpha_{23})^2 - \lambda^2 (h^\alpha_{14})^2 - \lambda^2 (h^\alpha_{24})^2 \\
&\qquad - |\lambda|(h^\alpha_{13})^2 - |\lambda|(h^\alpha_{24})^2 - |\lambda|(h^\alpha_{23})^2 - |\lambda|(h^\alpha_{14})^2 \\
&\qquad - \tfrac{1}{2}(1+\lambda^2)(h^\alpha_{11})^2 - \tfrac{1}{2}(1+\lambda^2)(h^\alpha_{22})^2 - (h^\alpha_{33})^2 - \lambda^2 (h^\alpha_{44})^2 \\
&\qquad + \tfrac{1}{2}(h^\alpha_{11} + h^\alpha_{33})^2 + \tfrac{1}{2}(h^\alpha_{22} + h^\alpha_{33})^2 + \tfrac{1}{2}\lambda^2 (h^\alpha_{11} + h^\alpha_{44})^2 + \tfrac{1}{2}\lambda^2 (h^\alpha_{22} + h^\alpha_{44})^2\big)
\end{aligned} \tag{3-2}$$

$$\geq 2(1+\lambda^2)\overline{K}_{\min} - \tfrac{2}{3}(1+\lambda^2)(\overline{K}_{\max} - \overline{K}_{\min})$$
$$+ \textstyle\sum_\alpha \big(-(h^\alpha_{31})^2 - (h^\alpha_{32})^2 - (h^\alpha_{33})^2 - \lambda^2(h^\alpha_{41})^2 - \lambda^2(h^\alpha_{42})^2 - \lambda^2(h^\alpha_{44})^2$$
$$- \tfrac{1}{2}(1+\lambda^2)(h^\alpha_{11})^2 - |\lambda|(h^\alpha_{13})^2 - |\lambda|(h^\alpha_{14})^2$$
$$- \tfrac{1}{2}(1+\lambda^2)(h^\alpha_{22})^2 - |\lambda|(h^\alpha_{23})^2 - |\lambda|(h^\alpha_{24})^2$$
$$+ \tfrac{1}{2}(h^\alpha_{11}+h^\alpha_{33})^2 + \tfrac{1}{2}(h^\alpha_{22}+h^\alpha_{33})^2 + \tfrac{1}{2}\lambda^2(h^\alpha_{11}+h^\alpha_{44})^2 + \tfrac{1}{2}\lambda^2(h^\alpha_{22}+h^\alpha_{44})^2\big)$$
$$\geq 2(1+\lambda^2)\overline{K}_{\min} - \tfrac{2}{3}(1+\lambda^2)(\overline{K}_{\max} - \overline{K}_{\min})$$
$$+ \textstyle\sum_\alpha \Big(-\sum_j \big((h^\alpha_{3j})^2 + \lambda^2(h^\alpha_{4j})^2 + \tfrac{1}{2}(1+\lambda^2)(h^\alpha_{1j})^2 + \tfrac{1}{2}(1+\lambda^2)(h^\alpha_{2j})^2\big)$$
$$+ \tfrac{1}{2}(h^\alpha_{11}+h^\alpha_{33})^2 + \tfrac{1}{2}(h^\alpha_{22}+h^\alpha_{33})^2 + \tfrac{1}{2}\lambda^2(h^\alpha_{11}+h^\alpha_{44})^2 + \tfrac{1}{2}\lambda^2(h^\alpha_{22}+h^\alpha_{44})^2\Big).$$

Substituting (2-3) into (3-2), we get

$$\begin{aligned}
&R_{1313} + \lambda^2 R_{1414} + R_{2323} + \lambda^2 R_{2424} - 2\lambda R_{1234}\\
&\geq 2(1+\lambda^2)\overline{K}_{\min} - \tfrac{2}{3}(1+\lambda^2)(\overline{K}_{\max} - \overline{K}_{\min})\\
&\quad + \mathrm{Ric}(e_3) - nHh^{n+1}_{33} - \textstyle\sum_j \overline{R}_{3j3j}\\
&\quad + \lambda^2\big(\mathrm{Ric}(e_4) - nHh^{n+1}_{44} - \textstyle\sum_j \overline{R}_{4j4j}\big)\\
&\quad + \tfrac{1}{2}(1+\lambda^2)\big(\mathrm{Ric}(e_1) - nHh^{n+1}_{11} - \textstyle\sum_j \overline{R}_{1j1j}\big)\\
&\quad + \tfrac{1}{2}(1+\lambda^2)\big(\mathrm{Ric}(e_2) - nHh^{n+1}_{22} - \textstyle\sum_j \overline{R}_{2j2j}\big)\\
&\quad + \tfrac{1}{2}(h^{n+1}_{11} + h^{n+1}_{33})^2 + \tfrac{1}{2}(h^{n+1}_{22} + h^{n+1}_{33})^2\\
&\quad + \tfrac{1}{2}\lambda^2(h^{n+1}_{11} + h^{n+1}_{44})^2 + \tfrac{1}{2}\lambda^2(h^{n+1}_{22} + h^{n+1}_{44})^2\\
&\geq 2(1+\lambda^2)\overline{K}_{\min} - \tfrac{2}{3}(1+\lambda^2)(\overline{K}_{\max} - \overline{K}_{\min}) - 2(1+\lambda^2)(n-1)\overline{K}_{\max}\\
&\quad + \mathrm{Ric}(e_3) + \lambda^2\,\mathrm{Ric}(e_4) + \tfrac{1}{2}(1+\lambda^2)\,\mathrm{Ric}(e_1) + \tfrac{1}{2}(1+\lambda^2)\,\mathrm{Ric}(e_2)\\
&\quad + \tfrac{1}{2}(h^{n+1}_{11} + h^{n+1}_{33})^2 - \tfrac{1}{2}nH(h^{n+1}_{11} + h^{n+1}_{33})\\
&\quad + \tfrac{1}{2}(h^{n+1}_{22} + h^{n+1}_{33})^2 - \tfrac{1}{2}nH(h^{n+1}_{22} + h^{n+1}_{33})\\
&\quad + \tfrac{1}{2}\lambda^2(h^{n+1}_{11} + h^{n+1}_{44})^2 - \tfrac{1}{2}nH\lambda^2(h^{n+1}_{11} + h^{n+1}_{44})\\
&\quad + \tfrac{1}{2}\lambda^2(h^{n+1}_{22} + h^{n+1}_{44})^2 - \tfrac{1}{2}nH\lambda^2(h^{n+1}_{22} + h^{n+1}_{44})\\
&= 2(1+\lambda^2)\overline{K}_{\min} - \tfrac{2}{3}(1+\lambda^2)(\overline{K}_{\max} - \overline{K}_{\min}) - 2(1+\lambda^2)(n-1)\overline{K}_{\max}\\
&\quad + \mathrm{Ric}(e_3) + \lambda^2\,\mathrm{Ric}(e_4) + \tfrac{1}{2}(1+\lambda^2)\,\mathrm{Ric}(e_1) + \tfrac{1}{2}(1+\lambda^2)\,\mathrm{Ric}(e_2)\\
&\quad + \tfrac{1}{2}\big(h^{n+1}_{11} + h^{n+1}_{33} - \tfrac{1}{2}nH\big)^2 + \tfrac{1}{2}\big(h^{n+1}_{22} + h^{n+1}_{33} - \tfrac{1}{2}nH\big)^2\\
&\quad + \tfrac{1}{2}\lambda^2\big(h^{n+1}_{11} + h^{n+1}_{44} - \tfrac{1}{2}nH\big)^2 + \tfrac{1}{2}\lambda^2\big(h^{n+1}_{22} + h^{n+1}_{44} - \tfrac{1}{2}nH\big)^2\\
&\quad - \tfrac{1}{4}(1+\lambda^2)n^2H^2.
\end{aligned} \tag{3-3}$$

By assumption, we have

$$\mathrm{Ric}(e_i) > \big((n-\tfrac{2}{3})\overline{K}_{\max} - \tfrac{4}{3}\overline{K}_{\min}\big) + \tfrac{1}{8}n^2H^2. \tag{3-4}$$

Substituting (3-4) into (3-3), we have

$$\begin{aligned} &R_{1313} + \lambda^2 R_{1414} + R_{2323} + \lambda^2 R_{2424} - 2\lambda R_{1234} \\ &\quad > 2(1+\lambda^2)\overline{K}_{\min} - \tfrac{2}{3}(1+\lambda^2)(\overline{K}_{\max} - \overline{K}_{\min}) - 2(1+\lambda^2)(n-1)\overline{K}_{\max} \\ &\qquad + 2(1+\lambda^2)\big((n-\tfrac{2}{3})\overline{K}_{\max} - \tfrac{4}{3}\overline{K}_{\min} + \tfrac{1}{8}n^2H^2\big) - (1+\lambda^2)\tfrac{1}{4}n^2H^2 \\ &\quad = 2(1+\lambda^2)\big(-(n-\tfrac{2}{3}) + (n-\tfrac{2}{3})\big)\overline{K}_{\max} \\ &\qquad + 2(1+\lambda^2)(1+\tfrac{1}{3}-\tfrac{4}{3})\overline{K}_{\min} = 0. \end{aligned} \tag{3-5}$$

It follows from Theorem 2.2 that the normalized Ricci flow with initial metric g_0,

$$\frac{\partial}{\partial t}g(t) = -2\,\mathrm{Ric}_{g(t)} + \frac{2}{n}r_{g(t)}g(t),$$

exists for all time and converges to a constant curvature metric as $t \to \infty$. Also, M is diffeomorphic to a spherical space form. In particular, if M is simply connected, then M is diffeomorphic to S^n. This completes the proof of Theorem 1.1. □

Corollary 3.1. *Suppose M is an n-dimensional compact submanifold* ($n \ge 4$) *in an* $(n+p)$*-dimensional pinched Riemannian manifold N^{n+p} whose sectional curvature satisfies $b \le \overline{K}_N \le c$. If the Ricci curvature of M satisfies*

$$\mathrm{Ric}_M > \big(n-\tfrac{2}{3}\big)c - \tfrac{4}{3}b + \tfrac{1}{8}n^2H^2,$$

then M is diffeomorphic to a spherical space form. In particular, *if M is simply connected*, *then M is diffeomorphic to S^n.*

Corollary 3.2. *Suppose M is an n-dimensional compact submanifold* ($n \ge 4$) *in an* $(n+p)$*-dimensional pointwise δ-pinched Riemannian manifold N^{n+p}*, *where* $\delta > 0$. *If the Ricci curvature of M satisfies*

$$\mathrm{Ric}_M > \big(n-\tfrac{2}{3}-\tfrac{4}{3}\delta\big)\overline{K}_{\max} + \tfrac{1}{8}n^2H^2,$$

then M is diffeomorphic to a spherical space form. In particular, *if M is simply connected*, *then M is diffeomorphic to S^n.*

Proof of Theorem 1.2. It follows from the Gauss equation (2-1) that

$$\mathrm{Ric}(e_i) = (n-1)c + \sum_{\alpha,k}\big(h^\alpha_{ii}h^\alpha_{kk} - (h^\alpha_{ik})^2\big). \tag{3-6}$$

This, together with the assumption, implies that

$$\begin{aligned}(3\text{-}7)\quad &\sum_{k=2}^{n}\big(2|h(e_1,e_k)|^2-\langle h(e_1,e_1),h(e_k,e_k)\rangle\big)\\ &=2\sum_{\alpha}\sum_{k=2}^{n}(h_{1k}^{\alpha})^2-\sum_{\alpha}\sum_{k=2}^{n}h_{11}^{\alpha}h_{kk}^{\alpha}\\ &=-2\operatorname{Ric}(e_1)+2(n-1)c-\sum_{\alpha}(h_{11}^{\alpha})^2+nHh_{11}^{n+1}\\ &=-2\operatorname{Ric}(e_1)+2(n-1)c-\sum_{\alpha\neq n+1}(h_{11}^{\alpha})^2-\big(h_{11}^{n+1}-\tfrac{1}{2}nH\big)^2+\tfrac{1}{4}n^2H^2\\ &\le -2\operatorname{Ric}(e_1)+2(n-1)c+\tfrac{1}{4}n^2H^2<(n-1)c,\end{aligned}$$

for any orthonormal basis $\{e_i\}$ of T_xM at any point $x\in M$.

Suppose that $\pi_1(M)\neq 0$. Because M is compact, it follows from a classical theorem due to Cartan and Hadamard that there exists a minimal closed geodesic in any nontrivial homotopy class in $\pi_1(M)$. However, by Theorem 2.3, we know that there do not exist any stable integral 1-currents on M. This contradicts the hypothesis. Therefore, $\pi_1(M)=0$. This proves Theorem 1.2. □

Applying Theorem 1.2 to Theorem C, we get a refined version of Ejiri's rigidity theorem.

Theorem 3.3. *Suppose M is an n-dimensional compact orientable minimal submanifold in an $(n+p)$-dimensional unit sphere S^{n+p}, where $n\ge 4$. If the Ricci curvature of M satisfies*

$$\operatorname{Ric}_M\ge n-2,$$

then M is either

- *the totally geodesic submanifold S^n, or*
- *the Clifford torus $S^m(\sqrt{1/2})\times S^m(\sqrt{1/2})$ in S^{n+1} with $n=2m$, or*
- *$\mathbb{C}P^2(\frac{4}{3})$, the 2-dimensional complex projective space minimally immersed in S^7 with constant holomorphic sectional curvature $\frac{4}{3}$.*

Proof of Theorem 1.3. When $n=3$, we have

$$(3\text{-}8)\qquad \operatorname{Ric}_M>c+\tfrac{9}{8}H^2\ge 0.$$

From Hamilton's convergence theorem [1982] for Ricci flow in dimension 3, it follows that M is diffeomorphic to a 3-dimensional spherical space form. When $n\ge 4$, it follows from Theorem 1.1 that M is diffeomorphic to a spherical space form.

On the other hand, it follows from the assumption that

$$(3\text{-}9)\qquad \operatorname{Ric}_M>(n-2)c+\tfrac{1}{8}n^2H^2\ge \tfrac{1}{2}(n-1)c+\tfrac{1}{8}n^2H^2\quad\text{for } n\ge 3.$$

This together with Theorem 1.2 implies that M is simply connected. Therefore, M is diffeomorphic to S^n. This completes the proof of Theorem 1.3. □

The following example shows that the pinching condition in Theorem 1.3 is sharp.

Example 3.4. (i) Let $M = \mathbb{C}P^2\big(\frac{4}{3}(c+H^2)\big)$ be a 2-dimensional complex projective space minimally immersed into $S^7\big(1/\sqrt{c+H^2}\,\big)$ with constant holomorphic sectional curvature $\frac{4}{3}(c+H^2)$, and let $S^7\big(1/\sqrt{c+H^2}\,\big)$ be the totally umbilical sphere in $F^{4+p}(c)$. Here H is a nonnegative constant and $c+H^2>0$. Then M is a compact submanifold in $F^{4+p}(c)$ with constant mean curvature H and constant Ricci curvature $\mathrm{Ric}_M \equiv 2c+2H^2$. M is not a topological sphere.

(ii) Let $M = S^2\big(1/\sqrt{2(c+H^2)}\,\big) \times S^2\big(1/\sqrt{2(c+H^2)}\,\big)$ be a minimal Clifford hypersurface in $S^5\big(1/\sqrt{c+H^2}\,\big)$, and let $S^5\big(1/\sqrt{c+H^2}\,\big)$ be the totally umbilical sphere in $F^{4+p}(c)$. Here H is a nonnegative constant and $c+H^2>0$. Then M is a compact submanifold in $F^{4+p}(c)$ with constant mean curvature H and constant Ricci curvature $\mathrm{Ric}_M \equiv 2c+2H^2$ that is not a topological sphere.

(iii) Let $M = S^m\big(1/\sqrt{2c}\,\big) \times S^m\big(1/\sqrt{2c}\,\big)$ be a minimal Clifford hypersurface in $F^{n+1}(c)$ with $c>0$ and $n=2m\ge 6$, and let $F^{n+1}(c)$ be the totally geodesic submanifold in $F^{n+p}(c)$. Then M is a compact minimal submanifold in $F^{n+p}(c)$ with $\mathrm{Ric}_M \equiv (n-2)c$, and is not homeomorphic to S^n.

Remark 3.5. Using the nonexistence theorem for stable currents on submanifolds in hyperbolic spaces [Fu and Xu 2008] and Theorem 1.1, one can also extend Theorem 1.3 to the case of compact submanifolds in hyperbolic spaces.

Motivated by Theorem 1.3 and the convergence results for mean curvature flow in higher codimensions [Andrews and Baker 2010; Liu et al. 2011], we would like to propose the following conjecture on mean curvature flow in higher codimensions.

Conjecture 3.6. *Let $M_0 = F_0(M)$ be an n-dimensional compact submanifold in an $(n+p)$-dimensional space form $F^{n+p}(c)$, with $c+H^2>0$. If the Ricci curvature of M_0 satisfies*

$$\mathrm{Ric}_{M_0} > (n-2)c + \tfrac{1}{8}n^2H^2,$$

then there exists for the mean curvature flow

$$\begin{cases} \dfrac{\partial}{\partial t}F(x,t) = n\xi(x,t), & x\in M, \quad t\ge 0,\\ F(\,\cdot\,,0) = F_0(\,\cdot\,) \end{cases} \tag{3-10}$$

a unique smooth solution $F_t(\,\cdot\,)$, and either $F_t(\,\cdot\,)$ converges to a round point in finite time, or $c>0$ and $F_t(\,\cdot\,)$ converges to a totally geodesic sphere as $t\to\infty$.

Acknowledgement

The authors thank Dr. Juan-Ru Gu and Dr. En-Tao Zhao for their helpful discussions. They also thank the referee for valuable suggestions.

References

[An-Min and Jimin 1992] L. An-Min and L. Jimin, "An intrinsic rigidity theorem for minimal submanifolds in a sphere", *Arch. Math.* (*Basel*) **58**:6 (1992), 582–594. MR 93b:53050 Zbl 0767.53042

[Andrews and Baker 2010] B. Andrews and C. Baker, "Mean curvature flow of pinched submanifolds to spheres", *J. Differential Geom.* **85**:3 (2010), 357–395. MR 2012a:53122 Zbl 05842818

[Berger 2000] M. Berger, *Riemannian geometry during the second half of the twentieth century*, University Lecture Series **17**, American Mathematical Society, Providence, RI, 2000. MR 2000h:53002 Zbl 0944.53001

[Brendle 2008] S. Brendle, "A general convergence result for the Ricci flow in higher dimensions", *Duke Math. J.* **145**:3 (2008), 585–601. MR 2010a:53132 Zbl 1161.53052

[Brendle 2010] S. Brendle, *Ricci flow and the sphere theorem*, Graduate Studies in Mathematics **111**, American Mathematical Society, Providence, RI, 2010. MR 2011e:53035 Zbl 1196.53001

[Brendle and Schoen 2008] S. Brendle and R. M. Schoen, "Classification of manifolds with weakly 1/4-pinched curvatures", *Acta Math.* **200**:1 (2008), 1–13. MR 2009h:53087 Zbl 1157.53020

[Brendle and Schoen 2009] S. Brendle and R. Schoen, "Manifolds with 1/4-pinched curvature are space forms", *J. Amer. Math. Soc.* **22**:1 (2009), 287–307. MR 2010a:53045 Zbl 05859406

[Chern et al. 1970] S. S. Chern, M. do Carmo, and S. Kobayashi, "Minimal submanifolds of a sphere with second fundamental form of constant length", pp. 59–75 in *Functional Analysis and Related Fields*, edited by F. E. Browder, Springer, New York, 1970. MR 42 #8424 Zbl 0216.44001

[Ejiri 1979] N. Ejiri, "Compact minimal submanifolds of a sphere with positive Ricci curvature", *J. Math. Soc. Japan* **31**:2 (1979), 251–256. MR 80h:53052 Zbl 0396.53026

[Fu and Xu 2008] H. Fu and H. Xu, "Vanishing and topological sphere theorems for submanifolds in a hyperbolic space", *Internat. J. Math.* **19**:7 (2008), 811–822. MR 2009f:53075 Zbl 1191.49045

[Gu and Xu 2011] J. R. Gu and H. W. Xu, "The sphere theorems for manifolds with positive scalar curvature", preprint, 2011. arXiv 1102.2424v1

[Hamilton 1982] R. S. Hamilton, "Three-manifolds with positive Ricci curvature", *J. Differential Geom.* **17**:2 (1982), 255–306. MR 84a:53050 Zbl 0504.53034

[Huisken 1987] G. Huisken, "Deforming hypersurfaces of the sphere by their mean curvature", *Math. Z.* **195**:2 (1987), 205–219. MR 88d:53058 Zbl 0626.53039

[Lawson 1969] H. B. Lawson, Jr., "Local rigidity theorems for minimal hypersurfaces", *Ann. of Math.* (2) **89** (1969), 187–197. MR 38 #6505 Zbl 0174.24901

[Lawson and Simons 1973] H. B. Lawson, Jr. and J. Simons, "On stable currents and their application to global problems in real and complex geometry", *Ann. of Math.* (2) **98** (1973), 427–450. MR 48 #2881 Zbl 0283.53049

[Liu et al. 2011] K. F. Liu, H. W. Xu, F. Ye, and E. T. Zhao, "The extension and convergence of mean curvature flow in higher codimension", preprint, 2011. arXiv 1104.0971v1

[Petersen and Tao 2009] P. Petersen and T. Tao, "Classification of almost quarter-pinched manifolds", *Proc. Amer. Math. Soc.* **137**:7 (2009), 2437–2440. MR 2010m:53046 Zbl 1168.53020

[Shen 1989] Z. M. Shen, “A sphere theorem for manifolds of positive Ricci curvature”, *Indiana Univ. Math. J.* **38**:1 (1989), 229–233. MR 90h:53044 Zbl 0678.53032

[Shen 1992] Y. B. Shen, “Curvature pinching for three-dimensional minimal submanifolds in a sphere”, *Proc. Amer. Math. Soc.* **115**:3 (1992), 791–795. MR 92i:53058 Zbl 0760.53038

[Shiohama 2000] K. Shiohama, “Sphere theorems”, pp. 865–903 in *Handbook of differential geometry*, vol. 1, edited by F. J. E. Dillen and L. C. A. Verstraelen, North-Holland, Amsterdam, 2000. MR 2001c:53051 Zbl 0968.53003

[Shiohama and Xu 1997] K. Shiohama and H. Xu, “The topological sphere theorem for complete submanifolds”, *Compositio Math.* **107**:2 (1997), 221–232. MR 98i:53080 Zbl 0905.53038

[Simons 1968] J. Simons, “Minimal varieties in riemannian manifolds”, *Ann. of Math.* (2) **88** (1968), 62–105. MR 38 #1617 Zbl 0181.49702

[Xin 1984] Y. L. Xin, “An application of integral currents to the vanishing theorems”, *Sci. Sinica Ser. A* **27**:3 (1984), 233–241. MR 86b:49060 Zbl 0554.49021

[Xu 2012] H.-W. Xu, “Recent developments in differentiable sphere theorems”, pp. 415–430 in *Fifth International Congress of Chinese Mathematicians*, vol. 1, edited by L. Ji et al., AMS/IP Studies in Advanced Mathematics **51**, Amer. Math. Soc., Providence, RI, 2012.

[Xu and Gu 2010] H.-W. Xu and J.-R. Gu, “An optimal differentiable sphere theorem for complete manifolds”, *Math. Res. Lett.* **17**:6 (2010), 1111–1124. MR 2012a:53054 Zbl 1221.53068

[Xu and Zhao 2009] H.-W. Xu and E.-T. Zhao, “Topological and differentiable sphere theorems for complete submanifolds”, *Comm. Anal. Geom.* **17**:3 (2009), 565–585. MR 2010k:53067 Zbl 1185.53033

[Yau 1974] S. T. Yau, “Submanifolds with constant mean curvature, I”, *Amer. J. Math.* **96**:2 (1974), 346–366. MR 51 #6670 Zbl 0304.53041

[Yau 1975] S. T. Yau, “Submanifolds with constant mean curvature, II”, *Amer. J. Math.* **97**:1 (1975), 76–100. MR 51 #6670 Zbl 0304.53042

[Yau 1993] S.-T. Yau, “Open problems in geometry”, pp. 1–28 in *Differential geometry: partial differential equations on manifolds* (Los Angeles, 1990), edited by R. Greene and S.-T. Yau, Proc. Sympos. Pure Math. **54**, Amer. Math. Soc., Providence, RI, 1993. MR 94k:53001 Zbl 0801.53001

Received April 21, 2011. Revised July 29, 2011.

HONG-WEI XU
CENTER OF MATHEMATICAL SCIENCES
ZHEJIANG UNIVERSITY
HANGZHOU, 310027
CHINA
xuhw@cms.zju.edu.cn

LING TIAN
CENTER OF MATHEMATICAL SCIENCES
ZHEJIANG UNIVERSITY
HANGZHOU, 310027
CHINA
tianling@cms.zju.edu.cn

CONTENTS

Volume 254, no. 1 and no. 2

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 254 No. 2 December 2011

0030-8730(201112)254:2;1-A