Progress in Mathematics
333

Jingyi Chen
Peng Lu
Zhiqin Lu
Zhou Zhang
Editors

Geometric Analysis

In Honor of Gang Tian's 60th Birthday

Birkhäuser

Birkhäuser

Progress in Mathematics

Volume 333

More information about this series at http://www.springer.com/series/4848

Jingyi Chen • Peng Lu • Zhiqin Lu • Zhou Zhang
Editors

Geometric Analysis

In Honor of Gang Tian's 60th Birthday

Birkhäuser

Editors
Jingyi Chen
Department of Mathematics
The University of British Columbia
Vancouver, BC, Canada

Peng Lu
Department of Mathematics
University of Oregon
Eugene, OR, USA

Zhiqin Lu
Department of Mathematics
University of California, Irvine
Irvine, CA, USA

Zhou Zhang
School of Mathematics and Statistics
University of Sydney
Sydney, NSW, Australia

ISSN 0743-1643 ISSN 2296-505X (electronic)
Progress in Mathematics
ISBN 978-3-030-34952-3 ISBN 978-3-030-34953-0 (eBook)
https://doi.org/10.1007/978-3-030-34953-0

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Preface

This volume of articles is dedicated to Gang Tian on the occasion of his sixtieth birthday.

Born in Nanjing, China, in 1958, Gang Tian received his B.S. from Nanjing University (1982), his MSc from Peking University under the supervision of Kung-Ching Chang (1984), and his Ph.D. from Harvard University under the supervision of Shing-Tung Yau (1988). After holding positions at the State University of New York at Stony Brook, New York University, Massachusetts Institute of Technology (Simons Professor of Mathematics) and Princeton University (Eugene Higgins Professor of Mathematics), Tian retired from Princeton University in 2017 and returned to Peking University as a full-time chair professor. More than 50 people received their Ph.D. under his supervision and a large number of postdocs benefit from his mentoring. Tian has been a major contributor to the development of modern mathematics in China. He has held a professorship at Peking University since 1991, and currently is a vice president of the university. Tian has been the director of the Beijing International Center for Mathematical Research (BICMR), since its founding in 2005. He is also a member of the Scientific Council of the Abdus Salam International Centre for Theoretical Physics in Italy.

Among many honors, Tian was awarded a research fellowship from the Alfred P. Sloan Foundation (1991–1993), the Alan T. Waterman Award from the National Science Foundation in 1994, and the Veblen Prize from the American Mathematical Society in 1996. He was an invited speaker at the ICM in Kyoto (1990) and a plenary speaker at the ICM in Beijing (2002). He was elected into the Chinese Academy of Sciences in 2001 and into the American Academy of Arts and Sciences in 2004.

As a world leader in geometric analysis, complex geometry and symplectic geometry, Tian made fundamental contributions to these fields.

In complex geometry, Tian proved the convergence of the Bergman metrics in his thesis. He settled the existence question for Kähler–Einstein metrics on compact complex surfaces with positive first Chern class. In what is now known as the Bogomolov–Tian–Todorov theorem, he proved the smoothness of the moduli space of compact Calabi–Yau manifolds and gave a new formulation of its Weil–Petersson metric. Jointly with Shing-Tung Yau, Tian proved the important existence result of Kähler–Einstein metrics on complete open Kähler manifolds with first Chern class equal to zero (Calabi's conjecture for the noncompact case). He introduced the notion of K-stability, which has become a very important and active topic in algebraic geometry. He gave a complete proof of the Yau–Tian–Donaldson conjecture on Fano manifolds, which relates the K-stability to the existence of Kähler–Einstein metrics. In Kähler geometry, the Cheeger–Colding–Tian theory plays a very important role. He initiated the Analytic Minimal Model program in the study of the Kähler–Ricci flow, now known as Tian–Song program in (birational) complex geometry.

In symplectic geometry, together with Yongbin Ruan, Tian established a mathematical theory for quantum cohomology and gave the first proof of the associativity of the quantum cohomology ring on semi-positive symplectic manifolds. Together with Jun Li, Tian obtained deep results on Gromov–Witten invariants via the construction of virtual fundamental cycles for algebraic varieties and symplectic manifolds. Jointly with Jeff Streets, Tian introduced and studied the symplectic curvature flow, among other geometric flows.

Tian developed a compactness theory of Yang–Mills fields in high dimension related to calibrated geometry.

Tian and John Morgan, among others, gave a detailed proof for Perelman's solution of the Poincaré Conjecture and the Thurston Geometrization Conjecture.

Many authors of this volume lectured on topics closely related to their articles at conferences held in Fukuoka (March 2017), Beijing (May 2017), Sydney (January 2018), and Vancouver (July 2017). These wonderful events attracted a large number of researchers from all over the world, including many of Tian's mentees, colleagues, collaborators and friends. We would like to thank the Pacific Institute for Mathematical Sciences and The University of Sydney for their warm hospitality and generous supports. Our thanks go also to the National Science Foundation.

Finally we would like to thank the authors for their invaluable contributions and to the referees for their dedicated work. Also we would like to give special thanks to Responsible Editor Thomas Hempfling, Associate Editor Sarah Goob, Editorial Assistant Sabrina Hoecklin, and the editor board of Progress in Mathematics series at Birkhäuser for their help and support.

June 19, 2019 *Jingyi Chen, Peng Lu, Zhiqin Lu, and Zhou Zhang*

A Brief Description of the Volume

The present volume covers a vast range of topics and results of a continuously expanding field belonging to differential geometry and partial differential equations, called geometric analysis, with demonstrated capability of uncovering deep and important relations between geometry and topology.

As evidenced by the names of the authors of the twenty four articles, this volume consists of both original and survey papers[1] penned by leading experts in their corresponding fields. In particular, the several long survey articles provide both a way for beginners to ease into the corresponding sub-fields and the up-to-date status in the sub-fields, e.g., [LiLiu], [Rubin], [Riviè][2], while the original articles give the readers a glimpse of the current research in geometric analysis and related PDEs.

To facilitate the readers' quick access to the articles related to the research themes which they are interested in, we spell out the subject theme(s) for each article. Here we just list a few field names and assign each article roughly to one or more fields in the next four paragraphs.

Complex differential geometry has deep connections with complex analysis, algebraic geometry, and mathematical physics. Over the last 50 years or so, Kähler geometry has made spectacular advances. The existence of Kähler–Einstein metrics has been a central problem and is one of the driving force of the progresses. The research topics have also been extended to extremal Kähler metrics, Kähler–Einstein metrics on manifolds with certain singularities, and other "canonical" metrics arising from Kähler geometry. The papers in this thematic field include [Arezz], [Futak], [HanVi], [LiLiu], [PaulS], [Rubin], and [Zhu]. In the foreseeable future complex differential geometry will continue to be one of the most dynamic research fields.

The 2002 breakthrough by G. Perelman brought new vigor into the flow method employed in searching for "canonical" geometric-topological objects. New flows, such as the Chern flow, Hermitian curvature flow, pluriclosed flow, and Sasaki–Ricci Flow, were introduced, and related interesting questions are raised. The parabolic flow approach to studying the classical and fundamental minimal

[1] All the papers were refereed.

[2] Here we use the first author's last name to identify each article in the volume; in the case of the same last name, we append the first few letters from the second author's last name.

model program in algebraic geometry, is proposed; it significantly extends the earlier treatement of the Kähler–Ricci flow for the Fano case. The papers in the thematic field of geometric flow include [Klein], [Sesum], [Song], and [Stree].

Historically, many interesting differential equations and key results have a geometric origin. A famous example is provided by the minimal surface equation and Bernstein's theorem for hypersurfaces in Euclidean spaces. Another classic example is the Monge–Ampère equation and the optimal transport. In this volume, we include papers on conformal geometry, gauge theory, degenerate elliptic equations, the fully nonlinear elliptic equation from calibrated geometry, and generalized complex Monge–Ampère type equations. The papers in the thematic field of geometric elliptic pdes with applications include [Chang], [Coldi], [Kołod], [LiNgu], [MaQin], [Riviè], [Trudi], and [Yuan].

A common feature of the remaining articles in the volume is their strong topological flavor, while each approaches its subject from innovative angles. They are about index theorem, K-theory, Gromov–Witten theory in symplectic topology, and the classical theme of curvature and topology related to the eigenvalues of the Laplace operator. The papers addressing the above topics are: [Ballm], [Dimak], [Ma], [Zhang], and [Zinge].

Many authors of this volume lectured on the subjects of their articles at various conferences held in the honor of Tian's sixtieth birthday. Ii is to be hoped first that many people have already benefited from these articles, since the conferences had a large audience.

Finally, we hope that the readers will enjoy the volume as much as we did.

Progress in Mathematics, Vol. 333, 1–16

Big and Nef Classes, Futaki Invariant and Resolutions of Cubic Threefolds

Claudio Arezzo and Alberto Della Vedova

To Gang Tian for his 60th Birthday!

Abstract. In this note we revisit and extend few classical and recent results on the definition and use of the Futaki invariant in connection with the existence problem for Kähler constant scalar curvature metrics on polarized algebraic manifolds, especially in the case of resolution of singularities. The general inspiration behind this work is no doubt the beautiful paper by Ding and Tian [16] which contains the germs of a huge amount of the successive developments in this fundamental problem, and it is a great pleasure to dedicate this to Professor G. Tian on the occasion of his birthday.

Mathematics Subject Classification (2010). Primary 53C55; Secondary 32Q15.

Keywords. Constant scalar curvature, Kähler metrics, Futaki invariant.

1. Introduction

Let X be a normal projective variety of dimension n, let L be an ample line bundle on X, and let be fixed a $\mathbf{C}^*$ action on X together with a linearization to L, that is a lifting of the given action on X to an action on L which is linear among the fibers. Up to replace L by some sufficiently large positive power L^m (always possible for our purposes), one can suppose with no loss that X is a subvariety of some complex projective space $\mathbf{CP}^d$, the line bundle L is the restriction to X of the hyperplane bundle, and the $\mathbf{C}^*$-action is induced by some one-parameter subgroup of $SL(d+1,\mathbf{C})$ acting linearly on $\mathbf{CP}^d$ and leaving X invariant.

Associated with these data there is a numerical invariant $F(X,L)$, named after Futaki, who introduced it as an obstruction to the existence of Kähler–Einstein metrics on Fano manifolds [20, 21]. Since then it has been widely generalized [13, 16, 32, 18, 10]. A crucial step towards a definition of stability for Fano manifolds was the extension of Futaki invariant to singular varieties. This was done by Ding–Tian, who defined a Futaki invariant for $\mathbf{Q}$-Fano varieties [16, 32]. Later, Donaldson defined a Futaki invariant for polarized varieties in purely algebraic terms [18]. As noticed in [33], the equivalence of all these extensions follows by results of Paul–Tian [26].

Furthermore, the concept of Futaki invariant has been conveniently extended to the case when instead of a polarization – that is an ample line bundle – on X, one is given a line bundle that is just big and nef [4, 5].

This last extension is of particular relevance when looking at the problem of degenerating the Kähler classes of canonical metrics towards the boundary of the Kähler cone, hence looking at possible convergence of such metrics towards singular ones.

In fact the above idea can be reversed in the hope that the existence of a singular cscK metric in a big and nef class would provide a good starting point for some deformation argument to get also smooth ones in the interior of the Kähler cone nearby the singular one. This turned out to be a successful strategy in a number of important situations, such as blow-ups of smooth points [7, 8, 29, 30], blow-ups of smooth submanifolds [31], smoothings of isolated singularities [12, 27] and resolutions of isolated quotient singularities [6, 3, 2].

Besides some general observations of possible intrinsic interest, the situation studied in this note is the following:

- the singular set S of X is finite (so that each point of S is fixed by the $\mathbf{C}^*$-action);
- $\pi : M \to X$ is an equivariant (log) resolution of singularities, i.e., π restricts to a biholomorphism from $M \setminus \pi^{-1}(S)$ to $X \setminus S$, and for all $p \in S$ the exceptional divisor $E_p = \pi^{-1}(p)$ is simple normal crossing;
- given $p \in S$ and a collection of numbers $b_p > 0$, we look at the line bundle $L_r = \pi^* L^r \otimes \mathcal{O}(-\sum_{p\in S} b_p E_p)$, assuming it is ample for r sufficiently large.

Our main results, Theorem 3.3 and Corollary 3.4, provide general formulae relating the Futaki invariant of (X, L), Futaki of (M, L_r), b_p, the behaviour of a potential for the $\mathbf{C}^*$-action at the singular points and intersection numbers of M.

This results extends all known instances where a similar problem has been attacked (blow-ups at smooth points and resolutions of isolated quotient singularities in the above-mentioned works), and provides many families of examples of new K-unstable polarized manifolds, even as resolutions of K-polystable normal varieties.

Two comments are in order:

1. The assumption on the normality of X is not always necessary for our analysis. Yet, being the final motivation the (non-)existence of cscK metrics, we might as well assume it right away, thanks to [24];
2. we just recall the reader that K-instability is indeed an obstruction to the existence of cscK metrics thanks to [19, Theorem 1].

We end this note with the discussion of few explicit examples. Of course we need to go in dimension at least three to find non-quotient isolated singularities. In particular the case of cubic threefolds is discussed in Section 4. Thanks to Allcock [1] and Liu–Xu [25], as recalled in Theorem 4.2, K-polystable cubic threefolds are now classified, and for example among them it appears the zero locus X of

$$F_{\Delta} = x_0 x_1 x_2 + x_3^3 + x_4^3$$

which has three D_4 singularities, and continuous families of automorphisms.

Now consider a resolution $\pi : M \to X$, and let $E_j, j = 0, 1, 2$, be the exceptional divisors. Chosen integers $b_j > 0$, consider the line bundle

$$L_r = \pi^* L^r \otimes \mathcal{O}(-\sum_{j=0}^{2} b_j E_j),$$

and assume it is ample for all r sufficiently large.

By applying our general computation of the Futaki invariant, we will show (see Proposition 4.3) that any polarized resolution (M, L_r) of the cubic threefold $F_{\Delta} = 0$ is K-unstable for r sufficiently large as soon as the intersection numbers $K_M \cdot (b_0 E_0)^2, K_M \cdot (b_1 E_1)^2, K_M \cdot (b_2 E_2)^2$ are not all the same.

The same strategy can be applied for other examples as discussed in Section 4.

2. Futaki invariant

In this section we give an account of the extension of the Futaki invariant to big and nef classes developed in [5, 4].

Recall that a line bundle B on a projective variety X of dimension n is said to be big when it has positive volume, the latter being the limit of $\dim H^0(X, B^k)/k^n$ as $k \to +\infty$. On the other hand, B is said to be nef if, for any irreducible curve $\Sigma \subset X$, the restriction of B to Σ has non-negative degree. By Kleiman's theorem, nefness is the closure of ampleness condition, meaning that B turns out to be nef if and only if for any ample line bundle A there is $k > 0$ such that $B^k \otimes A$ is ample. On a smooth projective manifold, a line bundle is big and nef if and only if its first Chern class lies at the boundary of the Kähler cone and has positive self-intersection.

Definition 2.1. Let X be a normal projective variety endowed with a $\mathbf{C}^*$-action and let B a big and nef line bundle on M. Choose a linearization on B and for all $k \geq 0$ consider the virtual $\mathbf{C}^*$-representation $H_k = \sum_{q \geq 0} (-1)^q H^q(X, B^k)$. Let $\chi(X, B^k) = \dim(H_k)$ be the Euler characteristic of B^k and let $w(X, B^k)$ be the trace of the infinitesimal generator of the representation H_k. For $k \to \infty$ we have an asymptotic expansion

$$\frac{w(X, B^k)}{\chi(X, B^k)} = F_0 k + F_1 + O(k^{-1}), \tag{2.1}$$

and the Futaki invariant $F(X, B)$ of the given $\mathbf{C}^*$-action on X is defined to be the constant term F_1 of expansion above.

A few comments on this definition are in order.

Firstly, note that given X acted on by $\mathbf{C}^*$ and B as in the definition, one can always find a linearization of the action to B [17, Theorem 7.2]. Actually, in order to do this, perhaps one should replace B with a $\mathbf{C}^*$-invariant line bundle B' isomorphic to B. Since this replacement has no effect for our purposes, from now on we implicitly assume that any line bundle on X is endowed with a linearization of the given $\mathbf{C}^*$-action on X. On the other hand, $F(X,B)$ does not depend on the chosen linearization, whereas the representation H_k and the weight $w(X,B^k)$ do depend on it. In fact, one can check that altering the linearization has the effect of adding $\lambda k\chi(X,B^k)$ to the weight $w(X,B^k)$ for some $\lambda \neq 0$, so that F_1 in expansion (2.1) stay unchanged.

Secondly, note that whenever B is ample, B^k has no higher cohomology for k positive and sufficiently large. Therefore H_k is a genuine representation of $\mathbf{C}^*$, and finally one recovers the Donaldson's definition of Futaki invariant [18, Subsection 2.1].

Thirdly, in the general case one has $\lim_{k\to\infty} k^{-n} \dim H^0(X,B^k) > 0$ by definition of bigness, and $\dim H^q(X,B^k) = O(k^{n-q})$ as a consequence of nefness [23, Theorem 1.4.40]. Hence, even in the more general case, in order to compute $F(X,B)$, one has to consider cohomology groups of B up to order $q=1$.

Finally, note that for any fixed $m>0$ replacing k with mk in (2.1) yields the identity

$$F(X,B^m) = F(X,B). \tag{2.2}$$

One advantage of definition above is that it extends the classical Futaki invariant continuously up to points of the boundary of the ample cone having non-zero volume. More specifically, it holds the following

Proposition 2.2. *Let X be a normal projective variety endowed with a $\mathbf{C}^*$-action. For all line bundles B big and nef, and F invariantly effective, as $r\to\infty$ one has*

$$F(X,B^r\otimes F) = F(X,B) + O(1/r). \tag{2.3}$$

Remark 2.3. By invariantly effective line bundle, we mean a line bundle F such that some positive power F^m posses a $\mathbf{C}^*$-invariant non-zero section. For example, any ample line bundle on X is invariantly effective. Another example is the line bundle $\mathcal{O}(-D)$ associated with a $\mathbf{C}^*$-invariant hypersurface $D\subset X$. In particular, the line bundle associated with an exceptional divisor of a blow-up is invariantly effective.

Proof of Proposition 2.2. For ease of notation let $B_r = B^r\otimes F$. Note that by (2.2) one can replace B_r with an arbitrary large power without altering $F(X,B_r)$. Therefore we can assume that there is an invariant section of F, and let $D\subset X$ be its null locus. Multiplication by kth power of the chosen section gives an equivariant sequence of sheaves on X

$$0\to B^{rk}\to B_r^k\to B_r^k\big|_{kD}\to 0 \tag{2.4}$$

which induces a sequence of (virtual) representation of $\mathbf{C}^*$, whence one has

$$\chi(X, B_r^k) = \chi(X, B^{rk}) + k\chi(D, B_r^k\big|_D). \tag{2.5}$$

and

$$w(X, B_r^k) = w(X, B^{rk}) + kw(D, B_r^k\big|_D). \tag{2.6}$$

Note that by bigness and nefness of B and by asymptotic Riemann–Roch theorem there is a polynomial $q(t) = q_0 t^n + \cdots + q_n$ with $q_0 > 0$ such that $\chi(X, B^{rk}) = q(rk)$ [23, Theorems 1.1.24 and 2.2.16]. Similarly, $w(X, B^{rk}) = p(rk)$ for some polynomial $p(t) = p_0 t^{n+1} + \cdots + p_{n+1}$. For the same reasons, since D has dimension $n-1$, the Euler characteristic $\chi(D, B_r^k\big|_D) = \tilde{q}(r,k)$ is a polynomial of the form $\tilde{q}_0(r)k^{n-1} + \cdots + \tilde{q}_{n-1}$ with $\tilde{q}_i(r)$ which are polynomials of degree at most $n - 1 - i$ and $\tilde{q}_0(r) > 0$ for $r > 0$. A similar situation stands for the total weight $w(D, B_r^k\big|_D) = \tilde{p}(r,k)$ with all degrees raised by one. The upshot is that

$$\frac{w(X, B_r^k)}{\chi(X, B_r^k)} = \frac{p(rk) + k\tilde{p}(r,k)}{q(rk) + k\tilde{q}(r,k)}. \tag{2.7}$$

Expanding the polynomials, by definition of Futaki invariant one finds

$$F(X, B_r) = \frac{p_1 + \tilde{p}_1(r)r^{-n}}{q_0 + \tilde{q}_0(r)r^{-n}} - \frac{\left(p_0 + \tilde{p}_0(r)/r^{n+1}\right)\left(q_1 + \tilde{q}_1(r)/r^{n-1}\right)}{\left(q_0 + \tilde{q}_0(r)r^{-n}\right)^2}. \tag{2.8}$$

At this point, note that $F(X, B) = p_1/q_0 - p_0 q_1/q_0^2$. On the other hand, by discussion above we know that $\tilde{p}_i(r)/r^{n+1-i}$ and $\tilde{q}_i(r)/r^{n-i}$ are $O(1/r)$ for large r. Therefore $F(X, B_r) = F(X, B) + O(1/r)$ as $r \to \infty$, which is the thesis. □

Thanks to Definition 2.1, one can equally work on a singular projective variety endowed with an ample line bundle, or on a smooth variety endowed with a big and nef line bundle, as shown by the following

Proposition 2.4. *Let X be a normal variety endowed with a $\mathbf{C}^*$-action and an ample line bundle L. Let $\pi : M \to X$ be an equivariant resolution of singularities. One has*

$$F(M, \pi^* L) = F(X, L).$$

Proof. Note that $\pi^* L$ is big and nef on M, so that the l.h.s of the identity in the statement makes sense. Now observe that there is an equivariant sequence of sheaves on X

$$0 \to \mathcal{O}_X \to \pi_* \mathcal{O}_M \to \eta \to 0 \tag{2.9}$$

where the support of η has co-dimension at least two. Indeed, the support of η is contained in the singular locus of X, and the latter has co-dimension at least two by normality assumption. After twisting by L^k, by the projection formula one then sees that

$$w(M, \pi^* L^k) = w(X, L^k) + O(k^{n-1}), \qquad \chi(M, \pi^* L^k) = \chi(M, L^k) + O(k^{n-2}),$$

whence the thesis follows by the definition of the Futaki invariant. □

Combining Propositions 2.2 and 2.4 one readily gets the following

Corollary 2.5. *In the situation of Proposition* 2.4*, let F be an invariantly effective line bundle on M (cf. Remark* 2.3*). For $r \to \infty$ one has*

$$F(M, \pi^* L^r \otimes F) = F(X, L) + O(1/r).$$

In the next section, we shall make more explicit the error term $O(1/r)$, at least when the singularities of X are not too bad.

3. Resolutions of isolated singularities

In this section we consider the Futaki invariant of adiabatic polarizations (i.e., making small the volume of exceptional divisors) on resolution of isolated singularities.

As above, consider a normal projective variety X of dimension n endowed with a $\mathbf{C}^*$-action, and let L be an ample line bundle on X. In this section we make the additional assumptions that X is $\mathbf{Q}$-Gorenstein with at most isolated singularities [22]. This means that the singular set $S \subset X$ is finite and each $p \in S$ is a fixed point for the $\mathbf{C}^*$-action. Moreover, some tensor power of the canonical bundle of the smooth locus $X \setminus S$ extends to a line bundle on X. Note that this makes the canonical bundle K_X of X a $\mathbf{Q}$-line bundle, meaning that K_X^m is a genuine line bundle for some integer $m > 0$.

Now consider an equivariant (log) resolution of singularities $\pi : M \to X$. By definition, π restricts to a biholomorphism from $M \setminus \pi^{-1}(S)$ to $X \setminus S$, and for all $p \in S$ the exceptional divisor $E_p = \pi^{-1}(p)$ is simple normal crossing.

Given a positive constant b_p for each $p \in S$, assume there is r sufficiently large such that the line bundle

$$L_r = \pi^* L^r \otimes \mathcal{O}(-\textstyle\sum_{p \in S} b_p E_p) \tag{3.1}$$

is ample on M. Moreover, $\pi^* L$ is big and nef, and each line bundle $\mathcal{O}(-E_p)$ is invariantly effective (cf. Remark 2.3) for E_p is invariant. Note that Corollary 2.5 applies, so that for large r it holds

$$F(M, L_r) = F(X, L) + O(1/r). \tag{3.2}$$

In order to make somehow more explicit the error term, consider the virtual representation $H_k = \sum_{q \geq 0} (-1)^q H^q(M, L_r^k)$. Since M is smooth, at least for $t \in \mathbf{R}$ sufficiently small, the character χ_{H_k} of such representation satisfies [11, Theorem 8.2]

$$\chi_{H_k}(e^{it}) = \int_M e^{c_1(L_r^k)} \operatorname{Td}(M), \tag{3.3}$$

where $c_1(L_r^k)$ and $\operatorname{Td}(M)$ are equivariant characteristic classes. To be more specific, consider the unit circle inside $\mathbf{C}^*$ and let $V \in \Gamma(TM)$ be the infinitesimal generator of the induced circle action on M. Moreover, let ω_r be a circle-invariant Kähler form representing the first Chern class of L_r, and let $u_r \in C^\infty(M)$ be a potential for the circle action on M, so that

$$i_V \omega_r = du_r. \tag{3.4}$$

Denoting by Δ_r the Laplace operator of the Kähler metric ω_r, then (3.3) reduces to

$$\chi_{H_k}(e^{it}) = \int_M e^{k(\omega_r+tu_r)}\left(1+\frac{1}{2}(\mathrm{Ric}(\omega_r)-t\Delta_r u_r)+\cdots\right), \tag{3.5}$$

where dots stand for higher-order terms that are irrelevant for our purposes, and the integral of any differential form of degree different form $2n$ is defined to be zero.

In order to determine the Futaki invariant $F(M,L_r)$, we need to consider the asymptotic behavior for large k of the Euler characteristic $\chi(M,L_r^k)$ and the trace $w(M,L_r^k)$ of the infinitesimal generator of the virtual representation H_k. Note that by definition of χ_{H_k} one has $\chi(M,L_r^k)=\chi_{H_k}(1)$ and $w(M,L_r^k)=\left.\frac{d\chi_{H_k}(e^{it})}{dt}\right|_{t=0}$. Therefore formula (3.5) gives $w(M,L_r^k)=a(r)k^{n+1}+b(r)k^n+O(k^{n-1})$, and $\chi(M,L_r^k)=c(r)k^n+d(r)k^{n-1}+O(k^{n-2})$ where

$$\begin{aligned}
a(r) &= \int_M \frac{(\omega_r+u_r)^{n+1}}{(n+1)!} & b(r) &= \int_M \frac{(\omega_r+u_r)^n\wedge(\mathrm{Ric}(\omega_r)-\Delta_r u_r)}{2n!}\\
c(r) &= \int_M \frac{(\omega_r+u_r)^n}{n!} & d(r) &= \int_M \frac{(\omega_r+u_r)^{n-1}\wedge(\mathrm{Ric}(\omega_r)-\Delta_r u_r)}{2(n-1)!}
\end{aligned} \tag{3.6}$$

are polynomial functions of r. Note that $b(r)$ could be simplified a bit by showing that the summand involving $\Delta_r u_r$ vanishes. On the other hand, u_r and $\Delta_r u_r$ do not affect the value of $c(r)$ and $d(r)$. However it will be apparent in a moment that is convenient to keep the integrands expressed as polynomials in ω_r+u_r and $\mathrm{Ric}(\omega_r)-\Delta_r u_r$. Indeed both of these differential forms turn out to be equivariantly closed, meaning that they are circle-invariant and belong to the kernel of the differential operator

$$d_V = d - i_V. \tag{3.7}$$

Note that one has $d_V^2=0$ on the space of circle-invariant differential forms. As a consequence d_V defines a cohomology, which is sometimes called (the Cartan model of) the equivariant cohomology of M with respect to the given circle action. The equivariant characteristic classes appearing in (3.3) belong to this cohomology.

Apart the deep result represented by (3.3), we need just some basic features of equivariant cohomology. In particular, below we repeatedly make use of the following integration by part formula, whose proof is a quite direct application of Stokes' theorem.

Lemma 3.1. *For all circle invariant inhomogeneous differential forms α, β on M one has*

$$\int_M d_V\alpha\wedge\beta = \int_M (\alpha_{odd}-\alpha_{even})\wedge d_V\beta,$$

where $\alpha=\alpha_{even}+\alpha_{odd}$ with obvious meaning.

At this point we come back to our problem of finding an asymptotic expansion for $F(M,L_r)$. By Definition 2.1 of Futaki invariant one readily sees that

$$F(M,L_r) = b(r)/c(r) - a(r)d(r)/c(r)^2. \tag{3.8}$$

Therefore we are lead to express most of coefficients of polynomials in (3.6) in terms of geometric data on X and M. In order to do this we need to introduce more notation.

For any exceptional divisor E_p let $\xi_p \in \Omega^{1,1}(M)$ be a closed form which represents the Poincaré dual and it is positive along E_p. If E_p is smooth, the latter requirement simply means that ξ_p restricts to a Kähler metric on E_p. In general, it means that $\int_\Sigma \gamma^*\xi_p > 0$ for any non-constant holomorphic curve $\gamma : \Sigma \to M$ whose image is contained in E_p.

We can assume that the supports of ξ_p and ξ_q are disjoint whenever $p, q \in S$ are distinct. Even more, we can assume that ξ_p has support contained in a circle-invariant open set W_p and that W_p and W_q are disjoint whenever $p, q \in S$ are distinct. Therefore, perhaps after averaging over the circle, we can also assume that ξ_p is circle-invariant. Moreover, let u_p be a potential for the vector field V with respect to ξ_p, meaning that $i_V\xi_p = du_p$. Note that u_p is defined up to an additive constant, and that it is constant in the complement of the support of ξ_p. Therefore, by fixing the additive constant, we can assume that the support of u_p is contained W_p. Summarizing, for any $p \in S$ there is an equivariantly closed differential form $\xi_p + u_p$ supported inside W_p such that $[\xi_p] \in H^{1,1}(M)$ is Poincaré dual to E_p.

We already observed in the previous section that for our purposes we can assume with no loss that X is an invariant subvariety of some complex projective space $\mathbf{CP}^d$ acted on linearly by some one-parameter subgroup of $SL(d+1, \mathbf{C})$, and L is the restriction of the hyperplane bundle to X. Therefore, if

$$\iota : X \to \mathbf{CP}^d$$

denotes the inclusion, then the composition $\iota \circ \pi$ is a smooth equivariant map form M to $\mathbf{CP}^d$ which pulls back the hyperplane bundle to π^*L.

Thanks to the inclusion ι we can equip X (or more correctly its smooth locus $X \setminus S$) with a Kähler metric ω and a Hamiltonian potential u for the circle action induced by the unit circle of $\mathbf{C}^*$. To see this, let $V_{FS} \in \Gamma(T\mathbf{CP}^d)$ be its infinitesimal generator of such circle action. Moreover, let ω_{FS} be a circle-invariant Fubini–Study metric on $\mathbf{CP}^d$. Now a potential u_{FS} for V_{FS} is a smooth function on $\mathbf{CP}^d$ satisfying $i_{V_{FS}}\omega_{FS} = du_{FS}$. Finally we define the Kähler form ω and the potential u as the restriction to X of ω_{FS} and u_{FS} respectively. We can think of $\omega + u$ as an equivariantly closed differential form on X. Whereas $\omega + u$ is a genuine equivariantly closed differential form on the smooth locus of X, it is delicate to specify what is ω at singular points of X. On the other hand, it is clear that u is a continuous function on X. However, the pull-back $\pi^*(\omega + u)$ is smooth on M since it is nothing but the pull-back of $\omega_{FS} + u_{FS}$ via the composition of π with the inclusion ι of X into $\mathbf{CP}^d$.

At this point, note that we are free to shrinking the set W_p in order to assume that it is contained in $(\iota \circ \pi)^{-1}(B_p)$ for some small ball $B_p \subset \mathbf{CP}^d$ centered at p. As a consequence $\pi^*(\omega + u)$ turns out to be equivariantly exact in W_p since $\omega_{FS} + u_{FS}$ is equivariantly exact in B_p (in fact one can check that $\omega_{FS} + u_{FS} = d_V d^c \log(1 + |z|^2)$

in affine coordinates making diagonal the circle action). More specifically, there is a circle-invariant function ϕ_p on M such that

$$\pi^*(\omega + u) = d_V d^c \phi_p \qquad \text{in } W_p \tag{3.9}$$

Given all of this, we can assume that the Kähler metric ω_r and the potential function u_r satisfy

$$\omega_r + u_r = r\pi^*(\omega + u) + \sum_{p\in S} b_p(\xi_p + u_p). \tag{3.10}$$

Finally we recall a result that will be useful in the following [3, p. 6].

Lemma 3.2. *Any equivariantly closed differential form α on M which is exact on W_p and restricts to the zero form on the exceptional divisor E_p satisfies $\int_M \alpha \wedge (\xi_p + u_p) = 0$.*

Now we are ready to make explicit coefficients of polynomials appearing in (3.8). Starting with $a(r)$, note that our assumption that $\xi_p + u_p$ is supported inside W_p yields

$$a(r) = r^{n+1}\int_{M\setminus\bigcup_p W_p} \frac{\pi^*(\omega+u)^{n+1}}{(n+1)!} + \sum_{p\in S}\int_{W_p} \frac{(r\pi^*(\omega+u) + b_p(\xi_p+u_p))^{n+1}}{(n+1)!}.$$

Moreover, observing that $\pi^*(\omega + u) - u(p)$ restricts to zero on E_p, by (3.9) and Lemmata 3.1, 3.2 equation above reduces to

$$a(r) = a_0 r^{n+1} + r\sum_{p\in S} b_p^n u(p)\int_M \frac{\xi_p^n}{n!} + \sum_{p\in S} b_p^{n+1}\int_M u_p \frac{\xi_p^n}{n!}, \tag{3.11}$$

where $a_0 = \int_X u\,\omega^n/n!$ coincides with the integral on M of the pull-back via $\iota\circ\pi$ of the smooth differential form $u_{FS}\omega_{FS}^n/n!$. Similarly, for $c(r)$ one finds

$$c(r) = c_0 r^n + \sum_{p\in S} b_p^n \int_M \frac{\xi_p^n}{n!}, \tag{3.12}$$

where $c_0 = \int_X \omega^n/n!$ is the volume of the line bundle L on X, or equivalently the volume of π^*L on M.

Now pass to consider $b(r)$. Arguing precisely as above we can write

$$\begin{aligned} b(r) = r^n &\int_{M\setminus\bigcup_p W_p} \frac{\pi^*(\omega+u)^n \wedge \pi^*(\mathrm{Ric}(\omega) - \Delta u)}{2n!} \\ &+ \sum_{p\in S}\int_{W_p} \frac{(r\pi^*(\omega+u) + b_p(\xi_p+u_p))^n \wedge (\mathrm{Ric}(\omega_r) - \Delta_r u_r)}{2n!}, \end{aligned} \tag{3.13}$$

whence, again by summing and subtracting $u(p)$ to $\pi^*(\omega+u)$ and using (3.9) and Lemmata 3.1, 3.2 as before, it follows

$$b(r) = b_0 r^n + r \sum_{p\in S} u(p) b_p^{n-1} \int_M \frac{(\xi_p+u_p)^{n-1}\wedge(\mathrm{Ric}(\omega_r)-\Delta_r u_r)}{2(n-1)!} + \sum_{p\in S} b_p^n \int_M \frac{(\xi_p+u_p)^n\wedge(\mathrm{Ric}(\omega_r)-\Delta_r u_r)}{2n!}, \quad (3.14)$$

where $b_0 = \int_M \pi^*(\omega+u)^n\wedge(\mathrm{Ric}(\omega_r)-\Delta_r u_r)/(2n!)$ does not depend on r. This follows by integration by parts (Lemma 3.1) and the fact that for all $r,s>0$ it holds

$$\mathrm{Ric}(\omega_r)-\Delta_r u_r = \mathrm{Ric}(\omega_s)-\Delta_s u_s - d_V d^c \log(\omega_r^n/\omega_s^n). \quad (3.15)$$

For the same reason, both integrals of formula (3.14) do not depend on r. In fact, the one of the first line reduces to

$$\int_M \frac{(\xi_p+u_p)^{n-1}\wedge(\mathrm{Ric}(\omega_r)-\Delta_r u_r)}{2(n-1)!} = \int_M \frac{\xi_p^{n-1}\wedge \mathrm{Ric}(\omega_r)}{2(n-1)!}.$$

Moreover, focusing on the second line of (3.14), let $I = \int_M (\xi_p+u_p)^n\wedge(\mathrm{Ric}(\omega_r)-\Delta_r u_r)/(2n!)$. In order to find a simpler expression for it, let $B_\varepsilon \subset M$ be the pull-back via $\iota\circ\pi$ of a small ball in $\mathbf{CP}^d$ of radius ε and centered at p. Since $\pi^*\omega$ is a Kähler metric on $W_p\setminus B_\varepsilon$, there one can write

$$\mathrm{Ric}(\omega_r)-\Delta_r u_r = \pi^*(\mathrm{Ric}(\omega)-\Delta u) - d_V d^c \log(\omega_r^n/\pi^*\omega^n).$$

Therefore, being ξ_p+u_p supported in W_p, by Stokes' theorem it follows

$$I = \int_{M\setminus B_\varepsilon} \frac{(\xi_p+u_p)^n\wedge\pi^*(\mathrm{Ric}(\omega)-\Delta u)}{2n!} + \int_{\partial B_\varepsilon} \frac{(\xi_p+u_p)^n\wedge d^c\log(\omega_r^n/\pi^*\omega^n)}{2n!} + \int_{B_\varepsilon} \frac{(\xi_p+u_p)^n\wedge(\mathrm{Ric}(\omega_r)-\Delta_r u_r)}{2n!}.$$

As we already observed after equation (3.15), I does not depend on r. On the other hand, note that $d^c\log(\omega_r^n/\pi^*\omega^n)$ is smooth on ∂B_ε for all r and is $O(1/r)$ for large r. Similarly, $\mathrm{Ric}(\omega_r)-\Delta_r u_r$ is smooth on B_ε. Therefore, passing to the limit $r\to\infty$ in the equation above yields

$$I = \int_M \frac{(\xi_p+u_p)^n\wedge\pi^*(\mathrm{Ric}(\omega)-\Delta u)}{2n!}. \quad (3.16)$$

Note that Δu is a continuous function on X. This can be checked after noting that Δu equals the ratio of the restrictions to X of $nL_{JV_{FS}}\omega_{FS}\wedge\omega_{FS}^{n-1}$ and ω_{FS}^n. On the other hand, note that $\pi^*(\mathrm{Ric}(\omega)-\Delta u)$ represents the first Chern class of the line bundle $\pi^* K_X^{-1}$. At this point consider the shifted form $\alpha = \pi^*(\mathrm{Ric}(\omega)-\Delta u)+\Delta u(p)$ so that one can rewrite

$$I = -\Delta u(p)\int_M \frac{\xi_p^n}{2n!} + \int_M \frac{(\xi_p+u_p)^n\wedge\alpha}{2n!}. \quad (3.17)$$

Since α vanishes on E_p, by Lemma 3.2 it follows that I reduces to the first summand of the equation above. As a consequence, (3.14) reduces to

$$b(r) = b_0 r^n + r \sum_{p\in S} u(p) b_p^{n-1} \int_M \frac{\xi_p^{n-1} \wedge \mathrm{Ric}(\omega_r)}{2(n-1)!} - \frac{1}{2} \sum_{p\in S} \Delta u(p) b_p^n \int_M \frac{\xi_p^n}{n!}, \quad (3.18)$$

Finally, a similar and easier argument for $d(r)$ gives the expansion

$$d(r) = d_0 r^{n-1} + \sum_{p\in S} b_p^{n-1} \int_M \frac{\xi_p^{n-1} \wedge \mathrm{Ric}(\omega_r)}{2(n-1)!}, \quad (3.19)$$

where $d_0 = \int_M \pi^*\omega^{n-1} \wedge \mathrm{Ric}(\omega_r)/(2(n-1)!)$ does not depend on r, and by the asymptotic Riemann–Roch theorem it is equal to $K_X \cdot L^{n-1}(2(n-1)!)$

At this point, note that we found a geometric meaning for all coefficients appearing in polynomials

$$\begin{aligned} a(r) &= a_0 r^{n+1} + a_n r + a_{n+1} & \qquad b(r) &= b_0 r^n + b_{n-1} r + b_n \\ c(r) &= c_0 r^n + c_n & d(r) &= d_0 r^{n-1} + d_{n-1}. \end{aligned}$$

By direct calculation starting from (3.8) one finds

$$\begin{aligned} F(M, L_r) = {} & \frac{b_0}{c_0} - \frac{a_0 d_0}{c_0^2} + \left(\frac{b_{n-1}}{c_0} - \frac{a_0 d_{n-1}}{c_0^2} \right) r^{1-n} \\ & + \left(\frac{b_n}{c_0} + \frac{d_0}{c_0} \frac{a_0 c_n - c_0 a_n}{c_0^2} - \frac{c_n}{c_0} \left(\frac{b_0}{c_0} - \frac{a_0 d_0}{c_0^2} \right) \right) r^{-n} + O(r^{-n-1}), \end{aligned} \quad (3.20)$$

as $r \to \infty$. By Proposition 2.2 we can recognize $F(X, L)$ in the leading term. Therefore, substituting coefficients calculated above yields the following

Theorem 3.3. *Let $\pi : M \to X$ be an equivariant log resolution of a $\mathbf{Q}$-Gorenstein polarized variety (X, L) acted on by $\mathbf{C}^*$. Assume that the singular locus $S \subset X$ is finite and choose a rational constant $b_p > 0$ for all $p \in S$. Assume moreover that $L_r = \pi^* L^r \otimes \mathcal{O}(-\sum_{p\in S} b_p E_p)$ is ample for r sufficiently large. With the notation introduced above, the Futaki invariant of L_r for $r \to \infty$ is given by*

$$\begin{aligned} F(M, L_r) = {} & F(X, L) + r^{1-n} \frac{n}{2} \sum_{p\in S} (u(p) - \underline{u}) b_p^{n-1} \frac{\int_M \xi_p^{n-1} \wedge \mathrm{Ric}(\omega_r)}{\int_X \omega^n} \\ & - \frac{1}{2} r^{-n} \sum_{p\in S} \left(\underline{s}(u(p) - \underline{u}) + \Delta u(p) + 2F(X, L) \right) b_p^n \frac{\int_M \xi_p^n}{\int_X \omega^n} + O(r^{-n-1}), \end{aligned} \quad (3.21)$$

where $\underline{s} = \frac{n}{2} \int_M \pi^*\omega^{n-1} \wedge \mathrm{Ric}(\omega_r) / \int_X \omega^n$ does not depend on r.

This result should be considered as an extension of a similar result for isolated quotient singularities [3, Theorem 2.3]. Some differences with the formula appearing there are due to a different normalization in the definition of the Futaki invariant.

On the other hand, note that at least the first error term in (3.21) can be expressed almost entirely in terms of intersection numbers on M. Therefore we have the following

Corollary 3.4. *In the situation above, as $r \to \infty$ one has*

$$F(M, L_r) = F(X, L) - r^{1-n}\frac{n}{2L^n}\sum_{p\in S}(u(p) - \underline{u})K_M \cdot (b_p E_p)^{n-1} + O(r^{-n}).$$

This result will be useful in order to produce several examples of K-unstable resolutions in the next section.

4. Resolutions of semi-stable cubic threefolds

In this section we show that most resolutions of semi-stable but not stable cubic threefolds are K-unstable. Here we do not need to recall the full definition of K-stability. Instead it is enough to recall that it is a GIT stability notion for polarized varieties (when no polarization is specified, it is assumed to be the anti-canonical bundle), and that the Hilbert–Mumford criterion for K-stability implies the following elementary

Fact 4.1. A polarized variety is K-unstable as soon as it carries a $\mathbf{C}^*$-action with non-zero Futaki invariant.

To begin with observe that by results of Allcock [1] and Liu–Xu [25] we have the following clear picture of K-stability of cubic threefolds.

Theorem 4.2. *Let $X \subset \mathbf{CP}^4$ be a cubic threefold.*

- *X is K-stable if and only if it is smooth or it has isolated singularities of type A_k with $k \leq 4$.*
- *X is K-polystable with non-discrete automorphism group if and only if it is projectively equivalent to the zero locus of one of the following cubic polynomials:*

$$F_\Delta = x_0x_1x_2 + x_3^3 + x_4^3, \qquad F_{A,B} = Ax_2^3 + x_0x_3^2 + x_1^2x_4 - x_0x_2x_4 + Bx_1x_2x_3,$$

with A and B which are not both zero.

Resolutions of K-stable cubic threefolds have no non-trivial holomorphic vector fields. Therefore, in order to study K-instability of their resolutions one should consider test configuration along the lines of [28, 14, 15]. On the other hand, studying K-instability of resolutions of (strictly) K-polystable cubic threefolds is more direct thanks to Corollary 3.4. In view of this application, observe that any K-polystable cubic threefold $X \subset \mathbf{CP}^4$ is $\mathbf{Q}$-Gorenstein, in that the anti-canonical bundle of the smooth locus extends to K_X^{-1}. Moreover, the latter is (very) ample and the restriction L of hyperplane bundle to X satisfies $L^2 = K_X^{-1}$. We consider separately the cases F_Δ and $F_{A,B}$.

4.1. F_Δ

Let $X \subset \mathbf{CP}^4$ be the zero locus of $F_\Delta = x_0x_1x_2 + x_3^3 + x_4^3$. As one can readily check, the singular locus of X is constituted by three coordinate points

$$S = \{p_0 = (1:0:0:0:0), p_1 = (0:1:0:0:0), p_2 = (0:0:1:0:0)\}\,. \quad (4.1)$$

Each of them is a D_4 singularity, since X is locally equivalent to $z_1^2 + z_2^2 + z_3^3 + z_4^3$ around any $p \in S$. Now pick $\alpha_0, \alpha_1, \alpha_2 \in \mathbf{Z}$ such that $\alpha_0+\alpha_1+\alpha_2 = 0$ and consider the diagonal action of $\mathbf{C}^*$ on $\mathbf{CP}^4$ induced by $\mathrm{diag}(t^{\alpha_1}, t^{\alpha_2}, t^{\alpha_3}, 1, 1)$, where $t \in \mathbf{C}^*$. Clearly X is invariant with respect to this action. A potential with respect to the Fubini–Study metric ω_{FS} for the generator of the induced circle action is given by

$$u_{FS} = \frac{\alpha_0|x_0|^2 + \alpha_1|x_1|^2 + \alpha_2|x_2|^2}{|x|^2}.$$

By direct calculation, one can check that the average $\underline{u} = \int_X u_{FS}\omega_{FS}^3 / \int_X \omega_{FS}^3$ is zero. Now consider a resolution $\pi : M \to X$ and let, as in the general case discussed above, E_j be the exceptional divisor over $p_j \in S$. Chosen an integer $b_j > 0$ for each $p_j \in S$, consider the line bundle

$$L_r = \pi^* L^r \otimes \mathcal{O}(-\sum_{j=0}^{2} b_j E_j),$$

and assume it is ample for all r sufficiently large. By Corollary 3.4 we get

$$F(M, L_r) = -\frac{1}{2r^2}\sum_{j=0}^{2} \alpha_j K_M \cdot (b_j E_j)^2 + O(1/r^3),$$

where we used that $F(X, L) = 0$ thanks to K-polystability of X, that $\underline{u} = 0$ as discussed above, and that $L^3 = 3$. As a consequence, as soon as b_j are chosen so that $K_M \cdot (b_0E_0)^2, K_M \cdot (b_1E_1)^2, K_M \cdot (b_2E_2)^2$ are not all the same, one can choose the α_j's so that $F(M, L_r)$ is non-zero for large r. Therefore we proved the following

Proposition 4.3. *With the notation above, any polarized log resolution (M, L_r) of the cubic threefold $F_\Delta = 0$ is K-unstable for r sufficiently large as soon as the intersection numbers $K_M \cdot (b_0E_0)^2, K_M \cdot (b_1E_1)^2, K_M \cdot (b_2E_2)^2$ are not all the same.*

4.2. $F_{A,B}$

Let $X \subset \mathbf{CP}^4$ be the zero locus of $F_{A,B} = Ax_2^3 + x_0x_3^2 + x_1^2x_4 - x_0x_2x_4 + Bx_1x_2x_3$ where at least one of A and B is non-zero. As described by Allcock [1], different choices of the pair A, B give projectively equivalent threefolds if and only if they give the same $\beta = 4A/B^2 \in \mathbf{C} \cup \{\infty\}$. In other words, β is a moduli parameter. The singularities of X depend on β. If $\beta \neq 0, 1$ then X has precisely two singular points of type A_5. If $\beta = 0$ then an additional singular point of type A_1 appears. If $\beta = 1$ then the singular locus of X is a rational curve. We drop the latter case since singularities are non-isolated. On the other hand, remaining cases are quite similar each other. Therefore we consider in some detail the case $\beta = 0$ and we leave the other ones as an exercise for the reader.

Thus, from now on, $X \subset \mathbf{CP}^4$ will be the zero locus of $F_{0,1} = x_0x_3^2 + x_1^2x_4 - x_0x_2x_4 + x_1x_2x_3$. One can directly check that the singular locus of X is constituted by three coordinate points

$$S = \{p_0 = (1:0:0:0:0), p_2 = (0:0:1:0:0), p_4 = (0:0:0:0:1)\}. \quad (4.2)$$

The points p_0, p_4 turn out to be singularities of type A_5, whereas p_2 is an A_1 singularity. Looking for $\mathbf{C}^*$-actions on $\mathbf{CP}^4$ which preserve X, one find that all of them are coverings of the one induced by $\mathrm{diag}(t^{-2}, t^{-1}, 1, t, t^2)$, where $t \in \mathbf{C}^*$. A potential with respect to the Fubini–Study metric ω_{FS} for the generator of the induced circle action is given by

$$u_{FS} = \frac{-2|x_0|^2 - |x_1|^2 + |x_3|^2 + 2|x_4|^2}{|x|^2}.$$

Note that the transformation which maps $(x_0 : \cdots : x_4)$ to $(x_4 : \cdots : x_0)$ is a holomorphic isometry of $\mathbf{CP}^4$ that preserves X and transforms u_{FS} into $-u_{FS}$. As a consequence, the average $\underline{u} = \int_X u_{FS}\omega_{FS}^3 / \int_X \omega_{FS}^3$ is zero. Now let $\pi : M \to X$ be a (log) resolution and let E_j be the exceptional divisor over $p_j \in S$. Choose an integer $b_j > 0$ for each $p_j \in S$, consider the line bundle

$$L_r = \pi^* L^r \otimes \mathcal{O}\Big(-\sum_{j=0}^{2} b_{2j} E_{2j}\Big),$$

and assume it is ample for all r sufficiently large. By Corollary 3.4 we get

$$F(M, L_r) = \frac{1}{r^2} \sum_{j=0}^{2} (1-j) K_M \cdot (b_{2j}E_{2j})^2 + O(1/r^3), \quad (4.3)$$

where we used that $F(X, L) = 0$ thanks to K-polystability of X, that $\underline{u} = 0$ as discussed above, and that $L^3 = 3$. Note that the local resolution chosen for the A_1 singularity p_2 does not affect the stability of (M, L_r). On the other hand, $F(M, L_r)$ is non-zero for all r sufficiently large whenever b_0, b_4 are chosen so that $K_M \cdot (b_0E_0)^2 + K_M \cdot (b_4E_4)^2 \neq 0$.

A minor adjustment of argument above extends the result above for resolutions of the zero locus of $F_{A,B}$ with $B^2 \neq 4A$. Summarizing we have the following

Proposition 4.4. *With the above notation, any polarized* (log) *resolution* (M, L_r) *of the cubic threefold* $F_{A,B} = 0$ *with* $4A \neq B^2$ *is* K*-unstable for* r *sufficiently large as soon as* $K_M \cdot (b_0E_0)^2 + K_M \cdot (b_4E_4)^2 \neq 0$.

Acknowledgment

Both authors at different times and places have benfited from hundreds of conversations with Prof. G. Tian on topics related to the ones studied in this note. It is a great pleasure to dedicate this paper to him, with our best wishes for his birthday! Moreover, we thanks Prof. J. Kollár for pointing out an inaccuracy in a previous version of this note.

References

[1] D. Allcock. *The moduli space of cubic threefolds*. J. Algebraic Geom. **12** (2003), no. 2, 201–223.

[2] C. Arezzo, A. Della Vedova, R. Lena and L. Mazzieri. *On the Kummer construction for cscK metrics*. To appear in the special volume of Bollettino Unione Matematica Italiana, in memory of Paolo De Bartolomeis. Boll Unione Mat Ital (2018). https://doi.org/10.1007/s40574-018-0170-4.

[3] C. Arezzo, A. Della Vedova and L. Mazzieri. *K-stability, Futaki invariants and cscK metrics on orbifold resolutions*. arXiv:1808.08420.

[4] C. Arezzo, A. Della Vedova and G. La Nave. *Singularities and K-semistability*. Int. Math. Res. Not. IMRN 2012, no. 4, 849–869.

[5] C. Arezzo, A. Della Vedova and G. La Nave. *K-destabilizing test configurations with smooth central fiber*. Variational problems in differential geometry, 24–36, London Math. Soc. Lecture Note Ser., 394, Cambridge Univ. Press, Cambridge, 2012.

[6] C. Arezzo, R. Lena and L. Mazzieri. *On the resolution of extremal and constant scalar curvature Kähler orbifold*. Int. Math. Res. Not. IMRN 2016, no. 21, 6415–6452.

[7] C. Arezzo and F. Pacard. *Blowing up and desingularizing constant scalar curvature Kähler manifolds*. Acta Math. **1**96 (2006), no. 2, 179–228.

[8] C. Arezzo and F. Pacard. *Blowing up Kähler manifolds with constant scalar curvature. II*. Ann. of Math. (2) **170** (2009), no. 2, 685–738.

[9] C. Arezzo, F. Pacard, and M. Singer. *Extremal metrics on blowups*. Duke Math. J. **157** (2011), no. 1, 1–51.

[10] S. Bando. *An obstruction for Chern class forms to be harmonic*. Kodai Math. J. **2**9 (2006), no. 3, 337–345.

[11] N. Berline, E. Getzler and M. Vergne. *Heat kernels and Dirac operators*. Corrected reprint of the 1992 original. Grundlehren Text Editions. Springer-Verlag, Berlin, 2004.

[12] O. Biquard and Y. Rollin. *Smoothing singular constant scalar curvature Kähler surfaces and minimal Lagrangians*. Adv. Math. **2**85 (2015), 980–1024.

[13] E. Calabi. *Extremal Kähler metrics, II*. In Differential Geometry and Complex Analysis, eds. I. Chavel and H.M. Farkas, Springer-Verlag (1985), 95–114.

[14] A. Della Vedova. *CM-stability of blow-ups and canonical metric*. arXiv:0810.5584.

[15] A. Della Vedova and F. Zuddas. *Scalar curvature and asymptotic chow stability of projective bundles and blowups*. Trans. Amer. Math. Soc., **3**64, (2012), 6495–6511.

[16] W. Ding and G. Tian. *Kähler–Einstein metrics and the generalized Futaki invariants*. Invent. Math., **110** (1992), 315–335.

[17] I. Dolgachev. *Lectures on invariant theory*. London Mathematical Society Lecture Note Series, 296. Cambridge University Press, Cambridge, 2003.

[18] S.K. Donaldson. *Scalar curvature and stability of toric varieties*. J. Differential Geom. **62** (2002), no. 2, 289–349.

[19] S.K. Donaldson. *Lower bounds on the Calabi functional*. J. Differential Geom. **70** (2005), no. 3, 453–472.

[20] A. Futaki. *An obstruction to the existence of Einstein Kähler metrics*, Invent. Math. **7**3 (1983), 437–443

[21] A. Futaki. *Kähler–Einstein metrics and Integral Invariants.* Springer-LNM 1314, Springer-Verlag (1988).

[22] S. Ishii. *Isolated Q-Gorenstein singularities of dimension three.* Complex analytic singularities, 165–198, Adv. Stud. Pure Math., 8, North-Holland, Amsterdam, 1987.

[23] R. Lazarsfeld. *Positivity in algebraic geometry – I* Ergebnisse der Mathematik und ihrer Grenzgebiete, 48. Springer-Verlag, Berlin, 2004.

[24] C. Li and C. Xu. *Special test configuration and K-stability of Fano varieties.* Ann. of Math. (2) **1**80 (2014), no. 1, 197–232.

[25] Y. Liu and C. Xu. *K-stability of cubic threefolds.* arXiv:1706.01933v2.

[26] S. Paul and G. Tian. *CM Stability and the Generalized Futaki Invariant I.* arXiv:math/0605278.

[27] C. Spotti. *Deformations of nodal Kähler–Einstein del Pezzo surfaces with discrete automorphism groups.* J. Lond. Math. Soc. (2), **8**9 539–558, 2014.

[28] J. Stoppa. *Unstable blowups.* J. Algebraic Geom. **1**9 (2010), no. 1, 1–17.

[29] G. Székelyhidi. *On blowing up extremal Kähler manifolds.* Duke Math. J. **161** (2012), no. 8, 1411–1453.

[30] G. Székelyhidi. *On blowing up extremal Kähler manifolds II.* Invent. Math. **200** (2015), no. 3, 925–977.

[31] R. Seyyedali and G. Székelyhidi. *Extremal metrics on blowups along submanifolds.* arXiv:1610.06865.

[32] G. Tian. *Kähler–Einstein metrics with positive scalar curvature.* Invent. Math. **1**30 (1997), no. 1, 1–37.

[33] G. Tian. *Existence of Einstein metrics on Fano manifolds.* Metric and differential geometry, 119–159, Progr. Math., **2**97, Birkhäuser/Springer, Basel, 2012.

Claudio Arezzo
ICTP Trieste and
Università di Parma, Italy
e-mail: `arezzo@ictp.it`

Alberto Della Vedova
Università di Milano-Bicocca, Italy
e-mail: `alberto.dellavedova@unimib.it`

Progress in Mathematics, Vol. 333, 17–35

Bottom of Spectra and Amenability of Coverings

Werner Ballmann, Henrik Matthiesen and Panagiotis Polymerakis

Dedicated to Gang Tian on the occasion of his sixtieth birthday

Abstract. For a Riemannian covering $\pi\colon M_1 \to M_0$, the bottoms of the spectra of M_0 and M_1 coincide if the covering is amenable. The converse implication does not always hold. Assuming completeness and a lower bound on the Ricci curvature, we obtain a converse under a natural condition on the spectrum of M_0.

Mathematics Subject Classification (2010). 58J50, 35P15, 53C99.

Keywords. Bottom of spectrum, amenable covering.

1. Introduction

We are interested in the behaviour of the bottom of the spectrum of Laplace and Schrödinger operators under coverings. To set the stage, let M be a simply connected and complete Riemannian manifold and $\pi_0\colon M \to M_0$ and $\pi_1\colon M \to M_1$ be Riemannian subcovers of M. Let Γ_0 and Γ_1 be the groups of covering transformations of π_0 and π_1, respectively, and assume that $\Gamma_1 \subseteq \Gamma_0$. Then the resulting Riemannian covering $\pi\colon M_1 \to M_0$ satisfies $\pi \circ \pi_1 = \pi_0$. Under these circumstances, we always have

$$\lambda_0(M_1) \geq \lambda_0(M_0), \tag{1.1}$$

see, e.g., [1, Theorem 1.1] (and Section 2 for notions and notations). Recall also that any local isometry between complete and connected Riemannian manifolds is a Riemannian covering and, therefore, fits into our schema.

A *mean* on a countable set X is a linear functional $\mu\colon L^\infty(X) \to \mathbb{R}$ with $\mu(f) \geq 0$ if $f \geq 0$ and $\mu(1) = 1$, where the 1 on the left stands for the constant function with value 1. An action of a countable group Γ on X is called *amenable* if X admits a Γ-invariant mean. The group Γ is called *amenable* if the right action of Γ on itself is amenable. Finite groups and actions of finite groups are amenable.

Amenability is a way of expressing a kind of smallness in the context of infinite groups. There are numerous characterizations of amenability. The most useful one for our purposes seems to be Følner's criterion (see 4.2).

We say that the covering π is *amenable* if the right action of Γ_0 on $\Gamma_1\backslash\Gamma_0$ is amenable. If π is normal, that is, if Γ_1 is a normal subgroup of Γ_0, then this holds if and only if $\Gamma_1\backslash\Gamma_0$ is an amenable group. If π is amenable, then

$$\lambda_0(M_1) = \lambda_0(M_0), \tag{1.2}$$

see [1, Theorem 1.2]. The problem whether, conversely, equality implies amenability of the covering is quite sophisticated, as Theorems 1.6 and 1.10, Example 1.12, and the examples on pages 104–105 in [3] show. In the case where M_0 is compact and π is the universal covering (that is, $\pi = \pi_0$), amenability has been established by Brooks [2, Theorem 1]. (A proof avoiding geometric measure theory is contained in [11].) Theorem 2 of Brooks in [3] and Théorème 4.3 of Roblin and Tapie in [13] include normal Riemannian coverings of non-compact manifolds, but impose spectral conditions on M_0 and π, which it might be difficult to verify, and restrictions on the topology of M_0. At the expense of requiring a lower bound on the Ricci curvature, we eliminate topological assumptions altogether and replace the spectral assumptions in [3] and [13] by a weaker and natural condition on the bottom $\lambda_{\mathrm{ess}}(M_0)$ of the essential spectrum of M_0.

Theorem 1.3. *Suppose that the Ricci curvature of M is bounded from below and that $\lambda_{\mathrm{ess}}(M_0) > \lambda_0(M_0)$. Then*

$$\lambda_0(M_1) = \lambda_0(M_0)$$

if and only if the covering $\pi\colon M_1 \to M_0$ is amenable.

Theorem 1.3 gives a positive answer to the speculations of Brooks on page 102 of [3]. Theorems 1.6 and 1.10 and Example 1.12 show that the assumption $\lambda_{\mathrm{ess}}(M_0) > \lambda_0(M_0)$ is sensible.

Examples 1.4. 1) If M_0 is compact, then the Ricci curvature of M_0 is bounded and $\lambda_{\mathrm{ess}}(M_0) = \infty > 0 = \lambda_0(M_0)$.

2) If M_0 is non-compact, of finite volume, and with sectional curvature $-b^2 \leq K_M \leq -a^2$, where $b > a > 0$, then $\lambda_0(M_0) = 0$ and $\mathrm{Ric}_M \geq (1-m)b^2$, where m denotes the dimension of M. Moreover,

$$\lambda_{\mathrm{ess}}(M_0) \geq a^2(m-1)^2/4, \tag{1.5}$$

and hence $\lambda_{\mathrm{ess}}(M_0) > \lambda_0(M_0)$. For the convenience of the reader, we will present a short proof of (1.5) at the end of the article.

A hyperbolic manifold M of dimension m is called *geometrically finite* if the action of its covering group Γ on the hyperbolic space H^m admits a fundamental domain $F \subseteq H^m$ which is bounded by finitely many totally geodesic hyperplanes. By the work of Lax and Phillips ([9, p. 281]), $\lambda_{\mathrm{ess}}(M) = (m-1)^2/4$ if M is geometrically finite of infinite volume.

Theorem 1.6. *Let $\pi\colon M_1 \to M_0$ be a Riemannian covering of hyperbolic manifolds of dimension m with corresponding covering groups $\Gamma_1 \subseteq \Gamma_0$ of isometries of H^m. Assume that M_0 is geometrically finite of infinite volume. Then we have:*

1. *If $\lambda_0(M_0) < (m-1)^2/4$, then $\lambda_0(M_1) = \lambda_0(M_0)$ if and only if π is amenable.*
2. *If $\lambda_0(M_0) = (m-1)^2/4$, then $\lambda_0(M_1) = \lambda_0(M_0)$.*

The first assertion of Theorem 1.6 follows immediately from Theorem 1.3 and the identification $\lambda_{\mathrm{ess}}(M_0) = (m-1)^2/4$ by Lax and Phillips quoted above, the second is an incarnation of the general observation stated in Proposition 1.13.2 below, using that $\lambda_0(H^m) = (m-1)^2/4$.

Remarks 1.7. 1) We say that a geometrically finite hyperbolic manifold $M = \Gamma\backslash H^m$ is *convex cocompact* if it does not have cusps or, equivalently, if Γ does not contain parabolic isometries. Theorem 1.6.1 is due to Brooks in the case where the covering is normal and M_0 is convex cocompact. See [3, Theorem 3] and also [13, Théorème 0.2].

2) The *critical exponent* $\delta(\Gamma)$ of a discrete group Γ of isometries of H^m is the infimum of the set of $s \in \mathbb{R}$ such that the Poincaré series

$$g(x,y,s) = \sum_{\gamma\in\Gamma} e^{-sd(x,\gamma(y))}$$

converges for all $x, y \in H^m$. Using Sullivan's [14, Theorem 2.17], the assumptions on $\lambda_0(M_0)$ in Theorem 1.6 may be reformulated in terms of the critical exponent of Γ_0. Namely

$$\lambda_0(M_0) = \delta(\Gamma_0)(m-1-\delta(\Gamma_0)) < (m-1)^2/4 = \lambda_{\mathrm{ess}}(M_0)$$

if $\delta(\Gamma_0) > (m-1)/2$ and $\lambda_0(M_0) = (m-1)^2/4$ if $\delta(\Gamma_0) \le (m-1)/2$.

3) Let M_0 be a geometrically finite hyperbolic manifold of infinite volume and M_0' be a connected Riemannian manifold such that $M_0 \setminus K$ is isometric to $M_0' \setminus K'$ for some compact domains $K \subseteq M_0$ and $K' \subseteq M_0'$. Since the essential spectrum is determined by the geometry at infinity, we then still have $\lambda_{\mathrm{ess}}(M_0') = (m-1)^2/4$, and therefore Theorem 1.3 applies to Riemannian coverings of M_0' if $\lambda_0(M_0') < (m-1)^2/4$. This latter condition is easy to achieve by choosing the metric on K' appropriately.

Let M be the interior of a compact and connected manifold N with non-empty boundary and h be a Riemannian metric on N. Let $\rho \ge 0$ be a smooth non-negative function on N defining ∂N, that is,

$$\partial N = \{\rho = 0\} \text{ and } \partial_\nu \rho > 0 \tag{1.8}$$

along ∂N, where ν denotes the inner normal of N along ∂N with respect to h. Consider the conformally equivalent metric

$$g = \rho^{-2} h \tag{1.9}$$

on M. The metric g is complete since the factor ρ^{-2} causes ∂N to have infinite distance to any point in M. Metrics of this kind were introduced by Mazzeo, who named them *conformally compact*. In [10, Theorem 1.3], he obtains that the essential spectrum of g is $[a^2(m-1)^2/4, \infty)$, where $a = \min \partial_\nu \rho > 0$ and $m = \dim M$. In particular, $\lambda_{\text{ess}}(g) = a^2(m-1)^2/4$.

Theorem 1.10. *Let $\pi\colon M_1 \to M_0$ be a Riemannian covering of manifolds of dimension m with corresponding covering groups $\Gamma_1 \subseteq \Gamma_0$ of isometries of their universal covering space M. Assume that M_0 is conformally compact with $a = \min \partial_\nu \rho$ as above. Then we have:*

1. *If $\lambda_0(M_0) < a^2(m-1)^2/4$, then $\lambda_0(M_1) = \lambda_0(M_0)$ if and only if π is amenable.*
2. *If $\lambda_0(M_0) = a^2(m-1)^2/4$, then $\lambda_0(M_1) = \lambda_0(M_0)$.*

The first assertion of Theorem 1.10 follows immediately from Theorem 1.3 together with Mazzeo's $\lambda_{\text{ess}}(M_0) = a^2(m-1)^2/4$ quoted above, where we note that the sectional curvature of M_0 is bounded from above and below. The second assertion of Theorem 1.10 is proved in Section 4.

Remark 1.11. By changing the metric on a compact part of M_0 appropriately, it is easy to obtain examples which satisfy the first assertion of Theorem 1.10.

Example 1.12 (concerning Theorem 1.3). Let P be a compact and connected manifold of dimension m with connected boundary $\partial P =: N_0$. Assume that the fundamental group of N_0 is amenable; e.g., $N_0 = S^{m-1}$. Let $U \cong [0,\infty) \times N_0$ be a collared neighborhood of $N_0 \cong \{0\} \times N_0$ in P. Let g_0 be a Riemannian metric on $M_0 = P \setminus N_0$, which is equal to $dx^2 + h_0$ along $V_0 = U \setminus N_0 \cong (0,\infty) \times N_0$, where we write elements of V_0 as pairs (x,y) with $x \in (0,\infty)$ and $y \in N_0$ and where h_0 is a Riemannian metric on N_0. Since N_0 is compact, we have $\lambda_0(V_0) = 0$. Since $\lambda_0(M_0) \le \lambda_0(V_0)$, we conclude that $\lambda_0(M_0) = 0$.

The volume of g_0 is infinite, and the sectional curvature of g_0 is bounded.

Let $\pi\colon M_1 \to M_0$ be a Riemannian covering and V_1 be a connected component of $\pi^{-1}(V_0)$. Then $\pi_1\colon V_1 \to V_0$ is a Riemannian covering, and it is amenable since the fundamental group of V_0 is amenable. Therefore $\lambda_0(V_1) = \lambda_0(V_0)$, by [1, Theorem 1.2]. Since $\lambda_0(M_1) \le \lambda_0(V_1) = 0$, we conclude that $\lambda_0(M_1) = 0$. It follows that $\lambda_0(M_1) = \lambda_0(M_0) = 0$, regardless of whether π is amenable or not.

The example is very much in the spirit of the surface S_α (for $0 < \alpha < 1$), discussed on page 104 of [3]. Note that S_α is complete with finite area and bounded curvature.

We see in Theorem 1.6.2 and Theorem 1.10.2 that the essential spectrum can be in the way of the bottom of the spectrum to grow. One aspect of this is revealed in the first of the following two observations.

Proposition 1.13. *In our setup of Riemannian coverings,*

1. *if π is infinite and $\lambda_0(M_1) = \lambda_0(M_0)$, then $\lambda_0(M_1) = \lambda_{\text{ess}}(M_1)$.*
2. *if $\lambda_0(M_0) = \lambda_0(M)$, then $\lambda_0(M_1) = \lambda_0(M_0)$.*

The case in Proposition 1.13.1, where the deck transformation group of π is infinite, is also a consequence of [11, Corollary 1.3]. The proof of Proposition 1.13.2 is trivial: By applying (1.1) to π and π_1, we see that $\lambda_0(M_1)$ is pinched between $\lambda_0(M_0)$ and $\lambda_0(M)$.

The lower bound on the Ricci curvature, required in Theorem 1.3, is used in two instances. First, we need that positive eigenfunctions of the Laplacian satisfy a Harnack inequality. To that end, we employ the Harnack inequality of Cheng and Yau (see (2.23)). Second, in the proof of Lemma 3.1, we use Buser's Lemma 2.16 below. Both, the Harnack inequality of Cheng and Yau and Buser's lemma, require a lower bound on the Ricci curvature.

Remark 1.14. In the first submitted version of this article, we asked the question whether the assumption on the lower Ricci curvature bound in Theorem 1.3 is necessary. Very recently, the third named author used Theorem 1.3 to show that the analog of Theorem 1.3 holds for compact M_0 with boundary with respect to the Neumann boundary condition and concluded from there that Theorem 1.3 actually also holds without assuming the lower Ricci curvature bound [12].

Structure of the article

In Section 2, we collect some preliminaries about Schrödinger operators and the geometry of Riemannian manifolds. The volume estimate in Section 3 is the basis of our discussion of the amenability of coverings. Much of the argumentation in this section follows Buser's [4, Section 4]. In Section 4, we prove a generalized version of Theorem 1.3 for Schrödinger operators, where the potential V and its derivative dV are assumed to be bounded. Furthermore, Section 4 contains the outstanding proofs of (1.5), Theorem 1.10.2, and Proposition 1.13.1.

2. Preliminaries

Let M be a Riemannian manifold of dimension m and $V\colon M \to \mathbb{R}$ be a smooth potential. We denote by Δ the Laplace operator of M and by $S = \Delta + V$ the Schrödinger operator associated to V. We say that a smooth function φ on M (not necessarily square integrable) is a λ-*eigenfunction* if it solves $S\varphi = \lambda\varphi$.

For a point $x \in M$, subset $A \subseteq M$, and radius $r > 0$, we denote by $B(p, r)$ the open geodesic ball of radius r around x and by

$$A^r = \{p \in M \mid d(p, A) < r\} \tag{2.1}$$

the open neighborhood of radius r around A, respectively.

For a Lipschitz function $f \neq 0$ on M with compact support, we call

$$R(f) = \frac{\int_M \|\operatorname{grad} f\|^2 + Vf^2}{\int_M f^2} \tag{2.2}$$

the *Rayleigh quotient* of f and

$$\lambda_0(M, V) = \inf R(f) \tag{2.3}$$

the *bottom of the spectrum of* (M, V). Here the infimum is taken over all Lipschitz functions $f \neq 0$ on M with compact support. In the case of the Laplacian, that is, $V = 0$, we write $\lambda_0(M)$ instead of $\lambda_0(M, 0)$ and call $\lambda_0(M)$ the *bottom of the spectrum of* M. If M is complete and V is bounded from below, then $\lambda_0(M, V)$ is the minimum of the spectrum of S, more precisely, of the closure of S on $C_c^\infty(M)$ in $L^2(M)$. We call

$$\lambda_{\mathrm{ess}}(M, V) = \sup_K \lambda_0(M \setminus K, V), \tag{2.4}$$

where the supremum is taken over all compact subsets K of M, the *bottom of the essential spectrum of* (M, V). In the case of the Laplacian, that is, $V = 0$, we write $\lambda_{\mathrm{ess}}(M)$ instead of $\lambda_{\mathrm{ess}}(M, 0)$ and call $\lambda_{\mathrm{ess}}(M)$ the *bottom of the essential spectrum of* M. If M is complete and V is bounded from below, then $\lambda_{\mathrm{ess}}(M, V)$ is the minimum of the essential spectrum of S.

For a Borel subset $A \subseteq M$, we denote by $|A|$ the volume of A. Similarly, for a submanifold N of M of dimension $n < m$, we let $|N|$ be the n-dimensional Riemannian volume of N. We call

$$h(M) = \inf \frac{|\partial A|}{|A|} \quad \text{and} \quad h_{\mathrm{ess}}(M) = \sup_K h(M \setminus K) \tag{2.5}$$

the *Cheeger constant* and *asymptotic Cheeger constant of* M, respectively. Here the infimum is taken over all compact domains $A \subseteq M$ with smooth boundary ∂A and the supremum over all compact subsets K of M. The respective *Cheeger inequality* asserts that

$$\lambda_0(M) \geq \frac{1}{4} h^2(M) \quad \text{and} \quad \lambda_{\mathrm{ess}}(M) \geq \frac{1}{4} h_{\mathrm{ess}}^2(M). \tag{2.6}$$

The *Buser inequality* is a converse to Cheeger's inequality. In the case where M is non-compact, complete, and connected with $\mathrm{Ric}_M \geq (1 - m)b^2$, where $b \geq 0$, it asserts that

$$\lambda_0(M) \leq C_{1,m} b h(M). \tag{2.7}$$

See [4, Theorem 7.1]. Here and below, indices attached to constants indicate the dependence of the constants on parameters. Thus $C_{1,m}$ indicates that the constant depends on m and that a constant $C_{2,m}$ is to be expected.

For a bounded domain $D \subseteq M$ with smooth boundary, we call

$$h^N(D) = \inf_A \frac{|\partial A \cap \operatorname{int} D|}{|A|} \tag{2.8}$$

the *Cheeger constant of* D *with respect to the Neumann boundary condition*. Here $\operatorname{int} D$ denotes the interior of D, and the infimum is taken over all domains $A \subseteq D$ with smooth intersection $\partial A \cap \operatorname{int} D$ such that $|A| \leq |D|/2$.

2.1. Renormalizing the Schrödinger operator

The idea of renormalizing the Laplacian occurs in [14, Section 8] and [3, Section 2]. The idea also works for Schrödinger operators, as explained in [11, Section 7]. More details about what we discuss here can be found in the latter article.

Let M be a Riemannian manifold and $V\colon \mathbb{R} \to M$ be a smooth potential. Let φ be a positive λ-eigenfunction of $S = \Delta + V$ on M. For a Borel subset $A \subseteq M$, we denote by $|A|_\varphi$ the φ-volume of A,

$$|A|_\varphi = \int_A \varphi^2. \tag{2.9}$$

Similarly, for a submanifold N of M of dimension $n < m$, we let $|N|_\varphi$ be the n-dimensional φ-volume of N.

We renormalize the Schrödinger operator $S = \Delta + V$ of M and consider

$$S_\varphi = m_{1/\varphi}(S - \lambda)m_\varphi \tag{2.10}$$

instead, where m_φ and $m_{1/\varphi}$ denote multiplication by φ and $1/\varphi$ respectively. Now S with domain $C_c^\infty(M)$ is formally and essentially self-adjoint in $L^2(M, dx)$, where dx denotes the Riemannian volume element of M, and S_φ is obtained from $S - \lambda$ by conjugation with $m_{1/\varphi}$. Hence S_φ with domain $C_c^\infty(M)$ is formally and essentially self-adjoint in $L^2(M, \varphi^2 dx)$. By [11, Proposition 7.3], we have

$$\lambda_0(M,V) - \lambda = \inf \frac{\int_M \|\operatorname{grad} f\|^2 \varphi^2}{\int_M f^2 \varphi^2}, \tag{2.11}$$

where the infimum is taken over all non-vanishing smooth functions f on M with compact support. By approximation, it follows easily that we obtain the same infimum by considering non-vanishing Lipschitz functions on M with compact support.

For a bounded domain $A \subseteq M$ with smooth boundary ∂A, we set

$$h_\varphi(M,A) = \frac{|\partial A|_\varphi}{|A|_\varphi}. \tag{2.12}$$

and call

$$h_\varphi(M) = \inf_A h_\varphi(M,A), \quad \text{and} \quad h_{\varphi,\mathrm{ess}}(M) = \sup_K h_\varphi(M \setminus K) \tag{2.13}$$

the *modified Cheeger constant* and *modified asymptotic Cheeger constant of* M, respectively. Here the infimum is taken over all compact domains $A \subseteq M$ with smooth boundary ∂A and the supremum over all compact subsets of M. The Cheeger constants in (2.5) correspond to the case $\varphi = 1$. By [11, Corollaries 7.4 and 7.5], we have the *modified Cheeger inequalities*

$$\lambda_0(M,V) - \lambda \geq h_\varphi(M)^2/4 \quad \text{and} \quad \lambda_{\mathrm{ess}}(M,V) - \lambda \geq h_{\varphi,\mathrm{ess}}(M)^2/4. \tag{2.14}$$

In particular, if $\lambda = \lambda_0(M,V)$, then $h_\varphi(M) = 0$.

2.2. Volume comparison

Let H^m be the hyperbolic space of dimension m and sectional curvature -1, and denote by $\beta_m(r)$ the volume of geodesic balls of radius r in H^m.

Theorem 2.15 (Bishop–Gromov inequality). *Let M be a complete Riemannian manifold of dimension m and $\mathrm{Ric}_M \geq 1-m$, and let x be a point in M. Then*

$$\frac{|B(x,R)|}{|B(x,r)|} \leq \frac{\beta_m(R)}{\beta_m(r)}$$

for all $0 < r < R$. In particular, $|B(x,r)| \leq \beta_m(r)$ for all $r > 0$.

We say that a subset $D \subseteq M$ is *star-shaped with respect to $x \in D$* if, for any $z \in D$ and minimal geodesic $\gamma\colon [0,1] \to M$ from x to z, we have $\gamma(t) \in D$ for all $0 \leq t \leq 1$. Observing that Buser's proof of Lemma 5.1 in [4] does not use the compactness of the ambient manifold M, but only the lower bound for its Ricci curvature, his arguments yield the following estimate.

Lemma 2.16 (Buser). *Let M be a complete Riemannian manifold of dimension m and $\mathrm{Ric}_M \geq 1-m$. Let $D \subseteq M$ be a domain which is star-shaped with respect to $x \in D$. Suppose that $B(x,r) \subseteq D \subseteq B(x,2r)$ for some $r > 0$. Then*

$$h^N(D) \geq C_{m,r} = \frac{1}{r} C_{2,m}^{1+r},$$

where $0 < C_{2,m} < 1$.

2.3. Separated sets

Given $r > 0$, we say that a subset $X \subseteq M$ is *r-separated* if $d(x,y) \geq r$ for all points $x \neq y$ in X. An r-separated subset $X \subseteq M$ is said to be *complete* if $\cup_{x \in X} B(x,r) = M$. Any r-separated subset $X \subseteq M$ is contained in a complete one.

We assume now again that M is complete of dimension m with $\mathrm{Ric}_M \geq 1-m$. For $r > 0$ given, we let $X \subseteq M$ be a complete $2r$-separated subset. For $x \in X$, we call

$$D_x = \{z \in M \mid d(z,x) \leq d(z,y) \text{ for all } y \in X\} \tag{2.17}$$

the *Dirichlet domain* about x. Since X is complete as a $2r$-separated subset of M,

$$B(x,r) \subseteq D_x \subseteq B(x,2r) \tag{2.18}$$

for all $x \in X$. We therefore get from Theorem 2.15 that

$$|D_x| \leq |B(x,2r)| \leq \frac{\beta_m(2r)}{\beta_m(r)} |B(x,r)|. \tag{2.19}$$

Furthermore, for any $x \in X$, $z \in D_x$, and minimal geodesic $\gamma\colon [0,1] \to M$ from x to z, we have the strict inequality $d(\gamma(t),x) < d(\gamma(t),y)$ for all $0 \leq t < 1$ and $y \in X$ different from x. In particular, D_x is star-shaped. Using Lemma 2.16, we conclude that

$$h^N(D_x) \geq C_{m,r} \quad \text{for all } x \in X. \tag{2.20}$$

2.4. Distance functions

Suppose that M is complete and connected. Let $K \subseteq M$ be a closed subset and $r > 0$. Define a function $f = f_{K,r}$ on M by

$$f(x) = \begin{cases} d(x,K) & \text{if } d(x,K) \le r, \\ r & \text{if } d(x,K) \ge r. \end{cases}$$

Then f is a Lipschitz function with Lipschitz constant 1. A theorem of Rademacher says that the set $\mathcal{R}$ of points $x \in M$, such that f is differentiable at x, has full measure in M. Clearly, $\|\operatorname{grad} f(x)\| \le 1$ for all $x \in \mathcal{R}$.

Lemma 2.21. *If x is a point in $\mathcal{R}$ such that* $\operatorname{grad} f(x) \neq 0$, *then x belongs to $K^r \setminus K$,* $\operatorname{grad} f(x)$ *has norm one, and there is a unique minimizing geodesic from x to K. Moreover, ∂K^r is disjoint from $\mathcal{R}$.*

Proof. Let c be a smooth curve through x such that $c'(0) = \operatorname{grad} f(x)$. Then $(f \circ c)(t) < f(x)$ for all $t < 0$ sufficiently close to 0 and $(f \circ c)(t) > f(x)$ for all $t > 0$ sufficiently close to 0. Hence $x \notin K$ since $f \ge 0$ and $x \notin M \setminus K^r$ since $f \le r$. Therefore $x \in K^r \setminus K$, that is, $0 < f(x) = d(x,K) < r$. Let $\gamma\colon [0, f(x)] \to M$ be a minimizing unit speed geodesic from x to K. Then $(f \circ \gamma)(t) = f(x) - t$ for all $0 \le t \le f(x)$, hence

$$\langle \operatorname{grad} f(x), \gamma'(0) \rangle = (f \circ \gamma)'(0) = -1.$$

Since $\|\operatorname{grad} f(x)\| \le 1$ and $\|\gamma'(0)\| = 1$, we get that $\operatorname{grad} f(x) = -\gamma'(0)$ and hence that γ is unique and that $\|\operatorname{grad} f(x)\| = 1$.

For $x \in \partial K^r \cap \mathcal{R}$ and $\gamma\colon [0, f(x)] \to M$ a minimizing unit speed geodesic from x to K, we would have $-1 = (f \circ \gamma)'(0) = \langle \operatorname{grad} f(x), \gamma'(0) \rangle$, hence that $\operatorname{grad} f(x) \neq 0$, contradicting the first part of the lemma. □

By the same reason as in the last part of the above proof, we get that a point on the boundary of K, which is the endpoint of a minimizing geodesic from some point $x \in M \setminus K$ to K, does not belong to $\mathcal{R}$.

2.5. Harnack inequalities

We say that a positive function φ on M satisfies a *Harnack estimate* if there is a constant $C_\varphi \ge 1$ such that

$$\sup_{B(x,r)} \varphi^2 \le C_\varphi^{r+1} \inf_{B(x,r)} \varphi^2 \tag{2.22}$$

for all $x \in M$ and $r > 0$.

Suppose now that M is complete with $\operatorname{Ric}_M \ge (1-m)b^2$, that $|V|$ and $\|\nabla V\|$ are bounded, and that φ is a positive λ-eigenfunction of $S = \Delta + V$ on M. By the estimate of Cheng and Yau [6, Theorem 6], we then have

$$\frac{\|\nabla \varphi(x)\|}{\varphi(x)} \le C_{3,m} \max\{\|V - \lambda\|_\infty / b, \|\nabla V\|_\infty^{1/3}, b\} \tag{2.23}$$

for all $x \in M$ (with $m_1 = m_4 = c = 0$, $m_2 = m_5 = \|V - \lambda\|_\infty$, $m_3 = \|\nabla V\|_\infty$, and $a = \infty$ in loc. cit.). In particular, φ satisfies a Harnack estimate (2.22). Notice that Δ and λ rescale by $1/s$ if the Riemannian metric of M is scaled by $s > 0$. To keep φ as an eigenfunction, V must therefore also be rescaled by $1/s$.

3. Modified Buser inequality

Following Buser's arguments in [4, Section 4], we prove the following estimate.

Lemma 3.1. *Let M be a complete and connected Riemannian manifold with Ricci curvature bounded from below and $\varphi > 0$ be a smooth function on M which satisfies a Harnack inequality. Suppose that $h_\varphi(M) = 0$, and let $\varepsilon, r > 0$ be given. Then there exists a bounded open subset $A \subseteq M$ such that*

$$|A^r \setminus A|_\varphi < \varepsilon |A|_\varphi.$$

Proof. Renormalizing the metric of M if necessary, we assume throughout the proof that $\mathrm{Ric}_M \geq 1 - m$ and let $\beta = \beta_m$ (see Section 2.2), where $m = \dim M$.

Let $\varepsilon, r > 0$ be given. Recall the constants $C_{m,r}$ and C_φ from Lemma 2.16 and (2.22). Let $A \subseteq M$ be a (non-empty) bounded domain with smooth boundary such that

$$\frac{2\beta(4r) C_\varphi^{6r+3}}{\beta(r) C_{m,r}} h_\varphi(M, A) < \varepsilon, \tag{3.2}$$

where $h_\varphi(M, A)$ is the isoperimetric ratio of A as in (2.12). We partition M into the sets

$$A_+ = \left\{ x \in M \mid |A \cap B(x,r)|_\varphi > \frac{1}{2C_\varphi^{r+1}} |B(x,r)|_\varphi \right\}, \tag{3.3}$$

$$M_0 = \left\{ x \in M \mid |A \cap B(x,r)|_\varphi = \frac{1}{2C_\varphi^{r+1}} |B(x,r)|_\varphi \right\}, \tag{3.4}$$

$$M_- = \left\{ x \in M \mid |A \cap B(x,r)|_\varphi < \frac{1}{2C_\varphi^{r+1}} |B(x,r)|_\varphi \right\}. \tag{3.5}$$

Clearly, $|A \cap D_x| \neq 0$ for all $x \in A_+ \cup M_0$. Since $|B(x,r)|_\varphi$ and $|A \cap B(x,r)|_\varphi$ depend continuously on x, a path from M_- to A_+ will pass through M_0. Since A is bounded, A_+ and M_0 are bounded. Moreover, $\partial A_+ \subseteq M_0$, A_+ and M_- are open, and M_0 is closed, hence compact. We will show that A_+ satisfies an inequality as required in Lemma 3.1. By passing from A to A_+, we get rid of a possibly "hairy structure" along the "outer part" of A. We pay by possibly losing regularity of the boundary.

We now choose a $2r$-separated subset X of M as follows. We start with a $2r$-separated subset $X_0 \subseteq M_0$ such that M_0 is contained in the union of the balls $B(x, 2r)$ with $x \in X_0$. (If $M_0 = \emptyset$, then $X_0 = \emptyset$.) We extend X_0 to a $2r$-separated subset $X_0 \cup X_+$ of $M_0 \cup A_+$ such that $M_0 \cup A_+$ is contained in the union of the balls $B(x, 2r)$ with $x \in X_0 \cup X_+$. (If $A_+ = \emptyset$, then $X_+ = \emptyset$.) We finally extend

$X_0 \cup X_+$ to a complete $2r$-separated subset $X = X_0 \cup X_+ \cup X_-$ of M. (If $M_- = \emptyset$, then $X_- = \emptyset$.) By definition, $X_+ \subseteq A_+$ and $X_- \subseteq M_-$. Since A is bounded and $|A \cap B(x,r)|_\varphi \neq 0$ for all $x \in X_0 \cup X_+$, the sets X_0 and X_+ are finite. By the same reason, the set Y of $x \in X_-$ with $|A \cap B(x,r)|_\varphi \neq 0$ is finite.

The neighborhood M_0^{2r} is covered by the balls $B(x, 4r)$ with $x \in X_0$. Using Theorem 2.15, (2.22), and (3.4), we therefore get

$$\begin{aligned} |M_0^{2r}|_\varphi &\leq \sum_{x \in X_0} |B(x,4r)|_\varphi \\ &\leq \frac{\beta(4r) C_\varphi^{4r+1}}{\beta(r)} \sum_{x \in X_0} |B(x,r)|_\varphi \\ &= \frac{2\beta(4r) C_\varphi^{5r+2}}{\beta(r)} \sum_{x \in X_0} |A \cap B(x,r)|_\varphi. \end{aligned}$$

For $x \in X_0 \subseteq M_0$, we have $|A \cap B(x,r)| \leq |B(x,r)|/2$ and hence

$$\frac{|\partial A \cap B(x,r)|}{|A \cap B(x,r)|} \geq h^N(B(x,r))$$

with $h^N(B(x,r))$ according to (2.8). Applying Lemma 2.16 to $D = B(x,r)$, we therefore obtain

$$\frac{|\partial A \cap B(x,r)|_\varphi}{|A \cap B(x,r)|_\varphi} \geq \frac{1}{C_\varphi^{r+1}} \frac{|\partial A \cap B(x,r)|}{|A \cap B(x,r)|} \geq \frac{C_{m,r}}{C_\varphi^{r+1}}.$$

Hence

$$\begin{aligned} |M_0^{2r}|_\varphi &\leq \frac{2\beta(4r) C_\varphi^{6r+3}}{\beta(r) C_{m,r}} \sum_{x \in X_0} |\partial A \cap B(x,r)|_\varphi \\ &\leq \frac{2\beta(4r) C_\varphi^{6r+3}}{\beta(r) C_{m,r}} |\partial A|_\varphi \\ &= \frac{2\beta(4r) C_\varphi^{6r+3}}{\beta(r) C_{m,r}} h_\varphi(M,A) |A|_\varphi \leq \varepsilon |A|_\varphi, \end{aligned} \tag{3.6}$$

where we use that $h_\varphi(M,A)$ satisfies (3.2).

Since any curve from A_+ to M_- passes through M_0, A_+ has distance at least $2r$ to $M_- \setminus M_0^{2r}$. Hence $M_- \setminus M_0^{2r}$ is covered by the Dirichlet domains D_x with $x \in X_-$.

With Y as above, we let $Z = X_0 \cup Y$. Using (3.4) and (3.5), we have

$$\frac{|A \cap B(x,r)|}{|B(x,r)|} \leq C_\varphi^{r+1} \frac{|A \cap B(x,r)|_\varphi}{|B(x,r)|_\varphi} \leq \frac{1}{2}$$

for any $x \in Z$. Letting $A^c = M \setminus A$, we obtain

$$\begin{aligned}|A^c \cap D_x| &\geq |A^c \cap B(x,r)| \geq \frac{1}{2}|B(x,r)| \\ &\geq \frac{\beta(r)}{2\beta(2r)}|D_x| \geq \frac{\beta(r)}{2\beta(2r)}|A \cap D_x| > 0\end{aligned}$$

for any $x \in Z$, where we use in the third inequality that $D_x \subseteq B(x, 2r)$. With the constant $C_{m,r}$ as in Lemma 2.16, we therefore get

$$\begin{aligned}C_{m,r} \leq h^N(D_x) &\leq \frac{|\partial A \cap \operatorname{int} D_x|}{\min\{|A \cap D_x|, |A^c \cap D_x|\}} \\ &\leq \frac{2\beta(2r)}{\beta(r)} \frac{|\partial A \cap \operatorname{int} D_x|}{|A \cap D_x|} \\ &\leq \frac{2\beta(2r)C_\varphi^{2r+1}}{\beta(r)} \frac{|\partial A \cap \operatorname{int} D_x|_\varphi}{|A \cap D_x|_\varphi}\end{aligned} \tag{3.7}$$

for any $x \in Z$, where we use again, now in the last inequality, that $D_x \subseteq B(x, 2r)$. Using (3.7) and (3.2), we conclude that

$$\begin{aligned}|A \cap (M_- \setminus M_0^{2r})|_\varphi &\leq \sum_{x \in Z} |A \cap D_x|_\varphi \leq \frac{2\beta(2r)C_\varphi^{2r+1}}{\beta(r)C_{m,r}} \sum_{x \in Z} |\partial A \cap \operatorname{int} D_x|_\varphi \\ &\leq \frac{2\beta(2r)C_\varphi^{2r+1}}{\beta(r)C_{m,r}} |\partial A|_\varphi \\ &= \frac{2\beta(2r)C_\varphi^{2r+1}}{\beta(r)C_{m,r}} h_\varphi(M, A)|A|_\varphi \leq \varepsilon |A|_\varphi,\end{aligned} \tag{3.8}$$

where we use (3.2) in the last step, recalling that $C_\varphi \geq 1$.

Since $A \subseteq A_+ \cup M_0^{2r} \cup (A \cap (M_- \setminus M_0^{2r}))$, we obtain

$$\begin{aligned}|A_+|_\varphi &\geq |A|_\varphi - |M_0^{2r}|_\varphi - |A \cap (M_- \setminus M_0^{2r})|_\varphi \\ &\geq (1 - 2\varepsilon)|A|_\varphi.\end{aligned}$$

In particular, A_+ is not empty. Since $A_+^{2r} \setminus A_+ \subseteq M_0^{2r}$, we conclude that

$$|A_+^{2r} \setminus A_+|_\varphi \leq |M_0^{2r}|_\varphi \leq \varepsilon |A|_\varphi \leq \frac{\varepsilon}{1 - 2\varepsilon}|A_+|_\varphi.$$

In conclusion, A_+ is a bounded open subset of M that satisfies an inequality as asserted in Lemma 3.1, albeit with $2r$ and 2ε in place of r and ε (assuming w.l.o.g. that $\varepsilon < 1/4$). □

Whereas $\varepsilon > 0$ should be viewed as small, the number r is large in our application of Lemma 3.1 (see (4.9)). The difference to Buser's discussion lies in the fact that in Lemma 3.1, for ε and r are given, the domain A is chosen according to (3.2).

Remark 3.9. Let M be a non-compact, complete, and connected Riemannian manifold of dimension m with $\mathrm{Ric}_M \geq (1-m)b^2$. Let $V\colon M \to \mathbb{R}$ be a smooth potential on M, and assume that V and ∇V are bounded. Let φ be a positive λ-eigenfunction of the associated Schrödinger operator S on M. Following the above line of proof and Buser's arguments at the end of his short proof of Theorem 1.2 in [4], one obtains inequalities of the form

$$\begin{aligned}\lambda_0(M,V)-\lambda &\le C'_{m,\|V-\lambda\|_\infty,\|\nabla V\|_\infty}\max\{bh_\varphi(M),h_\varphi(M)^2\},\\ \lambda_{\mathrm{ess}}(M,V)-\lambda &\le C'_{m,\|V-\lambda\|_\infty,\|\nabla V\|_\infty}\max\{bh_{\varphi,\mathrm{ess}}(M),h_{\varphi,\mathrm{ess}}(M)^2\}.\end{aligned}\tag{3.10}$$

To get rid of the squares $h_\varphi(M)^2$ and $h_{\varphi,\mathrm{ess}}(M)^2$, respectively, we change Buser's argument at the end of his proof of [4, Theorem 7.1] and estimate

$$\begin{aligned}h_\varphi(M),h_{\varphi,\mathrm{ess}}(M) &\le \sup_{x\in M} h_\varphi(B(x,1)) \le C_\varphi^2 \sup_{x\in M} h(B(x,1))\\ &\le 2C_\varphi^2 \sup_{x\in M}\lambda_0(B(x,1))^{1/2} \le 2C_\varphi^2\lambda_0(B)^{1/2}\\ &\le bC''_{m,\|V-\lambda\|_\infty,\|\nabla V\|_\infty},\end{aligned}$$

where we use the definition of h_φ and $h_{\varphi,\mathrm{ess}}$ as in (2.13), the Harnack constant of φ as in (2.22), the Cheeger inequality (2.6), and Cheng's [5, Theorem 1.1], where B denotes a ball of radius 1 in the m-dimensional hyperbolic space of sectional curvature $-b^2$. We finally arrive at the inequalities

$$\begin{aligned}\lambda_0(M,V)-\lambda &\le C_{m,\|V-\lambda\|_\infty,\|\nabla V\|_\infty} bh_\varphi(M),\\ \lambda_{\mathrm{ess}}(M,V)-\lambda &\le C_{m,\|V-\lambda\|_\infty,\|\nabla V\|_\infty} bh_{\varphi,\mathrm{ess}}(M),\end{aligned}\tag{3.11}$$

which extend Buser's [4, Theorem 7.1]. The dependence of $C_{m,\|V-\lambda\|_\infty,\|\nabla V\|_\infty}$ on $C_{3,m}$ (as in (2.23)), $\|V-\lambda\|_\infty$, and $\|\nabla V\|_\infty$ is exponential in our approach and, in particular, exponential in λ. Therefore the use of the estimates seems to be restricted. However, together with (2.14), they have at least the consequence that $\lambda_0(M,V)=\lambda$ if and only if $h_\varphi(M)=0$ and that $\lambda_{\mathrm{ess}}(M,V)=\lambda$ if and only if $h_{\varphi,\mathrm{ess}}(M)=0$.

4. Back to Riemannian coverings

We return to the situation of a Riemannian covering as in the introduction. Suppose that the Ricci curvature of M_0 is bounded from below. Let V_0 be a smooth potential on M_0 with $\|V_0\|_\infty,\|\nabla V_0\|_\infty<\infty$ and set $V_1=V_0\circ\pi$. Let $\lambda=\lambda_0(M_0,V_0)$ and φ_0 be a positive λ-eigenfunction of $S_0=\Delta+V_0$ on M_0. Then $\varphi=\varphi_0\circ\pi$ is a positive λ-eigenfunction of $S_1=\Delta+V_1$ on M_1.

Theorem 4.1. *If $\lambda_{\mathrm{ess}}(M_0,V_0)>\lambda_0(M_0,V_0)$, then $\lambda_0(M_1,V_1)=\lambda_0(M_0,V_0)$ if and only if the covering $\pi\colon M_1\to M_0$ is amenable.*

Consider the following three implications:

1. If $\pi\colon M_1\to M_0$ is amenable, then $\lambda_0(M_1,V_1)=\lambda_0(M_0,V_0)$.
2. If $\lambda_0(M_1,V_1)=\lambda_0(M_0,V_0)$, then $h_\varphi(M_1)=0$.
3. If $h_\varphi(M_1)=0$, then $\pi\colon M_1\to M_0$ is amenable.

The first one is [1, Theorem 1.2] and the second is an immediate consequence of (2.14). These two assertions hold without any assumptions on the curvature of M and the potential V. The third one does not hold without any further assumptions. We require that the Ricci curvature of M_0 is bounded from below, that the potential V_0 and its derivative dV_0 are bounded, and that $\lambda_{\mathrm{ess}}(M_0, V_0) > \lambda_0(M_0, V_0)$. To prove Theorem 4.1, and therewith also Theorem 1.3, it remains to establish the third implication under these additional assumptions. We need to prove that the right action of Γ_0 on $\Gamma_1 \backslash \Gamma_0$ is amenable. To that end, we will show that the Følner criterion for amenability is satisfied.

Følner criterion 4.2. *The right action of a countable group Γ on a countable set X is amenable if and only if, for any finite subset $G \subseteq \Gamma$ and $\varepsilon > 0$, there is a finite subset $F \subseteq X$ such that*

$$\#(F \setminus Fg) < \varepsilon \#(F) \quad \textit{for all } g \in G.$$

Proof of Theorem 4.1. Since $\lambda_{\mathrm{ess}}(M_0, V_0) > \lambda_0(M_0, V_0)$, there is a compact domain $K \subseteq M_0$ such that

$$\lambda_0(M_0 \setminus K, V_0) > \lambda_0(M_0, V_0). \tag{4.3}$$

Since $\pi\colon M_1 \setminus \pi^{-1}(K) \to M_0 \setminus K$ is a Riemannian covering, we have

$$\lambda_0(M_1 \setminus \pi^{-1}(K), V_1) \geq \lambda_0(M_0 \setminus K, V_0). \tag{4.4}$$

Note that the manifolds $M_0 \setminus K$ and $M_1 \setminus \pi^{-1}(K)$ might not be connected, but the assertion still holds since the inequality applies to each component of $M_0 \setminus K$ and connected component of $M_1 \setminus \pi^{-1}(K)$ over it.

Let χ_0 be a smooth cut-off function on M_0 which is equal to 0 on a neighborhood of K in M_0 and equal to 1 outside a compact domain $K_0 \subseteq M_0$ and set $\chi = \chi_0 \circ \pi$.

Lemma 4.5. *For all $r, \varepsilon > 0$, there is a bounded open subset $A \subseteq M_1$ and a point $x \in K_0$ such that $\pi^{-1}(x) \cap A \neq \emptyset$ and*

$$\frac{\#\big(\pi^{-1}(x) \cap (A^r \setminus A)\big)}{\#\big(\pi^{-1}(x) \cap A\big)} < \varepsilon.$$

Proof. Since M_1 is complete with Ricci curvature bounded from below and $h_\varphi(M_1) = 0$, Lemma 3.1 implies that there exist bounded open subsets $A_n \subseteq M_1$ such that

$$\frac{|A_n^r \setminus A_n|_\varphi}{|A_n|_\varphi} < \frac{1}{n}. \tag{4.6}$$

Let f_n be the Lipschitz function on M_1 with compact support defined by

$$f_n(x) = \begin{cases} 1 - d(x, A_n)/r & \text{for } x \in A_n^r, \\ 0 & \text{for } x \in M_1 \setminus A_n^r. \end{cases} \tag{4.7}$$

For the φ-Rayleigh quotient of f_n, we have

$$\begin{aligned} R_\varphi(f_n) &= \frac{\int_{M_1} \|\operatorname{grad} f_n\|^2\varphi^2}{\int_{M_1} f_n^2\varphi^2} \le \frac{\int_{A_n^r\setminus A_n} \|\operatorname{grad} f_n\|^2\varphi^2}{\int_{A_n} f_n^2\varphi^2} \\ &= \frac{1}{r^2}\frac{|A_n^r \setminus A_n|_\varphi}{|A_n|_\varphi} \le \frac{1}{nr^2}. \end{aligned} \tag{4.8}$$

Normalize f_n to $g_n = f_n/\|f_n\|$, where $\|f_n\|$ denotes the modified L^2-norm of f_n, that is, $\|f_n\|^2 = \int_{M_1} f_n^2\varphi^2$. Then

$$R_\varphi(g_n) = R_\varphi(f_n) \le 1/nr^2 \to 0.$$

Let $\mathcal{R} \subseteq M_0$ be the subset of full measure such that all g_n are differentiable at all $y \in \pi^{-1}(\mathcal{R})$. Suppose now that

$$\sum_{y\in\pi^{-1}(x)} \|\operatorname{grad} g_n(y)\|^2 \ge \varepsilon \sum_{y\in\pi^{-1}(x)} g_n(y)^2$$

for all $n \in \mathbb{N}$ and $x \in K_0 \cap \mathcal{R}$. Since π is a Riemannian covering and φ is constant along the fibers of π, we then have

$$\int_{\pi^{-1}(K_0)} \|\operatorname{grad} g_n\|^2\varphi^2 \ge \varepsilon \int_{\pi^{-1}(K_0)} g_n^2\varphi^2.$$

Since $\|g_n\| = 1$ and $R_\varphi(g_n) \le 1/nr^2 \to 0$, we get that

$$\int_{\pi^{-1}(K_0)} g_n^2\varphi^2 \to 0 \ \text{ and, as a consequence, } \int_{M_1\setminus\pi^{-1}(K_0)} g_n^2\varphi^2 \to 1.$$

Consider now $h_n = \chi g_n$ with χ as further up. Then h_n has compact support in $M_1 \setminus \pi^{-1}(K)$. Furthermore,

$$\int_{M_1} h_n^2\varphi^2 = \int_{\pi^{-1}(K_0)} h_n^2\varphi^2 + \int_{M_1\setminus\pi^{-1}(K_0)} g_n^2\varphi^2 \to 0 + 1$$

and

$$\begin{aligned} \int_{M_1} \|\operatorname{grad} h_n\|^2\varphi^2 &\le 2\int_{\pi^{-1}(K_0)} \big(g_n^2\|\operatorname{grad}\chi\|^2 + \chi^2\|\operatorname{grad} g_n\|^2\big)\varphi^2 \\ &\quad + \int_{M\setminus\pi^{-1}(K_0)} \|\operatorname{grad} g_n\|^2\varphi^2 \to 0, \end{aligned}$$

where we use that $0 \le \chi \le 1$, that $\operatorname{grad}\chi$ is uniformly bounded, and that $\int_{M_1} \|\operatorname{grad} g_n\|^2\varphi^2 \to 0$. Hence the modified Rayleigh quotients $R_\varphi(h_n) \to 0$. This is in contradiction to (4.4) since the h_n are Lipschitz functions on M_1 with compact support in $M_1 \setminus \pi^{-1}(K_0)$. It follows that there are an n and an $x \in K_0 \cap \mathcal{R}$ such that

$$\sum_{y\in\pi^{-1}(x)} \|\operatorname{grad} g_n(y)\|^2 < \varepsilon \sum_{y\in\pi^{-1}(x)} g_n(y)^2.$$

Since $g_n = 0$ on $M_1 \setminus A_n^r$, we must have $\pi^{-1}(x) \cap A_n^r \neq \emptyset$. Furthermore, since $0 \le g_n \le 1/\|f_n\|$ and $\|\operatorname{grad} g_n\| = 1/r\|f_n\|$ on $\pi^{-1}(\mathcal{R}) \cap (A_n^r \setminus A_n)$, we conclude that

$$\frac{1}{r^2\|f_n\|^2}\#(\pi^{-1}(x) \cap (A_n^r \setminus A_n)) \le \frac{\varepsilon}{\|f_n\|^2}\#(\pi^{-1}(x) \cap A_n^r).$$

This yields that

$$\#(\pi^{-1}(x) \cap (A_n^r \setminus A_n)) < \varepsilon r^2 \#(\pi^{-1}(x) \cap A_n^r).$$

Since A_n^r is the disjoint union of A_n with $A_n^r \setminus A_n$, we conclude that

$$\#(\pi^{-1}(x) \cap (A_n^r \setminus A_n)) < \frac{\varepsilon r^2}{1 - \varepsilon r^2}\#(\pi^{-1}(x) \cap A_n)$$

as long as $\varepsilon < 1/r^2$. In particular, $\pi^{-1}(x) \cap A_n \neq \emptyset$ if $\varepsilon < 1/r^2$. □

We return to the proof of the amenability of the right action of Γ_0 on $\Gamma_1 \backslash \Gamma_0$. We will use Følner's criterion 4.2 and let $G \subseteq \Gamma_0$ be a finite subset and $\varepsilon > 0$. We need to show that there is a non-empty finite subset $F \subseteq \Gamma_1 \backslash \Gamma_0$ such that

$$\#(F \setminus Fg) < \varepsilon \#(F) \text{ for all } g \in G.$$

Write K_0 as the union of finitely many compact and connected domains $D_i \subseteq M_0$ which are evenly covered with respect to the universal covering $\pi_0 \colon M \to M_0$ of M_0. For each i, let B_i be a lift of D_i to a leaf of π_0 over D_i. Then each B_i is a compact subset of M with $\pi_0(B_i) = D_i$. Since there are only finitely many B_i and all of them are compact, there is a number $r > 0$ such that

$$d(u, g^{-1}u) < r \text{ for all } g \in G \text{ and } u \in \cup_i B_i. \tag{4.9}$$

Let $R \subseteq \Gamma_0$ be a set of representatives of the right cosets of Γ_1 in Γ_0, that is, of the elements of $\Gamma_1 \backslash \Gamma_0$. Corresponding to ε and r, choose $x \in K_0$ and A as in Lemma 4.5. Fix preimages $u \in M$ and $y = \pi_1(u) \in M_1$ of x under π_0 and π, respectively, and write $\pi_0^{-1}(x) = \Gamma_0 u$ as the union of Γ_1-orbits $\Gamma_1 g u$. Then $\pi^{-1}(x) = \{\pi(gu) \mid g \in R\}$. Set

$$F = \{\Gamma_1 h \mid h \in R \text{ and } \pi_1(hu) \in \pi^{-1}(x) \cap A\}.$$

Then $\#(F) = \#(\pi^{-1}(x) \cap A) \neq 0$.

Let now $g \in G$ and $h \in R$ with $\Gamma_1 h \in F \setminus Fg$. Then

$$\pi_1(hu) \in \pi^{-1}(x) \cap A \text{ and } \pi_1(hg^{-1}u) \in \pi^{-1}(x) \setminus A.$$

Since

$$d(\pi_1(hu), \pi_1(hg^{-1}u)) \le d(hu, hg^{-1}u) = d(u, g^{-1}u) < r$$

for all $g \in G$, we get that $\pi_1(hg^{-1}u) \in A^r$. Hence $\pi_1(hg^{-1}u)$ belongs to $A^r \setminus A$ and therefore

$$\begin{aligned}\#(F \setminus Fg) &\le \#\big(\pi^{-1}(x) \cap A^r \setminus A)\big) \\ &< \varepsilon \#\big(\pi^{-1}(x) \cap A\big) = \varepsilon \#(F).\end{aligned}$$

Since G and ε were arbitrary, we conclude from the Følner criterion 4.2 that the right action of Γ_0 on $\Gamma_1\backslash\Gamma_0$ is amenable. $\square$

Proof of Theorem 1.10.2. Let M_0 be the interior of a compact manifold N_0 as in the definition of conformally compact (in the introduction), and denote by g_0, h_0, and ρ_0 the corresponding Riemannian metrics and defining function ρ_0 of ∂N_0. Let $X = \operatorname{grad}\rho_0/\|\operatorname{grad}\rho_0\|^2$, where the gradient of ρ_0 is taken with respect to h_0. Since ∂N_0 is compact, the flow of X leads to a diffeomorphism of a neighborhood of ∂N_0 in N_0 with $\partial N_0 \times [0, y_0)$ with respect to which $\rho(x, y) = y$ for $(x, y) \in \partial N_0 \times [0, y_0)$. Then

$$g_0(x, y) = \frac{1}{y^2} h_0(x, y)$$

on $\partial N_0 \times (0, y_0) = \partial N_0 \times [0, y_0) \cap M_0$. This is reminiscent of the upper half-space model of the hyperbolic space H^m.

From standard formulas for conformal metrics it is now easy to see that, for all $x_0 \in \partial N_0$ and $\varepsilon > 0$, there exists a neighborhood U of $(x_0, 0) \in \partial N_0 \times [0, y_0)$ such that the sectional curvature of each tangent plane at each (x, y) in $U \cap M_0$ is in

$$(-(\partial_\nu \rho(x, 0))^2 - \varepsilon, -(\partial_\nu \rho(x, 0))^2 + \varepsilon),$$

where ν denotes the inner normal of N_0 along ∂N_0 with respect to h_0. Note that, for any $r > 0$, the g_0-ball $B((x_0, y), r))$ is contained in $U \cap M_0$ for all sufficiently small $y > 0$. From Cheng's [5, Theorem 1.1], we conclude that $\lambda_0(M_0) \le a^2(m-1)^2/4$.

Since M_0 is homotopy equivalent to N_0, there is a covering $\pi_1 \colon N_1 \to N_0$ which restricts to the covering $M_1 \to M_0$ and such that M_1 is the interior of the manifold N_1, but where the boundary ∂N_1 of N_1 need not be compact anymore. Nevertheless, lifting g_0, h_0, and ρ_0 to Riemannian metrics g_1 on M_1, h_1 on N_1, and defining function $\rho_1 = \rho_0 \circ \pi_1$ of ∂N_1, the above statement about sectional curvature remains valid for

$$\partial N_1 \times [0, y_0) = \pi^{-1}(\partial N_0 \times [0, y_0)).$$

In particular, we have $\lambda_0(M_1) \le a^2(m-1)^2/4$.

Now we are ready for the final step of the proof. By assumption and (1.1),

$$a^2(m-1)^2/4 = \lambda_0(M_0) \le \lambda_0(M_1) \le a^2(m-1)^2/4.$$

Hence $\lambda_0(M_0) = \lambda_0(M_1)$ as asserted. $\square$

Proof of Proposition 1.13.1. By definition, $\lambda_{\text{ess}}(M_1) > \lambda_0(M_1) =: \lambda$ would imply that λ does not belong to the essential spectrum of M_1. Hence λ would be an eigenvalue of M_1 with a square integrable positive eigenfunction φ. On the other hand, the lift ψ of a positive λ-eigenfunction from M_0 to M_1 is also a positive λ-eigenfunction, but definitely not square integrable since π is an infinite covering. Now by Sullivan's [14, Theorems 2.7 and 2.8], the space of positive, but not necessarily square integrable, λ-eigenfunctions on M_1 is of dimension one. Hence ψ would be a multiple of φ, a contradiction. $\square$

Proof of (1.5). By [7, Theorem 3.1], each end of M_0 has a neighborhood of the form $U = \Gamma_\infty \backslash B$, where B is a horoball in the universal covering space M of M_0 and $\Gamma_\infty \subseteq \Gamma_0$ is the stabilizer of the center ξ of B in the sphere of M at infinity. Furthermore, Γ_ξ leaves the Busemann functions associated to ξ invariant. We let b be the one such that $\{b = 0\}$ is the horosphere ∂B. Then the level sets $\{b = -y\}$, $y > 0$, are horospheres foliating B. They are perpendicular to the unit speed geodesics γ_z starting in $z \in \{b = 0\}$ and ending in ξ. Moreover, $b(\gamma_z(y)) = -y$ and $\operatorname{grad} b(\gamma_z(y)) = -\dot\gamma_z(y)$. Since Busemann functions are C^2 (see [8, Proposition 3.1]), we obtain a C^2-diffeomorphism

$$\{b = 0\} \times (0, \infty) \to B, \quad (z, y) \mapsto \gamma_z(y).$$

Since Γ_ξ leaves b invariant, we arrive at a C^2-diffeomorphism $U \cong N \times (0, \infty)$, where $N = \Gamma_\xi \backslash \{b = 0\}$ and where the curves $\gamma_x = \gamma_x(y) = (x, y)$ are unit speed geodesics perpendicular to the cross sections $\{y = \text{const}\}$. The latter lift to the horospheres $\{b = \text{const}\}$ in B and, therefore, have second fundamental form $\le -a$ with respect to the unit normal field $Y = \partial/\partial y$. In particular, their mean curvature is $\le (1 - m)a$ with respect to Y. For the divergence of Y, we have

$$\operatorname{div} Y = \sum \langle \nabla_{E_i} Y, E_i \rangle = -\sum \langle Y, \nabla_{E_i} E_i \rangle,$$

where (E_i) is a local orthonormal frame. We choose it such that $E_1 = Y$. Then $\nabla_{E_1} E_1 = 0$, and we see that $\operatorname{div} Y$ is the mean curvature of the corresponding cross section with respect to the unit normal field Y, hence is $\le (1 - m)a$. All this is well known, but we recall it for convenience.

For a compact domain A in U with smooth boundary ∂A and outer unit normal field ν, we obtain from the above that

$$|\partial A| \ge -\int_{\partial A} \langle Y, \nu \rangle = -\int_A \operatorname{div} Y \ge a(m - 1)|A|.$$

Hence the Cheeger constant of U is at least $a(m - 1)$. The claim about $\lambda_{\mathrm{ess}}(M_0)$ now follows from the Cheeger inequality (2.6). □

Acknowledgment

We are grateful to the Max Planck Institute for Mathematics and the Hausdorff Center for Mathematics for their support and hospitality.

References

[1] W. Ballmann, H. Matthiesen, P. Polymerakis, On the bottom of spectra under coverings. *Math. Zeitschrift* **288** (2018), 1029–1036, MR3778988, Zbl 1394.58012.

[2] R. Brooks, The fundamental group and the spectrum of the Laplacian. *Comment. Math. Helv.* **56** (1981), no. 4, 581–598, MR0656213, Zbl 0495.58029.

[3] R. Brooks, The bottom of the spectrum of a Riemannian covering. *J. Reine Angew. Math.* **357** (1985), 101–114, MR0783536, Zbl 0553.53027.

[4] P. Buser, *A note on the isoperimetric constant.* Ann. Sci. École Norm. Sup. (4) 15 (1982), no. 2, 213–230, MR0683635, Zbl 0501.53030.

[5] S.Y. Cheng, Eigenvalue comparison theorems and its geometric applications. *Math. Z.* **143** (1975), no. 3, 289–297, MR0378001, Zbl 0329.53035.

[6] S.Y. Cheng, S.T. Yau, Differential equations on Riemannian manifolds and their geometric applications. *Comm. Pure Appl. Math.* **28** (1975), no. 3, 333–354, MR0385749, Zbl 0312.53031.

[7] P. Eberlein, Lattices in spaces of nonpositive curvature, *Ann. of Math.* **111** (1980), 435–476, MR0577132, Zbl 0401.53015.

[8] E. Heintze, H.-C. Im Hof, Geometry of horospheres. *J. Differential Geom.* **12** (1977), no. 4, 481–491 (1978), MR0512919, Zbl 0434.53038.

[9] P.D. Lax, R.S. Phillips, The asymptotic distribution of lattice points in Euclidean and non-Euclidean spaces. *Toeplitz centennial* (Tel Aviv, 1981), pp. 365–375, Operator Theory: Adv. Appl., 4, Birkhäuser, Basel-Boston, Mass., 1982, MR0650290, Zbl 0497.52007.

[10] R. Mazzeo, The Hodge cohomology of a conformally compact metric. *J. Differential Geom.* **28** (1988), no. 2, 309–339, MR0961517, Zbl 0656.53042.

[11] P. Polymerakis, On the spectrum of differential operators under Riemannian coverings. *J. Geom. Anal.* (2019), doi.org/10.1007/s12220-019-00196-1.

[12] P. Polymerakis, Coverings preserving the bottom of the spectrum, MPIM-Preprint 2019-3, arxiv.org/abs/1811.07844.

[13] T. Roblin, S. Tapie, *Exposants critiques et moyennabilité.* Géométrie ergodique, 61–92, Monogr. Enseign. Math., 43, Enseignement Math., Geneva, 2013, MR3220551, Zbl 1312.53060.

[14] D. Sullivan, Related aspects of positivity in Riemannian geometry. *J. Differential Geom.* **25** (1987), no. 3, 327–351, MR0882827, Zbl 0615.53029.

Werner Ballmann and Henrik Matthiesen
Max Planck Institute for Mathematics
Vivatsgasse 7
D-53111 Bonn, Germany

and

Hausdorff Center for Mathematics
Endenicher Allee 60
D-53115 Bonn, Germany
e-mail: `hwbllmnn@mpim-bonn.mpg.de`
`hematt@mpim-bonn.mpg.de`

Panagiotis Polymerakis
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
D-10099 Berlin, Germany
e-mail: `polymerp@hu-berlin.de`

Progress in Mathematics, Vol. 333, 37–56

Some Remarks on the Geometry of a Class of Locally Conformally Flat Metrics

Sun-Yung A. Chang, Zheng-Chao Han and Paul Yang

In honor of Gang Tian on the occasion of his sixtieth birthday

Abstract. We prove that conformal metrics on domains of the round sphere, with non-negative constant Q-curvature, and non-negative scalar curvature, has positive mean curvature on the boundary of embedded balls (in the round metric). As a result, such metrics have certain reflection symmetries if the complement of the domain is contained in a lower-dimensional round sphere. We also prove that the development map of a locally conformally flat metric with non-positive Schouten tensor is an embedding.

Mathematics Subject Classification (2010). Primary 58J06; 53C21; Secondary 35B06.

Keywords. Conformal metrics, σ_k-curvature, Q-curvature, moving planes, moving spheres.

1. Introduction

An important question in conformal geometry is to understand under what conditions the development map of a locally conformally flat manifold into the standard sphere is an embedding. Another related question is to understand conditions on a domain $\Omega \subset \mathbb{S}^n$ under which a complete conformal metric exists on Ω with constant scalar curvature; also relevant is the uniqueness of such a metric, or possibly cataloging such metrics when uniqueness fails and Ω is some canonical domain in $\mathbb{S}^n$ such as $\mathbb{S}^n \setminus \mathbb{S}^l$ for some $0 \leq l < n$.

In [SY88] Schoen and Yau found some sufficient conditions for the development map of a locally conformally flat manifold to be an embedding. In particular they proved that the answer is positive if the Yamabe constant of (M, g) is non-negative. No positive result is known, as far as we are aware, in the case when the Yamabe constant of (M, g) is negative. In general the development map may not be an embedding, as shown by the elementary examples $\mathbb{S}^1_r \times \mathbb{H}^{n-1}$ where r

is the radius of the circle; moreover, these examples also show that the holonomy representation of the fundamental group of M under the development map may not be discrete. However, Kulkarni and Pinkall showed in [KP86] that for a closed conformally flat n-manifold M with infinite fundamental group, its development map $d : M \mapsto \mathbb{S}^n$ is a covering map, *iff* d is not surjective.

In [SY88], Schoen–Yau also proved that if a complete metric $g = v^{-2}(x)|dx|^2$ exists on a domain $\Omega \subset \mathbb{R}^n$ with its scalar curvature having a positive lower bound, then the Hausdorff dimension of $\partial\Omega$ has to be $\leq \frac{n-2}{2}$. In another direction, Schoen constructed in [S88] complete conformal metrics with scalar curvature 1 on $\mathbb{S}^n \setminus \Lambda$ when Λ is a certain subset of $\mathbb{S}^n$, including the case when it is any finite set with at least two points. Later Mazzeo and Parcard [MP96] [MP99] proved that if $\Omega \subset \mathbb{S}^n$ is a domain such that $\mathbb{S}^n \setminus \Omega$ consists a finite number of smooth submanifold of dimension $< \frac{n-2}{2}$, then one can find a complete metric $g = v^{-2}(x)|dx|^2$ on Ω with its scalar curvature identical to $+1$.

For the negative scalar curvature case, the works of Löwner–Nirenberg [LN75], Aviles [A82], and Veron [V81] imply that if $\Omega \subset \mathbb{S}^n$ admits a complete, conformal metric with negative constant scalar curvature, then the Hausdorff dimension of $\partial\Omega > \frac{n-2}{2}$. Löwner–Nirenberg [LN75] also proved that if $\Omega \subset \mathbb{S}^n$ is a domain with smooth boundary $\partial\Omega$ of dimension $> \frac{n-2}{2}$, then there exists a complete metric $g = v^{-2}(x)|dx|^2$ on Ω with its scalar curvature $= -1$; such a metric is unique when $\partial\Omega$ consists of hypersurfaces. This result was later generalized by D. Finn [F95] to the case of $\partial\Omega$ consisting of smooth submanifolds of dimension $> \frac{n-2}{2}$ and with boundary.

It is natural to ask whether some kind of additional curvature condition in the negative Yamabe constant case would force the development map to be an embedding as well, and whether further curvature conditions would improve the estimate on the Hausdorff dimension of $\partial\Omega$?

The additional curvature conditions are often imposed in terms of the σ_k or Q-curvature of a representative metric g. The σ_k-curvature, denoted as $\sigma_k(A_g)$, refers to the kth elementary symmetric functions of the eigenvalues of the 1-1 tensor derived from the Weyl–Schouten tensor A_g of the conformal metric g,

$$A_g = \frac{1}{n-2}\left\{\mathrm{Ric} - \frac{R}{2(n-1)}g\right\}.$$

Note that the $\sigma_1(A_g)$ curvature is simply the scalar curvature of g, up to a dimensional constant.

The condition involving the σ_k-curvature often assumes that the Weyl–Schouten tensor A_g is in the Γ_k^+ class for some $k > 1$, *i.e.*, the eigenvalues, $\lambda_1 \leq \cdots \leq \lambda_n$, of A_g at each x satisfy $\sigma_j(\lambda_1, \ldots, \lambda_n) > 0$ for all $j, 1 \leq j \leq k$. It is also natural to consider metrics whose Weyl–Schouten tensor A_g is in Γ_k^- class, namely, $(-1)^j\sigma_j(\lambda_1, \ldots, \lambda_n) > 0$ for all $j, 1 \leq j \leq k$. It is known that the operator $w \mapsto \sigma_k(A_{e^{2w}g_0})$ is elliptic when the Weyl–Schouten tensor of $g = e^{2w}g_0$ is in either Γ_k^+ or Γ_k^- class.

The Q-curvature of a metric g is defined through

$$Q_g = c_n |Rc_g|^2 + d_n |R_g|^2 - \frac{\Delta_g R_g}{2(n-1)},$$

with c_n and d_n being some dimensional constants: $c_n = -\frac{2}{(n-2)^2}$ and $d_n = \frac{n^3-4n^3+16n-16}{8(n-1)^2(n-2)^2}$. Note that Q_g involves 4th-order derivatives of the metric. The Q-curvatures of two conformally related metrics g and $g_u = u^{\frac{n+4}{n-4}} g$ (for $n \neq 4$) have the following relation through a 4th-order differential operator P_g, called the Panietz-type operator:

$$P_g(u) = \frac{n-4}{2} Q_{g_u} u^{\frac{n+4}{n-4}} \quad \text{for } n \neq 4,$$

where $P_g(u) = (-\Delta_g)^2 u + \operatorname{div}[(a_n R_g g + b_n \operatorname{Ric}_g) du] + \frac{n-4}{2} Q_g u$ for some dimensional constants a_n and b_n. For $n = 4$, $P_g(u) = (-\Delta_g)^2 u + \operatorname{div}[(\frac{2}{3} R_g - 2 \operatorname{Ric}_g) du]$, and the relation between the Q-curvatures takes the following form:

$$P_g(w) + 2Q_g = 2Q_{e^{2w}g} e^{4w}.$$

P_g enjoys certain conformal covariance properties much like those of the conformal Laplace operator; see [CY97] for more details.

In [CHgY04], Chang, Hang, and Yang proved that *if* $\Omega \subset \mathbb{S}^n$ $(n \geq 5)$ *admits a complete, conformal metric* g *with*

$$\sigma_1(A_g) \geq c_1 > 0, \quad \sigma_2(A_g) \geq 0, \quad \text{and}$$
$$|R_g| + |\nabla_g R|_g \leq c_0, \tag{1.1}$$

then $\dim(\mathbb{S}^n \setminus \Omega) < \frac{n-4}{2}$. This has been generalized by M. Gonzáles [G04] to the case of $2 < k < n/2$: *if* $\Omega \subset \mathbb{S}^n$ *admits a complete, conformal metric* g *with*

$$\sigma_1(A_g) \geq c_1 > 0, \quad \sigma_2(A_g), \; \ldots, \; \sigma_k(A_g) \geq 0, \quad \text{and (1.1)},$$

then $\dim(\mathbb{S}^n \setminus \Omega) < \frac{n-2k}{2}$. See also the work of Guan, Lin and Wang [GLW04].

[CHgY04] also contains a result involving conditions on the Q-curvature: *if* $\Omega \subset \mathbb{S}^n$ $(n \geq 3)$ *admits a complete, conformal metric* g *with* $R_g \geq c_1 > 0$ *and* $Q_g \geq c_2 > 0$*, then* $\dim(S^n \setminus \Omega) < \frac{n-4}{2}$*. In particular, this means* $\Omega = S^n$ *when* $n \leq 4$*. If we replace* $Q_g \geq c_2 > 0$ *by* $Q_g \geq 0$*, then when* $n \geq 5$*, we have* $\dim(S^n \setminus \Omega) \leq \frac{n-4}{2}$.

There are earlier results involving the Q-curvature that are relevant to the discussion here: they concern the radial symmetry and classification of solutions to constant Q-curvature equations on $\mathbb{R}^n$. When $g = u^{\frac{4}{n-4}} |dx|^2$ on a domain in $\mathbb{R}^n$, $u > 0$, $n \neq 4$, the Q-curvature Q_g of g is computed through

$$(-\Delta)^2 u = \frac{n-4}{2} Q_g u^{\frac{n+4}{n-4}};$$

while on a domain in $\mathbb{R}^4$, if $g = e^{2w} |dx|^2$, then

$$(-\Delta)^2 w = 2Q_g e^{4w}.$$

In [CY97] Chang and Yang proved that any entire smooth solution $u(x)$ to

$$(-\Delta)^{\frac{n}{2}} u(x) = (n-1)! e^{nu(x)} \quad \text{on} \quad \mathbb{R}^n \tag{1.2}$$

with the asymptotic behavior as $x \to \infty$:

$$\begin{aligned} u(x) = \log \frac{2}{1+|x|^2} + w\left(\frac{x}{1+|x|^2}\right) \text{ for some} \\ \text{smooth function } w \text{ defined near } 0, \end{aligned} \tag{1.3}$$

must be rotationally symmetric with respect to some point in $\mathbb{R}^n$, and of the form

$$\log \frac{2\lambda}{1+\lambda^2|x-x_0|^2} \text{ for some } x_0 \in \mathbb{R}^n \text{ and constant } \lambda > 0. \tag{1.4}$$

In [L98] C.S. Lin obtained related results. For (1.2) in the case $n = 4$, Lin obtained the same result as in [CY97] under an integral assumption

$$\int_{\mathbb{R}^4} e^{4u(y)} dy = \frac{8\pi^2}{3}. \tag{1.5}$$

Lin's result actually implies that any solution u to (1.2) for the $n = 4$ case with $\int_{\mathbb{R}^4} e^{4u(y)} dy < \infty$ must satisfy $\int_{\mathbb{R}^4} e^{4u(y)} dy \le \frac{8\pi^2}{3}$, with equality iff u is of the form (1.4).

Lin also obtained a related result for the positive constant Q-curvature equation on $\mathbb{R}^n$, $n > 4$,

$$\begin{cases} \Delta^2 u(x) = u^{\frac{n+4}{n-4}}(x), & x \in \mathbb{R}^n; \\ u(x) > 0, & x \in \mathbb{R}^n. \end{cases} \tag{1.6}$$

His result implies that any solution to (1.6) must be rotationally symmetric with respect to some point $x_0 \in \mathbb{R}^n$, and of the form

$$u(x) = c_n \left(\frac{\lambda}{1+\lambda^2|x-x_0|^2}\right)^{\frac{n-4}{2}}, \tag{1.7}$$

for some $\lambda > 0$ and a dimensional constant c_n.

The proofs in both [CY97] and [L98] involve the method of moving planes. In [X00] X. Xu provided a proof for the rotational symmetry of solutions to (1.6) using the method of moving spheres.

Considering also that the canonical locally conformally flat metric on $\mathbb{S}^n \setminus \mathbb{S}^{n-k} \cong \mathbb{S}^{k-1} \times \mathbb{H}^{n-k+1}$, given via stereographic coordinates of $\mathbb{S}^n \setminus \mathbb{S}^{n-k}$ as

$$\begin{aligned} \frac{dx_1^2 + \cdots + dx_n^2}{x_1^2 + \cdots + x_k^2} &= \frac{d\rho^2 + \rho^2 d\omega^2_{\mathbb{S}^{k-1}} + dx_{k+1}^2 + \cdots + dx_n^2}{\rho^2} \\ &= d\omega^2_{\mathbb{S}^{k-1}} + \frac{d\rho^2 + dx_{k+1}^2 + \cdots + dx_n^2}{\rho^2} \end{aligned}$$

with $\rho^2 = x_1^2 + \cdots + x_k^2$, has its scalar curvature equal to $(n-1)(2k-n-2)$ and its Q-curvature equal to $8(2k-n)(2k-n-4)/n$, the sign of the Q-curvature alone is a poor indicator of how the metric behaves. Our first result, stated below, stems from this observation that it is natural to impose some additional condition involving the scalar curvature when considering global properties of solutions to the Q-curvature equation.

Theorem 1.1. *Let g be a conformal metric on $\Omega \subset \mathbb{S}^n$ such that*

$$Q_g \equiv 1 \ or \ \ 0 \quad in\ \Omega, \tag{1.8}$$

and

$$R_g \geq 0 \quad in\ \Omega. \tag{1.9}$$

(i) *If $\mathbb{S}^n \setminus \Omega$ contains more than one point and g is complete on Ω, then for any ball $B \subset\subset \Omega$ in the canonical metric $g_{\mathbb{S}^n}$, the mean curvature of its boundary ∂B in metric g with respect to its inner normal is positive;*

(ii) *If $\mathbb{S}^n \setminus \Omega$ is empty or consists of one point, then g is the round metric on $\mathbb{S}^n$ in the case $Q_g \equiv 1$; and the flat metric on $\mathbb{S}^n \setminus \{\infty\} \sim \mathbb{R}^n$ in the case $Q_g \equiv 0$ and $\mathbb{S}^n \setminus \Omega = \{\infty\}$.*

A corollary of Theorem 1.1 is the following

Corollary 1.2. *Suppose that $\Gamma \subset \mathbb{S}^l$ for $l \leq \frac{n-2}{2}$ and contains more than one point. Then any complete, conformal metric g on $\mathbb{S}^n \setminus \Gamma$ satisfying* (1.8) *and* (1.9) *has to be symmetric with respect to rotations of $\mathbb{S}^n$ which leave $\mathbb{S}^l$ invariant.*

A second corollary of Theorem 1.1 is the following

Corollary 1.3. *Suppose that $g = u(x)^{4/(n-4)} g_{\mathbb{S}^n}$ is a conformal metric on $\Omega \subsetneqq \mathbb{S}^n$ such that* (1.8) *and* (1.9) *hold, and that g is a complete metric on Ω or $u(x) \to \infty$ as $x \to \partial\Omega$, then there exists a constant $C > 0$ such that $u(x)^{2/(n-4)} \leq C\delta(x, \partial\Omega)^{-1}$, where $\delta(x, \partial\Omega)$ is the distance from x to $\partial\Omega$ in the metric $g_{\mathbb{S}^n}$.*

Remark 1.4. Corollary 1.2 can be considered as extending the consideration in [CY97, L98, X00] to cases where the solutions are not defined on $\mathbb{S}^n$ or $\mathbb{R}^n$, but on some more general $\Omega \subsetneqq \mathbb{S}^n$. Note that the corresponding (classification) results on $\mathbb{S}^n$ or $\mathbb{R}^n$ in [CY97, L98, X00] hold without assuming (1.9), but assuming (1.5) only when $n = 4$. We will see below – Remark 1.7 and the last 2 paragraphs for the $n = 4$ case in the proof of Theorem1.1 in the next section – that, in the $n = 4$ context, (1.5) is equivalent to (1.9).

In [Sc88] Schoen proved a version of Theorem 1.1 and Corollary 1.2 for a conformal, complete metric on $\Omega \subsetneqq \mathbb{S}^n$ with non-negative constant scalar curvature using a moving spheres argument, and proved a version of Corollary 1.3 in the same setting using a blow up argument. There have been many similar symmetry results on entire solutions or entire solutions with one point deleted to the constant σ_k curvature equation in the positive Γ_k class, which are generalizations of the Yamabe equation; a partial list of work in this direction includes those of Viaclovsky [V00a][V00b], Chang, Gursky and Yang [CGY02b][CGY03], Li and Li [LL03][LL05], Guan, Lin and Wang [GLW04], Li[L06].

Remark 1.5. In Theorem 1.1 we use the sign convention for the mean curvature as in [Sc88], namely, the mean curvature of the boundary of round Euclidean balls with respect to their inner normals is positive, as is the mean curvature of the boundary of round balls in $\mathbb{S}^n$ when they are confined to a hemisphere – note that as soon as a round ball in $\mathbb{S}^n$ contains a hemisphere, the mean curvature of its boundary with respect to its inner normal becomes negative in our convention.

Since umbilicity is invariant under a conformal change of metric, and round balls are umbilic in the canonical metric, our theorem implies that all principal curvatures of ∂B in metric g are positive.

Remark 1.6. The canonical locally conformally flat metric on $\mathbb{S}^n \setminus \mathbb{S}^{n-k} \cong \mathbb{S}^{k-1} \times \mathbb{H}^{n-k+1}$, for appropriate range of k, provides examples of metrics satisfying the assumptions in Theorem 1 with $\Gamma = \mathbb{S}^{n-k}$. The existence of conformal metrics satisfying the assumptions in Theorem 1 for more general Γ is an interesting question, but will not be addressed here.

In a local conformal representation for $g(x) = u(x)^{\frac{4}{n-4}}|dx|^2$ when $n \neq 4$, we have

$$(n-4)R_g(x) = -4(n-1)u(x)^{-\frac{n}{n-4}}\left(\Delta u(x) + \frac{2}{n-4}\frac{|\nabla u(x)|^2}{u(x)}\right). \tag{1.10}$$

We see that the condition (1.9) for $n > 4$ implies that

$$\Delta u(x) \leq 0. \tag{1.11}$$

The analog of (1.10) when $n = 4$ and $g = e^{2w}|dx|^2$ is

$$R_g(x) = -6e^{-2w(x)}\left(\Delta w(x) + |\nabla w(x)|^2\right). \tag{1.12}$$

We remark that in Y. Li's joint work [LL05] with A. Li on the study of entire solutions to a class of conformally invariant PDEs, and later on in his study of local behavior near isolated singularities of such solutions in [L06], condition (1.11) was used. One important ingredient in [LL05, L06] is that they work with the equations of u as well as of its classical Kelvin transforms with respect to the spheres $\partial B(x_0, R)$:

$$u_{x_0,R}(x) := \frac{R^{n-2}}{|x-x_0|^{n-2}}u\left(x_0 + \frac{R^2(x-x_0)}{|x-x_0|^2}\right),$$

and that $\Delta u(x) \leq 0$ for $|x - x_0| < (\geq)R$ is equivalent to $\Delta u_{x_0,R}(x) \leq 0$ for $|x - x_0| > (\leq)R$. This comes from computing the scalar curvature of a conformal metric g in the set up: $g(x) = u(x)^{\frac{4}{n-2}}|dx|^2$

$$R_g(x) = -4\frac{n-1}{n-2}u(x)^{-\frac{n+2}{n-2}}\Delta u(x). \tag{1.13}$$

This is different from (1.10). Note that under the inversion in $\partial B(x_0, R) : x \mapsto x_0 + \frac{R^2(x-x_0)}{|x-x_0|^2}$, the same metric g is represented as

$$\begin{aligned} g(x_0 + \frac{R^2(x-x_0)}{|x-x_0|^2}) &= u(x_0 + \frac{R^2(x-x_0)}{|x-x_0|^2})^{\frac{4}{n-2}}\left(\frac{R}{|x-x_0|}\right)^4|dx|^2 \\ &= u_{x_0,R}(x)^{\frac{4}{n-2}}|dx|^2, \end{aligned}$$

so

$$R_g\left(x_0+\frac{R^2(x-x_0)}{|x-x_0|^2}\right)=-4\frac{n-1}{n-2}u_{x_0,R}(x)^{-\frac{n+2}{n-2}}\Delta u_{x_0,R}(x).$$

In dealing with the Q-curvature equations, the metric g is often represented as $g(x)=u(x)^{\frac{4}{n-4}}|dx|^2$ (when $n\neq 4$), so the corresponding transformation under the inversion in $\partial B(x_0,R)$ is

$$u_{Q;x_0,R}(x)=\frac{R^{n-4}}{|x-x_0|^{n-4}}u\left(x_0+\frac{R^2(x-x_0)}{|x-x_0|^2}\right),$$

which gives

$$u(y)^{\frac{4}{n-4}}|dy|^2\Big|_{y=x_0+\frac{R^2(x-x_0)}{|x-x_0|^2}}=u_{Q;x_0,R}(x)^{\frac{4}{n-4}}|dx|^2.$$

In this set up, $u_{Q;x_0,R}(x)$ would satisfy (1.6) if $u(x)$ does, although $\Delta u(x_0+\frac{R^2(x-x_0)}{|x-x_0|^2})\leq 0$ is not equivalent to $\Delta u_{Q;x_0,R}(x)\leq 0$. However, condition (1.9) is a geometric condition and would imply (1.11) (when $n>4$) for any of its local representation in the form above, namely, (1.9) would imply $\Delta u(x_0+\frac{R^2(x-x_0)}{|x-x_0|^2})\leq 0$ as well as $\Delta u_{Q;x_0,R}(x)\leq 0$ – this is essential for our argument; (1.11) itself is not a geometrically invariant condition.

These discussions have their analogs in the $n=4$ case, where we write $g(x)=e^{2w(x)}|dx|^2$, and under the inversion in $\partial B(x_0,R)$, the metric g is represented as $g(x_0+\frac{R^2(x-x_0)}{|x-x_0|^2})=e^{2w_{Q;x_0,R}(x)}|dx|^2$, where

$$w_{Q;x_0,R}(x)=w\left(x_0+\frac{R^2(x-x_0)}{|x-x_0|^2}\right)+\ln\left(\frac{R^2}{|x-x_0|^2}\right).$$

Remark 1.7. Theorem 1 implies that $\mathbb{S}^n\setminus\Omega$ cannot be a single point for the case of $Q_g=1$ and g complete on Ω. For, if that were the case, we could place that single point at ∞ so as to obtain an entire solution on $\mathbb{R}^n$, and the conclusion in (i) of Theorem 1 would imply that, at any $x\in\mathbb{R}^n$, for any unit vector e in $\mathbb{R}^n$, and for any $r>0$, $\nabla_e u(x)+\frac{n-4}{2r}u(x)\geq 0$ for the $n>4$ case – see the set up in the next section, and this would imply $\nabla_e u(x)=0$, which would imply that u is a constant in $\mathbb{R}^n$, but constants are not solutions to (1.2) or (1.6).

In [CY97, L98, X00] cited above, one technical step is to establish that entire solutions u to (1.2) under their respective assumptions and to (1.6) are superharmonic on $\mathbb{R}^n$. In the $n>4$ case these authors proved that all entire solutions to (1.6) are of the standard form (1.7). But in the $n=4$ case the superharmonicity of u is not enough to lead to the same classification; condition (1.5) is needed. In fact, Theorem 1.2 in [L98] Lin gives some properties of entire solutions to (1.2) for $n=4$ satisfying $\int_{\mathbb{R}^4}e^{4u(y)}dy<\frac{8\pi^2}{3}$; and in [CC01] Chang and Chen constructed such solutions. These solutions are not of the form (1.7), and from the perspective of Theorem 1, the conclusions of Theorem 1 does not hold on all Euclidean spheres for these solutions. This means that condition (1.9) is related to, but different from, the superharmonicity of a particular representation of the metric, that it is needed for the conclusions of Theorem 1, and that these solutions do not satisfy (1.9). These properties can be verified directly on large Euclidean spheres using the expansion (1.10) and (1.11) of [L98].

Remark 1.8. Corollary 1.3 follows from Theorem 1.1 as Schoen did in [Sc88] in the case of constant positive scalar curvature equation. The outline of the argument goes as follows. If the upper bound does not hold, then a sequence of *rescaled solutions* centered along a sequence of points approaching $\partial\Omega$ would converge to an entire solution on $\mathbb{R}^n$ to the same equation. The solutions of the latter equation are completely classified; they correspond to the round metric on $\mathbb{S}^n$, therefore their mean curvatures (and principal curvatures) along large Euclidean spheres become negative. But in the closure of any such a large Euclidean ball, this metric is the uniform limit of a sequence of metrics whose principal curvatures along its boundary sphere is positive by the version of Theorem 1.1. This would cause a contradiction; therefore Corollary 1.3 must hold. We will not supply a detailed proof of Corollary 1.3 here.

Remark 1.9. Although Schoen's result in [Sc88] corresponding to Theorem 1.1 was stated and proved for a constant *positive* scalar curvature metric on a domain $\Omega \subsetneqq \mathbb{S}^n$, an examination of the proof indicates that, as long as the three main ingredients for the moving plane/sphere arguments are valid, the same conclusion can be drawn, namely, the same conclusion as given in Theorem 1 continues to hold if the following three steps are still valid: (i). the initiation of the inequality between a solution in a half-space/ball enclosing its singular set and its reflected solution; (ii). the above inequality is a strict point wise inequality unless it becomes a point wise equality in the entire comparison domain; and (iii). the strict inequality continues to hold if the half-space/ball is moved in a small open neighborhood.

Both (i) and (iii) involve proving that the solution in a neighborhood of its singular set stays above its reflected solution (which is a smooth solution to the equation near the singular set) by a positive amount – we did this here by using the maximum principle for superharmonic functions in a domain with a boundary component having zero Newtonian capacity, without imposing an explicit growth condition of the solution toward its singular set.

Both (ii) and (iii) involve using the strong maximum principle and the Hopf boundary lemma for the difference between the solution and its reflected solution, when this difference is assumed to be non-negative. But this part works for solutions to the constant scalar curvature equation, even if the constant is non-positive; in fact, it works even for the constant σ_k curvature equation, as long as the equation is elliptic. (i) and (iii) can be established if we assume that the conformal factor tends to ∞ uniformly upon approaching the boundary of its domain. We thus have

Theorem 1.10. *Let g be a conformal metric on $\Omega \subsetneqq \mathbb{S}^n$ such that* (a) $\sigma_k(A_g) =$ *a constant in Ω,* (b) $A_g \in \Gamma_k^+$ (*or Γ_k^- respectively*) *pointwise in Ω, and* (c) *if we write $g = e^{2w} g_{\mathbb{S}^n}$, then $w \to \infty$ uniformly upon approaching $\partial\Omega$. Then for any ball*

$B \subset\subset \Omega$ in the canonical metric $g_{\mathbb{S}^n}$, the mean curvature of its boundary ∂B in metric g with respect to its inner normal is positive.

Theorem 1.10 essentially appeared in earlier work, maybe not in such explicit formulation, see, for example, estimate (27) in [LL05]. Some computational sketches will be provided in the next section to illustrate the implementation of the argument outlined in the previous remark.

Remark 1.11. Our background discussion mentioned results which construct conformal metrics satisfying the assumptions in Theorem 1.10 for the $k = 1$ case. The construction of such conformal metrics for $k > 1$ and more general Ω (subject to the constraints on the dimension of $\mathbb{S}^n \setminus \Omega$) is an interesting question. In a recent work [GLN18] González, Li, and Nguyen construct viscosity solutions to a class of conformally invariant equations, which include the equations in (a) of Theorem 1.10, such that these solutions are in Γ_k^- when they solve the equations in (a), and these solutions satisfy (c) – under appropriate dimensional constraints on $\partial\Omega$. Maximum principle, the key tool in proving Theorems 1.1 and 1.10, is valid for viscosity solutions, see [LNW18]; so Theorems 1.10 applies to solutions in [GLN18].

Our next result provides a criterion for the development map of a locally conformally flat manifold to be an embedding in the negative Yamabe constant case.

Theorem 1.12. *Let (M, g) be a complete, locally conformally flat manifold, and $F : (M, g) \mapsto (\mathbb{S}^n, g_{\mathbb{S}^n})$ be a conformal immersion. If the Schouten tensor A_g of some metric in the conformal class of g is non-positive point wise on M, then F is an imbedding.*

Based on the following algebraic property that $\sigma_1(A_g) \leq 0$ and

$$\sigma_2(A_g) \geq \frac{(n-2)}{2(n-1)}(\sigma_1(A_g))^2 \tag{1.14}$$

imply $A_g \leq 0$, we have

Corollary 1.13. *If (M, g) is a complete, locally conformally flat manifold, and satisfies $\sigma_1(A_g) \leq 0$ and (1.14), and $F : (M, g) \mapsto (\mathbb{S}^n, g_{\mathbb{S}^n})$ is a conformal immersion, then F is an imbedding.*

Remark 1.14. Based on the following relation between the Schouten tensor A_g and the Einstein tensor E_g

$$A_g = \frac{E_g}{n-2} + \frac{R_g}{2n(n-1)}g = \frac{E_g}{n-2} + \frac{\sigma_1(A_g)}{n}g, \tag{1.15}$$

we have

$$2\sigma_2(A_g) = \frac{n-1}{n}(\sigma_1(A_g))^2 - \frac{||E_g||^2}{(n-2)^2}, \tag{1.16}$$

where $||E_g||$ is the metric norm of E with respect to g.

It then follows that condition (1.14) is a kind of pinching condition, as it is equivalent to

$$\frac{||E_g||^2}{(n-2)^2} \le \frac{(\sigma_1(A_g))^2}{(n-1)n}, \tag{1.17}$$

using (1.16). (1.14) is also equivalent to

$$(n-1)||A_g||^2 \le (\sigma_1(A_g))^2 \tag{1.18}$$

Remark 1.15. Theorem 1.12 and its corollary were obtained in the early 2000's, and were lectured by the second author in several seminar talks, including the fall 2003 CUNY Graduate Center Differential Geometry and Analysis Seminar.

When the condition $A_g \le 0$ is not satisfied, F may not be an embedding, as shown by the canonical locally conformally flat metric on $\mathbb{S}^1_r \times \mathbb{H}^{n-1}$, whose Schouten tensor is $\mathrm{diag}(\frac{1}{2}, -\frac{1}{2}, \ldots, -\frac{1}{2})$.

We will provide a proof of Theorem 1.1 and Corollary 1.2 in Section 2. The proof of Theorem 1.1 will follow the outlines as was done in [CY97], [L98], and [X00]. We will only sketch the main steps of the proof, in particular, indicate how to handle the behavior of the solution near $\partial\Omega$. Some computational sketches will also be provided for a proof of Theorem 1.10.

We will provide a proof of Theorem 1.12 in Section 3.

2. Proof of Theorems 1.1 and 1.10

Proof of Theorem 1.1. We first set up a stereographic coordinate for proving Theorem 1.1. Let $B \subset\subset \Omega$ be as given in Theorem 1.1. We can choose a stereographic coordinate such that B is mapped onto $\{x \in \mathbb{R}^n : x_1 < \lambda_0\}$ – this amounts to choosing coordinate such that the north pole lies on ∂B, and is equivalent to working with an appropriately transformed $u_{Q;x_0,R}$ in place of u. Define $\Sigma_\lambda = \{x \in \mathbb{R}^n : x_1 > \lambda\}$, and $T_\lambda = \partial(\Sigma_\lambda) = \{x \in \mathbb{R}^n : x_1 = \lambda\}$. Let Γ be the image of $\mathbb{S}^n \setminus \Omega$ under this stereographic map. Then Γ is a compact subset in Σ_{λ_0}. Define $\Sigma'_\lambda = \Sigma_\lambda \setminus \Gamma$. In this stereographic coordinate we can write

$$g(x) = u(x)^{\frac{4}{n-4}}|dx|^2 \quad \text{for } x \in \mathbb{R}^n \setminus \Gamma. \tag{2.1}$$

Here, we first provide the details for *the* $n > 4$ *case*; the modifications needed for the $n = 3, 4$ cases will be sketched at the end.

The statement that the mean curvature of ∂B in metric g with respect to its inner normal is positive in the $n > 4$ case is equivalent to

$$u_{x_1}(x) > 0 \text{ for all } x \in T_{\lambda_0}. \tag{2.2}$$

Remark 2.1. If one represents B by a Euclidean ball $B(x_0, r)$ with x_0 as center and $r > 0$ as radius, then the statement that the mean curvature of ∂B at $x \in \partial B$ in metric g with respect to its inner normal is positive (when $n > 4$) is equivalent to

$$\nabla_\theta u(x) + \frac{n-4}{2r}u(x) := \frac{\partial u(x_0 + r\theta)}{\partial r} + \frac{n-4}{2r}u(x) > 0 \quad \text{for } x = x_0 + r\theta, \theta \in \mathbb{S}^{n-1}, \tag{2.3}$$

as the mean curvature in metric $g = u(x)^{4/(n-4)}|dx|^2$ at a point x on ∂B is given by

$$\frac{2u(x)^{\frac{2-n}{n-4}}}{n-4}\left[\nabla_\theta u(x) + \frac{n-4}{2r}u(x)\right].$$

It follows from this set up that (2.3) implies gradient estimate for u, for, if $\nabla u(x) = |\nabla u(x)|e$ for some $e \in \mathbb{S}^{n-1}$, then, with $x_0 = x + \frac{\delta(x,\Gamma)}{2}e$, $B_{\frac{\delta(x,\Gamma)}{2}}(x_0) \subset \Omega$, and $x \in \partial B_{\frac{\delta(x,\Gamma)}{2}}(x_0)$, thus (2.3) at x implies that $|\nabla u(x)| = -\nabla_{-e}u(x) < \frac{n-4}{\delta(x,\Gamma)}u(x)$. Y. Li and his collaborators also used estimates like (2.3) (see, for example [LL05] and [LN14]), or rather an inequality of the form $u(y) \geq u_{x_0,r}(y)$ (or $u(y) \geq u_{Q;x_0,r}(y)$ in our setting, which is used in deriving (2.3)), to obtain gradient estimates.

Remark 2.2. It follows from Theorem 2.7 in [SY88] that, in the situation of our Theorem 1.1, the Newtonian capacity $\operatorname{cap}(\mathbb{S}^n \setminus \Omega) = 0$, which implies that $\operatorname{cap}(\Gamma) = 0$. We will use this to deal with the behavior of $u(x)$ and that of $\Delta u(x)$ near Γ.

Let $v(x) = -\Delta u(x)$. Then based on our set up, we have

$$\begin{cases} \Delta v(x) = -u^{\frac{n+4}{n-4}}(x) \leq 0 & \text{in } \mathbb{R}^n \setminus \Gamma, \\ \quad v(x) \geq 0 & \text{in } \mathbb{R}^n \setminus \Gamma, \end{cases} \tag{2.4}$$

(the $Q \equiv 0$ case can be handled by a straightforward modification) and that $u(x)$ has an expansion at $x = \infty$:

$$\begin{cases} u(x) = c_1|x|^{4-n} + \sum_{j=1}^n \frac{b_j x_j}{|x|^{n-2}} + O\left(\frac{1}{|x|^{n-2}}\right) \\ u_{x_i} = -(n-2)c_1 x_i |x|^{2-n} + O\left(\frac{1}{|x|^{n-2}}\right) \\ u_{x_i x_j}(x) = O\left(\frac{1}{|x|^{n-2}}\right) \end{cases} \tag{2.5}$$

for some constants $c_1 > 0$ and b_j's. It follows from this expansion that $v(x) = -\Delta u(x)$ has the following expansion at $x = \infty$:

$$\begin{cases} v(x) = c_0|x|^{2-n} + \sum_{j=1}^n \frac{a_j x_j}{|x|^{n}} + O\left(\frac{1}{|x|^{n}}\right) \\ v_{x_i} = -(n-2)c_0 x_i |x|^{-n} + O\left(\frac{1}{|x|^{n}}\right) \\ v_{x_i x_j}(x) = O\left(\frac{1}{|x|^{n}}\right) \end{cases} \tag{2.6}$$

for some constants $c_0 > 0$ and a_j's.

Set $x^\lambda = (2\lambda - x_1, x_2, \ldots, x_n)$, which is the reflection of x with respect to T_λ, and

$$w_\lambda(x) = u(x) - u(x^\lambda) \quad \text{for } x \in \Sigma_\lambda^{'}.$$

We will prove using the moving planes method that, when g cannot be extended as a smooth metric across Γ,

$$u(x) - u(x^\lambda) > 0 \text{ and } v(x) - v(x^\lambda) > 0, \text{ for all } x \in \Sigma_\lambda^{'} \text{ and } \lambda \leq \lambda_0. \tag{2.7}$$

It would then follow from (2.7) that

$$u_{x_1}(x) \geq 0 \text{ and } \partial_{x_1}(\Delta u(x)) \leq 0, \text{ for any } x \text{ with } x_1 \leq \lambda_0, \tag{2.8}$$

which, together with the strong maximum principle applied to $u(x) - u(x^\lambda)$ and $v(x) - v(x^\lambda)$, would conclude our proof.

In our setting it is impossible for $v(x) \equiv 0$ on $\mathbb{R}^n \setminus \Gamma$ due to (2.6). Then it follows from (2.4) and the strong maximum principle that $v(x) > 0$ in $\mathbb{R}^n \setminus \Gamma$.

We may suppose that $\Gamma \subset B(0, R_0)$ for some $R_0 > 0$. Now for any $R \geq R_0$, since $v > 0$ in $\overline{B(0,R)} \setminus \Gamma$, there exists $\delta > 0$ depending on R such that $v(x) \geq \delta$ for all $x \in \partial B(0, R)$. It now follows, using $\mathrm{cap}(\Gamma) = 0$ and (2.4), that

$$v(x) \geq \delta \quad \text{ for all } x \in B(0, R) \setminus \Gamma. \tag{2.9}$$

A reference for this kind of extended maximum principle is [L72, Chap. III, Thm. 3.4]. A formulation of this kind extended maximum principle in our setting is

Lemma 2.3. *Suppose that* (i) Ω *is a bounded domain in* $\mathbb{R}^n$ *and that* $\Gamma \subset \Omega$ *has capacity* 0, (ii) v *is superharmonic in* $\Omega \setminus \Gamma$, *and* (iii) v *is bounded below in* $\Omega \setminus \Gamma$, *and there exists* M *such that for any* $z \in \partial\Omega$, $\liminf_{x \in \Omega, x \to z} v(x) \geq M$. *Then* $v(x) \geq M$ *in* $\Omega \setminus \Gamma$.

The expansion (2.6) of $v(x)$ at ∞ and Lemma 2.3 in [CGS89] implies that

$$\begin{cases} \text{there exists } \lambda_1 \leq \lambda_0 \text{ and } R_1 \geq R_0 \text{ such that} \\ v(x) > v(x^\lambda) \quad \text{for all } x \in \Sigma_\lambda^{'} \text{ with } |x| \geq R_1, \text{ and } \lambda \leq \lambda_1. \end{cases} \tag{2.10}$$

Then using (2.9) and the expansion (2.6) of $v(x)$ at ∞, we conclude that there exists $\lambda_2 \leq \lambda_1$ such that

$$v(x) > v(x^\lambda) \quad \text{for all } x \in \Sigma_\lambda^{'},\ \lambda \leq \lambda_2. \tag{2.11}$$

Next, $w_\lambda(x)$ satisfies

$$\Delta w_\lambda(x) = v(x^\lambda) - v(x) \leq 0 \quad \text{for all } x \in \Sigma_\lambda^{'}, \tag{2.12}$$

and $\lambda \leq \lambda_2$. The expansion (2.5) of $u(x)$ at ∞ implies that

$$w_\lambda(x) \to 0 \quad \text{ as } x \to \infty. \tag{2.13}$$

Using (2.12), (2.13), $w_\lambda(x) = 0$ for all $x \in T_\lambda$, and the observation that $w_\lambda(x) = u(x) - u(x^\lambda) \geq -u(x^\lambda)$ is bounded below in a neighborhood of Γ and the information that $\mathrm{cap}(\Gamma) = 0$, we conclude that $w_\lambda(x) \geq 0$ for all $x \in \Sigma_\lambda^{'}$, $\lambda \leq \lambda_2$.

In the situation of (i), the completeness assumption on g and $\Omega \neq \mathbb{S}^n$ imply that $w_\lambda(x)$ cannot be $\equiv 0$, so with the strong maximum principle, we conclude that

$$w_\lambda(x) > 0 \quad \text{for all } x \in \Sigma_\lambda^{'}, \tag{2.14}$$

and $\lambda \leq \lambda_2$.

We now define

$$\lambda_* = \sup\{\lambda \le \lambda_0 : v(x^\mu) < v(x) \quad \text{for all } x \in \Sigma_\mu^{'}, \text{ and all } \mu \le \lambda,\}$$

and proceed to prove that $\lambda_* = \lambda_0$.

By continuity (together with strong maximum principle and completeness of g), (2.12) and (2.14) continue to hold for λ_* replacing λ. We now have, using (2.14) for λ_* replacing λ, that

$$\Delta\left[v(x^{\lambda_*}) - v(x)\right] = u^{\frac{n+4}{n-4}}(x) - u^{\frac{n+4}{n-4}}(x^{\lambda_*}) \ge 0 \quad \text{for all } x \in \Sigma_{\lambda_*}^{'}. \tag{2.15}$$

$v(x^{\lambda_*}) - v(x) \le 0$ for all $x \in \Sigma_{\lambda_*}^{'}$. Now strong maximum principle, (2.14) and (2.15) imply that $v(x^{\lambda_*}) - v(x) < 0$ for all $x \in \Sigma_{\lambda_*}^{'}$ – the $Q \equiv 0$ case would need a modified argument to rule out $v(x^{\lambda_*}) - v(x) \equiv 0$ using the Liouville theorem on the harmonic function $v(x)$ and (2.6). Furthermore, using $\mathrm{cap}(\Gamma) = 0$, there exists some $\delta_* > 0$ such that

$$v(x^{\lambda_*}) - v(x) \le -\delta_* \quad \text{for } x \text{ in a neighborhood of } \Gamma.$$

This, together with (2.15) and Lemma 2.4 in [CGS89], implies that $\lambda_* = \lambda_0$, and concludes the case for (i).

In the $\Omega = \mathbb{S}^n$ subclass of (ii), the set up in the proof of (i) is used to prove, in a more standard fashion as in [CY97], that $u(x)$ is rotationally symmetric about some point; then in the $Q \equiv 1$ case the argument in [L98] proves that $u(x)$ is of the standard form; while in the $Q \equiv 0$ case standard properties on entire positive harmonic functions implies that $u(x)$ must be a positive constant, but the associated metric would not be a smooth metric over $\Omega = \mathbb{S}^n$, so this latter case cannot occur.

In the remaining case of (ii): $\Omega = \mathbb{S}^n \setminus \{\text{a point}\}$, the set up in the proof of (i) works identically, and proves that the solution is rotationally symmetric about the image point of ∞ under the inversion used in the set up. But the sphere with respect to which the inversion is done can be chosen arbitrarily, so the solution is shown to be rotationally symmetric about any point, therefore is a positive constant. This cannot happen in the $Q \equiv 1$ case, and in the $Q \equiv 0$ case leads to the conclusion that the metric is the flat one on $\Omega = \mathbb{S}^n \setminus \{\text{a point}\}$.

We now indicate the modifications needed for the $n = 3$ case. (2.2) turns into

$$u_{x_1} < 0 \quad \text{for all } x \in T_{\lambda_0}; \tag{2.16}$$

(2.3) turns into

$$\nabla_\theta u(x) - \frac{u(x)}{2r} < 0; \tag{2.17}$$

The condition $R_g \ge 0$ turns into

$$\Delta u(x) - \frac{2|\nabla u(x)|^2}{u(x)} \ge 0; \tag{2.18}$$

and the three-dimensional version of (1.6) for $Q = 2$ is

$$(-\Delta)^2 u = -u^{-7}, \quad x \in \Omega \subset \mathbb{R}^3. \tag{2.19}$$

Setting $\tilde{v}(x) = \Delta u(x)$, we find that under (2.18), $\tilde{v}(x) \geq 0$; and $\eta(x) := \tilde{v}(x) - \tilde{v}(x^\lambda)$ satisfies $\eta(x) \geq -\Delta u(x^\lambda)$, as well as $\Delta\eta(x) \leq 0$ whenever $u(x) \leq u(x^\lambda)$. The version of (2.7) that we need to establish in 3 dimension is

$$u(x) - u(x^\lambda) < 0 \text{ and } \eta(x) = \tilde{v}(x) - \tilde{v}(x^\lambda) > 0 \quad \text{for all } x \in \Sigma_\lambda' \text{ and } \lambda \leq \lambda_0. \quad (2.20)$$

Given (2.19) and the information on η above, (2.20) is established in almost identical way as in the $n > 4$ case.

For the $n = 4$ case, we use $g(x) = e^{2w(x)}|dx|^2$; (2.18) is replaced by $\Delta w(x) + |\nabla w(x)|^2 \leq 0$; (2.3) is replaced by $\frac{\partial w}{\partial r} + \frac{1}{r} \geq 0$; in place of (2.5), we have a similar expansion for $e^{w(x)}$ at ∞ (in an appropriately chosen stereographic coordinate) whose leading order term is $2|x|^{-2}$, or equivalently, an expansion for $w(x)$ whose leading order term is $-2\ln|x|$ – the expansions for $w_{x_j}(x)$ and $\Delta w(x)$ come as consequences of the expansion for $w(x)$. Our objective in this set up is still to establish (2.7) with $w(x)$ replacing $u(x)$ there. (2.9) is established in the same way for $v(x) = -\Delta w(x)$, as we still have $\Delta w(x) \leq 0$ (< 0 in fact) based on $R_g \geq 0$, and $\Delta v(x) \leq 0$ in $\mathbb{R}^4 \setminus \Gamma$. The analog of (2.14) we need is $w(x) - w(x^\lambda) > 0$ for $x \in \Sigma_\lambda'$ and all $\lambda \leq \lambda_0$, and one key ingredient in proving this is a lower bound for $w(x)$ in a neighborhood of Γ. This is done using

$$\Delta e^{w(x)} = e^{w(x)}\left[\Delta w(x) + |\nabla w(x)|^2\right] \leq 0 \quad \text{for } x \in \mathbb{R}^4 \setminus \Gamma,$$

from which it follows from the extended maximum principle applied to $e^{w(x)}$ over $B_R \setminus \Gamma$ for a fixed, sufficiently large $R > 0$ that $e^{w(x)}$ has a positive lower bound in $B_R \setminus \Gamma$, which then implies a lower bound for $w(x)$ in $B_R \setminus \Gamma$. These modifications suffice to complete a proof for the $n = 4$ case.

For the $n = 4$ and $\Gamma = \{\text{one point}\}$ case, we can arrange coordinates such that $\Gamma = \{0\}$. The argument in the above paragraph applies, except that it is possible that $w(x) - w(x^\lambda) \equiv 0$ for some λ – in fact, this will always happen. Then it's easy to see that g must be the round metric, and as a consequence, (1.5) holds. If (1.5) is assumed in place of (1.9), then it follows from [L98] that g must be the round metric, and as a consequence, (1.9) holds. Thus in this context, (1.9) and (1.5) are equivalent. □

Remark 2.4. An examination of the proof shows three crucial ingredients for completing the proof for Theorem 1:

(i) (2.9) for initiating of the relation that $v(x) - v(x^\lambda) = -\Delta u(x) + \Delta u(x^\lambda) \geq 0$ for $x \in \Sigma_\lambda'$ and $\lambda \leq \lambda_1$ for $|\lambda_1|$ large (for the $n > 4$ cases; the $n = 3, 4$ cases can be formulated appropriately);
(ii) $u(x) - u(x^\lambda) \equiv 0$ cannot happen in Σ_λ'; and
(iii) once $u(x) - u(x^\lambda) \geq 0$ in Σ_λ' is established, there exists $\delta > 0$ such that

$$u(x) - u(x^\lambda) > \delta \text{ and } -\Delta u(x) + \Delta u(x^\lambda) \geq \delta \text{ in a neighborhood of } \Gamma.$$

(2.9) is proved using the equation for $v(x) = -\Delta u(x)$, the property $v(x) \geq 0$, which follows from $R_g \geq 0$, and $cap(\Gamma) = 0$; (iii) also relies on $R_g \geq 0$ and $cap(\Gamma) = 0$ crucially; while (ii) relies on the assumption that g is a complete metric on $\Omega = \mathbb{S}^n \setminus \Gamma$ – this assumption, together with $R_g \geq 0$ (and locally conformal flatness of g), implies $cap(\Gamma) = 0$, based on [SY88]. Based on this examination, the assumption in Theorem 1 that g is a complete metric on $\mathbb{S}^n \setminus \Gamma$ can be replaced by the assumption that g cannot be extended as a smooth metric over Γ and that $cap(\Gamma) = 0$.

Remark 2.5. Here is another illustration why the assumption that $R_g \geq 0$ cannot be dropped: in the case of $Q \equiv 0$ on $\mathbb{S}^n \setminus \{\text{a point}\}$, which we identify as $\mathbb{R}^n$, we would be studying positive solutions $u(x)$ on $\mathbb{R}^n$ to $\Delta^2 u(x) = 0$. A simple argument using the Green's formula to $\Delta u(x)$:

$$\begin{aligned}\Delta u(x) &= \frac{1}{|B_r(x)|}\int_{B_r(x)} \Delta u(y)dy \\ &= \frac{1}{|B_r(x)|}\int_{\partial B_r(x)} \frac{\partial u(y)}{\partial \nu(y)} d\sigma(y) = \frac{n}{r}\partial_r \left(\int_{\mathbb{S}^{n-1}} u(x+r\omega)d\omega\right),\end{aligned}$$

and $u > 0$ on $\mathbb{R}^n$ shows that $\Delta u(x) \geq 0$. This then makes $\Delta u(x)$ a non-negative entire harmonic function, so $\Delta u(x) = c$ for some non-negative constant c. $u(x) = u_0 + \sum_{j=1}^n a_j x_j^2$, for $u_0 > 0$ and appropriately chosen $a_j \geq 0$, are positive solutions. These solutions have reflection symmetries, but do not have rotational symmetry unless a_j's are all equal; and, in any case, do not satisfy the conclusions of Theorem 1.1 unless a_j's are all 0.

Unless $c = 0$, these solutions do not correspond to metrics with $R_g \geq 0$. If we were to follow the set up in the proof of (i) of Theorem 1.1, we would work with

$$u_{Q;0,1}(x) = |x|^{4-n} u\left(\frac{x}{|x|^2}\right) = u_0 |x|^{4-n} + |x|^{-n} \sum_{j=1}^n a_j x_j^2.$$

But $\Delta u_{Q;0,1}(x)$ may become unbounded near $x = 0$ when $a_j \neq a_k$ for some $j \neq k$. This would prevent an estimate like (2.9) for $v(x) := -\Delta u_{Q;0,1}(x)$, which is needed for the initiation of step (i) alluded to in the previous remark.

Proof of Corollary 1.2. It suffices to prove that, when $\mathbb{S}^l \setminus \{\infty\}$ is represented via a stereographic projection as $\mathbb{R}^l = \{x \in \mathbb{R}^n : x_{l+1} = \cdots = x_n = 0\}$, and for any $x \in \mathbb{R}^n \setminus \mathbb{R}^l$, and for any (unit) vector $e = (0, \ldots, 0, e_{l+1}, \ldots, e_n) \perp x$, we have $\nabla_e u(x) = 0$ – this set up would require $n - l \geq 2$, which we have from $l \leq (n-2)/2$. This would imply that, in this set up, $u = u(x_1, \ldots, x_n)$ depends on $x_{l+1}, \ldots, x_n$ only through $\sqrt{x_{l+1}^2 + \cdots + x_n^2}$.

For any $r > 0$, we see that $B(x - re, r) \subset\subset \mathbb{R}^n \setminus \mathbb{R}^l$, so the conclusion of Theorem 1 is valid on $\partial B(x - re, r)$. In particular, for the $n > 4$ case and at $x \in \partial B(x - re, r)$, we have, by (2.3)

$$\nabla_e u(x) + \frac{n-4}{2r} u(x) > 0. \tag{2.21}$$

Since we can take $r > 0$ arbitrarily large, we conclude that $\nabla_e u(x) \geq 0$. Repeating this argument with $-e$ replacing e, we obtain $\nabla_{-e} u(x) \geq 0$, and therefore conclude that $\nabla_e u(x) = 0$. The $n = 3, 4$ cases need only minor modifications. $\square$

Sketch of proof of Theorem 1.10. Here we will express the metric g in the form of $e^{2w(x)}|dx|^2$, and express the equation in the form of $f(\lambda(A[w])) = 1$, where $A[w] = -\nabla^2 w + \nabla w \otimes \nabla w - \frac{|\nabla w|^2}{2} I$ denotes the matrix representing the Schouten tensor, and $\lambda(A[w])$ refers to the eigenvalues of $A[w]$. Again we have set up coordinates such that $\Gamma \subset \mathbb{R}^n$, and that $w(x)$ has an expansion at ∞ in the spirit of (2.5), but with $-2\ln|x|$ as the leading order term. $w(x^\lambda)$ satisfies the same equation.

To initiate the moving plane method, we need a positive lower bound for $e^{w(x)}$ near Γ. This is provided for by assumption (c). Then traditional method is used to establish $w(x) - w(x^\lambda) \geq 0$ for $x \in \Sigma'_\lambda$ for $\lambda \leq \lambda_1$ for some large $|\lambda_1|$. To carry through the moving plane method, namely, to establish the above inequality for all expected range of λ, we use the equations for $w(x)$ and $w(x^\lambda)$ to obtain a linear, second-order, elliptic equation for $w(x) - w(x^\lambda)$: $L[w(x) - w(x^\lambda)] = 0$ in Σ'_λ, thanks to assumption (b). Using assumption (c), $w(x) - w(x^\lambda) \geq 0$, and the strong maximum principle, we obtain $w(x) - w(x^\lambda) > 0$ – the version used here is for non-negative solutions, which can be derived from Lemma 3.4 in [GT], and does not require a condition on the sign of the coefficient of the zeroth-order term in L; an explicitly formulated version for such a setting appears, e.g., as Lemma 3.5 in [CY97]; it is for this reason that the argument for Theorem 1.10 does not distinguish between the solutions in Γ_k^+ class from those in the Γ_k^- class. Assumption (c) further implies that there exists $\delta > 0$ such that $w(x) - w(x^\lambda) > \delta$ in a neighborhood of Γ. This, the Hopf Lemma and Lemma 2.4 in [CGS89], imply that $w(x) - w(x^\lambda) > 0$ holds for all expected range of λ. $\square$

3. Proof of Theorem 1.12

Proof of Theorem 1.12 . We may assume that g itself satisfies that its Schouten tensor A_g is non-positive point wise on M. For any point $z_0 \in F(M) \subset \mathbb{S}^n$, using stereographic coordinates, there is a smooth function u on $\mathbb{S}^n$ such that $u > 0$ on $\mathbb{S}^n \setminus \{z_0\}$, $u(z_0) = 0$, and $u^{-2} g_0$ is flat. Writing $F^*(u^{-2} g_0) = v^{-2} g \overset{\text{def}}{=} \widehat{g}$ on $M \setminus F^{-1}(z_0)$, then $\widehat{g}$ is flat. Hence, on $M \setminus F^{-1}(z_0)$, $\widehat{E} = 0$, $\widehat{R} = 0$.

Under a (pointwise) conformal change of the metric g, $\widehat{g} = v^{-2} g$, the Einstein tensor and scalar curvature transform as follows.

$$\widehat{E} = E + \frac{n-2}{v}\left\{\nabla^2 v - \frac{\Delta v}{n} g\right\}, \tag{3.1}$$

$$\widehat{R} = v^2 \left\{R + 2(n-1)\frac{\Delta v}{v} - n(n-1)\frac{|\nabla v|^2}{v^2}\right\}. \tag{3.2}$$

Thus in the situation here, we have, by (3.1) and (3.2),

$$E = -\frac{n-2}{v}\left\{\nabla^2 v - \frac{\Delta v}{n} g\right\}, \tag{3.3}$$

$$R = -(n-1)\left\{2\frac{\Delta v}{v} - n\frac{|\nabla v|^2}{v^2}\right\}. \tag{3.4}$$

It now follows that

$$A = -\frac{\nabla^2 v}{v} + \frac{|\nabla v|^2}{2v^2} g. \tag{3.5}$$

Under our assumption that $A \leq 0$, we therefore have

$$\nabla^2 v \geq \frac{|\nabla v|^2}{2v} g, \tag{3.6}$$

on $M \setminus F^{-1}(z_0)$. By a limiting argument, $v(\gamma(s))$ is a non-negative convex function along any geodesic (in metric g) $\gamma(s)$ on M.

If $P_0 \neq P_1 \in M$ are such that $F(P_0) = F(P_1)$, we set $z_0 = F(P_0) = F(P_1)$ and carry out the computations in the paragraph above. Since (M, g) is assumed to be complete, we may joint P_0 and P_1 by a geodesic (in metric g) $\gamma(s)$ parametrized over $s \in [0, 1]$ with $\gamma(0) = P_0$ and $\gamma(1) = P_1$, then $v(\gamma(0)) = v(\gamma(1)) = 0$. Since $v > 0$ on $M \setminus F^{-1}(z_0)$, this would imply that $\gamma(s) \in F^{-1}(z_0)$ for all $s \in [0, 1]$, using the convexity of v. But this is not possible, and this contradiction implies that F must be an imbedding. □

Proof of Corollary 1.13. We just need to establish the algebraic property that $\sigma_1(A_g) \leq 0$ and (1.14) imply $A_g \leq 0$. Since E is trace free, we have the sharp inequality

$$-\sqrt{\frac{n-1}{n}}||E_g||g \leq E \leq \sqrt{\frac{n-1}{n}}||E_g||g. \tag{3.7}$$

Thus, when (1.14) holds, we have (1.17), as remarked earlier, which then implies that

$$\frac{||E_g||}{(n-2)} \leq \frac{|\sigma_1(A_g)|}{\sqrt{n(n-1)}}. \tag{3.8}$$

It now follows from (3.8), (1.15) and (3.7) that $A_g \leq 0$. □

Note added to galley proof

After our submission was accepted in 2018 for publication in this volume, we became aware of the paper "Asymptotic symmetry and local behavior of solutions of higher order conformally invariant equations with isolated singularities" by Tianling Jin and Jingang Xiong (https://arxiv.org/abs/1901.01678), which, among other results, proves sharp blow up rates of solutions of higher order conformally invariant equations in a bounded domain with an isolated singularity, and the asymptotic radial symmetry of the solutions near the singularity. It also contains some recent relevant references.

Acknowledgement

Chang's work was partially supported by NSF through grant DMS-1509505. Part of Han's work was completed in summer 2015 when the author was on a visiting professorship appointment at North University of China. Yang's work was partially supported by NSF through grant DMS-1509505.

References

[A82] P. Aviles, *A study of the singularities of solutions of a class of nonlinear partial differential equations*, Comm. PDE **7** (1982), 609–643.

[CGS89] Luis A. Caffarelli, Basilis Gidas and Joel Spruck, *Asymptotic symmetry and local behavior of semilinear elliptic equations with critical Sobolev growth*, Communications on Pure and Applied Mathematics **42**(1989), 271–297.

[CC01] S.-Y.A. Chang and W. Chen, *A note on a class of Higher order conformally covariant equations*, Discreet and Continuous Dynamical systems, Vol. 7, No. 2, (April 2001), pp. 275–281.

[CGY02b] S.-Y.A. Chang, M. Gursky and P. Yang, *An a priori estimate for a fully nonlinear equation on four-manifolds*, J. Anal. Math., **87** (2002), 151–186.

[CGY03] S.-Y.A. Chang, M. Gursky and P. Yang, *Entire solutions of a fully nonlinear equation*, Lectures in Partial Differential Equations in honor of Louis Nirenberg's 75th birthday, chapter 3. International Press, 2003.

[CY97] S.-Y.A. Chang, and P. Yang, *On uniqueness of solution of nth order differential equation in conformal geometry*, Math. Research Letters, **4** (1997) pp. 91–102.

[CHY03] S.-Y.A. Chang, Z. Han, and P. Yang, *Injectivity criteria for the developing map of a locally conformally flat manifold*, preprint, 2003.

[CHgY04] S.-Y.A. Chang, F. Hang, and P. Yang, *On a class of locally conformally flat manifolds*, Int Math Res Not, **2004**:4 (2004), 185–209.

[F95] D. Finn, *Positive Solutions to Nonlinear Elliptic Equations with Prescribed Singularities*, Ph.D. thesis, Northeastern University, 1995.

[FM94] D. Finn and R. McOwen, *Positive solutions of $\Delta_g u = \mu q + Su$ singular at submanifolds with boundary*, Indiana Univ. Math. J. **43**(1994), no. 4, 1359–1397.

[GT] D. Gilberg, and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer; Corr. 3rd printing 1998 edition (March 1, 2001).

[G04] M. González, Ph. D. thesis, Princeton University, 2004.

[G05] M. González, *Singular sets of a class of locally conformally flat manifolds*, Duke Math. J. **129** (2005), no. 3, 551–572.

[GLN18] M. González, Y.-Y. Li, L. Nguyen, *Existence and uniqueness to a fully nonlinear version of the Löwner–Nirenberg problem*, arXiv:1804.08851v1 [math.AP], 24 Apr. 2018.

[GLW04] P. Guan, C.S. Lin and G. Wang, *Application of The Method of Moving Planes to Conformally Invariant Equations*, Mathematische Zeitschrift. **247** (2004), 1–19.

[KP86] Ravi S. Kulkarni and Ulrich Pinkall, *Uniformization of geometric structures with applications to conformal geometry.* Differential geometry, Peñíscola 1985, 190–209, Lecture Notes in Math., 1209, Springer, Berlin, 1986.

[L72] N.S. Landkof, *Foundations of modern potential theory*, Die Grundlehren der mathematischen Wissenschaften, Band 180. Springer-Verlag, New York-Heidelberg, 1972. x+424 pp.

[LL03] Aobing Li and YanYan Li, *On some conformally invariant fully nonlinear equations*, Comm. Pure Appl. Math. **56** (2003), 1414–1464.

[LL05] Aobing Li and YanYan Li, *On some conformally invariant fully nonlinear equations, Part II: Liouville, Harnack and Yamabe*, Acta Math. **195** (2005), 117–154.

[L06] YanYan Li, *Conformally invariant fully nonlinear elliptic equations and isolated singularities*, Journal of Functional Analysis **233** (2006), 380–425.

[LN14] YanYan Li and Luc Nguyen, *A fully nonlinear version of the Yamabe problem on locally conformally flat manifolds with umbilic boundary*, Adv. Math. **251** (2014), 87–110.

[LNW18] YanYan Li, Luc Nguyen, and Bo Wang, *Comparison principles and Lipschitz regularity for some nonlinear degenerate elliptic equations*, Calc. Var. Partial Differential Equations **57** (2018), no. 4.

[L98] C.S. Lin, *A classification of solutions of a conformally invariant fourth order equation in R^n*, Commentarii Mathematici Helvetici, **73** (1998), 206–231.

[LN75] C. Löwner and L. Nirenberg, *Partial Differential Equations invariant under conformal and projective transformation*, Contributions to Analysis, Academic Press, 245–275, 1975.

[MP96] R. Mazzeo and F. Pacard, *A construction of singular solutions for a semilinear elliptic equation using asymptotic analysis*, J. Diff. Geom. **44**(1996), no.2, 331–370.

[MP99] R. Mazzeo and F. Pacard, *Constant scalar curvature metrics with isolated singularities*, Duke Math. J. 99 (1999), no. 3, 353–418.

[MS88] R. Mazzeo and N. Smale, *Conformally flat metrics of constant positive scalar curvature on subdomains of the sphere*, J. Diff. Geom. **34**(1988), 581–621.

[Mc90] R. McOwen, *Singularities and the conformal scalar curvature equation*, Geometric Analysis and Nonlinear Partial Differential Equations (Denton, Tex., 1990), Lecture Notes in Pure and Appl. Math. **144**, Dekker, New York, 1993, 221–233.

[P95] F. Pacard, *The Yamabe problem on subdomains of even-dimensional spheres*, Topol. Methods Nonlinear Anal. **6** (1995), 137–150.

[S88] R. Schoen, *The existence of weak solutions with prescribed singular behavior for a conformally invariant scalar equation*, Comm. Pure Appl. Math. **41** (1988), 317–392.

[Sc88] R. Schoen, *Graduate Course at Stanford*, 1988, https://sites.math.washington.edu/ pollack/research/Schoen-1988-notes.html.

[SY88] R. Schoen and S.T. Yau, *Conformally flat manifolds, Klein groups and scalar curvature*, Invent. Math. **92**(1988), 47–72.

[V81] L. Veron, *Singularités éliminables d'équations elliptiques non linéaires*, J. Diff. Eq., **41** (1981), 225–242.

[V00a] J. Viaclovsky, *Conformal geometry, contact geometry, and the calculus of variations*, Duke Math. J. **101, No. 2** (2000), 283–316.

[V00b] J. Viaclovsky, *Conformally invariant Monge–Ampère equations: global solutions.* Trans. Amer. Math. Soc., **352** (2000), 4371–4379.

[X00] X. Xu, *Uniqueness theorem for the entire positive solutions of biharmonic equations in* $\mathbb{R}^n$, Proc. Roy. Soc. Edinburgh Sect. A **130** (2000), no. 3, 651–670.

Sun-Yung A. Chang and Paul Yang
Department of Mathematics
Fine Hall
Princeton University
Princeton, NJ 08540, USA
e-mail: chang@math.princeton.edu
yang@math.princeton.edu

Zheng-Chao Han
Department of Mathematics
Rutgers University
110 Frelinghuysen Road
Piscataway, NJ 08854, USA
e-mail: zchan@math.rutgers.edu

Progress in Mathematics, Vol. 333, 57–70

Analytical Properties for Degenerate Equations

Tobias Holck Colding and William P. Minicozzi II

Abstract. By a classical result, solutions of analytic elliptic PDEs, like the Laplace equation, are analytic. In many instances, the properties that come from being analytic are more important than analyticity itself. Many important equations are degenerate elliptic and solutions have much lower regularity. Still, one may hope that solutions share properties of analytic functions. These properties are closely connected to important open problems. In this survey, we will explain why solutions of an important degenerate elliptic equation have analytic properties even though the solutions are not even C^3.

Mathematics Subject Classification (2010). 53C44 (35J60).

Keywords. Mean curvature flow, arrival time, degenerate elliptic equations, Thom's gradient conjecture.

0. Introduction

By a classical result, solutions of analytic elliptic PDEs, like the Laplace equation, are analytic. In many instances, the properties that come from being analytic are more important than analyticity itself. Many important equations are degenerate elliptic and solutions have much lower regularity. Still, one may hope that solutions share properties of analytic functions. These properties are closely connected to important open problems.

In this survey, we will explain why solutions of an important degenerate elliptic equation have analytic properties even though the solutions are not even C^3. This equation, known as the *arrival time equation*, is

$$-1 = |\nabla u| \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) . \tag{0.1}$$

Here u is defined on a compact connected subset of $\mathbf{R}^{n+1}$ with smooth mean convex boundary and u is constant on the boundary. Equation (0.1) is the prototype for a family of equations, see, e.g., [OsSe], used for tracking moving interfaces in complex situations. These equations have been instrumental in applications, including semiconductor processing, fluid mechanics, medical imaging, computer graphics, and material sciences.

Even though solutions of (0.1) are a priori only in the viscosity sense, they are always twice differentiable by [CM5], though not necessarily C^2; see [CM6], [H2], [I], [KS]. Even when a solution is C^2, it still might not be C^3, Sesum, [S], let alone analytic. However, solutions have the following property conjectured for analytic functions:

Theorem 0.2 ([CM9]). *The Arnold–Thom conjecture holds for C^2 solutions of* (0.1). *Namely, if $x(t)$ is a gradient flow line for u, then $x(t)$ has finite length and $\frac{x'(t)}{|x'(t)|}$ has a limit.*

The theorem applies, for instance, to solutions of (0.1) on closed convex domains since these are C^2 by a 1990 result of Huisken, [H1] (though not necessarily C^3, [S]).

As we will see in Section 1, the Arnold–Thom conjecture states that $\frac{x'(t)}{|x'(t)|}$ has a limit whenever u is analytic and the gradient flow line itself has a limit. Thus, Theorem 0.2 shows that, in this way, C^2 solutions of (0.1) behave like analytic functions are expected to.

0.1. The arrival time

The geometric meaning of (0.1) is that the level sets $u^{-1}(t)$ are mean convex and evolve by mean curvature flow. One says that u is the *arrival time* since $u(x)$ is the time the hypersurfaces $u^{-1}(t)$ arrive at x under the mean curvature flow; see Chen–Giga–Goto, [ChGG], Evans–Spruck, [ES], Osher–Sethian, [OsSe], and [CM3].

Conjecturally, the Arnold–Thom conjecture holds even for solutions that are not C^2, but merely twice differentiable:

Conjecture 0.3 ([CM9]). *Lojasiewicz's inequalities and the Arnold–Thom conjecture hold for all solutions of* (0.1).

If so, this would explain various conjectured phenomena. For example, this would imply that the associated mean curvature flow is singular at only finitely many times as has been conjectured, [W3], [AAG], [Wa], [M].

We believe that the principle that solutions of degenerate equations behave as though they are analytic, even when they are not, should be quite general. For instance, there should be versions for other flows, including Ricci flow.

0.2. Ideas in the proof

To explain the ideas in the proof of Theorem 0.2, suppose that the unit speed curve $\gamma(s)$ traces out a gradient flow line for u that limits to a critical point x_0. The simplest way to prove that the unit tangent γ_s has a limit would be to prove that

$$\int |\gamma_{ss}| \, ds < \infty \, . \tag{0.4}$$

However, this is not necessarily true. It turns out that γ_{ss} is better behaved in some directions than in others, depending on the geometry of the level sets of u.

By [CM5], the level sets of u near x_0 are, in a scale-invariant way, converging to either spheres or cylinders. This comes from a blow up analysis for the singularities of an associated mean curvature flow. The spherical case is easy to handle and one can show that $|\gamma_{ss}|$ is integrable in this case. However, the cylindrical case is more subtle and γ_{ss} behaves quite differently. In particular, the estimates in the direction of the axis of the cylinder are not strong enough to give (0.4).

There is a good reason that the estimates are not strong enough here: the presence of a "non-integrable" kernel for a linearized operator. Here "non-integrable" means that there are infinitesimal variations that do not arise as the derivative of an actual one-parameter family of solutions. As is well known, this corresponds to a slow rate of convergence to the limiting blow up. Overcoming this requires a careful analysis of this kernel using the rate of growth (the frequency function) for the drift Laplacian.

1. Gradient flows in finite dimensions

Given a function f, a gradient flow line $x(t)$ is a solution of the ODE

$$x'(t) = \nabla f \circ x(t) \tag{1.1}$$

with the initial condition $x(0) = \bar{x}$. The chain rule gives that

$$(f \circ x(t))' = |\nabla f|^2 \circ x(t)\,, \tag{1.2}$$

so we see that $f \circ x(t)$ is increasing unless $x(t) \equiv \bar{x}$ is a critical point of f.

It is possible that $x(t)$ runs off to infinity (e.g., if $f(x, y) = x$ on $\mathbf{R}^2$), but we are interested in the case where there is a limit point x_∞. That is, where there exist $t_i \to \infty$ so that $x(t_i) \to x_\infty$. It follows easily that $\lim_{t\to\infty} f \circ x(t) = f(x_\infty)$, x_∞ is a critical point, and $|x'|^2$ is integrable. This raises the obvious question:

Question 1.3. *Does $x(t)$ converge to x_∞?*

Perhaps surprisingly, there are examples where $x(t)$ does not converge; see, e.g., Fig. 3.5 in [Si] or Fig. 1 in [CM8]. However, if f is real analytic, Lojasiewicz, [L1], proved that $x(t)$ has finite length and, thus, converges. This is known as *Lojasiewicz's theorem*. The proof relied on two *Lojasiewicz inequalities* for analytic functions.

1.1. Lojasiewicz inequalities

In real algebraic geometry, the Lojasiewicz inequality, [L3], bounds the distance from a point to the nearest zero of a given real analytic function. Namely, if $Z \neq \emptyset$ is the zero set of f and K is a compact set, then there exist $\alpha \geq 2$ and a positive constant C such that for $x \in K$

$$\inf_{z \in Z} |x - z|^\alpha \leq C\,|f(x)|\,. \tag{1.4}$$

The exponent α can be arbitrarily large, depending on the function f.

Equation (1.4) was the main ingredient in Lojasiewicz' proof of Laurent Schwarz' division conjecture[1] in analysis. Around the same time, Hörmander, [Hö], independently proved Schwarz' division conjecture in the special case of polynomials and a key step in his proof was also (1.4) when f is a polynomial.

Lojasiewicz solved a conjecture of Whitney[2] in [L4] using a second inequality – known as the gradient inequality: Given a critical point z, there is neighborhood W of z and constants $p > 1$ and $C > 0$ such that for all $x \in W$

$$|f(x) - f(z)| \leq C\, |\nabla_x f|^p \,. \tag{1.5}$$

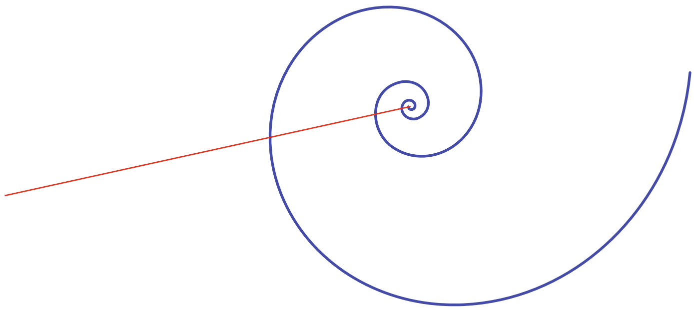
An immediate consequence of (1.5) is that f takes the same value at every critical point in W. It is easy to construct smooth functions where this is not the case.

This gradient inequality (1.5) was the key ingredient in the proof of the Lojasiewicz theorem. The idea is that (1.5) and (1.2) give a differential inequality for f along the gradient flow line that leads to a rate of convergence; see, e.g., [L1], [CM2], [CM8] and [Si].

1.2. Arnold–Thom conjectures

Around 1972, Thom, [T], [L2], [Ku], [A], [G], conjectured a strengthening of Lojasiewicz' theorem, asserting that each gradient flow line $x(t)$ of an analytic function f approaches its limit from a unique limiting direction:

Conjecture 1.6. *If $x(t)$ has a limit point, then the limit of secants $\lim_{t\to\infty} \frac{x(t)-x_\infty}{|x(t)-x_\infty|}$ exists.*



The figure illustrates in $\mathbf{R}^3$ a situation conjectured to be impossible. The Arnold–Thom conjecture asserts that a blue integral curve does not spiral as it approaches the critical set (illustrated in red, orthogonal to the plane where the curve spirals).

[1] L. Schwartz conjectured that if f is a non-trivial real analytic function and T is a distribution, then there exists a distribution S satisfying $f\,S = T$.

[2] Whitney conjectured that if f is analytic in an open set U of $\mathbf{R}^n$, then the zero set Z is a deformation retract of an open neighborhood of Z in U.

This conjecture arose in Thom's work on catastrophe theory and singularity theory and became known as *Thom's gradient conjecture*. The conjecture was finally proven in 2000 by Kurdyka, Mostowski, and Parusinski in [KMP], but the following stronger conjecture remains open (see page 282 in Arnold's problem list, [A]):

Conjecture 1.7. *If $x(t)$ has a limit point, then the limit of the unit tangents $\frac{x'(t)}{|x'(t)|}$ exists.*

It is easy to see that the *Arnold–Thom conjecture* 1.7 implies Thom's conjecture 1.6.

2. Lojasiewicz theorem for the arrival time

The arrival time u is a solution of the degenerate elliptic equation (0.1) and, in particular, it is not smooth in general, let alone real analytic. However, if u is C^2, then it satisfies the following gradient Lojasiewicz inequality:

Theorem 2.1 ([CM9]). *If u is a C^2 solution of* (0.1) *and* $\sup u = 0$*, then* 0 *is the only critical value and*

$$\frac{|\nabla u|^2}{-u} \to \frac{2}{n-k} \text{ as } u \to 0\,. \tag{2.2}$$

In particular, there exists $C > 0$ so that $C^{-1}\,|\nabla u|^2 \leq -u \leq C\,|\nabla u|^2$.

In particular, (1.5) holds with $p = 2$ for a C^2 solution u of (0.1). Given any $p > 1$, there are solutions of (0.1) where (1.5) fails for p; obviously, these are not C^2. Namely, for any odd $m \geq 3$, Angenent and Velázquez, [AV], construct rotationally symmetric examples with

$$|u - u(y)| \approx |\nabla u|^{\frac{m}{m-1}} \tag{2.3}$$

for a sequence of points tending to y. The examples in [AV] were constructed to analyze so-called type II singularities that were previously observed by Hamilton and proven rigorously to exist by Altschuler–Angenent–Giga, [AAG]; cf. also [GK].

2.1. The flow lines approach the critical set orthogonally

Let u be a C^2 solution to (0.1) with $\sup u = 0$ and $\mathcal{S}$ its critical set

$$\mathcal{S} = \{x \,|\, \nabla u(x) = 0\}\,. \tag{2.4}$$

The mean curvature flow given by the level sets of u is smooth away from $\mathcal{S}$ and each point in $\mathcal{S}$ has a cylindrical singularity; see, [W1], [W2], [H1], [HS1], [HS2], [HaK], [An]; cf. [B], [CM1]. Moreover, [CM5] and [CM6] give:

($\mathcal{S}1$) $\mathcal{S}$ is a closed embedded connected k-dimensional C^1 submanifold whose tangent space is the kernel of Hess_u. Moreover, $\mathcal{S}$ lies in the interior of the region where u is defined.

($\mathcal{S}2$) If $q \in \mathcal{S}$, then $\text{Hess}_u(q) = -\frac{1}{n-k}\,\Pi$ and $\Delta u(q) = -\frac{n+1-k}{n-k}$, where Π is the orthogonal projection onto the orthogonal complement of the kernel.

The next theorem shows that the gradient flow lines of u have finite length (this is the Lojasiewicz theorem for u), converge to points in $\mathcal{S}$, and approach $\mathcal{S}$ orthogonally. The first claims follow immediately from the gradient Lojasiewicz inequality of Theorem 2.1. Let Π_{axis} denote the orthogonal projection onto the kernel of Hess_u.

Theorem 2.5 ([CM9]). *Each flow line γ for ∇u has finite length and limits to a point in $\mathcal{S}$. Moreover, if we parametrize γ by $s \geq 0$ with $|\gamma_s| = 1$ and $\gamma(0) \in \mathcal{S}$, then*

$$u(\gamma(s)) \approx \frac{-s^2}{2(n-k)}\,, \tag{2.6}$$

$$|\nabla u(\gamma(s))|^2 \approx \frac{s^2}{(n-k)^2}\,, \tag{2.7}$$

$$\Pi_{axis}(\gamma_s) \to 0\,. \tag{2.8}$$

In particular, for s small, we have that $\gamma(s) \subset B_{2n\sqrt{-u(\gamma(s))}}(\gamma(0))$.

3. Theorem 0.2 and an estimate for rescaled MCF

The Arnold–Thom conjecture for the arrival time is phrased as an analytic question about solutions to a degenerate elliptic partial differential equation. Yet, we will see that the key is understanding the geometry of an associated mean curvature flow.

3.1. Rescaled mean curvature flow

A one-parameter family of hypersurfaces M_τ evolves by *mean curvature flow* (or *MCF*) if each point $x(\tau)$ evolves by $\partial_\tau x = -H\,\mathbf{n}$. Here H is the mean curvature and $\mathbf{n}$ a unit normal. The arrival time gives a mean curvature flow

$$\Sigma_\tau = \{x \,|\, u(x) = \tau\}\,. \tag{3.1}$$

As τ goes to the extinction time (the supremum of u), the level sets contract and eventually disappear. To capture the structure near the extinction, we consider the rescaled level sets

$$\Sigma_t = \frac{1}{\sqrt{-u}}\,\{x \,|\, u(x) = -\mathrm{e}^{-t}\}\,. \tag{3.2}$$

This is equivalent to simultaneously running MCF and rescaling space and reparameterizing time. The one-parameter family Σ_t satisfies the *rescaled MCF*

$$\partial_t x = -\left(H - \frac{1}{2}\,\langle x, \mathbf{n}\rangle\right)\,\mathbf{n}\,. \tag{3.3}$$

The rescaled MCF is the negative gradient flow for the Gaussian area

$$F(\Sigma) \equiv \int_{\Sigma} \mathrm{e}^{-\frac{|x|^2}{4}} . \tag{3.4}$$

In particular, $F(\Sigma_t)$ is non-increasing.

It will be convenient to set $\phi = H - \frac{1}{2}\langle x, \mathbf{n}\rangle$. The fixed points for rescaled MCF are shrinkers where $\phi = 0$; the most important shrinkers are cylinders $\mathcal{C} = \mathbf{S}^{n-k}_{\sqrt{2(n-k)}} \times \mathbf{R}^k$ where $k = 0, \ldots, n-1$.

3.2. Rate of convergence of the rescaled MCF

The F functional is nonincreasing along the rescaled MCF Σ_t and it is constant only when Σ_t is also constant. Furthermore, the distance between Σ_j and Σ_{j+1} is bounded by

$$\delta_j \equiv \sqrt{F(\Sigma_{j-1}) - F(\Sigma_{j+2})} . \tag{3.5}$$

We refer to [CM9] (cf. [CM2]) for the precise statement, but the idea is simple. To see this, consider the analogous question for a finite-dimensional gradient flow $x(t)$. In this case, the fundamental theorem of calculus and the Cauchy–Schwarz inequality give

$$|x(j+1) - x(j)| \leq \int_j^{j+1} |x'(t)| \, dt \leq \left(\int_j^{j+1} |x'(t)|^2 \, dt \right)^{\frac{1}{2}} = (f(j+1) - f(j))^{\frac{1}{2}} . \tag{3.6}$$

Existence of $\lim_{t\to\infty} \Sigma_t$ is proven in [CM2] by showing that $\sum \delta_j < \infty$. In [CM9], we prove that δ_j is summable even after being raised to some power less than one:

Proposition 3.7 ([CM9]). *There exists $\bar{\beta} < 1$ so that $\sum_{j=1}^{\infty} \delta_j^{\bar{\beta}} < \infty$.*

3.3. A strong cylindrical approximation

Since Σ_t converges to a limit $\mathcal{C}$, Σ_j is close to $\mathcal{C}$ for j large. However, we will construct cylinders $\mathcal{C}_j$, varying with j, that are even closer. We need some notation: Π_j is projection orthogonal to axis of $\mathcal{C}_j$, $\mathcal{L}$ is the drift Laplacian on $\mathcal{C}_j$, and

$$\|g\|^p_{L^p(\Sigma_j)} \equiv \int_{\Sigma_j} |g|^p \, \mathrm{e}^{-\frac{|x|^2}{4}} . \tag{3.8}$$

The precise statement of the approximation is technical (see [CM9]), but it roughly says:

Proposition 3.9 ([CM9]). *Given $\beta < 1$, there exist C, radii R_j, and cylinders $\mathcal{C}_j$ with:*

1. *For $t \in [j, j+1]$, Σ_t is a graph over $B_{R_j} \cap \mathcal{C}_{j+1}$ of a function w with*
$$\|w\|^2_{W^{3,2}} + \|\phi\|_{W^{3,2}(B_{R_j})} + e^{-\frac{R_j^2}{4}} \leq C \, \delta_j^{\beta} .$$
2. *w is almost an eigenfunction; i.e., $|\phi - (\mathcal{L}+1)\, w|$ is quadratic in w.*

3. $|\Pi_j - \Pi_{j+1}| \leq C\,\delta_j^\beta$.
4. *The higher derivatives of w and ϕ are bounded.*

3.4. Reduction

The next theorem reduces Theorem 0.2 to an estimate for rescaled MCF.

Theorem 3.10 ([CM9]). *Theorem 0.2 holds if*

$$\sum_{j=1}^{\infty} \int_j^{j+1} \left(\sup_{B_{2n} \cap \Sigma_t} |\Pi_{j+1}(\nabla H)| \right) dt < \infty\,. \tag{3.11}$$

To explain Theorem 3.10, let $\gamma(s)$ be a unit speed parameterization of a gradient flow line with $\gamma(0) \in \mathcal{S}$. We will show that γ_s has a limit as $s \to 0$. The derivative of $\gamma_s = -\frac{\nabla u}{|\nabla u|}$ is

$$\gamma_{ss} = -\frac{1}{|\nabla u|} \left(\mathrm{Hess}_u(\gamma_s) - \gamma_s \left\langle \mathrm{Hess}_u(\gamma_s), \gamma_s \right\rangle \right) = -\frac{\left(\mathrm{Hess}_u(\gamma_s)\right)^T}{|\nabla u|} = \nabla^T \log |\nabla u|\,, \tag{3.12}$$

where $(\cdot)^T$ is the tangential projection onto the level set of u.

The simplest way to prove that $\lim \gamma_s$ exists would be to show that $\int |\gamma_{ss}| < \infty$, which is related to the rate of convergence for an associated rescaled MCF. While this rate fails to give integrability of $|\gamma_{ss}|$, it does give the following:

Lemma 3.13. *Given any $\Lambda > 1$, we have $\lim_{s \to 0} \int_s^{\Lambda s} |\gamma_{ss}|\, ds = 0$.*

Proof. Using Theorem 2.5 and the fact that $\mathrm{Hess}_u \to -\frac{1}{n-k}\,\Pi$, (3.12) implies that $s\,|\gamma_{ss}| \to 0$. The lemma follows immediately from this. $\square$

To get around the lack of integrability, we will decompose γ_s into two pieces – the parts tangent and orthogonal to the axis – and deal with these separately. The tangent part goes to zero by (2.8) in Theorem 2.5. We will use (3.11) to control the orthogonal part.

Translate so that $\gamma(0) = 0$ and let $\bar{H} = \frac{1}{|\nabla u|}$ be the mean curvature of the level set of u. The mean curvature H of Σ_t at time $t = -\log(-u)$ is given by

$$\bar{\nabla} \log \bar{H} = \frac{\nabla \log H}{\sqrt{-u}} \approx \frac{\sqrt{2(n-k)}}{s}\; \nabla \log H\,. \tag{3.14}$$

Note that $u(\gamma(s))$ is decreasing and Theorem 2.5 gives

$$t(s) \approx -2\,\log s + \log(2(n-k))\,, \tag{3.15}$$

$$t'(s) = -\partial_s \left(\log(-u(\gamma(s))) \right) = \frac{-\partial_s u(\gamma(s))}{u(\gamma(s))} \approx -\frac{2}{s}\,. \tag{3.16}$$

Given a positive integer j, define s_j so that $t(s_j) = j$. Note that $\left| \log \frac{s_{j+1}}{s_j} \right|$ is uniformly bounded. Therefore, by Lemma 3.13, it suffices to show that γ_{s_j} has a limit.

We can write $\gamma_{s_j} = \Pi_{\text{axis},j}(\gamma_{s_j}) + \Pi_j(\gamma_{s_j})$. We have $\Pi_{\text{axis},j}(\gamma_{s_j}) \to 0$ since $\Pi_{\text{axis},j} \to \Pi_{\text{axis}}$ and $\Pi_{\text{axis}}(\gamma_s) \to 0$. Thus, we need that $\lim_{j\to\infty} \Pi_j(\gamma_{s_j})$ exists; this will follow from

$$\sum_j \left|\Pi_j(\gamma_{s_j}) - \Pi_{j+1}(\gamma_{s_{j+1}})\right| < \infty\,. \tag{3.17}$$

Theorem 2.5 gives (for s small) that $\gamma(s) \subset B_{2n\sqrt{-u(\gamma(s))}}$ and, thus, (3.12) gives

$$\begin{aligned}\left|\Pi_{j+1}(\gamma_{s_j}) - \Pi_{j+1}(\gamma_{s_{j+1}})\right| &\leq \int_{s_{j+1}}^{s_j} \left|\Pi_{j+1}(\gamma_{ss})\right|\, ds \\ &= \int_{s_{j+1}}^{s_j} \left|\Pi_{j+1}\left(\bar\nabla \log \bar H(\gamma(s))\right)\right|\, ds \\ &\leq C \int_{s_{j+1}}^{s_j} \sup_{B_{2n\sqrt{-u(\gamma(s))}}} \left|\Pi_{j+1}(\nabla \log \bar H)\right|(\cdot, -u)\, ds\,.\end{aligned} \tag{3.18}$$

Using (3.14) and (3.16) in (3.18) and then applying Theorem 3.10 gives

$$\sum_j \left|\Pi_{j+1}(\gamma_{s_j}) - \Pi_{j+1}(\gamma_{s_{j+1}})\right| \leq C \sum_j \int_j^{j+1} \sup_{B_{2n}\cap\Sigma_t} \left|\Pi_{j+1}(\nabla H)\right|\, dt < \infty\,. \tag{3.19}$$

On the other hand, $\sum_j \left|\Pi_j(\gamma_{s_j}) - \Pi_{j+1}(\gamma_{s_j})\right| < \infty$ by (3) in Proposition 3.9 and Proposition 3.7. The triangle inequality gives (3.17), so we conclude that γ_s has a limit.

3.5. The summability condition (3.11)

We have seen that the key is to prove (3.11). This summability is plausible since Σ_t is converging to a cylinder where H is constant and, thus, ∇H is going to zero. The rate of convergence then becomes critical. If the convergence was fast enough, then $|\nabla H|$ would be summable even without the projection Π_{j+1}.

The mean curvature H of the graph of w is given at each point explicitly as a function of w, ∇w and Hess_w; see corollary $A.30$ in [CM2]. We can write this as the first-order part (in $w, \nabla w, \text{Hess}_w$) plus a quadratic remainder

$$H = H_{\mathcal{C}} + \left(\Delta_\theta + \Delta_x + \frac{1}{2}\right) w + O(w^2)\,. \tag{3.20}$$

Here $O(w^2)$ is a term that depends at least quadratically on $w, \nabla w, \text{Hess}_w$ and the constant $H_{\mathcal{C}} = \frac{\sqrt{n-k}}{\sqrt{2}}$ is the mean curvature of $\mathcal{C}$.

The bound for w^2 in (1) from Proposition 3.9 is summable by Proposition 3.7, but the bound for w is not. In particular, (1) gives a bound for ∇H that is not summable.

4. Approximate eigenfunctions on cylinders

Proposition 3.9 shows the graph function w is an approximate eigenfunction on the cylinder $\mathcal{C}$. Namely, (2) gives that

$$|(\mathcal{L}+1)w - \phi| = O(w^2)\,, \tag{4.1}$$

where $O(w^2)$ is a term that is quadratically bounded in w and its derivatives. Note that ϕ itself is bounded by (1) and, moreover, ϕ is of the same order as w^2.

Even though the bound for w is not summable, we will see that there is a function $\tilde{w}$ in the kernel of $\mathcal{L}+1$ so that $|w-\tilde{w}|$ is summable. Moreover, the contribution of $\tilde{w}$ to ∇H goes away once we project orthogonally to the axis. Putting this together gives (3.11) and, thus, completes the proof of Arnold–Thom. The arguments needed for this decomposition of w are technically complicated because of higher-order "error" terms; see [CM9]. However, the idea is clear. We will explain this in a model case next.

4.1. Eigenfunctions on cylinders

The eigenfunctions on the cylinder are built out of spherical eigenfunctions on the cross-section and eigenfunctions for the Euclidean drift Laplacian on the axis. Namely, by lemma 3.26 in [CM2], the kernel of $\mathcal{L}+1$ on the weighted Gaussian space on $\mathcal{C}$ consists of quadratic polynomials and "infinitesimal rotations"

$$\tilde{w} = \sum_i a_i(x_i^2 - 2) + \sum_{i<j} a_{ij}x_i x_j + \sum_k x_k h_k(\theta)\,, \tag{4.2}$$

where a_i, a_{ij} are constants and each h_k is a Δ_θ-eigenfunction with eigenvalue $\frac{1}{2}$.

To illustrate the ideas involved, it is helpful to recall the Euclidean case:

Lemma 4.3. *If $\mathcal{L}v = -\lambda v$ on $\mathbf{R}^n$ and $\int_{\mathbf{R}^n} v^2\, e^{-\frac{|x|^2}{4}} < \infty$, then $2\,\lambda$ is a nonnegative integer and v is a polynomial of degree 2λ.*

When $n=1$, these are the Hermite polynomials (up to a scaling normalization).

Sketch of the proof of Lemma 4.3. There are two ingredients:

- Each partial derivative $v_i = \frac{\partial v}{\partial x_i}$ satisfies $\mathcal{L}\,v_i = \left(\frac{1}{2} - \lambda\right)\, v_i$.
- $\int_{\mathbf{R}^n} |\nabla v|^2\, \mathrm{e}^{-\frac{|x|^2}{4}} \leq 2\lambda \int_{\mathbf{R}^n} v^2\, \mathrm{e}^{-\frac{|x|^2}{4}} < \infty$.

The second property implies that $\lambda \geq 0$ and v is constant if $\lambda = 0$. The lemma follows by applying this to 2λ derivatives of v. $\square$

The next theorem approximates w in $|x| \leq 3n$ by $\tilde{w}$ as in (4.2) (we state the theorem in the model case where w is an eigenfunction; see [CM9] for approximate eigenfunctions).

Theorem 4.4 ([CM9]). *Given $\nu < 1$, there exists C so that if $(\mathcal{L}+1)w = 0$ on B_R with $e^{-\frac{R^2}{4}} + \|w\|^2_{W^{3,2}} \leq \delta$ and $w^2 \leq \delta\, e^f$, then there is a function $\tilde{w}$ as in (4.2) with*

$$\sup_{|x|\leq 3n} |w - \tilde{w}| \leq C\, \delta^{\nu}\,. \tag{4.5}$$

This gives the improved estimate that we need. Namely, (1) in Proposition 3.9 gives $|w| \leq C\, \delta_j^{\frac{1}{2}}$, while (4.5) gives $|w - \tilde{w}| \leq C\, \delta_j^{\nu}$ with $\nu \approx 1$. The first bound is not summable, but the second bound is by Proposition 3.7.

The L^2 methods for Lemma 4.3 yield sharp global results, but are not sharp enough for (4.5). We will need a different approach – the frequency – that is explained next.

4.2. The frequency

The key to understanding the growth of eigenfunctions for $\mathcal{L}$ is a frequency function inspired by Almgren's frequency for harmonic functions, [Al], cf. [GL], [HaS], [Ln], [CM10], [D]. The frequency was used by Bernstein, [Be], to study the ends of shrinkers and in [CM7] to study the growth of approximate eigenfunctions.

To explain the frequency, set $f = \frac{|x|^2}{4}$ on $\mathbf{R}^n$ and define quantities $I(r)$ and $D(r)$

$$I(r) = r^{1-n} \int_{\partial B_r} u^2\,, \tag{4.6}$$

$$D(r) = r^{2-n} \int_{\partial B_r} u u_r = r^{2-n}\, \mathrm{e}^{f(r)} \int_{B_r} \left(|\nabla u|^2 - V\, u^2\right) \mathrm{e}^{-f}\,, \tag{4.7}$$

and the *frequency* $U(r) = \frac{D}{I}$. Thus, $(\log I)' = \frac{2U}{r}$, so U measures the polynomial rate of growth of $\sqrt{I}$. For example, if $u(x) = |x|^d$, then $U = d$.

There is a dichotomy where eigenfunctions of $\mathcal{L}$ are polynomial or grow exponentially:

Theorem 4.8. [CM7] *Given $\epsilon > 0$ and $\delta > 0$, there exist $r_1 > 0$ such that if $\mathcal{L}u = -\lambda u$ and $U(\bar{r}_1) \geq \delta + 2 \sup\{0, \lambda\}$ for some $\bar{r}_1 \geq r_1$, then for all $r \geq R(\bar{r}_1)$*

$$U(r) > \frac{r^2}{2} - n - 2\,\lambda - \epsilon\,. \tag{4.9}$$

This theorem explains why we expect a good approximation when the eigenfunction is defined (and bounded) on a large ball. This is easiest to explain in the case $\lambda = 0$ (we can reduce to this after taking 2λ partial derivatives). Namely, subtracting a constant to make the average zero on a ball, we can use the Poincaré inequality to get a positive lower bound for the frequency on a fixed inner ball. Theorem 4.8 then implies extremely rapid growth out to the boundary of the ball. Since the function is bounded, we conclude that it must be small on the inner ball as claimed. See Theorem 4.1 in [CM7] for details (compare [CM9]).

Acknowledgment

The authors were partially supported by NSF Grants DMS 1404540 and DMS 1707270.

References

[AAG] S. Altschuler, S. Angenent, and Y. Giga, *Mean curvature flow through singularities for surfaces of rotation*. J. Geom. Anal. 5(3), 293–358 (1995).

[An] B. Andrews, *Noncollapsing in mean-convex mean curvature flow*. G&T 16 (2012), no. 3, 1413–1418.

[AV] S. Angenent and J.J.L. Velázquez, *Degenerate neckpinches in mean curvature flow*. J. Reine Angew. Math. 482 (1997), 15–66.

[Al] F. Almgren, Jr., *Q-valued functions minimizing Dirichlet's integral and the regularity of area minimizing recti able currents up to codimension two*, preprint.

[A] V.I. Arnold, *Arnold's problems*. Translated and revised edition of the 2000 Russian original. Springer-Verlag, Berlin; PHASIS, Moscow, 2004.

[Be] J. Bernstein, *Asymptotic structure of almost eigenfunctions of drift Laplacians on conical ends*, preprint, https://arxiv.org/pdf/1708.07085.pdf.

[B] S. Brendle, *Embedded self-similar shrinkers of genus* 0, Annals of Math. 183, 715–728 (2016).

[ChGG] Y.G. Chen, Y. Giga, and S. Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations*. J. Differential Geom. 33 (1991), no. 3, 749–786.

[CM1] T.H. Colding and W.P. Minicozzi II, *Generic mean curvature flow I; generic singularities*, Annals of Math., 175 (2012), 755–833.

[CM2] ———, *Uniqueness of blowups and Łojasiewicz inequalities*, Annals of Math., 182 (2015), 221–285.

[CM3] ———, *Level set method for motion by mean curvature*, Notices of the AMS, (2016) Vol. 63, No. 10, 1148–1153.

[CM4] ———, *The singular set of mean curvature flow with generic singularities*, Inventiones Math., 204 (2) (2016), 443–471.

[CM5] ———, *Differentiability of the arrival time*, CPAM, Vol. LXIX, (2016) 2349–2363.

[CM6] ———, *Regularity of the level set flow*, CPAM, Vol. LXXIV, (2018) 814–824.

[CM7] ———, *Sharp frequency bounds for eigenfunctions of the Ornstein–Uhlenbeck operator*, Calc. Var. Partial Differential Equations 57 (2018), no. 5, Art. 138, 16 pp.

[CM8] ———, *Łojasiewicz inequalities and applications*. Surveys in differential geometry 2014. Regularity and evolution of nonlinear equations, 63–82, Surv. Differ. Geom., 19, Int. Press, Somerville, MA, 2015.

[CM9] ———, *Arnold–Thom gradient conjecture for the arrival time*, Comm. Pure Appl. Math. 72 (2019), no. 7, 1548–1577.

[CM10] ———, *Harmonic functions of polynomial growth*, JDG 46 (1997), 1–77.

[D] C. De Lellis, *The size of the singular set of area-minimizing currents*, Surveys in differential geometry 2016. Advances in geometry and mathematical physics, 1–83, Surv. Differ. Geom., 21, Int. Press, Somerville, MA, 2016.

[ES] L.C. Evans and J. Spruck, *Motion of level sets by mean curvature I*, JDG 33 (1991) 635–681.

[GK] Z. Gang and D. Knopf, *Universality in mean curvature flow neckpinches*. Duke Math. J. 164 (2015), no. 12, 2341–2406.

[GL] N. Garofalo and F.H. Lin, *Monotonicity properties of variational integrals, A_p weights and unique continuation*, Indiana Univ. Math. J. 35 (1986), no. 2, 245–268.

[G] P. Goldstein, *Gradient flow of a harmonic function in* $\mathbf{R}^3$. J. Diff. Eq. 247 (2009), no. 9, 2517–2557.

[HaS] R. Hardt and L. Simon, *Nodal sets for solutions of elliptic equations*, JDG 30 (1989), 505–522.

[HaK] R. Haslhofer and B. Kleiner, *Mean curvature flow of mean convex hypersurfaces*, CPAM, (2017), 511–546.

[Hö] L. Hörmander, *On the division of distributions by polynomials*, Ark. Mat. 3 (1958) 555–568.

[H1] G. Huisken, *Asymptotic behavior for singularities of the mean curvature flow.* JDG 31 (1990), 285–299.

[H2] ———, *Local and global behaviour of hypersurfaces moving by mean curvature.* Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990), 175–191, Proc. Sympos. Pure Math., 54, Part 1, AMS, Providence, RI, 1993.

[HS1] G. Huisken and C. Sinestrari, *Mean curvature flow singularities for mean convex surfaces*, Calc. Var. Partial Differ. Equ. 8 (1999), 1–14.

[HS2] ———, *Convexity estimates for mean curvature flow and singularities of mean convex surfaces*, Acta Math. 183 (1999), no. 1, 45–70.

[I] T. Ilmanen, *Generalized Flow of Sets by Mean Curvature on a Manifold*, Indiana J. Math. 41 (1992) 671–705.

[Ku] N. Kuiper, *On gradient curves of an analytic function near a critical point*, preprint IHES, 1991.

[KMP] K. Kurdyka, T. Mostowski, and A. Parusinski, *Proof of the gradient conjecture of R. Thom*, Annals of Math., 152 (2000), 763–792.

[KS] R.V. Kohn and S. Serfaty, *A deterministic control based approach to motion by curvature*. CPAM. 59 (2006), no. 3, 344–407.

[Ln] F.H. Lin, *Nonlinear theory of defects in nematic liquid crystals; phase transition and flow phenomena*, CPAM. 42 (1989), 789–814.

[L1] S. Lojasiewicz, *Ensembles semi-analytiques*, IHES notes (1965).

[L2] ———, *Sur la géometrie semi- et sous-analytique*, Ann. Inst. Fourier 43 (1993), 1575–1595.

[L3] ———, *Sur le problème de la division*, Studia Math. 18 (1959) 87–136.

[L4] ———, *Une propriété topologique des sous-ensembles analytiques réels*, Les Équations aux Dérivées Partielles (1963) 87–89, CNRS, Paris.

[M] T. Miura, *An example of a mean-convex mean curvature flow developing infinitely many singular epochs*. J. Geom. Anal. 26 (2016), no. 4, 3019–3026.

[OsSe] S. Osher and J.A. Sethian, *Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations*. J. Comput. Phys. 79 (1988), no. 1, 12–49.

[S] N. Sesum, *Rate of convergence of the mean curvature flow*. CPAM 61 (2008), no. 4, 464–485.

[Si] L. Simon, *Theorems on Regularity and Singularity of Energy Minimizing Maps*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1996.

[T] R. Thom, *Problèmes rencontrés dans mon parcours mathématique: un bilan*, Publ. Math. IHES 70 (1989), 200–214.

[Wa] X.-J. Wang, *Convex solutions to the mean curvature flow*. Ann. of Math. (2) 173 (2011), 1185–1239.

[W1] B. White, *The size of the singular set in mean curvature flow of mean-convex sets*. JAMS 13 (2000), 665–695.

[W2] ———, *The nature of singularities in mean curvature flow of mean-convex sets*. JAMS 16 (2003), 123–138.

[W3] ———, *Evolution of curves and surfaces by mean curvature*, in Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002), Higher Ed. Press, Beijing, 2002, 525–538.

Tobias Holck Colding and William P. Minicozzi II
MIT, Dept. of Math.
77 Massachusetts Avenue
Cambridge, MA 02139-4307, USA
e-mail: `colding@math.mit.edu`
`minicozz@math.mit.edu`

Progress in Mathematics, Vol. 333, 71–92

Equivariant K-theory and Resolution I: Abelian Actions

Panagiotis Dimakis and Richard Melrose

Abstract. The smooth action of a compact Lie group on a compact manifold can be resolved to an iterated space, as made explicit by Pierre Albin and the second author. On the resolution the lifted action has fixed isotropy type, in an iterated sense, with connecting fibrations and this structure descends to a resolution of the quotient. For an Abelian group action the equivariant K-theory can then be described in terms of bundles over the base with morphisms covering the connecting maps. A similar model is given, in terms of appropriately twisted de Rham forms over the base as an iterated space, for delocalized equivariant cohomology in the sense of Baum, Brylinski and MacPherson. This approach allows a direct proof of their equivariant version of the Atiyah–Hirzebruch isomorphism.

Mathematics Subject Classification (2010). 19L47, 55N91, 58E40.

Keywords. Equivariant, K-theory, resolution, radial blow up, group action, Atiyah–Hirzebruch isomorphism.

Introduction

One intention of this note is to demonstrate that real blow-up can be an effective tool in the analysis of smooth group actions, particularly in the compact case. To do so, we describe equivariant K-theory in terms of resolved spaces and in consequence introduce (here only in the Abelian case) a geometric model for the delocalized equivariant cohomology of Baum, Brylinski and MacPherson [2], designed to realize an equivariant form of the Atiyah–Hirzebruch isomorphism

$$\mathrm{Ch} : K^*_G(M) \otimes \mathbb{C} \xrightarrow{\simeq} H^*_{\mathrm{dl},G}(M). \tag{1}$$

The more general case of the action by a non-Abelian compact Lie group will be treated subsequently. That the non-Abelian case is more intricate can be seen from the computation of the equivariant K-theory in case of an action with single isotropy type by Wassermann [7]. See also the paper of Rosu [5].

Resolution of a group action, as described by Pierre Albin and the second author in [1], replaces it by a tree of actions each with unique isotropy type and with connecting equivariant fibrations. This results in a similar resolution of the quotient, which we call an 'iterated space' corresponding to its smooth stratification. The description given here of the various cohomology theories is directly in terms of smooth 'iterated' objects, bundles or forms, over these iterated spaces with augmented 'pull-back' morphisms covering the connecting fibrations. Resolution may be thought of as replacing the 'analytic complexity' of strata by the 'combinatorial complexity' of iterated fibrations. The perceived advantage of this is that many standard arguments can be transferred directly to this iterated setting, since the spaces are smooth. The objects which appear here have local product structures.

The case of a compact Abelian group, G, acting, with single isotropy group, on a compact manifold (with corners), M, is relatively simple and forms the core of our iterative approach.

If the action is free then each equivariant bundle is equivariantly isomorphic to the pull-back of a bundle over the base; thus equivariant bundles descend to bundles. Equivariant K-theory is then identified, as a ring, with the ordinary K-theory of the base. However the structure of $K_G(M)$ as a module over the representation ring of G is lost in this identification. With $\widehat{G}$ the dual group, tensor product and descent defines an action of irreducible representations of G on smooth bundles over the base

$$\sigma : \widehat{G} \times \mathfrak{B}\mathrm{un}(Y) \longrightarrow \mathfrak{B}\mathrm{un}(Y) \tag{2}$$

which projects to give the action of $\widehat{G}$ on $K_G(M)$. In realizing equivariant K-theory and delocalized equivariant cohomology over the resolved space we need to retain aspects of σ.

For an Abelian action with fixed, but non-trivial, isotropy group $B \subset G$ there is a similar reduction to objects on the base. Equivariant bundles may be decomposed over the dual group, $\widehat{B}$, giving a finite number of coefficient bundles. Lifting an element of $\widehat{B}$ into $\widehat{G}$ and taking the tensor product with the inverse gives the coefficient bundle an action of $\widehat{G/B}$. The case of a principal action then applies and results in a collection of bundles $W_{\widehat{g}}$ over the base, Y, indexed by $\widehat{g} \in \widehat{G}$. We assemble these into a bundle over $\widehat{G} \times Y$ – allowed to have different dimensions over different components – with two additional properties. First its support projects to a finite subset of $\widehat{B}$ and more significantly it is 'twisted' under the action of $\widehat{G/B}$ on $\widehat{G}$ in the sense that

$$\sigma(\widehat{h}) \otimes W_{\widehat{h}\widehat{g}} = W_{\widehat{g}}. \tag{3}$$

In this setting of a single isotropy group, the delocalized equivariant cohomology is given in terms of a twisted de Rham complex. The forms are finite sums of formal products

$$\sum_i \widehat{g}_i \otimes u_i,\ u_i \in \mathcal{C}^\infty(Y; \Lambda^*)$$

where the twisting law (3) is replaced by its cohomological image

$$(\widehat{h}\widehat{g}) \otimes \mathrm{Ch}(\widehat{h}) \wedge v \simeq \widehat{g} \otimes v,\ \widehat{h} \in \widehat{G/B},\ v \in \mathcal{C}^\infty(Y;\Lambda^*). \tag{4}$$

Here $\mathrm{Ch}(\widehat{h})$ is the Chern character of the bundle, with connection, given by descent from the representation $\widehat{h}$ interpreted as a trivial bundle with equivariant action and with product connection. The reduced bundles may be given connections, consistent with the connection on $\widehat{h}$ and (3) for which the Chern character is a delocalized form in the sense of (4). For discussions of the equivariant Chern character see the book [3] of Berline, Getzler and Vergne and the paper of Getzler [4].

These definitions of reduced bundles and delocalized de Rham forms are extended to iterated objects over the resolution of the quotient, Y_*, by adding morphisms covering the connecting fibrations. This leads directly to the Atiyah–Hirzebruch–Baum–Brylinski–MacPherson isomorphism (1), proved here using the six-term exact sequences which result from successive pruning of the isotropy tree.

In outline the paper proceeds as follows. In §1 we recall from [1] the resolution X_* of any compact Lie group action on a compact manifold, with the quotient an iterated space Y_*. The lifting of equivariant bundles to iterated equivariant bundles on X_* is described in §2 and the reduction to twisted iterated bundles over Y_* is discussed, for Abelian actions, in §3 – the non-Abelian case is much more intricate because of the appearance of 'Mackey twisting'. The realization of equivariant K-theory in terms of reduced bundles is contained in §4 and this leads to the geometric model for delocalized (Abelian) equivariant cohomology in §5. The relative sequences obtained by successive pruning of the isotropy tree are introduced in §6 and used to establish (1) in §7. Examples of circle actions are considered in §8.

The authors are happy to thank Pierre Albin, Victor Guillemin, Eckhard Meinrenken, Michèle Vergne and David Vogan for helpful conversations.

1. Resolution

In [1], the resolution of the smooth action of a compact Lie group on a compact manifold, M, was described. The action stratifies M, into smooth submanifolds, with the isotropy group lying in a fixed conjugacy class of closed subgroups of G on each stratum. For convenience we shall assume, without loss of generality, that the quotient, M/G, is connected. If M is not connected then G acts on the set of components and we may consider each orbit separately and so assume that G acts transitively on the set of components. Similarly, we declare the strata, $M_\alpha \subset M$, to consist of the images under the action of G of the individual components of the manifolds where the isotropy class is fixed. Thus the labeling index, $\alpha \in A$, records a little more than the isotropy type since different (collective) components may have the same isotropy type. The strata are partially ordered, by the condition that the closure of one is contained in the other. This partial order is consistent with the codimension of the strata and the strata containing a point of M in their closures form a chain in both senses, under inclusion and under the order corresponding to codimension.

We recall both the resolution of such a group action and the consequent resolution of the quotient in terms of 'iterated spaces'. This is essentially the notion of a 'resolved stratified space'.

For present purposes the category $\mathfrak{Man}$ has as objects the *compact* manifolds with corners, not necessarily connected. Each such manifold has a finite collection, $\mathcal{M}_1(M)$, of boundary hypersurfaces $H \subset M$. By definition of a 'manifold with corners' we require that these boundary hypersurfaces are embedded – they are themselves manifolds with corners having no boundary faces identified in M. As a result each boundary hypersurface has a global defining function $0 \leq \rho \in \mathcal{C}^\infty(M)$, vanishing simply and precisely on H. As morphisms we will take 'smooth interior b-maps' which is to say smooth maps in the usual sense $M_1 \longrightarrow M_2$ such that the pull-back of a boundary defining function for a boundary hypersurface of M_2 is the product of powers of boundary defining functions for hypersurfaces of M_1 (including the case that the pull-back is strictly positive). Certainly all smooth diffeomorphisms are interior b-maps. A smooth G action on X is required to be *boundary free* in the sense that

$$g \in G,\ H \in \mathcal{M}_1(M) \Longrightarrow \text{ either } gH = H \text{ or } gH \cap H = \emptyset. \tag{1.1}$$

In fact the morphisms we are most concerned with here are fibre bundles, which we call 'fibrations'. In this compact context, these are simply the surjective interior b-maps with surjective differentials. The implicit function theorem applies to show that for such a map each point in the base has an open neighbourhood U with inverse image diffeomorphic to the product $U \times Z$ with Z a fixed (over components of the base) compact manifold with corners with the map reducing to projection. Note that the b-map condition is used here; without such an assumption the fibres can be cut off by boundaries.

Definition 1. The category, $\mathfrak{IMan}$, has as objects, X_*, *iterated spaces* in the following sense. There is a 'principal' manifold with corners $X = X_0$ which is the root of a tree X_α of manifolds corresponding to a partial order ('depth') $\alpha \leq \beta \in A$. The boundary hypersurfaces of X_0 are partitioned into subsets, with elements which do not intersect, forming 'collective boundary hypersurfaces' $H_\alpha(X_0) \subset \mathcal{M}_1(X_0)$. These carry fibrations

$$\psi_\alpha : H_\alpha(X_0) \longrightarrow X_\alpha. \tag{1.2}$$

Under the partial order on the H_α two (always collective) hypersurfaces are related if and only if they intersect, and any collection with non-trivial total intersection forms a chain. For each α the set of boundary hypersurfaces of X_α is also partitioned into collective boundary hypersurfaces

$$H_\beta(X_\alpha) = \psi_\alpha(H_\beta),\ \beta > \alpha \tag{1.3}$$

and ψ_β restricted to H_α factors through a fibration

$$\psi_{\beta,\alpha} : H_\beta(X_\alpha) \longrightarrow X_\beta,\ \ \beta > \alpha; \tag{1.4}$$

with the base index denoted 0, $\psi_\alpha = \psi_{\alpha,0}$.

A *smooth G-action* on an iterated space is a boundary free G action on each X_α with respect to which all the fibrations $\psi_{\alpha,\beta}$ are G-equivariant.

It follows that in an iterated space, for any chain

$$\alpha_1 < \alpha_2 < \cdots < \alpha_k \tag{1.5}$$

there is a sequence of fibrations

$$\bigcap_{1\le j\le k} H_{\alpha_j}(X_0) \xrightarrow{\psi_{\alpha_1}} \bigcap_{2\le k\le k} H_{\alpha_j}(X_{\alpha_1}) \xrightarrow{\psi_{\alpha_2,\alpha_1}} \cdots \longrightarrow H_{\alpha_k}(X_{\alpha_{k-1}}) \xrightarrow{\psi_{\alpha_k,\alpha,k-1}} X_{\alpha_k} \tag{1.6}$$

with composite the restriction of ψ_{α_k}. It is also follows that the fibres of the restricted fibrations have strictly increasing codimension as submanifolds of the fibres in the hypersurfaces.

Resolution is accomplished in [1] by radial blow up (which corresponds to a sequence of interior b-maps) of successive smooth centres corresponding to the tree of isotropy types, in (any) order of decreasing codimension. This results in a well-defined iterated space, X_*, with G-action in the sense described above with principal space $X = X_0$ and iterated blow-down map

$$\beta : X \longrightarrow M \tag{1.7}$$

giving the resolution of M. The X_α are the resolutions of the isotropy types $\overline{M_\alpha}$ in the same sense. The important property of the resolution is that the G-action on each (smooth, compact) X_α now has fixed isotropy type and the 'change of isotropy type' occurs within the fibrations $\psi_{\alpha,\beta}$.

Since the action on each X_α has fixed isotropy type the quotients

$$Y_\alpha = X_\alpha/G \tag{1.8}$$

are all smooth manifolds with corners having boundary hypersurfaces $H_\beta(Y_\alpha)$, $\beta > \alpha$, labeled by the index set

$$A_\alpha = \{\beta \in A; \beta > \alpha\} \tag{1.9}$$

and forming a tree with the corresponding intersection relations and base α. The G-equivariant fibrations (1.4) descend to give Y_* the structure of an iterated space

$$\begin{tikzcd} H_\beta(X_\alpha) \arrow[r, "/G"] \arrow[d, "\psi_{\beta,\alpha}"'] & H_\beta(Y_\alpha) \arrow[d, "\phi_{\beta,\alpha}"] \\ X_\beta \arrow[r, "/G"'] & Y_\beta \end{tikzcd} \ ,\ \beta > \alpha,\ \phi_\alpha = \phi_{\alpha,0}. \tag{1.10}$$

2. Lifting

Let $\mathfrak{Bun}(M)$ denote the category of finite-dimensional, smooth, complex, vector bundles over a compact manifold M, with bundle maps as morphisms. Similarly if M is a smooth G-space let $\mathfrak{Bun}_G(M)$ denote the category of bundles with equivariant G-action covering the action on M and with morphisms the bundle maps intertwining the actions. Thus the equivariant K-theory of M can be realized (see Segal [6]) as the Grothendieck group

$$K_G(M) = \mathfrak{Bun}_G(M) \ominus \mathfrak{Bun}_G(M)/ \simeq \tag{2.1}$$

with the relation of stable G-equivariant bundle isomorphism.

In general if $F : M \longrightarrow N$ is a smooth G-equivariant map of G-spaces then pull-back defines a functor

$$F^* : \mathfrak{Bun}_G(N) \longrightarrow \mathfrak{Bun}_G(M). \tag{2.2}$$

In particular this applies to the blow-down map in the resolution of the action.

Definition 2. If X_* is an iterated space we denote by $\mathfrak{Bun}(X_*)$ the category with objects 'iterated bundles' consisting of a bundle $B_\alpha \in \mathfrak{Bun}(X_\alpha)$ for each $\alpha \in A$ and with pull-back isomorphisms specified over each $H_\alpha(X_0)$,

$$\mu_\alpha : \phi_\alpha^* B_\alpha \simeq B_0\big|_{H_\alpha(X_0)} \tag{2.3}$$

which factor through intermediate bundle isomorphisms $\mu_{\alpha,\beta}$, $\alpha < \beta$, covering the sequence (1.6) over each boundary face of X_0. The morphisms are bundle maps between the corresponding bundles which commute with the connecting morphisms (2.3).

If X_* is an iterated space with G-action, $\mathfrak{Bun}_G(X_*)$ denotes the category in which the bundles carry G-actions covering the actions on the X_α and the connecting isomorphisms, (2.3), are G-equivariant; morphisms are then required to be G-equivariant.

Lemma 1. *If the iterated G-space X_* is the resolution of M, with compact G action, then pull-back under the iterated blow-down map defines a functor*

$$\beta^* : \mathfrak{Bun}_G(M) \longrightarrow \mathfrak{Bun}_G(X_*) \tag{2.4}$$

and every iterated bundle in $\mathfrak{Bun}_G(X_*)$ is isomorphic to the image of a bundle in $\mathfrak{Bun}_G(M)$.

Proof. The lifting of the objects, G-equivariant bundles, and corresponding morphisms under β is simply iterated pull-back. It only remains to show that every G-equivariant iterated bundle in $\mathfrak{Bun}_G(X_*)$ is isomorphic to such a pull-back. As shown in [1] the resolution X_* can be 'rigidified' by choosing product decompositions near all boundary hypersurfaces with G-invariant smooth defining functions consistent near all corners, i.e., so that the various retractions commute.

In the simple setting of a compact manifold with boundary, M, suppose V is a smooth vector bundle over M, U is a vector bundle over the boundary H and $T : V\big|_H \longrightarrow U$ is a bundle isomorphism. Then V can be modified near H

to an isomorphic bundle $\tilde{V}$ which has fibres over H identified with those of U and outside a small collar neighbourhood of H has fibres identified with V. This can be accomplished by a rotation in the isomorphism bundle of $V \oplus U$ and in particular carries over to the equivariant case. Indeed the standard construction has the virtue of leaving the original bundle unchanged over any set in the collar over an open set on which T is already an identification. This allows the bundle isomorphisms to be 'removed' inductively over the isotropy tree.

Once the isomorphisms are reduced to the identity, the bundles themselves can be similarly modified in equivariant collars around the boundary hypersurfaces of X_0 to be constant along the normal fibrations and hence to be the pull-backs of smooth bundles on the base. Alternatively the topological bundles obtained by direct projection can be smoothed over M. □

Pulling back a G-connection from a bundle on M we find:

Corollary 2. *For a G-bundle $W_* \in \mathfrak{Bun}_G(X_*)$ there are a G-equivariant connections on each W_α which are intertwined by the $\mu_{\alpha,\beta}$.*

Now, we can therefore identify

$$K_G(M) = \mathfrak{Bun}_G(X_*) \ominus \mathfrak{Bun}_G(X_*)/\simeq \tag{2.5}$$

as the Grothendieck group of iterated G-bundles on the resolution up to stable isomorphism.

Finite-dimensional representations of a compact Lie group, G, can be decomposed into direct sums of tensor products with respect to a fixed set $\widehat{G}$ of irreducibles, which can be identified with the set of characters. This allows the representation category to be identified with the category $\mathfrak{Bun}_c(\widehat{G})$, with objects the 'bundles' over $\widehat{G}$ with finite support and morphisms being bundle maps. Here, for a non-connected space, the objects in $\mathfrak{Bun}_c$ are permitted to have different dimensions over different components but in this case, where there may be infinitely many components, the bundles must have dimension zero in the complement of a compact set. So the objects consist of a finite number of characters, each associated to a (complex) vector space. Each object in $\mathfrak{Bun}_c(\widehat{G})$ defines an equivariant bundle over any G-space and tensor product with these bundles induces an action of the representation ring, $R(G) = \widehat{G}(\mathbb{Z})$ on $K_G(M)$. Aspects of this action are particularly important in the sequel.

Proposition 3. *For the action of a compact Lie group on an iterated space X_*, taking the tensor product with a (finite-dimensional) representation gives a functor*

$$\sigma : \mathfrak{Bun}(\widehat{G}) \times \mathfrak{Bun}_G(X_*) \longrightarrow \mathfrak{Bun}_G(X_*). \tag{2.6}$$

Proof. Given an element $(V, E) \in \mathfrak{Bun}(\widehat{G}) \times \mathfrak{Bun}_G(X_*)$, the corresponding object in $\mathfrak{Bun}_G(X_*)$ is the tensor product of E and V, with V thought as the trivial iterated bundle over X_* with the implied G-action. Given an element V of $\mathfrak{Bun}(\widehat{G})$ and an equivariant iterated bundle map $E \to F$, we obtain an equivariant bundle map $V \otimes E \to V \otimes F$. Similarly for a morphism of representations $V_1 \to V_2$ and an equivariant iterated bundle E, we obtain an induced equivariant bundle map $V_1 \otimes E \to V_2 \otimes E$. □

3. Reduction

The Abelian case is considerably simpler than the general one and has been more widely studied. From this point on, in this paper, we shall assume that G is compact and Abelian. One fundamental simplification is that all (complex) irreducible representations in the Abelian case are one-dimensional (and of course all one-dimensional representations are irreducible). In this case $\widehat{G}$ is a discrete Abelian group.

As recalled in the Introduction, if a compact Lie group acts freely on a compact manifold, X, then the quotient, Y, is a compact manifold and X is a principal bundle over it. For an equivariant bundle over X, the action over each orbit gives descent data for the bundle, defining a vector bundle over the base. This gives an equivalence of categories

$$\mathfrak{Bun}_G(X) \cong \mathfrak{Bun}(Y) \text{ if } G \text{ acts freely.} \tag{3.1}$$

For such a free action, tensor product with representations gives a 'quantization' of the dual group

$$\sigma : \widehat{G} \longrightarrow \mathfrak{Bun}(Y) \tag{3.2}$$

corresponding to (2.6).

We need to understand this operation in the more general case of an action with a fixed isotropy group $B \subset G$, necessarily a closed subgroup. There is then a short exact sequence

$$B \longrightarrow G \longrightarrow G/B \tag{3.3}$$

that is split since the groups are Abelian. The dual sequence

$$\begin{tikzcd} \widehat{G/B} \arrow[r] & \widehat{G} \arrow[r, shift left] & \widehat{B} \arrow[l, shift left, "\tau"] \end{tikzcd} \tag{3.4}$$

is also exact and split so there exists a group homomorphism τ as indicated, giving a right inverse. Two such maps τ, τ' are related by a group homomorphism

$$\mu : \widehat{B} \longrightarrow \widehat{G/B}. \tag{3.5}$$

with $\tau' = m(\mu, \tau)$, where $m : \widehat{G/B} \times \widehat{G} \longrightarrow \widehat{G}$ is the multiplication map.

For an action with isotropy group B, the quotient G/B acts freely on X and the discussion above gives the equivalence of categories and shift functor

$$\mathfrak{Bun}_{G/B}(X) \cong \mathfrak{Bun}(Y),\ \sigma : \widehat{G/B} \longrightarrow \mathfrak{Bun}(Y),$$
$$Y = X/G = X/(G/B). \tag{3.6}$$

It is still the case that G-equivariant bundles descend to the quotient but only after decomposition under the action of B. Consider the space $\widehat{G} \times Y$, which has a natural action by $\widehat{G/B}$, with $\widehat{g} \times Y$ mapped to $(\widehat{h} \otimes \widehat{g}) \times Y$.

Definition 3. For a compact Abelian G-action on a compact manifold X with fixed isotropy group $B \subset G$ and base $Y = X/G$, let $\mathfrak{Bun}_c^B(\widehat{G} \times Y)$ denote the category of bundles over $\widehat{G} \times Y$ with support which is finite when projected to $\widehat{B}$ and which satisfy the transformation law

$$\sigma(\widehat{h}) \otimes W_{\widehat{h} \otimes \widehat{g}} = W_{\widehat{g}} \ \forall \ \widehat{h} \in \widehat{G/B},\ \widehat{g} \in \widehat{G}. \tag{3.7}$$

Morphisms are bundle maps over each $\widehat{g} \times Y$ which are natural with respect to (3.7).

Note that we could eliminate the action (3.7) at the expense of choosing a splitting group homomorphism $\tau : \widehat{B} \longrightarrow \widehat{G}$ as in (3.4), reducing elements of $\mathfrak{Bun}_c^B(\widehat{G} \times Y)$ to arbitrary elements of $\mathfrak{Bun}_c(\widehat{B} \times Y)$.

Proposition 4. *For an action of a compact Abelian group with fixed isotropy group $B \subset G$ there is an equivalence of categories*

$$R : \mathfrak{Bun}_G(X) \cong \mathfrak{Bun}_c^B(\widehat{G} \times Y),\ Y = X/G \tag{3.8}$$

where $W \in \mathfrak{Bun}_c^B(\widehat{G} \times Y)$ corresponds to the G-equivariant bundle

$$\bigoplus_{\widehat{b} \in \widehat{B}} \tau(\widehat{b}) \otimes \pi^*(W_{\tau(\widehat{b})}) \tag{3.9}$$

for a splitting homomorphism τ as in (3.4).

Elements of $\mathfrak{Bun}_c^B(\widehat{G} \times Y)$ are our 'reduced bundles' in this simple case.

In order to define (3.9) we pass to the restriction of an element of $\mathfrak{Bun}_c^B(\widehat{G} \times Y)$ to the image $\tau(\widehat{B}) \times U$ given by a splitting homomorphism τ. As a consequence of (3.7) the final result is independent of the choice of τ.

Proof. The isotropy group at each point acts on the fibres of an equivariant bundle $U \in \mathfrak{Bun}_G(X)$ which therefore decomposes into a direct sum of B-equivariant bundles

$$U = \bigoplus_{\widehat{b} \in \widehat{B},\text{finite}} U_{\widehat{b}} \tag{3.10}$$

where the action of B on each term factors through the irreducible representation $\widehat{b}$. If $\widehat{g} \in \widehat{G}$ is a representation which restricts to $\widehat{b}$ then the action of B on $\widehat{g}^{-1} \otimes U_{\widehat{b}}$ is trivial. This bundle therefore has an equivariant G/B-action and so descends to a bundle $W_{\widehat{g}}$ over Y. Doing this for every $\widehat{g}$, we define a bundle over $\widehat{G} \times Y$, supported over a finite subset of $\widehat{B}$. Clearly these bundles satisfy (3.7). Conversely each element of $\mathfrak{Bun}_c^B(\widehat{G} \times Y)$ defines an element of $\mathfrak{Bun}_G(X)$. □

Note that the category $\mathfrak{Bun}_c^B(\widehat{G} \times Y)$ is not determined by the groups and base Y alone since it depends on the 'shift' isomorphism σ which retains some information about the principal bundle, namely the images under descent to Y of the trivial G-bundles given by elements of $\widehat{G/B}$.

4. Reduced K-theory

Consider next a principal G-bundle, for G compact Abelian, and a G-equivariant fibration giving a commutative diagram

$$\begin{tikzcd} X \arrow[r, "\pi"] \arrow[d, "/G"'] & X_1 \arrow[d, "/G"] \\ Y \arrow[r, "\pi_1"'] & Y_1 \end{tikzcd} \tag{4.1}$$

where the G-action on X_1 has fixed isotropy group B; thus π_1 is a fibration of smooth compact manifolds. In view of the identification of equivariant bundles in Proposition 4, the pull-back map descends to an 'augmented pull-back map'

$$\begin{tikzcd} \mathfrak{B}\mathrm{un}_G(X) \arrow[d, hook] & \mathfrak{B}\mathrm{un}_G(X_1) \arrow[l, "\pi^*"'] \arrow[d, hook] \\ \mathfrak{B}\mathrm{un}(Y) & \mathfrak{B}\mathrm{un}_c^B(\widehat{G}\times Y_1) \arrow[l, "\pi_1^{\#}"'] \end{tikzcd} \tag{4.2}$$

given by pull-back followed by summation over a splitting $\tau : \widehat{B} \longrightarrow \widehat{G}$:

$$\pi_1^{\#} : \mathfrak{B}\mathrm{un}_c^B(\widehat{G}\times Y_1) \xrightarrow{\pi_1^*} \mathfrak{B}\mathrm{un}_c^B(\widehat{G}\times Y) \xrightarrow{\sigma^{\#}} \mathfrak{B}\mathrm{un}(Y) = \mathfrak{B}\mathrm{un}_c^{\{e\}}(\widehat{G}\times Y),$$
$$\sigma^{\#}(V) = \bigoplus_{\widehat{b}\in\widehat{B}} \sigma(\tau(\widehat{b}))V_{\tau(\widehat{b})}.$$

As implicitly indicated by the notation, $\sigma^{\#}(V)$ is independent of the section τ.

We need this in the more general case of an equivariant fibration between two actions with fixed isotropy groups. For nested closed subgroups, $K \subset B \subset G$, we choose iterated splittings

$$\tau' : \widehat{K} \longrightarrow \widehat{B},\ \tau_1 : \widehat{B} \longrightarrow \widehat{G} \Longrightarrow \tau = \tau_1\tau' : \widehat{K} \longrightarrow \widehat{G}. \tag{4.3}$$

Then

$$\widehat{b}',\ \widehat{b} \in \widehat{B},\ \widehat{b}'\big|_K = \widehat{b} \Longleftrightarrow\ \exists\,!\ \widehat{h} \in \widehat{G/K} \text{ s.t. } \tau_1(\widehat{b}') = \widehat{h}\tau(\widehat{b}). \tag{4.4}$$

Proposition 5. *If* (4.1) *is an equivariant fibration between actions of a compact Abelian Lie group G with fixed isotropy groups $B \supset K$ then pull back of equivariant bundles descends to the augmented pull-back map*

$$\pi_1^{\#} : \mathfrak{B}\mathrm{un}_c^B(\widehat{G}\times Y_1) \longrightarrow \mathfrak{B}\mathrm{un}_c^K(\widehat{G}\times Y) \tag{4.5}$$

given by pull back on the fibres

$$\pi_1^{*} : \mathfrak{B}\mathrm{un}_c^B(\widehat{G}\times Y_1) \longrightarrow \mathfrak{B}\mathrm{un}_c^B(\widehat{G}\times Y) \tag{4.6}$$

followed by summation to give the value at the image $\tau(\widehat{k})$ *using* (4.4)

$$\begin{gathered}\left(\sigma^{\#}(V)\right)_{\tau(\widehat{k})} = \bigoplus_{\{\widehat{b}\in\widehat{B};\widehat{b}\big|_K=\widehat{k}\}} \sigma(\widehat{h})V_{\widehat{h}\tau'(\widehat{k})},\\ \forall\ V \in \mathfrak{B}\mathrm{un}_c^B(\widehat{G}\times Y),\ \widehat{k}\in\widehat{K}.\end{gathered} \tag{4.7}$$

Proof. An equivariant fibration can be factored through the fibre product

$$\tilde{X}_1 = X_1 \times_{Y_1} Y \longrightarrow Y,\ X \longrightarrow \tilde{X}_1 \longrightarrow X_1$$

where the G-action on $\tilde{X}_1$ has isotropy group B. Thus it suffices to consider the two cases of the pull-back of an action under a fibration and the quotient of an action with isotropy group K by a larger subgroup B. In the first case the augmented pull-back is simply the pull-back as in (4.6) with (4.7) being the identity. In the second case the base is unchanged, so (4.6) is the identity and the summation is over those elements of $\widehat{B}$ with fixed restriction to K. □

Now we pass to the general case of the action of a compact Abelian Lie group G on a compact manifold M with resolution X_* and resolved quotient Y_* as discussed above. The isotropy groups $B_\alpha \subset G$ form a tree with root B_0 the principal isotropy group. Generalizing the choice (4.3) we can choose iterative splittings by proceeding step-wise along chains

$$\tau_{\beta,\alpha} : \widehat{B_\alpha} \longrightarrow \widehat{B_\beta}\ \forall\ \beta > \alpha,\ \tau_{\gamma,\beta}\circ\tau_{\beta,\alpha} = \tau_{\gamma,\alpha},\ \gamma > \beta > \alpha. \tag{4.8}$$

Using notation as for the fibration maps we set $\tau_\alpha = \tau_{\alpha,0}$. Then the formulæ (4.5) and (4.7) are valid for any pair and are consistent along chains.

Definition 4. Reduced bundles W_* in the case of an Abelian action, consist of the following data

1. A bundle $W_\alpha \in \mathfrak{B}\mathrm{un}_c^{B_\alpha}(\widehat{G}\times Y_*)$ for each element of the tree.
2. For each non-principal isotropy type $\alpha > 0$ (so $B_\alpha \supset B_0$) a bundle isomorphism
$$T_\alpha : \pi_\alpha^{\#} W_\alpha \simeq W_0\big|_{H_\alpha(Y_0)}. \tag{4.9}$$
3. The consistency conditions that for any chain α_*, $\alpha_k > \cdots > \alpha_1 > 0$ the isomorphisms (4.9) restricted to the boundary face, of codimension k,
$$H_{\alpha_*}(Y_0) = \bigcap_j H_{\alpha_j}(Y_0)$$
form a chain, corresponding to isomorphisms for each $\alpha < \beta$
$$T_{\alpha,\beta} : \pi_{\alpha\beta}^{\#} W_\beta \simeq W_\alpha\big|_{H_\beta(Y_\alpha)}. \tag{4.10}$$

Morphisms between such data consist of bundle maps at each level of the tree intertwining the isomorphisms T_α in (4.9).

We denote by $\mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G}\times Y_*)$ the category of such reduced bundles and the corresponding Grothendieck group of pairs of reduced bundles up to stable isomorphism by

$$K_{\mathrm{red}}(Y_*)=\mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G}\times Y_*)\ominus\mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G}\times Y_*)/\simeq. \tag{4.11}$$

Theorem 6. *The equivariant K-theory for the action of a compact Abelian group on a compact manifold M is naturally identified with the reduced K-theory* (4.11) *of the resolved quotient.*

Proof. This follows from the equivalence of the categories of G-equivariant iterated bundles over X_* and reduced bundles over Y_* which in turn follows from Propositions 4 and 5. □

Definition 5. An iterated connection ∇_* on a reduced bundle $W_*\in\mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G}\times Y_*)$ is a connection $\nabla_{\widehat{g},\alpha}$ on each bundle $W_{\widehat{g},\alpha}\in\mathfrak{B}\mathrm{un}(Y_\alpha)$ satisfying

$$\nabla_{\widehat{h}}\otimes\nabla_{\widehat{h}\widehat{g},\alpha}=\nabla_{\widehat{g},\alpha},\ \widehat{h}\in\widehat{G/B_\alpha}\ \forall\,\alpha\in A,\ \widehat{g}\in\widehat{G} \tag{4.12}$$

under the transformation law (3.7) and compatible under augmented pull-back isomorphisms.

Lemma 7. *Any reduced bundle can be equipped with an iterated connection in the sense of Definition* 5*.*

Proof. Such a connection can be obtained following the reduction procedure from a G-connection on the corresponding iterated G-bundle over X_*. It is also straightforward to construct such a connection directly. □

The odd version of reduced bundles may be defined by 'suspension' – simply taking the product with an interval and demanding that all bundles be trivialized over the end points leading to a category

$$\mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G}\times([0,1]\times Y_*;(\{0\}\cup\{1\})\times Y_*). \tag{4.13}$$

This leads to the odd version of equivariant K-theory

$$\begin{aligned}&K^1_G(M)=K^1_{\mathrm{red}}(Y_*)=\\&\quad\mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G}\times Y_*;(\{0\}\cup\{1\})\times Y_*)\ominus\mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G}\times Y_*;(\{0\}\cup\{1\})\times Y_*)/\simeq.\end{aligned} \tag{4.14}$$

The isotropy tree can also be 'pruned' by considering any subtree $P\subset A$, so

$$\alpha,\ \beta\in P,\ \alpha\le\beta\Longrightarrow\beta\in P. \tag{4.15}$$

Then reduced bundles which are trivialized on the elements of P form a subcategory

$$\mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G}\times Y_*;P). \tag{4.16}$$

These correspond to G bundles over M which are trivialized in a neighborhood of isotropy types indexed by P. We denote by $K^*_{\mathrm{red}}(Y_*;P)$ the Grothendieck groups of these relative spaces of bundles and their suspended versions.

5. Delocalized equivariant cohomology

If $\rho \in \widehat{G}$ is an irreducible representation of a compact Abelian group on a complex line, E, then, the corresponding trivial line bundle over a G-space, X, is G-equivariant,

$$E \in \mathfrak{B}\mathrm{un}_G(X). \tag{5.1}$$

The de Rham differential defines a G-equivariant connection on E. If the action of G is free, so $X \longrightarrow Y$ is a principal G-bundle, then E descends to a bundle, $\widetilde{E}$, with connection. The Chern character therefore defines a multiplicative map

$$\mathrm{Ch} : \widehat{G} \longrightarrow \mathcal{C}^\infty(Y; \Lambda^{2*}). \tag{5.2}$$

Let $R(G)$ be the representation algebra with complex coefficients, so the vector space of formal finite linear combinations of elements of $\widehat{G}$. Then the map (5.2) extends to a map of algebras

$$\mathrm{Ch} : R(G) \longrightarrow \mathcal{C}^\infty(Y; \Lambda^{2*}). \tag{5.3}$$

For a closed subgroup, $B \subset G$, $R(G/B) \longrightarrow R(G)$ gives a multiplicative action

$$R(G/B) \times R(G) \longrightarrow R(G). \tag{5.4}$$

This and (5.2), for G/B lead to:

Definition 6. For a compact Abelian group G acting with fixed isotropy group B on a compact manifold X the space of twisted forms over the base Y is defined as

$$\mathcal{C}^\infty(Y; \Lambda^*_{\mathrm{dl}}) = \mathcal{C}^\infty(Y; \Lambda^*) \otimes_{\mathrm{Ch}} R(G). \tag{5.5}$$

Thus an element of this space is a finite linear combination of formal products

$$u_i \otimes \widehat{g}_i,\ u_i \in \mathcal{C}^\infty(Y; \Lambda^*),\ \widehat{g}_i \in \widehat{G}$$

under the equivalence relation

$$u \otimes \widehat{g} \simeq (\mathrm{Ch}(\widehat{h}) \wedge u) \otimes \widehat{h}\widehat{g},\ \forall\ \widehat{h} \in \widehat{G/B},\ u \in \mathcal{C}^\infty(Y; \Lambda^*). \tag{5.6}$$

Since the Chern character is closed, the de Rham differential descends

$$d : \mathcal{C}^\infty(Y; \Lambda^*_{\mathrm{dl}}) \longrightarrow \mathcal{C}^\infty(Y; \Lambda^*_{\mathrm{dl}}),\ d^2 = 0. \tag{5.7}$$

Lemma 8. *Suppose that $\pi_1 : X \longrightarrow X_1$ is a G-equivariant fibration for actions with fixed isotropy groups $K \subset B$ and $\tilde{\pi}_1 : Y \longrightarrow Y_1$ is the induced fibration, then there is a natural augmented pull-back*

$$\pi_1^\# : \mathcal{C}^\infty(Y_1; \Lambda^*_{\mathrm{dl}}) \longrightarrow \mathcal{C}^\infty(Y; \Lambda^*_{\mathrm{dl}}) \tag{5.8}$$

which intertwines the action of d.

Proof. To define (5.8) it suffices to consider three elementary cases.

First suppose that π is simply an isomorphism of principle bundles covering the identity map $\tilde{\pi}_1$. The only appearance of the bundle in (5.5), (5.6) is through the Chern character and this is invariant under such a transformation.

Secondly suppose that $K = B$ but that π_1 is a G-equivariant fibration. Then, after a bundle isomorphism, this corresponds to X_1 being the pull-back of the principal G/B bundle over Y_1 under a fibration $\tilde{\pi}_1$. The bundles E corresponding to representations of G/B and their connections pull back naturally and in this case (5.8) corresponds to the pull-back of the coefficient forms.

Finally then consider the case that X is a principal G/K bundle and that $K \subset B \subset G$ is a second closed subgroup with

$$\pi_1 : X \longrightarrow X_1 = X/B, \tag{5.9}$$

so $Y = Y_1$. The equivalence relation (5.6), now for $\widehat{i} \in \widehat{G/B}$ means that any element of $\mathcal{C}^\infty(X_1; \Lambda^*_{\mathrm{dl}})$ can be represented by a finite sum

$$u_i \otimes \widehat{g}_i \tag{5.10}$$

where the $\widehat{g}_i \in \widehat{G}$ exhaust $\widehat{B}$ under restriction. These can be chosen, and relabeled, to be $\widehat{f}_{kj}\widehat{g}_j$ where the $\widehat{g}_j \in \widehat{G}$ restrict to exhaust $\widehat{K}$ and $\widehat{f}_{kj} \in \widehat{G/B}$. Then

$$\pi_1^\# : \sum_{\text{finite}} u_{jk} \otimes \widehat{f}_{kj}\widehat{g}_j = \sum_{\text{finite}} \left(\sum\nolimits_k \mathrm{Ch}(f_{kj})^{-1} \wedge u_{jk} \right) \otimes \widehat{g}_j. \tag{5.11}$$

For elements of $\widehat{G/B}$ the construction of the Chern character factors through the projection to X_1.

The general case corresponds to a composite of these three cases. □

Our model for the delocalized equivariant cohomology of Baum, Brylinski and MacPherson in the case of a smooth action of a compact Abelian Lie group G on a compact manifold M is the following data on the resolved quotient.

Definition 7. An element of the delocalized de Rham complex $\mathcal{C}^\infty(Y_*; \Lambda^*_{\mathrm{dl}})$ consists of

1. For each $\alpha \in A$ a twisted smooth form $u_\alpha \in \mathcal{C}^\infty(Y_\alpha; \Lambda^*_{\mathrm{dl}})$.
2. Compatibility conditions at all boundary faces

$$u_\alpha \big|_{H_\beta(Y_\alpha)} = \pi^\#_{\alpha\beta} u_\beta, \ \beta > \alpha. \tag{5.12}$$

including the boundary hypersurfaces of the principal quotient corresponding to $\alpha = 0$.

Again the relative versions corresponding to a subtree $P \subset A$, are similarly defined by demanding that the forms vanish over the spaces indexed by P.

If ∇_* is an iterated connection on an iterated bundle $W_* \in \mathfrak{B}\mathrm{un}_c^{B_*}(\widehat{G} \times Y_*)$, as in Definition 5 and Lemma 7 then the Chern character of each bundle W_α is a form on $\widehat{G} \times Y_\alpha$:

$$\mathrm{Ch}(W_\alpha, \nabla_\alpha)_{\widehat{g}} = \mathrm{Ch}((W_\alpha)_{\widehat{g}}, \nabla_\alpha) \text{ on } \{\widehat{g}\} \times Y_\alpha. \tag{5.13}$$

Proposition 9. *The Chern character of a reduced bundle with compatible connection is an element of* $\mathcal{C}^\infty(Y_*; \Lambda^*_{\mathrm{dl}})$.

Proof. The forms (5.13) shift correctly under the action of $\widehat{G/B}$ in view of the corresponding property for the connections and the iterative relations over the boundary fibrations similarly follow from the standard properties of the Chern character under pull-back. □

Definition 8. The *delocalized equivariant cohomology* $H^*_{\mathrm{dl},G}(M)$ of a compact manifold with smooth action by a compact Abelian group is identified with $H^*_{\mathrm{dl}}(Y_*)$, the de Rham cohomology of the complex $\mathcal{C}^\infty(Y_*;\Lambda^*_{\mathrm{dl}})$ of the reduced space; the relative cohomology $H^*_{\mathrm{dl}}(Y_*;P)$ with respect to a subtree is defined similarly.

That this cohomology theory is identified with that given by Baum, Brylinski and MacPherson can be proved directly but in any case follows from the isomorphism with $K_G(M)\otimes\mathbb{C}$ discussed below.

The verification of this 'Atiyah–Hirzebruch–Baum–Brylinski–MacPherson' isomorphism (7.1) is given in §7 below. The iterative proof uses the six-term exact sequences arising from pruning the isotropy tree at successive points of a sequence which is increasing with respect to the partial order. As with the whole approach here, this is based on reduction to the case of a fixed isotropy group where the result reduces in essence to the Atiyah–Hirzebruch isomorphism.

Proposition 10. *If G is a compact Abelian group acting on a compact manifold with fixed isotropy group B then the Chern character gives an isomorphism*

$$K_G(M)\otimes\mathbb{C}\longrightarrow H^{\mathrm{even}}_{\mathrm{dl},G}(M)=H^{\mathrm{even}}_{\mathrm{dl}}(Y_*). \tag{5.14}$$

Proof. The Atiyah–Hirzebruch isomorphism is valid rationally. This amounts to the two statements that for a compact manifold (with corners) the range of the Chern character

$$\mathrm{Ch}:K(Y)\longrightarrow H^{\mathrm{even}}(Y) \tag{5.15}$$

spans the cohomology (with complex coefficients) and that the null space consists of torsion elements. At the bundle level this means that if the Chern character for a pair of bundles $V_+\ominus V_-$ is exact then for some integers p and N there is an isomorphism

$$I:V_+^p\oplus\mathbb{C}^N\longrightarrow V_-^p\oplus\mathbb{C}^N. \tag{5.16}$$

A given connection on the $V_\pm$ lifts to a connection which can then be deformed to commute with I and so have zero Chern character.

Now, in the equivariant case we can consider a splitting homomorphism $\tau:\widehat{B}\longrightarrow\widehat{G}$ and then pull a pairs of bundles $V_\pm\in\mathfrak{Bun}_c^B(\tilde{G}\times Y)$ back to $\widehat{B}\times Y$ where the Chern character is given by

$$\sum_{\widehat{b}\in\widehat{B}}\tau(\widehat{g})\otimes(\mathrm{Ch}(V_{+,\tau(\widehat{b})})-\mathrm{Ch}(V_{-,\tau(\widehat{b})})). \tag{5.17}$$

The vanishing of the class $H^{\mathrm{even}}(Y;\Lambda_{\mathrm{dl}})$ is equivalent to the exactness of each of the de Rham classes $\mathrm{Ch}(V_{+,\tau(\widehat{b})})-\mathrm{Ch}(V_{-,\tau(\widehat{b})})$. Thus the vanishing of the Chern character in delocalized cohomology implies that each of the pairs $V_{\pm,\tau(\widehat{b})}$ is stably trivial in the sense of (5.16).

Since $\widehat{B}$ is finite and we may always further stabilize (5.16) by taking powers and adding trivial bundles, we may take p to be the product of the integers for each $\widehat{b}$ and similarly increase N. This however amounts to a stable trivialization of the whole bundle $V_+^p \ominus V_-^p$ as an element $\mathfrak{Bun}_c^B(\widehat{G} \times Y)$ and proves the injectivity of (5.14) in this case.

The surjectivity of (5.14) is a direct consequence of the Atiyah–Hirzebruch isomorphism and the definition of delocalized forms. □

The proof extends readily to give a relative version of this when we consider a particular element Y_α in the tree as the base and only consider bundles and forms which are trivial over 'deeper' faces Y_β, $\beta \geq \alpha$.

Proposition 11. *If $\alpha \in A$ for the action of a compact Abelian Lie group on a smooth compact manifold M then the augmented Chern character induces an isomorphism*

$$K_{\mathrm{red}}(Y_\alpha; A'_\alpha) \otimes \mathbb{C} \longrightarrow H^{\mathrm{even}}_{\mathrm{dl}}(Y_\alpha; A'_\alpha). \tag{5.18}$$

6. The relative sequences

The proof of Theorem 14 is based on induction over pruning and the six-term exact sequences which results from adding a minimal element $\alpha \in P$ in a subtree. Thus

$$P = \{\alpha\} \sqcup Q,\ Q = \{\beta \in P; \beta > \alpha\},\ \gamma \in P,\ \gamma \leq \alpha \Longrightarrow \gamma = \alpha. \tag{6.1}$$

The relative K-groups are introduced above.

Proposition 12. *For a minimal element in a subtree $\alpha \in P \subset A$ as in* (6.1) *there is a six term exact sequence*

$$\begin{tikzcd} K^0_{\mathrm{red}}(Y_*; P) \arrow[r] & K^0_{\mathrm{red}}(Y_*; Q) \arrow[r] & K^0_{\mathrm{red}}(Y_\alpha; Q) \arrow[d] \\ K^1_{\mathrm{red}}(Y_\alpha; Q) \arrow[u] & K^1_{\mathrm{red}}(Y_*; Q) \arrow[l] & K^1_{\mathrm{red}}(Y_*; P). \arrow[l] \end{tikzcd} \tag{6.2}$$

Proof. The upper left arrow is given by inclusion and the upper right arrow by restriction of the reduced bundle data to be non-trivial only on Y_α. The arrows in the bottom row are defined accordingly. Exactness in the middle of the top and the bottom row are immediate from the definitions.

To define the connecting homomorphisms on the left, consider an element in $K^1_{\mathrm{red}}(Y_\alpha; P)$. Choosing splittings as in (4.8) this can be represented by a pair of bundles over $\widehat{B_\alpha} \times Y_\alpha \times [0,1]$ trivial at all boundary hypersurfaces of Y_α, since these correspond to strata labeled by Q, and also trivial at the ends of the interval. Under the augmented pull-back map, this lifts to a pair of bundles over $\widehat{B} \times H_\alpha \times [0,1]$. Now, we can identify $\widehat{B} \times H_\alpha \times [0,1]$ with a collar neighborhood of $\widehat{B} \times H_\alpha$ in $\widehat{B} \times Y$. Since the bundles are trivial over the ends of the interval, this defines an element of $K^0_{\mathrm{red}}(Y; P)$ which is independent of choices so defines a homomorphism.

For the connecting homomorphism on the right the construction is the same after tensoring with the Bott bundle on $[0,1]^2$ and using one variable as the normal to the boundary and the other as the suspension variable.

To check exactness at the top left corner, suppose an element of $K^0_{\mathrm{red}}(Y_*;P)$ maps to the trivial element of $K^0_{\mathrm{red}}(Y_*;Q)$ under inclusion of reduced bundle data. Then, for a stabilized representative $V_\pm$, inside Y_α there is a homotopy of the trivial bundle to itself (respecting triviality of the bundle over deeper strata) that lifts to a homotopy from $V_\pm$ to the reduced bundle data corresponding to the trivial element. Such a homotopy induces a pair of bundles over $\widehat{B}_\alpha \times Y_\alpha \times [0,1]$ trivial at the endpoints and trivial at all strata deeper than Y_α and hence an element of $K^1_{\mathrm{red}}(Y_\alpha;Q)$. A similar argument shows that any element in the kernel of the upper left arrow is of the form discussed in the construction of the connecting homomorphism.

Finally we prove exactness at the bottom left corner. Suppose that an element in $K^1_{\mathrm{red}}(Y_\alpha;Q)$ inserted into the neck near $\widehat{B} \times H_\alpha$ is homotopic to the trivial element, the homotopy preserving the appropriate trivializations corresponding to greater depth. This is equivalent to the existence of a bundle $V_{[0,1]}$ over $\widehat{B} \times Y \times [0,1]$ trivial at $\widehat{B} \times Y \times \{1\}$ and equal to the given element at $\widehat{B} \times Y \times \{0\}$. Denote by V_t the bundle above $\widehat{B} \times Y \times \{t\}$; we need to show is that this data allows the lift of the given bundle to be extended from $\widehat{B} \times H_\alpha \times [0,1]$ to $\widehat{B} \times Y \times [0,1]$ with the extended bundle trivial everywhere (after a homotopy) except at $\widehat{B_\alpha} \times Y_\alpha \times [0,1]$. Fixing a collar neighborhood $\widehat{B} \times H_\alpha \times [0,1] \times [0,1] \subset \widehat{B} \times Y \times [0,1]$ and extending the lifted bundle over $\widehat{B} \times Y \times \{t\}$ by embedding it into the bundle $V_0\big|_{\widehat{B} \times H_\alpha \times [t,1]}$. This can be seen to be a bundle over $\widehat{B} \times Y \times [0,1]$ extending the lift and to be homotopically trivial due to the existence of the initial homotopy. The only issue is that the bundle is not trivial over $\widehat{B} \times Y \times \{0\}$. To ensure this, we use the homotopy to deform the bundle over $\widehat{B} \times Y \times \{t\}$ to $V_{1-t}\big|_{\widehat{B} \times H_\alpha \times [t,1]}$. This satisfies the appropriate triviality conditions and it is still a lift because the homotopy does not affect $\widehat{B} \times H_\alpha$. □

Proposition 13. *For any subtrees $A' = A \setminus P$ and $A' \setminus \{\alpha\}$ there is a six term exact sequence*

$$\begin{array}{ccccc} H^0_{\mathrm{dl}}(Y_*;P) & \longrightarrow & H^0_{\mathrm{dl}}(Y_*;Q) & \longrightarrow & H^0_{\mathrm{dl}}(Y_\alpha;Q) \\ \uparrow & & & & \downarrow \\ H^1_{\mathrm{dl}}(Y_\alpha;Q) & \longleftarrow & H^1_{\mathrm{dl}}(Y_*;Q) & \longleftarrow & H^1_{\mathrm{dl}}(Y_*;P). \end{array} \tag{6.3}$$

Proof. This can be proved by combining standard arguments for the long exact sequence for the cohomology of a manifold relative to its boundary with the arguments as in the case of K-theory above. □

7. The isomorphism

Theorem 14 (See [2]). *For the action of a compact Abelian Lie group on a compact manifold the Chern character defines an isomorphism*

$$K^*_G(X) \otimes \mathbb{C} \longrightarrow H^*_{\mathrm{dl},G}(X). \tag{7.1}$$

Proof. For a minimal element in a subtree as in (6.1) the exact sequences (6.2) and (6.3) combine to form a commutative diagram

$$\begin{tikzcd}
K^0_{\mathrm{red}}(Y_*;P) \arrow[rr] \arrow[dr, "\mathrm{Ch}"] & & K^0_{\mathrm{red}}(Y_*;Q) \arrow[rr] \arrow[d, "\mathrm{Ch}"] & & K^0_{\mathrm{red}}(Y_\alpha;Q) \arrow[dl, "\mathrm{Ch}"'] \arrow[ddd] \\
& H^0_{\mathrm{dl}}(Y_*;P) \arrow[r] & H^0_{\mathrm{dl}}(Y_*;Q) \arrow[r] & H^0_{\mathrm{dl}}(Y_\alpha;Q) \arrow[d] & \\
& H^1_{\mathrm{dl}}(Y_\alpha;Q) \arrow[u] & H^1_{\mathrm{dl}}(Y_*;Q) \arrow[l] & H^1_{\mathrm{dl}}(Y_*;P). \arrow[l] & \\
K^1_{\mathrm{red}}(Y_\alpha;Q) \arrow[uuu] \arrow[ur, "\mathrm{Ch}"'] & & K^1_{\mathrm{red}}(Y_*;Q) \arrow[ll] \arrow[u, "\mathrm{Ch}"] & & K^1_{\mathrm{red}}(Y_*;P). \arrow[ll] \arrow[ul, "\mathrm{Ch}"]
\end{tikzcd} \tag{7.2}$$

Tensoring the K-theory part with $\mathbb{C}$ therefore also gives a commutative diagram with both six-term sequences exact.

Now we may proceed by induction, starting with Proposition 11 applied to the open isotropy type M_0 Then we consider an increasing sequence of subtrees, starting with $P_0 = \{0\}$, the base index and successively adding minimal element in the remainder of the tree

$$P_0 = \{0\} \subset P_1 \cdots \subset P_N = A. \tag{7.3}$$

Thus at each step the commutative diagram (7.2) applies with $P = P_i$ and $Q = P_{i+1}$. The inductive hypothesis is that, after tensoring the K-groups with $\mathbb{C}$, the left upper and right lower Chern characters are isomorphisms. Proposition 11 shows that the vertical maps are isomorphisms so an application of the Fives Lemma shows that the top right and lower left maps are isomorphisms and hence (7.1) follows. $\square$

8. Examples

The examples considered here are covered by Proposition 3.19 in Segal's paper [6]. The aim of this section is to illustrate how the same conclusions can be reached directly by resolution. For the first example we also compute the delocalized cohomology to demonstrate the equality of Theorem 14.

Consider first the standard circle action on the 2-sphere given by rotation around an axis. For $G = \mathrm{U}(1)$, $\widehat{G} \cong \mathbb{Z}$ with the representation at k being the k-fold rotation action, so with $e^{i\theta} \in \mathrm{U}(1)$ acting as $e^{ik\theta}$ on $\mathbb{C}$ for $k \in \mathbb{Z}$. This representation will be denoted L^k.

The two poles on the sphere are fixed points of the action and on the complement the action is free. Radial blow up of the two poles replaces the 2-sphere by a compact cylinder $I \times \mathbb{S}$ with free circle action and quotient an interval, $I = [-1, 1]$. Thus the reduced, iterated manifold consists of a tree with two branches, the base being I with trivial isotropy group and with the other nodes corresponding to the end-points where the isotropy group is $\mathrm{U}(1)$. A reduced bundle on this iterated space consists of a bundle, W, over the interval I, a pair of bundles over the two end-points with decompositions

$$W_{\pm} = \sum_{\text{finite}} L^j \otimes W_{\pm,j}. \tag{8.1}$$

and pull-back (compatibility) isomorphisms to W; in this case these are the forgetful maps sending L^j to the trivial representation, the circle action on the cylinder being free. Since bundles over the interval are trivial, the reduced bundle data amounts to the $W_{\pm}$ constrained to have equal dimensions. Thus the equivariant K-theory is

$$K^0_{\mathrm{U}(1)}(\mathbb{S}^2) = \mathcal{R}(\mathrm{U}(1)) \oplus \mathcal{R}(\mathrm{U}(1))/\simeq \tag{8.2}$$

where the relation corresponds to the equality of the dimensions. The action of $\widehat{\mathrm{U}(1)}$ is the diagonal action on the representation rings

The odd equivariant K-theory is given by the even equivariant K-theory with compact supports of $\mathbb{S}^2 \times (-1, 1)$. Radial blow up of the two poles crossed with the interval replaces $\mathbb{S}^2 \times (-1, 1)$ with $\mathbb{S}^1 \times I \times (-1, 1)$. As above, the reduced manifold consists of a tree with two branches, the base being $I \times (-1, 1)$ and the other nodes being $(-1, 1)$ corresponding to the endpoints with isotropy group $\mathrm{U}(1)$. Since $K^0(\mathbb{R}) = K^0_{\mathrm{U}(1)}(\mathbb{R}) = 0$, the reduced bundle data must be trivial. Thus

$$K^1_{\mathrm{U}(1)}(\mathbb{S}^2) = 0 \tag{8.3}$$

Following Definition 7 the reduced forms are differential forms over I, differential forms tensored with representations over the two nodes related the pull-back, forgetful, maps from the forms over the nodes to the forms over I. Thus $H^0_{dl,\mathrm{U}(1)}(\mathbb{S}^2)$ is given by the closed reduced even forms modulo the image of the reduced odd forms. In particular, since in this case there are only 0 and 1-forms, it suffices to describe the closed 0-forms. A reduced 0-form is given by a smooth function $f(x)$ on I and elements

$$\omega_\pm = \sum_{\text{finite}} c_{\pm,j} L^j \in \mathcal{R}(\mathrm{U}(1)),\ c_{\pm,j} \in \mathbb{R} \tag{8.4}$$

over the two nodes with the pull-back condition being $f(\pm 1) = \sum c_{\pm,j}$. Such a reduced form is closed if and only if $f(x)$ is constant. Thus the even cohomology group is

$$H^0_{\mathrm{dl},\mathrm{U}(1)}(\mathbb{S}^2) = \mathbb{C} \otimes_{\mathbb{Z}} (\mathcal{R}(\mathrm{U}(1)) \oplus \mathcal{R}(\mathrm{U}(1))/\simeq). \tag{8.5}$$

A reduced odd form is fdx for a smooth function $f(x)$ on I which vanishes at the endpoints. Clearly the integral of any such function is a reduced 0-form. Thus

$$H^1_{\mathrm{dl},\mathrm{U}(1)}(\mathbb{S}^2) = 0. \tag{8.6}$$

Lemma 15. *For a group action by a compact Abelian group G with subgroup $A \subset G$ acting trivially, the equivariant K-theory is equal to*

$$K_G(X) = \mathcal{R}(A) \otimes K_{G/A}(X). \tag{8.7}$$

Proof. This follows immediately from the decomposition of bundles under the action of A and the naturality of the lift of representations for a product. □

This lemma in combination with the calculation above shows that for the n-fold rotation action on the sphere the even equivariant K-theory is

$$\mathcal{R}(\mathbb{Z}_n) \otimes \big(\mathcal{R}(\mathrm{U}(1)) \oplus \mathcal{R}(\mathrm{U}(1))/\simeq\big) \tag{8.8}$$

and the odd equivariant K-theory vanishes.

As a final example, consider the two-dimensional complex projective space with the following circle action

$$\mathrm{U}(1) \ni e^{i\theta} : \mathbb{P}^2 \ni [z_1 : z_2 : z_3] \longmapsto [e^{i\theta} z_1 : z_2 : e^{-i\theta} z_3] \in \mathbb{P}^2. \tag{8.9}$$

In general $\mathbb{P}^n = (\mathbb{C}^{n+1} \setminus \{0\})/\mathbb{C}^*$, each of the coordinate planes $\{z_j = 0\}$ in $\mathbb{C}^{n+1}$ projects to $\mathbb{P}^{n-1}_j \subset \mathbb{P}^n$ and $\mathbb{P}^n \setminus \mathbb{P}^{n-1}_j = \mathbb{C}^n$ is covered by the projective coordinate system z'/z_j, $z' = (z_k)_{k \neq j}$. Real blow-up of $\mathbb{P}^{n-1}_j$ replaces $\mathbb{P}^n$ by the radial compactification $\overline{\mathbb{C}^n}$ with the bounding sphere forming the Hopf fibration

$$\partial\overline{\mathbb{C}^n} = \mathbb{S}^{2n-1} \ni \frac{Z}{|Z|} \longrightarrow \mathbb{P}^{n-1}_j,\ Z = z'/z_j. \tag{8.10}$$

The action (8.9) has three fixed points at $[0:0:1]$, $[0:1:0]$ and $[1:0:0]$ and the isotropy group on the sphere $\mathbb{P}_2 = \{[z_1 : 0 : z_3]\}$, outside the first and third of these points, is $\mathbb{Z}_2 = \{1, -1\} \subset \mathrm{U}(1)$. Otherwise the action is free. Since these are smooth submanifolds the resolution of the action, which in principal starts with the fixed submanifolds, can be equivalently produced by first blowing up $\mathbb{P}_2$ and then blowing up the lifts of the preimages of the fixed points.

As noted above the real blow-up $[\mathbb{P}^2; \mathbb{P}_2]_{\mathbb{R}} = \overline{\mathbb{C}^2}$ is, a manifold with boundary, diffeomorphic to the closed real four-ball and the induced action is

$$\mathbb{C}^2 \ni (Z_1, Z_3) \longmapsto (e^{i\theta} Z_1, e^{-i\theta} Z_3). \tag{8.11}$$

Blow up of the fixed point, $[0:1:0]$, at the center of this ball, replaces it by a spherical shell $[0,1]\times\mathbb{S}^3$ with the radial variable compactified to an interval. The action on $\mathbb{S}^3$ projects under the Hopf fibration to the action $\mathbb{P}_2 \ni [Z_1, Z_3] \longrightarrow [e^{i\theta}Z_1, e^{-i\theta}Z_3]$.

The lifts of the first and third fixed points are non-intersecting circles, $Z_1 = 0$ and $Z_3 = 0$ in the 'outer' sphere and blowing these up gives the full resolution which is a manifold with corners up to codimension two. The outer spherical boundary is replaced by three boundary hypersurfaces. The first is the 3-sphere with two circles blown up, so this is the product of a closed interval and a 2-torus; it fibres (via the lift of the Hopf fibration) over the 2-cylinder, the blow-up of $\mathbb{P} = \mathbb{S}^2$ at two poles. The action on the boundary hypersurface factors through that on the 2-torus by opposite rotations on the two circles and so fibres over the double rotation action on the circle in the base. The other two boundary hypersurfaces are solid 3-tori given as the product a circle and a disk, with (opposite) rotation actions on the circle and the 2-disk. At the outer boundary this fibres via the lift of the Hopf fibration over the circle with double rotation action, and overall fibres over a point with trivial action.

The action on the resolved manifold is free with quotient a spherical shell, the product of an interval and a 2-sphere, with the 'outer' boundary blown up at antipodal points on the sphere. Thus it is a 3-manifold with corners of codimension two having boundary hypersurfaces an 'inner' 2-sphere, fibering over a point, an outer 2-cylinder, fibering over an interval, and two outer 2-disks, fibering over points. Each of the 2-disks meets the 2-cylinder in circles with the fibrations consistent, mapping these codimension two bounding circles to the end-points of the interval. This then is the reduced iterated manifold which as a tree has base node, M, the three manifold connected to two nodes, one, P_0, corresponding to the second fixed point, the other, C, the 2-cylinder which in turn is connected to the other two fixed point nodes, $P_\pm$.

$$\begin{tikzcd}[column sep=small] & M \arrow[dl, no head] \arrow[dr, no head] & & \\ P_0 & & C \arrow[dl, no head] \arrow[dr, no head] & \\ & P_+ & & P_-. \end{tikzcd}$$

Reduced bundles consist of bundles over $\mathbb{Z}\times Y_*$ for all five nodes. Over the base, corresponding to the free action, the bundles are of the form $L_{-k}\otimes W$ for a fixed bundle W over the 3-manifold. Over the inner boundary the restriction of W is identified as the augmented pull back of representations of $\mathrm{U}(1)$; this is no restriction since every bundle over the 2-sphere is of this form.

From the discussion above we already know that the K-theory of $\mathbb{P} = [z_1 : 0 : z_3]$, since it is the sphere with double rotation, and the K-theory of $[0:1:0]$ is just $\mathcal{R}(U(1))$. The lifts of a bundle from $[0:1:0]$ and a bundle from $\mathbb{P}$ extend to a bundle on all of $[0,1]\times\mathbb{S}^3$ if and only if they are $\mathbb{S}$-equivariantly isomorphic over

$\mathbb{S}^3$ for the action of $\mathbb{S}$ on $\mathbb{S}^3$ given above. Since the action is free, this is equivalent to two bundles being isomorphic over $\mathbb{S}^2$. Since the K-theory of $\mathbb{S}^2$ is just $\mathbb{Z}$, this is the case if and only if the two bundles have the same dimension. Therefore,

$$K^0_{U(1)}(\mathbb{P}^2) = (\mathcal{R}(U(1)) \oplus (\mathcal{R}(Z_2) \otimes (\mathcal{R}(\mathrm{U}(1)) \oplus \mathcal{R}(\mathrm{U}(1))/ \simeq) / \simeq . \tag{8.12}$$

Unsurprisingly, since $\mathbb{P}^2$ only has cells in even dimensions, the odd equivariant K-theory is zero. This can be seen directly from the analysis of the previous examples since the odd K-theory at the two endpoints is zero, as is the even K-theory of $\mathbb{S}^2 \times I \times (0, 1)$. Thus

$$K^1_{U(1)}(\mathbb{P}^2) = 0. \tag{8.13}$$

References

[1] Pierre Albin and Richard Melrose, *Resolution of smooth group actions*, Spectral theory and geometric analysis, Contemp. Math., vol. 535, Amer. Math. Soc., Providence, RI, 2011, pp. 1–26. MR **2560748 (2012j:**58008)

[2] Paul Baum, Jean-Luc Brylinski, and Robert MacPherson, *Cohomologie équivariante délocalisée*, C. R. Acad. Sci. Paris Sér. I Math. **300** (1985), no. 17, 605–608. MR **86g:**55006

[3] Nicole Berline, Ezra Getzler, and Michèle Vergne, *Heat kernels and Dirac operators*, Springer-Verlag, Berlin, 1992.

[4] Ezra Getzler, *The equivariant Chern character for non-compact Lie groups*, Adv. Math. **109** (1994), no. 1, 88–107. MR 1302758

[5] Ioanid Rosu, *Equivariant K-theory and equivariant cohomology*, Math. Z. **243** (2003), no. 3, 423–448, with an appendix by Allen Knutson and Rosu.

[6] Graeme Segal, *Equivariant K-theory*, Inst. Hautes Études Sci. Publ. Math. (1968), no. 34, 129–151. MR 0234452

[7] Antony Wasserman, *Equivariant K-theory I: compact transformation groups with one orbit type*, IHES preprint, 1985.

Panagiotis Dimakis
Department of Mathematics
Stanford University
e-mail: pdimakis@stanford.edu

Richard Melrose
Department of Mathematics
Massachusetts Institute of Technology
e-mail: rbm@math.mit.edu

Progress in Mathematics, Vol. 333, 93–111

On the Existence Problem of Einstein–Maxwell Kähler Metrics

Akito Futaki and Hajime Ono

On the occasion of Professor Gang Tian's sixtieth birthday

Abstract. In this expository paper we review on the existence problem of Einstein–Maxwell Kähler metrics, and make several remarks. Firstly, we consider a slightly more general set-up than Einstein–Maxwell Kähler metrics, and give extensions of volume minimization principle, the notion of toric K-stability and other related results to the general set-up. Secondly, we consider the toric case when the manifold is the one point blow-up of the complex project plane and the Kähler class Ω is chosen so that the area of the exceptional curve is sufficiently close to the area of the rational curve of self-intersection number 1. We observe by numerical analysis that there should be a Killing vector field K which gives a toric K-stable pair (Ω, K) in the sense of Apostolov–Maschler.

Mathematics Subject Classification (2010). 32Q15, 53C55.

Keywords. Einstein–Maxwell Kähler metric, volume minimization, K-stability, toric manifold.

1. Introduction

Let (M, J) be a compact Kähler manifold of complex dimension m. A Hermitian metric $\tilde{g}$ of constant scalar curvature on (M, J) is said to be a conformally Kähler, Einstein–Maxwell (cKEM for short) metric if there exists a positive smooth function f on M such that $g = f^2\tilde{g}$ is Kähler and that the Hamiltonian vector field $K = J\mathrm{grad}_g f$ of f with respect to the Kähler form ω_g of g is a Killing vector field for both g and $\tilde{g}$. In this case we call the Kähler metric g an Einstein–Maxwell Kähler (EMK for short) metric. Let ω_0 be a Kähler form, and consider $\Omega = [\omega_0] \in H^2_{\mathrm{DR}}(M, \mathbf{R})$ as a fixed Kähler class. We look for an Einstein–Maxwell Kähler metric g such that the Kähler form ω_g belongs to Ω.

Let G be a maximal torus of the reduced automorphism group, and pick $K \in \mathfrak{g} := \mathrm{Lie}(G)$. Then the problem is to find a G-invariant Kähler metric g with its Kähler form $\omega_g \in \Omega$ such that

(i) $\tilde{g} = f^{-2} g$ is a cKEM metric,
(ii) $J\mathrm{grad}_g f = K$.

The scalar curvature $s_{\tilde{g}}$ of $\tilde{g} = f^{-2} g$ is given by

$$s_{\tilde{g}} = f^2 s_g - 2(2m-1) f \Delta_g f - 2m(2m-1)|df|_g^2 \tag{1}$$

where s_g is the scalar curvature of g and Δ_g is the Hodge Laplacian with respect to g.

Now, starting with a Kähler metric g and a Killing potential f, for any real number $n \in \mathbf{R}$ with $n \neq 0,\ 1,\ 2$ and $k \in \mathbf{R}$ with $k \neq 0$ we define the (g, f, k, n)-scalar curvature $s_{g,f,k,n}$ by

$$s_{g,f,k,n} = f^{-k} \left\{ s_g + k(n-1)\frac{1}{f}\Delta_g f + \frac{k}{4}(n-1)(4+2k-kn)\frac{1}{f^2}|df|_g^2 \right\}. \tag{2}$$

The case $n = 2m$ is the scalar curvature $s_{\tilde{g}}$ of the conformal metric $\tilde{g} = f^k g$, and for other values of n such a meaning is lost. However, the cases of general values of n appear in natural contexts such as in [2] and [13]. Moreover, Lahdili proves in [10] and [11] results for cKEM metrics can be generalized to constant $(g, f, -2, n)$-scalar curvature.

In this expository paper we give extensions of the volume minimization principle [8], [9], the notion of toric K-stability [3] for $k = -2$ and other related results for cKEM metrics to the general set-up of constant (g, f, k, n)-scalar curvature. We consider the toric case where the manifold is the one point blow-up of the complex project plane and the Kähler class Ω is chosen so that the area of the exceptional curve is sufficiently close to the area of the rational curve of self-intersection number 1. We observe by numerical analysis that there should be a Killing vector field K which gives a toric K-stable pair (Ω, K) in the sense of Apostolov–Maschler. For this purpose we show in Theorem 5.3 that we have only to consider the simple test configurations to test toric K-stability, extending the earlier works of Donaldson [4], Wang and Zhou [14], [15].

The rest of this paper is organized as follows. In section 2 we extend the volume minimization for Einstein–Maxwell Kähler metrics, see Theorem 2.1. In Section 3 we review the normalized Einstein–Hilbert functional, and study its relation to the volume functional and the Futaki invariant. In Section 4 we consider the normalized Einstein–Hilbert functional on toric Kähler manifolds. In Section 5 we review toric K-stability, and prove Theorem 5.3. We then review the result of our paper [8] on the one-point blow-up of $\mathbf{CP}^2$ and show the graphics of the results of the numerical analysis which indicate that this case should be K-stable and there should be a conformally Kähler, Einstein–Maxwell metric.

2. Volume minimization for Einstein–Maxwell Kähler metrics

In this section we review the results in [8] and extend them to constant (g,f,k,n)-scalar curvature. Let M be a compact smooth manifold. We denote by $\mathrm{Riem}(M)$ the set of all Riemannian metrics on M, by s_g the scalar curvature of g, and by dv_g the volume form of g. For any given positive smooth function f and real numbers $n \in \mathbf{R}$ with $n \neq 0,\ 1,\ 2$ and $k \in \mathbf{R}$ with $k \neq 0$, we define $s_{g,f,k,n}$ by the same formula as (2). We put

$$S(g,f,k,n) := \int_M s_{g,f,k,n}\, f^{\frac{nk}{2}}\, dv_g \tag{3}$$

and call it the total (g,f,k,n)-scalar curvature, and put

$$\mathrm{Vol}(g,f,k,n) := \int_M f^{\frac{nk}{2}}\, dv_g \tag{4}$$

and call it the (g,f,k,n)-volume.

Let f_t be a smooth family of positive functions such that $f_0 = f$, $d/dt|_{t=0} f_t = \phi$. Then by straightforward computations we have

$$\left.\frac{d}{dt}\right|_{t=0} S(g,f_t,k,n) = \frac{k}{2}(n-2)\int_M s_{g,f,k,n}\,\phi\, f^{\frac{nk}{2}-1}\, dv_g \tag{5}$$

and

$$\left.\frac{d}{dt}\right|_{t=0} \mathrm{Vol}(g,f_t,k,n) = \frac{nk}{2}\int_M \phi\, f^{\frac{nk}{2}-1}\, dv_g. \tag{6}$$

Now we consider a compact Kähler manifold (M,J) of complex dimension m. As in section 1, let G be a maximal torus of the reduced automorphism group, and take $K \in \mathfrak{g} := \mathrm{Lie}(G)$. Consider a fixed Kähler class Ω on (M,J), and denote by $\mathcal{K}_\Omega^G$ the space of G-invariant Kähler metrics ω in Ω. For any $(K,a,g) \in \mathfrak{g} \times \mathbf{R} \times \mathcal{K}_\Omega^G$, there exists a unique function $f_{K,a,g} \in C^\infty(M,\mathbf{R})$ satisfying the following two conditions:

$$\iota_K\omega = -df_{K,a,g}, \quad \int_M f_{K,a,g}\frac{\omega^m}{m!} = a. \tag{7}$$

By (7), it is easy to see that $f_{K,a,g}$ has the following properties:

$$f_{K+H,a+b,g} = f_{K,a,g} + f_{H,b,g} \tag{8}$$

$$f_{0,a,g} = \frac{a}{\mathrm{Vol}(M,\omega)} \tag{9}$$

$$f_{CK,Ca,g} = Cf_{K,a,g} \tag{10}$$

Hereafter the Kähler metric g and its Kähler form ω_g are often identified, and ω_g is often denoted by ω. Noting that $\min\{f_{K,a,g}\,|\,x \in M\}$ is independent of $g \in \mathcal{K}_\Omega^G$ (this follows from the convexity of moment map images and the fact that the vertices do not move even if we change the Kähler metric in the fixed Kähler class Ω), we put

$$\mathcal{P}_\Omega^G \quad := \quad \{(K,a) \in \mathfrak{g} \times \mathbf{R}\,|\, f_{K,a,g} > 0\}. \tag{11}$$

Note that the right-hand side of (11) is independent of $g \in \mathcal{K}_\Omega^G$ again since the moment polytope is independent of $g \in \mathcal{K}_\Omega^G$. Fixing $(K,a) \in P_\Omega^G$, $n \in \mathbf{R}$ and $k \in \mathbf{R}$, put

$$c_{\Omega,K,a,k,n} := \frac{\displaystyle\int_M s_{g,f_{K,a,g},k,n}\, f_{K,a,g}^{\frac{kn}{2}-1}\, \frac{\omega^m}{m!}}{\displaystyle\int_M f_{K,a,g}^{\frac{kn}{2}-1}\, \frac{\omega^m}{m!}} \tag{12}$$

and

$$d_{\Omega,K,a,k,n} := \frac{S(g,f_{K,a,g},k,n)}{\mathrm{Vol}(g,f_{K,a,g},k,n)} = \frac{\displaystyle\int_M s_{g,f_{K,a,g},k,n}\, f_{K,a,g}^{\frac{kn}{2}}\, \frac{\omega^m}{m!}}{\displaystyle\int_M f_{K,a,g}^{\frac{kn}{2}}\, \frac{\omega^m}{m!}} \tag{13}$$

Then $c_{\Omega,K,a,k,n}$ and $d_{\Omega,K,a,k,n}$ are constants independent of the choice of $g \in \mathcal{K}_\Omega^G$ since the integrands of (12) and (13) are part of equivariant cohomology, see, e.g., [6], [5], [7]. Since P_Ω^G is a cone in $\mathfrak{g} \times \mathbf{R}$ by (10), with n and k fixed we consider its slice

$$\tilde{\mathcal{P}}_\Omega^G := \left\{ (K,a) \in \mathcal{P}_\Omega^G \;\middle|\; d_{\Omega,K,a,k,n} = \gamma \right\} \tag{14}$$

where γ is chosen to be -1, 0 or 1 depending on the sign of $d_{\Omega,K,a,k,n}$. Let $(K(t),a(t))$, $t \in (-\varepsilon,\varepsilon)$ be a smooth curve in $\tilde{\mathcal{P}}_\Omega^G$ such that $(K(0),a(0)) = (K,a), (K'(0),a'(0)) = (H,b)$. Then

$$S(g,f_{K(t),a(t),g},k,n) = \gamma\, \mathrm{Vol}(g,f_{K(t),a(t),g},k,n)$$

holds for any $t \in (-\varepsilon,\varepsilon)$. By differentiating this equation at $t=0$ and noting $k \neq 0$, we have

$$(n-2)\int_M s_{g,f_{K,a,g},k,n} f_{H,b,g}\, f_{K,a,g}^{\frac{nk}{2}-1} \frac{\omega^m}{m!} = n\gamma \int_M f_{H,b,g}\, f_{K,a,g}^{\frac{nk}{2}-1} \frac{\omega^m}{m!}. \tag{15}$$

The linear function $\mathrm{Fut}^G_{\Omega,K,a,k,n} : \mathfrak{g} \to \mathbf{R}$ defined by

$$\mathrm{Fut}^G_{\Omega,K,a,k,n}(H) := \int_M (s_{g,K,a,k,n} - c_{\Omega,K,a,k,n})\, f_{H,b,g}\, f_{K,a,g}^{\frac{nk}{2}-1} \frac{\omega_g^m}{m!} \tag{16}$$

is independent of the choice of Kähler metric $g \in \mathcal{K}_\Omega^G$ and $b \in \mathbf{R}$ ([3]). If there exists a Kähler metric $g \in \mathcal{K}_\Omega^G$ such that $\tilde{g} = f_{K,a,g}^k g$ is a constant (g,f,k,n)-scalar curvature metric, then $\mathrm{Fut}^G_{\Omega,K,a,k,n}$ vanishes identically.

For the path $(K(t),a(t))$, $t \in (-\varepsilon,\varepsilon)$ in $\tilde{\mathcal{P}}_\Omega^G$ with $(K(0),a(0)) = (K,a)$, $(K'(0),a'(0)) = (H,b)$ we have from (15)

$$\begin{aligned} \mathrm{Fut}^G_{\Omega,K,a,k,n}(H) &= \left(\frac{n\gamma}{n-2} - c_{\Omega,K,a,k,n}\right) \int_M f_{H,b,g}\, f_{K,a,g}^{\frac{nk}{2}-1} \frac{\omega^m}{m!} \\ &= \left(\frac{n\gamma}{n-2} - c_{\Omega,K,a,k,n}\right) \frac{2}{nk} \left.\frac{d}{dt}\right|_{t=0} \mathrm{Vol}(g,f_{K(t),a(t),g},k,n). \end{aligned} \tag{17}$$

If there exists a constant (g, f, k, n)-scalar curvature metric $\tilde{g} = f_{K,a,g}^{k} g$ with $g \in \mathcal{K}_{\Omega}^{G}$, then

$$c_{\Omega,K,a,k,n} = d_{\Omega,K,a,k,n} = \gamma$$

and

$$\mathrm{Fut}_{\Omega,K,a,k,n}^{G}(H) = 0.$$

Therefore for $\gamma = \pm 1$ we have

$$\left.\frac{d}{dt}\right|_{t=0} \mathrm{Vol}(g, f_{K(t),a(t),g}, k, n) = 0. \tag{18}$$

The case of $\gamma = 0$ can be treated separately, see [8].

We summarize the result as follows.

Theorem 2.1. *Let Ω be a fixed Kähler class, and $n \neq 0,\ 1,\ 2$ and $k \neq 0$ be fixed real numbers. Suppose that the pair (K, a) of Killing vector field K and normalization constant a belongs to $\tilde{\mathcal{P}}_{\Omega}^{G}$. If there exists a G-invariant Kähler metric g in the Kähler class Ω, i.e., $g \in \mathcal{K}_{\Omega}^{G}$, such that the (g, f, k, n)-scalar curvature is constant for the Killing Hamiltonian function $f = f_{K,a,g}$ then (K, a) is a critical point of $\mathrm{Vol}_{n,k} : \tilde{\mathcal{P}}_{\Omega}^{G} \to \mathbf{R}$ given by*

$$\begin{aligned} \mathrm{Vol}_{n,k}(K, a) &:= \mathrm{Vol}(g, f_{K,a,g}, k, n) \\ &= \int_M f_{K,a,g}^{\frac{nk}{2}} dv_g \end{aligned}$$

for $(K, a) \in \tilde{\mathcal{P}}_{\Omega}^{G}$. Further, (K, a) is a critical point of $\mathrm{Vol}_{n,k} : \tilde{\mathcal{P}}_{\Omega}^{G} \to \mathbf{R}$ if and only if $\mathrm{Fut}_{\Omega,K,a,k,n}^{G} \equiv 0$.

Corollary 2.2. *Let Ω be a fixed Kähler class. Take $n = 2m$ and $k = -2$, and let $(K, a) \in \tilde{\mathcal{P}}_{\Omega}^{G}$. If there exists a conformally Kähler, Einstein–Maxwell metric $\tilde{g} = f_{K,a,g}^{-2} g$ with $g \in \mathcal{K}_{\Omega}^{G}$, then (K, a) is a critical point of $\mathrm{Vol} : \tilde{\mathcal{P}}_{\Omega}^{G} \to \mathbf{R}$ given by $\mathrm{Vol}(K, a) := \mathrm{Vol}(g, f_{K,a,g}, -2, 2m)$ for $(K, a) \in \tilde{\mathcal{P}}_{\Omega}^{G}$. Further, (K, a) is a critical point of $\mathrm{Vol} : \tilde{\mathcal{P}}_{\Omega}^{G} \to \mathbf{R}$ if and only if $\mathrm{Fut}_{\Omega,K,a,-2,2m}^{G} \equiv 0$.*

For a given Kähler class Ω the critical points of $\mathrm{Vol} : \tilde{\mathcal{P}}_{\Omega}^{G} \to \mathbf{R}$ are not unique in general as can be seen from LeBrun's construction [12].

3. The normalized Einstein–Hilbert functional

In the previous section we confined ourselves to the view point from the volume functional. In the present section we see that, when restricted to $\tilde{\mathcal{P}}_{\Omega}^{G}$, considering the volume functional is essentially the same as considering the normalized Einstein–Hilbert functional. The normalized Einstein–Hilbert functional $EH : \mathrm{Riem}\,(M) \to \mathbf{R}$ on an n-dimensional compact Riemannian manifold is the functional on $\mathrm{Riem}(M)$ defined by

$$EH(g) := \frac{S(g)}{(\mathrm{Vol}(g))^{\frac{n-2}{n}}}$$

where $S(g)$ and $\mathrm{Vol}(g)$ are respectively the total scalar curvature and the volume of g. It is a standard fact that the critical points of EH are Einstein metrics, and that, when restricted to a conformal class, the critical points are metrics of constant scalar curvature.

Let us see this in a slightly different setting. In the equation (2), let us replace s_g by a smooth function φ, and put

$$s_{g,f,k,n,\varphi} = f^{-k}\left\{\varphi + k(n-1)\frac{1}{f}\Delta_g f + \frac{k}{4}(n-1)(4+2k-kn)\frac{1}{f^2}|df|_g^2\right\}. \tag{19}$$

Accordingly, we may replace (3) by

$$S(g,f,k,n,\varphi) := \int_M s_{g,f,k,n,\varphi}\, f^{\frac{nk}{2}}\, dv_g, \tag{20}$$

and replace the normalized Einstein–Hilbert functional by

$$EH(g,f,k,n,\varphi) := \frac{S(g,f,k,n,\varphi)}{(\mathrm{Vol}(g,f,k,n,\varphi))^{\frac{n-2}{n}}}.$$

As before, let f_t be a smooth family of positive functions such that

$$f_0 = f,\ d/dt|_{t=0} f_t = \phi.$$

Then one can show

$$\begin{aligned} &\left.\frac{d}{dt}\right|_{t=0} EH(g,f_t,k,n,\varphi) \\ &\quad = \frac{(n-2)k}{2}\,\mathrm{Vol}(g,f,k,n,\varphi)^{\frac{2-n}{n}} \\ &\qquad \cdot \left\{\int_M \left(s_{g,f,k,n,\varphi} - \frac{S(g,f,k,n,\varphi)}{\mathrm{Vol}(g,f,k,n)}\right)\phi\, f^{\frac{nk}{2}-1} dv_g\right\}. \end{aligned} \tag{21}$$

Thus we have shown

Proposition 3.1. *The function $s_{g,f,k,n,\varphi}$ satisfies*

$$s_{g,f,k,n,\varphi} = \text{constant}$$

if and only if f is a critical point of the functional $f \mapsto EH(g,f,k,n,\varphi)$.

Let us return to the situation of the previous section where we considered a compact Kähler manifold with a maximal torus G of the reduced automorphisms group, with a fixed Kähler class Ω. Taking φ to be the (g,f,k,n)-scalar curvature, we consider the Einstein–Hilbert functional $EH(g,f,k,n) := EH(g,f,k,n,s_{g,f,k,n})$. By the same reasoning from equivariant cohomology again, for a fixed (K,a), $EH(g,f_{K,a,g},k,n)$ is independent of the choice of $g \in \mathcal{K}_\Omega^G$. Set $EH_{k,n}(K,a) := EH(g,f_{K,a,g},k,n)$. Then using (8), (9) and (10), we see

$$\begin{aligned} &\left.\frac{d}{dt}\right|_{t=0} EH_{k,n}(K+tH,a) \\ &\quad = \frac{(n-2)k}{2\,\mathrm{Vol}_n(K,a)^{\frac{n-2}{n}}}\int_M \left(s_{g,f_{K,a,g},k,n} - d_{\Omega,K,a,k,n}\right) f_{K,a,g}^{\frac{nk}{2}-1} f_{H,0,g}\frac{\omega_g^m}{m!} \end{aligned} \tag{22}$$

and

$$\left.\frac{d}{dt}\right|_{t=0} EH_{k,n}(K, a+tb) \tag{23}$$
$$= \frac{(n-2)kb}{2\,\mathrm{Vol}_n(K,a)^{\frac{n-2}{n}+1}}(c_{\Omega,K,a,n} - d_{\Omega,K,a,n}) \int_M f_{K,a,g}^{\frac{nk}{2}-1} \frac{\omega_g^m}{m!}.$$

If there exist $g \in \mathcal{K}_\Omega^G$, K and a such that $s_{g,f_{K,a,g},k,n}$ is constant, then

$$s_{g,f_{K,a,g},k,n} = c_{\Omega,K,a,k,n} = d_{\Omega,K,a,k,n}, \tag{24}$$

and thus the pair (K,a) is a critical point of the function $EH_{k,n} : \mathcal{P}_\Omega^G \to \mathbf{R}$ given by

$$(K,a) \mapsto EH_{k,n}(K,a) := EH(g, f_{K,a,g}, k, n). \tag{25}$$

Conversely, suppose that (K,a) is a critical point of $EH_{k,n} : \mathcal{P}_\Omega^G \to \mathbf{R}$. Then one can see (K,a) satisfies $c_{\Omega,K,a,k,n} = d_{\Omega,K,a,k,n}$. Hence, by (16) and (22), $\mathrm{Fut}_{\Omega,K,a,k,n}^G$ vanishes. More direct relation between the volume functional and the Einstein–Hilbert functional can be seen as follows.

Remark 3.2. Since $EH_{k,n}$ is homogeneous of degree 0 on $\mathcal{P}_\Omega^G$ we may restrict $EH_{k,n}$ to the slice

$$\tilde{\mathcal{P}}_{\Omega,n}^G := \{(K,a) \in \mathcal{P}_{\Omega,n}^G \mid d_{\Omega,k,a,k,n} = \gamma\} \tag{26}$$

Then

$$EH_{k,n}(K,a) = \gamma\, \mathrm{Vol}_{k,n}(K,a)^{\frac{2}{n}} \tag{27}$$

on $\tilde{\mathcal{P}}_{\Omega,n}^G$. This shows that the volume minimization Theorem 2.1 is equivalent to finding a critical point of the Einstein–Hilbert functional.

4. The normalized Einstein–Hilbert functional for toric Kähler manifolds

In this section, we give the explicit formula for the Futaki invariant and the normalized Einstein–Hilbert functional when (M, J, ω) is a compact toric Kähler manifold and $k = -2$.

Let (M, ω) be a $2m$-dimensional compact toric manifold and $\mu : M \to \mathbf{R}^m$ the moment map. It is well known that the image of μ, $\Delta := \mathrm{Image}\,\mu$, is an m-dimensional Delzant polytope in $\mathbf{R}^m$. A T^m-invariant, ω-compatible complex structure J on M gives a convex function u, called a symplectic potential, on Δ as follows. For the action-angle coordinates $(\mu_1, \ldots, \mu_m, \theta_1, \ldots, \theta_m) \in \Delta \times T^m$, there exists a smooth convex function u on Δ which satisfies

$$J\frac{\partial}{\partial \mu_i} = \sum_{j=1}^m u_{,ij} \frac{\partial}{\partial \theta_j}, \quad J\frac{\partial}{\partial \theta_i} = \sum_{j=1}^m \mathbf{H}_{ij}^u \frac{\partial}{\partial \mu_j},$$

where, for a smooth function φ of $\mu = (\mu_1, \ldots, \mu_m)$, we denote by $\varphi_{,i}$ the partial derivative $\partial\varphi/\partial\mu_i$ and by $\mathbf{H}^u = (\mathbf{H}_{ij}^u)$ the inverse matrix of the Hessian $(u_{,ij})$ of u. Conversely, if we give a smooth convex function u on Δ satisfying some boundary conditions, by the formula above, we can recover a T^m-invariant ω-compatible complex structure on M, see [1] for more details.

Let u be a symplectic potential on Δ. Then the toric Kähler metric $g_J = \omega(\cdot, J\cdot)$ is represented as

$$g_J = \sum_{i,j=1}^{m} u_{,ij} d\mu_i d\mu_j + \sum_{i,j=1}^{m} \mathbf{H}^u_{ij} d\theta_i d\theta_j. \tag{28}$$

According to Abreu [1], the scalar curvature s_J of g_J is

$$s_J = -\sum_{i,j=1}^{m} \mathbf{H}^u_{ij,ij}. \tag{29}$$

In this case, a Killing potential is an affine linear function positive on Δ. Fix a Killing potential f. Then (g_J, f, k, n)-scalar curvature $s_{J,f,k,n}$ is given by

$$s_{J,f,k,n} = f^{-k} s_J + \frac{4(n-1)}{n-2} f^{-\frac{k(n+2)}{4}} \Delta_J f^{\frac{k(n-2)}{4}}, \tag{30}$$

where $\Delta_J = \Delta_{g_J}$. For a smooth function φ of $\mu_1, \dots, \mu_m$,

$$\Delta_J \varphi = -\sum_{i,j=1}^{m} \{\varphi_{,ij} \mathbf{H}^u_{ij} + \varphi_{,i} \mathbf{H}^u_{ij,j}\}$$

holds (see the equation (20) in [3]). Since f is affine linear, we have

$$\begin{aligned} &\Delta_J f^{\frac{k(n-2)}{4}} \\ &= -\frac{k(n-2)}{4} f^{\frac{k(n-2)}{4}} \sum_{i,j=1}^{m} \left\{ \left(\frac{k(n-2)}{4} - 1 \right) \frac{f_{,i} f_{,j}}{f^2} \mathbf{H}^u_{ij} + \frac{f_{,i}}{f} \mathbf{H}^u_{ij,j} \right\}. \end{aligned} \tag{31}$$

By (29), (30) and (31), the (g_J, f, k, n)-scalar curvature is

$$\begin{aligned} &s_{J,f,k,n} \\ &= -f^{-k} \sum_{i,j=1}^{m} \left\{ \mathbf{H}^u_{ij,ij} + \frac{k(n-1)}{f} f_{,i} \mathbf{H}^u_{ij,j} + \frac{k(n-1)}{f^2} \left(\frac{k(n-2)}{4} - 1 \right) f_{,i} f_{,j} \mathbf{H}^u_{ij} \right\}. \end{aligned} \tag{32}$$

On the other hand, for any $\alpha \in \mathbf{R}$,

$$\sum_{i,j=1}^{m} \left(f^\alpha \mathbf{H}^u_{ij} \right)_{,ij} = f^\alpha \sum_{i,j=1}^{m} \left\{ \mathbf{H}^u_{ij,ij} + \frac{2\alpha}{f} f_{,i} \mathbf{H}^u_{ij,j} + \frac{\alpha(\alpha-1)}{f^2} f_{,i} f_{,j} \mathbf{H}^u_{ij} \right\} \tag{33}$$

holds. We easily see that $2\alpha = k(n-1)$ and $\alpha(\alpha-1) = k(n-1)(k(n-2)/4 - 1)$ hold if and only if $k = -2$ and $\alpha = 1-n$. In this case, we have

$$s_{J,f,-2,n} f^{-1-n} = -\sum_{i,j=1}^{m} \left(f^{1-n} \mathbf{H}^u_{ij} \right)_{,ij}. \tag{34}$$

By Lemma 2 in [3], for any smooth function ϕ on $\mathbf{R}^m$,

$$\int_\Delta \phi \sum_{i,j=1}^m \left(f^{1-n}\mathbf{H}^u_{ij}\right)_{,ij} d\mu = \int_\Delta f^{1-n} \sum_{i,j=1}^m \mathbf{H}^u_{ij}\phi_{,ij}\, d\mu - 2\int_{\partial\Delta} f^{1-n}\phi\, d\sigma. \tag{35}$$

In particular, when ϕ is an affine function

$$\int_\Delta \phi \sum_{i,j=1}^m \left(f^{1-n}\mathbf{H}^u_{ij}\right)_{,ij} d\mu = -2\int_{\partial\Delta} f^{1-n}\phi\, d\sigma \tag{36}$$

holds. Hence, if we define the constant $c_{\Delta,f,-2,n}$ as

$$c_{\Delta,f,-2,n} = 2\,\frac{\displaystyle\int_{\partial\Delta} f^{1-n}\, d\sigma}{\displaystyle\int_\Delta f^{-1-n}\, d\mu},$$

the Futaki invariant (16) is given by

$$\mathrm{Fut}_{\Delta,f,-2,n}(\phi) = 2\int_{\partial\Delta} f^{1-n}\phi\, d\sigma - c_{\Delta,f,-2,n}\int_\Delta f^{-1-n}\phi\, d\mu \tag{37}$$

for any linear function ϕ on $\mathbf{R}^m$.

By (34) and (36), $EH(g_J, f, -2, n)$ is given by

$$EH_{-2,n}(f) := EH(g_J, f, -2, n) = \mathrm{Const.}\frac{\displaystyle\int_{\partial\Delta} f^{2-n}\, d\sigma}{\left(\displaystyle\int_\Delta f^{-n}\, d\mu\right)^{\frac{n-2}{n}}}. \tag{38}$$

If there exists a symplectic potential u such that the $(g_J, f, -2, n)$-scalar curvature is constant, then $\mathrm{Fut}_{\Delta,f,-2,n}$ vanishes identically and f is a critical point of $EH_{-2,n}$.

5. Toric K-stability

Let (M, ω) be a $2m$-dimensional compact toric manifold with the moment image $\Delta \subset \mathbf{R}^m$. Following the argument by Donaldson in [4], we may define the Donaldson–Futaki invariant with respect to a positive affine function f on Δ as

$$\mathrm{DF}_{\Delta,f,n}(\phi) = 2\int_{\partial\Delta} f^{1-n}\phi\, d\sigma - c_{\Delta,f,-2,n}\int_\Delta f^{-1-n}\phi\, d\mu \tag{39}$$

for a convex function ϕ on Δ, see also [3]. For any affine function ϕ,

$$\mathrm{Fut}_{\Delta,f,-2,n}(\phi) = \mathrm{DF}_{\Delta,f,n}(\phi).$$

We can prove the following straightforward analogue of the results in [4]:

Theorem 5.1. *Suppose that there exists a symplectic potential u on Δ such that the $(g_J, f, -2, n)$-scalar curvature is a constant c. Then*

$$c = c_{\Delta,f,-2,n} \quad \textit{and} \quad \mathrm{DF}_{\Delta,f,n}(\phi) \geq 0$$

for any smooth convex function ϕ on Δ. Equality holds if and only if ϕ is affine.

Proof. Suppose that $s_{J,f,-2,n} = c$. Then

$$c\int_\Delta f^{-1-n}\, d\mu = -\int_\Delta \sum_{i,j=1}^m (f^{1-n}\mathbf{H}^u_{ij})_{,ij}\, d\mu = 2\int_{\partial\Delta} f^{1-n}\, d\sigma$$

by (34) and (36). Hence $c = c_{\Delta,f,-2,n}$. By (35),

$$\begin{aligned}\mathrm{DF}_{\Delta,f,n}(\phi) =\ & -\int_\Delta \left(c_{\Delta,f,-2,n} f^{-1-n} + \sum_{i,j=1}^m \left(f^{1-n}\mathbf{H}^u_{ij}\right)_{,ij} \right) \phi\, d\mu \\ & + \int_\Delta f^{1-n} \sum_{i,j=1}^m \mathbf{H}^u_{ij}\phi_{,ij}\, d\mu \qquad (40) \\ =\ & \int_\Delta f^{1-n} \sum_{i,j=1}^m \mathbf{H}^u_{ij}\phi_{,ij}\, d\mu \geq 0. \qquad \square\end{aligned}$$

Definition 5.2. Let $\Delta \subset \mathbf{R}^m$ be a Delzant polytope, $n \neq 0, 1, 2$ and f a positive affine function on Δ. (Δ, f, n) is K-semistable if $\mathrm{DF}_{\Delta,f,n}(\phi) \geq 0$ for any piecewise linear convex function ϕ on Δ. (Δ, f, n) is K-polystable if it is K-semistable and the equality $\mathrm{DF}_{\Delta,f,n}(\phi) = 0$ is only possible for ϕ affine linear.

Since any piecewise linear convex function on Δ can be approximated by smooth convex functions on Δ, the existence of a constant $(g_J, f, -2, n)$-scalar curvature metric implies the K-semistability of (Δ, f, n).

We next consider compact toric surfaces and prove that the positivity of Donaldson–Futaki invariant for simple piecewise linear functions implies K-polystability. This is a generalization of the result by Donaldson [4] and Wang–Zhou [14, 15]. The proof is similar to the one given in [15], but to make this paper as self-contained as possible, we give a proof here.

Let $P \subset \mathbf{R}^m$ be an m-dimensional open convex polytope, P^* a union of P and the facets of P. Denote

$$\mathcal{C}_1 := \left\{ u : P^* \to \mathbf{R}, \text{ convex} \,\middle|\, \int_{\partial P} u\, d\sigma < \infty \right\}.$$

For positive bounded functions α, β on $\bar{P}$ and an affine function A on $\mathbf{R}^m$, we define the linear functional $\mathcal{L}$ on $\mathcal{C}_1$ as

$$\mathcal{L}(u) := \int_{\partial P} \alpha u\, d\sigma - \int_P A\beta u\, d\mu. \tag{41}$$

Theorem 5.3. *Suppose that $\mathcal{L}(f) = 0$ for any affine function f on $\mathbf{R}^m$. When $m = 2$, the following two conditions are equivalent.*

(1) $\mathcal{L}(u) \geq 0$ *for any* $u \in \mathcal{C}_1$ *and the equality holds if and only if* u *is affine.*
(2) $\mathcal{L}(u) > 0$ *for any simple piecewise linear convex function* u *with nonempty crease.*

Here a convex function u is simple piecewise linear, sPL for short, if $u = \max\{L, 0\}$ for a non-zero affine function L. The crease of sPL convex function u is the intersection of P and $\{L = 0\}$.

Proof. It is sufficient to prove that (2) implies (1). Suppose that $\mathcal{L}$ is positive for any sPL convex function with nonempty crease. Moreover we assume the case (1) does not occur, that is, one of the following holds:

∘ There exists $v \in \mathcal{C}_1$ such that $\mathcal{L}(v) < 0$.
∘ For any $u \in \mathcal{C}_1$, $\mathcal{L}(u) \geq 0$ and there exists $v \in \mathcal{C}_1 \setminus \{\text{affine function}\}$ such that $\mathcal{L}(v) = 0$.

We fix $p_0 \in P$ and denote

$$\tilde{\mathcal{C}}_1 := \left\{ u \in \mathcal{C}_1 \mid \int_{\partial P} \alpha u \, d\sigma = 1, \inf_P u = u(p_0) = 0 \right\}.$$

Since $\mathcal{L}$ vanishes on the set of affine functions and $\mathcal{L}(cu) = c\mathcal{L}(u)$ for any $c > 0$ and $u \in \mathcal{C}_1$, we may assume v in the condition above is an element of $\tilde{\mathcal{C}}_1$.

Lemma 5.4. *The functional* $\mathcal{L} : \tilde{\mathcal{C}}_1 \to \mathbf{R}$ *is bounded from below.*

Proof. By Lemma 5.1.3 in [4], there exists a constant $C > 0$ such that

$$\int_P u \, d\mu \leq C \int_{\partial P} u \, d\sigma$$

for all $u \in \tilde{\mathcal{C}}_1$. Since α, β are positive and bounded on $\bar{P}$

$$\int_P \beta u \, d\mu \leq \sup_{\bar{P}} \beta \int_P u \, d\mu \leq \frac{C \sup_{\bar{P}} \beta}{\inf_{\bar{P}} \alpha} =: C'$$

for $u \in \tilde{\mathcal{C}}_1$. Hence, on $\tilde{\mathcal{C}}_1$,

$$\mathcal{L}(u) = 1 - \int_P A\beta u \, d\mu \geq 1 - \max_{\bar{P}} |A| \int_P \beta u \, d\mu \geq 1 - \max_{\bar{P}} |A| C'. \qquad \square$$

By assumption, $\inf_{\tilde{\mathcal{C}}_1} \mathcal{L} \leq 0$. Moreover we see that there exists $u_0 \in \tilde{\mathcal{C}}_1$ which attains the infimum of $\mathcal{L}$ on $\tilde{\mathcal{C}}_1$ by the same argument with the proof of Lemma 4.2 in [15] as follows. Let $\{u_k\}$ be a sequence in $\tilde{\mathcal{C}}_1$ with $\lim_{k\to\infty} \mathcal{L}(u_k) = \inf_{\tilde{\mathcal{C}}_1} \mathcal{L}$. By Lemma 5.4 above and Corollary 5.2.5 in [4], there is the limit function u_0 convex on P^*. More precisely,

$$u_0(p) = \begin{cases} \lim\limits_{k\to\infty} u_k(p) & \text{if } p \in P \\ \lim\limits_{t \nearrow 1} u_0((1-t)p_0 + tp) & \text{if } p \text{ is in a facet of } P. \end{cases}$$

The limit function u_0 satisfies

$$\int_P A\beta u_0\,d\mu = \lim_{k\to\infty}\int_P A\beta u_k\,d\mu \text{ and } \inf_P u_0 = u_0(p_0) = 0.$$

By convexity, $\int_{\partial P}\alpha u_0\,d\sigma \le 1$. Suppose that $\int_{\partial P}\alpha u_0\,d\sigma < 1$. Then

$$\begin{aligned}\mathcal{L}(u_0) &= \int_{\partial P}\alpha u_0\,d\sigma - \int_P A\beta u_0\,d\mu < 1 - \int_P A\beta u_0\,d\mu\\ &= \lim_{k\to\infty}\mathcal{L}(u_k) = \inf_{\tilde{\mathcal{C}}_1}\mathcal{L} \le 0.\end{aligned}$$

On the other hand, since $\tilde{u}_0 := \left(\int_{\partial P}\alpha u_0\,d\sigma\right)^{-1}u_0 \in \tilde{\mathcal{C}}_1$,

$$\left(\int_{\partial P}\alpha u_0\,d\sigma\right)^{-1}\mathcal{L}(u_0) = \mathcal{L}(\tilde{u}_0) \ge \inf_{\tilde{\mathcal{C}}_1}\mathcal{L}.$$

Hence $\mathcal{L}(u_0) < \left(\int_{\partial P}\alpha u_0\,d\sigma\right)^{-1}\mathcal{L}(u_0)$. Since $\mathcal{L}(u_0) < 0$, $\int_{\partial P}\alpha u_0\,d\sigma > 1$. It is a contradiction. Therefore $u_0 \in \tilde{\mathcal{C}}_1$ and it attains the infimum of $\mathcal{L}$ on $\tilde{\mathcal{C}}_1$.

By the same argument with the proof of Lemma 4.3 in [15], we see that u_0 is a generalized solution to the degenerate Monge–Ampère equation

$$\det D^2 u = 0.$$

By convexity, $\mathcal{T} = \{x \in P \mid u_0(x) = 0\}$ is convex. Moreover any extreme point of $\mathcal{T}$ is a boundary point of P by Lemma 4.1 in [14]. Since P is two-dimensional, $\mathcal{T}$ is either a line segment through p_0 with both endpoints on ∂P or a convex polygon with vertices on ∂P. Note here that if the dimension of P is greater than two the convex set $\mathcal{T}$ may be more complicated. We set an affine function L on $\mathbf{R}^2$ as follows. When $\mathcal{T}$ is a line segment,

$$L(x) := \langle n, x - p_0\rangle,$$

where n is a unit normal vector of $\mathcal{T}$. When $\mathcal{T}$ is a polygon,

$$L(x) := \langle n, x - p_1\rangle,$$

where $p_1 \in \partial\mathcal{T} \setminus \partial P$ and n is the outer unit normal vector of $\partial\mathcal{T}$ at p_1. In either case, $\psi = \max\{0, L\}$ is a sPL convex function with nonempty crease.

We next define a function a as

$$a(p) = \lim_{t\searrow 0}\frac{u_0(p+tn) - u_0(p)}{t}.$$

Here $p \in \mathcal{T}$ when $\mathcal{T}$ is a line segment or p is in the edge of $\mathcal{T}$ containing p_1 when $\mathcal{T}$ is a polygon. By convexity of u_0, the limit exists and is nonnegative for any p.

Lemma 5.5. $a_0 := \inf a = 0$.

Proof. We give a proof only when $\mathcal{T}$ is a line segment since the case when $\mathcal{T}$ is a polygon is similar. Suppose $a_0 > 0$. Denote $u' := u_0 - a_0\psi$. Then $\int_{\partial P} \alpha u' \, d\sigma < 1$. By the definition of a_0, u' is convex on P^* and

$$\inf_P u' = u'(p_0) = u_0(p_0) - a_0\psi(p_0) = 0.$$

Since $\mathcal{L}(\psi) > 0$ by assumption,

$$\mathcal{L}(u_0) = \mathcal{L}(u') + a_0\mathcal{L}(\psi) > \mathcal{L}(u').$$

Hence, since $\tilde{u'} := \left(\int_{\partial P} \alpha u' \, d\sigma\right)^{-1} u' \in \tilde{\mathcal{C}}_1$,

$$0 \geq \mathcal{L}(u_0) > \mathcal{L}(u') > \mathcal{L}(\tilde{u'}).$$

This is a contradiction. $\square$

By the definition of $\mathcal{T}$ and L, u_0 is positive on $P \cap \{L > 0\}$. For any $\varepsilon > 0$, $G_\varepsilon := \{x \in P \,|\, u_0(x) < \varepsilon\psi(x)\}$ is nonempty because $a_0 = 0$. Since $\mathcal{T} \subset \{L \leq 0\}$, there exists $\delta(\varepsilon) > 0$ such that $G_\varepsilon \subset \{0 \leq L < \delta(\varepsilon)\}$ and $\lim_{\varepsilon \searrow 0} \delta(\varepsilon) = 0$. Denote

$$u_1 := u_0\chi_-, \quad u_2 := (u_0 - \varepsilon\psi)\chi_+, \quad \tilde{u_2} := \max\{0, u_2\},$$

where

$$\chi_-(x) = \begin{cases} 1 & \text{when } L(x) < 0 \\ 0 & \text{otherwise} \end{cases}, \quad \chi_+ = 1 - \chi_-.$$

It is easy to see that $u_1 + \tilde{u_2} \geq 0$ is convex and $(u_1 + \tilde{u_2})(p_0) = 0$. Denote $\tilde{u} := u_1 + \tilde{u_2} + \varepsilon\psi$. Then we have

$$\tilde{u} - u_0 = \tilde{u_2} - u_2 = \begin{cases} -u_2 = \varepsilon L - u_0 \leq \varepsilon\delta(\varepsilon) & \text{on } G_\varepsilon, \\ 0 & \text{on } G_\varepsilon^c. \end{cases}$$

Hence there exists a positive constant C such that

$$\mathcal{L}(\tilde{u} - u_0) = \int_{\partial P} \alpha(\tilde{u} - u_0) \, d\sigma - \int_P A\beta(\tilde{u} - u_0) \, d\mu < C\varepsilon\delta(\varepsilon).$$

Therefore we have

$$\mathcal{L}(u_1 + \tilde{u_2}) = \mathcal{L}(\tilde{u}) - \varepsilon\mathcal{L}(\psi) < \mathcal{L}(u_0) + \varepsilon(C\delta(\varepsilon) - \mathcal{L}(\psi)) < \mathcal{L}(u_0).$$

for any sufficiently small $\varepsilon > 0$. Denote $u_3 := \left(\int_{\partial P} \alpha(u_1 + \tilde{u_2}) \, d\sigma\right)^{-1} (u_1 + \tilde{u_2}) \in \tilde{\mathcal{C}}_1$. Since $u_1 + \tilde{u_2} \leq u_0$, $\int_{\partial P} \alpha(u_1 + \tilde{u_2}) \, d\sigma \leq 1$. Therefore we obtain $\mathcal{L}(u_3) \leq \mathcal{L}(u_1 + \tilde{u_2}) < \mathcal{L}(u_0)$. This is a contradiction. This completes the proof of Theorem 5.3. $\square$

Finally we observe by numerical analysis that there exists a Killing vector field which gives a toric K-stable pair in the sense of Apostolov–Maschler.

Let Δ_p be the convex hull of $(0,0),(p,0),(p,1-p)$ and $(0,1)$ for $0<p<1$. By Delzant construction, the Kähler class of a toric Kähler metric on the one point blow up of $\mathbf{C}P^2$ corresponds to Δ_p up to multiplication of a positive constant.

Denote

$$\mathcal{P} := \{(a,b,c)\in \mathbf{R}^3 \,|\, c>0,\ ap+c>0,\ ap+b(1-p)+c>0,\ b+c>0\}.$$

An affine function $a\mu_1+b\mu_2+c$ is positive on Δ_p if and only if $(a,b,c)\in\mathcal{P}$. By the argument in Sections 3 and 4, $\mathrm{Fut}_{\Delta_p, a\mu_1+b\mu_2+c,-2,n}$ vanishes if and only if $(a,b,c)\in\mathcal{P}$ is a critical point of

$$EH_n(a,b,c) := \frac{\displaystyle\int_{\partial\Delta_p}(a\mu_1+b\mu_2+c)^{2-n}\,d\sigma}{\left(\displaystyle\int_{\Delta_p}(a\mu_1+b\mu_2+c)^{-n}\,d\mu\right)^{\frac{n-2}{n}}}.$$

For $n=4$, the authors identified in [8] such critical points as follows:

(a) $C\left(1,0,\dfrac{p(1-\sqrt{1-p})}{2\sqrt{1-p}+p-2}\right),\ C>0,\ 0<p<1,$

(b) $C\left(-1,0,\dfrac{p(3p\pm\sqrt{9p^2-8p})}{2(p\pm\sqrt{9p^2-8p})}\right),\quad C>0,\ \dfrac{8}{9}<p<1,$

(c) $C\left(-p^2+4p-2\pm\sqrt{F(p)}, \pm 2\sqrt{F(p)}, -p^2-2p+2\mp\sqrt{F(p)}\right),$
$C>0,\ 0<p<\alpha,$

where $\alpha\approx 0.386$ is a real root of

$$F(x) := x^4-4x^3+16x^2-16x+4=0.$$

For the affine functions corresponding to (a) and (b), LeBrun gave concrete examples of cKEM metrics in [12]. Hence $(\Delta_p, a\mu_1+b\mu_2+c,4)$ is K-polystable by Corollary 3 in [3]. On the other hand, in case (c), we do not know whether there exists cKEM metrics. Denote

$$\begin{aligned} f_p^\pm &= (-p^2+4p-2\pm\sqrt{F(p)})\mu_1 \pm 2\sqrt{F(p)}\mu_2 - p^2-2p+2\mp\sqrt{F(p)} \\ &=: a_p^\pm\mu_1 + b_p^\pm\mu_2 + c_p^\pm. \end{aligned}$$

By Theorem 5.3, if $\mathrm{DF}_{\Delta_p, f_p^\pm,4}(\phi)$ is positive for any sPL convex function ϕ, $(\Delta_p, f_p^\pm, 4)$ is K-polystable. According to the position of the boundary points $\mathbf{u},\mathbf{v}$ of creases, we divide into the following six cases.

1. $\mathbf{u} = (0, e)$, $\mathbf{v} = (p, f)$ $(0 \le e \le 1, 0 \le f \le 1 - p)$: In this case, the corresponding sPL convex function is $\phi = \max\{(f - e)\mu_1 - p\mu_2 + pe, 0\}$. Then

$$\int_{\partial\Delta_p} \frac{\phi}{(f_p^\pm)^3}\, d\sigma = \int_0^p \frac{(f-e)\mu_1 + pe}{(a_p^\pm \mu_1 + c_p^\pm)^3}\, d\mu_1 + \int_0^f \frac{p(f-\mu_2)}{(a_p^\pm p + b_p^\pm \mu_2 + c_p^\pm)^3}\, d\mu_2 + \int_0^e \frac{p(e-\mu_2)}{(b_p^\pm \mu_2 + c_p^\pm)^3}\, d\mu_2$$

and

$$\int_{\Delta_p} \frac{\phi}{(f_p^\pm)^5}\, d\mu = \int_0^p d\mu_1 \int_0^{\frac{f-e}{p}\mu_1 + e} \frac{(f-e)\mu_1 - p\mu_2 + pe}{(f_p^\pm)^5}\, d\mu_2$$

It is too long and complicated to give the full description of $\mathrm{DF}_{\Delta_p, f_p^\pm, 4}(\phi)$. We put the graph of $\mathrm{DF}_{\Delta_{0.1}, f_{0.1}^-, 4}$, as a function of (e, f), instead (Fig. 1). All graphics in this article are drawn by *Mathematica*.

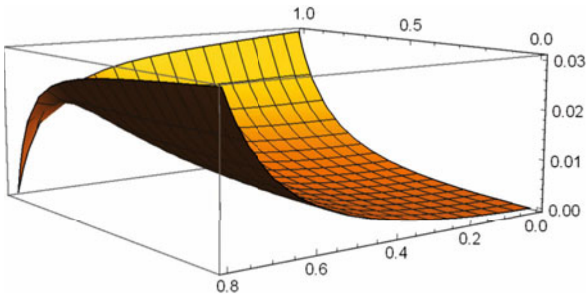


FIGURE 1.

2. $\mathbf{u} = (e, 0)$, $\mathbf{v} = (f, 1 - f)$ $(0 \le e \le p, 0 \le f \le p)$: In this case, the corresponding sPL convex function is $\phi = \max\{(f-1)\mu_1 + (f-e)\mu_2 + (1-f)e, 0\}$. Then

$$\int_{\partial\Delta_p} \frac{\phi}{(f_p^\pm)^3}\, d\sigma = \int_0^e \frac{(1-f)(e-\mu_1)}{(a_p^\pm \mu_1 + c_p^\pm)^3}\, d\mu_1 + \int_0^1 \frac{(f-e)\mu_2 + (1-f)e}{(b_p^\pm \mu_2 + c_p^\pm)^3}\, d\mu_2 + \int_0^f \frac{(f-1)\mu_1 + (f-e)(1-\mu_1) + (1-f)e}{(a_p^\pm \mu_1 + b_p^\pm (1-\mu_1) + c_p^\pm)^3}\, d\mu_1$$

and

$$\int_{\Delta_p} \frac{\phi}{(f_p^\pm)^5}\, d\mu = \int_0^{1-f} d\mu_2 \int_0^{\frac{f-e}{1-f}\mu_2 + e} \frac{(f-1)\mu_1 + (f-e)\mu_2 + (1-f)e}{(f_p^\pm)^5}\, d\mu_1 + \int_{1-f}^1 d\mu_2 \int_0^{1-\mu_2} \frac{(f-1)\mu_1 + (f-e)\mu_2 + (1-f)e}{(f_p^\pm)^5}\, d\mu_1 .$$

The graph of $\mathrm{DF}_{\Delta_{0.1}, f_{0.1}^-, 4}$ is as shown in Figure 2.

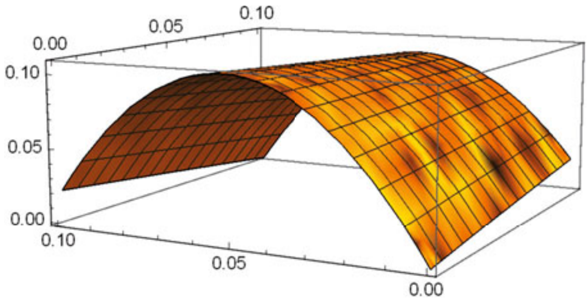


Figure 2.

3. $\mathbf{u} = (0, e)$, $\mathbf{v} = (f, 1-f)$ $(0 \le e \le 1, 0 \le f \le p)$: In this case, the corresponding sPL convex function is $\phi = \max\{(f+e-1)\mu_1 + f\mu_2 - fe, 0\}$. Then

$$\int_{\partial \Delta_p} \frac{\phi}{(f_p^{\pm})^3}\, d\sigma = \int_e^1 \frac{f(\mu_2 - e)}{(b_p^{\pm}\mu_2 + c_p^{\pm})^3}\, d\mu_2 + \int_0^f \frac{(f+e-1)\mu_1 + f(1-\mu_1) - fe}{(a_p^{\pm}\mu_1 + b_p^{\pm}(1-\mu_1) + c_p^{\pm})^3}\, d\mu_2$$

and

$$\int_{\Delta_p} \frac{\phi}{(f_p^{\pm})^5}\, d\mu = \int_0^f d\mu_1 \int_{\frac{1-f-e}{f}\mu_1 + e}^{1-\mu_1} \frac{(f+e-1)\mu_1 + f\mu_2 - fe}{(f_p^{\pm})^5}\, d\mu_2\,.$$

The graph of $\mathrm{DF}_{\Delta_{0.1}, f_{0.1}^-, 4}$ is as shown in Figure 3.

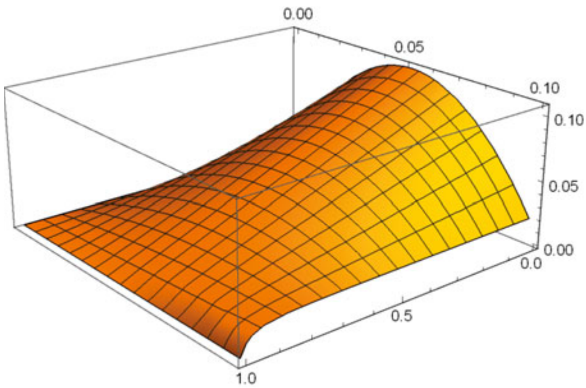


Figure 3.

4. $\mathbf{u} = (0, e)$, $\mathbf{v} = (f, 0)$ $(0 \le e \le 1, 0 \le f \le p)$: In this case, the corresponding sPL convex function is $\phi = \max\{-e\mu_1 - f\mu_2 + fe, 0\}$. Then

$$\int_{\partial\Delta_p} \frac{\phi}{(f_p^{\pm})^3}\, d\sigma = \int_0^f \frac{e(f-\mu_1)}{(a_p^{\pm}\mu_1 + c_p^{\pm})^3}\, d\mu_1 + \int_0^e \frac{f(e-\mu_2)}{(b_p^{\pm}\mu_2 + c_p^{\pm})^3}\, d\mu_2$$

and

$$\int_{\Delta_p} \frac{\phi}{(f_p^{\pm})^5}\, d\mu = \int_0^f d\mu_1 \int_0^{-\frac{e}{f}\mu_1 + e} \frac{-e\mu_1 - f\mu_2 + fe}{(f_p^{\pm})^5}\, d\mu_2\,.$$

The graph of $\mathrm{DF}_{\Delta_{0.1}, f_{0.1}^-, 4}$ is as shown in Figure 4.

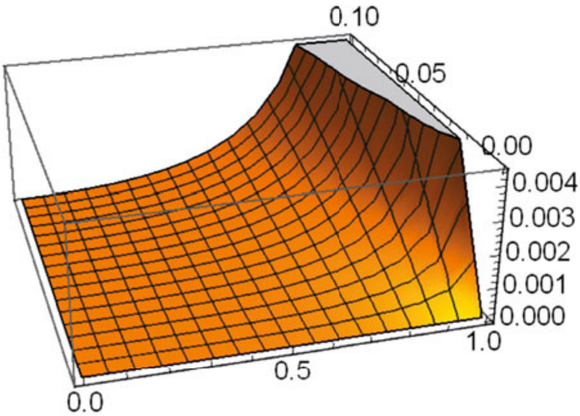


Figure 4.

5. $\mathbf{u} = (p, e)$, $\mathbf{v} = (f, 0)$ $(0 \le e \le 1-p, 0 \le f \le p)$: In this case, the corresponding sPL convex function is $\phi = \max\{e\mu_1 + (f-p)\mu_2 - fe, 0\}$. Then

$$\int_{\partial\Delta_p} \frac{\phi}{(f_p^{\pm})^3}\, d\sigma = \int_f^p \frac{e(\mu_1 - f)}{(a_p^{\pm}\mu_1 + c_p^{\pm})^3}\, d\mu_1 + \int_0^e \frac{(p-f)(e-\mu_2)}{(a_p^{\pm}p + b_p^{\pm}\mu_2 + c_p^{\pm})^3}\, d\mu_2$$

and

$$\int_{\Delta_p} \frac{\phi}{(f_p^{\pm})^5}\, d\mu = \int_f^p d\mu_1 \int_0^{\frac{e}{p-f}(\mu_1 - p) + e} \frac{e\mu_1 + (f-p)\mu_2 - fe}{(f_p^{\pm})^5}\, d\mu_2\,.$$

The graph of $\mathrm{DF}_{\Delta_{0.1}, f_{0.1}^-, 4}$ is as shown in Figure 5.

6. $\mathbf{u} = (p, e)$, $\mathbf{v} = (f, 1-f)$ $(0 \le e \le 1-p, 0 \le f \le p)$: In this case, the corresponding sPL convex function is $\phi = \max\{(1-e-f)(\mu_1 - p) + (p -$

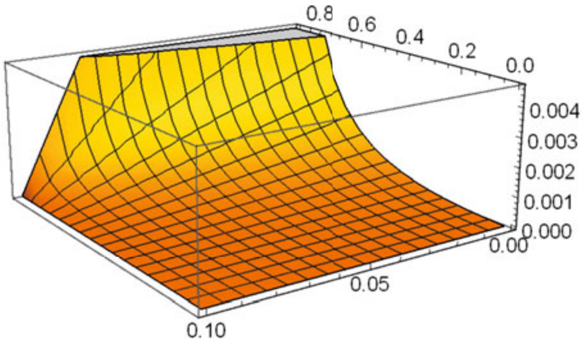


FIGURE 5.

$f)\mu_2 + (f-p)e, 0\}$. Then

$$\int_{\partial\Delta_p} \frac{\phi}{(f_p^{\pm})^3}\,d\sigma = \int_f^p \frac{(1-e-f)(\mu_1-p)+(p-f)(1-\mu_1)+(f-p)e}{(a_p^{\pm}\mu_1 + b_p^{\pm}(1-\mu_1)+c_p^{\pm})^3}\,d\mu_1 + \int_e^{1-p} \frac{(p-f)(\mu_2-e)}{(a_p^{\pm}p + b_p^{\pm}\mu_2 + c_p^{\pm})^3}\,d\mu_2$$

and

$$\int_{\Delta_p} \frac{\phi}{(f_p^{\pm})^5}\,d\mu = \int_f^p d\mu_1 \int_{\frac{e+f-1}{p-f}(\mu_1-p)+e}^{1-\mu_1} \frac{(1-e-f)(\mu_1-p)+(p-f)\mu_2+(f-p)e}{(f_p^{\pm})^5}\,d\mu_2\,.$$

The graph of $\mathrm{DF}_{\Delta_{0.1}, f_{0.1}^-, 4}$ is as shown in Figure 6.

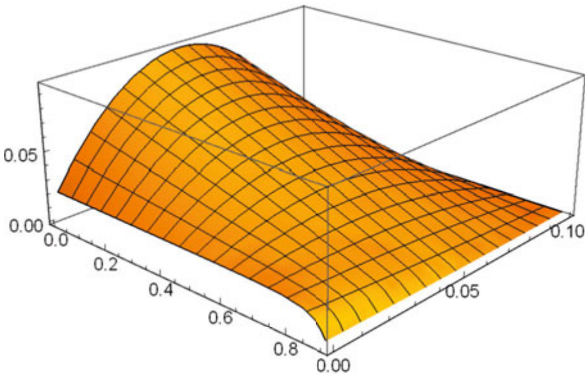


FIGURE 6.

Looking at the graphs, $(\Delta_p, f_p^{\pm}, 4)$ must be K-polystable. By Theorem 5 in [3], cKEM metrics with Killing potential $f_p^{\pm}$ ought to exist. We leave this problem to the interested readers.

References

[1] M. Abreu: Kähler geometry of toric varieties and extremal metrics. Int. J. Math., 9(1998), 641–651.

[2] V. Apostolov, D.M.J. Calderbank, P. Gauduchon and E. Legendre: Levi–Kähler reduction of CR structures, products of spheres, and toric geometry, arXiv:1708.05253.

[3] V. Apostolov and G. Maschler: Conformally Kähler, Einstein–Maxwell geometry, arXiv preprint arXiv:1512.06391.

[4] S.K. Donaldson: Scalar curvature and stability of toric varieties, J. Differential Geometry, 62(2002), 289–349.

[5] A. Futaki: Kähler–Einstein metrics and integral invariants, Lecture Notes in Math., vol. 1314, Springer-Verlag, Berlin-Heidelberg-New York,(1988).

[6] A. Futaki and S. Morita: Invariant polynomials of the automorphism group of a compact complex manifold, J. Diff. Geom., 21(1985), 135–142.

[7] A. Futaki and T. Mabuchi: Moment maps and symmetric multilinear forms associated with symplectic classes, Asian J. Math., 6(2002), 349–372.

[8] A. Futaki and H. Ono: Volume minimization and Conformally Kähler, Einstein–Maxwell geometry. To appear in J. Math. Soc. Japan. preprint, arXiv:1706.07953.

[9] A. Futaki and H. Ono: Volume minimization and obstructions to solving some problems in Kähler geometry, preprint.

[10] A. Lahdili: Automorphisms and deformations of conformally Kähler, Einstein–Maxwell metrics. arXiv:1708.01507.

[11] A. Lahdili: Conformally Kähler, Einstein–Maxwell metrics and boundedness of the modified Mabuchi-functional. arXiv:1710.00235

[12] C. LeBrun: The Einstein–Maxwell equations and conformally Kähler geometry, Commun. Math. Phys., 344, 621–653 (2016).

[13] M. Lejmi and M. Upmeier: Integrability theorems and conformally constant Chern scalar curvature metrics in almost Hermitian geometry, arXiv:1703.01323.

[14] X.-J. Wang and B. Zhou: On the existence and nonexistence of extremal metrics on toric Kähler surfaces. Adv. Math. 226 (2011), no. 5, 4429–4455.

[15] X.-J. Wang and B. Zhou: K-stability and canonical metrics on toric manifolds. Bull. Inst. Math. Acad. Sin. (N.S.) 9(2014), no. 1, 85–110.

Akito Futaki
Graduate School of Mathematical Sciences
The University of Tokyo
3-8-1 Komaba Meguro-ku
Tokyo 153-8914, Japan
e-mail: `afutaki@ms.u-tokyo.ac.jp`

Current address:
Yau Mathematical Sciences Center
Tsinghua University
Hai Dian District
Beijing 100084, P.R. China
e-mail: `futaki@tsinghua.edu.cn`

Hajime Ono
Department of Mathematics
Saitama University
255 Shimo-Okubo, Sakura-Ku
Saitama 380-8570, Japan
e-mail: `hono@rimath.saitama-u.ac.jp`

Progress in Mathematics, Vol. 333, 113–135

Local Moduli of Scalar-flat Kähler ALE Surfaces

Jiyuan Han and Jeff A. Viaclovsky

Abstract. In this article, we give a survey of our construction of a local moduli space of scalar-flat Kähler ALE metrics in complex dimension 2. We also prove an explicit formula for the dimension of this moduli space on a scalar-flat Kähler ALE surface which deforms to the minimal resolution of $\mathbb{C}^2/\Gamma$, where Γ is a finite subgroup of U(2) without complex reflections, in terms of the embedding dimension of the singularity.

Mathematics Subject Classification (2010). 53C55, 53C25.

Keywords. Scalar-flat Kähler, asymptotically locally Euclidean.

1. Introduction

In this article, the main objects of interest will be a certain class of complete non-compact Kähler metrics. In the following, Γ will always be a finite subgroup of U(2) containing no complex reflections.

Definition 1.1. Let (X, g, J) be a Kähler surface (X, g, J) of complex dimension 2, with metric g and complex structure J. We say that (X, g, J) is asymptotically locally Euclidean (ALE) if there exists a compact subset $K \subset X$, a real number $\mu > 0$, and a diffeomorphism $\psi : X \setminus K \to (\mathbb{R}^4 \setminus \overline{B})/\Gamma$, such that for each multi-index $\mathcal{I}$ of order $|\mathcal{I}|$

$$\partial^{\mathcal{I}}(\psi_*(g) - g_{\mathrm{Euc}}) = O(r^{-\mu-|\mathcal{I}|}), \tag{1.1}$$

as $r \to \infty$. In the above, B denotes a ball centered at the origin, and g_{Euc} denotes the Euclidean metric.

The number μ is referred to as the order of g. It was shown in [HL16] that for any ALE Kähler metric of order μ, there exist ALE coordinates for which

$$\partial^{\mathcal{I}}(J - J_{\mathrm{Euc}}) = O(r^{-\mu-|\mathcal{I}|}), \tag{1.2}$$

The authors were partially supported by NSF Grant DMS-1811096.

for any multi-index $\mathcal{I}$ as $r \to \infty$, where J_{Euc} is the standard complex structure on Euclidean space. This follows because the Kähler assumption implies that J is parallel.

In this definition, we only assumed that the metric is Kähler. A natural condition is that the metric be in addition scalar-flat. Such metrics are then *extremal* in the sense of Calabi [Cal85]. These spaces arise naturally as "bubbles" in orbifold compactness theorems for sequences of extremal Kähler metrics [And89, BKN89, CW11, CLW08, Nak94, Tia90, TV05a, TV05b, TV08]. Furthermore, they arise in a number of natural gluing constructions for extremal Kähler metrics [ALM15, ALM16, AP06, APS11, BR15, Szé12, RS05, RS09].

We note that in the case of scalar-flat Kähler ALE metrics, it is known that there exists an ALE coordinate system for which the order of such a metric is at least 2 [LM08].

There are many known examples of scalar-flat Kähler ALE metrics:

- SU(2) case: when $\Gamma \subset \mathrm{SU}(2)$, Kronheimer has constructed families of hyperkähler ALE metrics [Kro89a] on manifolds diffeomorphic to the minimal resolution of $\mathbb{C}^2/\Gamma$. In [Kro89b], Kronheimer also proved a Torelli-type theorem classifying hyperkähler ALE surfaces. In the A_k case, these metrics were previously discovered by Eguchi–Hanson for $k = 1$ [EH79], and by Gibbons–Hawking for all $k \geq 1$ [GH78].
- Cyclic case: For the $\frac{1}{p}(1, q)$-action, Calderbank–Singer constructed a family of scalar-flat Kähler ALE metrics on the minimal resolution of any cyclic quotient singularity [CS04]. These metrics are toric and come in families of dimension $k-1$, where k is the length of the corresponding Hirzebruch–Jung algorithm. For $q = 1$ and $q = p - 1$, these metrics are the LeBrun negative mass metrics and the toric multi-Eguchi–Hanson metrics, respectively [LeB88, GH78].
- Non-cyclic non-SU(2) case: The existence of scalar-flat Kähler metrics on the minimal resolution of $\mathbb{C}^2/\Gamma$, was shown by Lock–Viaclovsky [LV14].

A natural question is whether the scalar-flat Kähler property is preserved under small deformations of complex structure. In [HV16], we showed that for any scalar-flat Kähler ALE surface, all small deformations of complex structure admit scalar-flat Kähler ALE metrics, and so do all small deformations of the Kähler class. An informal statement is the following.

Theorem 1.2. *Let (X, g, J) be a scalar-flat Kähler ALE surface. Then there is a finite-dimensional family $\mathfrak{F}$ of scalar-flat Kähler ALE metrics near g, parametrized by a small ball in $\mathbb{R}^d$, for some integer d. This family $\mathfrak{F}$ is "versal" in the following sense: it contains all possible scalar-flat Kähler ALE metrics "near" to the given scalar-flat Kähler ALE metric, up to diffeomorphisms which are sufficiently close to the identity.*

A more precise statement of this theorem can be found in Section 2 below. The family $\mathfrak{F}$ is not "universal" since it is possible that 2 metrics in $\mathfrak{F}$ could be isometric. However, the orbit space of the group of biholomorphic isometries does give a universal moduli space, an informal statement of which is the following.

Theorem 1.3. *The group $\mathfrak{G}$ of holomorphic isometries of (X, g, J) acts on $\mathfrak{F}$, and each orbit represents a unique isometry class of metric up to the action of diffeomorphisms which are sufficiently close the identity.*

Again, a more precise statement can be found in Section 2 below. As a consequence, the quotient space $\mathfrak{M} = \mathfrak{F}/\mathfrak{G}$ is the "local moduli space of scalar-flat Kähler ALE metrics near g." The local moduli space $\mathfrak{M}$ is not a manifold in general, but its dimension is in fact well defined, and we define $m = \dim(\mathfrak{M})$.

1.1. Deformations of the minimal resolution

As mentioned above, there are families of examples of scalar-flat Kähler ALE metrics on minimal resolutions of isolated quotient singularities. We next recall the definition of a minimal resolution.

Definition 1.4. Let $\Gamma \subset \mathrm{U}(2)$ be as above. A smooth complex surface X is called a *minimal resolution* of $\mathbb{C}^2/\Gamma$ if there is a holomorphic mapping $\pi : X \to \mathbb{C}^2/\Gamma$ such that the restriction $\pi : X \setminus \pi^{-1}(0) \to \mathbb{C}^2/\Gamma \setminus \{0\}$ is a biholomorphism, and the set $\pi^{-1}(0)$ is a divisor in X containing no -1 curves.

The divisor $\pi^{-1}(0)$ is called the *exceptional divisor* of the resolution. In the cyclic case, the exceptional divisor is a string of rational curves with normal crossing singularities, and these are known are Hirzebruch–Jung strings. In the case that Γ is non-cyclic, the exceptional divisor is a tree of rational curves with normal crossing singularities [Bri68]. There are three Hirzebruch–Jung strings attached to a single curve, called the *central rational curve*. The self-intersection number of this curve will be denoted $-b_\Gamma$, and the total number of rational curves will be denoted by k_Γ.

In the special case of a minimal resolution, our main result can be stated as follows.

Theorem 1.5. *Let (X, g, J) be any scalar-flat Kähler ALE metric on the minimal resolution of $\mathbb{C}^2/\Gamma$, where $\Gamma \subset \mathrm{U}(2)$ is as above. Define*

$$j_\Gamma = 2\sum_{i=1}^{k_\Gamma}(e_i - 1), \tag{1.3}$$

where $-e_i$ is the self-intersection number of the ith rational curve, and k_Γ is the number of rational curves in the exceptional divisor, and let

$$d_\Gamma = j_\Gamma + k_\Gamma. \tag{1.4}$$

Then there is a family, $\mathfrak{F}$, parametrized by a ball in $\mathbb{R}^{d_\Gamma}$, of scalar-flat Kähler metrics near g which is "versal". The group $\mathfrak{G}$ of holomorphic isometries of (X, g, J) acts on $\mathfrak{F}$, and the dimension m_Γ of the local moduli space $\mathfrak{M} = \mathfrak{F}/\mathfrak{G}$ is given in Table 1.1, where e_Γ is the embedding dimension of $\mathbb{C}^2/\Gamma$.

TABLE 1.1. Dimension of local moduli space of scalar-flat Kähler metrics

$\Gamma \subset \mathrm{U}(2)$	d_Γ	m_Γ
$\frac{1}{3}(1,1)$	5	2
$\frac{1}{p}(1,1), p \geq 4$	$2p-1$	$2p-5$
$\frac{1}{p}(1,q), q \neq 1, p-1$	$j_\Gamma + k_\Gamma$	$2e_\Gamma + 3k - 8$
non-cyclic, not in $\mathrm{SU}(2)$	$j_\Gamma + k_\Gamma$	$2e_\Gamma + 3k - 7$

A description of the possible groups Γ and other explicit formulas for m_Γ can be found in Section 4 below.

Remark 1.6. We did not include the SU(2) case in the above since the dimension of the moduli space of hyperkähler metrics is known to be $3k-3$ in the A_k, D_k and E_k cases for $k \geq 2$, and equal to 1 in the A_1 case [Kro86]. Our method of parametrizing by complex structures and Kähler classes overcounts in this case, since a hyperkähler metric is Kähler with respect to a 2-sphere of complex structures, see Section 4 for some further remarks. For other related results in the Ricci-flat case, see [CH15, Şuv12].

It turns out that the moduli count in Theorem 1.5 is correct not just for the minimal resolution, but for any generic scalar-flat Kähler ALE surface which can be continuously deformed to the minimal resolution.

Theorem 1.7 ([HV16]). *Let (X, g, J) be any scalar-flat Kähler ALE surface which deforms to the minimal resolution of $\mathbb{C}^2/\Gamma$ through a path (X, g_t, J_t) $(0 \leq t \leq 1)$, where $g_1 = g$, g_0 is the minimal resolution, and $\|g_t - g_s\|_{C^{k,\alpha}_\delta(g_0)} \leq C \cdot |s-t|$ with $C > 0$ a uniform constant for any $0 \leq s, t \leq 1$, $k \geq 4$, $-2 < \delta < -1$. If $\mathfrak{G}(g) = \{e\}$ then the local moduli space $\mathfrak{F}$ is smooth near g and is a manifold of dimension $m = m_\Gamma$.*

The proof of this theorem is more or less a direct application of Theorem 1.5 together with the basic fact that the index of a strongly continuous family of Fredholm operators is constant.

Remark 1.8. It was recently shown that Kähler ALE surface with group $\Gamma \subset \mathrm{U}(2)$ is birational to a deformation of $\mathbb{C}^2/\Gamma$ [HRŞ16]. There are several possible components of the deformation of such a cone, so the above result gives the dimension of the moduli space for the "Artin component" of deformations of $\mathbb{C}^2/\Gamma$, which is the component with maximal dimension.

2. Construction of the local moduli space

In this section, we will give a survey of the main results in [HV16]. We first recall some basic facts regarding deformations of complex structures. For a complex manifold (X, J), let $\Lambda^{p,q}$ denote the bundle of (p,q)-forms, and let Θ denote the holomorphic tangent bundle. The deformation complex corresponds to a real complex as shown in the commutative diagram

$$
\begin{tikzcd}
\Gamma(\Theta) \arrow[r, "\overline{\partial}"] \arrow[d, "Re"] & \Gamma(\Lambda^{0,1} \otimes \Theta) \arrow[r, "\overline{\partial}"] \arrow[d, "Re"] & \Gamma(\Lambda^{0,2} \otimes \Theta) \arrow[d, "Re"] \\
\Gamma(TX) \arrow[r, "Z \mapsto -\frac{1}{2} J \circ \mathcal{L}_Z J"] & \Gamma(\mathrm{End}_a(TX)) \arrow[r, "I \mapsto \frac{1}{4} J \circ N'_J(I)"] & \Gamma\big(\{\Lambda^{0,2} \otimes \Theta \oplus \Lambda^{2,0} \otimes \overline{\Theta}\}_{\mathbb{R}}\big),
\end{tikzcd}
\tag{2.1}
$$

where $\mathfrak{L}_Z J$ is the Lie derivative of J,

$$\mathrm{End}_a(TX) = \{I \in \mathrm{End}(TX) : IJ = -JI\}, \tag{2.2}$$

and N'_J is the linearization of Nijenhuis tensor

$$N(X, Y) = 2\{[JX, JY] - [X, Y] - J[X, JY] - J[JX, Y]\} \tag{2.3}$$

at J. Each isomorphism Re is simply taking the real part of a section. If g is a Hermitian metric compatible with J, then let $\square$ denote the $\bar{\partial}$-Laplacian

$$\square \equiv \bar{\partial}^* \bar{\partial} + \bar{\partial} \bar{\partial}^*, \tag{2.4}$$

where $\bar{\partial}^*$ denotes the formal L^2-adjoint. Each complex bundle in the diagram (2.1) admits a $\square$-Laplacian, and these correspond to real Laplacians on each real bundle in (2.1). We will use the same $\square$-notation for these real Laplacians.

We next define the spaces of harmonic sections which will appear in the statement of the main result.

Definition 2.1. Let (X, g, J) be a Kähler ALE surface. For any bundle E in the diagram (2.1), and $\tau \in \mathbb{R}$, define

$$\mathcal{H}_\tau(X, E) = \{\theta \in \Gamma(X, E) : \square\theta = 0, \theta = O(r^\tau) \text{ as } r \to \infty\}. \tag{2.5}$$

Define

$$\mathbb{W} = \{Z \in \mathcal{H}_1(X, TX) \mid \mathfrak{L}_Z g = O(r^{-1}),\ \mathfrak{L}_Z J = O(r^{-3}),\ \text{as } r \to \infty\}. \tag{2.6}$$

Finally, define the real subspace

$$\mathcal{H}_{\mathrm{ess}}(X, \mathrm{End}_a(TX)) \subset \mathcal{H}_{-3}(X, \mathrm{End}_a(X)) \tag{2.7}$$

to be the L^2-orthogonal complement in $\mathcal{H}_{-3}(X, \mathrm{End}_a(X))$ of the subspace

$$\mathbb{V} = \{\theta \in \mathcal{H}_{-3}(X, \mathrm{End}_a(TX)) \mid \theta = J \circ \mathfrak{L}_Z J,\ Z \in \mathbb{W}\}. \tag{2.8}$$

The subscript *ess* in (2.7) is short for essential, and is necessary because there is a gauge freedom of Euclidean motions in the definition of ALE coordinates, so that element of $\mathbb{V}$ are not really essential deformations, i.e., they can be gauged away.

To state the main result precisely, we need to define weighted Hölder spaces.

Definition 2.2. Let E be a tensor bundle on X, with Hermitian metric $\|\cdot\|_h$. Let φ be a smooth section of E. We fix a point $p_0 \in X$, and define $r(p)$ to be the distance between p_0 and p. Then define

$$\|\varphi\|_{C^0_\delta} := \sup_{p\in X}\left\{\|\varphi(p)\|_h \cdot (1+r(p))^{-\delta}\right\} \tag{2.9}$$

$$\|\varphi\|_{C^k_\delta} := \sum_{|\mathcal{I}|\le k}\sup_{p\in X}\left\{\|\nabla^{\mathcal{I}}\varphi(p)\|_h \cdot (1+r(p))^{-\delta+|\mathcal{I}|}\right\}, \tag{2.10}$$

where $\mathcal{I} = (i_1, \dots, i_n), |\mathcal{I}| = \sum_{j=1}^n i_j$. Next, define

$$[\varphi]_{C^\alpha_{\delta-\alpha}} := \sup_{0<d(x,y)<\rho_{\mathrm{inj}}}\left\{\min\{r(x), r(y)\}^{-\delta+\alpha}\frac{\|\varphi(x)-\varphi(y)\|_h}{d(x,y)^\alpha}\right\}, \tag{2.11}$$

where $0 < \alpha < 1$, ρ_{inj} is the injectivity radius, and $d(x, y)$ is the distance between x and y. The meaning of the tensor norm is to use parallel transport along the unique minimal geodesic from y to x, and then take the norm of the difference at x. The weighted Hölder norm is defined by

$$\|\varphi\|_{C^{k,\alpha}_\delta} := \|\varphi\|_{C^k_\delta} + \sum_{|\mathcal{I}|=k}[\nabla^{\mathcal{I}}\varphi]_{C^\alpha_{\delta-k-\alpha}}, \tag{2.12}$$

and the space $C^{k,\alpha}_\delta(X, E)$ is the closure of $\{\varphi \in C^\infty(X, E) : \|\varphi\|_{C^{k,\alpha}_\delta} < \infty\}$.

The main result of [HV16] is the following.

Theorem 2.3 ([HV16]). *Let (X, g, J) be a scalar-flat Kähler ALE surface. Let $-2 < \delta < -1$, $0 < \alpha < 1$, and k an integer with $k \ge 4$ be fixed constants. Let $B^1_{\epsilon_1}$ denote an ϵ_1-ball in $\mathcal{H}_{\mathrm{ess}}(X, \mathrm{End}_a(TX))$, $B^2_{\epsilon_2}$ denote an ϵ_2-ball in $\mathcal{H}_{-3}(X, \Lambda^{1,1})$ (both using the L^2-norm). Then there exists $\epsilon_1 > 0$ and $\epsilon_2 > 0$ and a family $\mathfrak{F}$ of scalar-flat Kähler metrics near g, parametrized by $B^1_{\epsilon_1} \times B^2_{\epsilon_2}$, that is, there is a differentiable mapping*

$$F : B^1_{\epsilon_1} \times B^2_{\epsilon_2} \to Met(X), \tag{2.13}$$

into the space of smooth Riemannian metrics on X, with $\mathfrak{F} = F(B^1_{\epsilon_1} \times B^2_{\epsilon_2})$ satisfying the following "versal" property: there exists a constant $\epsilon_3 > 0$ such that for any scalar-flat Kähler metric $\tilde{g} \in B_{\epsilon_3}(g)$, there exists a diffeomorphism $\Phi : X \to X$, $\Phi \in C^{k+1,\alpha}_{\mathrm{loc}}$, such that $\Phi^*\tilde{g} \in \mathfrak{F}$, where

$$B_{\epsilon_3}(g) = \{g' \in C^{k,\alpha}_{\mathrm{loc}}(S^2(T^*X)) \mid \|g - g'\|_{C^{k,\alpha}_\delta(S^2(T^*X))} < \epsilon_3\}. \tag{2.14}$$

2.1. Outline of Proof of Theorem 2.3

The main steps in the proof of Theorem 2.3 are the following.

Step I: One first analyzes deformations of complex structures using an adaptation of Kuranishi's theory [Kur65], to ALE spaces. To first order, the almost complex structures near a given ALE Kähler metric are in correspondence with sections in $\Gamma(\Lambda^{0,1} \otimes \Theta)$. The integrable complex structures solve a nonlinear elliptic equation, modulo diffeomorphisms. By imposing a divergence-free gauging condition, we obtain a finite-dimensional Kuranishi family which is parametrized by decaying harmonic sections in $\mathcal{H}_{-3}(X, \Lambda^{0,1} \otimes \Theta)$. Unobstructedness follows from a vanishing theorem, which relies on some analysis of the complex analytic compactifications of Kähler ALE spaces, due to Hein–LeBrun–Maskit [HL16, LM08]. In the appendix of this paper, we provide some details of a decay estimate needed in the proof of [HV16, Proposition 3.3]. Another important point is that since the manifold is non-compact, the sheaf cohomology group $H^1(X, \Theta)$, which vanishes in the Stein case, should be replaced by an appropriate space of decaying harmonic forms.

Step II: Several key results about gauging and diffeomorphisms are needed to prove "versality" of the family constructed. Our main infinitesimal slicing result is the following.

Lemma 2.4. *Let (X, J_0, g_0) be a Kähler ALE surface with $J_0, g_0 \in C^\infty$. There exists an $\epsilon_1' > 0$ such that for any complex structure $\|J_1 - J_0\|_{C^{k,\alpha}_\delta} < \epsilon_1'$, where $k \geq 3, \alpha \in (0,1), \delta \in (-2,-1)$, there exists a unique diffeomorphism Φ, of the form Φ_Y (see (2.17) below) for $Y \in C^{k+1,\alpha}_{\delta+1}(TX)$ such that $\Phi_Y^*(J_1)$ is in the divergence-free gauged Kuranishi family.*

Essentially, this shows that the divergence-free gauge gives a local slice transverse to the "small" diffeomorphism group action. However, a more refined gauging procedure is needed in order to construct the Kuranishi family of "essential" deformations. As stated above, this refined gauging is necessary because of the freedom of Euclidean motions in the definition of an ALE metric, which means that there are decaying elements in the kernel of the linearized operator which can be written as Lie derivatives of linearly growing vector fields. These directions are not true moduli directions, and we show that they can be ignored modulo diffeomorphisms. Thus we can restrict attention to the subspace of essential deformations defined in (2.7) above.

Step III: Next, one needs to generalize Kodaira–Spencer's stability theorem for Kähler structures [KS60] to the ALE setting, to prove that the above deformations retain the ALE Kähler property. This was proved using some arguments similar to that of Biquard–Rollin [BR15].

Step IV: To study the deformations of the scalar-flat Kähler structure, we then adapted the LeBrun–Singer–Simanca theory of deformations of extremal Kähler metrics to the ALE setting [LS93, LS94]. Denote $S(\omega_0 + \sqrt{-1}\partial\bar{\partial}f)$ as the scalar curvature of X with metric $\omega_0 + \sqrt{-1}\partial\bar{\partial}f$. We consider S as mapping between weighted Hölder spaces,

$$\begin{aligned} S : C^{k,\alpha}_\epsilon(X) &\to C^{k-4,\alpha}_{\epsilon-4}(X) \\ f &\mapsto S(\omega_0 + \sqrt{-1}\partial\bar{\partial}f). \end{aligned} \tag{2.15}$$

If ω_0 is scalar-flat, the linearized operator is $L(f) = -(\bar{\partial}\bar{\partial}^\#)^*(\bar{\partial}\bar{\partial}^\#)(f)$, where the operator $\bar{\partial}^\# f = g_0^{i,\bar{j}}\bar{\partial}_j f$. We showed that the linearized map is surjective for

$0 < \epsilon < 1$, and then an application of the implicit function theorem completes the proof.

2.2. Universality

As mentioned above, the family $\mathfrak{F}$ is not necessarily "universal", because some elements in $\mathfrak{F}$ might be isometric. To construct a universal moduli space, we need to describe a neighborhood of the identity in the space of diffeomorphisms. If (X, g) is an ALE metric, and Y is a vector field on X, the Riemannian exponential mapping $\exp_p : T_pX \to X$ induces a mapping

$$\Phi_Y : X \to X \tag{2.16}$$

by

$$\Phi_Y(p) = \exp_p(Y). \tag{2.17}$$

If $Y \in C^{k,\alpha}_s(TX)$ has sufficiently small norm, ($s < 0$ and k will be determined in specific cases) then Φ_Y is a diffeomorphism. We will use the correspondence $Y \mapsto \Phi_Y$ to parametrize a neighborhood of the identity, analogous to [Biq06].

Definition 2.5. We say that $\Phi : X \to X$ is a *small diffeomorphism* if Φ is of the form $\Phi = \Phi_Y$ for some vector field Y satisfying

$$\|Y\|_{C^{k+1,\alpha}_{\delta+1}} < \epsilon_4 \tag{2.18}$$

for some $\epsilon_4 > 0$ sufficiently small which depends on ϵ_3.

The following result shows that after taking a quotient by an action of the holomorphic isometries of the central fiber (X, g, J), the family $\mathfrak{F}$ is in fact universal (up to small diffeomorphisms).

Theorem 2.6 ([HV16]). *Let (X, g, J) be as in Theorem 2.3, and let $\mathfrak{G}$ denote the group of holomorphic isometries of (X, g, J). Then there is an action of $\mathfrak{G}$ on $\mathfrak{F}$ with the following properties.*

- *Two metrics in $\mathfrak{F}$ are isometric if they are in the same orbit of $\mathfrak{G}$.*
- *If two metrics in $\mathfrak{F}$ are isometric by a small diffeomorphism then they must be the same.*

This theorem was proved in [HV16] more or less by keeping track of the action of $\mathfrak{G}$ in every step of the proof of Theorem 2.3. Since each orbit represents a unique isometry class of metric (up to small diffeomorphism), we will refer to the quotient $\mathfrak{M} = \mathfrak{F}/\mathfrak{G}$ as the "local moduli space of scalar-flat Kähler ALE metrics near g." The local moduli space $\mathfrak{M}$ is not a manifold in general, but since $\mathfrak{F}$ is of finite dimension, and $\mathfrak{G}$ is a compact group action on $\mathfrak{F}$, the dimension m of $\mathfrak{M} = \mathfrak{F}/\mathfrak{G}$ is well defined. In the non-hyperkähler case,

$$m = d - (\text{the dimension of a maximal orbit of } \mathfrak{G}), \tag{2.19}$$

where

$$d = \dim_{\mathbb{R}} \big(\mathcal{H}_{\mathrm{ess}}(X, \mathrm{End}_a(TX))\big) + b_2(X), \tag{2.20}$$

where $b_2(X)$ is the second Betti number of X. (For the hyperkähler case, recall Remark 1.6.)

Remark 2.7. We note that the local moduli space of metrics contains small rescalings, i.e., $g \mapsto \frac{1}{c^2} g(c\cdot, c\cdot)$ for c close to 1. If one considers scaled metrics as equivalent (which we do not), then the dimension would decrease by 1.

3. The case of the minimal resolution

Let X denote the minimal resolution of $\mathbb{C}^2/\Gamma$, where Γ is a finite subgroup of $\mathrm{U}(2)$ without complex reflections. The divisor $E = \cup_i E_i$ is a union of irreducible components which are rational curves, with only normal crossing singularities. Let $\mathrm{Der}_E(X)$ denote the sheaf dual to logarithmic 1-forms along E (see [Kaw78]). We note that $\mathrm{Der}_E(X)$ is a locally free sheaf of rank 2, see [Wah75]. Away from E, this is clear. If $p \in E_i$, we can choose a holomorphic coordinate chart $\{z_1, z_2\}$ such that near p, $E_i = \{z_1 = 0\}$. Then local sections of $\mathrm{Der}_E(X)$ are generated by $\{z_1 \frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2}\}$.

Since E is composed of rational curves whose self-intersection numbers are negative, we have $H^0(E, \mathcal{O}_E(E)) = 0$. The short exact sequence

$$0 \to \mathrm{Der}_E(X) \to \Theta_X \to \mathcal{O}_E(E) \to 0, \tag{3.1}$$

then induces an exact sequence of cohomologies

$$0 \to H^1(X, \mathrm{Der}_E(X)) \to H^1(X, \Theta) \to H^1(E, \mathcal{O}_E(E)) \to H^2(X, \mathrm{Der}_E(X)). \tag{3.2}$$

By Siu's vanishing theorem ([Siu69]), since X is a non-compact σ-compact complex manifold, for any coherent analytic sheaf $\mathscr{F}$ on X, the top degree sheaf cohomology $H^2(X, \mathscr{F})$ is trivial. Consequently, $H^2(X, \mathrm{Der}_E(X)) = 0$,

In [HV16] we cited several papers from algebraic geometry [BKR88, Bri68, Lau73, Wah75], to conclude that $H^1(X, \mathrm{Der}_E(X)) = 0$. In this section, we will give a different proof of the following result, using some tools from geometric analysis.

Theorem 3.1. *For X the minimal resolution of $\mathbb{C}^2/\Gamma$, we have*

$$\dim_{\mathbb{C}}(H^1(X, \Theta)) = \sum_{j=1}^{k_\Gamma} (e_j - 1). \tag{3.3}$$

In relation to the construction of the moduli space of scalar-flat Kähler ALE metrics, we need to construct a weighted version of Hodge theory, that links the sheaf cohomology with the decaying harmonic forms. Recall that in Theorem 2.3, the deformations of complex structure are parametrized by decaying harmonic sections in $\mathcal{H}_{-3}(X, \Lambda^{0,1} \otimes \Theta)$ which are "essential", that is, they are in the subspace $\mathbb{V}$. But, in the case of the minimal resolution, it turns out the dimension of this space is equal to the dimension of $H^1(X, \Theta)$.

Theorem 3.2 ([HV16]). *Let (X, g, J) denote the minimal resolution of $\mathbb{C}^2/\Gamma$ with any ALE Kähler metric g of order $\tau > 1$. Then*

$$H^1(X, \Theta) \cong \mathcal{H}_{-3}(X, \Lambda^{0,1} \otimes \Theta) \cong \mathcal{H}_{\mathrm{ess}}(X, \Lambda^{0,1} \otimes \Theta) \tag{3.4}$$

As a consequence, the dimension of the space of essential deformations is given by (3.3). This is very special to the case of the minimal resolution. In the Stein case, Theorem 3.2 is not true in general because the sheaf cohomology group necessarily vanishes.

3.1. Cyclic quotient singularity

We will first prove Theorem 3.1 in the case of a cyclic group, using a direct arguments involving only sheaf theory. Consider a cyclic quotient singularity of the form $\Gamma = \frac{1}{p}(1, q)$ $(p \geq q)$. We will first give some additional detail regarding the Hirzebruch–Jung resolutions. Details can be found in [Rei, Kol07].

The continued fraction described below in formula (4.2), can also be represented by lattice points

$$c_0 = (1, 0), c_1 = \frac{1}{p}(1, q), \ldots, c_{m+1} = (0, 1), \tag{3.5}$$

with iterative relation

$$\begin{pmatrix} c_i \\ c_{i+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & e_i \end{pmatrix} \cdot \begin{pmatrix} c_{i-1} \\ c_i \end{pmatrix}. \tag{3.6}$$

Meanwhile, the dual continued fraction $\frac{p}{p-q} = [a_1, \ldots, a_k]$ can be used to give the invariant polynomials:

$$u_0 = x^p, u_1 = x^{p-q}y, u_2, \ldots, u_k, u_{k+1} = y^p, \tag{3.7}$$

which satisfy the relation $u_{i-1}u_{i+1} = u_i^{a_i}$.

The polynomials $\{u_0, \ldots, u_{k+1}\}$ give an embedding of the cone in $\mathbb{C}^{k+2}$. Let

$$c_0 = (s_0, t_0), \ldots, c_{m+1} = (s_{m+1}, t_{m+1}) \tag{3.8}$$

be lattice points, where $s_0 = 0, t_0 = 1, s_m = 1, t_m = 0, s_{i+1} > s_i, t_{i+1} < t_i$. Let $\{\eta_i, \xi_i\}$ $(0 \leq i \leq m+1)$ be monomials forming the dual basis to $\{c_i, c_{i+1}\}$, i.e.,

$$c_i(\eta_i) = 1, c_i(\xi_i) = 0, c_{i+1}(\eta_i) = 0, c_{i+1}(\xi_i) = 1. \tag{3.9}$$

Proposition 3.3. *The numbers s_i, t_i satisfy the relation*

$$t_i s_{i+1} - t_{i+1} s_i = \frac{1}{p}. \tag{3.10}$$

Proof. We prove it by induction. First, note that $c_0 = (0, 1), c_1 = \frac{1}{p}(1, q)$. Then $t_0 s_1 - t_1 s_0 = \frac{1}{p}$. Next, assume that $t_{i-1}s_i - t_i s_{i-1} = \frac{1}{p}$. By the recursive formula $c_{i+1} + c_{i-1} = a_i c_i$, it follows that

$$(s_{i+1}, t_{i+1}) + (s_{i-1}, t_{i-1}) = a_i(s_i, t_i). \tag{3.11}$$

Then we have

$$s_{i+1} = a_i s_i - s_{i-1},\ t_{i+1} = a_i t_i - t_{i-1}. \tag{3.12}$$

So finally,

$$t_i s_{i+1} - t_{i+1} s_i = t_i(a_i s_i - s_{i-1}) - (a_i t_i - t_{i-1}) s_i = t_{i-1} s_i - t_i s_{i-1} = \frac{1}{p}. \tag{3.13}$$

□

By the formula (3.10), we have that $\eta_i = p \cdot (-t_{i+1}, s_{i+1}), \xi_i = p \cdot (t_i, -s_i)$. Then

$$\xi_i = \frac{x^{pt_i}}{y^{ps_i}}, \eta_i = \frac{y^{ps_{i+1}}}{x^{pt_{i+1}}}, \xi_{i+1} = \frac{x^{pt_{i+1}}}{y^{ps_{i+1}}}, \eta_{i+1} = \frac{y^{ps_{i+2}}}{x^{pt_{i+2}}}. \tag{3.14}$$

It follows that the coordinate transition from $\{\eta_i, \xi_i\}$ to $\{\eta_{i+1}, \xi_{i+1}\}$ for $\xi_{i+1} \neq 0$, is given by

$$\eta_i = \xi_{i+1}^{-1},\ \eta_{i+1} = \eta_i^{e_{i+1}} \xi_i,\ (0 \leq i \leq m-1) \tag{3.15}$$

which defines an acyclic cover $Y = Y_0 \cup Y_1 \cdots \cup Y_m$ of X satisfying

$$Y_i \cap Y_{i+1} \simeq \mathbb{C} \times \mathbb{C}^*,\ Y_i \cap Y_{i+k} = Y_i \cap Y_{i+1} \cdots \cap Y_{i+k}, \tag{3.16}$$

see [Rei, Theorem 3.2]. For use below, we record the following formulae:

$$\begin{aligned}
\frac{\partial}{\partial \eta_i} &= \frac{1}{\eta_i}\left(s_i x \frac{\partial}{\partial x} + t_i y \frac{\partial}{\partial y}\right)\\
\frac{\partial}{\partial \xi_i} &= \frac{1}{\xi_i}\left(s_{i+1} x \frac{\partial}{\partial x} + t_{i+1} y \frac{\partial}{\partial y}\right)\\
\frac{\partial}{\partial x} &= \frac{p}{x}\left(t_i \xi_i \frac{\partial}{\partial \xi_i} - t_{i+1} \eta_i \frac{\partial}{\partial \eta_i}\right)\\
\frac{\partial}{\partial y} &= \frac{p}{y}\left(-s_i \xi_i \frac{\partial}{\partial \xi_i} + s_{i+1} \eta_i \frac{\partial}{\partial \eta_i}\right).
\end{aligned} \tag{3.17}$$

With these preliminaries, we can now prove the following result.

Lemma 3.4. *When Γ is cyclic, $H^1(X, \mathrm{Der}_E(X)) = 0$ for the minimal resolution X of $\mathbb{C}^2/\Gamma$.*

Proof. From (3.15) above,

$$\frac{\partial}{\partial \xi_{i+1}} = -\eta_i^2 \frac{\partial}{\partial \eta_i} + e_{i+1} \eta_i \xi_i \frac{\partial}{\partial \xi_i}, \tag{3.18}$$

$$\frac{\partial}{\partial \eta_{i+1}} = \eta_i^{-e_{i+1}} \frac{\partial}{\partial \xi_i}. \tag{3.19}$$

The sections of $\mathrm{Der}_E(X)$ are generated by

$$\left\{\frac{\partial}{\partial \eta_i}, \xi_i \frac{\partial}{\partial \xi_i}\right\},\ \left\{\frac{\partial}{\partial \xi_{i+1}}, \eta_{i+1} \frac{\partial}{\partial \eta_{i+1}}\right\} \tag{3.20}$$

on Y_i, Y_{i+1} respectively. For $\theta_i \in \Gamma(Y_i, \mathrm{Der}_E(X))$, θ_i can be expanded as a Laurent series:

$$\theta_i = \sum_{k\geq 0, l\geq 0} a^i_{k,l}\eta_i^k\xi_i^l\frac{\partial}{\partial\eta_i} + b^i_{k,l}\eta_i^k\xi_i^{l+1}\frac{\partial}{\partial\xi_i}. \tag{3.21}$$

For $\theta_{i+1} \in \Gamma(Y_{i+1}, \mathrm{Der}_E(X))$,

$$\theta_{i+1} = \sum_{k\geq 0, l\geq 0} a^{i+1}_{k,l}\xi_{i+1}^k\eta_{i+1}^l\frac{\partial}{\partial\xi_{i+1}} + b^{i+1}_{k,l}\xi_{i+1}^k\eta_{i+1}^{l+1}\frac{\partial}{\partial\eta_{i+1}}. \tag{3.22}$$

For $\theta_{i,i+1} \in \Gamma(Y_i\cap Y_{i+1}, \mathrm{Der}_E(X))$ on the intersection $Y_i \cap Y_{i+1}$ where $\eta_i \neq 0$,

$$\theta_{i,i+1} = \sum_{k\in\mathbb{Z}, l\geq 0} a^{i,i+1}_{k,l}\eta_i^k\xi_i^l\frac{\partial}{\partial\eta_i} + b^{i,i+1}_{k,l}\eta_i^k\xi_i^{l+1}\frac{\partial}{\partial\xi_i}. \tag{3.23}$$

By the transition formula (3.15),

$$\begin{aligned}\theta_{i+1} = \sum_{k\geq 0, l\geq 0}\Big\{ &- a^{i+1}_{k,l}\eta_i^{-k+le_{i+1}+2}\xi_i^l\frac{\partial}{\partial\eta_i}\\ &+ (a^{i+1}_{k,l}e_{i+1}\eta_i^{-k+le_{i+1}+1}\xi_i^{l+1} + b^{i+1}_{k,l}\eta_i^{-k+le_{i+1}}\xi_i^{l+1})\frac{\partial}{\partial\xi_i}\Big\},\end{aligned} \tag{3.24}$$

which shows that the exponents of η_i in θ_{i+1} can be any negative integers. Then it is clear that for any $\theta_{i,i+1}$, there exist θ_i, θ_{i+1} such that on $Y_i\cap Y_{i+1}$, $\theta_{i,i+1} = \theta_{i+1} - \theta_i$. Furthermore, if $\{\theta_{k,l}\}$ $(k < l)$ is closed, then $\theta_{k,l} = \theta_{k,l+1} + \cdots + \theta_{l-1,l}$. Then $\{\theta_{k,l}\}$ is determined if and only if the set of consecutive elements $\{\theta_{i,i+1}\}$ is determined. These arguments imply that any closed $\{\theta_{k,l}\}$ is exact, so $H^1(X, \mathrm{Der}_E(X)) = 0$. □

3.2. Non-cyclic quotient singularities

In Lemma 3.4, we have shown that when Γ is cyclic, $H^1(X, \mathrm{Der}_E(X)) = 0$, which implies that $H^1(X, \Theta) \simeq H^1(E, \mathcal{O}_E(E))$. In the following, we will use a relative index theorem to show this also holds for the general case.

Proposition 3.5. *Let $X \mapsto \mathbb{C}^2/\Gamma$ be a minimal resolution, where $\Gamma \subset \mathrm{U}(2)$ with no complex reflections. Then $H^1(X, \Theta) \simeq H^1(E, \mathcal{O}_E(E))$ and $H^1(X, \mathrm{Der}_E(X)) = 0$.*

Proof. Assume Γ is non-cyclic. We will first construct a Kähler form on the minimal resolution X of $\mathbb{C}^2/\Gamma$ by gluing Calderbank–Singer ALE surfaces Y_j $(j = 1, 2, 3)$ to a quotient of a LeBrun orbifold X_0, which has three cyclic quotient singularities on the central rational curve. The following is a sketch of the gluing procedure, details of a similar construction can be found in [LV14]. Note that in that paper, this gluing was used to proceed scalar-flat Kähler metrics, while in the following argument we will be considering a different operator P, see (3.31) below.

Let x_i, $i = 1, 2, 3$ be the cyclic quotient singularities of X_0 with group Γ_i. Let (z_i^1, z_i^2) be local holomorphic coordinates on $U_i \setminus \{x_i\}$. Let ω_{X_0} be the Kähler form of the LeBrun metric. Then ω_{X_0} admits an expansion

$$\omega_{X_0} = \frac{\sqrt{-1}}{2}(\partial\bar{\partial}|z_i|^2 + \partial\bar{\partial}\xi_i) \tag{3.25}$$

on $U_i \backslash \{x_i\}$, where ξ_i is a potential function satisfying $\xi_i = O(|z_i|^4)$. For the LeBrun orbifold X_0, outside of a compact subset, it admits a holomorphic coordinate (v_1, v_2). Let Y_i denote the minimal resolution of $\mathbb{C}^2/\Gamma_i$. Outside of a compact subset of Y_i, there exist holomorphic coordinates (u_i^1, u_i^2). Let ω_{Y_i} be a Kähler form on Y_i corresponding to any Calderbank–Singer metric on Y_i. From [RS09], the Kähler form admits an expansion

$$\omega_{Y_i} = \frac{\sqrt{-1}}{2}(\partial\bar{\partial}|u_i|^2 + \partial\bar{\partial}\eta_i), \tag{3.26}$$

where $\eta_i - c\log(|u|^2) = O(|u_i|^{-1})$, for some constant c. In fact, a similar expansion holds for any scalar-flat Kähler ALE surface on a resolution [ALM16].

Next, we construct a Kähler form on X. Choose two small positive numbers a, b, we glue the regions $\frac{1}{a} \le |u_i| \le \frac{4}{a}$ and $b \le |z_i| \le 4b$, by letting $z_i = ab \cdot u_i$. This mapping is biholomorphic in the intersection. Let ρ be a smooth cutoff function satisfying $\rho(t) = 1$ when $t \le 1$, $\rho = 0$ when $t \ge 2$. Let

$$\begin{aligned}
\omega_b &= \begin{cases} \frac{\sqrt{-1}}{2}(\partial\bar{\partial}|z_i|^2 + \partial\bar{\partial}((1-\rho(\frac{|z_i|}{2b}))\xi_i(z_i))) & \text{if } |z_i| \le b \\ \omega_{X_0} & \text{if } |z_i| \ge 4b \end{cases} \\
\omega_a &= \begin{cases} \frac{\sqrt{-1}}{2}(\partial\bar{\partial}|u_i|^2 + \partial\bar{\partial}(\rho(a|u_i|)\eta_i(u_i))) & \text{if } |u_i| \ge 4a^{-1} \\ \omega_{Y_i} & \text{if } |u_i| \le a^{-1} \end{cases}.
\end{aligned} \tag{3.27}$$

Then we define

$$\omega_{a,b} = \begin{cases} a^{-2}b^{-2}\omega_b & \text{if } |z_i| \ge 2b \\ \omega_a & \text{if } |u_i| \le 2a^{-1} \end{cases}. \tag{3.28}$$

For a, b sufficiently small, $\omega_{a,b}$ is a Kähler form on X. Since ω_{X_0} was ALE of order 2, the Kähler metric $\omega_{a,b}$ is also ALE of order 2.

Next, choose R_1, R_2, R_3, such that $0 < 2R_1 < R_2, R_3 > 0$, and define smooth functions r_1, r_2, r_3 as:

$$\begin{aligned}
r_1(x) &= \begin{cases} |z_i| & \text{if } |z_i| \le R_1 \\ 1 & \text{if } |z_i| \ge 2R_1 \end{cases} \\
r_2(x) &= \begin{cases} 1 & \text{if } |v| \le R_2 \\ |v| & \text{if } |v| \ge 2R_2 \end{cases} \\
r_3(x) &= \begin{cases} 1 & \text{if } |u_i| \le R_3 \\ |u| & \text{if } |u_i| \ge 2R_3 \end{cases}
\end{aligned} \tag{3.29}$$

For $\delta \in \mathbb{R}$, and weight function $\gamma > 0$, define the weighted Hölder space $C^{k,\alpha}_{\delta,\gamma}(M, T)$ of sections of any vector bundle T over M as the closure of the space of C^∞-sections in the norm

$$\begin{aligned}\|\sigma\|_{C^{k,\alpha}_{\delta,\gamma}(M,T)} &= \sum_{|\mathcal{I}|\le k} |\gamma^{-\delta+|\mathcal{I}|}\nabla^{\mathcal{I}}\sigma| \\ &+ \sum_{|\mathcal{I}|=k} \sup_{0<d(x,y)<\rho_{\mathrm{inj}}} \Big(\min\{\gamma(x),\gamma(y)\}^{-\delta+k+\alpha} \frac{|\nabla^{\mathcal{I}}\sigma(x) - \nabla^{\mathcal{I}}\sigma(y)|}{d(x,y)^{\alpha}} \Big).\end{aligned} \tag{3.30}$$

Lemma 3.6. *Let X_0 be the LeBrun orbifold with quotient singularities x_1, x_2, x_3. The elliptic operator*

$$P : C^{k,\alpha}_{\delta,r_1r_2}(X_0,\Lambda^{0,1}\otimes\Theta) \xrightarrow{(\bar\partial^*,\bar\partial)} C^{k-1,\alpha}_{\delta-1,r_1r_2}(X_0,\Theta) \oplus C^{k-1,\alpha}_{\delta-1,r_1r_2}(X_0,\Lambda^{0,2}\otimes\Theta) \tag{3.31}$$

is Fredholm and surjective, where $\delta \in (-2,-1), k \ge 3$.

Proof. First, note that any element of the kernel and cokernel is $\square$-harmonic. By the standard theory of harmonic functions, any $\square$-harmonic element which is $O(r_1^{\delta})$ as $r_1 \to 0$ has a removable singularity. The remainder of the proof is almost the same as the proof of [HV16, Lemma 4.2], and is omitted. $\square$

We will now consider the weight function $\gamma : X \to \mathbb{R}_+$ by the following:

$$\gamma = \begin{cases} a^{-1}b^{-1}r_1r_2 & \text{if } |z_i| \ge 2b \\ r_3 & \text{if } |u_i| \le 2a^{-1} \end{cases}. \tag{3.32}$$

Define the elliptic operator P as $(\bar\partial^*, \bar\partial)$ with respect to the glued metric $\omega_{a,b}$, on the weighted space $C^{k,\alpha}_{\delta,\gamma}(X,\Lambda^{0,1}\otimes\Theta)$. Because $\omega_{a,b}$ is a Kähler ALE metric, by [HV16, Lemma 4.2], P is Fredholm and surjective. Since each P_{Y_i} has a bounded right inverse for $i = 1,2,3$, and P_{X_0} has a bounded right inverse, a standard argument (see for example [RS05]) shows that there is a uniformly bounded right inverse of P, for a, b sufficiently small.

In [LV14, Prop. 6.1], it was shown that $\dim(\ker(P_{X_0})) = b_\Gamma - 1$, where $-b_\Gamma$ is the self-intersection number of the central divisor in X. For each Calderbank–Singer ALE space Y_j, by Lemma 3.4, we have $\dim(\ker(P_{Y_j})) = \sum_{i=1}^{k_j}(e_i^j - 1)$, where $-e_i^j$ is the self-intersection number of each irreducible exceptional divisor in Y_j.

In the proof of [HV16, Theorem 10.2], it was shown that the natural mapping

$$\mathcal{H}_{-3}(X,\Lambda^{0,1}\otimes\Theta) \to H^1(X,\Theta) \tag{3.33}$$

is surjective, which implies that $\dim(H^1(X,\Theta)) \le \dim(\mathcal{H}_{-3}(X,\Lambda^{0,1}\otimes\Theta))$. Also, from (3.2) above, we have $\dim(H^1(E,\mathcal{O}_E(E))) \le \dim(H^1(X,\Theta))$. Combining these, we have that

$$\begin{aligned}\dim(H^1(X,\Theta)) &\le \dim(\mathcal{H}_{-3}(X,\Lambda^{0,1}\otimes\Theta)) \\ &= \dim(\ker(P_X)) = b_\Gamma - 1 + \sum_{j=1}^{3}\sum_{i=1}^{k_j}(e_i^j - 1) \\ &= \dim(H^1(E,\mathcal{O}_E(E))) \le \dim(H^1(X,\Theta)).\end{aligned} \tag{3.34}$$

This implies the isomorphism $H^1(X, \Theta) \simeq H^1(E, \mathcal{O}_E(E))$. Then by the exact sequence (3.2), $H^1(X, \mathrm{Der}_E(X)) = 0$. $\square$

4. Dimension of the moduli space

We will next discuss the dimension of the generic orbit in the cases in Table 1.1.

4.1. Discussion of Table 1.1

Cases 1 and 2: $\Gamma = \frac{1}{p}(1,1)$. This case has been studied in [Hon13]. The group of biholomorphic automorphisms is $\mathrm{GL}(2, \mathbb{C})$, and the identity component of the holomorphic isometry group is $\mathrm{U}(2)$. When $p = 3$, the action of $\mathrm{U}(2)$ coincides with the action of $\mathrm{SU}(2)$, and the dimension of generic orbits is 3. Then $\dim(\mathfrak{M}) = 5 - 3 = 2$. When $p > 3$, the dimension of each orbit is 4, and $\dim(\mathfrak{M}) = 2p - 5$.

Case 3: $\Gamma = \frac{1}{p}(1,q)$, where $q \neq 1, p-1$. In this case, by direct calculation, the subgroup in $\mathrm{U}(2)$ that commutes with Γ is isomorphic to $S^1 \times S^1$. By [ALM16, Proposition 3.3] the identity component of the holomorphic isometry group must be $S^1 \times S^1$. Using the fact that the cyclic quotient singularity is characterized by the invariant polynomials in (3.7) above, it is easy to show that the dimension of the generic orbit of $\mathfrak{G}$ is 2, and therefore $\dim(\mathfrak{M}) = j_\Gamma + k_\Gamma - 2$.

Case 4: Γ is non-cyclic and not in $\mathrm{SU}(2)$. In this case, the subgroup in $\mathrm{U}(2)$ that commutes with Γ is isomorphic to S^1, so by [ALM16, Proposition 3.3] the identity component of the holomorphic isometry group must be S^1. Since the Hopf action is always nontrivial on the normal bundle of the central divisor E, the dimension of the generic orbit of the Hopf action on $H^1(E, \mathcal{O}_E(E)) \cong H^0(E, \mathcal{O}(b_\Gamma - 2))$ is 1, and therefore $\dim(\mathfrak{M}) = j_\Gamma + k_\Gamma - 1$.

4.2. Cyclic case

Any cyclic action without complex reflections is conjugate to the action generated by

$$(z_1, z_2) \mapsto (\xi_p z_1, \xi_p^q z_2), \tag{4.1}$$

where ξ_p is a pth root of unity, and q is relatively prime to p, which we will call a $\frac{1}{p}(1,q)$ action. Define the integers $e_i \geq 2$, and k by the continued fraction expansion

$$\frac{p}{q} = e_1 - \cfrac{1}{e_2 - \cdots \cfrac{1}{e_k}} \equiv [e_1, \ldots, e_k]. \tag{4.2}$$

The singularity of $\mathbb{C}^2/\Gamma$ is known as a Hirzebruch–Jung singularity, and the exceptional divisor is a string of rational curves with normal crossing singularities.

If $1 \leq q < p$, then let $q' = p - q$. Let $e'_i \geq 2$, and k' denote integers arising in the Hirzebruch–Jung algorithm for the $\frac{1}{p}(1,q')$-action. In [Rie74], Riemenschneider proved the formulas

$$\sum_{i=1}^{k}(e_i - 1) = \sum_{i=1}^{k'}(e'_i - 1), \tag{4.3}$$

$$k' = e - 2, \tag{4.4}$$

$$e = 3 + \sum_{i=1}^{k}(e_i - 2), \tag{4.5}$$

where e is the embedding dimension. In particular, these formulas give that

$$\sum_{i=1}^{k}(e_i - 1) = e + k - 3. \tag{4.6}$$

From Subsection 4.1 above, for $q \neq 1, p-1$, it follows that

$$m_\Gamma = 2e_\Gamma + 3k_\Gamma - 8. \tag{4.7}$$

4.3. Non-cyclic cases

The non-cyclic finite subgroups of U(2) without complex reflections are given in Table 4.1, where the binary polyhedral groups (dihedral, tetrahedral, octahedral, icosahedral) are respectively denoted by D^*_{4n}, T^*, O^*, I^*, and the map $\phi : \mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$ denotes the usual double cover, see [Bri68, BKR88, LV14] for more details.

Table 4.1. Non-cyclic finite subgroups of U(2) containing no complex reflections

$\Gamma \subset \mathrm{U}(2)$	Conditions	Order
$\phi(L(1,2l) \times D^*_{4n})$	$(l, 2n) = 1$	$4ln$
$\phi(L(1,2l) \times T^*)$	$(l, 6) = 1$	$24l$
$\phi(L(1,2l) \times O^*)$	$(l, 6) = 1$	$48l$
$\phi(L(1,2l) \times I^*)$	$(l, 30) = 1$	$120l$
Index-2 diagonal $\subset \phi(L(1,4l) \times D^*_{4n})$	$(l, 2) = 2, (l, n) = 1$	$4ln$
Index-3 diagonal $\subset \phi(L(1,6l) \times T^*)$	$(l, 6) = 3$	$24l$.

First, consider the case of $\phi(L(1,2l) \times D^*_{4n})$, where $(l, 2n) = 1$. In [BR77], it is shown that

$$e_\Gamma = \sum_{i=1}^{k}(e_i - 2) + 3, \tag{4.8}$$

which implies that

$$2e_\Gamma + 3k_\Gamma - 7 = 2\sum_{i=1}^{k}(e_i - 1) + k - 1, \tag{4.9}$$

which, from above, is equal to m_Γ. For the index 2 subgroup which is contained in $\phi(L(1,4l) \times D^*_{4n})$, where $(l,2) = 2, (l,n) = 1$, a similar computation shows that the same formula (4.9) holds in this case as well.

TABLE 4.2. Cases with T^*, O^*, I^* for $l > 1$

$\Gamma \subset \mathrm{U}(2)$	m_Γ
$\phi(L(1,2l) \times T^*)$	
$l \equiv 1 \mod 6$	$\frac{1}{3}(l-1) + 17$
$l \equiv 5 \mod 6$	$\frac{1}{3}(l-5) + 15$
Index–3 diagonal $\subset \phi(L(1,6l) \times T^*)$	
$(l,6) = 3$	$\frac{1}{3}(l-3) + 16$
$\phi(L(1,2l) \times O^*)$	
$l \equiv 1 \mod 12$	$\frac{1}{6}(l-1) + 20$
$l \equiv 5 \mod 12$	$\frac{1}{6}(l-5) + 19$
$l \equiv 7 \mod 12$	$\frac{1}{6}(l-7) + 18$
$l \equiv 11 \mod 12$	$\frac{1}{6}(l-11) + 17$
$\phi(L(1,2l) \times I^*)$	
$l \equiv 1 \mod 30$	$\frac{1}{15}(l-1) + 23$
$l \equiv 7 \mod 30$	$\frac{1}{15}(l-7) + 19$
$l \equiv 11 \mod 30$	$\frac{1}{15}(l-11) + 22$
$l \equiv 13 \mod 30$	$\frac{1}{15}(l-13) + 19$
$l \equiv 17 \mod 30$	$\frac{1}{15}(l-17) + 18$
$l \equiv 19 \mod 30$	$\frac{1}{15}(l-19) + 20$
$l \equiv 23 \mod 30$	$\frac{1}{15}(l-23) + 18$
$l \equiv 29 \mod 30$	$\frac{1}{15}(l-29) + 19$

Table 4.2 lists the dimension of the moduli space for subgroups of U(2) for finite subgroups involving T^*, O^*, I^*. The papers of [Bri68, BKR88, LV14] give a complete description of the exceptional divisors and self-intersection numbers, and it is a straightforward computation to obtain the right column in the table, which, from Subsection 4.1 is equal $j_\Gamma + k_\Gamma - 1$. Furthermore, using the formulas for the embedding dimension given in [BKR88], it can be easily checked that

$$m_\Gamma = j_\Gamma + k_\Gamma - 1 = 2e_\Gamma + 3k_\Gamma - 7 \tag{4.10}$$

in all of these cases (the computations are omitted). We point out that there is a typo in case I_7 in [BKR88]: in our notation, this case should have $l = 30(b-2)+7$ and $e = b + 2$.

4.4. Hyperkähler case

Recall that a hyperkähler metric is Kähler with respect to a 2-sphere of complex structures $S^2 = \{aI + bJ + cK : a^2 + b^2 + c^2 = 1\}$.

A_1-type: the Eguchi–Hanson metric is an ALE Ricci-flat Kähler metric on $X = T^*S^2$. Since $k = 1$, $B^1 \subset \mathbb{R}^2$, $B^2 \subset \mathbb{R}$, $d_\Gamma = 2+1 = 3$. With respect to the complex structure I (the complex structure arising as the total space of a holomorphic line bundle), the biholomorphic isometry group is $\mathrm{U}(2)$. The quotient $\mathfrak{F}/\mathfrak{G}$ has two orbit-types. The orbit of $(0, \rho)$ is one-dimensional. The orbit of (t, ρ) where t is non-zero, is also one-dimensional. Consequently, $\mathfrak{M}$ is isomorphic to the two-dimensional upper half-space. The remaining parameter of complex structures just corresponds to a hyperkähler rotation, so the metrics obtained are all just scalings of the Eguchi–Hanson metric.

Instead, consider the complex structure J. The biholomorphic isometry group is $\mathrm{SU}(2)$. The subspace $\mathbb{V}$ is now of dimension 1, so our parameter space is now $\mathbb{R} \times \mathbb{R}$. The group $\mathfrak{G}$ now acts trivially, so our parameter space is a ball in $\mathbb{R}^2$. The remaining parameter of complex structures again just corresponds to a hyperkähler rotation, so the metrics obtained are all just scalings of the Eguchi–Hanson metric.

ADE-type: For the general A_k, D_k, E_k $(k \geq 2)$ type ALE minimal resolution, the dimension of local moduli space of Ricci-flat Kähler metrics is $3k - 3$.

For A_k $(k \geq 2)$, this is the case of Gibbons–Hawking ALE hyperkähler surface. $Aut(X) = \mathbb{C}^* \times S^1 \subset \mathbb{R}_+ \times \mathrm{U}(2)$. Recall that $\mathrm{U}(2) = \mathrm{U}(1) \times_{\mathbb{Z}_2} \mathrm{SU}(2)$, which acts on $\vec{v} \in \mathbb{C}^2$ as $g_L \cdot \vec{v} \cdot g_R$, where $g_L \in \mathrm{U}(1)$ is the left action and $g_R \in \mathrm{SU}(2)$ is the right action. The $\mathbb{C}^*$-action is generated by $(v_1, v_2) \to (\lambda v_1, \lambda v_2)$ where $\lambda \in \mathbb{C}^*$; the S^1-action is generated by $(v_1, v_2) \to (\lambda v_1, \lambda^{-1} v_2))$ where $|\lambda| = 1$. The $\mathbb{C}^*$ action induces a two-dimensional action on the hyperkähler sphere, while the S^1 action preserves the hyperkähler structure. Then $m_\Gamma = d_\Gamma - 3 = 3k - 3$.

For the case of D_k, E_k, $Aut(X) = \mathbb{C}^*$. The $\mathbb{C}^*$ action can be interpreted as follows: let $\mathfrak{g}_{\mathbb{C}^*}$ denote the set of real vector fields which correspond to the Lie algebra of $\mathbb{C}^*$. For any $Y \in \mathfrak{g}_{\mathbb{C}^*}$, Φ_Y^* acts on the complex structures which gives an action on B^1. Since Y is a real vector field, Φ_Y^* is transverse to the action on the hyperkähler sphere (it is not transverse only in the A_k case). Then the dimension of the maximal orbit generated by the $\mathbb{C}^*$ action and the action on hyperkähler sphere is 3, so $m_\Gamma = d_\Gamma - 3$.

5. Appendix

In this appendix, we provide some details of a decay estimate needed in the proof of [HV16, Proposition 3.3]. Let $h = h_1 d\bar{z}_1 + h_2 d\bar{z}_2$ be a $\bar{\partial}$-closed $(0,1)$-form on $\mathbb{C}^2$ with $h \in C^1(\mathbb{C}^2)$, and

$$|h| < C(1 + |z|)^{-\mu} \tag{5.1}$$

$$|\nabla h| < C(1 + |z|)^{-\mu-1}. \tag{5.2}$$

for some $\mu > 1$. Define

$$q = \frac{1}{2\pi\sqrt{-1}} \int_{\mathbb{C}} h_2(z_1, \zeta_2) \frac{d\zeta_2 \wedge d\bar{\zeta}_2}{\zeta_2 - z_2}. \tag{5.3}$$

Proposition 5.1. *The function $q \in C^1_{\mathrm{loc}}(\mathbb{C}^2)$, satisfies*

$$\bar{\partial} q = h, \tag{5.4}$$

and obeys the estimate

$$|q|(z) < C \cdot (1 + |z|)^{-\mu+1}. \tag{5.5}$$

Proof. Since $\mu > 1$, the integrand in (5.3) is in $L^1(\mathbb{C})$ (in the ζ_2 variable), so q is well defined. A standard argument shows that $q \in C^1_{\mathrm{loc}}(\mathbb{C}^2)$. We claim that

$$\frac{\partial}{\partial \bar{z}_2} q = h_2. \tag{5.6}$$

To see this, make the change of variables $w_2 = \zeta_2 - z_2$, and write

$$q = \frac{1}{2\pi\sqrt{-1}} \int_{\mathbb{C}} \frac{h_2(z_1, w_2 + z_2)}{w_2} dw_2 \wedge d\bar{w}_2. \tag{5.7}$$

Notice that

$$\frac{\partial}{\partial \bar{z}_2} \frac{h_2(z_1, w_2 + z_2)}{w_2} = \frac{1}{w_2} \frac{\partial}{\partial \bar{z}_2} h_2(z_1, w_2 + z_2), \tag{5.8}$$

and using (5.2), this is uniformly in L^1 in the w_2 variable, so we can differentiate under the integral sign to obtain

$$\begin{aligned} \frac{\partial}{\partial \bar{z}_2} q &= \frac{1}{2\pi\sqrt{-1}} \int_{\mathbb{C}} \frac{1}{w_2} \frac{\partial}{\partial \bar{z}_2} h_2(z_1, w_2 + z_2) dw_2 \wedge d\bar{w}_2 \\ &= \frac{1}{2\pi\sqrt{-1}} \int_{\mathbb{C}} \frac{1}{w_2} \frac{\partial}{\partial \bar{w}_2} h_2(z_1, w_2 + z_2) dw_2 \wedge d\bar{w}_2 \\ &= \frac{1}{2\pi\sqrt{-1}} \int_{\mathbb{C}} \frac{1}{\zeta_2 - z_2} \frac{\partial}{\partial \bar{\zeta}_2} h_2(z_1, \zeta_2) d\zeta_2 \wedge d\bar{\zeta}_2 = h_2, \end{aligned} \tag{5.9}$$

by Cauchy's integral formula.

Next, by a similar argument, we can differentiate under the integral sign in the z_1-variable to obtain

$$\begin{aligned} \frac{\partial}{\partial \bar{z}_1} q &= \frac{1}{2\pi\sqrt{-1}} \int_{\mathbb{C}} \frac{\partial}{\partial \bar{z}_1} h_2(z_1, \zeta_2) \frac{d\zeta_2 \wedge d\bar{\zeta}_2}{\zeta_2 - z_2} \\ &= \frac{1}{2\pi\sqrt{-1}} \int_{\mathbb{C}} \frac{\partial}{\partial \bar{z}_2} h_1(z_1, \zeta_2) \frac{d\zeta_2 \wedge d\bar{\zeta}_2}{\zeta_2 - z_2} = h_1, \end{aligned} \tag{5.10}$$

with the middle equality holding since $\bar{\partial} h = 0$ by assumption.

We next claim that estimate (5.5) is satisfied for any $z \in E_1$, where E_1 is defined by

$$E_1 = \left\{ (z_1, z_2) \,|\, |z_1| > \frac{1}{4} |z_2| \right\}. \tag{5.11}$$

To see this, observe that if $z \in E_1$, then

$$|z_1| > \frac{1}{\sqrt{17}} |z|. \tag{5.12}$$

So if $|\zeta_2 - z_2| < 10|z_2|$, then

$$\begin{aligned}
&(1+|z|)^{\mu-1}\left|\frac{1}{2\pi\sqrt{-1}}\int_{B(z_2,10|z_2|)} h_2(z_1,\zeta_2)\frac{d\zeta_2\wedge d\bar{\zeta_2}}{\zeta_2 - z_2}\right| \\
&\quad\le C(1+|z|)^{\mu-1}\int_{B(z_2,10|z_2|)}(1+|z_1|+|\zeta_2|)^{-\mu}\frac{1}{|\zeta_2 - z_2|}|d\zeta_2\wedge d\bar{\zeta_2}| \\
&\quad\le C(1+|z|)^{\mu-1}(1+|z_1|)^{-\mu}\int_{B(z_2,10|z_2|)}\frac{1}{|\zeta_2 - z_2|}|d\zeta_2\wedge d\bar{\zeta_2}| \\
&\quad\le C(1+|z|)^{\mu-1}\left(1+\frac{|z|}{\sqrt{17}}\right)^{-\mu}10|z_2| \le C\frac{|z_2|}{1+|z|}\le C.
\end{aligned}\tag{5.13}$$

Next, if $|\zeta_2 - z_2| \ge 10|z_2|$, first note that

$$B(z_2,10|z_2|)^c \subset B(0,2|z_2|)^c. \tag{5.14}$$

So if $\zeta_2 \in B(z_2,10|z_2|)^c$ then $|\zeta_2 - z_2| > |\zeta_2| - |z_2| > 0$, and we estimate

$$\begin{aligned}
I &\equiv (1+|z|)^{\mu-1}\left|\frac{1}{2\pi\sqrt{-1}}\int_{B(z_2,10|z_2|)^c} h_2(z_1,\zeta_2)\frac{d\zeta_2\wedge d\bar{\zeta_2}}{\zeta_2 - z_2}\right| \\
&\le C(1+|z|)^{\mu-1}\int_{B(0,2|z_2|)^c}(1+|z_1|+|\zeta_2|)^{-\mu}\frac{1}{|\zeta_2 - z_2|}|d\zeta_2\wedge d\bar{\zeta_2}| \\
&\le C(1+|z|)^{\mu-1}\int_{B(0,2|z_2|)^c}(1+|z_1|+|\zeta_2|)^{-\mu}\frac{1}{|\zeta_2| - |z_2|}|d\zeta_2\wedge d\bar{\zeta_2}| \\
&\le C(1+|z|)^{\mu-1}\int_{2|z_2|}^{\infty}(1+|z_1|+r)^{-\mu}\frac{r}{r-|z_2|}dr.
\end{aligned}\tag{5.15}$$

Note that the function $\frac{r}{r-|z_2|}$ is decreasing for $r > |z_2|$ and its maximum value on the domain of integration is at the lower endpoint, which is equal to 2, so we have

$$\begin{aligned}
I &\le C(1+|z|)^{\mu-1}\int_{2|z_2|}^{\infty}(1+|z_1|+r)^{-\mu}dr \\
&= C(1+|z|)^{\mu-1}\frac{1}{1-\mu}(1+|z_1|+r)^{1-\mu}\Big|_{2|z_2|}^{\infty} \\
&\le C(1+|z|)^{\mu-1}(1+|z_1|+2|z_2|)^{1-\mu}\le C,
\end{aligned}\tag{5.16}$$

and therefore estimate (5.5) is satisfied for any $z \in E_1$.

Next, define $\tilde{q}$ by

$$\tilde{q} = \frac{1}{2\pi\sqrt{-1}}\int_{\mathbb{C}} h_1(\zeta_1,z_2)\frac{d\zeta_1\wedge d\bar{\zeta_1}}{\zeta_1 - z_1}. \tag{5.17}$$

An identical argument to the above shows that $\tilde{q} \in C^1_{\mathrm{loc}}(\mathbb{C}^2)$, satisfies

$$\bar{\partial}\tilde{q} = h \tag{5.18}$$

on $\mathbb{C}^2$, and

$$|\tilde{q}|(z) < C\cdot(1+|z|)^{-\mu+1} \tag{5.19}$$

for any $z \in E_2$, where E_2 is defined by

$$E_2 = \left\{ (z_1, z_2) \,|\, |z_2| > \tfrac{1}{4}|z_1| \right\}. \tag{5.20}$$

Since $\bar{\partial}(q - \tilde{q}) = 0$, $q - \tilde{q}$ is a holomorphic function on $\mathbb{C}^2$. Note that

$$E_1 \cap E_2 = \left\{ (z_1, z_2) \mid \tfrac{1}{4}|z_2| < |z_1| < 4|z_2| \right\} \tag{5.21}$$

together with the origin contains any complex line of the form $z_1 = cz_2$ with $1/4 < |c| < 4$. The function $q - \tilde{q}$ restricted to any such line is a decaying holomorphic function on $\mathbb{C}$ and must therefore vanish by Liouville's Theorem. Consequently, $q = \tilde{q}$ on $E_1 \cap E_2$. By unique continuation, it follows that $q = \tilde{q}$ on $\mathbb{C}^2$. Since $\mathbb{C}^2 = E_1 \cup E_2$, by the decay estimates for q and $\tilde{q}$ above, we conclude that (5.5) is satisfied on $\mathbb{C}^2$. □

Acknowledgement

This article is dedicated to Gang Tian on the occasion of his 60th birthday. The second author is extremely grateful to Tian for his collaboration, friendship, and generosity throughout the years since we first met over 20 years ago.

References

[ALM15] Claudio Arezzo, Alberto Della Vedova, Riccardo Lena, and Lorenzo Mazzieri, *On the Kummer construction for Kcsc metrics.* Boll. Unione Mat. Ital. 12 (2019), no. 1-2, 83–132.

[ALM16] ______, *On the resolution of extremal and constant scalar curvature Kähler orbifolds*, Int. Math. Res. Not. IMRN (2016), no. 21, 6415–6452.

[And89] Michael T. Anderson, *Ricci curvature bounds and Einstein metrics on compact manifolds*, J. Amer. Math. Soc. **2** (1989), no. 3, 455–490.

[AP06] Claudio Arezzo and Frank Pacard, *Blowing up and desingularizing constant scalar curvature Kähler manifolds*, Acta Math. **196** (2006), no. 2, 179–228.

[APS11] Claudio Arezzo, Frank Pacard, and Michael Singer, *Extremal metrics on blowups*, Duke Math. J. **157** (2011), no. 1, 1–51.

[Biq06] Olivier Biquard, *Asymptotically symmetric Einstein metrics*, SMF/AMS Texts and Monographs, vol. 13, American Mathematical Society, Providence, RI; Société Mathématique de France, Paris, 2006, Translated from the 2000 French original by Stephen S. Wilson.

[BKN89] Shigetoshi Bando, Atsushi Kasue, and Hiraku Nakajima, *On a construction of coordinates at infinity on manifolds with fast curvature decay and maximal volume growth*, Invent. Math. **97** (1989), no. 2, 313–349.

[BKR88] Kurt Behnke, Constantin Kahn, and Oswald Riemenschneider, *Infinitesimal deformations of quotient surface singularities*, Singularities (Warsaw, 1985), Banach Center Publ., vol. 20, PWN, Warsaw, 1988, pp. 31–66.

[BR77] Kurt Behnke and Oswald Riemenschneider, *Infinitesimale Deformationen von Diedersingularitäten*, Manuscripta Math. **20** (1977), no. 4, 377–400.

[BR15] Olivier Biquard and Yann Rollin, *Smoothing singular extremal Kähler surfaces and minimal Lagrangians*, Advances in Math. **285** (2015), 980–1024.

[Bri68] Egbert Brieskorn, *Rationale Singularitäten komplexer Flächen*, Invent. Math. **4** (1967/1968), 336–358.

[Cal85] Eugenio Calabi, *Extremal Kähler metrics. II*, Differential geometry and complex analysis, Springer, Berlin, 1985, pp. 95–114.

[CH15] Ronan J. Conlon and Hans-Joachim Hein, *Asymptotically conical Calabi–Yau metrics on quasi-projective varieties*, Geom. Funct. Anal. **25** (2015), no. 2, 517–552.

[CLW08] Xiuxiong Chen, Claude Lebrun, and Brian Weber, *On conformally Kähler, Einstein manifolds*, J. Amer. Math. Soc. **21** (2008), no. 4, 1137–1168.

[CS04] David M.J. Calderbank and Michael A. Singer, *Einstein metrics and complex singularities*, Invent. Math. **156** (2004), no. 2, 405–443.

[CW11] Xiuxiong Chen and Brian Weber, *Moduli spaces of critical Riemannian metrics with $L^{\frac{n}{2}}$ norm curvature bounds*, Adv. Math. **226** (2011), no. 2, 1307–1330.

[EH79] Tohru Eguchi and Andrew J. Hanson, *Self-dual solutions to Euclidean gravity*, Ann. Physics **120** (1979), no. 1, 82–106.

[GH78] G.W. Gibbons and S.W. Hawking, *Gravitational multi-instantons*, Physics Letters B **78** (1978), no. 4, 430–432.

[HL16] Hans-Joachim Hein and Claude LeBrun, *Mass in Kähler geometry*, Comm. Math. Phys. **347** (2016), no. 1, 183–221.

[Hon13] Nobuhiro Honda, *Deformation of LeBrun's ALE metrics with negative mass*, Comm. Math. Phys. **322** (2013), no. 1, 127–148.

[HRŞ16] Hans-Joachim Hein, Rares Rasdeaconu, and Ioana Şuvaina, *On the classification of ALE Kähler manifolds*, arXiv:1610.05239, 2016, to appear in International Mathematics Research Notices.

[HV16] Han, Jiyuan, and Jeff A. Viaclovsky, *Deformation theory of scalar-flat Kähler ALE surfaces.* American Journal of Mathematics 141, no. 6 (2019): 1547–1589. doi:10.1353/ajm.2019.0043.

[Kaw78] Yujiro Kawamata, *On deformations of compactifiable complex manifolds*, Math. Ann. **235** (1978), no. 3, 247–265.

[Kol07] János Kollár, *Lectures on resolution of singularities*, Annals of Mathematics Studies, vol. 166, Princeton University Press, Princeton, NJ, 2007.

[Kro86] Peter Benedict Kronheimer, *Instantons gravitationnels et singularités de Klein*, C. R. Acad. Sci. Paris Sér. I Math. **303** (1986), no. 2, 53–55.

[Kro89a] P.B. Kronheimer, *The construction of ALE spaces as hyper-Kähler quotients*, J. Differential Geom. **29** (1989), no. 3, 665–683.

[Kro89b] ______, *A Torelli-type theorem for gravitational instantons*, J. Differential Geom. **29** (1989), no. 3, 685–697.

[KS60] K. Kodaira and D.C. Spencer, *On deformations of complex analytic structures. III. Stability theorems for complex structures*, Ann. of Math. (2) **71** (1960), 43–76.

[Kur65] M. Kuranishi, *New proof for the existence of locally complete families of complex structures*, Proc. Conf. Complex Analysis (Minneapolis, 1964), Springer, Berlin, 1965, pp. 142–154.

[Lau73] Henry B. Laufer, *Taut two-dimensional singularities*, Math. Ann. **205** (1973), 131–164.

[LeB88] Claude LeBrun, *Counter-examples to the generalized positive action conjecture*, Comm. Math. Phys. **118** (1988), no. 4, 591–596.

[LM08] Claude LeBrun and Bernard Maskit, *On optimal 4-dimensional metrics*, J. Geom. Anal. **18** (2008), no. 2, 537–564.

[LS93] Claude LeBrun and Michael Singer, *Existence and deformation theory for scalar-flat Kähler metrics on compact complex surfaces*, Invent. Math. **112** (1993), no. 2, 273–313.

[LS94] C. LeBrun and S.R. Simanca, *Extremal Kähler metrics and complex deformation theory*, Geom. Funct. Anal. **4** (1994), no. 3, 298–336.

[LV14] Michael T. Lock, and Jeff A. Viaclovsky, *A smörgåsbord of scalar-flat Kähler ALE surfaces.* J. Reine Angew. Math. 746 (2019), 171–208.

[Nak94] Hiraku Nakajima, *A convergence theorem for Einstein metrics and the ALE spaces*, Selected papers on number theory, algebraic geometry, and differential geometry, Amer. Math. Soc. Transl. Ser. 2, vol. 160, Amer. Math. Soc., Providence, RI, 1994, pp. 79–94.

[Rei] M. Reid, *Surface cyclic quotient singularities and Hirzebruch–Jung resolutions*, lecture notes.

[Rie74] Oswald Riemenschneider, *Deformationen von Quotientensingularitäten* (*nach zyklischen Gruppen*), Math. Ann. **209** (1974), 211–248.

[RS05] Yann Rollin and Michael Singer, *Non-minimal scalar-flat Kähler surfaces and parabolic stability*, Invent. Math. **162** (2005), no. 2, 235–270.

[RS09] ———, *Constant scalar curvature Kähler surfaces and parabolic polystability*, J. Geom. Anal. **19** (2009), no. 1, 107–136.

[Siu69] Yum-tong Siu, *Analytic sheaf cohomology groups of dimension n of n-dimensional noncompact complex manifolds*, Pacific J. Math. **28** (1969), 407–411.

[Şuv12] Ioana Şuvaina, *ALE Ricci-flat Kähler metrics and deformations of quotient surface singularities*, Ann. Global Anal. Geom. **41** (2012), no. 1, 109–123.

[Szé12] Gábor Székelyhidi, *On blowing up extremal Kähler manifolds*, Duke Math. J. **161** (2012), no. 8, 1411–1453.

[Tia90] G. Tian, *On Calabi's conjecture for complex surfaces with positive first Chern class*, Invent. Math. **101** (1990), no. 1, 101–172.

[TV05a] Gang Tian and Jeff Viaclovsky, *Bach-flat asymptotically locally Euclidean metrics*, Invent. Math. **160** (2005), no. 2, 357–415.

[TV05b] ———, *Moduli spaces of critical Riemannian metrics in dimension four*, Adv. Math. **196** (2005), no. 2, 346–372.

[TV08] ———, *Volume growth, curvature decay, and critical metrics*, Comment. Math. Helv. **83** (2008), no. 4, 889–911.

[Wah75] Jonathan M. Wahl, *Vanishing theorems for resolutions of surface singularities*, Invent. Math. **31** (1975), no. 1, 17–41.

Jiyuan Han
Department of Mathematics
Purdue University
West Lafayette, IN 47907, USA
e-mail: `han556@purdue.edu`

Jeff A. Viaclovsky
Department of Mathematics
University of California
Irvine, CA 92697, USA
e-mail: `jviaclov@uci.edu`

Progress in Mathematics, Vol. 333, 137–155

Singular Ricci Flows II

Bruce Kleiner and John Lott

Abstract. We establish several quantitative results about singular Ricci flows, including estimates on the curvature and volume, and the set of singular times.

Mathematics Subject Classification (2010). 53C44.

Keywords. Ricci flow, geometric flow, weak solution, singular.

1. Introduction

In [KL17, BKb], it was shown that there exists a canonical Ricci flow through singularities starting from an arbitrary compact Riemannian 3-manifold, and that this flow may obtained as a limit of a sequence of Ricci flows with surgery. These results confirmed a conjecture of Perelman [Per02, Per03], and were used in the proof of the Generalized Smale Conjecture in [BKa].

The purpose of this paper, which is a sequel to [KL17], is to further study Ricci flow through singularities.

We recall that the basic object introduced in [KL17] is a *singular Ricci flow*, which is a Ricci flow spacetime subject to several additional conditions; see Definition 2.2 of Section 2 or [KL17, Def. 1.6].

In the following, we let $\mathcal{M}$ be a singular Ricci flow with parameter functions κ and r, and we let $\mathcal{M}_t$ denote a time slice. The main results of the paper are the following.

Theorem 1.1. *For all $p \in (0,1)$ and all t, the scalar curvature is L^p on $\mathcal{M}_t$.*

Theorem 1.2. *The volume function $\mathcal{V}(t) = \mathrm{vol}(\mathcal{M}_t)$ has a locally bounded upper-right derivative and is locally α-Hölder in t for some exponent $\alpha \in (0,1)$.*

The first assertion of the theorem was shown in [KL17, Proposition 5.5 and Corollary 7.7], so the issue here is to prove Hölder continuity in the opposite direction.

We state the next result loosely, with a more precise formulation given later.

Theorem 1.3. *The a priori assumptions in Definition* 2.2 *of a singular Ricci flow are really conditions on the spacetime near infinity, in the sense that if the conditions hold outside of compact subsets then, after redefinition of κ and r, they hold everywhere.*

Finally, we estimate the size of the set of singular times in a singular Ricci flow.

Theorem 1.4. *For any $T < \infty$, the set of times $t \in [0, T]$ for which $\mathcal{M}_t$ is noncompact has Minkowski dimension at most $\frac{1}{2}$.*

The structure of the paper is as follows. In Section 2 we recall some notation and terminology from [KL17]. In Section 3 we prove some needed results about compact κ-solutions. Section 4 has the proofs of Theorems 1.1 and 1.2. In Section 5 we prove Theorem 1.3 and in Section 6 we prove Theorem 1.4.

2. Notation and terminology

We will assume some familiarity with [KL17], but in order to make this paper as self-contained as possible, we will give precise references for all results from [KL17] that are used here. We follow the notation and terminology of [KL17]. All manifolds that arise will be taken to be orientable. A *κ-solution* is a special type of ancient Ricci flow solution, for which we refer to [KL17, Appendix A.5]. The function $r : [0, \infty) \to (0, \infty)$ is the parameter r of the canonical neighborhood assumption [KL17, Appendix A.8].

Definition 2.1. A *Ricci flow spacetime* is a tuple $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ where:

- $\mathcal{M}$ is a smooth manifold-with-boundary.
- $\mathfrak{t}$ is the *time function* – a submersion $\mathfrak{t} : \mathcal{M} \to I$ where $I \subset \mathbb{R}$ is a time interval; we will usually take $I = [0, \infty)$.
- The boundary of $\mathcal{M}$, if it is nonempty, corresponds to the endpoint(s) of the time interval: $\partial\mathcal{M} = \mathfrak{t}^{-1}(\partial I)$.
- $\partial_{\mathfrak{t}}$ is the *time vector field*, which satisfies $\partial_{\mathfrak{t}}\mathfrak{t} \equiv 1$.
- g is a smooth inner product on the spatial subbundle $\ker(d\mathfrak{t}) \subset T\mathcal{M}$, and g defines a Ricci flow: $\mathcal{L}_{\partial_{\mathfrak{t}}} g = -2\,\mathrm{Ric}(g)$.

For $0 \leq a < b$, we write $\mathcal{M}_a = \mathfrak{t}^{-1}(a)$, $\mathcal{M}_{[a,b]} = \mathfrak{t}^{-1}([a, b])$ and $\mathcal{M}_{\leq a} = \mathfrak{t}^{-1}([0, a])$. Henceforth, unless otherwise specified, when we refer to geometric quantities such as curvature, we will implicitly be referring to the metric on the time slices.

Definition 2.2. A Ricci flow spacetime $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ is a *singular Ricci flow* if it is four-dimensional, the initial time slice $\mathcal{M}_0$ is a compact normalized Riemannian manifold and

a. The scalar curvature function $R : \mathcal{M}_{\leq T} \to \mathbb{R}$ is bounded below and proper for all $T \geq 0$.

b. $\mathcal{M}$ satisfies the Hamilton–Ivey pinching condition of [KL17, (A.14)].
c. For a global parameter $\epsilon > 0$ and decreasing functions $\kappa, r : [0, \infty) \to (0, \infty)$, the spacetime $\mathcal{M}$ is κ-noncollapsed below scale ϵ in the sense of [KL17, Appendix A.4] and satisfies the r-canonical neighborhood assumption in the sense of [KL17, Appendix A.8].

Here "normalized" means that at each point m in the initial time slice, the eigenvalues of the curvature operator $\mathrm{Rm}(m)$ are bounded by one in absolute value, and the volume of the unit ball $B(m, 1)$ is at least half the volume of the Euclidean unit ball. By rescaling, any compact Riemannian manifold can be normalized. "Proper" has the usual meaning, that the preimage of a compact set is compact. Since R is bounded below, its properness means that as one goes out an end of $\mathcal{M}_{\leq T}$, the function R goes to infinity.

Let $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ be a Ricci flow spacetime (Definition 2.1). For brevity, we will often write $\mathcal{M}$ for the quadruple.

Given $s > 0$, the rescaled Ricci flow spacetime is $\hat{\mathcal{M}}(s) = (\mathcal{M}, \frac{1}{s}\mathfrak{t}, s\partial_{\mathfrak{t}}, \frac{1}{s}g)$.

Definition 2.3. Let $\mathcal{M}$ be a Ricci flow spacetime. A path $\gamma : I \to \mathcal{M}$ is *time-preserving* if $\mathfrak{t}(\gamma(t)) = t$ for all $t \in I$. The *worldline* of a point $m \in \mathcal{M}$ is the maximal time-preserving integral curve $\gamma : I \to \mathcal{M}$ of the time vector field $\partial_{\mathfrak{t}}$, which passes through m.

If $\gamma : I \to \mathcal{M}$ is a worldline then we may have $\sup I < \infty$. In this case, the scalar curvature blows up along $\gamma(t)$ as $t \to \sup I$, and the worldline encounters a singularity. An example would be a shrinking round space form, or a neckpinch. A worldline may also encounter a singularity going backward in time.

Definition 2.4. A worldline $\gamma : I \to \mathcal{M}$ is *bad* if $\inf I > 0$, i.e., if it is not defined at $t = 0$.

Given $m \in M_t$, we write $B(m, r)$ for the open metric ball of radius r in $\mathcal{M}_t$. We write $P(m, r, \Delta t)$ for the parabolic neighborhood, i.e., the set of points m' in $\mathcal{M}_{[t,t+\Delta t]}$ if $\Delta t > 0$ (or $\mathcal{M}_{[t+\Delta t,t]}$ if $\Delta t < 0$) that lie on the worldline of some point in $B(m, r)$. We say that $P(m, r, \Delta t)$ is *unscathed* if $B(m, r)$ has compact closure in $\mathcal{M}_t$ and for every $m' \in P(m, r, \Delta t)$, the maximal worldline γ through m' is defined on a time interval containing $[t, t + \Delta t]$ (or $[t + \Delta t, t]$). We write $P_+(m, r)$ for the forward parabolic ball $P(m, r, r^2)$ and $P_-(m, r)$ for the backward parabolic ball $P(m, r, -r^2)$.

3. Compact κ-solutions

In this section we prove some structural results about compact κ-solutions. The main result of this section, Corollary 3.3, will be used in the proof of Proposition 4.2.

We recall from [KL17, Appendix A.5] that if $\mathcal{M}$ is a κ-solution then $\mathcal{M}_{t,\hat{\epsilon}}$ denotes the points in $\mathcal{M}_t$ that are not centers of $\hat{\epsilon}$-necks.

Lemma 3.1. *There is some $\overline{\epsilon} > 0$ so that for any $0 < \hat{\epsilon} < \overline{\epsilon}$, there are $\epsilon' = \epsilon'(\hat{\epsilon}) > 0$ and $\alpha = \alpha(\hat{\epsilon}) < \infty$ with the following property. Let $\mathcal{M}$ be a compact κ-solution. Suppose that $\mathcal{M}_t$ contains an ϵ'-neck. Then there are points $m_1, m_2 \in \mathcal{M}_t$ so that $\mathcal{M}_{t,\hat{\epsilon}}$ is covered by disjoint balls $B(m_1, \alpha R(m_1,t)^{-\frac{1}{2}})$ and $B(m_2, \alpha R(m_2,t)^{-\frac{1}{2}})$, whose intersections with $\mathcal{M}_t - \mathcal{M}_{t,\hat{\epsilon}}$ are nonempty.*

Proof. Let $\alpha = \alpha(\hat{\epsilon})$ be the parameter of [KL08, Corollary 48.1]. Suppose that there is some point $x \in \mathcal{M}_{t,\hat{\epsilon}}$ so that $R(x)\,\mathrm{Diam}^2(\mathcal{M}_t) < \alpha$. By the compactness of the space of pointed κ-solutions, it follows that there is an upper bound on $(\sup_{\mathcal{M}_t} R) \cdot \mathrm{Diam}^2(\mathcal{M}_t)$, depending only on α. If ϵ' is sufficiently small then we obtain a contradiction. Hence we are in case C of [KL08, Corollary 48.1], so there are points $m_1, m_2 \in \mathcal{M}_t$ such that $M_{t,\hat{\epsilon}} \subset B(m_1, \alpha R(m_1,t)^{-\frac{1}{2}}) \cup B(m_2, \alpha R(m_2,t)^{-\frac{1}{2}})$. If ϵ' is sufficiently small then a cross-section of the ϵ'-neck separates $\mathcal{M}_t$ into two connected components, each of which must have a cap region. Hence if ϵ' is sufficiently small then $B(m_1, \alpha R(m_1,t)^{-\frac{1}{2}})$ and $B(m_2, \alpha R(m_2,t)^{-\frac{1}{2}})$ are disjoint. As $\mathcal{M}_{t,\hat{\epsilon}}$ is closed, both $B(m_1, \alpha R(m_1,t)^{-\frac{1}{2}})$ and $B(m_2, \alpha R(m_2,t)^{-\frac{1}{2}})$ intersect $\mathcal{M}_t - \mathcal{M}_{t,\hat{\epsilon}}$. □

Lemma 3.2. *Given $\hat{\epsilon} > 0$ and a compact family $\mathcal{F}$ of pointed κ-solutions, with basepoints at time zero, there is some $T = T(\hat{\epsilon}, \mathcal{F}) < 0$ such that for each $\mathcal{M} \in \mathcal{F}$, there is a point $(m,t) \in \mathcal{M}$ with $t \in [-T, 0]$ so that $\left(\hat{\mathcal{M}}(-t), m\right)$ is $\hat{\epsilon}$-close to a pointed gradient shrinking soliton which is a κ-solution.*

Proof. Suppose that the lemma fails. Then for each $j \in \mathbb{Z}^+$, there is some $\mathcal{M}^j \in \mathcal{F}$ so that for each $(m,t) \in \mathcal{M}^j$ with $t \in [-j, 0]$, there is no pointed gradient shrinker (which is a κ-solution) that is $\hat{\epsilon}$-close to $\left(\hat{\mathcal{M}}^j(-t), m\right)$. After passing to a subsequence, we can assume that $\lim_{j \to \infty} \mathcal{M}^j = \mathcal{M}^\infty \in \mathcal{F}$. From the existence of an asymptotic soliton for $\mathcal{M}^\infty$, there is some $(m_\infty, t_\infty) \in \mathcal{M}^\infty$ so that $\left(\hat{\mathcal{M}}^\infty(-t_\infty), m_\infty\right)$ is $\frac{\hat{\epsilon}}{2}$-close to a gradient shrinking soliton (which is a κ-solution). Then for large j, there is some $(m_j, t_\infty) \in \mathcal{M}^j$ so that $\left(\hat{\mathcal{M}}^j(-t_\infty), m_j\right)$ is $\hat{\epsilon}$-close to the gradient shrinking soliton. This is a contradiction. □

Corollary 3.3. *Let $\mathcal{F}$ be a compact family of compact κ-solutions that does not have any constant curvature elements. Then for each $\hat{\epsilon} > 0$, there is some $\mathcal{T} = \mathcal{T}(\hat{\epsilon}, \mathcal{F}) < 0$ such that for each $\mathcal{M} \in \mathcal{F}$, there is a point $(m,t) \in \mathcal{M}_{[\mathcal{T},0]}$ which is the center of an $\hat{\epsilon}$-neck.*

Proof. By assumption, there is some $\sigma = \sigma(\mathcal{F}) > 0$ so that no time-zero slice of an element of $\mathcal{F}$ is σ-close to a constant curvature manifold. By Lemma 3.2, for each $\epsilon' > 0$, there is some $\mathcal{T} = \mathcal{T}(\epsilon', \mathcal{F}) < 0$ such that for each $\mathcal{M} \in \mathcal{F}$, there is some $(m,t) \in \mathcal{M}_{[\mathcal{T},0]}$ so that $\left(\hat{\mathcal{M}}(-t), m\right)$ is ϵ'-close to a gradient shrinking soliton (which is a κ-solution). If ϵ' is sufficiently small, in terms of σ, then by the local stability of Ricci flows of constant positive curvature, this soliton cannot have constant curvature. Hence it is either a round shrinking cylinder or a $\mathbb{Z}_2$-quotient of a round shrinking cylinder. If it is a round shrinking cylinder then as long as $\epsilon' \leq \hat{\epsilon}$, we are done. If it is a $\mathbb{Z}_2$-quotient of a round shrinking cylinder then if ϵ' is sufficiently small, by moving the basepoint we can find a point $(m', t) \in \mathcal{M}$ that is the center of an ϵ-neck. $\square$

4. Curvature and volume estimates

In this section we establish curvature and volume estimates for singular Ricci flows. There are two main results. In Proposition 4.2, we show that $|R|^p$ is integrable on each time slice, for each $p \in (0, 1)$. In Proposition 4.36 we give an estimate on how the volume $\mathcal{V}(t)$ can change as a function of t. When combined with part (5) of [KL17, Proposition 5.5], it shows that $\mathcal{V}(t)$ is $\frac{1}{\eta}$-Hölder in t, where $\eta \geq 1$ is the constant in the estimate

$$|\nabla R(x,t)| < \eta R(x,t)^{\frac{3}{2}}, \qquad \left| \frac{\partial R}{\partial t}(x,t) \right| < \eta R(x,t)^2, \tag{4.1}$$

for canonical neighborhoods from [KL17, (A.8)].

Proposition 4.2. *Let $\mathcal{M}$ be a singular Ricci flow. Then for all $p \in (0, 1)$ and $T < \infty$, there is a bound*

$$\int_{\mathcal{M}_t} |R|^p \ \mathrm{dvol}_{g(t)} \leq \mathrm{const.}(p, T)\, \mathrm{vol}_{g(0)}(\mathcal{M}_0) \tag{4.3}$$

for all $t \in [0, T]$.

Proof. Before entering into the details, we first give a sketch of the proof.

Due to the bounds on $\mathcal{V}(t)$ from [KL17, Proposition 5.5], it suffices to control the contribution to the left-hand side of (4.3) from the points with large scalar curvature. Such points fall into three types, according to the geometry of the canonical neighborhoods: (a) neck points, (b) cap points, and (c) points p whose connected component in $\mathcal{M}_t$ is compact and has diameter comparable to $R(p)^{-\frac{1}{2}}$. If $(p, t) \in \mathcal{M}_t$ is a neck point with worldline $\gamma : [0, t] \to \mathcal{M}$ then thanks to the stability of necks going backward in time, the scale-invariant time derivative $R^{-1}\frac{\partial R}{\partial t}$ will remain very close to the cylindrical value along γ, until R falls down to a value comparable to $(r(t))^{-2}$. Combining this with previous estimates on the Jacobian as in [KL17, Section 5], we can bound the contribution from the neck points to the left-hand side of (4.3) in terms of the volume of the corresponding set of points in the time zero slice. To control the contribution from points of type (b), we show that it is dominated by that of the neck points. To control the contribution from the points of type (c) we use a similar approach. We again analyze the geometry going backward in time along worldlines, except that in this case there are three stages: one where the components are nearly round, one when they are no longer nearly round but still have diameter comparable to $R^{-\frac{1}{2}}$, and one when they have a large necklike region.

We now start on the proof. With the notation of the proof of [KL17, Proposition 5.5], let $X_3 \subset \mathcal{M}_t$ be the complement of the set of points in $\mathcal{M}_t$ with a bad worldline. From [KL17, Theorem 7.1], it has full measure in $\mathcal{M}_t$. Given $x \in X_3$, let $\gamma_x : [0,t] \to \mathcal{M}_{[0,t]}$ be the restriction of its worldline to the interval $[0,t]$. Define $J_t(x)$ as in [KL17, (5.8)], with $t_1 = 0$. That is,

$$J_t(x) = \frac{i_t^* \operatorname{dvol}_{g(t)}}{\operatorname{dvol}_{g(0)}}(x) \tag{4.4}$$

is the pointwise volume distortion of the inclusion map i_t that goes from (a subset of) the time-zero slice to the time-t slice. From [KL17, (5.9)], we have

$$J_t(x) = e^{-\int_0^t R(\gamma_x(u))\, du}. \tag{4.5}$$

Given $T > 0$, we consider times t in the range $[0,T]$. Let $\hat{\epsilon}$, C_1, $\overline{R}$ and $\overline{R}'$ be as in [KL17, Proposition 5.16]. We take $\overline{R} > r(T)^{-2}$. From [KL17, Proposition 5.15] we can assume that the $\hat{\epsilon}$-canonical neighborhood assumption holds on the superlevel set $\mathcal{M}_{[0,T]}^{>\overline{R}}$ of the scalar curvature function. We will further adjust the parameters $\hat{\epsilon}$ and $\overline{R}'$ later.

For any $\widehat{R} \geq \overline{R}$, write

$$\mathcal{M}_t^{>\widehat{R}} = \mathcal{M}_{t,\mathrm{neck}}^{>\widehat{R}} \cup \mathcal{M}_{t,\mathrm{cap}}^{>\widehat{R}} \cup \mathcal{M}_{t,\mathrm{closed}}^{>\widehat{R}}, \tag{4.6}$$

where

- $\mathcal{M}_{t,\mathrm{neck}}^{>\widehat{R}}$ consists of the points in $\mathcal{M}_t^{>\widehat{R}}$ that are centers of $\widehat{\epsilon}$-necks,
- $\mathcal{M}_{t,\mathrm{cap}}^{>\widehat{R}}$ consists of the points $x \in \mathcal{M}_t^{>\widehat{R}} - \mathcal{M}_{t,\mathrm{neck}}^{>\widehat{R}}$ so that after rescaling by $R(x)$, the pair $(\mathcal{M}_t, x)$ is $\hat{\epsilon}$-close to a pointed noncompact κ-solution, and
- $\mathcal{M}_{t,\mathrm{closed}}^{>\widehat{R}} = \mathcal{M}_t^{>\widehat{R}} - \left(\mathcal{M}_{t,\mathrm{neck}}^{>\widehat{R}} \cup \mathcal{M}_{t,\mathrm{cap}}^{>\widehat{R}}\right)$.

Taking $\widehat{R} = \overline{R}'$, there is a compact set $\mathcal{C}$ of κ-solutions so that for $x \in \mathcal{M}_{t,\mathrm{closed}}^{>\overline{R}'}$, after rescaling by $R(x)$ the connected component of $\mathcal{M}_t$ containing x is $\hat{\epsilon}$-close to an element of $\mathcal{C}$ (cf. Step 1 of the proof of [KL17, Theorem 7.1]). In particular, before rescaling, the diameter of the component is bounded above by $CR(x)^{-\frac{1}{2}}$ and the scalar curvature on the component satisfies

$$C^{-1} R_{\mathrm{av}} \leq R \leq C R_{\mathrm{av}}, \tag{4.7}$$

for an appropriate constant $C = C(\hat{\epsilon}) < \infty$, where R_{av} denote the average scalar curvature on the component.

By the pointed compactness of the space of normalized κ-solutions, and the diameter bound on the caplike regions in normalized pointed noncompact κ-solutions, there is a $C' = C'(\hat{\epsilon}, \overline{R}') < \infty$ so that

$$\int_{\mathcal{M}_{t,\mathrm{cap}}^{>\overline{R}'}} |R|^p \operatorname{dvol}_{g(t)} \leq C' \int_{\mathcal{M}_{t,\mathrm{neck}}^{>\overline{R}'}} |R|^p \operatorname{dvol}_{g(t)}. \tag{4.8}$$

Hence we can restrict our attention to $\mathcal{M}_{t,\mathrm{neck}}^{>\overline{R}'}$ and $\mathcal{M}_{t,\mathrm{closed}}^{>\overline{R}'}$.

Consider $x \in \mathcal{M}^{>\overline{R}'}_{t,\mathrm{neck}} \cap X_3$. With δ_{neck}, δ_0 and δ_1 being parameters of [KL17, Theorem 6.1], we assume that $\hat{\epsilon} < \frac{\delta_0}{100}$ and $\delta_1 < \frac{\delta_0}{100}$. Using [KL17, Theorem 6.1] and the $\hat{\epsilon}$-canonical neighborhood assumption, there are $T'' < T' < 0$ so that for $s \in [T'', T']$, the rescaled solution $\left(\hat{\mathcal{M}}(-sR(x)^{-1}), \gamma_x(t + sR(x)^{-1})\right)$ is $\frac{\delta_0}{10}$-close to $(Cyl, (y_0, -1))$. By reducing $\hat{\epsilon}$, we can make T'' arbitrarily negative.

The gradient bound (4.1) gives

$$\frac{1}{R(\gamma_x(u))} \leq \frac{1}{R(x)} + \eta(t-u). \tag{4.9}$$

as long as $\gamma_x(u)$ stays in a canonical neighborhood. If

$$\overline{R}' \geq (1 - \eta T'')\overline{R} \tag{4.10}$$

then for all $u \in [t + T''R(x)^{-1}, t]$, we have $R(\gamma_x(u)) \geq \overline{R}$ and (4.9) holds, so

$$\begin{aligned} \int_{t+T'R(x)^{-1}}^{t} R(\gamma_x(u))\, du \;&\geq\; \int_{t+T'R(x)^{-1}}^{t} \frac{1}{\frac{1}{R(x)} + \eta(t-u)}\, du \\ &=\; \frac{1}{\eta}\log(1 - \eta T'). \end{aligned} \tag{4.11}$$

For a round shrinking cylinder, the sharp value for η in (4.9) is 1. For any $q > 1$, if δ_0 is sufficiently small then we are ensured that

$$\begin{aligned} \int_{t+T''R(x)^{-1}}^{t+T'R(x)^{-1}} R(\gamma_x(u))\, du \;&\geq\; \int_{t+T''R(x)^{-1}}^{t+T'R(x)^{-1}} \frac{1}{\frac{1}{R(x)} + q(t-u)}\, du \\ &=\; \frac{1}{q}\log\frac{1 - qT''}{1 - qT'}. \end{aligned} \tag{4.12}$$

In all,

$$e^{-\int_{t+T''R(x)^{-1}}^{t} R(\gamma_x(u))\, du} \leq (1 - \eta T')^{-\frac{1}{\eta}}(1 - qT')^{\frac{1}{q}}(1 - qT'')^{-\frac{1}{q}}. \tag{4.13}$$

Because of the cylindrical approximation,

$$\frac{1}{2} \leq \frac{(1 - T'')R(\gamma_x(t + T''R(x)^{-1}))}{R(x)} \leq 2, \tag{4.14}$$

and so there is a constant $C'' = C''(q, \eta, T') < \infty$ such that for very negative T'', we have

$$e^{-\int_{t+T''R(x)^{-1}}^{t} R(\gamma_x(u))\, du} \leq C'' \left(\frac{R(x)}{R(\gamma(t + T''R(x)^{-1}))}\right)^{-\frac{1}{q}}. \tag{4.15}$$

We now replace t by $t + T''R(x)^{-1}$ and iterate the argument. Eventually, there will be a first time t_x when we can no longer continue the iteration because the curvature has gone below $\overline{R}$. Suppose that there are N such iterations. Then

$$e^{-\int_{t_x}^{t} R(\gamma_x(u))\, du} \leq (C'')^N \left(\frac{R(x)}{\overline{R}}\right)^{-\frac{1}{q}}. \tag{4.16}$$

From (4.14),

$$\left(\frac{1-T''}{2}\right)^{N-1} \leq \frac{R(x)}{\overline{R}}, \tag{4.17}$$

so

$$\begin{aligned} e^{-\int_{t_x}^{t} R(\gamma_x(u))\,du} &\leq C''\,(C'')^{N-1}\left(\frac{R(x)}{\overline{R}}\right)^{-\frac{1}{q}} \\ &\leq C''\left(\frac{R(x)}{\overline{R}}\right)^{\frac{\log C''}{\log\left(\frac{1-T''}{2}\right)}-\frac{1}{q}}. \end{aligned} \tag{4.18}$$

Put $p = \frac{1}{q} - \frac{\log C''}{\log\left(\frac{1-T''}{2}\right)}$. By choosing q sufficiently close to 1 (from above) and T'' sufficiently negative, we can make p arbitrarily close to 1 (from below). Using the lower scalar curvature bound [KL17, Lemma 5.2], we have

$$\int_0^{t_x} R(\gamma_x(u))\,du \geq -\int_0^T \frac{3}{1+2u}\,du = -\frac{3}{2}\log(1+2T). \tag{4.19}$$

Then

$$e^{-\int_0^{t_x} R(\gamma_x(u))\,du} \leq (1+2T)^{\frac{3}{2}}. \tag{4.20}$$

As

$$e^{-\int_0^{t} R(\gamma_x(u))\,du} = e^{-\int_0^{t_x} R(\gamma_x(u))\,du} e^{-\int_{t_x}^{t} R(\gamma_x(u))\,du}, \tag{4.21}$$

by combining (4.18) and (4.20) we obtain

$$\frac{\mathrm{dvol}_{g(t)}}{\mathrm{dvol}_{g(0)}}(x) = J_t(x) \leq C''(1+2T)^{\frac{3}{2}}\left(\frac{R(x)}{\overline{R}}\right)^{-p}. \tag{4.22}$$

Then

$$\begin{aligned} \int_{\mathcal{M}_{t,\mathrm{neck}}^{>\overline{R}'}} R^p\ \mathrm{dvol}_{g(t)} &\leq C''\overline{R}^p(1+2T)^{\frac{3}{2}} \int_{\mathcal{M}_{t,\mathrm{neck}}^{>\overline{R}'}} \mathrm{dvol}_{g(0)} \\ &\leq C''\overline{R}^p(1+2T)^{\frac{3}{2}}\,\mathrm{vol}_{g(0)}(\mathcal{M}_0). \end{aligned} \tag{4.23}$$

This finishes the discussion of the neck points.

Let $\overline{R}'' > \overline{R}'$ be a new parameter. Given $\sigma > 0$ small, let $\mathcal{M}_{t,\mathrm{round}}$ be the connected components of $\mathcal{M}_t$ that intersect $\mathcal{M}_{t,\mathrm{closed}}^{>\overline{R}''}$ and are σ-close to a constant curvature metric, and let $\mathcal{M}_{t,\mathrm{nonround}}$ be the other connected components of $\mathcal{M}_t$ that intersect $\mathcal{M}_{t,\mathrm{closed}}^{>\overline{R}''}$. Using [KL17, Proposition 5.17], a connected component $\mathcal{N}_t$ in $\mathcal{M}_t$ determines a connected component $\mathcal{N}_{t'}$ in $\mathcal{M}_{t'}$ for all $t' \leq t$.

Let $\mathcal{N}_t$ be a component in $\mathcal{M}_{t,\mathrm{nonround}}$. From (4.7), we have $R_{\mathrm{av}} \geq C^{-1}\overline{R}''$. Using the compactness of the space of approximating κ-solutions, we can apply Lemma 3.1 and Corollary 3.3. Then for ϵ' small and $\mathcal{T} = \mathcal{T}(\epsilon') < \infty$, there is some $t' \in [t, t - 10C\mathcal{T}R_{\mathrm{av}}^{-1}]$ so that $\mathcal{N}_{t'}$ consists of centers of ϵ'-necks and two caps. From (4.9), if $x \in \mathcal{N}_t$ and

$$R\Big|_{\gamma_x([t',t])} \geq \overline{R} \tag{4.24}$$

then

$$\frac{1}{R(\gamma_x(t'))} \leq \frac{C}{R_{\mathrm{av}}} + 10C\eta\mathcal{T}R_{\mathrm{av}}^{-1}, \tag{4.25}$$

so

$$R(\gamma_x(t')) \geq \frac{R_{\mathrm{av}}}{C(1+10\eta\mathcal{T})} \geq \frac{\overline{R}''}{C^2(1+10\eta\mathcal{T})}. \tag{4.26}$$

If $\overline{R}'' > C^2(1+10\eta\mathcal{T})\overline{R}'$ then (4.24) holds and from (4.26), $\mathcal{N}_{t'} \subset \mathcal{M}_{t'}^{>\overline{R}'}$. From (4.7) and (4.26), we also have

$$R(x) \leq CR_{\mathrm{av}} \leq C^2(1+10\eta\mathcal{T})R(\gamma_x(t')). \tag{4.27}$$

Since the volume element at $\gamma_u(x)$ is nonincreasing as a function of $u \in [t', t]$, we obtain

$$\int_{\mathcal{N}_t} R \ \mathrm{dvol}_{g(t)} \leq C^2(1+10\eta\mathcal{T}) \int_{\mathcal{N}_{t'}} R \ \mathrm{dvol}_{g(t')}. \tag{4.28}$$

We now apply the argument starting with (4.8) to $\mathcal{N}_{t'}$. Taking $\frac{\overline{R}'}{\overline{R}}$ large compared to $\frac{\overline{R}''}{\overline{R}'}$, in order to ensure many iterations in the earlier-neck argument, we get a bound

$$\int_{\mathcal{N}_t} |R|^p \ \mathrm{dvol}_{g(t)} \leq \mathrm{const.}(p,T)\,\mathrm{vol}_{g(0)}(\mathcal{N}_t). \tag{4.29}$$

This takes care of the components in $\mathcal{M}_{t,\mathrm{nonround}}$.

Let $\mathcal{N}_t$ be a component of $\mathcal{M}_t$ in $\mathcal{M}_{t,\mathrm{round}}$. Let τ be the infimum of the u's so that for all $t' \in [u,t]$, the metric on $\mathcal{N}_{t'}$ is σ-close to a constant curvature metric. For a Ricci flow solution with time slices of constant positive curvature, R is strictly increasing along forward worldlines but $\int R \ \mathrm{dvol}$ is strictly decreasing in t. Hence if σ is sufficiently small then we are ensured that

$$\int_{\mathcal{N}_t} R \ \mathrm{dvol}_{g(t)} \leq \int_{\mathcal{N}_\tau} R \ \mathrm{dvol}_{g(\tau)}. \tag{4.30}$$

If N_τ has a point with scalar curvature at most $\overline{R}''$ and σ is small then

$$\int_{\mathcal{N}_\tau} R \ \mathrm{dvol} \leq 2\overline{R}''\,\mathrm{vol}_{g(\tau)}(\mathcal{N}_\tau) \leq 2\overline{R}''(1+2T)^{\frac{3}{2}}\,\mathrm{vol}_{g(0)}(\mathcal{N}_t). \tag{4.31}$$

If, on the other hand, $N_\tau \subset \mathcal{M}_\tau^{>\overline{R}''}$ then we can apply the preceding argument for $\mathcal{M}_{t,\mathrm{nonround}}$, replacing t by τ. The conclusion is that

$$\int_{\mathcal{M}_{t,\mathrm{closed}}^{>\overline{R}''}} |R|^p \ \mathrm{dvol}_{g(t)} \leq \mathrm{const.}(p,T)\,\mathrm{vol}_{g(0)}(\mathcal{M}_{t,\mathrm{closed}}^{>\overline{R}''}). \tag{4.32}$$

Since

$$\int_{\mathcal{M}_{t,\mathrm{neck}}^{>\overline{R}''}} R^p \ \mathrm{dvol}_{g(t)} \leq \int_{\mathcal{M}_{t,\mathrm{neck}}^{>\overline{R}'}} R^p \ \mathrm{dvol}_{g(t)}, \tag{4.33}$$

$$\int_{\mathcal{M}_{t,\mathrm{cap}}^{>\overline{R}''}} R^p \ \mathrm{dvol}_{g(t)} \leq \int_{\mathcal{M}_{t,\mathrm{cap}}^{>\overline{R}'}} R^p \ \mathrm{dvol}_{g(t)} \tag{4.34}$$

and

$$\int_{\mathcal{M}_t^{\leq \overline{R}''}} |R|^p \operatorname{dvol}_{g(t)} \leq \left(\overline{R}''\right)^p (1+2T)^{\frac{3}{2}} \operatorname{vol}_{g(0)} \left(\mathcal{M}_t^{\leq \overline{R}''}\right), \tag{4.35}$$

the proposition follows from (4.8), (4.23), (4.32), (4.33), (4.34) and (4.35). □

Proposition 4.36. *Let $\mathcal{M}$ be a singular Ricci flow. Let η be the constant from (4.1). We can assume that $\eta \geq 1$. Then whenever $0 \leq t_1 \leq t_2 < \infty$ satisfies $t_2 - t_1 < \frac{1}{\eta} r(t_2)^2$ and $t_1 > \frac{1}{100\eta}$, we have*

$$\begin{aligned} &\mathcal{V}(t_2) - \mathcal{V}(t_1) \\ &\geq -\eta^{\frac{1}{\eta}} \left(2 \int_{\mathcal{M}_{t_1}} |R|^{\frac{1}{\eta}} \operatorname{dvol}_{g(t_1)} + r(t_2)^{-\frac{2}{\eta}} \mathcal{V}(t_1)\right) (t_2 - t_1)^{\frac{1}{\eta}} \\ &\geq -5\eta^{\frac{1}{\eta}} r(t_2)^{-\frac{2}{\eta}} (1+2t_1)^{\frac{3}{2}} \mathcal{V}(0) \cdot (t_2 - t_1)^{\frac{1}{\eta}}. \end{aligned} \tag{4.37}$$

Proof. Let $X_1 \subset \mathcal{M}_{t_1}$ be the set of points $x \in \mathcal{M}_{t_1}$ whose worldline γ_x extends forward to time t_2 and let $X_2 \subset \mathcal{M}_{t_1}$ be the points x whose worldline γ_x does not extend forward to time t_2. Put

$$X_1' = \left\{ x \in X_1 \ : \ R(x) > \frac{1}{\eta(t_2 - t_1)} \right\}, \tag{4.38}$$

$$X_1'' = \left\{ x \in X_1 \ : \ r(t_2)^{-2} < R(x) \leq \frac{1}{\eta(t_2 - t_1)} \right\} \tag{4.39}$$

and

$$X_1''' = \left\{ x \in X_1 \ : \ R(x) \leq r(t_2)^{-2} \right\}. \tag{4.40}$$

Then

$$\begin{aligned} \operatorname{vol}(\mathcal{M}_{t_2}) - \operatorname{vol}(\mathcal{M}_{t_1}) \geq\, & \operatorname{vol}_{t_2}(X_1') - \operatorname{vol}_{t_1}(X_1') \\ & + \operatorname{vol}_{t_2}(X_1'') - \operatorname{vol}_{t_1}(X_1'') \\ & + \operatorname{vol}_{t_2}(X_1''') - \operatorname{vol}_{t_1}(X_1''') - \operatorname{vol}_{t_1}(X_2) \\ \geq\, & \operatorname{vol}_{t_2}(X_1'') - \operatorname{vol}_{t_1}(X_1'') + \operatorname{vol}_{t_2}(X_1''') \\ & - \operatorname{vol}_{t_1}(X_1''') - \operatorname{vol}_{t_1}(X_2) - \operatorname{vol}_{t_1}(X_1'). \end{aligned} \tag{4.41}$$

Suppose that $x \in X_2$.

Lemma 4.42. *Let $[t_1, t_x)$ be the domain of the forward extension of γ_x, with $t_x < t_2$. For all $u \in [t_1, t_x)$, we have*

$$R(\gamma_x(u)) \geq \frac{1}{\eta(t_x - u)}. \tag{4.43}$$

Proof. If the lemma is not true, put

$$u' = \sup \left\{ u \in [t_1, t_x) \ : \ R(\gamma_x(u)) < \frac{1}{\eta(t_x - u)} \right\}. \tag{4.44}$$

Then $u' > t_1$.

From the gradient estimate (4.1) and the fact that $\lim_{u\to t_x} R(\gamma_x(u)) = \infty$, we know that $u' < t_x$. Whenever $u \geq u'$, we have

$$R(\gamma_x(u)) \geq \frac{1}{\eta(u_x - u')} \geq \frac{1}{\eta(t_2 - t_1)} > r(t_2)^{-2}, \tag{4.45}$$

so there is some $\mu > 0$ so that the gradient estimate (4.1) holds on the interval $(u' - \mu, t_x)$. This implies that (4.43) holds for all $u \in (u' - \mu, t_x)$, which contradicts the definition of u'. This proves the lemma. □

Hence

$$(X_2 \cup X_1') \subset \left\{ x \in \mathcal{M}_{t_1} \ : \ R(x) \geq \frac{1}{\eta(t_2 - t_1)} \right\} \tag{4.46}$$

and

$$\begin{aligned} \mathrm{vol}_{t_1}(X_2) + \mathrm{vol}_{t_1}(X_1') &\leq \mathrm{vol}\left\{ x \in \mathcal{M}_{t_1} \ : \ R(x) \geq \frac{1}{\eta(t_2 - t_1)} \right\} \\ &\leq \eta^{\frac{1}{\eta}} (t_2 - t_1)^{\frac{1}{\eta}} \int_{\mathcal{M}_{t_1}} |R|^{\frac{1}{\eta}} \ \mathrm{dvol}_{g(t_1)}, \end{aligned} \tag{4.47}$$

since $\eta^{\frac{1}{\eta}} (t_2 - t_1)^{\frac{1}{\eta}} |R|^{\frac{1}{\eta}} \geq 1$ on the set $\left\{ x \in \mathcal{M}_{t_1} : R(x) \geq \frac{1}{\eta(t_2 - t_1)} \right\}$.

Suppose now that $x \in X_1''$.

Lemma 4.48. *For all $u \in [t_1, t_2]$, we have*

$$R(\gamma_x(u)) \leq \frac{1}{\frac{1}{R(x)} - \eta(u - t_1)} < \infty. \tag{4.49}$$

Proof. If the lemma is not true, put

$$u'' = \inf \left\{ u \in [t_1, t_2] \ : \ R(\gamma_x(u)) > \frac{1}{\frac{1}{R(x)} - \eta(u - t_1)} \right\}. \tag{4.50}$$

Then $u'' < t_2$ and the gradient estimate (4.1) implies that $u'' > t_1$. Now

$$R(\gamma_x(u'')) = \frac{1}{\frac{1}{R(x)} - \eta(u'' - t_1)} > R(x) > r(t_2)^{-2}. \tag{4.51}$$

Hence there is some $\mu > 0$ so that $R(\gamma_x(u)) \geq r(t_2)^{-2}$ for $u \in [u'', u'' + \mu]$. If $R(\gamma_x(u)) \geq r(t_2)^{-2}$ for all $u \in [t_1, u'']$ then (4.1) implies that (4.49) holds for $u \in [t_1, u'' + \mu]$, which contradicts the definition of u''. On the other hand, if it is not true that $R(\gamma_x(u)) \geq r(t_2)^{-2}$ for all $u \in [t_1, u'']$, put

$$v'' = \sup \left\{ u \in [t_1, u''] \ : \ R(\gamma_x(u)) < r(t_2)^{-2} \right\}. \tag{4.52}$$

Then $v'' > t_1$ and $R(\gamma_x(v'')) = r(t_2)^{-2}$. Equation (4.1) implies that

$$R(\gamma_x(u'')) \leq \frac{1}{r(t_2)^2 - \eta(u'' - v'')} < \frac{1}{\frac{1}{R(x)} - \eta(u'' - t_1)}, \tag{4.53}$$

which contradicts (4.51). This proves the lemma. □

Hence if $x \in X_1''$ then

$$\int_{t_1}^{t_2} R(\gamma_x(u))\, du \le \int_{t_1}^{t_2} \frac{1}{\frac{1}{R(x)} - \eta(u-t_1)}\, du \tag{4.54}$$
$$= -\frac{1}{\eta} \log\left(1 - \eta R(x) \cdot (t_2 - t_1)\right),$$

so

$$\frac{\mathrm{dvol}_{g(t_2)}}{\mathrm{dvol}_{g(t_1)}}(x) = J_{t_2}(x) \ge (1 - \eta R(x) \cdot (t_2 - t_1))^{\frac{1}{\eta}}. \tag{4.55}$$

Thus

$$\mathrm{vol}_{t_2}(X_1'') - \mathrm{vol}_{t_1}(X_1'') \ge \tag{4.56}$$
$$\int_{X_1''} \left((1 - \eta R \cdot (t_2 - t_1))^{\frac{1}{\eta}} - 1 \right) \mathrm{dvol}_{g(t_1)}.$$

Since $\eta \ge 1$, if $z \in [0,1]$ then $\left(z^{\frac{1}{\eta}}\right)^\eta + \left(1 - z^{\frac{1}{\eta}}\right)^\eta \le 1$, so

$$(1-z)^{\frac{1}{\eta}} - 1 \ge -z^{\frac{1}{\eta}}. \tag{4.57}$$

Then

$$\mathrm{vol}_{t_2}(X_1'') - \mathrm{vol}_{t_1}(X_1'') \ge -\eta^{\frac{1}{\eta}} (t_2 - t_1)^{\frac{1}{\eta}} \int_{X_1''} R^{\frac{1}{\eta}}\ \mathrm{dvol}_{g(t_1)}. \tag{4.58}$$

Now suppose that $x \in X_1'''$.

Lemma 4.59. *For all $u \in [t_1, t_2]$, we have*

$$R(\gamma_x(u)) \le \frac{1}{r(t_2)^2 - \eta(u - t_1)} < \infty. \tag{4.60}$$

Proof. If the lemma is not true, put

$$u''' = \inf \left\{ u \in [t_1, t_2] \, : \, R(\gamma_x(u)) > \frac{1}{r(t_2)^2 - \eta(u - t_1)} \right\}. \tag{4.61}$$

Then $u''' < t_2$. If $R(x) < r(t_2)^{-2}$ then clearly $u''' > t_1$. If $R(x) = r(t_2)^{-2}$ then since $r(t_1) > r(t_2)$, there is some $\nu > 0$ so that $R(\gamma_x(u)) > r(u)^{-2}$ for $u \in [t_1, t_1 + \nu]$; then (4.1) gives the validity of (4.60) for $u \in [t_1, t_1 + \nu]$, which implies that $u''' > t_1$. In either case, $t_1 < u''' < t_2$. Now

$$R(\gamma_x(u''')) = \frac{1}{r(t_2)^2 - \eta(u''' - t_1)} > r(t_2)^{-2}. \tag{4.62}$$

Hence there is some $\mu > 0$ so that $R(\gamma_x(u)) \ge r(t_2)^{-2}$ for $u \in [u''', u''' + \mu]$. If $R(\gamma_x(u)) \ge r(t_2)^{-2}$ for all $u \in [t_1, u''']$ then (4.1) implies that (4.60) holds for $u \in [t_1, u''' + \mu]$, which contradicts the definition of u'''. On the other hand, if it is not true that $R(\gamma_x(u)) \ge r(t_2)^{-2}$ for all $u \in [t_1, u''']$, put

$$v''' = \sup \left\{ u \in [t_1, u'''] \, : \, R(\gamma_x(u)) < r(t_2)^{-2} \right\}. \tag{4.63}$$

Then $v''' > t_1$ and $R(\gamma_x(v''')) = r(t_2)^{-2}$. The gradient estimate (4.1) implies that

$$R(\gamma_x(u''') \leq \frac{1}{r(t_2)^2 - \eta(u''' - v''')} < \frac{1}{r(t_2)^2 - \eta(u''' - t_1)}, \tag{4.64}$$

which contradicts (4.62). This proves the lemma. □

Hence if $x \in X_1'''$ then

$$\begin{aligned} \int_{t_1}^{t_2} R(\gamma_x(u))\, du &\leq \int_{t_1}^{t_2} \frac{1}{r(t_2)^2 - \eta(u - t_1)}\, du \\ &= -\frac{1}{\eta} \log\left(1 - \eta r(t_2)^{-2} \cdot (t_2 - t_1)\right), \end{aligned} \tag{4.65}$$

so

$$\frac{\mathrm{dvol}_{g(t_2)}}{\mathrm{dvol}_{g(t_1)}}(x) = J_{t_2}(x) \geq \left(1 - \eta r(t_2)^{-2} \cdot (t_2 - t_1)\right)^{\frac{1}{\eta}}. \tag{4.66}$$

Thus

$$\begin{aligned} &\mathrm{vol}_{t_2}(X_1''') - \mathrm{vol}_{t_1}(X_1''') \geq \\ &\int_{X_1'''} \left(\left(1 - \eta r(t_2)^{-2} \cdot (t_2 - t_1)\right)^{\frac{1}{\eta}} - 1\right) \mathrm{dvol}_{g(t_1)}. \end{aligned} \tag{4.67}$$

Since $\eta r(t_2)^{-2} \cdot (t_2 - t_1) \in [0, 1]$, we can apply (4.57) to conclude that

$$\begin{aligned} \mathrm{vol}_{t_2}(X_1''') - \mathrm{vol}_{t_1}(X_1''') &\geq -\eta^{\frac{1}{\eta}} r(t_2)^{-\frac{2}{\eta}} \cdot (t_2 - t_1)^{\frac{1}{\eta}} \mathrm{vol}_{t_1}(X_1''') \\ &\geq -\eta^{\frac{1}{\eta}} r(t_2)^{-\frac{2}{\eta}} \mathcal{V}(t_1) \cdot (t_2 - t_1)^{\frac{1}{\eta}}. \end{aligned} \tag{4.68}$$

Combining (4.41), (4.47), (4.58) and (4.68) gives (4.37). □

5. Asymptotic conditions

In this section we show that the *a priori* assumptions in Definition 2.2 are really conditions on the spacetime near infinity. That is, given $\epsilon > 0$ and a decreasing function $r : [0, \infty) \to (0, \infty)$, there are decreasing functions $\kappa' = \kappa'(\epsilon) : [0, \infty) \to (0, \infty)$ and $r' = r'(\epsilon, r) : [0, \infty) \to (0, \infty)$ with the following property. Let $\mathcal{M}$ be a Ricci flow spacetime with normalized initial condition, on which condition (a) of Definition 2.2 holds. Suppose that for each $T \geq 0$ there is a compact subset of $\mathcal{M}_{\leq T}$ so that condition (b), and the r-canonical neighborhood assumption of condition (c), hold on the part of $\mathcal{M}_{\leq T}$ outside of the compact subset. Then $\mathcal{M}$ satisfies Definition 2.2 globally with parameters ϵ, κ' and r'.

If $\mathcal{M}$ is a Ricci flow spacetime and $m_0 \in \mathcal{M}$, put $t_0 = \mathfrak{t}(m_0)$. We define Perelman's l-function using curves emanating backward from m_0, as in [KL08, Section 15]. That is, given $m \in \mathcal{M}$ with $\mathfrak{t}(m) < t_0$, consider a time-preserving map $\gamma : [\mathfrak{t}(m), t_0] \to \mathcal{M}$ from m to m_0. We reparametrize $[\mathfrak{t}(m), t_0]$ by $\tau(t) = t_0 - t$. Then

$$\mathcal{L}(\gamma) = \int_0^{t_0 - \mathfrak{t}(m)} \sqrt{\tau}\left(R(\gamma(\tau)) + |\dot{\gamma}(\tau)|^2\right)\, d\tau, \tag{5.1}$$

where $\dot{\gamma}$ is the spatial projection of the velocity vector of γ and $|\dot{\gamma}(\tau)|$ is computed using the metric on $\mathcal{M}_{t_0-\tau}$ at $\gamma(\tau)$. Let $L(m)$ be the infimal $\mathcal{L}$-length of such curves γ. The reduced length is

$$l(m) = \frac{L(m)}{2\sqrt{t_0 - \mathfrak{t}(m)}}. \tag{5.2}$$

Proposition 5.3. *Given $\overline{\Delta r}, \overline{\Delta t} > 0$ there are $\Delta r = \Delta r(m_0) < \overline{\Delta r}$ and $\Delta t = \Delta t(m_0) < \overline{\Delta t}$ with the following property. For any $m \in \mathcal{M}$ with $\mathfrak{t}(m) < t_0 - \Delta t$, let $d_{qp}(m, P(m_0, \Delta r, -\Delta t))$ denote the $g^{qp}_{\mathcal{M}}$-distance from m to the set*

$$P(m_0, \Delta r, -\Delta t).$$

Then

$$L(m) \geq \min\left(\frac{(\Delta r)^2}{4\sqrt{\Delta t}}, \frac{\sqrt{\Delta t}}{10}\, d_{qp}(m, P(m_0, \Delta r, -\Delta t))\right) - \frac{8}{3} t_0^{\frac{3}{2}}. \tag{5.4}$$

Proof. With a slight variation on Perelman's definition [KL08, Definition 79.1], we put

$$\mathcal{L}_+(\gamma) = \int_0^{t_0 - \mathfrak{t}(m)} \sqrt{\tau}\left(R_+(\gamma(\tau)) + |\dot{\gamma}(\tau)|^2\right)\, d\tau, \tag{5.5}$$

where $R_+(m) = \max(R(m), 1)$. We define $L_+(m)$ using $\mathcal{L}_+(\gamma)$ instead of $\mathcal{L}(\gamma)$. Applying the lower curvature bound [KL17, (5.3)] (with $C = n = 3$), we know that $R \geq -3$ and so

$$\begin{aligned} \mathcal{L}(\gamma) - \mathcal{L}_+(\gamma) &= \int_0^{t_0 - \mathfrak{t}(m)} \sqrt{\tau}\left(R(\gamma(\tau)) - R_+(\gamma(\tau))\right)\, d\tau \\ &\geq -\int_0^{t_0 - \mathfrak{t}(m)} \sqrt{\tau} \cdot 4\, d\tau \geq -\frac{8}{3} t_0^{\frac{3}{2}}. \end{aligned} \tag{5.6}$$

Hence it suffices to estimate $\mathcal{L}_+(\gamma)$ from below.

Given numbers $\Delta r, \Delta t > 0$, if $\mathfrak{t}(m) < t_0 - \Delta t$ then

$$\mathcal{L}_+(\gamma) \geq \int_0^{\Delta t} \sqrt{\tau}|\dot{\gamma}(\tau)|^2\, d\tau = \frac{1}{2}\int_0^{\sqrt{\Delta t}} \left|\frac{d\gamma}{ds}\right|^2\, ds. \tag{5.7}$$

Suppose first that γ leaves m_0 and exits $P(m_0, \Delta r, -\Delta t)$ at some time $t \in (t_0 - \Delta t, t_0)$. If the parabolic ball were Euclidean then we could say from (5.7) that $\mathcal{L}_+(\gamma) \geq \frac{1}{2}\frac{(\Delta r)^2}{\sqrt{\Delta t}}$. If Δr and Δt are small enough, depending on m_0, then we can still say that $P(m_0, \Delta r, -\Delta t)$ is unscathed and $\mathcal{L}_+(\gamma) \geq \frac{1}{4}\frac{(\Delta r)^2}{\sqrt{\Delta t}}$. Given such values of Δr and Δt, suppose now that γ does not exit $P(m_0, \Delta r, -\Delta t)$ in the time interval $(t_0 - \Delta t, t_0)$. Then $\gamma(t_0 - \Delta t) \in P(m_0, \Delta r, -\Delta t)$. Now

$$\mathcal{L}_+(\gamma) \geq \sqrt{\Delta t}\int_{\Delta t}^{t_0 - \mathfrak{t}(m)} \left(R_+(\gamma(\tau)) + |\dot{\gamma}(\tau)|^2\right)\, d\tau, \tag{5.8}$$

Since $R \geq -3$, it follows that along γ, we have

$$\sqrt{1 + R^2} \leq 10 R_+. \tag{5.9}$$

Then

$$\sqrt{1+R^2}|\dot\gamma|^2 + 1 + R^2 \le 10R_+|\dot\gamma|^2 + 100R_+^2 \tag{5.10}$$
$$\le 100\left(|\dot\gamma|^4 + 2R_+|\dot\gamma|^2 + R_+^2\right),$$

so

$$\sqrt{\sqrt{1+R^2}|\dot\gamma|^2 + 1 + R^2} \le 10\left(|\dot\gamma|^2 + R_+\right). \tag{5.11}$$

Thus

$$\mathcal{L}_+(\gamma) \ge \frac{\sqrt{\Delta t}}{10}\int_{\Delta t}^{t_0-\mathfrak{t}(m)} \left|\frac{d\gamma}{d\tau}\right|_{g^{qp}_{\mathcal{M}}} d\tau \tag{5.12}$$
$$\ge \frac{\sqrt{\Delta t}}{10} d_{qp}(m, P(m_0, \Delta r, -\Delta t)).$$

The proposition follows. □

Corollary 5.13. *Suppose that $(\mathcal{M}_{\le t_0}, g^{qp}_{\mathcal{M}})$ is complete away from the time-zero slice and the time-t_0 slice. Given $t' < t_0$, the restriction of L to $\mathcal{M}_{\le t'}$ is proper and bounded below.*

Proof. From (5.4), the function L is bounded below on $\mathcal{M}_{\le t'}$. Suppose that it is not proper. Then for some $C < \infty$, there is a sequence $\{m_i\}_{i=1}^\infty$ in $\mathcal{M}_{\le t'}$ going to infinity with $L(m_i) < C$ for all i. We can choose $\Delta r, \Delta t > 0$ with $\Delta t < t_0 - t'$ and

$$\frac{(\Delta r)^2}{4\sqrt{\Delta t}} - \frac{8}{3}t_0^{\frac{3}{2}} \ge C. \tag{5.14}$$

By the completeness of $g^{qp}_{\mathcal{M}}$, we have $\lim_{i\to\infty} d_{qp}(m_i, P(m_0, \Delta r, -\Delta t)) = \infty$. Then (5.4) gives a contradiction. □

It is not hard to see that L is continuous on $\mathcal{M}_{<t_0}$. From the proof of Proposition 5.3, given $m \in \mathcal{M}_{<t_0}$ and $K < \infty$, the time-preserving curves $\gamma : [\mathfrak{t}(m), t_0] \to \mathcal{M}$ from m to m_0 with $\mathcal{L}(\gamma) < K$ lie in a compact subset of $\mathcal{M}$. From standard arguments [KL08, Section 17], it follows that there is an $\mathcal{L}$-minimizer from m_0 to m.

Since L is bounded below and time-slices have finite volume from [KL17, Corollary 7.7], the reduced volume $\tilde{V}(\tau) = \tau^{-\frac{3}{2}}\int_{\mathcal{M}_{t_0-\tau}} e^{-l}\,\mathrm{dvol}$ exists. The results of [KL08, Sections 17–29] go through in our setting. In particular, $\tilde{V}(\tau)$ is nonincreasing in τ.

Proposition 5.15. *Suppose that $(\mathcal{M}_{\le t_0}, g^{qp}_{\mathcal{M}})$ is complete away from the time-zero slice and the time-t_0 slice. For every $t \in [0, t_0)$, there is some $m \in \mathcal{M}_t$ with $l(m) \le \frac{3}{2}$.*

Proof. Putting $\overline{L} = 2\sqrt{t_0 - \mathfrak{t}}\, L$, we have

$$\partial_{\mathfrak{t}}(-\overline{L} + 6(t_0 - \mathfrak{t})) \le \triangle(-\overline{L} + 6(t_0 - \mathfrak{t})) \tag{5.16}$$

in the barrier sense [KL08, Section 24]. From Corollary 5.13, for each $\widetilde{t'} \in [0, t_0)$, the function $-\overline{L} + 6(t_0 - \mathfrak{t})$ is proper and bounded above on $\mathcal{M}_{\le \widetilde{t'}}$. In particular, for each $t \in [0, t_0)$, the maximum of $-\overline{L} + 6(t_0 - t)$ exists on $\mathcal{M}_t$. We want to show that the maximum is nonnegative. By way of contradiction, suppose that for some $\widetilde{t} \in [0, t_0)$ and some $\alpha < 0$, we have $-\overline{L}(m) + 6(t_0 - \widetilde{t}) \leq \alpha$ for all $m \in \mathcal{M}_{\widetilde{t}}$. Given $\widetilde{t}' \in (\widetilde{t}, t_0)$, we can apply [KL17, Lemma 5.1] on the interval $[\widetilde{t}, \widetilde{t}']$ to conclude that $-\overline{L}(m) + 6(t_0 - \widetilde{t}') \leq \alpha$ for all $m \in \mathcal{M}_{\widetilde{t}'}$. However, along the worldline γ going through m_0, for small $\tau > 0$ we have

$$\overline{L}(\gamma(t_0 - \tau)) \leq \text{const.}\, \tau^2. \tag{5.17}$$

Then for small τ, we have $-\overline{L}(\gamma(t_0 - \tau)) + 6\tau > 0$. Taking $\widetilde{t}' = t_0 - \tau$ gives a contradiction and proves the proposition. □

In his first Ricci flow paper, Perelman showed that there is a decreasing function $\kappa' : [0, \infty) \to (0, \infty)$ with the property that if $\mathcal{M}$ is a smooth Ricci flow solution, with normalized initial conditions, then $\mathcal{M}$ is κ'-noncollapsed at scales less than ϵ [KL08, Theorem 26.2].

Proposition 5.18. *Let $\mathcal{M}$ be a Ricci flow spacetime with normalized initial condition. Given $t' > 0$, suppose that $(\mathcal{M}_{\leq t'}, g^{qp}_{\mathcal{M}})$ is complete away from the time-zero slice and the time-t' slices. Then $\mathcal{M}_{\leq t'}$ is κ'-noncollapsed at scales less than ϵ.*

Proof. The proof is along the lines of that of [KL08, Theorem 26.2]. We can assume that $t' > \frac{1}{100}$. To prove κ'-noncollapsing near $m_0 \in \mathcal{M}_{\leq t'}$, we consider $\mathcal{L}$-curves emanating backward in time from m_0 to a fixed time slice $\mathcal{M}_{\overline{t}}$, say with $\overline{t} = \frac{1}{100}$. By Proposition 5.15, there is some $m \in \mathcal{M}_{\overline{t}}$ with $l(m) \leq \frac{3}{2}$. Using the bounded geometry near m and the monotonicity of $\widetilde{V}$, the κ'-noncollapsing follows as in [KL08, Pf. of Theorem 26.2]. □

We now show that the conditions in Definition 2.2, to define a singular Ricci flow, are actually asymptotic in nature.

Proposition 5.19. *Given $\epsilon > 0$, $t' < \infty$ and a decreasing function $r : [0, t'] \to (0, \infty)$, there is some $r' = r'(\epsilon, r) > 0$ with the following property. Let $\mathcal{M}$ be a Ricci flow spacetime such that $R : \mathcal{M}_{\leq t'} \to \mathbb{R}$ is bounded below and proper, and there is a compact set $K \subset \mathcal{M}_{\leq t'}$ so that for each $m \in \mathcal{M}_{\leq t'} - K$,*

(a) *The Hamilton–Ivey pinching condition of* [KL17, (A.14)] *is satisfied at m, with time parameter $\mathfrak{t}(m)$, and*
(b) *The r-canonical neighborhood assumption of* [KL17, Appendix A.8] *is satisfied at m.*

Then the conditions of Definition 2.2 *hold on $\mathcal{M}_{\leq t'}$, with parameters ϵ, κ' and r'.*

Proof. Condition (a) of Definition 2.2 holds on $\mathcal{M}_{\leq t'}$ by assumption.

Also by assumption, for $m \in \mathcal{M}_{\leq t'} - K$, the curvature operator at m lies in the convex cone of [KL17, (A.13)]. The proof of Hamilton–Ivey pinching in [CLN06, Pf. of Theorem 6.44], using the vector-valued maximum principle, now goes through since any violations in $\mathcal{M}_{\leq t'}$ of [KL17, (A.14)] would have to occur in K. This shows that condition (b) of Definition 2.2 holds on $\mathcal{M}_{\leq t'}$.

Since the r-canonical neighborhood assumption holds on $\mathcal{M}_{\leq t'} - K$ the proof of [KL17, Lemma 5.13] shows that $g^{qp}_{\mathcal{M}}$ is complete on $\mathcal{M}_{\leq t'}$ away from the time-zero slice and the time-t' slice. Proposition 5.18 now implies that $\mathcal{M}_{\leq t'}$ is κ'-noncollapsed at scales less than ϵ.

To show that condition (c) of Definition 2.2 holds on $\mathcal{M}_{\leq t'}$, with parameters ϵ and κ', and some parameter $r' > 0$, we apply the method of proof of [KL08, Theorem 52.7] for smooth Ricci flow solutions. Suppose that there is no such r'. Then there is a sequence $\{\mathcal{M}^k\}_{k=1}^{\infty}$ of Ricci flow spacetimes satisfying the assumptions of the proposition, and a sequence $r'_k \to 0$, so that for each k there is some $m_k \in \mathcal{M}^k_{\leq t}$ where the r'_k-canonical neighborhood assumption does not hold. The first step in [KL08, Pf. of Theorem 52.7] is to find a point of violation so that there are no nearby points of violation with much larger scalar curvature, in an earlier time interval which is long in a scale-invariant sense. The proof of this first step uses point selection. Because of our assumption that the r-canonical neighborhood assumption holds in $\mathcal{M}^k_{\leq t'} - K^k$, as soon as $r'_k < r(t')$ we know that any point of violation lies in K^k. Thus this point selection argument goes through. The second step in [KL08, Pf. of Theorem 52.7] is a bounded-curvature-at-bounded-distance statement that uses Hamilton–Ivey pinching and κ'-noncollapsing. Since we have already proven that the latter two properties hold, the proof of the second step goes through. The third and fourth steps in [KL08, Pf. of Theorem 52.7] involve constructing an approximating κ'-solution. These last two steps go through without change. □

Proposition 5.19 shows that the r'-canonical neighborhood assumption holds with parameter $r' = r'(t')$. We can assume that r' is a decreasing function of t'. Hence $\mathcal{M}$ is a singular Ricci flow with parameters ϵ, $\kappa' = \kappa'(\epsilon)$ and $r' = r'(\epsilon, r)$.

6. Dimension of the set of singular times

In this section we give an upper bound on the Minkowski dimension of the set of singular times for a Ricci flow spacetime.

The geometric input comes from the proofs of Propositions 4.2 and 4.36. We isolate it in the following lemma. The lemma states that any point with large curvature determines a region in backward spacetime on which the scalar curvature behaves nicely (i.e., R^{-1} grows with upper and lower linear bounds as one goes backward in time), and which carries a controlled amount of volume.

Lemma 6.1. *For every $\lambda > 0$, $t < \infty$ there is a constant $C = C(\lambda, t) < \infty$ with the following property.*

Let $\mathcal{M}$ be a singular Ricci flow and suppose $x \in \mathcal{M}_t$ is a point with $\rho(x) := R^{-\frac{1}{2}}(x) \leq C^{-1} r(t)$. Then there is a product domain $U \subset \mathcal{M}$ defined on the time interval $[t_-, t]$, where $t_- := t - C^{-1} r^2(t)$, with the following properties:

1. *$U_t \subset B(x, C\rho(x))$.*

2. (*Scalar curvature control*) *For all* $t' \in [t_-, t]$, $x' \in U_{t'}$, *we have*

$$C^{-1}R^{-1}(x) + \eta_-(t - t') \le R^{-1}(x') \le CR^{-1}(x) + \eta_+(t - t') . \tag{6.2}$$

Here $\eta_\pm$ *are constants coming from the geometry of* κ*-solutions.*

3. (*Volume control*) *For* $t' \in [t - C^{-1}r^2(t), t - \frac{1}{2}C^{-1}r^2(t)]$ *we have*

$$\mathrm{vol}(U_{t'}) \ge C^{-1}r^{2-\lambda}(t)\rho^{1+\lambda}(x).$$

In particular the spacetime volume of U *is at least* $\frac{1}{2}C^{-2}r^{4-\lambda}(t)\rho^{1+\lambda}(x)$.

Proof. The proof of the lemma is based on arguments similar to those in the proofs of Propositions 4.2 and 4.36. We give an outline of the proof. The details are similar to those for Propositions 4.2 and 4.36.

Case 1. *x is sufficiently neck-like that we can apply the neck stability result.* Then we let U be a product domain with $U_t = B(x, \rho(x))$. The scalar curvature estimate (6.2) then follows from the fact that the worldline of every $y \in U_t$ remains neck-like until its scale becomes comparable to the canonical neighborhood scale. The volume estimates in (3) follow using the Jacobian estimate, as in [KL17, Section 5] or in the proof of Proposition 4.2.

Case 2. *The canonical neighborhood of x is neither sufficiently neck-like, nor nearly round.* Then for a constant c independent of x, we find that $P(x, c\rho(x)) \cap \mathcal{M}_{t-c\rho^2(x)}$ contains a necklike point y to which the previous argument applies. If the product domain U_y associated with y is defined on the time interval $[t_0, t - c\rho^2(x)]$, then we let U be the result of extending U_y to the interval $[t_0, t]$.

Case 3. *The canonical neighborhood of x is nearly round, i.e., $B := B(x, 100\rho(x))$ is nearly isometric to a spherical space form, modulo rescaling.* We follow this region backward in time, and have two subcases:

3(a). The region remains nearly round until its scale becomes comparable to $r(t)$. Then we take U to be the product region with $U_t = B$, and the scalar curvature and volume estimates follow readily from the fact that time slices of U are nearly round.

3(b). For some $t_0 < t$, and every $t' \in [t_0, t]$, the image $B_{t'}$ of B in $\mathcal{M}_{t'}$ under the flow of the time vector field ∂_t is δ-close to round, but $B_{t_0} \subset \mathcal{M}_{t_0}$ is not $\frac{\delta}{2}$-close to round. Then we can apply Case 2 to B_{t_0} to obtain a product region U', and we define U by extending U' forward in time over the time interval $[t_0, t]$. □

As a corollary of this lemma, we get:

Theorem 6.3. *If $\mathcal{M}$ is a singular Ricci flow, $T < \infty$, then the set of times $t \in [0, T]$ such that $\mathcal{M}_t$ is noncompact has Minkowski dimension $\le \frac{1}{2}$.*

Proof. Choose $\lambda > 0$, and let $C = C(\lambda, T) < \infty$ be the constant from Lemma 6.1.

Pick $A > C_\lambda^2 r^{-2}(T)$. Let $\mathcal{T}_A$ be the set of times $t \in [0, T]$ such that the time slice $\mathcal{M}_t$ contains a point with $R > A$.

Let $\{t_i^0\}_{i \in I}$ be a maximal A^{-1}-separated subset of $\mathcal{T}_A$. For every $i \in I$, we may find $t_i \in [0, T]$ with

$$|t_i - t_i^0| \le \mathrm{const.}\, A^{-1} \tag{6.4}$$

such that $\mathcal{M}_{t_i}$ contains a point x_i with $R(x_i) = A$; the existence of such a point t_i follows by iterating [KL17, Lemma 3.3]. Now we apply Lemma 6.1 to x_i, for every $i \in I$, to obtain a collection $\{U_i\}_{i\in I}$ of product domains in $\mathcal{M}$.

Note that if $i, j \in I$ and $U_i \cap U_j \neq \emptyset$, then comparing the scalar curvature using Lemma 6.1(2), we get that $|t_i - t_j| < C_1 A^{-1}$ for some $C_1 = C_1(\lambda, T)$. Hence the collection $\{U_i\}_{i\in I}$ has intersection multiplicity $< N = N(\lambda, T)$. Now Lemma 6.1(3) implies that the spacetime volume of each U_i is at least

$$\frac{1}{2} C^{-2} r^{4-\lambda}(T) A^{-\frac{1}{2}(1+\lambda)} .$$

Using the multiplicity bound and the bound on spacetime volume we get

$$|I| \leq C_2 A^{\frac{1}{2}(1+\lambda)} ,$$

for $C_2 = C_2(\lambda, T)$. Since by (6.4) we can cover $\mathcal{T}_A$ with at most $|I|$ intervals of length comparable to A^{-1}, this implies that $\cap_{A>0} \mathcal{T}_A$ has Minkowski dimension $\leq \frac{1}{2} + \frac{\lambda}{2}$. As λ is arbitrary, this proves the theorem. □

Acknowledgment

Bruce Kleiner was supported by NSF grants DMS-1405899 and DMS-1711556, and a Simons Collaboration grant. John Lott was supported by NSF grants DMS-1510192 and DMS-1810700.

References

[BKa] R. Bamler and B. Kleiner, *Ricci flow and diffeomorphism groups of 3-manifolds*, http://lanl.arxiv.org/abs/1712.06197.

[BKb] ———, *Uniqueness and stability of Ricci flow through singularities*, http://lanl.arxiv.org/abs/1709.04122.

[CLN06] B. Chow, P. Lu, and L. Ni, *Hamilton's Ricci flow*, Graduate Studies in Mathematics, vol. 77, American Mathematical Society, Providence, RI; Science Press, New York, 2006.

[KL08] B. Kleiner and J. Lott, *Notes on Perelman's papers*, Geom. Topol. **12** (2008), no. 5, 2587–2855.

[KL17] ———, *Singular Ricci flows I*, Acta Math. (2017), no. 219, 65–134.

[Per02] G. Perelman, *The entropy formula for the Ricci flow and its geometric applications*, http://arxiv.org/abs/math/0211159, 2002.

[Per03] ———, *Ricci flow with surgery on three-manifolds*, http://arxiv.org/abs/math/0303109, 2003.

Bruce Kleiner
Courant Institute
of Mathematical Sciences
251 Mercer St.
New York, NY 10012, USA
e-mail: bkleiner@cims.nyu.edu

John Lott
Department of Mathematics
University of California at Berkeley
Berkeley, CA 94720, USA
e-mail: lott@berkeley.edu

Progress in Mathematics, Vol. 333, 157–166

An Inequality Between Complex Hessian Measures of Hölder Continuous m-subharmonic Functions and Capacity

Sławomir Kołodziej and Ngoc Cuong Nguyen

Dedicated to Professor Gang Tian on the occasion of his 60th birthday

Abstract. We show that the complex m-Hessian operator of a Hölder continuous m-subharmonic function is well dominated by the corresponding capacity. As a consequence we obtain the Hölder continuous subsolution theorem for the complex m-Hessian equation.

Mathematics Subject Classification (2010). 53C55, 35J96, 32U40.

Keywords. Dirichlet problem, weak solutions, Hölder continuous, complex Hesssian equation, subsolution problem.

1. Introduction

Consider Ω an open subset of $\mathbb{C}^n$ and a real-valued function $u \in C^2(\Omega)$. Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ be the set of eigenvalues of the complex Hessian of u at a given point $z \in \Omega$. By $S_m(u(z))$ we denote the mth elementary symmetric function of λ:

$$S_m(u(z)) = \sum_{0<j_1<\cdots<j_m\leq n} \lambda_{j_1}\,\lambda_{j_2}, \ldots, \lambda_{j_m}\,.$$

The m-Hessian equation for unknown function u is

$$S_m(u(z)) = f(z),$$

where f is a given function. The equation can be rewritten in terms of differential forms. For $d = \partial + \bar\partial,\ d^c = i(\bar\partial - \partial)$ and $\beta = dd^c\|z\|^2$ we have

$$(dd^c u)^m \wedge \beta^{n-m} = \frac{m!(n-m)!}{n!} S_m(u(z))\beta^n.$$

We shall study the m-Hessian equation in the form

$$(dd^c u)^m \wedge \beta^{n-m} = f\beta^n.$$

This allows to consider weak solutions of the equation when the equation is understood in the sense of currents.

We call a C^2 function $u : \Omega \to \mathbb{R} \cup \{-\infty\}$ *m-subharmonic* (*m-sh*) if the forms

$$(dd^c u)^k \wedge \beta^{n-k}$$

are positive for $k = 1, \ldots, m$ (in particular u is subharmonic). If u is subharmonic but not smooth then one can define m-sh function by $(dd^c u)^k \wedge \beta^{n-k} \geq 0$ in the sense of currents (see [3]). $SH_m(\Omega)$ stands for the set of m-subharmonic functions in Ω. Throughout the paper we assume $m < n$, since for $m = n$ (the Monge–Ampère equation) the results are known and some of them sharper.

Unlike the real Hessian equations the complex m-Hessian equation has not been studied until recently. In 2005, S.Y. Li [11] proved that if Ω is smoothly bounded and strictly m-pseudoconvex (as defined in the next section) then, given a smooth boundary data and a smooth, positive function f there exists a unique smooth m-sh solution of the Dirichlet problem for the m-Hessian equation. The proof is inspired by the one in [4]. In the same year Z. Błocki published [3], where he considered weak solutions of the equation, for possibly vanishing right-hand side. He proved that the m-sh function u is maximal in this class if and only if

$$(dd^c u)^m \wedge \beta^{n-m} = 0.$$

He also described the maximal domain of definition of the m-Hessian operator. This pluripotential approach was also taken by S. Dinew and the first author in [7], where the Dirichlet problem in smoothly bounded m-pseudoconvex domains was solved for a continuous boundary data and $f \in L^p(\Omega)$ for $p > n/m$. The bound for the exponent p is sharp and related to the volume-capacity estimate [7] saying that for any $0 < \alpha < \frac{n}{n-m}$ there exists $C_\alpha > 0$ such that for all compact $K \subset \Omega$

$$V_{2n}(K) \leq C_\alpha [\mathrm{cap}_m(K)]^\alpha, \tag{1.1}$$

where V_{2n} denotes the Euclidean volume and the m-capacity is defined by

$$\mathrm{cap}_m(K) = \sup \left\{ \int_K (dd^c w)^m \wedge \beta^{n-m} : w \in SH_m(\Omega), 0 \leq w \leq 1 \right\}.$$

The weak solution theory in the Cegrell classes (unbounded functions) for a more general right-hand side has been developed by H.-C. Lu [13, 15] (see [12] [9], [20] for recent results). On the other hand the second author [16] proved that the Dirichlet problem described above has a bounded solution if a bounded subsolution exists. In the present paper we provide an analogue of this result in the category of Hölder continuous solutions.

Our first result says that if μ is compactly supported in Ω and upper bounded by $(dd^c u)^m \wedge \beta^{n-m}$, where $u \in SH_m(\Omega)$ is Hölder continuous then the inequality (1.1) remains true for some $\alpha > 1$ if we replace the Euclidean volume V_{2n} by μ.

Having this we can adopt the method of [18] and [19], which deal with the Monge–Ampère equation, and show the existence of continuous solutions of the Dirichlet problem for such measures. Furthermore, for Hölder continuous boundary data the solutions are actually also Hölder continuous. Let us give the precise statements.

Theorem 1.1. *Suppose that μ is a positive Borel, compactly supported, measure in Ω such that*

$$\mu \leq (dd^c\varphi)^m \wedge \beta^{n-m} \tag{1.2}$$

for some $\varphi \in SH_m(\Omega) \cap C^\alpha(\bar{\Omega})$ with $0 < \alpha \leq 1$. Then, there exist uniform constants $C, \alpha_0 > 0$ depending only on φ, Ω, such that for every compact set $K \subset \Omega$,

$$\mu(K) \leq C[\mathrm{cap}_m(K)]^{1+\alpha_0}. \tag{1.3}$$

Let us fix the function φ in Theorem 1.1 throughout the paper. Denote by $\mathcal{S}$ the set of all positive Borel measures with compact support in Ω which satisfy the inequality (1.3). For $\mu \in \mathcal{S}$ we consider the Dirichlet problem for the complex m-Hessian equation. Let ψ be a continuous function on the boundary of Ω. We seek for the solution to the equation.

$$\begin{aligned} &u \in SH_m(\Omega) \cap C^0(\bar{\Omega}), \\ &(dd^c u)^m \wedge \beta^{n-m} = \mu, \\ &u = \psi \quad \text{on } \partial\Omega. \end{aligned} \tag{1.4}$$

We obtain the following subsolution theorem for this problem.

Theorem 1.2. *Let Ω be a bounded strictly m-pseudoconvex domain. Then for $\mu \in \mathcal{S}$, we have the following:*

(a) *There exists a unique solution to the Dirichlet problem* (1.4)*.*
(b) *If ψ is Hölder continuous on $\partial\Omega$, then the solution is also Hölder continuous on $\bar{\Omega}$.*

The modulus of continuity and Hölder continuity of solutions to the complex m-Hessian equation with the right-hand side $\mu = f dV_{2n}$ has been investigated by Charabati [6] (see also [14], [17]) for a continuous density and a L^p density in Ω, respectively. The following L^p-property result can be considered as a generalization. The proof follows from a simple application of the Hölder inequality and Theorem 1.2.

Corollary 1.3. *Let $\mu \in \mathcal{S}$ and α_0 be the constant in Theorem* 1.1*. Assume that $f \in L^p(\Omega, d\mu)$ with $p > \frac{1}{\alpha_0}$. Then, $f d\mu$ is the complex m-Hessian operator of a Hölder continuous m-subharmonic function in Ω.*

2. Preliminaries

In this section we collect several results needed in our proofs. They are mostly directly adapted from the case of $m = n$. Where the proofs follow precisely the ones for the Monge–Ampère equation we just give reference to the latter.

Let Ω be a strictly m-pseudoconvex domain. Then, by definition, there exists $\rho \in C^2(\bar{\Omega})$ a defining function for Ω, strictly m-subharmonic and such that

$$\Omega = \{\rho < 0\}, \quad \nabla\rho \neq 0 \quad \text{on } \partial\Omega. \tag{2.1}$$

Since it is strictly m-subharmonic, we can assume (after multiplying by a big constant) that for any integer $1 \leq k \leq m$

$$(dd^c\rho)^k \wedge \beta^{n-k} \geq \beta^n. \tag{2.2}$$

Let us first make a reduction which will be used in what follows. Since Ω is strictly m-pseudoconvex and μ is compactly supported in Ω we can modify φ near the boundary so that $\varphi = 0$ on the boundary and the total mass of its complex m-Hessian operator is finite (see, e.g., [18, Corollary 2.13]). Therefore, from now on we have

$$\varphi_{|\partial\Omega} = 0 \quad \text{and} \quad \int_\Omega (dd^c\varphi)^m \wedge \beta^{n-m} < +\infty. \tag{2.3}$$

We use the notation

$$\|\cdot\|_\infty := \sup_\Omega |\cdot|, \tag{2.4}$$

and the convention that the constants $C > 0$ which appear below are uniform constants. They may differ from place to place.

Lemma 2.1 (Błocki [2]). *Let $v_1, \ldots, v_m, v, h \in SH_m \cap L^\infty(\Omega)$ be such that $v_i \leq 0$ for $i = 1, \ldots, m$, and $v \leq h$. Assume that $\lim_{z\to\partial\Omega}[h(z) - v(z)] = 0$. Then, for any integer $1 \leq k \leq m$,*

$$\begin{aligned} &\int_\Omega (h-v)^k dd^c v_1 \wedge \cdots \wedge dd^c v_m \wedge \beta^{n-m} \\ &\leq k!\|v_1\|_\infty \cdots \|v_k\|_\infty \int_\Omega (dd^c v)^k \wedge dd^c v_{k+1} \wedge \cdots \wedge dd^c v_m \wedge \beta^{n-m}. \end{aligned} \tag{2.5}$$

As a consequence one obtains a bound for the Laplacian of a bounded m-subharmonic function. This kind of estimate was first given in [1].

Corollary 2.2. *Let $v \in SH_m \cap L^\infty(\Omega)$. Then, for every small $\delta > 0$,*

$$\int_{\Omega_\delta} dd^c v \wedge \beta^{n-1} \leq \frac{C\|v\|_\infty}{\delta}, \tag{2.6}$$

where $\Omega_\delta = \{z \in \Omega : \operatorname{dist}(z, \partial\Omega > \delta)\}$.

Proof. Since $\|v\|_\infty \leq \sup_\Omega v - \inf_\Omega v \leq 2\|v\|_\infty$, we may assume that $v \leq 0$ in Ω. Observe that for a defining function there exists a uniform constant $c_0 > 0$ such that

$$|\rho(z)| \geq c_0 \operatorname{dist}(z, \partial\Omega). \tag{2.7}$$

Therefore, $\Omega_\delta \subset \{\rho(z) < -c_0\delta\}$. It follows that

$$\begin{aligned}\int_{\Omega_\delta} dd^c v \wedge \beta^{n-1} &\leq \frac{2}{c_0\delta}\int_\Omega \left(\max\{\rho, -\frac{c_0\delta}{2}\} - \rho\right) dd^c v \wedge \beta^{n-1} \\ &\leq \frac{C\|v\|_\infty}{\delta}\int_\Omega dd^c\rho \wedge \beta^{n-1},\end{aligned} \tag{2.8}$$

where we used Lemma 2.1 for the second inequality. The last integral is bounded as $\rho \in C^2(\bar{\Omega})$. □

Let us recall the definition of the Cegrell classes for complex m-Hessian operators:

$$\mathcal{E}_0(m) := \left\{ v \in SH_m \cap L^\infty(\Omega) : \begin{array}{l} \lim_{z\to\partial\Omega} v(z) = 0, \\ \int_\Omega (dd^c v)^m \wedge \beta^{n-m} < +\infty \end{array} \right\}. \tag{2.9}$$

We refer the reader to [13] for more information on this class and its generalization. The Cegrell inequality in this class reads:

Lemma 2.3 (Cegrell [5]). *Let $v_1, \ldots, v_n \in \mathcal{E}_0(m)$. Then,*

$$\begin{aligned}&\int_\Omega dd^c v_1 \wedge \cdots \wedge dd^c v_m \wedge \beta^{n-m} \\ &\quad\leq \left(\int_\Omega (dd^c v_1)^m \wedge \beta^{n-m}\right)^{\frac{1}{m}} \cdots \left(\int_\Omega (dd^c v_n)^m \wedge \beta^{n-m}\right)^{\frac{1}{m}}.\end{aligned} \tag{2.10}$$

We also consider a subclass

$$\mathcal{E}_0'(m) := \left\{ v \in \mathcal{E}_0(m) : \int_\Omega (dd^c v)^m \wedge \beta^{n-m} \leq 1 \right\}. \tag{2.11}$$

Lemma 2.4. *Let $v, w \in \mathcal{E}_0'(m)$ be such that $v = w$ near $\partial\Omega$. Then, there exists a uniform constant $0 < \alpha_1 \leq 1$ such that*

$$\int_\Omega |v-w|(dd^c\varphi)^m \wedge \beta^{n-m} \leq C\left(\int_\Omega |v-w| dV_{2n}\right)^{\alpha_1}. \tag{2.12}$$

In particular, for $\mu \in \mathcal{S}$,

$$\mu(|v-w|) := \int_\Omega |v-w| d\mu \leq C\|v-w\|_{L^1}^{\alpha_1}. \tag{2.13}$$

Here, the constants C and α_1 are independent of v, w.

Proof. See Lemma 2.7 in [18]. Notice that the additional assumption $v = w$ near the boundary $\partial\Omega$ allows to avoid the use the notion of the boundary measure for complex m-Hessian operators. The Cegrell inequality (Lemma 2.3) is sufficient for our need here. □

We can follow the proofs in [8] to derive the stability estimate for the measures which are dominated by capacity.

Proposition 2.5. *Suppose that* $\mu = (dd^c u)^m \wedge \beta^{n-m}$ *for* $u \in SH_m \cap L^\infty(\Omega)$ *and* μ *satisfies the inequality* (1.3). *Let* $v \in SH_m \cap L^\infty(\Omega)$ *be such that* $\liminf_{z \to \partial\Omega}(u - v) \geq 0$. *Then, there exist constants* $C > 0$ *and* $0 < \alpha_2 < 1$ *independent of* v *such that*

$$\sup_{\Omega}(v-u) \leq C\left(\int_{\Omega} \max\{v-u, 0\} d\mu\right)^{\alpha_2}. \tag{2.14}$$

Proof. It readily follows from the one of [8, Theorem 1.1] with obvious adjustments. □

3. The Dirichlet problem

In this section we prove Theorem 1.1 and then we study the Dirichlet problem. The existence of a continuous solution is proved by applying the stability estimate (see [7]). Next, we prove that this solution is Hölder continuous via the strategy from [8].

We will prove a slightly stronger statement than the one in Theorem 1.1. Since φ is uniformly continuous on $\bar{\Omega}$, it admits a modulus of continuity $\varpi : [0, \infty] \to [0, \infty]$ – an increasing function defined by

$$\varpi(t) := \sup\left\{|\varphi(z) - \varphi(w)| : z, w \in \bar{\Omega}, \quad |z-w| \leq t\right\}. \tag{3.1}$$

Lemma 3.1. *Fix* $1 < p < n/(n-m)$ *and* $0 < \tau < (p-1)/2m$. *There exists a uniform constant* C *such that for every compact set* $K \subset \Omega$,

$$\mu(K) \leq C\left\{\varpi\left([\mathrm{cap}(K)]^\tau\right) + [\mathrm{cap}_m(K)]^{p-1-2m\tau}\right\} \cdot \mathrm{cap}_m(K), \tag{3.2}$$

where $\mathrm{cap}_m(K) := \mathrm{cap}_m(K, \Omega)$.

Proof. Fix a compact subset $K \subset\subset \Omega$. Without loss of generality we may assume that K is regular (in the sense that the relative extremal function u_K is continuous) as μ is a Radon measure. Denote by φ_ε the standard regularization of φ defined via the convolution with smooth, symmetric mollifier. We choose $\varepsilon > 0$ so small that

$$\mathrm{supp}\ \mu \subset \Omega'' \subset\subset \Omega' \subset \Omega_\varepsilon \subset \Omega, \tag{3.3}$$

where $\Omega_\varepsilon = \{z \in \Omega : \mathrm{dist}(z, \partial\Omega) > \varepsilon\}$. Since for every $K \subset \Omega''$ we have

$$\mathrm{cap}_m(K, \Omega') \sim \mathrm{cap}_m(K, \Omega) \tag{3.4}$$

in what follows we will write $\mathrm{cap}_m(K)$ for either one of these capacities. We have

$$0 \leq \varphi_\varepsilon - \varphi \leq \varpi(\varepsilon) := \delta \quad \text{on } \Omega'. \tag{3.5}$$

Let u_K be the relative extremal function for K with respect to Ω'. Consider the set $K' = \{3\delta u_K + \varphi_\varepsilon < \varphi - 2\delta\}$. Then,

$$K \subset K' \subset \left\{u_K < -\frac{1}{2}\right\} \subset \Omega'. \tag{3.6}$$

Then, it follows from [7, Proposition 2.2] and [13, Théorème 1.4.12] (see also [21]) that

$$\begin{aligned}\operatorname{cap}_m(K') &\le \operatorname{cap}_m\left(\{u_K < -1/2\}\right)\\ &\le 2^m \int_{\{u_K<0\}} (dd^c u_K)^m \wedge \beta^{n-m} \\ &= 2^m \operatorname{cap}_m(K).\end{aligned}\tag{3.7}$$

Note that

$$dd^c \varphi_\varepsilon \le \frac{C\beta}{\varepsilon^2}, \quad \|\varphi_\varepsilon + u_K\|_\infty =: M \le \|\varphi\|_\infty + 1 \tag{3.8}$$

(where the first inequality follows by differentiating twice the mollifier in the convolution defining φ_ε). Let us also use the short notation $H_m(\cdot)$ for the complex m-Hessian operators. The comparison principle [7, Theorem 1.4] and above estimates give us that

$$\begin{aligned}\int_{K'} H_m(\varphi) &\le \int_{K'} H_m(3\delta u_K + \varphi_\varepsilon)\\ &\le \int_{K'} 3\delta H_m(u_K + \varphi_\varepsilon) + \int_{K'} H_m(\varphi_\varepsilon)\\ &\le 3\delta M^m \operatorname{cap}_m(K') + C\varepsilon^{-2m}\left[\operatorname{cap}_m(K')\right]^p,\end{aligned}\tag{3.9}$$

where in the last inequality we used the volume-capacity inequality (1.1).

We may assume that ε is so small that (3.6) is satisfied, otherwise the inequality (3.2) holds true by increasing the constant. Then, choose $\varepsilon := [\operatorname{cap}_m(K')]^\tau$ and use the definition of δ to get

$$\begin{aligned}\mu(K) &\le \int_{K'} H_m(\varphi)\\ &\le 3M^m \varpi\left([\operatorname{cap}_m(K')]^\tau\right) \operatorname{cap}_m(K')\\ &\quad + C\left[\operatorname{cap}_m(K')\right]^{p-2m\tau}.\end{aligned}\tag{3.10}$$

This combined with (3.7) proves the desired inequality. $\square$

Now we are going to conclude the proof of the first theorem. It says that under the assumption $\varphi \in C^{0,\alpha}(\bar{\Omega})$, $0 < \alpha \le 1$, there exist uniform constants $C, \alpha_0 > 0$ such that

$$\mu(K) \le C\left[\operatorname{cap}_m(K)\right]^{1+\alpha_0}. \tag{3.11}$$

Proof of Theorem 1.1. We choose $\tau > 0$ in Lemma 3.1 such that

$$\tau\alpha = p - 1 - 2m\tau \Leftrightarrow \tau = \frac{p-1}{2m+\alpha}. \tag{3.12}$$

Thus, the theorem follows with $\alpha_0 = \frac{\alpha(p-1)}{2m+\alpha}$ as $\varpi(t) \le Ct^\alpha$ for $t \ge 0$. $\square$

Thanks to the subsolution theorem in [16], which was inspired by [10], if $\mu \in \mathcal{S}$, then there exists a unique $u \in SH_m \cap L^\infty(\Omega)$ solving

$$(dd^c u)^m \wedge \beta^{n-m} = \mu, \quad \lim_{\zeta\to z} u(\zeta) = \psi(z) \quad \forall z \in \partial\Omega. \tag{3.13}$$

Thus to prove the second statement we need to show the continuity and Hölder continuity of this solution.

Proof of Theorem 1.2. Extend ψ onto Ω so that $\psi \in SH_m(\Omega) \cap C^0(\bar{\Omega})$. By the comparison principle
$$\varphi + \psi \leq u \leq \tilde{\psi}, \tag{3.14}$$
where $\tilde{\psi}$ is the harmonic extension of ψ onto Ω. This implies that u is continuous on the boundary $\partial\Omega$. It remains to show that u is continuous in Ω. Take $u_j := u * \chi_j(z)$, where $\chi_j(z) := \chi(|z|/j)/j^{2n}$ and $\chi(z)$ is the standard smoothing kernel. Then, u_j is defined on
$$\Omega_j := \left\{ z \in \Omega : \mathrm{dist}(z, \partial\Omega) > \frac{1}{j} \right\}. \tag{3.15}$$
Since u is continuous on $\partial\Omega$, it follows that
$$\sup_{\partial\Omega_j} |u_j - u| = \varepsilon_j \to 0 \quad \text{as } j \to +\infty. \tag{3.16}$$
Applying Proposition 2.5 on Ω_j we get that
$$\sup_{\Omega_j} (u_j - u) \leq \varepsilon_j + C \left(\int_{\Omega_j} \max\{u_j - u - \varepsilon_j, 0\} d\mu \right)^{\alpha_2}. \tag{3.17}$$
The Lebesgue dominated convergence theorem implies that u_j converges uniformly to u on every compact subset of Ω. Thus, $u \in C^0(\Omega)$. Combining this with the fact that $u \in C^0(\partial\Omega)$ we get the first part.

To prove the second part we follow the steps of [8]. Define for $\delta > 0$ small
$$\Omega_\delta := \{ z \in \Omega : \mathrm{dist}(z, \partial\Omega) > \delta \}; \tag{3.18}$$
and for $z \in \Omega_\delta$ we define
$$u_\delta(z) := \sup_{|\zeta| \leq \delta} u(z + \zeta), \tag{3.19}$$
$$\hat{u}_\delta(z) := \frac{1}{\sigma_{2n} \delta^{2n}} \int_{|\zeta| \leq \delta} u(z + \zeta) dV_{2n}(\zeta), \tag{3.20}$$
where σ_{2n} is the volume of the unit ball.

Using the Hölder continuity of ψ on the boundary we get that u is Hölder continuous on the boundary. Next, with the aid of Proposition 2.5, Lemma 2.4 and the Laplacian bound (Corollary 2.2) as in the proof of [18, Theorem 2.5] we have
$$\sup_{\Omega_\delta} (\hat{u}_\delta - u) \leq C \delta^{\alpha_3}. \tag{3.21}$$
This is the desired inequality. Thus, the Hölder continuity of u follows. □

Acknowledgement

The first author was partially supported by NCN grant 2013/ 08/A/ST1/00312. The second author was supported by the NRF Grant 2011-0030044 (SRC-GAIA) of The Republic of Korea. He also would like to thank Kang-Tae Kim for encouragement and support.

References

[1] L. Baracco, T.-V. Khanh, S. Pinton, G. Zampieri, *Hölder regularity of the solution to the complex Monge–Ampère equation with L^p density*, Calc. Var. Partial Differential Equations **55** (2016) no. 4. art 74, 8pp.

[2] Z. Błocki, *Estimates for the complex Monge–Ampère operator*. Bull. Polish Acad. Sci. Math. **41** (1993), no. 2, 151–157.

[3] Z. Błocki, *Weak solutions to the complex Hessian equation*, Ann. Inst. Fourier (Grenoble) **55** (2005), 1735–1756.

[4] L. Caffarelli, L. Nirenberg, J. Spruck, *The Dirichlet problem for nonlinear second order elliptic equations, III: Functions of the eigenvalues of the Hessian*, Acta Math. **155** (1985), 261–301.

[5] U. Cegrell,*The general definition of the complex Monge–Ampère operator*, Ann. Inst. Fourier (Grenoble), **54** (2004), 159–179.

[6] M. Charabati, *Modulus of continuity of solutions to complex Hessian equations*, Internat. J. Math. **27** (2016), no. 1.

[7] S. Dinew, S. Kołodziej, *A priori estimates for complex Hessian equations*, Anal. PDE **7** (2014), no. 1, 227–244.

[8] V. Guedj, S. Kołodziej, A. Zeriahi, *Hölder continuous solutions to Monge–Ampère equations*, Bull. Lond. Math. Soc. **40** (2008), 1070–1080.

[9] V.V. Hung and N. Van Phu, Hessian measures on m-polar sets and applications to the complex Hessian equations, Complex Var. Elliptic Equ. **62** (2017), no. 8, 1135–1164.

[10] S. Kołodziej, *The range of the complex Monge–Ampère operator. II*, Indiana Univ. Math. J.**44** (1995) no. 3, 765–782.

[11] S.-Y. Li, *On the Dirichlet problems for symmetric function equations of the eigenvalues of the complex Hessian*, Asian J. Math. **8** (2004), 87–106.

[12] H.M. Le, H.X. Nguyen and H.V. Vu, Some characterizations of the class $\mathcal{E}_m(\Omega)$ and applications, Ann. Polon. Math. **115** (2015), no. 2, 145–163.

[13] C.H. Lu, Équation hessiennes complexes, Thèse Doctorat. (2012) Toulouse.

[14] C.H. Lu, Viscosity solutions to complex Hessian equations, J. Funct. Anal. **264** (2013), no. 6, 1355–1379.

[15] C.H. Lu, A variational approach to complex Hessian equations in $\mathbb{C}^n$, J. Math. Anal. Appl. **431** (2015), no. 1, 228–259.

[16] N.-C. Nguyen, *Subsolution theorem for the complex Hessian equation*, Univ. Iagiell. Acta Math. **50** (2012), 69–88.

[17] N.-C. Nguyen, *Hölder continuous solutions to complex Hessian equations*, Potential Anal. **41** (2014), no. 3, 887–902.

[18] N.-C. Nguyen, *On the Hölder continuous subsolution problem for the complex Monge–Ampère equation*, Calc. Var. Partial Differential Equations **57** (2018), no. 1, Art. 8, 15 pp.

[19] N.-C. Nguyen, *On the Hölder continuous subsolution problem for the complex Monge–Ampère equation, II.* preprint arXiv:1803.02510

[20] V.-T. Nguyen, *A characterization of the Cegrell classes and generalized m-capacities*, Ann. Polon. Math, DOI: 10.4064/ap170728-26-1.

[21] A.S. Sadullaev, B.I. Abdullaev, *Potential theory in the class of m-subharmonic functions*, Tr. Mat. Inst. Steklova **279** (2012) 166–192.

Sławomir Kołodziej
Faculty of Mathematics and Computer Science
Jagiellonian University
Łojasiewicza 6
PL-30-348 Kraków, Poland
e-mail: Slawomir.Kolodziej@im.uj.edu.pl

Ngoc Cuong Nguyen
Faculty of Mathematics and Computer Science
Jagiellonian University
Łojasiewicza 6
PL-30-348 Kraków, Poland
and
Department of Mathematics
Center for Geometry and its Applications
Pohang University of Science and Technology
37673 Nam-gu Pohang-si, The Republic of Korea
Current address
Department of Mathematical Sciences, KAIST
291 Daehak-ro, Yuseong-gu, Daejeon 34141, South Korea
e-mail: Nguyen.Ngoc.Cuong@im.uj.edu.pl
cuongnn@kaist.ac.kr

Progress in Mathematics, Vol. 333, 167–219

A Guided Tour to Normalized Volume

Chi Li, Yuchen Liu and Chenyang Xu

Dedicated to Gang Tian's Sixtieth Birthday with admiration

Abstract. This is a survey on the recent theory on minimizing the normalized volume function attached to any klt singularities.

Mathematics Subject Classification (2010). Primary 14B05; Secondary 32S05.

Keywords. Local volumes, K-stability, stable degeneration.

1. Introduction

The development of algebraic geometry and complex geometry has interwoven in the history. One recent example is the interaction between the theory of higher-dimensional geometry centered around the minimal model program (MMP), and the existence of 'good' metrics on algebraic varieties. Both subjects have major steps forward, whose influences are beyond the subjects themselves, spurring out new progress in topics once people could not imagine. In this note, we will discuss a 'local stability theory' of singularities, which in our opinion provides an excellent example on the philosophy that there are many unexpected connections underlying these two different topics.

Ever since the starting of the theory of MMP in higher dimensions (that is, the dimension is at least three), people understand that a feature of such a theory is that we need to deal with singular varieties. Then it becomes very nature to investigate this class of singularities for people working on the MMP. To deal with singular varieties in complex geometry is a more recent trend, and it significantly improves people's knowledge on the existence of interesting metrics, even in situations which people originally only want to study smooth varieties.

It becomes clear now, Kawamata log terminal (klt) singularities form an exceptionally important class of singularities for many reasons: it is the natural class of singular varieties for people to inductively prove deep results in the MMP; it is the class of singularities appearing on degenerations in many natural settings and it carries properties which globally Fano varieties have.

What we want to survey here is a rather new theory on klt singularities. The picture consists of two closely related parts: firstly, we want to establish a structure which provides a canonically determined degeneration to a stable log Fano cone from each klt singularity; secondly, to construct the degeneration, we need a valuation which minimizes the normalized volume function on the 'non-archimedean link', and since such minimum is a deep invariant defined for all klt singularities, we want to explore more properties of this invariant, including calculating it in many cases.

1.1. History

The first prototype of the local stability theory underlies in [MSY06, MSY08]. They find that the existence of Ricci-flat cone metric on an affine variety with a good action by a torus group T is closely related to the normalized volume minimizing problem. In our language, they concentrate on the valuations induced by the vectors in the Reeb cone provided by the torus action. Later a systematic study of K-stability in the setting of Sasaki geometry is further explored in [CS18, CS15].

Consider klt singularities which appear on the Gromov–Hausdorff (GH) limit of Kähler–Einstein Fano manifolds. At the first sight, we do not know more algebraic structure for these singularities. Nevertheless, by looking at the metric tangent cone, it is shown in [DS17], built on the earlier works in [CCT02, DS14, Tia13], that the metric tangent cone of such singularities is an affine T-variety with a Ricci-flat cone metric. Furthermore, [DS17] gives a two-step degeneration description of the metric tangent cone. They further conjecture that this two-step degeneration should only depend on the algebraic structure of the singularity, but not the metric.

Then in [Li18a], the normalized volume function on the 'non-archimedean link' of a given klt singularity is defined, and a series of conjectures on normalized volume function are proposed. This attempt is not only to algebrize the work in [DS17] without invoking the metric, but it is also of a completely local nature. Since then, the investigation on this local stability theory points to different directions.

In [Blu18], the existence of a minimum (opposed to only infimum) which was conjectured in [Li18a] is affirmatively answered. The proof uses the properness estimates in [Li18a] and the observation in [Liu18] that the minimizer can be computed by the minimal normalized multiplicities, and then skillfully uses the techniques from the study of asymptotic invariants (see [Laz04]). Later in [BL18], lower semicontinuity of the volume of singularities are also established using this circle of ideas.

In [Li17, LL19], the case of a cone singularity over a Fano variety is intensively studied, and it was found if we translate the minimizing question for the canonical valuation into a question on the base Fano varieties, what appears is the sign of the β-invariant developed in [Fuj18, Fuj16, Li17].

Built on the previous study of cone singularities, implementing the ideas circled around the MMP in birational geometry, an effective process of degenerating a general singularity to a cone singularity is established in [LX16], provided the minimizer is a divisorial valuation. In [LX17a], a couple of conjectural properties are added to complete the picture proposed in [Li18a], and now the package is called '*stable degeneration conjecture*', see Conjecture 4.1. The investigation in [LX16] is also extended in [LX17a] to the case when the minimizer is a quasi-monomial valuation with a possibly higher rational rank, where the study involves a considerable amount of new techniques. As a corollary, the first part of Donaldson–Sun's conjecture in [DS17] is answered affirmatively in [LX17a]. Later the work is extended in [LWX18] and a complete solution of Donaldson–Sun's conjecture is found.

Applications to global questions, especially the existence of KE metrics on Fano varieties, are also explored. In [Liu18], built on the work of [Fuj18], an inequality to connect the local volume and the global one is proved. Then in [SS17, LX17b], via the approach of the 'comparison of moduli', complete moduli spaces parametrizing explicit Fano varieties with a KE metric are established by studying the local constraint posted by the lower bound of the local volumes.

1.2. Outline

In the note, we will survey a large part of the results mentioned above. From the perspective of techniques, there are three closely related ways to think about the volume of a singularity: the infima of the normalized volume of valuations, of the normalized multiplicity of primary ideals or of the volume of models. The viewpoint using valuations gives the most canonical picture, e.g., the stable degeneration conjecture, but there are less techniques available to directly study the space. The viewpoint using ideals is flexible for many purposes, e.g., taking degenerations. Moreover, though usually working on a single ideal does not give too much advantage over others, working on a graded sequence of ideals really enables one to use the powerful theory on asymptotical invariants for such setting. The third viewpoint of using models allows us to apply the machinery from the MMP theory, and it is the key to degenerate the underlying singularities into cone singularities. The interplay among these three circle of techniques is fruitful, and we expect further insight can be made in the future.

In Section 2, we give the definition of the function of the normalized volumes and sketch the basic properties of its minimizer, including the existence. In Section 3, we discuss the theory on searching for Sasaki–Einstein metrics on a Fano cone singularities. The algebraic side, namely the K-stability notions on a Fano cone plays an important role as we try to degenerate any klt singularity to a K-semistable Fano cone. Such an attempt is formulated in the stable degeneration conjecture, which is the focus of Section 4. In Section 5, we present some applications, including the torus equivariant K-stability (Section 5.1), a solution of Donaldson–Sun's conjecture (Section 5.2) and the K-stability of cubic threefolds (Section 5.3). In the last Section 6, we discuss many unsolved questions, which we hope will lead to some future research. Some of them give new approaches to attack existing problems.

2. Definitions and first properties

2.1. Definitions

In this section, we give the definition of the normalized volume $\widehat{\mathrm{vol}}_{(X,D),x}(v)$ (or abbreviated as $\widehat{\mathrm{vol}}(v)$ if there is no confusion) for a valuation v centered on a klt singularity $x \in (X, D)$ as in [Li18a]. It consists of two parts: the volume $\mathrm{vol}(v)$ (see Definition 2.1) and the log discrepancy $A_{X,D}(v)$ (see Definition 2.2).

Let X be a reduced, irreducible variety defined over $\mathbb{C}$. A *real valuation* of its function field $K(X)$ is a non-constant map $v\colon K(X)^{\times} \to \mathbb{R}$, satisfying:

- $v(fg) = v(f) + v(g)$;
- $v(f + g) \geq \min\{v(f), v(g)\}$;
- $v(\mathbb{C}^*) = 0$.

We set $v(0) = +\infty$. A valuation v gives rise to a valuation ring $\mathcal{O}_v := \{f \in K(X) \mid v(f) \geq 0\}$. We say a real valuation v is *centered at* a scheme-theoretic point $\xi = c_X(v) \in X$ if we have a local inclusion $\mathcal{O}_{\xi,X} \hookrightarrow \mathcal{O}_v$ of local rings. Notice that the center of a valuation, if exists, is unique since X is separated. Denote by Val_X the set of real valuations of $K(X)$ that admits a center on X. For a closed point $x \in X$, we denote by $\mathrm{Val}_{X,x}$ the set of real valuations of $K(X)$ centered at $x \in X$. It is well known that $v \in \mathrm{Val}_X$ is centered at $x \in X$ if $v(f)$ for any $f \in \mathfrak{m}_x$.

For each valuation $v \in \mathrm{Val}_{X,x}$ and any integer m, we define the valuation ideal $\mathfrak{a}_m(v) := \{f \in \mathcal{O}_{x,X} \mid v(f) \geq m\}$. Then it is clear that $\mathfrak{a}_m(v)$ is an $\mathfrak{m}_x$-primary ideal for each $m > 0$.

Given a valuation $v \in \mathrm{Val}_X$ and a nonzero ideal $\mathfrak{a} \subset \mathcal{O}_X$, we may evaluate $\mathfrak{a}$ along v by setting $v(\mathfrak{a}) := \min\{v(f) \mid f \in \mathfrak{a} \cdot \mathcal{O}_{c_X(v),X}\}$. It follows from the above definition that if $\mathfrak{a} \subset \mathfrak{b} \subset \mathcal{O}_X$ are nonzero ideals, then $v(\mathfrak{a}) \geq v(\mathfrak{b})$. Additionally, $v(\mathfrak{a}) > 0$ if and only if $c_X(v) \in \mathrm{Cosupp}(\mathfrak{a})$. We endow Val_X with the weakest topology such that, for every ideal $\mathfrak{a}$ on X, the map $\mathrm{Val}_X \to \mathbb{R} \cup \{+\infty\}$ defined by $v \mapsto v(\mathfrak{a})$ is continuous. The subset $\mathrm{Val}_{X,x} \subset \mathrm{Val}_X$ is endowed with the subspace topology. In some literatures, the space $\mathrm{Val}_{X,x}$ is called the *non-Archimedean link* of $x \in X$. When $X = \mathbb{C}^2$, the geometry of $\mathrm{Val}_{X,x}$ is understood well (see [FJ04]). For higher dimension, its structure is much more complicated but can be described as an inverse limit of dual complexes (see [JM12, BdFFU15]).

Let $Y \xrightarrow{\mu} X$ be a proper birational morphism with Y a normal variety. For a prime divisor E on Y, we define a valuation $\mathrm{ord}_E \in \mathrm{Val}_X$ that sends each rational function in $K(X)^{\times} = K(Y)^{\times}$ to its order of vanishing along E. Note that the center $c_X(\mathrm{ord}_E)$ is the generic point of $\mu(E)$. We say that $v \in \mathrm{Val}_X$ is a *divisorial valuation* if there exists E as above and $\lambda \in \mathbb{R}_{>0}$ such that $v = \lambda \cdot \mathrm{ord}_E$.

Let $\mu : Y \to X$ be a proper birational morphism and $\eta \in Y$ a point such that Y is regular at η. Given a system of parameters $y_1, \ldots, y_r \in \mathcal{O}_{Y,\eta}$ at η and $\alpha = (\alpha_1, \ldots, \alpha_r) \in \mathbb{R}^r_{\geq 0} \setminus \{0\}$, we define a valuation v_α as follows. For $f \in \mathcal{O}_{Y,\eta}$

we can write it as $f = \sum_{\beta \in \mathbb{Z}^r_{\geq 0}} c_\beta y^\beta$, with $c_\beta \in \widehat{\mathcal{O}_{Y,\eta}}$ either zero or unit. We set

$$v_\alpha(f) = \min\{\langle \alpha, \beta \rangle \mid c_\beta \neq 0\}.$$

A *quasi-monomial valuation* is a valuation that can be written in the above form.

Let $(Y, E = \sum_{k=1}^N E_k)$ be a log smooth model of X, i.e., $\mu : Y \to X$ is an isomorphism outside of the support of E. We denote by $\mathrm{QM}_\eta(Y, E)$ the set of all quasi-monomial valuations v that can be described at the point $\eta \in Y$ with respect to coordinates $(y_1, \ldots, y_r)$ such that each y_i defines at η an irreducible component of E (hence η is the generic point of a connected component of the intersection of some of the divisors E_i). We put $\mathrm{QM}(Y, E) := \bigcup_\eta \mathrm{QM}_\eta(Y, E)$ where η runs over generic points of all irreducible components of intersections of some of the divisors E_i.

Given a valuation $v \in \mathrm{Val}_{X,x}$, its *rational rank* $\mathrm{rat.rank}\, v$ is the rank of its value group. The *transcendental degree* $\mathrm{trans.deg}\, v$ of v is the transcendental degree of the field extension $\mathbb{C} \hookrightarrow \mathcal{O}_v/\mathfrak{m}_v$. The Zariski–Abhyankar Inequality says that

$$\mathrm{trans.deg}\, v + \mathrm{rat.rank}\, v \leq \dim X.$$

A valuation satisfying the equality is called an *Abhyankar valuation.* By [ELS03], we know that a valuation $v \in \mathrm{Val}_X$ is Abhyankar if and only if it is quasi-monomial.

Definition 2.1. Let X be an n-dimensional normal variety. Let $x \in X$ be a closed point. We define the *volume of a valuation* $v \in \mathrm{Val}_{X,x}$ following [ELS03] as

$$\mathrm{vol}_{X,x}(v) = \limsup_{m \to \infty} \frac{\ell(\mathcal{O}_{x,X}/\mathfrak{a}_m(v))}{m^n/n!}.$$

where ℓ denotes the length of the Artinian module.

Thanks to the works of [ELS03, LM09, Cut13] the above limsup is actually a limit.

Definition 2.2. Let (X, D) be a klt log pair. We define the *log discrepancy function of valuations* $A_{(X,D)} : \mathrm{Val}_X \to (0, +\infty]$ in successive generality.

(a) Let $\mu : Y \to X$ be a proper birational morphism from a normal variety Y. Let E be a prime divisor on Y. Then we define $A_{(X,D)}(\mathrm{ord}_E)$ as

$$A_{(X,D)}(\mathrm{ord}_E) := 1 + \mathrm{ord}_E(K_Y - \mu^*(K_X + D)).$$

(b) Let $(Y, E = \sum_{k=1}^N E_k)$ be a log smooth model of X. Let η be the generic point of a connected component of $E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_r}$ of codimension r. Let $(y_1, \ldots, y_r)$ be a system of parameters of $\mathcal{O}_{Y,\eta}$ at η such that $E_{i_j} = (y_j = 0)$. Then for any $\alpha = (\alpha_1, \ldots, \alpha_r) \in \mathbb{R}^r_{\geq 0} \setminus \{0\}$, we define $A_{(X,D)}(v_\alpha)$ as

$$A_{(X,D)}(v_\alpha) := \sum_{j=1}^r \alpha_j A_{(X,D)}(\mathrm{ord}_{E_{i_j}}).$$

(c) In [JM12], it was showed that there exists a retraction map $r_{Y,E} : \mathrm{Val}_X \to \mathrm{QM}(Y,E)$ for any log smooth model (Y,E) over X, such that it induces a homeomorphism $\mathrm{Val}_X \to \varprojlim_{(Y,E)} \mathrm{QM}(Y,E)$. For any real valuation $v \in \mathrm{Val}_X$, we define

$$A_{(X,D)}(v) := \sup_{(Y,E)} A_{(X,D)}(r_{(Y,E)}(v)).$$

where (Y,D) ranges over all log smooth models over X. For details, see [JM12] and [BdFFU15, Theorem 3.1]. It is possible that $A_{(X,D)}(v) = +\infty$ for some $v \in \mathrm{Val}_X$, see, e.g., [JM12, Remark 5.12].

Then we can define the main invariant in this paper. As we mentioned in Section 3, it is partially inspired the definition in [MSY08] for a valuation coming from the Reeb vector field.

Definition 2.3 ([Li18a]). Let (X,D) be an n-dimensional klt log pair. Let $x \in X$ be a closed point. Then the *normalized volume function of valuations* $\widehat{\mathrm{vol}}_{(X,D),x} : \mathrm{Val}_{X,x} \to (0,+\infty)$ is defined as

$$\widehat{\mathrm{vol}}_{(X,D),x}(v) = \begin{cases} A_{(X,D)}(v)^n \cdot \mathrm{vol}_{X,x}(v), & \text{if } A_{(X,D)}(v) < +\infty; \\ +\infty, & \text{if } A_{(X,D)}(v) = +\infty. \end{cases}$$

The *volume of the singularity* $(x \in (X,D))$ is defined as

$$\widehat{\mathrm{vol}}(x,X,D) := \inf_{v \in \mathrm{Val}_{X,x}} \widehat{\mathrm{vol}}_{(X,D),x}(v).$$

Since $\widehat{\mathrm{vol}}(v) = \widehat{\mathrm{vol}}(\lambda \cdot v)$ for any $\lambda \in \mathbb{R}_{>0}$, for any valuation $v \in \mathrm{Val}_{X,D}$ with a finite log discrepancy, we can rescale such that $\lambda \cdot v \in \mathrm{Val}^{=1}_{X,D}$ where $\mathrm{Val}^{=1}_{X,D}$ consists of all valuations $v \in \mathrm{Val}_{X,x}$ with $A_{(X,D)}(v) = 1$.

Remark 2.4. A definition of volume of singularities is also given in [BdFF12]. Their definition is the local analogue of the volume K_X whereas our definition is the one of the volume of $-K_X$. In particular, a singularity has volume 0 in the definition of [BdFF12] if it is log canonical.

2.2. Properties

In this section, we discuss some properties of $\widehat{\mathrm{vol}}$ on $\mathrm{Val}_{X,x}$. We start from the properness and Izumi estimates. As a corollary, we conclude that $\widehat{\mathrm{vol}}(x,X,D)$ is always positive for any klt singularity $x \in (X,D)$.

Theorem 2.5 ([Li18a]). *Let $(x \in (X,D))$ be a klt singularity. Then there exists positive constants C_1, C_2 which only depend on $x \in (X,D)$ (but not the valuation v) such that the following holds.*

1. (*Izumi-type inequality*) *For any valuation $v \in \mathrm{Val}_{X,x}$, we have*

$$v(\mathfrak{m}_x)\mathrm{ord}_x \le v \le C_2 \cdot A_{(X,D)}(v)\mathrm{ord}_x.$$

2. (*Properness*) *For any valuation* $v \in \mathrm{Val}_{X,x}$ *with* $A_{(X,D)}(v) < +\infty$, *we have*

$$C_1 \frac{A_{(X,D)}(v)}{v(\mathfrak{m}_x)} \leq \widehat{\mathrm{vol}}(v).$$

Note that since $x \in X$ is singular, ord_x in the above inequality might not be a valuation. In other words, for $f, g \in \mathcal{O}_{X,x}$, $\mathrm{ord}_x(fg) \geq \mathrm{ord}_x(f) + \mathrm{ord}_x(g)$ may be a strict inequality.

The above Izumi type inequality is well known when $x \in X$ is a smooth point. In the case of a general klt singularity, it can be reduced to the smooth case after a log resolution and decreasing the constant. Then for the properness, it follows from a more subtle estimate that there exists a positive constant c_2 that

$$\mathrm{vol}(v) \geq c_2 \left(\sup_{\mathfrak{m}_x} \frac{v}{\mathrm{ord}_x} \right)^{1-n} \cdot \frac{1}{v(\mathfrak{m})}.$$

Let $\mathfrak{a}_\bullet = \{\mathfrak{a}_m\}_{m\in\mathbb{Z}}$ be a graded sequence of $\mathfrak{m}_x$-primary ideals. By the works in [LM09, Cut13], the following identities hold true:

$$\mathrm{mult}(\mathfrak{a}_\bullet) := \lim_{m\to+\infty} \frac{\ell(\mathcal{O}_{X,x}/\mathfrak{a}_m)}{m^n/n!} = \lim_{m\to+\infty} \frac{\mathrm{mult}(\mathfrak{a}_m)}{m^n}.$$

In particular, the two limits exist. Note that, by definition, for any $v \in \mathrm{Val}_{X,x}$ and $\mathfrak{a}_\bullet(v) = \{\mathfrak{a}_m(v)\}$, we have $\mathrm{vol}(v) = \mathrm{mult}(\mathfrak{a}_\bullet(v))$.

The following observation on characterizing the normalized volumes by normalized multiplicities provides lots of flexibility in the study as we will see.

Theorem 2.6 ([Liu18]**).** *Let* $(x \in (X, D))$ *be an* n*-dimensional klt singularity. Then we have*

$$\begin{aligned}\widehat{\mathrm{vol}}(x, X, D) &= \inf_{\mathfrak{a}:\ \mathfrak{m}_x\text{-primary}} \mathrm{lct}(X, D; \mathfrak{a})^n \mathrm{mult}(\mathfrak{a}) \\ &= \inf_{\mathfrak{a}_\bullet:\ \mathfrak{m}_x\text{-primary}} \mathrm{lct}(X, D; \mathfrak{a}_\bullet)^n \mathrm{mult}(\mathfrak{a}_\bullet).\end{aligned}$$

We also set $\mathrm{lct}(X, D; \mathfrak{a}_\bullet)^n \mathrm{mult}(\mathfrak{a}_\bullet) = +\infty$ *if* $\mathrm{lct}(X, D; \mathfrak{a}_\bullet) = +\infty$.

Proof. Firstly, for any $\mathfrak{m}_x$-primary ideal $\mathfrak{a}$, we can take a divisorial valuation $v \in \mathrm{Val}_{X,x}$ computing $\mathrm{lct}(\mathfrak{a})$. In other words, $\mathrm{lct}(\mathfrak{a}) = A_X(v)/v(\mathfrak{a})$. We may rescale v such that $v(\mathfrak{a}) = 1$. Then clearly $\mathfrak{a}^m \subset \mathfrak{a}_m(v)$ for any $m \in \mathbb{N}$, hence $\mathrm{mult}(\mathfrak{a}) \geq \mathrm{vol}(v)$. Therefore, $\mathrm{lct}(\mathfrak{a})^n \mathrm{mult}(\mathfrak{a}) \geq A_X(v)^n \mathrm{vol}(v)$ which implies

$$\widehat{\mathrm{vol}}(x, X, D) \leq \inf_{\mathfrak{a}:\ \mathfrak{m}_x\text{-primary}} \mathrm{lct}(\mathfrak{a})^n \mathrm{mult}(\mathfrak{a}). \tag{1}$$

Secondly, for any graded sequence of $\mathfrak{m}_x$-primary ideals $\mathfrak{a}_\bullet$, we have

$$\mathrm{lct}(\mathfrak{a}_\bullet) = \lim_{m\to\infty} m \cdot \mathrm{lct}(\mathfrak{a}_m)$$

by [JM12, BdFFU15]. Hence

$$\mathrm{lct}(\mathfrak{a}_\bullet)^n \mathrm{mult}(\mathfrak{a}_\bullet) = \lim_{m\to\infty} (m \cdot \mathrm{lct}(\mathfrak{a}_m))^n \frac{\mathrm{mult}(\mathfrak{a}_m)}{m^n} = \lim_{m\to\infty} \mathrm{lct}(\mathfrak{a}_m)^n \mathrm{mult}(\mathfrak{a}_m).$$

As a result,

$$\inf_{\mathfrak{a}\colon \mathfrak{m}_x\text{-primary}} \mathrm{lct}(\mathfrak{a})^n \mathrm{mult}(\mathfrak{a}) \le \inf_{\mathfrak{a}_\bullet\colon \mathfrak{m}_x\text{-primary}} \mathrm{lct}(\mathfrak{a}_\bullet)^n \mathrm{mult}(\mathfrak{a}_\bullet). \tag{2}$$

Lastly, for any valuation $v \in \mathrm{Val}_{X,x}$, we consider the graded sequence of its valuation ideals $\mathfrak{a}_\bullet(v)$. Since $v(\mathfrak{a}_\bullet(v)) = 1$, we have $\mathrm{lct}(\mathfrak{a}_\bullet) \le A_X(v)$. We also have $\mathrm{mult}(\mathfrak{a}_\bullet(v)) = \mathrm{vol}(v)$. Hence $\mathrm{lct}(\mathfrak{a}_\bullet(v))^n \mathrm{mult}(\mathfrak{a}_\bullet(v)) \le A_X(v)^n \mathrm{vol}(v)$, which implies

$$\inf_{\mathfrak{a}_\bullet\colon \mathfrak{m}_x\text{-primary}} \mathrm{lct}(\mathfrak{a}_\bullet)^n \mathrm{mult}(\mathfrak{a}_\bullet) \le \widehat{\mathrm{vol}}(x, X, D). \tag{3}$$

The proof is finished by combining (1), (2), and (3). □

In general we have the following relation between a sequence of graded ideals and the one from a valuation: Let Φ^g be an ordered subgroup of the real numbers $\mathbb{R}$. Let $(R, \mathfrak{m})$ be the local ring at a normal singularity $o \in X$. A Φ^g-graded filtration of R, denoted by $\mathcal{F} := \{\mathfrak{a}^m\}_{m\in\Phi^g}$, is a decreasing family of $\mathfrak{m}$-primary ideals of R satisfying the following conditions:

- **(i)** $\mathfrak{a}^m \neq 0$ for every $m \in \Phi^g$, $\mathfrak{a}^m = R$ for $m \le 0$ and $\cap_{m\ge 0}\mathfrak{a}^m = (0)$;
- **(ii)** $\mathfrak{a}^{m_1} \cdot \mathfrak{a}^{m_2} \subseteq \mathfrak{a}^{m_1+m_2}$ for every $m_1, m_2 \in \Phi^g$.

Given such an $\mathcal{F}$, we get an associated order function

$$v = v_{\mathcal{F}} : R \to \mathbb{R}_{\ge 0} \qquad v(f) = \max\{m; f \in \mathfrak{a}^m\} \text{ for any } f \in R.$$

Using the above **(i)**–**(ii)**, it is easy to verify that v satisfies

$$v(f+g) \ge \min\{v(f), v(g)\} \quad \text{and} \quad v(fg) \ge v(f) + v(g).$$

We also have the associated graded ring:

$$\mathrm{gr}_{\mathcal{F}} R = \sum_{m\in\Phi^g} \mathfrak{a}^m/\mathfrak{a}^{>m}, \text{ where } \mathfrak{a}^{>m} = \bigcup_{m'>m} \mathfrak{a}^{m'}.$$

For any real valuation v with valuative group Φ^g, $\{\mathcal{F}^m\} := \{\mathfrak{a}_m(v)\}$ is a Φ^g-graded filtration of R. We will need the following facts.

Lemma 2.7 (see [Tei03, Tei14]**).** *With the above notations, the following statements hold true:*

1. ([Tei14, Page 8]) *If $\mathrm{gr}_{\mathcal{F}} R$ is an integral domain, then $v = v_{\mathcal{F}}$ is a valuation centered at $o \in X$. In particular, $v(fg) = v(f) + v(g)$ for any $f, g \in R$.*
2. (*Piltant*) *A valuation v is quasi-monomial if and only if the Krull dimension of $\mathrm{gr}_v R$ is the same as the Krull dimension of R.*

The existence of a minimizer for $\widehat{\mathrm{vol}}_{(X,D),x}$ was conjectured in the first version of [Li18a] and then proved in [Blu18].

Theorem 2.8 ([Blu18]**).** *For any klt singularity $x \in (X, D)$, there exists a valuation $v_{\min} \in \mathrm{Val}_{X,x}$ that minimizes the function $\widehat{\mathrm{vol}}_{(X,D),x}$.*

Let us sketch the idea of proving the existence of $\widehat{\mathrm{vol}}$-minimizer. We first take a sequence of valuations $(v_i)_{i\in\mathbb{N}}$ such that

$$\lim_{i\to\infty}\widehat{\mathrm{vol}}(v_i)=\widehat{\mathrm{vol}}(x,X,D).$$

Then we would like to find a valuation v^* that is a limit point of the sequence $(v_i)_{i\in\mathbb{N}}$ and then show that v^* is a minimizer of $\widehat{\mathrm{vol}}$.

Instead of seeking a limit point v^* of $(v_i)_{i\in\mathbb{N}}$ in the space of valuations, we consider graded sequences of ideals. More precisely, each valuation v_i induces a graded sequence $\mathfrak{a}_\bullet(v_i)$ of $\mathfrak{m}_x$-primary ideals. By Theorem 2.6, we have

$$\widehat{\mathrm{vol}}(v_i)\geq \mathrm{lct}(\mathfrak{a}_\bullet(v_i))^n\mathrm{mult}(\mathfrak{a}_\bullet(v_i))\geq\widehat{\mathrm{vol}}(x,X,D).$$

Therefore, once we find a graded sequence of $\mathfrak{m}_x$-primary ideals $\tilde{\mathfrak{a}}_\bullet$ that is a 'limit point' of the sequence $(\mathfrak{a}_\bullet(v_i))_{i\in\mathbb{N}}$, a valuation v^* computing $\mathrm{lct}(\tilde{\mathfrak{a}}_\bullet)$ will minimizes $\widehat{\mathrm{vol}}$. The existence of such 'limits' relies on two ingredients: the first is an asymptotic estimate to control the growth for $\mathfrak{a}_k(v_i)$ for a fixed k; once the growth is controlled, we can apply the generic limit construction.

Proof. For simplicity, we will assume $D=0$. More details about log pairs can be found in [Blu18, Section 7].

Let us choose a sequence of valuations $v_i\in\mathrm{Val}_{X,x}$ such that

$$\lim_{i\to\infty}\widehat{\mathrm{vol}}(v_i)=\widehat{\mathrm{vol}}(x,X).$$

Since the normalized volume function is invariant after rescaling, we may assume that $v_i(\mathfrak{m})=1$ for all $i\in\mathbb{N}$ where $\mathfrak{m}:=\mathfrak{m}_x$. Our goal is to show that the family of graded sequences of $\mathfrak{m}$-primary ideals $(\mathfrak{a}_\bullet(v_i))_{i\in\mathbb{N}}$ satisfies the following conditions:

(a) For every $\epsilon>0$, there exists positive constants M,N so that

$$\mathrm{lct}(\mathfrak{a}_m(v_i))^n\mathrm{mult}(\mathfrak{a}_m(v_i))\leq\widehat{\mathrm{vol}}(x,X)+\epsilon \text{ for all } m\geq M \text{ and } i\geq N.$$

(b) For each $m,i\in\mathbb{N}$, we have $\mathfrak{m}^m\subset\mathfrak{a}_m(v_i)$.

(c) There exists $\delta>0$ such that $\mathfrak{a}_m(v_i)\subset\mathfrak{m}^{\lfloor m\delta\rfloor}$ for all $m,i\in\mathbb{N}$.

Part (b) follows easily from $v_i(\mathfrak{m})=1$. Hence $\mathrm{vol}(v_i)\leq\mathrm{mult}(\mathfrak{m})=:B$. For part (c), we need to use Theorem 2.5. By Part (2), there exists a positive constant C_1 such that

$$A_X(v)\leq C_1^{-1}\cdot v(\mathfrak{m})\widehat{\mathrm{vol}}(v) \text{ for all } v\in\mathrm{Val}_{X,x}.$$

Let $A:=C_1^{-1}\sup_{i\in\mathbb{N}}\widehat{\mathrm{vol}}(v_i)$, then $A_X(v_i)\leq A$ for any $i\in\mathbb{N}$. By Theorem 2.5(1), then there exists a positive constant C_2 such that

$$v(f)\leq C_2\cdot A_X(v)\mathrm{ord}_x(f) \text{ for all } v\in\mathrm{Val}_{X,x} \text{ and } f\in\mathcal{O}_{X,x}.$$

In particular, $v_i(f)\leq C_2A\cdot\mathrm{ord}_x(f)$ for all $i\in\mathbb{N}$ and $f\in\mathcal{O}_{X,x}$. Thus by letting $\delta:=(C_2A)^{-1}$ we have $\mathfrak{a}_m(v_i)\subset\mathfrak{m}^{\lfloor m\delta\rfloor}$ which proves part (c).

The proof of part (a) relies on the following result on uniform convergence of multiplicities of valuation ideals.

Proposition 2.9 ([Blu18]). *Let $(x \in X)$ be an n-dimensional klt singularity. Then for $\epsilon, A, B, r \in \mathbb{R}_{>0}$, there exists $M = M(\epsilon, A, B, r)$ such that for every valuation $v \in \mathrm{Val}_{X,x}$ with $A_X(v) \leq A$, $\mathrm{vol}(v) \leq B$, and $v(\mathfrak{m}) \geq 1/r$, we have*

$$\mathrm{vol}(v) \leq \frac{\mathrm{mult}(\mathfrak{a}_m(v))}{m^n} < \mathrm{vol}(v) + \epsilon \textit{ for all } m \geq M.$$

Proof. The first inequality is straightforward. When the point is smooth, the second inequality uses the inequality that for the graded sequence of ideals $\{\mathfrak{a}_\bullet\}$, there exists a k such that for any m and l

$$\mathfrak{a}_{ml} \subseteq \mathfrak{a}_{m-k}^l.$$

The proof of such result uses the multiplier ideal, see [ELS03]. For isolated klt singularity, then an estimate of a similar form in [Tak06] says

$$\mathcal{J}_X^{l-1} \cdot \mathfrak{a}_{ml} \subset \mathfrak{a}_{m-k}^l \tag{4}$$

suffices, where $\mathcal{J}_X$ is the Jacobian ideal of X. Finally, in the general case, an argument using (4) and interpolating $\mathcal{J}_X$ and a power of $\mathfrak{m}$ gives the proof. See [Blu18, Section 3] for more details. □

To continue the proof, let us fix an arbitrary $\epsilon \in \mathbb{R}_{>0}$. Since $A_X(v_i) \leq A$, $\mathrm{vol}(v_i) \leq B$, and $v_i(\mathfrak{m}) = 1$ for all $i \in \mathbb{N}$, Proposition 2.9 implies that there exists $M \in \mathbb{N}$ such that

$$\frac{\mathrm{mult}(\mathfrak{a}_m(v_i))}{m^n} \leq \mathrm{vol}(v_i) + \epsilon/(2A^n) \text{ for all } i \in \mathbb{N}.$$

We also have $\mathrm{lct}(\mathfrak{a}_m(v_i)) \leq A_X(v_i)/v_i(\mathfrak{a}_m(v_i)) \leq m \cdot A_X(v_i)$. Let us take $N \in \mathbb{N}$ such that $\widehat{\mathrm{vol}}(v_i) \leq \widehat{\mathrm{vol}}(x, X) + \epsilon/2$ for any $i \geq N$. Therefore,

$$\begin{aligned}
\mathrm{lct}(\mathfrak{a}_m(v_i))^n \mathrm{mult}(\mathfrak{a}_m(v_i)) &\leq A_X(v_i)^n(\mathrm{vol}(v_i) + \epsilon/(2A^n)) \\
&= \widehat{\mathrm{vol}}(v_i) + \epsilon \cdot A_X(v_i)^n/(2A^n) \\
&= \widehat{\mathrm{vol}}(v_i) + \epsilon/2 \\
&\leq \widehat{\mathrm{vol}}(x, X) + \epsilon.
\end{aligned}$$

So part (a) is proved.

Finally, (b) and (c) guarantee that we can apply a generic limit type construction (cf. [Blu18, Section 5]). Then (a) implies that a 'limit point' $\tilde{\mathfrak{a}}_\bullet$ of the sequence $(\mathfrak{a}_\bullet(v_i))_{i \in \mathbb{N}}$ satisfies that $\mathrm{lct}(\tilde{\mathfrak{a}}_\bullet)^n \mathrm{mult}(\tilde{\mathfrak{a}}_\bullet) \leq \widehat{\mathrm{vol}}(x, X)$. Thus a valution v^* computing the log canonical threshold of $\tilde{\mathfrak{a}}_\bullet$, whose existence follows from [JM12], necessarily minimizes the normalized volume. □

Theorem 2.10 ([LX17b]). *Let $x \in (X, D)$ be an n-dimensional klt singularity. Then $\widehat{\mathrm{vol}}(x, X, D) \leq n^n$ and the equality holds if and only if $x \in X \backslash \mathrm{Supp}(D)$ is a smooth point.*

Using the fact that we can specialize a graded sequence of ideals preserving the colength, and the lower semi-continuous of the log canonical thresholds, we easily get the inequality part of Theorem 2.10. Then the equality part gives us a characterization of the smooth point using the normalized volume. The following Theorem 2.11 on the semicontinuity needs a more delicate analysis. We conjecture that the normalized volume function is indeed constructible (see Conjecture 6.6).

Theorem 2.11 ([BL18]). *Let $\pi : (\mathcal{X}, D) \to T$ together with a section $t \in T \mapsto x_t \in \mathcal{X}_t$ be a $\mathbb{Q}$-Gorenstein flat family of klt singularities. Then the function $t \mapsto \widehat{\mathrm{vol}}(x_t, \mathcal{X}_t, D_t)$ is lower semicontinuous with respect to the Zariski topology.*

Now we introduce a key tool that the minimal model program provides to us to understand minimizing the normalized volume. For more discussions, see Section 4.3.

Definition 2.12 (Kollár component, [Xu14]**).** Let $x \in (X, D)$ be a klt singularity. We call a proper birational morphism $\mu : Y \to X$ provides a Kollár component S, if μ is isomorphic over $X \setminus \{x\}$, and $\mu^{-1}(x)$ is an irreducible divisor S, such that $(Y, S + \mu_*^{-1}D)$ is purely log terminal (plt) and $-S$ is $\mathbb{Q}$-Cartier and ample over X.

Theorem 2.13 ([LX16]). *We have the identity:*

$$\widehat{\mathrm{vol}}(x, X, D) = \inf_S \{\widehat{\mathrm{vol}}(\mathrm{ord}_S) \mid \textit{for all Kollár components } S \textit{ over } x\}. \tag{5}$$

For the explanation of proof, see the discussions for (27) in Section 4.3.

3. Stability in Sasaki–Einstein geometry

To proceed the study of normalized volumes, we will introduce the concept of K-stability. This is now a central notion in complex geometry, which serves as an algebraic characterization of the existence of some 'canonical metrics'.

In the local setting, such problem on an affine T-variety X with a unique fixed point x was first considered in [MSY08]. We can then vary the Reeb vector field $\xi \in \mathfrak{t}^+_{\mathbb{R}}$, and call such a structure (X, ξ) is a Fano cone if X only has klt log terminal singularities. The name is justified since if $\xi \in \mathfrak{t}^+_{\mathbb{Q}}$, let $\langle \xi \rangle$ be the $\mathbb{C}^*$ generated by ξ, then $X \setminus \{x\}/\langle \xi \rangle$ is a log Fano variety.

In [MSY08], the relation between the existence of Sasaki–Einstein metric along (X, ξ) and the K-stability of (X, ξ), a mimic of the absolute case, was explored. A key observation in [MSY08] is that we can define a normalized volume function $\widehat{\mathrm{vol}}_X(\xi)$ for $\xi \in \mathfrak{t}^+_{\mathbb{R}}$, and among all choices of ξ the one minimizing $\widehat{\mathrm{vol}}_X(\cdot)$ gives 'the most stable' direction.

Then an important step to advance such a picture is made in [CS18, CS15] by extending the definition of K-stability notions on (X, ξ) allowing degenerations, and showing that there is a Sasaki–Einstein metric along an isolated Fano cone singularity (X, ξ) if and only of (X, ξ) is K-polystable, extending the solution of the

Yau–Tian–Donaldson's conjecture in the Fano manifold case (see [CDS15, Tia15]) to the cone case.

In this section, we will briefly introduce these settings.

3.1. T-varieties

We first introduce the basic setting using T-varieties. For general results of T-varieties, see [AIP$^+$12].

Assume $X = \mathrm{Spec}_{\mathbb{C}}(R)$ is an affine variety with $\mathbb{Q}$-Gorenstein klt singularities. Denote by T the complex torus $(\mathbb{C}^*)^r$. Assume X admits a good T-action in the following sense.

Definition 3.1 (see [LS13, Section 4]**).** Let X be a normal affine variety. We say that a T-action on X is *good* if it is effective and there is a unique closed point $x \in X$ that is in the orbit closure of any T-orbit. We shall call x (sometimes also denoted by o_X) the vertex point of the T-variety X.

Let $N = \mathrm{Hom}(\mathbb{C}^*, T)$ be the co-weight lattice and $M = N^*$ the weight lattice. We have a weight space decomposition of the coordinate ring of X:

$$R = \bigoplus_{\alpha \in \Gamma} R_\alpha \text{ where } \Gamma = \{\alpha \in M |\ R_\alpha \neq 0\}.$$

The action being good implies $R_0 = \mathbb{C}$, which will always be assumed in the below. An ideal $\mathfrak{a}$ is called homogeneous if $\mathfrak{a} = \bigoplus_{\alpha \in \Gamma} \mathfrak{a} \cap R_\alpha$. Denote by $\sigma^\vee \subset M_{\mathbb{Q}}$ the cone generated by Γ over $\mathbb{Q}$, which will be called the *weight cone* or the *moment cone*. The cone $\sigma \subset N_{\mathbb{R}}$, dual to $\sigma^\vee$, is the same as the following conical set

$$\mathfrak{t}^+_{\mathbb{R}} := \{\ \xi \in N_{\mathbb{R}} \ \ | \ \langle \alpha, \xi \rangle > 0 \text{ for any } \alpha \in \Gamma \setminus \{0\}\}.$$

Motivated by notations from Sasaki geometry, we will introduce:

Definition 3.2. With the above notations, $\mathfrak{t}^+_{\mathbb{R}}$ will be called the Reeb cone of the T-action of X. A vector $\xi \in \mathfrak{t}^+_{\mathbb{R}}$ will be called a Reeb vector on the T-variety X.

To adapt this definition into our setting in Section 2.1, for any $\xi \in \mathfrak{t}^+_{\mathbb{R}}$, we can define a valuation

$$\mathrm{wt}_\xi(f) = \min_{\alpha \in \Gamma} \{\langle \alpha, \xi \rangle \mid f_\alpha \neq 0\}.$$

It is easy to verify that $\mathrm{wt}_\xi \in \mathrm{Val}_{X, o_X}$. The rank of ξ, denoted by $\mathrm{rk}(\xi)$, is the dimension of the subtorus T_ξ (as a subgroup of T) generated by $\xi \in \mathfrak{t}$. The following lemma can be easily seen.

Lemma 3.3. *For any $\xi \in \mathfrak{t}^+_{\mathbb{R}}$,* wt_ξ *is a quasi-monomial valuation of rational rank equal to the rank of ξ. Moreover, the center of* wt_ξ *is o_X.*

We recall the following structure results for any T-varieties.

Theorem 3.4 (see [AIP+12, Theorem 4]**).** *Let $X = \mathrm{Spec}(R)$ be a normal affine variety and suppose $T = \mathrm{Spec}\,(\mathbb{C}[M])$ has a good action on X with the weight cone $\sigma^\vee \subset M_{\mathbb{Q}}$. Then there exist a normal projective variety Y and a polyhedral divisor $\mathfrak{D}$ such that there is an isomorphism of graded algebras:*

$$R \cong H^0(X, \mathcal{O}_X) \cong \bigoplus_{u \in \sigma^\vee \cap M} H^0\big(Y, \mathcal{O}(\mathfrak{D}(u))\big) =: R(Y, \mathfrak{D}).$$

In other words, X is equal to $\mathrm{Spec}_{\mathbb{C}}\big(\bigoplus_{u\in\sigma^\vee\cap M} H^0(Y, \mathcal{O}(\mathfrak{D}(u)))\big)$.

In the above definition, a polyhedral divisor $\mathfrak{D} : u \to \mathfrak{D}(u)$ is a map from $\sigma^\vee$ to the set of $\mathbb{Q}$-Cartier divisors that satisfies:

1. $\mathfrak{D}(u) + \mathfrak{D}(u') \le \mathfrak{D}(u + u')$ for any $u, u' \in \sigma^\vee$;
2. $u \mapsto \mathfrak{D}(u)$ is piecewisely linear;
3. $\mathfrak{D}(u)$ is semiample for any $u \in \sigma^\vee$, and $\mathfrak{D}(u)$ is big if u is in the relative interior of $\sigma^\vee$.

Here Y is projective since from our assumption

$$H^0(Y, \mathcal{O}_Y) = R^T = R_0 = \mathbb{C}$$

(see [LS13]). We collect some basic results about valuations on T-varieties.

Theorem 3.5 (see [AIP+12]**).** *Assume a T-variety X is determined by the data $(Y, \sigma, \mathfrak{D})$ such that Y is a projective variety, where $\sigma = \mathfrak{t}^+_{\mathbb{R}} \subset N_{\mathbb{R}}$ and $\mathfrak{D}$ is a polyhedral divisor.*

1. *For any T-invariant quasi-monomial valuation v, there exist a quasi-monomial valuation $v^{(0)}$ over Y and $\xi \in M_{\mathbb{R}}$ such that for any $f \cdot \chi^u \in R_u$, we have:*
$$v(f \cdot \chi^u) = v^{(0)}(f) + \langle u, \xi\rangle.$$
We will use $(\xi, v^{(0)})$ to denote such a valuation.
2. *T-invariant prime divisors on X are either vertical or horizontal. Any vertical divisor is determined by a divisor Z on Y and a vertex v of $\mathfrak{D}_Z$, and will be denoted by $D_{(Z,v)}$. Any horizontal divisor is determined by a ray ρ of σ and will be denoted by E_ρ.*
3, *Let D be a T-invariant vertical effective $\mathbb{Q}$-divisor. If $K_X + D$ is $\mathbb{Q}$-Cartier, then the log canonical divisor has a representation $K_X + D = \pi^* H + \mathrm{div}(\chi^{-u_0})$ where $H = \sum_Z a_Z \cdot Z$ is a principal $\mathbb{Q}$-divisor on Y and $u_0 \in M_{\mathbb{Q}}$. Moreover, the log discrepancy of the horizontal divisor E_ρ is given by:*
$$A_{(X,D)}(E_\rho) = \langle u_0, n_\rho\rangle, \tag{6}$$
where n_ρ is the primitive vector along the ray ρ.

Sketch of the proof. For the first statement, the case of divisorial valuations follows from [AIP+12, Section 11]. It can be extended to the case of quasi-monomial valuations by the same proof. Note also that any T-invariant quasimonomial valuation can be approximated by a sequence of T-invariant divisorial valuations. The second statement is in [PS11, Proposition 3.13]. The absolute case (e.g., without boundary divisor D) for the third statement is from [LS13, Section 4] whose proof also works for the case of log pairs. $\square$

We will specialize the study of general affine T-varieties to case that the log pair is klt. Assume X is a normal affine variety with $\mathbb{Q}$-Gorenstein klt singularities and a good T-action. Let D be a T-invariant vertical divisor. Then there is a nowhere-vanishing T-equivariant section s of $m(K_X+D)$ where m is sufficiently divisible. The following lemma says that the log discrepancy of wt_ξ can indeed be calculated in a similar way as in the toric case (the toric case is well known). Moreover, it can be calculated by using the weight of T-equivariant pluri-log-canonical sections. The latter observation was first made in [Li18a].

Lemma 3.6. *Using the same notion as in the Theorem* 3.5*, the log discrepancy of* wt_ξ *is given by:* $A_{(X,D)}(\mathrm{wt}_\xi)=\langle u_0,\xi\rangle$. *Moreover, let s be a T-equivariant nowhere-vanishing holomorphic section of* $|-m(K_X+D)|$*, and denote* $\mathcal{L}_\xi$ *the Lie derivative with respect to the holomorphic vector field associated to* ξ. *Then* $A_{(X,D)}(\xi)=\lambda$ *if and only if*

$$\mathcal{L}_\xi(s)=m\lambda s \quad \textit{for} \quad \lambda>0.$$

As a consequence of the above lemma, we can formally extend $A_{(X,D)}(\xi)$ to a linear function on $\mathfrak{t}_\mathbb{R}$:

$$A_{(X,D)}(\eta)=\langle u_0,\eta\rangle. \tag{7}$$

for any $\eta\in\mathfrak{t}_\mathbb{R}$. By Lemma 3.6, $A_{(X,D)}(\eta)=\frac{1}{m}\mathcal{L}_\eta s/s$ where s is a T-equivariant nowhere-vanishing holomorphic section of $|-m(K_X+D)|$.

Definition 3.7 (Log Fano cone singularity). Let (X,D) be an affine pair with a good T action. Assume (X,D) is a normal pair with klt singularities. Then for any $\xi\in\mathfrak{t}^+_\mathbb{R}$, we call the triple (X,D,ξ) a *log Fano cone* structure that is polarized by ξ. We will denote by $\langle\xi\rangle$ the sub-torus of T generated by ξ. If $\langle\xi\rangle\cong\mathbb{C}^*$, then we call (X,D,ξ) to be quasi-regular. Otherwise, we call it irregular.

In the quasi-regular case, we can take the quotient $(X\setminus\{x\},D\setminus\{x\})$ by the $\mathbb{C}^*$-group generated by ξ in the sense of Seifert $\mathbb{C}^*$-bundles (see [Kol04]), and we will denote by $(X,D)/\langle\xi\rangle$, which is a log Fano variety, because of the assumption that (X,D) is klt at x (see [Kol13, Lemma 3.1]).

3.2. *K*-stability

In this section, we will discuss the K-stability notion of log Fano cones, which generalizes the K-stability of log Fano varieties originally defined by Tian and Donaldson. For irregular Fano cones, such a notion was first defined in [CS18].

Definition 3.8 (Test configurations). Let (X,D,ξ_0) be a log Fano cone singularity and T a torus containing $\langle\xi_0\rangle$.

A T-equivariant test configuration (or simply called a test configuration) of (X,D,ξ_0) is a quadruple $(\mathcal{X},\mathcal{D},\xi_0;\eta)$ with a map $\pi:(\mathcal{X},\mathcal{D})\to\mathbb{C}$ satisfying the following conditions:

(1) $\pi : \mathcal{X} \to \mathbb{C}$ is a flat family and $\mathcal{D}$ is an effective $\mathbb{Q}$-divisor such that $\mathcal{D}$ does not contain any component X_0, the fibres away from 0 are isomorphic to (X, D) and $\mathcal{X} = \mathrm{Spec}(\mathcal{R})$ is affine, where $\mathcal{R}$ is a finitely generate flat $\mathbb{C}[t]$ algebra. The torus T acts on $\mathcal{X}$, and we write $\mathcal{R} = \bigoplus_\alpha \mathcal{R}_\alpha$ as decomposition into weight spaces.

(2) η is a holomorphic vector field on $\mathcal{X}$ generating a $\mathbb{C}^*(= \langle\eta\rangle)$-action on $(\mathcal{X}, \mathcal{D})$ such that π is $\mathbb{C}^*$-equivariant where $\mathbb{C}^*$ acts on the base $\mathbb{C}$ by the multiplication (so that $\pi_*\eta = t\partial_t$ if t is the affine coordinate on $\mathbb{C}$) and there is a $\mathbb{C}^*$-equivariant isomorphism $\phi : (\mathcal{X}, \mathcal{D}) \times_{\mathbb{C}} \mathbb{C}^* \cong (X, D) \times \mathbb{C}^*$.

(3) The torus T-action commutes with η. The holomorphic vector field ξ_0 on $\mathcal{X} \times_{\mathbb{C}} \mathbb{C}^*$ (via the isomorphism ϕ) extends to a holomorphic vector field on $\mathcal{X}$ which we still denote to be ξ_0.

In most of our study, we only need to treat the case that test configuration $(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ of (X, D, ξ_0) satisfies that

(4) $K_{\mathcal{X}} + \mathcal{D}$ is $\mathbb{Q}$-Cartier and the central fibre (X_0, D_0) is klt.

In other words, we will mostly consider special test configurations (see [LX14, CS15]).

Condition (1) implies that each weight piece $\mathcal{R}_\alpha$ is a flat $\mathbb{C}[t]$-module. So X and X_0 have the same weight cone and Reeb cone with respect to the fiberwise T-action.

A test configuration $(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ is called a product one if there is a T-equivariant isomorphism $(\mathcal{X}, \mathcal{D}) \cong (X, D) \times \mathbb{C}$ and $\eta = \eta_0 + t\partial_t$ where η_0 is a holomorphic vector field on X that preserves D and commutes with ξ_0. In this case, we will denote $(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ by

$$(X \times \mathbb{C}, D \times \mathbb{C}, \xi_0; \eta) =: (X_{\mathbb{C}}, D_{\mathbb{C}}, \xi_0; \eta).$$

In [MSY08], only such test configurations are considered.

Definition 3.9 (K-stability). For any special test configuration $(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ of (X, D, ξ_0) with central fibre (X_0, D_0, ξ_0), its generalized Futaki invariant is defined as

$$\mathrm{Fut}(\mathcal{X}, \mathcal{D}, \xi_0; \eta) := \frac{D_{-T_{\xi_0}(\eta)}\mathrm{vol}_{X_0}(\xi_0)}{\mathrm{vol}_{X_0}(\xi_0)}$$

where we denote

$$T_{\xi_0}(\eta) = \frac{A(\xi_0)\eta - A(\eta)\xi_0}{n}. \tag{8}$$

Since the generalized Futaki invariant defined above only depends on the data on the central fibre, we will also denote it by $\mathrm{Fut}(X_0, D_0, \xi_0; \eta)$.

We say that (X, D, ξ_0) is K-semistable, if for any special test configuration, $\mathrm{Fut}(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ is nonnegative.

We say that (X, D, ξ_0) is K-polystable, if it is K-semistable, and any special test configuration $(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ with $\mathrm{Fut}(\mathcal{X}, \mathcal{D}, \xi_0; \eta) = 0$ is a product test configuration.

In the above definition, we used the notation (8) and the directional derivative:

$$D_{-T_{\xi_0(\eta)}}\mathrm{vol}_{X_0}(\xi_0) := \frac{d}{d\epsilon}\Big|_{\epsilon=0} \mathrm{vol}_{X_0}(\xi_0 - \epsilon T_{\xi_0(\eta)}).$$

Recall that the $\pi_*\eta = t\partial_t$. Then the negative sign in front of $T_{\xi_0}(\eta)$ in the above formula is to be compatible with our later computation. Using the rescaling invariance of the normalized volume, it is easy to verify that the following identity holds:

$$D_{-T_{\xi_0}\eta}\mathrm{vol}_{X_0}(\xi_0) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} \widehat{\mathrm{vol}}_{X_0}(\mathrm{wt}_{\xi_0-\epsilon\eta}) \cdot \frac{1}{nA(\xi_0)^{n-1}}, \tag{9}$$

where $A(\xi_0) = A_{(X_0,D_0)}(\mathrm{wt}_{\xi_0})$. As a consequence, we can rewrite the Futaki invariant of a special test configuration as:

$$\mathrm{Fut}(\mathcal{X}, \mathcal{D}, \xi_0; \eta) := D_{-\eta}\widehat{\mathrm{vol}}_{X_0}(\mathrm{wt}_{\xi_0}) \cdot \frac{1}{nA(\xi_0)^{n-1} \cdot \mathrm{vol}_{X_0}(\xi_0)}. \tag{10}$$

One can show that, up to a constant, the above definition of $\mathrm{Fut}(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ coincides with the one in [CS18, CS15] defined using index characters. For convenience of the reader, we recall their definition. It is enough to define the Futaki invariant for the central fibre which we just denote by X. For any $\xi \in \mathfrak{t}_{\mathbb{R}}^+$, the index character $F(\xi, t)$ is defined by:

$$F(\xi, t) := \sum_{\alpha \in \Gamma} e^{-t\langle \alpha, \xi\rangle} \dim_{\mathbb{C}} R_\alpha. \tag{11}$$

Then there is a meromorphic expansion for $F(\xi, t)$ as follows:

$$F(\xi, t) = \frac{a_0(\xi)(n-1)!}{t^n} + \frac{a_1(\xi)(n-2)!}{t^{n-1}} + O(t^{2-n}). \tag{12}$$

One always has the identity $a_0(\xi) = \mathrm{vol}(\xi)/(n-1)!$.

Definition 3.10 (see [CS18]**).** For any $\eta \in \mathfrak{t}_{\mathbb{R}}$, define:

$$\begin{aligned} \mathrm{Fut}_{\xi_0}(X, \eta) &= \frac{1}{n-1} D_{-\eta}(a_1(\xi_0)) - \frac{1}{n}\frac{a_1(\xi_0)}{a_0(\xi_0)} D_{-\eta} a_0(\xi_0) \\ &= \frac{a_0(\xi_0)}{n-1} D_{-\eta}\left(\frac{a_1}{a_0}\right)(\xi_0) + \frac{a_1(\xi_0) D_{-\eta} a_0(\xi_0)}{n(n-1)a_0(\xi_0)}. \end{aligned}$$

This is a complicated expression. But in [CS15, Proposition 6.4], it was shown that, when X is $\mathbb{Q}$-Gorenstein log terminal, there is an identity $a_1(\xi)/a_0(\xi) = A(\xi)(n-1)/2$ for any $\xi \in \mathfrak{t}_{\mathbb{R}}^+$ (by using our notation involving log discrepancies). Note that the rescaling properties $a_0(\lambda\xi) = \lambda^{-n}a_0(\xi)$ and $a_1(\lambda\xi) = \lambda^{-(n-1)}a_1(\xi)$ which imply $\mathrm{Fut}_{\xi_0}(X, \xi_0) = 0$. If we denote $\eta' = \eta - \frac{A(\eta)}{A(\xi_0)}\xi_0$, then we get:

$$\mathrm{Fut}_{\xi_0}(X, \eta) = \mathrm{Fut}_{\xi_0}(X, \eta') = \frac{A(\xi_0)}{2n} D_{-\eta'} a_0(\xi_0) = \frac{1}{2(n-1)!} D_{-T_{\xi_0}(\eta)}\mathrm{vol}(\xi_0). \tag{13}$$

So the definition in [CS18, CS15] differs from our notation by a constant $2(n-1)!/\mathrm{vol}_X(\xi_0)$.

Remark 3.11. More precisely, our notation differs from that in [CS18] by a sign. Our choice of minus sign for $-\eta$, besides being compatible with the sign choice in Tian's original definition of K-stability in [Tia97], is made for least two reasons. The first is that the careful calculation in [LX17a, Section 5.2] shows that the limiting slope of the Ding energy along the geodesic ray associated to any special test configuration is indeed the directional derivative of $\mathrm{vol}(\xi)$ along $-\eta$ instead of η. For the second reason, as we stressed in [LX17a, Remark 3.4], for the special test configuration coming from a Kollár component S, the $-\eta$ vector corresponds to ord_S. Since our goal is to compare $\widehat{\mathrm{vol}}(\mathrm{wt}_{\xi_0})$ and $\widehat{\mathrm{vol}}(\mathrm{ord}_S)$, $-\eta$ is the correct choice of sign (see [LX17a, Proof of Theorem 3.5]).

Remark 3.12. In fact, in a calculation, instead of the generalized Futaki invariant, it is the Berman–Ding invariant, denoted by $D^{\mathrm{NA}}(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$, where

$$D^{\mathrm{NA}}(\mathcal{X}, \mathcal{D}, \xi_0; \eta) := \frac{D_{-T_{\xi_0}(\eta)}\mathrm{vol}_{X_0}(\xi_0)}{\mathrm{vol}(\xi_0)} - (1 - \mathrm{lct}(\mathcal{X}, D; \mathcal{X}_0)).$$

appears more naturally, whenever we know

(D) there exists a nowhere vanishing section $s \in |m(K_{\mathcal{X}} + \mathcal{D})|$ such that we can use it to define $A(\cdot)$ as in the formula in Lemma 3.6.

Then we can similarly define Ding semi(poly)-stable, replacing $\mathrm{Fut}(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ by $D^{\mathrm{NA}}(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$. For a special test configuration, since

$$\mathrm{Fut}(\mathcal{X}, \mathcal{D}, \xi_0; \eta) = D^{\mathrm{NA}}(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$$

the two notions coincide.

If we specialize the above definitions to the case of quasi-regular log Fano cone (X, D, ξ_0), then we get the corresponding more familiar notions for the log Fano projective pair $(S, B) = (X, D)/\langle \xi_0 \rangle$.

3.3. Sasaki–Einstein geometry

The introduction of normalized volumes in [Li18a] was motivated by the minimization phenomenon in the study of Sasaki–Einstein metrics. The latter was discovered in [MSY06, MSY08] and was motivated by the so-called AdS/CFT correspondence from mathematical physics. Here we give a short account on this. For the reader who are mostly interested in the algebraic part of the theory, one can skip this section. The results will only be used in Section 5.2.

Classically, a Sasaki manifold is defined as an odd-dimensional Riemannian manifold (M^{2n-1}, g_M) such that metric cone over it, defined as:

$$(X, g_X) := ((M \times \mathbb{R}_{>0}) \cup \{o_X\}, dr^2 + r^2 g_M)$$

is Kähler. It's convenient to work directly on $X = X^\circ \cup o_X$ which is an affine variety with the Kähler metric $\sqrt{-1}\partial\bar{\partial}r^2$. The Reeb vector field of (X°, g_X) is usually defined as $J(r\partial_r)$ where J is the complex structure on X°. The corresponding holomorphic vector field $\xi = r\partial_r - iJ(r\partial_r)$, which we also call Reeb vector field, generates a $T_\xi \cong (\mathbb{C}^*)^{\mathrm{rk}(\xi)}$-action on X where $r(\xi) \geq 1$. For simplicity, we will denote such a torus by $\langle\xi\rangle$. Moreover the corresponding element in $(\mathfrak{t}_\xi)_{\mathbb{R}}$, also denoted by ξ is in the Reeb cone: $\xi \in (\mathfrak{t}_\xi)^+_{\mathbb{R}}$. The volume of ξ is defined to be the volume density of g_X:

$$\begin{aligned}\mathrm{vol}(\xi) := \mathrm{vol}(r^2) &= \frac{1}{(2\pi)^n n!}\int_X e^{-r^2}(\sqrt{-1}\partial\bar{\partial}r^2)^n \\ &= \frac{1}{(2\pi)^n}\int_M (-Jdr)\wedge(-dJdr)^{n-1} \\ &= \frac{(n-1)!}{2\pi^n}\mathrm{vol}(M, g_M) = \frac{\mathrm{vol}(M, g_M)}{\mathrm{vol}(\mathbb{S}^{2n-1})} \\ &= \frac{\mathrm{vol}(B_1(X), g_X)}{\mathrm{vol}(B_1(\underline{0}), g_{\mathbb{C}^n})}. \end{aligned} \tag{14}$$

Here $g_X = \frac{1}{2}\sqrt{-1}\partial\bar{\partial}r^2(\cdot, J\cdot)$ and $g_M = g_X|_M$ are the Riemannian metric on X and M respectively, $\mathbb{S}^{2n-1}$ is the standard unit sphere in $\mathbb{C}^n$ with volume $\mathrm{vol}(\mathbb{S}^{2n-1}) = 2\pi^n/(n-1)!$.

This is well defined because if two Sasaki metrics have the same Reeb vector field, then their volumes are the same. Indeed, $\omega_1 = \sqrt{-1}\partial\bar{\partial}r_1$ and $\omega_2 = \sqrt{-1}\partial\bar{\partial}r_2$ have the same Reeb vector field if $r_2 = r_1 e^{\varphi}$ for a function φ satisfying $\mathcal{L}_{r\partial_r}\varphi = \mathcal{L}_\xi\varphi = 0$ (i.e., φ is a horizontal function on M with respect to the foliation defined by $\mathrm{Im}(\xi_0)$). Letting $r_t^2 = r^2 e^{t\varphi}$ and differentiating the volume we get:

$$\begin{aligned} &C\cdot\frac{d}{dt}\mathrm{vol}(r_t^2) = \int_X e^{-r_t^2}(-r_t^2\varphi)\sqrt{-1}\partial\bar{\partial}r_t^2)^n + e^{-r_t^2}n\sqrt{-1}\partial\bar{\partial}(r_t^2\varphi)\wedge(\sqrt{-1}\partial\bar{\partial}r_t^2)^{n-1} \\ &= \int_X -e^{-r_t^2}r_t^2\varphi(\sqrt{-1}\partial\bar{\partial}r_t^2)^n + e^{-r_t^2}n\sqrt{-1}\partial r_t^2\wedge(\varphi\bar{\partial}r_t^2)\wedge(\sqrt{-1}\partial\bar{\partial}r_t^2)^{n-1} \\ &\quad + \int_X e^{-r_t^2}n\sqrt{-1}\partial r_t^2\wedge(r_t^2\bar{\partial}\varphi)\wedge(\sqrt{-1}\partial\bar{\partial}r_t^2)^{n-1} \\ &= 0. \end{aligned}$$

The second equality follows from integration by parts. The last equality follows by substituting $f = r_t^2$ and $f = \varphi$ into the following identities and using the fact that φ is horizontal (so that $\xi_t(\varphi) = 0$):

$$n\sqrt{-1}\partial r_t^2\wedge\bar{\partial}f\wedge(\sqrt{-1}\partial\bar{\partial}r_t^2)^{n-1} = \xi_t(f)(\sqrt{-1}\partial\bar{\partial}r_t^2)^n.$$

One should compare this to the fact that two Kähler metrics in the same Kähler class have the same volume.

The Reeb vector field associated to a Ricci-flat Kähler cone metric satisfies the minimization principle in [MSY08]. To state it in general, we assume X is a T-variety with the Reeb cone $\mathfrak{t}^+_{\mathbb{R}}$ with respect to T and recall the variation formulas of volumes of Reeb vector fields from [MSY08]. For any $\xi \in \mathfrak{t}^+_{\mathbb{R}}$, we can find a radius function $r : X \to \mathbb{R}_+$ such that $\mathrm{vol}(\xi)$ is given by the formula (14).

Lemma 3.13. *The first-order derivative of* $\mathrm{vol}_X(\xi)$ *is given by:*

$$D\mathrm{vol}(\xi)\cdot\eta_1 = \frac{1}{(2\pi)^n(n-1)!}\int_X \theta_1 e^{-r^2}(\sqrt{-1}\partial\bar{\partial}r^2)^n, \tag{15}$$

where $\theta_i = \eta_i(\log r^2)$. *The second-order variation of* $\mathrm{vol}_X(\xi)$ *is given by:*

$$D^2\mathrm{vol}(\xi)(\eta_1,\eta_2) = \frac{n+1}{(2\pi)^n(n-1)!}\int_X \theta_1\theta_2 e^{-r^2}(\sqrt{-1}\partial\bar{\partial}r^2)^n.$$

Now we fix a $\xi_0 \in \mathfrak{t}^+_{\mathbb{R}}$ and a radius function $r : X \to \mathbb{R}_+$ (by using equivariant embedding of X into $\mathbb{C}^N$ for example), we define:

Definition 3.14. $PSH(X,\xi_0)$ is the set of bounded real functions φ on X° that satisfies:

(1) $\varphi\circ\tau = \varphi$ for any $\tau \in \langle\xi_0\rangle$, the torus generated by ξ_0;
(2) $r^2_\varphi := r^2 e^\varphi$ is a proper plurisubharmonic function on X.

To write down the equation of Ricci-flat Kähler-cone equation, we fix a T-equivariant nowhere vanishing section $s \in H^0(X, mK_X)$ as in the last section and define an associated volume form on X:

$$dV_X := \left((\sqrt{-1})^{mn^2} s\wedge\bar{s}\right)^{1/m}. \tag{16}$$

Definition 3.15. We say that $r^2_\varphi := r^2e^\varphi$ where $\varphi \in PSH(X,\xi_0)$ is the radius function of a Ricci-flat Kähler cone metric on (X,ξ_0) if φ is smooth on X^{reg} and there exists a positive constant $C > 0$ such that

$$(\sqrt{-1}\partial\bar{\partial}r^2_\varphi)^n = C\cdot dV, \tag{17}$$

where the constant C is equal to:

$$C = \frac{\int_X e^{-r^2_\varphi}(\sqrt{-1}\partial\bar{\partial}r^2_\varphi)^n}{\int_X e^{-r^2_\varphi}dV_X} = \frac{(2\pi)^n n!\mathrm{vol}(\xi_0)}{\int_X e^{-r^2_\varphi}dV_X}.$$

Motivated by standard Kähler geometry, one defines the Monge–Ampère energy $E(\varphi)$ using either its variations or the explicit expression on the link $M := X\cap\{r=1\}$:

$$\delta E(\varphi)\cdot\delta\varphi = -\frac{1}{(n-1)!(2\pi)^n\mathrm{vol}(\xi_0)}\int_X \delta\varphi e^{-r^2_\varphi}(\sqrt{-1}\partial\bar{\partial}r^2_\varphi)^n.$$

Then the equation (16) is the Euler–Lagrange equation of the following Ding–Tian-typed functional:

$$D(\varphi) = E(\varphi) - \log\left(\int_X e^{-r^2_\varphi}dV_X\right).$$

This follows from the identity:

$$\begin{aligned}\delta D(\varphi)\cdot\delta\varphi &= \frac{1}{(2\pi)^n(n-1)!\mathrm{vol}(\xi_0)}\int_X \delta\varphi e^{-r_\varphi^2}(\sqrt{-1}\partial\bar\partial r_\varphi^2)^n - \frac{\int_X r_\varphi^2\delta\varphi e^{-r_\varphi^2}dV_X}{\int_X e^{-r_\varphi^2}dV_X}\\ &= n\int_X e^{-r_\varphi^2}\delta\varphi\left(\frac{(\sqrt{-1}\partial\bar\partial r_\varphi^2)^n}{(2\pi)^n n!\mathrm{vol}(\xi_0)} - \frac{dV_X}{\int_X e^{-r_\varphi^2}dV_X}\right).\end{aligned}$$

Compared with the weak Kähler–Einstein case, it is expected that the regularity condition in the above definition is automatically satisfied. With this regularity assumption, on the regular part X^{reg}, both sides of (17) are smooth volume forms and we have $r_\varphi\partial_{r_\varphi} = 2\mathrm{Re}(\xi_0)$ or, equivalently, $\xi_0 = r_\varphi\partial_{r_\varphi} - iJ(r_\varphi\partial_{r_\varphi})$. Moreover, taking $\mathcal{L}_{r_\varphi\partial_{r_\varphi}}$ on both sides gives us the identity $\mathcal{L}_{r_\varphi\partial_{r_\varphi}}dV = 2n\ dV$. Equivalently we have:

$$\mathcal{L}_{\xi_0}s = mn\cdot s,$$

where $s \in |-mK_X|$ is the chosen T-equivariant non-vanishing holomorphic section. By Lemma 3.6, this implies $A_X(\mathrm{wt}_{\xi_0}) = n$ (see [HS17, LL19] for this identity in the quasi-regular case). The main result of [MSY08] can be stated as follows.

Theorem 3.16. *If (X,ξ_0) admits a Ricci-flat Kähler cone metric, then $A_X(\xi_0) = n$ and wt_{ξ_0} obtains the minimum of vol on $\mathfrak{t}_{\mathbb{R}}^+$.*

The following result partially generalizes Berman's result on K-polystability of Kähler–Einstein Fano varieties to the more general case of Ricci-flat Fano cones. Together with Theorem 4.6, it is used to show a generalization the minimization result [MSY08]: the valuation wt_{ξ_0} minimizes $\widehat{\mathrm{vol}}$ where ξ_0 is the Reeb vector field of the Ricci-flat Fano cone.

Theorem 3.17 (see [CS15, LX16, LX17a]). *Assume (X,ξ_0) admits a Ricci-flat Kähler cone metric. Then $A_X(\mathrm{wt}_{\xi_0}) = n$ and (X,ξ_0) is K-polystable among all special test configurations.*

Proof. Fix any smooth Kähler cone metric $\sqrt{-1}\partial\bar\partial r^2$ on X. Any special test configuration determines a geodesic ray $\{r_t^2 = r^2e^{\varphi_t}\}_{t>0}$ of Kähler cone metrics. Denote $D(t) = D(\varphi_t)$. Then we have the following formula:

$$\lim_{t\to 0}\frac{D(t)}{-\log|t|^2} = \frac{D_{-\eta}\mathrm{vol}(\xi_0)}{\mathrm{vol}(\xi_0)} - (1-\mathrm{lct}(\mathcal{X},\mathcal{X}_0)) = D^{\mathrm{NA}}(\chi,\xi_0;\eta), \tag{18}$$

which is a combination of two ingredients:

1. The Fano cone version of an identity from Kähler geometry which combined with (15) gives the formula:
$$\lim_{t\to 0}\frac{E(\varphi_t)}{-\log|t|^2} = \frac{D_{-\eta}\mathrm{vol}(\xi_0)}{\mathrm{vol}(\xi_0)}. \tag{19}$$
2. $G(\varphi_t)$ is subharmonic in t (cone version of Berndtsson's result) and its Lelong number at $t=0$ is given by $1-\mathrm{lct}(\mathcal{X},\mathcal{X}_0)$ (cone version of Beman's result).

The other key result is the cone version of Berndtsson's subharmonicity and uniqueness result, which was used to characterize the case of vanishing Futaki invariant. □

Remark 3.18. The argument in [LX17a] gives a slightly more general result: Assume (X, ξ_0) admits a Ricci-flat Käler cone metric, then $A_X(\mathrm{wt}_{\xi_0}) = n$ and (X, ξ_0) is Ding-polystable among $\mathbb{Q}$-Gorenstein test configurations (see Remark 3.12).

4. Stable degeneration conjecture

In this section, we give a conjectural description of minimizers for general klt singularities, and explain various parts of the picture that we can establish.

4.1. Statement

For a klt singularity $x \in (X, D)$, one main motivation to study the minimizer v of $\widehat{\mathrm{vol}}_{(X,D),x}$ is to establish a 'local K-stability' theory, guided by the local-to-global philosophy mentioned in the introduction. In particular, we propose the following conjecture for all klt singularities.

Conjecture 4.1 (Stable Degeneration Conjecture, [Li18a, LX17a]**).** *Given any arbitrary klt singularity $x \in (X = \mathrm{Spec}(R), D)$, there is a unique minimizer v up to rescaling. Furthermore, v is quasi-monomial, with a finitely generated associated graded ring $R_0 =_{\mathrm{defn}} \mathrm{gr}_v(R)$, and the induced degeneration*

$$(X_0 = \mathrm{Spec}(R_0), D_0, \xi_v)$$

is a K-semistable Fano cone singularity. (*See below for the definitions.*)

Let us explain the terminology in more details: First by the grading of R_0, there is a $T \cong \mathbb{C}^r$-action on X_0 where r is the rational rank of v, i.e., the valuative semigroup Φ of v generates a group $M \cong \mathbb{Z}^r$. Moreover, since the valuation v identifies M to a subgroup of $\mathbb{R}$ and sends Φ into $\mathbb{R}_{\geq 0}$, it induces an element in the Reeb cone ξ_v. If R_0 is finitely generated, then [LX17a] shows that we can embed $(x \in X) \subset (0 \in \mathbb{C}^N)$ and find an rational vector $\xi \in \mathfrak{t}^+_{\mathbb{R}} \cap N_{\mathbb{Q}}$ sufficiently close to ξ_v such that the $\mathbb{C}^*$-action generated by ξ degenerates X to X_0 with a good action. We denote by o (or o_{X_0}) the unique fixed point on X_0. Furthermore, the extended Rees algebra yielding the degeneration does not depend on the choice of ξ. So we can define D_0 as the degeneration of D.

Conjecture 4.1, if true, would characterize deep properties of a klt singularity. Various parts are known, see Theorem 4.14. However, the entire picture remains open in general.

4.2. Cone case

The study of the case of cone is not merely verifying a special case. In fact, since the stable degeneration conjecture predicts the degeneration of any klt singularities to a cone, understanding the cone case is a necessary step to attack the conjecture. Here we divide our presentations into two case: the rank one case and the general higher rank case. Although our argument in the higher rank case covers the rank one case with various simplifications, we believe it is easier for reader to first understand the rank one case, as it is equivalent to the more standard K-semistability theory of the base which is a log Fano pair. This connection is made via the theory of β-invariant, which is first introduced in [Fuj18] in terms of ideal sheaves and further developed in [Li17, Fuj16] via valuations.

4.2.1. Rank one case. The rank one Fano cone is just a cone over a log Fano pair. More precisely, let (S,B) be an $(n-1)$-dimensional log Fano pair, and r a positive integer such that $r(K_S+B)$ is Cartier. Then we can consider the minimizing problem of the normalized volume at the vertex of the cone

$$x \in (X,D) = C(S,B;-r(K_S+B)).$$

Such a question was first extensively studied in [Li17]. More precisely, there is a canonical divisorial valuation obtained by blowing up x to get a divisor S_0 isomorphic to S, which yields the degeneration of $x \in (X,D)$ to itself with ξ being the natural rescaling vector field from the cone structure. Therefore, the stable degeneration conjecture predicts $v_{S_0} = \mathrm{ord}_{S_0}$ is a minimizer of $\widehat{\mathrm{vol}}_{(X,D),x}$ if and only if (S,B) is K-semistable, and this is confirmed in [Li17, LL19, LX16].

Theorem 4.2. *The valuation v_{S_0} is a stabilizer of $\widehat{\mathrm{vol}}_{(X,D),x}$ if and only if (S,B) is K-semistable. Moreover, $\widehat{\mathrm{vol}}(S_0) < \widehat{\mathrm{vol}}(E)$ for any other divisor E over x.*

In the below, we will sketch the ideas of two slightly different proofs of Theorem 4.2.

In the first approach, we carry out a straightforward calculation as follows: Given a compactified nontrivial special test configuration $(\mathcal{S},\mathcal{B})$ of (S,B), then we obtain a valuation v^* by restricting the divisorial valuation of the special fiber S_0 to $K(S) \subset K(S\times\mathbb{A}^1)$, which is a multiple of some divisorial valuation (cf. [BHJ17]). Such a valuation v^* pull backs a valuation v^*_X on $K(X)$. Then we define a $\mathbb{C}^*$-valuation on $K(X)$ by $v_\infty(f_m) = v^*_X(f_m) - mra_S(v^*)$ over X for any $f_m \in H^0(S,-mr(K_S+B))$. In other words, $v_\infty = v^*_X - ra_S(v^*)v_{S_0}$, and we know that the induced filtration on R yields the Duistermaat–Heckman (DH) measure of $(\mathcal{S},\mathcal{B})$ (see [BHJ17, Definition 3.5]). We define the ray in

$$\left\{ v_t = v_{S_0} + t\cdot v_\infty \in \mathrm{Val}_{X,x} \mid t \in \left[0, \frac{1}{ra_S(v^*)}\right) \right\}.$$

Then the key computation in [Li17] is that

$$\frac{d}{dt}\widehat{\mathrm{vol}}(v_t)|_{t=0} = \frac{n}{r^n}(-K_S-B)^{n-1}\cdot \mathrm{Fut}(\mathcal{S},\mathcal{B}). \tag{20}$$

In fact, if for any valuation v over S, we denote by $R_m = H^0(S,-mr(-K_S-B))$ and define

$$\mathcal{F}^x_v R_m := \{ f\in R_m |\ \ f \in H^0(S, -mr(-K_S-B)\otimes \mathfrak{a}_x)\},$$

then we easily see

$$\mathfrak{a}_k(v_t) \cap R_m = \mathcal{F}_{v_\infty}^{\frac{k-m}{t}} H^0(S, -mr(-K_S - B)).$$

So

$$\begin{aligned}
\mathrm{vol}(v_t) &= \lim_k \frac{l_{\mathbb{C}}(R/\mathfrak{a}_k(v_t))}{k^n/n!} \\
&= \lim_{k\to\infty} \frac{n!}{k^n} \sum_{m=0} \left(\dim \mathcal{F}^0_{v_\infty} R_m - \dim \mathcal{F}_{v_\infty}^{\frac{k-m}{t}} H^0(S, -mr(-K_S - B))\right) \\
&= -\int_{-\infty}^{\infty} \frac{d\mathrm{vol}(\mathcal{F}_{v_\infty} R^{(x)})}{(1+tx)^n}, \qquad (21)
\end{aligned}$$

where $\mathcal{F}_{v_\infty} R^{(x)} := \bigoplus_m \mathcal{F}^{mx}_{v_\infty} R_m$ and the last equality is obtained by a change of variables (see Lemma [Li17, Lemma 4.5]).

Since $A(S_0) = \frac{1}{r}$ and $A(v_\infty) = 0$, $A_{v_t} = \frac{1}{r}$, so

$$\widehat{\mathrm{vol}}(v_t) = -\left(\frac{1}{r}\right)^n \int_{-\infty}^{\infty} \frac{d\mathrm{vol}(\mathcal{F}_{v_\infty} R^{(x)})}{(1+tx)^n},$$

and this implies that

$$\begin{aligned}
\frac{d}{dt}\widehat{\mathrm{vol}}(v_t)|_{t=0} &= \frac{n}{r^n} \int_{-\infty}^{\infty} x \cdot d\mathrm{vol}(\mathcal{F}_{v_\infty} R^{(x)}) \\
&= \frac{n}{r^n} \lim_{k\to\infty} \frac{w_k}{kN_k} \\
&= -\frac{1}{r^n}(-K_{\mathcal{S}} - \mathcal{B})^n, \\
&= \frac{n}{r^n}(-K_S - B)^{n-1} \cdot \mathrm{Fut}(\mathcal{S}, \mathcal{B}).
\end{aligned}$$

where for the second equality we use that v_∞ is the DH measure for $(\mathcal{S}, \mathcal{B})$.

It is not straightforward to reverse the argument to show that (S, B) is K-semistable implies that ord_{S_0} is a minimizer of $\widehat{\mathrm{vol}}_{(X,D),x}$, since a priori there could be more complicated valuations than those induced by central fibres of test configurations. In particular, originally in [Li17], the techniques of 'taking the limit of a sequence of filtered linear systems' developed in [Fuj18] were used in the case when the associated bigraded ring

$$\bigoplus_{m,k} H^0(S, -rm(K_S + B) \otimes \mathfrak{a}_k)$$

is not finitely generated, and this is enough to treat all $\mathbb{C}^*$-equivariant valuations.

In [LX16], after the MMP method was systematically applied, it was shown that

$$\inf_{v\in \mathrm{Val}_{X,x}} \widehat{\mathrm{vol}}(v) = \{\inf \widehat{\mathrm{vol}}(\mathrm{ord}_S) \mid \ \mathbb{C}^*\text{-equivariant Kollár components } S\} \qquad (22)$$

(see (27) and the discussion below it). Since Kollár components yield special degenerations, therefore, the above arguments can be essentially reversed. See Section 4.2.2.

Remark 4.3. In fact, we establish a one-to-one correspondence between special test configurations of (S, B) (up to a base change) and rays in $\mathrm{Val}_{X,x}$ emanating from v_{S_0} containing a Kollár component (different with v_{S_0}).

An interesting consequence is that the above argument indeed gives an alternative way to show that K-semistability implies the valuative criterion of K-semistability with β-invariant as in [Fuj16, Li17], but without using the arguments of 'taking a limit of filtered linear systems'.

The second approach to treat the cone singularity is developed in [LL19] (see also [LX16]). It is shown that K-semistablity of (S, B) is equivalent to that of $(\bar{X}, \bar{D} + (1 - \frac{1}{rn})S_\infty)$, where $(\bar{X}, \bar{D})$ is the projective cone of (X, D) with respect to $-r(K_X + D)$ and $S_\infty (= S)$ is the divisor at the infinity place. This follows from a straightforward Futaki invariant calculation as in [LX16, Proposition 5.3]. Applying the inequality 5.12 to $x \in (\bar{X}, \bar{D} + (1 - \frac{1}{rn})S_\infty)$, we immediately conclude that

$$\widehat{\mathrm{vol}}(x, \bar{X}, \bar{D}) \geq \frac{(-K_S - B)^{n-1}}{r^n} = \widehat{\mathrm{vol}}_{(X,D),x}(\mathrm{ord}_{S_0}). \tag{23}$$

To understand better the relation between the K-semistability of $(\bar{X}, \bar{D} + (1 - \frac{1}{rn})S_\infty)$ and of (S, B), we want to present a direct calculation which connects the calculation on β-invariant on $(\bar{X}, \bar{D} + \frac{1}{rn}S_\infty)$ and the one on (S, B).

Lemma 4.4. *Assume β-invariant is nonnegative for any divisorial valuation over S. Denote by $\hat{L} = \mathcal{O}(1) = \mathcal{O}(S_\infty)$ and $\delta = \frac{n+1}{rn}$. For any $\mathbb{C}^*$-invariant divisorial valuation E. We have the following*

$$\beta(E) := A_{(\bar{X}, \bar{D} + (1 - \frac{1}{rn})S_\infty)}(E) - \frac{\delta}{\hat{L}^n} \int_0^{+\infty} \mathrm{vol}(\mathcal{F}_{\mathrm{ord}_E} \hat{R}^{(x)}) dx \geq 0, \tag{24}$$

where $\hat{R} = \bigoplus_{m=0}^{+\infty} H^0(\bar{X}, m\hat{L})$.

The key of the proof is to relate the β-invariant for a $\mathbb{C}^*$-invariant valuation v over $\bar{X}$ to the β-invariant of the restriction of v over the base S.

Proof. We have $K_{\bar{X}} + \bar{D} + (1 - \frac{1}{rn})S_\infty = -\frac{n+1}{rn}\hat{L} = -\delta\hat{L}$, and define

$$\mathcal{F}_v^x \hat{R}_m := \{f \in \hat{R}_m | \ \ f \in \hat{R}_m = \oplus_{0 \leq k \leq m} H^0(S, kr(-K_S - B)) \text{ and } v(f) \geq x\},$$

For any $\mathbb{C}^*$-invariant divisorial valuation $v = \mathrm{ord}_E$ on $\bar{X}$, there exists $c_1 \in \mathbb{Z}$, $a \geq 0$ and a divisorial valuation ord_F over S such that for any $f \in H^0(S, mr(-K_S - B))$, we have

$$v(t) = c_1; \text{ and } v(f) = a \cdot \mathrm{ord}_F(f) =: \bar{v}(f).$$

We estimate $\beta(E)$ in three cases depending on the signs of a and c_1:

$(a = 0)$: The valuation v is associated to the canonical $\mathbb{C}^*$-action along the ruling of the cone, up to rescaling, then we easily get $\beta(E) = 0$

$(a > 0$ and $c_1 \geq 0)$: Then the center of v is contained in S_∞. In this case we can easily calculate:

$$\begin{aligned}\mathrm{vol}(\mathcal{F}\hat{R}^{(x)}) &= \lim_{m\to+\infty} \frac{\dim_{\mathbb{C}} \mathcal{F}^{xm}\hat{R}_m}{m^n/n!} = \lim_{m\to+\infty} \frac{1}{m^n/n!} \sum_{k=0}^{m} \dim_{\mathbb{C}} \mathcal{F}_{\bar{v}}^{xm-c_1(m-k)} R_k \\ &= n\int_0^1 \mathrm{vol}(\mathcal{F}_{\bar{v}} R^{(c_1+\frac{x-c_1}{\tau})})\tau^{n-1}d\tau,\end{aligned}$$

where the last identity can be proved in the same way as in (21). So we have:

$$\begin{aligned}\int_0^{+\infty} \mathrm{vol}(\mathcal{F}\hat{R}^{(x)})dx &= n\int_0^{+\infty} dx \int_0^1 \mathrm{vol}(\mathcal{F}_{\bar{v}} R^{(c_1+\frac{x-c_1}{\tau})})\tau^{n-1}d\tau \\ &= n\int_0^1 \tau^{n-1}d\tau \int_0^{+\infty} \mathrm{vol}(\mathcal{F}_{\bar{v}} R^{(c_1+\frac{x-c_1}{\tau})})dx \\ &= n\int_0^1 \tau^{n-1}d\tau \left[H^{n-1}c_1(1-\tau) + \tau\int_{c_1}^{+\infty} \mathrm{vol}\left(\mathcal{F}_{\bar{v}} R^{(y)}\right) dy\right] \\ &= \frac{c_1}{n+1} + n\int_0^{+\infty} \mathrm{vol}(\mathcal{F}_{\bar{v}} R^{(x)})dx \int_0^1 \tau^n d\tau \\ &= \frac{c_1}{n+1} + \frac{n}{n+1}\int_0^{+\infty} \mathrm{vol}(\mathcal{F}_{\bar{v}} R^{(x)})dx.\end{aligned}$$

On the other hand, we have $H^{n-1} = \hat{L}^n$ and:

$$A_{(\bar{X},\bar{D}+(1-\beta)S_\infty)}(\mathrm{ord}_E) = A_{(S,B)}(\bar{v}) + c_1 - (1-\beta)c_1 = A_{(S,B)}(\bar{v}) + \frac{c_1}{rn}.$$

So we get:

$$\begin{aligned}\beta(E) &= A_{(S,B)}(\bar{v}) + \frac{c_1}{rn} - \frac{\frac{n+1}{rn}}{H^{n-1}}\frac{n}{n+1}\left(\frac{c_1}{n+1} + \int_0^{+\infty} \mathrm{vol}(\mathcal{F}_{\bar{v}} R^{(x)})dx\right) \\ &= A_{(S,B)}(\bar{v}) - \frac{1}{rH^{n-1}}\int_0^{+\infty} \mathrm{vol}(\mathcal{F}_{\bar{v}} R^{(x)})dx = \beta(\bar{v}),\end{aligned}$$

which is nonnegative by our assumption.

$(a > 0$ and $c_1 < 0)$: In this case, the center of v is at the vertex. As a consequence we have:

$$\begin{aligned}A_{(\bar{X},\bar{D}+(1-\beta)S_\infty)}(v) &= A_{(S,B)}(\bar{v}) + (-c_1) + \left(\frac{1}{r} - 1\right)(-c_1) \\ &= A_{(S,B)}(\bar{v}) + \frac{-c_1}{r} \geq A_{(S,B)}(\bar{v}).\end{aligned}$$

The similar calculation as in the second case shows that $\beta(E) \geq \beta(\bar{v})$. $\square$

Finally, to show $\widehat{\mathrm{vol}}(S_0) < \widehat{\mathrm{vol}}(E)$ for $E \neq S_0$, in [LX16], it was first proved that if E is a minimizer then it has to be a $\mathbb{C}^*$-equivariant Kollár component. Then a careful study of the geometry of E using the equality condition in (23) implies $E = S$. This is similar to the analysis for the equality case in [Fuj18, Liu18] where they showed that the K-stable $\mathbb{Q}$-Fano variety with the maximal volume $(n+1)^n$ can only be $\mathbb{CP}^n$. We will leave the discussion on this uniqueness type result to the general case of cones of higher rational ranks, where we take a somewhat different approach, using more convex geometry.

Remark 4.5. It is worthy pointing out that there is another global invariant for an n-dimensional log Fano pair (S,B), defined as

$$\delta(S,B)=\inf_{v\in \mathrm{Val}_S}\frac{A_{(S,B)}(v)\cdot(-K_S-B)^n}{\int_0^\infty \mathrm{vol}(-K_S-B-tv)dt}$$

(see [FO16, BJ17]). δ-invariant shares lots of common properties with the normalized volume. For example, the existence of minimizers were proved using similar strategy. They both have differential geometric meanings. The minimizer of $\widehat{\mathrm{vol}}$ is related to the metric tangent cone (see Section 5.2); while the valuation on $K(S)$ yielding $\delta(S,B)$ is related to the existence of twisted Kähler–Einstein metrics (see [BJ18]).

For a log Fano pair (S,B) and a cone $x\in(X,D)=C(S,B;-r(K_S+B))$, if (S,B) is not K-semistable, or equivalently $\delta=\delta(S,B)<1$, then we have

$$\widehat{\mathrm{vol}}(x,X,D)\geq\frac{\delta^n\cdot(-K_S-B)^{n-1}}{r^n}.$$

This follows from our second proof by looking at $(\bar{X},\bar{D}+(1-\beta)S_\infty)$ and applying the inequality [BJ17, Theorem D] which can be written as

$$(K_{\bar{X}}+\bar{D}+(1-\beta)S_\infty)^n\leq\frac{(n+1)^n}{n^n}\cdot\widehat{\mathrm{vol}}(x,X,D)\cdot\bar{\delta}^n,$$

where $\bar{\delta}:=\delta(\bar{X},\bar{D}+(1-\beta)S_\infty)$. We claim $\min\{\bar{\delta},1\}=\delta$. In fact, by the argument in [BJ17, Section 7], we know that $\bar{\delta}$ is computed by a $\mathbb{C}^*$-invariant valuation and the claim follows from the calculation in the proof of Lemma 4.4.

4.2.2. Log Fano cone in general. We proceed to investigate a log Fano cone $o\in(X,D,\xi)$ where the torus T could have dimension larger than one. However, we consider not only the valuations in $\mathfrak{t}^+_{\mathbb{R}}(X)$ coming from the torus as in [MSY08] (see Section 3.1) but all valuations in $\mathrm{Val}_{X,o}$. Compared to the proof of Theorem 4.2, for the higher rational rank case, we rely more on the construction of Kollár components coming from the birational geometry. More explicitly, we use the relation between special test configurations and Kollár components (see [LX16, 2.3] and [LX17a, 3.1]).

By the results from the MMP (see (27) and the explanation below), to show a valuation is a minimizer in $\mathrm{Val}_{X,x}$, we only need to show its normalized volume is not greater than that of any T-invariant Kollár component. On the other hand, any T-equivariant Kollár component E in $\mathrm{Val}_{X,o}$ yields a special test configuration of $(\mathcal{X},\mathcal{D},\xi;\eta)$ of (X,D) such that $-\eta\in\mathfrak{t}^+_{\mathbb{R}}(X_0)$ and the valuation associated to $-\eta$ coincides with ord_E. We denote by (X_0,D_0) the fiber with a cone vertex o. Then we can compare the volumes as $\widehat{\mathrm{vol}}_X(\xi)=\widehat{\mathrm{vol}}_{X_0}(\xi)$ and $\widehat{\mathrm{vol}}_X(E)=\widehat{\mathrm{vol}}(-\eta)$. Since

$\xi, -\eta \in \mathfrak{t}^+_{\mathbb{R}}(X_0)$ we reduce the question to the set up of [MSY08] on X_0. Then we only need to each time treat one degeneration X_0 and try to understand how to pass properties between X_0 and X.

With this strategy, we can show the following generalization of Theorem 4.2.

Theorem 4.6 ([LX17a]). *Let $x \in (X, D, \xi)$ be a log Fano cone singularity. Then v_ξ is a minimizer of $\widehat{\mathrm{vol}}_{(X,D),x}$ if and only if (X, D, ξ) is K-semistable. In such case, $\widehat{\mathrm{vol}}(v_\xi) < \widehat{\mathrm{vol}}(v)$ for any quasi-monomial valuation v if v is not a rescaling of v_ξ.*

If (X,D,ξ) is K-semistable, then for each special test configuration $(\mathcal{X},\mathcal{D},\xi;\eta)$, on X_0, we can consider the ray $\xi_t = \xi - t\eta$ for $t \in [0, \infty)$. We know

$$\frac{d}{dt}\widehat{\mathrm{vol}}_{(X_0,D_0),o}(v_{\xi_t})|_{t=0} = c \cdot \mathrm{Fut}(\mathcal{X}, \mathcal{D}, \xi; \eta) \geq 0.$$

Moreover, when $(X_0, D_0, o) = (X_0, \emptyset, o)$ is an isolated singularity, it was shown in [MSY08] that $\widehat{\mathrm{vol}}(v_{\xi_t})$ is a convex function. We obtain a stronger result for any log Fano cone (X_0, D_0, ξ_0) (see Section 4.2.3). In particular, we conclude that $\widehat{\mathrm{vol}}(v_{\xi_t})$ is an increasing function of t, and its limit is $\widehat{\mathrm{vol}}(-\eta)$, thus the inequality in the following relation holds true:

$$\widehat{\mathrm{vol}}_{(X,D),x}(\xi) = \widehat{\mathrm{vol}}_{(X_0,D_0),o}(\xi) \leq \widehat{\mathrm{vol}}_{(X_0,D_0),o}(-\eta) = \widehat{\mathrm{vol}}_{(X,D),x}(E).$$

The first identity consists of two identities: $A_{(X,D)}(v_\xi) = A_{(X_0,D_0)}(v_\xi)$ and $\mathrm{vol}_X(v_\xi) = \mathrm{vol}_{X_0}(v_\xi)$, which essentially follow from the flatness of T-equivariant test configuration (see [LX17a, Lemma 3.2]). The last identity is because $v_{-\eta} = \mathrm{ord}_E$.

This argument is reversible since we can indeed attach to any special test configuration such a set of valuations (see Remark 4.3): if we consider the valuation w_t obtained by considering the vector field ξ_t as a valuation on $K(\mathcal{X})$ and then take its restriction on $K(X)$. The corresponding degeneration induces the test configuration. See [LX16, 6] and [LX17a, 4.2] for more details.

4.2.3. Uniqueness. We have seen the convexity of the normalized volume function in the Reeb cone plays a key role. In [MSY08], the strict convexity on the normalized function is established for the valuation varying inside the Reeb cone for an isolated singularity. This is the kind of property we need for the uniqueness of the minimizer of a K-semistable Fano cone singularity (X, D, ξ). However, as we do not know the associated graded ring of other minimizer is finitely generated, we cannot degenerate two minimizers into the Reeb cone. Thus we need develop a technique to deal with valuations outside the Reeb cone.

The idea of the argument in [LX17a, Section 3.2] is to use the theory of Newton–Okounkov bodies which was first developed in [LM09, KK12]) and in the local setting in [Cut13, KK14]. This is a theory which realizes the volumes in algebraic geometry with an asymptotic nature to the Euclidean volumes of some convex bodies in $\mathbb{R}^n$. So our aim is to apply the Newton–Okounkov body construction to translate the normalized volume of valuations into the volume of convex bodies, and then invoke a convexity property of the volumes functions known in the latter setting.

To start, we first need to set a valuation $\mathbb{V}$ with $\mathbb{Z}^n$-valued valuation, which sends the elements in R to the lattice points inside a convex region $\tilde{\sigma}$, so that later we can realize the normalized volumes of valuations as the volume of subsets in $\tilde{\sigma}$.

For any fixed $T \cong (\mathbb{C}^*)^r$-equivariant quasi-monomial valuation μ, we know it is of the form $(\xi_\mu, v^{(0)})$ where $\xi_\mu \in M_{\mathbb{R}}$ and $v^{(0)}$ is a quasi-monomial valuation over $K(Y)$, such that for any function $f \in R_u$,

$$\mu(f) = \langle \xi_\mu, u\rangle + v^{(0)}(f)$$

(see Theorem 3.5(1)). We fix a lexicographic order on $\mathbb{Z}^r$ and define for any $f \in R$,

$$\mathbb{V}_1(f) = \min\{u; f = \sum_u f_u \text{ with } f_u \neq 0\} = \mathbb{V}_1(f),$$

i.e., the first factor $\mathbb{V}_1$ comes from the toric part of μ.

We extend this $\mathbb{Z}^r$-valuation $\mathbb{V}_1$ to become a $\mathbb{Z}^n$-valued valuation in the following way: Denote $u_f = \mathbb{V}_1(f) \in \sigma^\vee$ and f_{u_f} the corresponding nonzero component. Define $\mathbb{V}_2(f) = v^{(0)}(f_{u_f})$. Because $\{\beta_i\}$ are $\mathbb{Q}$-linearly independent, we can write $\mathbb{V}_2(f) = \sum_{i=1}^s m_i^* \beta_i$ for a uniquely determined $m^* := m^*(f_{u_f}) = \{m_i^* := m_i^*(f_{u_f})\}$. Moreover, the Laurent expansion of f has the form:

$$f_{u_f} = z_1^{m_1^*} \cdots z_s^{m_s^*} \chi_{m^*}(z'') + \sum_{m \neq m^*} z_1^{m_1} \cdots z_s^{m_s} \chi_m(z''). \tag{25}$$

Then $\chi_{m^*}(z'')$ in the expansion of (25) is contained in $\mathbb{C}(Z)$, where on some model of Y, we have $Z = \{z_1 = 0\} \cap \cdots \{z_s = 0\} = D_1 \cap \cdots \cap D_s$ is the center of $v^{(0)}$.

Extend the set $\{\beta_1, \ldots, \beta_s\}$ to $d = n - r$ $\mathbb{Q}$-linearly independent positive real numbers $\{\beta_1, \ldots, \beta_s; \gamma_1, \ldots, \gamma_{d-s}\}$. Define $\mathbb{V}_3(f) = w_\gamma(\chi_{m^*}(z''))$ where w_γ is the quasi-monomial valuation with respect to the coordinates z'' and the $(d-s)$ tuple $\{\beta_1, \ldots, \beta_s; \gamma_1, \ldots, \gamma_{d-s}\}$.

Now we assign the lexicographic order on

$$\mathbb{G} := \mathbb{Z}^r \times G_2 \times G_3 \cong \mathbb{Z}^r \times \mathbb{Z}^s \times \mathbb{Z}^{n-r-s}$$

and define $\mathbb{G}$-valued valuation:

$$\mathbb{V}(f) = (\mathbb{V}_1(f), \mathbb{V}_2(f_{u_f}), \mathbb{V}_3(\chi_{m^*})). \tag{26}$$

Let $\mathcal{S}$ be the valuative semigroup of $\mathbb{V}$. Then $\mathcal{S}$ generates a cone $\tilde{\sigma}$ which is the one we are looking for. We also let $P_1 : \mathbb{R}^n \to \mathbb{R}^r$, $P_2 : \mathbb{R}^n \to \mathbb{R}^s$ and $P = (P_1, P_2) : \mathbb{R}^n \to \mathbb{R}^{r+s}$ be the natural projections. Then $P_1(\tilde{\sigma}) = \sigma \subset \mathbb{R}^r$.

To continue, we consider how to construct some subsets $\Delta_{\tilde{\Xi}_t} \subset \tilde{\sigma}$ whose Euclidean volume is the same as the normalized volumes of the valuations. For any $\xi \in \mathrm{int}(\sigma)$, denote by wt_ξ the valuation associated to ξ. We can connect wt_ξ and μ by a family of quasi-monomial valuations: $\mu_t = ((1-t)\xi + t\xi_\mu, tv^{(0)})$ defined as

$$\mu_t(f) = tv^{(0)}(f) + \langle u, (1-t)\xi + t\xi_\mu\rangle \quad \text{for any } f \in R_u.$$

So the vertical part of μ_t corresponds to the vector $\Xi_t := ((1-t)\xi+t\xi_\mu, t\beta) \in \mathbb{R}^{r+s}$. Extend Ξ_t to $\tilde{\Xi}_t := (\Xi_t, 0) \in \mathbb{R}^n$ and define the following set:

$$\Delta_{\tilde{\Xi}_t} = \left\{ y \in \tilde{\sigma}; \langle y, \tilde{\Xi}_t \rangle \leq 1 \right\} = \{ y \in \tilde{\sigma}; \langle P(y), \Xi_t \rangle \leq 1 \}.$$

Because $\widehat{\mathrm{vol}}$ is rescaling invariant, we can assume $A_{(X,D)}(v) = A_{(X,D)}(\xi) = 1$. Then by the T-invariance of v_t, we easily get:

$$A(v_t) = tA(v^{(0)}) + A_{(X,D)}((1-t)\xi + t\zeta) = tA_{(X,D)}(v) + (1-t)A_{(X,D)}(\xi) \equiv 1.$$

The Newton–Okounkov body theory implies that we have

$$\widehat{\mathrm{vol}}(v_t) = \mathrm{vol}(v_t) = \mathrm{vol}(\Delta_{\tilde{\Xi}_t}).$$

To finish the uniqueness argument, now we only need to look at the convex geometry of $\Delta_{\tilde{\Xi}_t}$. We note that $\tilde{\Xi}_t$ is linear with respect to t, and each region $\Delta_{\tilde{\Xi}_t}$ is cut out by a hyperplane H_t on the convex cone $\tilde{\sigma}$. Moreover, all H_t passes through a fixed point. A key result from convex geometry then shows that $\phi(t) := \mathrm{vol}(\Delta_{\tilde{\Xi}_t})$ is strictly convex as a function of $t \in [0,1]$ (see [MSY06, Gig78]). By the assumption $\phi(0) = \mathrm{vol}(v_0) = \widehat{\mathrm{vol}}(\mathrm{wt}_\xi)$ is a minimum. So the strict convexity implies

$$\phi(1) = \mathrm{vol}(\Delta_{\tilde{\Xi}_1}) = \widehat{\mathrm{vol}}(v) > \widehat{\mathrm{vol}}(\mathrm{wt}_\xi) = \phi(0).$$

4.3. Results on the general case

To treat the general case, the key idea, suggested by the degeneration conjecture, is to understand how an arbitrary klt singularity can be degenerated to a K-semistable Fano cone singularity. In [LX16], by localizing the setting of [LX14], the following approach of using Kollár components is developed.

From each ideal $\mathfrak{a}$, we can take a dlt modification of

$$f\colon (Y, D_Y) \to (X, D + \mathrm{lct}(X, D; \mathfrak{a}) \cdot \mathfrak{a}),$$

where $D_Y = f_*^{-1}D + \mathrm{Ex}(f)$ and for any component $E_i \subset \mathrm{Ex}(f)$ we have

$$A_{X,D}(E) = \mathrm{lct}(X, D; \mathfrak{a}) \cdot \mathrm{mult}_E f^*\mathfrak{a}.$$

There is a natural inclusion $\mathcal{D}(D_Y) \subset \mathrm{Val}^{=1}_{X,x}$, and using a similar argument as in [LX14], we can show that there exists a Kollár component S whose rescaling in $\mathrm{Val}^{=1}_{X,x}$ contained in $\mathcal{D}(D_Y)$ satisfies that

$$\widehat{\mathrm{vol}}(\mathrm{ord}_S) = \mathrm{vol}^{\mathrm{loc}}(-A_{X,D}(S) \cdot S) \leq \mathrm{vol}^{\mathrm{loc}}(-K_Y - D_Y) \leq \mathrm{mult}(\mathfrak{a}) \cdot \mathrm{lct}^n(X, D; \mathfrak{a}).$$

Here $\mathrm{vol}^{\mathrm{vol}}(\cdot)$ is the local volume of divisors over X as defined in [Ful13]. Then Theorem 2.6 immediately implies that

$$\widehat{\mathrm{vol}}(x, X, D) = \inf\{\widehat{\mathrm{vol}}(\mathrm{ord}_S) |\ S \text{ is a Kollár component over } x\}. \qquad (27)$$

Moreover, if $x \in (X, D)$ admits a torus group T-action, then by degenerating to the initial ideals, as the colengths are preserved and the log canonical thresholds may only decrease, the infimum of the normalized multiplicities in Theorem 2.6

can be only run over all T-equivariant ideals. Then the equivariant MMP allows us to make all the above data Y and S be T-equivariant.

In case a minimizer is divisorial, then the above discussion shows that

Lemma 4.7 ([LX16, Blu18]). *A divisorial minimizer of* $\widehat{\mathrm{vol}}_{X,D}$ *yields a Kollár component.*

In general, we know that the minimizer is a limit of a rescaling of Kollár components (see [LX16]). So understanding the limiting process is crucial. When the minimizer is quasi-monomial v of rational rank r, i.e., the valuation v is étale locally a monomial valuation with respect to a log resolution $(Y, E) \to X$, then a natural candidates will be the valuations given by taking rational approximations of the monomial coordinates $\alpha \in \mathbb{R}^r_{>0}$.

Our first observation in [LX17a] is using MMP results including the ACC of log canonical thresholds, we could construct a weak log canonical model which extracts divisors whose coordinates are good linear Diophantine approximations of the coordinates of v.

Proposition 4.8. *For any quasi-monomial valuation v computing a log canonical threshold of a graded sequence of ideals, we can find a sequence of divisors $S_1, \ldots, S_r$, such that*

1. *there is a model $Y \to X$ which precisely extracts $S_1, \ldots, S_r$ over x,*
2. *there exists a component Z of $\cap_{i=1}^r S_i$ such that $(Y, E := \sum_{i=1}^r S_i)$ is toroidal around the generic point $\eta(Z)$,*
3. *v is étale locally a monomial valuation over $\eta(Z)$ with respect to (Y, E) (see Section 2.1),*
4. *(Y, E) is log canonical, and $-K_Y - E$ is nef.*

Fix the first model $Y_0 = Y$, then one can construct a sequence of models (Y_j, E_j) satisfying Proposition 4.8 such that a suitable rescaling of the components of E_j become closer and closer to v. To make the notation easier, we rescale v into $\mathrm{Val}^{=1}_{X,x}$. Similarly, we can embed the dual complex of a dlt modification of (Y_j, E_j) into $\mathrm{Val}^{=1}_{X,x}$ (see [dFKX17]). Our construction moreover satisfies that

$$\mathcal{DR}(Y_0, E_0) \supset \mathcal{DR}(Y_1, E_1) \supset \cdots$$

Then the above discussion indeed implies that

Lemma 4.9. *A quasi-monomial minimizer $v \in \mathrm{Val}^{=1}_{X,x}$ can be written as a limit of $c_j \cdot \mathrm{ord}_{S_j} \in \mathcal{DR}(Y_j, E_j)$ where S_i are Kollár components.*

It would be natural to expect that $c_j \cdot \mathrm{ord}_{S_j}$ is indeed contained in the simplex $\sigma_{\eta(Z)} \subset \mathrm{Val}^{=1}_{X,x}$ which corresponds to all the monomial valuations in $\mathrm{Val}^{=1}_{X,x}$ over $\eta(Z)$ with respect to (Y, E). However, for now we cannot show it.

If we further assume $R_0 = \mathrm{gr}_v(R)$ is finitely generated, then we have the following

Proposition 4.10. *If $R_0 = \mathrm{gr}_v(R)$ is finitely generated, then $\mathrm{gr}_v(R) \cong \mathrm{gr}_{v_i}(R)$ for any $v_i \in \sigma_{\eta(Z)}$ sufficiently close to v.*

This immediately implies that $(X_0 := \mathrm{Spec}(R_0), D_0)$ is semi-log-canonical (slc). The final ingredient we need is the following

Proposition 4.11. *Under the above assumptions on (X, D) and its quasi-monomial minimizer v, then ξ_v is a minimizer of (X_0, D_0). In particular,*

$$\widehat{\mathrm{vol}}(x, X, D) = \widehat{\mathrm{vol}}(o, X_0, D_0).$$

Proof. We claim that ξ_v is indeed a minimizer of $\widehat{\mathrm{vol}}_{X_0,D_0}$. If not, we can find a degeneration (Y, D_Y, ξ_Y) induced by an irreducible anti-ample divisor E over $o' \in X_0$ with

$$\widehat{\mathrm{vol}}_Y(\xi_E) = \widehat{\mathrm{vol}}_{X_0}(\mathrm{ord}_E) < \widehat{\mathrm{vol}}_{X_0}(\xi_v) = \widehat{\mathrm{vol}}_Y(\xi_Y).$$

This is clear by our discussion when (X_0, D_0) is klt. The same thing still holds when the model extracting S_j is only log canonical but not plt, which implies that (X_0, D_0) is semi-log-canonical but not klt. In fact, denote by $(X_0^{\mathrm{n}}, D_0^{\mathrm{n}}) \to (X_0, D_0)$ the normalization, then Lemma 4.13 implies that

$$\widehat{\mathrm{vol}}(o', X_0, D_0) := \sum_{o_i \to o'} \widehat{\mathrm{vol}}(o_i, X_0^{\mathrm{n}}, D_0^{\mathrm{n}}) = 0$$

in this case. The argument in [LX17a, Lemma 4.13] then says in this case, we can still extract an equivariant anti-ample irreducible divisor E over $o' \in X_0$ with $\widehat{\mathrm{vol}}(\mathrm{ord}_E)$ arbitrarily small.

Then Lemma 4.12 shows that we can construct a degeneration from (X, D) to (Y, D_Y) and a family of valuations $v_t \in \mathrm{Val}_{X,x}$ for $t \in [0, \epsilon]$ (for some $0 < \epsilon \ll 1$), with the property that

$$\widehat{\mathrm{vol}}_X(v_t) = \widehat{\mathrm{vol}}_Y(\xi_Y - t\eta) < \widehat{\mathrm{vol}}_Y(\xi_Y) = \widehat{\mathrm{vol}}_{X_0}(\xi_v) = \widehat{\mathrm{vol}}_X(v),$$

where for the second inequality, we use again the fact that $\widehat{\mathrm{vol}}_Y(\xi_Y - t \cdot \eta)$ is a convex function in this setting as well. But this is a contradiction. □

Lemma 4.12. *Let $(x \in X) \subset (0 \in \mathbb{C}^N)$ be a closed affine variety. If $\lambda_1 \in \mathbb{N}^N$ is a coweight of $(\mathbb{C}^*)^N$ which gives an action degenerating X to X_0 when $t \to 0$, and $\lambda_2 \in \mathbb{N}^N$ degenerates X_0 to Y when $t \to 0$, then for $k \in \mathbb{N}$ sufficiently large, $k\lambda_1 + \lambda_2$ degenerates X to Y_0.*

The proof was essentially given in [LX16, Section 6] (see also [LWX18, Lemma 3.1]) and uses some argument in the study of toric degenerations (see, e.g., [And13, Section 5]).

Lemma 4.13. *If $o \in (X, D)$ is an lc but not klt point, then*

$$\widehat{\mathrm{vol}}(o, X, D) := \inf_{v \in \mathrm{Val}_{X,o}} \widehat{\mathrm{vol}}(v) = 0.$$

Proof. Let $\pi^{\mathrm{dlt}}\colon (X^{\mathrm{dlt}}, D^{\mathrm{dlt}}) \to (X, D)$ be a dlt modification and pick o'' a preimage of o under π^{dlt}, then $\widehat{\mathrm{vol}}(o'', X^{\mathrm{dlt}}, D^{\mathrm{dlt}}) \geq \widehat{\mathrm{vol}}(o, X, D)$, thus we can assume $(X^{\mathrm{dlt}}, D^{\mathrm{dlt}})$ is dlt $\mathbb{Q}$-factorial.

By specializing a sequence of points, and applying Theorem 2.11, we can assume $o \in (X, D)$ is a point on a smooth variety with a smooth reduced divisor D. Now we can take a weighted blow up of $(1, \epsilon, \ldots, \epsilon)$ where the first coordinate yields D. Then the exceptional divisor E has its normalized volume

$$\widehat{\mathrm{vol}}(E) = \frac{(n-1)^n \epsilon^n}{\epsilon^{n-1}} = (n-1)^n \epsilon \to 0 \text{ as } \epsilon \to 0. \qquad \square$$

This implies that (X_0, D_0) is klt and (X_0, D_0, ξ_v) is a K-semistable Fano cone. To summarize, we have shown Part (a) in the following theorem which characterize what we know about the Stable Degeneration Conjecture 4.1 for a general klt singularity.

Theorem 4.14 ([LX17a, Theorem 1.1])**.** *Let $x \in (X, D)$ be a klt singularity. Let v be a quasi-monomial valuation in $\mathrm{Val}_{X,x}$ that minimizes $\widehat{\mathrm{vol}}_{(X,D)}$ and has a finitely generated associated graded ring $\mathrm{gr}_v(R)$ (which is always true if the rational rank of v is one by Lemma* 4.7)*. Then the following properties hold:*

(a) *The degeneration $\big(X_0 =_{\mathrm{defn}} \mathrm{Spec}\big(\mathrm{gr}_v(R)\big), D_0, \xi_v\big)$ is a K-semistable Fano cone, i.e., v is a K-semistable valuation;*

(b) *Let v' be another quasi-monomial valuation in $\mathrm{Val}_{X,x}$ that minimizes $\widehat{\mathrm{vol}}_{(X,D)}$. Then v' is a rescaling of v.*

Conversely, any quasi-monomial valuation that satisfies (a) *above is a minimizer.*

Proof. We first show the uniqueness in general, under the assumption that it admits a degeneration (X_0, D_0, ξ_v) given by a K-semistable minimizer v. For another quasi-monomial minimizer v' of rank r', by a combination of the Diophantine approximation and an MMP construction including the application of ACC of log canonical thresholds (see Proposition 4.8), we can obtain a model $f\colon Z \to X$ which extracts r' divisors E_i $(i = 1, \ldots, r')$ such that $(Z, D_Z =_{\mathrm{defn}} \sum E_i + f_*^{-1} D)$ is log canonical. Moreover, the quasi-monomial valuation v' can be computed at the generic point of a component of the intersection of E_i, along which (Z, D_Z) is toroidal. Then with the help of the MMP, one can show $Z \to X$ degenerates to a birational morphism $Z_0 \to X_0$. Moreover, there exists a quasi-monomial valuation w computed on Y_0 which can be considered as a degeneration of v' with

$$\widehat{\mathrm{vol}}_{X_0}(w) = \widehat{\mathrm{vol}}_X(v') = \widehat{\mathrm{vol}}_X(v) = \widehat{\mathrm{vol}}_{X_0}(\xi_v).$$

Thus $w = \xi_v$ by Section 4.2.3 after a rescaling. Since $w(\mathbf{in}(f)) \geq v'(f)$ and $\mathrm{vol}(w) = \mathrm{vol}(v')$, we may argue this implies

$$\xi_v(\mathbf{in}(f)) = v'(f)$$

(see [LX17a, Section 4.3]). Therefore, v' is uniquely determined by ξ_v.

To show the last statement, we already know it for a cone singularity. For a valuation v on a general singularity X such that the degeneration (X_0, D_0, ξ_v) is K-semistable, since the degeneration to the initial ideal argument implies that $\widehat{\mathrm{vol}}(x, X, D) \geq \widehat{\mathrm{vol}}(o, X_0, D_0)$, then

$$\widehat{\mathrm{vol}}_X(v) = \widehat{\mathrm{vol}}_{X_0}(\xi_v) = \widehat{\mathrm{vol}}(o, X_0, D_0)$$

is equal to $\widehat{\mathrm{vol}}(x, X, D)$. $\square$

So in other words, the stable degeneration conjecture precisely predicts the following two sets coincide:

$$\Big\{ \text{ Minimizers of } \widehat{\mathrm{vol}} \Big\} \longleftrightarrow \Big\{ K\text{-semistable valuations } \Big\}.$$

Theorem 2.8 and Theorem 4.14 together imply the existence of the left-hand side and the uniqueness of the right-hand side, as well as the direction that any K-semistable valuation is a minimizer.

Finally, let us conclude this section with the two-dimensional case.

Theorem 4.15. *Let (X, D, x) be a two-dimensional log terminal singularity. The Stable Degeneration Conjecture* 4.1 *holds for (X, D). Moreover, if D is a $\mathbb{Q}$-divisor, then the minimizer of* $\widehat{\mathrm{vol}}_{(X,D)}$ *is always divisorial.*

Proof. We first consider the case when $X = \mathbb{C}^2$. Let v_* be a minimizer and denote $\mathfrak{a}_\bullet = \{\mathfrak{a}_m(v_*)\}_{m\in\mathbb{N}}$. Then it was known that v_* computes the log canonical threshold of $(X, D + \mathfrak{a}_\bullet)$. By a similar argument as in [JM12], we know that v_* must be quasi-monomial.

If v_* is divisorial, then we know that the associated divisor is a Kollár component. Otherwise, v_* satisfies $\mathrm{rat.rk.}(v_*) = 2$ and $\mathrm{tr.deg.}(v_*) = 0$. From the description of valuations on $\mathbb{C}^2$ using *sequences of key polynomials* (SKP), it was showed that the valuative semigroup Γ of v_* is finitely generated (see [FJ04, Theorem 2.28]). Since the residual field of v_* is $\mathbb{C}$, we know that $\mathrm{gr}_{v_*} R \cong \mathbb{C}[\Gamma]$, which is finitely generated. By [LX17a], we know that v_* is indeed the unique minimizer of $\widehat{\mathrm{vol}}$ (up to scaling) which is a K-semistable valuation.

If D is a $\mathbb{Q}$-divisor and v_* is not divisorial, then the pair (X_0, D_0) is a $\mathbb{Q}$-Gorenstein toric pair with $\mathbb{Q}$-boundary toric divisor and the associated Reeb vector field ξ_{v_*} solves the convex geometric problem. But in dimension two case (i.e., on the plane), it is easy to see that the corresponding convex geometric problem as discussed in section 4.2.3 for toric valuations always has a rational solution. This is a contradiction to v_* being non-divisorial.

More generally, we know that $X = \mathbb{C}^2/G$ where G is a finite group acting on $\mathbb{C}^2$ without pseudo-reflections. Consider the covering $(\mathbb{C}^2, \tilde{D}, 0) \to (X, D, x)$. Then by the above discussion, there exists a unique minimizer v_* of $\widehat{\mathrm{vol}}_{(\mathbb{C}^2, \tilde{D}, 0)}$. In particular, v_* is invariant under the G-action. So it descends to a minimizer of $\widehat{\mathrm{vol}}_{(X,D,x)}$ which is quasi-monomial and has a finitely generated associated graded ring. $\square$

5. Applications

In this section, we give some applications of the normalized volume. We have seen that the normalized volume question of a cone singularity is closely related the K-semistability of the base. Another situation where singularities naturally appear is on the limit of smooth Fano manifolds.

5.1. Equivariant K-semistability of Fano

An interesting application of the minimizing theory is to treat the equivariant K-semistability.

Definition 5.1. A log Fano pair (S, B) with a G-action is called G-equivariant K-semistable, if for any G-equivariant test configuration $(\mathcal{S}, \mathcal{B})$, the generalized Futaki invariant $\mathrm{Fut}(\mathcal{S}, \mathcal{B}) \geq 0$. We can similarly define G-equivariant K-polystability.

The notion of usual K-(semi,poly)stability trivially implies the equivariant one. It is a natural question to ask whether they are equivalent, and if it is confirmed it will reduce the problem of verifying K-stability into a much simpler ones if the log Fano pair carries a large symmetry. When S is smooth and $B = 0$, this is proved in [DS16], using an analytic argument. Here we want to explain how our approach can give a proof of such an equivalence when $G = T$ is a torus group.

The key is the fact we obtain in (22) and (27) : let $x \in (X, D)$ be a klt singularity which admits a T action for a torus group T, then

$$\inf_{v \in \mathrm{Val}_{X,x}} \widehat{\mathrm{vol}}(v) = \{\inf(\widehat{\mathrm{vol}}(\mathrm{ord}_S)) |\ T\text{-equivariant Kollár components } S\}. \tag{28}$$

So if (S, B) is not K-semistable, by Theorem 4.2, we know that over the cone $x \in (X, D)$, the valuation ord_{S_∞} obtained by the canonical blow up does not give a minimizer. By (28), there exists a T-equivariant valuation v such that $\widehat{\mathrm{vol}}(v) < \widehat{\mathrm{vol}}(\mathrm{ord}_{S_\infty})$. So we can find a T-equivariant Kollár component S such that $\widehat{\mathrm{vol}}(\mathrm{ord}_S) < \widehat{\mathrm{vol}}(\mathrm{ord}_{S_\infty})$. Then arguing as before, we can find a T-equivariant test configuration $(\mathcal{S}, \mathcal{B})$ with $\mathrm{Fut}(\mathcal{S}, \mathcal{B}) < 0$.

To prove a similar statement for K-polystability is more delicate. Assume a K-semistable log pair (S, B) admits a test configuration $(\mathcal{S}, \mathcal{B})$ with $\mathrm{Fut}(\mathcal{S}, \mathcal{B}) = 0$. We still take the cone construction of a K-semistable log Fano pair as before. The special test configuration determines a ray v_t of valuations in $\mathrm{Val}_{X,x}$, emanating from the canonical component $v_0 = \mathrm{ord}_{S_\infty}$. Using the fact that the Futaki invariant is 0, a minimal model program argument shows that this implies for $t \ll 1$, v_t is automatically $\mathbb{C}^*$-equivariant, which immediately implies the test configuration is $\mathbb{C}^*$-equivariant. Therefore we show the following result (also see [CS16] for an earlier attempt).

Theorem 5.2 ([LX16, LWX18]). *The K-semistability (resp. K-polystability) of a log Fano pair (S, B) is equivalent to the T-equivariant K-semistablity (T-equivariant K-polystablity) for any torus group T acting on (S, B).*

For other groups G, e.g., finite groups or general reductive groups, we have not proved the corresponding result as (28). It is a consequence of the uniqueness part of the stable degeneration conjecture. We also note that in [LX17a], it is proved that quasi-monomial minimizers over a T-equivariant klt singularity are automatically T-invariant.

5.2. Donaldson–Sun's Conjecture

One major application of what we know about the stable degeneration conjecture, formulated in Theorem 4.14, is the solution of [DS17, Conjecture 3.22] (see Conjecture 5.3), which predicts that for a singularity appearing on a Gromov–Hausdorff limit of Kähler–Einstein metrics, its metric tangent cone only depends on the algebraic structure of the singularity. In this section, we briefly explain the idea.

5.2.1. *K*-semistable degeneration. Let (M_k, g_k) be a sequence of Kähler–Einstein manifolds with positive curvature. Then possibly taking a subsequence, (M_k, g_k) converges in the Gromov–Hausdorff topology to a limit metric space (X, d_∞). By the work of Donaldson–Sun and Tian, X is homeomorphic to a $\mathbb{Q}$-Fano variety. For any point $x \in X$, a metric tangent cone $C_x X$ is defined as a pointed Gromov–Hausdorff limit:

$$C_x X = \lim_{r_k \to 0} \left(X, x, \frac{d_\infty}{r_k} \right). \tag{29}$$

By Cheeger–Colding's theory, $C_x X$ is always a metric cone. By [CCT02], the real codimension of singularity set of $C_x X$ is at least 4 and the regular part admits a Ricci-flat Kähler cone structure. In [DS17], it is further proved that $C_x X$ is an affine variety with an effective torus action. They proved that $C_x X$ is uniquely determined by the metric structure d_∞ and can be obtained in the following steps. In the first step, they defined a filtration $\{\mathcal{F}^\lambda\}_{\lambda \in \mathcal{S}}$ of the local ring $R = \mathcal{O}_{X,o}$ using the limiting metric structure d_∞. Here $\mathcal{S}$ is a set of positive numbers that they called the holomorphic spectrum which depends on the torus action on the metric tangent cone C. In the second step, they proved that the associated graded ring of $\{\mathcal{F}^\lambda\}$ is finitely generated and hence defines an affine variety, denoted by W. In the last step, they showed that W equivariantly degenerates to C. Notice that this process depends crucially on the limiting metric d_∞ on X. They then made the following conjecture.

Conjecture 5.3 (Donaldson–Sun). *Both W and C depend only on the algebraic germ structure of X near x.*

We made the following observations:

1. $\{\mathcal{F}^\lambda\}$ comes from a valuation v_0. This is due to the fact that W is a normal variety. More explicitly, since the question is local, we can assume $X = \mathrm{Spec}(R)$ with the germ of $x \in X$, by the work in [DS17], one can embed both X and C into a common ambient space $\mathbb{C}^N$, and v_0 on X is induced by the monomial valuation wt_{ξ_0} where ξ_0 is the linear holomorphic vector field with $2\mathrm{Im}(\xi_0)$ being the Reeb vector field of the Ricci flat Kähler cone metric on C. By this

construction, it is clear that the induced valuation by v_0 on W is nothing but wt_{ξ_0}.

2. v_0 is a quasi-monomial valuation. This follows from Lemma 2.7.

More importantly we conjectured in [Li18a] that v_0 can be characterized as the unique minimizer of $\widehat{\mathrm{vol}}_{X,x}$. As a corollary of the theory developed so far, we can already confirm [DS17, Conjecture 3.22] for W.

Theorem 5.4 ([LX17a]). *The semistable cone W in Donaldson–Sun's construction depends on the algebraic structure of (X, x).*

The proof consists of the following steps consisting of analytic and algebraic arguments:

1. By Theorem 3.17, (C, ξ_0) is K-polystable and in particular K-semistable. By Theorem 4.6, wt_{ξ_0} is a minimizer of $\widehat{\mathrm{vol}}_C$.
2. By Proposition 5.5, (W, ξ_0) is K-semistable. By Theorem 4.6 again, wt_{ξ_0} is a minimizer of $\widehat{\mathrm{vol}}_W$. Moreover, by Theorem 4.14, v_0 is a minimizer of $\widehat{\mathrm{vol}}_X$.
3. v_0 is a quasi-monomial minimizer of $\widehat{\mathrm{vol}}_X$ with a finitely generated associated graded ring. By Theorem 4.14, such a v_0 is indeed the unique minimizer of $\widehat{\mathrm{vol}}$ among all quasi-monomial valuations.

The following is an immediate consequence of Theorem 4.6.

Proposition 5.5. *Assume there is a special degeneration of a log-Fano cone (X, D, ξ_0) to (X_0, D_0, ξ_0). Assume that (X_0, D_0, ξ_0) is K-semistable, then (X, D, ξ_0) is also K-semistable, or equivalently, wt_{ξ_0} is the minimizer of $\widehat{\mathrm{vol}}_{(X,D,x)}$.*

Asssume (X, x) lives on a Gromov–Hausdorff limit of Kähler–Einstein Fano manifold. Then we can define the volume density in the sense of Geometric Measure Theory as the following quantity:

$$\Theta(x, X) = \lim_{r\to 0} \frac{\mathrm{Vol}(B_r(x))}{r^{2n}\mathrm{Vol}(B_1(\underline{0}))}. \tag{30}$$

Note that $n^n = \widehat{\mathrm{vol}}(0, \mathbb{C}^n)$. The normalized volumes of klt singularities on Gromov–Hausdorff limits have the following differential geometric meaning:

Theorem 5.6 ([LX17a]). *With the same notation as above, we have the identity:*

$$\widehat{\mathrm{vol}}(x, X)\,/n^n = \Theta(x, X). \tag{31}$$

Proof. From the standard metric geometry, we have $\Theta(x, X) = \Theta(o_C, C)$. Because C admits a Ricci-flat Kähler cone metric, by Theorem 3.17, (C, ξ_0) is K-semistable. $\widehat{\mathrm{vol}}(x, X) = \widehat{\mathrm{vol}}(o_C, C)$.

On the other hand, since C is a metric cone, from the definition of the volume of $\xi_0 = \frac{1}{2}\,(r\partial_r - iJ(r\partial_r))$ is equal to:

$$\Theta(o_C, C) = \frac{\mathrm{Vol}(C \cap \{r = 1\})}{\mathrm{Vol}(S^{2n-1})} = \mathrm{vol}(\xi_0).$$

By Theorem 3.16, $A(\mathrm{wt}_{\xi_0}) = n$ and $\widehat{\mathrm{vol}}(o_C, C) = n^n\mathrm{vol}(\xi_0) = n^n\Theta(o_C, C)$. □

5.2.2. Uniqueness of polystable degeneration. To confirm Donaldson–Sun's conjecture, we also need to prove the uniqueness of polystable degenerations for K-semistable Fano cones.

Since a Fano cone singularity (C, ξ) with a Ricci-flat Kähler cone metric is always K-polystable (see [CS15, Theorem 7.1] and also Theorem 3.17), once knowing that W only depends on the algebraic structure of $o \in M_\infty$, an affirmative answer to Conjecture 5.3 follows from the following more general result by letting $(X, D, \xi_0) = (W, \emptyset, \xi_0)$:

Theorem 5.7 ([LWX18]). *Given a K-semistable log Fano cone singularity (X, D, ξ_0), there always exists a special test configuration $(\mathcal{X}, \mathcal{D}, \xi_0; \eta)$ which degenerates (X, D, ξ_0) to a K-polystable log Fano cone singularity (X_0, D_0, ξ_0). Furthermore, such (X_0, D_0, ξ_0) is uniquely determined by (X, D, ξ_0) up to isomorphism.*

For the special case of smooth (or $\mathbb{Q}$-Gorenstein smoothable) Fano varieties, this was proved in [LWX14, 7.1] based on analytic results which also show the uniqueness of Gromov–Hausdorff limit for a flat family of Fano Kähler–Einstein manifolds. Our proof of Theorem 5.7 is however a completely new algebraic argument.

We briefly discuss the idea to prove Theorem 5.7 in [LWX18], which heavily depends on the study of normalized volumes as discussed in Section 4.2.

Let $(\mathcal{X}^{(i)}, D^{(i)}, \xi_0, \eta^{(i)}), (i = 1, 2)$, be two special test configurations of the log Fano cone (X, D, ξ_0) with the central fibre $(X_0^{(i)}, D_0^{(i)}, \xi_0)$. To show Theorem 5.7, the main step is to show that if $\mathrm{Fut}(\mathcal{X}^{(i)}, \mathcal{D}^{(i)}, \xi_0; \eta^{(i)}) = 0, (i = 1, 2)$, then there exist special test configurations $(\mathcal{X}'^{(i)}, \mathcal{D}'^{(i)})$ of $(X_0^{(i)}, D_0^{(i)})$ such that $(\mathcal{X}'^{(i)}, \mathcal{D}'^{(i)})$ have isomorphic central fibres, which we will describe below.

We consider the normalized volume functional defined on the valuation space $\mathrm{Val}_{X,x}$ over the vertex x of the cone X. Then $(\mathcal{X}^{(1)}, \mathcal{D}^{(1)}, \xi_0; \eta^{(1)})$ determines a "ray" of valuations emanating from the toric valuation wt_{ξ_0} and the generalized Futaki invariant $\mathrm{Fut}(\mathcal{X}^{(1)}, \mathcal{D}^{(1)}, \xi_0; \eta^{(1)})$ is the derivative of the normalized volume at wt_{ξ_0} along this ray.

$$\begin{tikzcd}
(X_0^{(2)}, D_0^{(2)}) \arrow[d, squiggly, "(\mathcal{X}'^{(2)}, \mathcal{D}'^{(2)})"'] & & (X, D) \arrow[ll, squiggly, "(\mathcal{X}^{(2)}, \mathcal{D}^{(2)}) \longleftarrow \mathcal{Y}_k^{(2)} \longleftarrow \mathcal{E}_k^{(2)}"'] \arrow[d, squiggly, "(\mathcal{X}^{(1)}, \mathcal{D}^{(1)}) \leftarrow \mathcal{Y}_k \leftarrow \mathcal{E}_k = E_k \times \mathbb{C}^1"] & Y_k \arrow[l] & E_k \arrow[l] \\
(X_0', D_0') & & (X_0^{(1)}, D_0^{(1)}) \arrow[ll, squiggly, "(\mathcal{X}'^{(1)}, \mathcal{D}'^{(1)})"'] & Y_{k,0} \arrow[l] & E_k \arrow[l]
\end{tikzcd} \tag{32}$$

We can approximate ξ_0 by a sequence of integral vectors $\tilde{\xi}_k$ such that $|\tilde{\xi}_k - k\xi_0| \le C$. For $k \gg 1$, the vector $\tilde{\xi}_k - \eta$ corresponds to a Kollár component E_k over X. Our key argument is to show that E_k can be degenerated along $(\mathcal{X}^{(2)}, \mathcal{D}^{(2)})$ to get a model $\mathcal{Y}_k^{(2)} \to \mathcal{X}^{(2)}$ with an exceptional divisor $\mathcal{E}_k^{(2)}$ such that $(\mathcal{Y}_k^{(2)}, \mathcal{E}_k^{(2)}) \times_{\mathbb{C}} \mathbb{C}^* \cong (Y_k, E_k) \times \mathbb{C}^*$ where the isomorphism is compatible with the equivariant isomorphism of the *second* special test configuration. Note that $E_k \times \mathbb{C}^*$ determines a divisorial valuation over $X \times \mathbb{C}^*$ and hence over $(\mathcal{X}^{(2)}, \mathcal{D}^{(2)})$. So the goal is to show that this divisorial valuation can be extracted as the only exceptional divisor over $\mathcal{X}^{(2)}$. By the work in the minimal model program (MMP) (see [BCHM10]), this would be true if there is a graded sequence of ideals $\mathfrak{A}_\bullet$ and a positive real number c'_k such that two conditions are satisfied:

$$(\mathcal{X}^{(2)}, \mathcal{D}^{(2)} + c'_k\mathfrak{A}_\bullet) \text{ is klt} \quad \text{and} \quad A(E_k \times \mathbb{C}; \mathcal{X}^{(2)}, \mathcal{D}^{(2)} + c'_k\mathfrak{A}_\bullet) < 1,$$

where $A(E_k \times \mathbb{C}; \mathcal{X}^{(2)}, \mathcal{D}^{(2)} + c'_k\mathfrak{A}_\bullet)$ is the log discrepancy of (the birational transform of) $E_k \times \mathbb{C}$ with respect to the triple $(\mathcal{X}^{(2)}, \mathcal{D}^{(2)} + c'_k\mathfrak{A}_\bullet)$.

To find such an $\mathfrak{A}_\bullet$, we look at the graded sequence of valuative ideals $\{\mathfrak{a}_\bullet\}$ of ord_{E_k} and its equivariant degeneration along the second special test configuration $(\mathcal{X}^{(2)}, \mathcal{D}^{(2)})$. The resulting graded sequence of ideals over $\mathcal{X}^{(2)}$ will be denoted by $\mathfrak{A}_\bullet$. Using the study in Section 4.2 one can show the assumptions that $(\mathcal{X}^{(1)}, \mathcal{D}^{(1)}; \xi_0)$ is K-semistable and $\mathrm{Fut}(\mathcal{X}^{(1)}, \mathcal{D}^{(1)}, \xi_0; \eta) = 0$ implies

$$f(k) := \widehat{\mathrm{vol}}(E_k) \text{ is of the order } f(0) + O(k^{-2}).$$

This in turn guarantees that we can find c'_k satisfying the above two conditions.

Applying the relative Rees algebra construction to $\mathcal{E}_k^{(2)} \subset \mathcal{Y}_k^{(2)}/\mathbb{C}$, we get a family over $\mathbb{C}^2$, which over $\mathbb{C} \times \{t\}$ is the same as $(\mathcal{X}^{(1)}, \mathcal{D}^{(1)})$ for $t \neq 0$ and gives a degeneration of $(X_0^{(1)}, D_0^{(1)})$ for $t = 0$. On the other hand, over $\{0\} \times \mathbb{C}$, we get a degeneration of $(X_0^{(2)}, D_0^{(2)})$. Therefore, we indeed show that the two special fibers of two special test configurations $(\mathcal{X}^{(i)}, \mathcal{D}^{(i)}, \xi_0; \eta^{(i)})$ $(i = 1, 2)$ with $\mathrm{Fut}(\mathcal{X}^{(i)}, \mathcal{D}^{(i)}, \xi_0; \eta^{(i)}) = 0$ will have a common degeneration.

5.3. Estimates in dimension three and *K*-stability of threefolds

In general, it is not so easy to find the minimizer of $\widehat{\mathrm{vol}}(\cdot)$ for a given singularity. A number of cases have been computed in [Li18a, LL19, LX16, LX17a, LX17b] including quotient singularities, ADE singularities in all dimensions (except four-dimensional D_4) etc.

Here we study normalized volumes of threefold klt singularities, and then give a global application where we show that all GIT semi-stable (resp. polystable) cubic threefolds are also K-semi-stable (resp. K-polystable). Our main estimate is in Theorem 5.8, which heavily depends on classifications of canonical threefold singularities.

Theorem 5.8 ([LX17b]). *Let $x \in X$ be a three-dimensional non-smooth klt singularity. Then $\widehat{\mathrm{vol}}(x, X) \leq 16$ and the equality holds if and only if it is an A_1 singularity;*

The proof of Theorem 5.8 heavily relies on the classification theory of three-dimensional canonical and terminal singularities, developed in the investigation of explicit three-dimensional MMP.

The idea goes as follows. Firstly, we reduce to the case of Gorenstein canonical singularity. If $x \in X$ is not Gorenstein, let us take the index one cover $\tilde{x} \in \widetilde{X}$ of

$x \in X$. Hence $\tilde{x} \in \widetilde{X}$ is a Gorenstein canonical singularity. If $\tilde{x} \in \widetilde{X}$ is smooth, then $\widehat{\mathrm{vol}}(x, X) = 27/\mathrm{ind}(x, K_X) \leq 13.5 < 16$. If $\tilde{x} \in \widetilde{X}$ is not smooth, the a weak version of finite degree formula (Proposition 5.10) implies that $\widehat{\mathrm{vol}}(x, X) < \widehat{\mathrm{vol}}(\tilde{x}, \widetilde{X})$.

Next, let us assume that $x \in X$ is Gorenstein canonical. By [KM98, Proposition 2.36], there exist only finitely many crepant exceptional divisors over X. By [BCHM10], we can extract these divisors simultaneously on a birational model $Y_1 \to X$. If none of these exceptional divisors are centered at x, then [KM98, Theorem 5.34] implies that $x \in X$ is a cDV singularity, hence $\mathrm{lct}(\mathfrak{m}_x) \leq 4 - \mathrm{mult}(\mathfrak{m}_x)$ which implies

$$\widehat{\mathrm{vol}}(x, X) \leq \mathrm{lct}(\mathfrak{m}_x)^3 \mathrm{mult}(\mathfrak{m}_x) \leq (4 - \mathrm{mult}(\mathfrak{m}_x))^3 \mathrm{mult}(\mathfrak{m}_x) \leq 16.$$

The equality case can be characterized using the volume of birational models approach in [LX16]. If some crepant exceptional divisor $E_1 \subset Y_1$ is centered at x, then let us run $(Y_1, \epsilon E_1)$-MMP over X for $0 \ll \epsilon < 1$. By [Kol13, 1.35], this MMP will terminate as $Y_1 \dashrightarrow Y \xrightarrow{g} Y'$, where $Y_1 \dashrightarrow Y$ is the composition of a sequence of flips, and $g : Y \to Y'$ contracts the birational transform E of E_1. If $g(E)$ is a curve, then Y' has cDV singularities along $g(E)$ by [KM98, Theorem 5.34]. By choosing a point $y' \in g(E)$, we have

$$\widehat{\mathrm{vol}}(x, X) < \widehat{\mathrm{vol}}(y', Y') \leq 16.$$

If $g(E) = y'$ is a point, then we still have $\widehat{\mathrm{vol}}(x, X) < \widehat{\mathrm{vol}}(y', Y')$. Thus it suffices to show $\widehat{\mathrm{vol}}(y', Y') < 16$.

If Y has a singular point $y \in E$, then we know that $y \in Y$ is a cDV singularity. Hence

$$\widehat{\mathrm{vol}}(y', Y') < \widehat{\mathrm{vol}}(y, Y) \leq 16.$$

So we may assume that Y is smooth along E. In particular, E is a (possibly non-normal) reduced Gorenstein del Pezzo surface. If E is normal, then classification of such surfaces show that $(-K_E)^2 \leq 9$. Thus

$$\widehat{\mathrm{vol}}(y', Y') \leq A_{Y'}(\mathrm{ord}_E)^3 \mathrm{vol}(\mathrm{ord}_E) = (-K_E)^2 \leq 9 < 16.$$

If E is non-normal, then from Reid's classification [Rei94] either $(-K_E)^2 \leq 4$ or the normalization of E is a Hirzebruch surface. In the former case, we have $\widehat{\mathrm{vol}}(y', Y') \leq 4$. In the latter case, we need to take a general fiber l of E and argue that $\widehat{\mathrm{vol}}_{Y', y'}(\mathrm{ord}_l) \leq 16$.

Here are some intermediate results in proving Theorem 5.8.

Proposition 5.9. *Let $\phi : (Y, y) \to (X, x)$ be a birational morphism of klt singularities such that $y \in \mathrm{Ex}(\phi)$. If $K_Y \leq \phi^* K_X$, then $\widehat{\mathrm{vol}}(x, X) < \widehat{\mathrm{vol}}(y, Y)$.*

See Conjecture 6.4 for more discussions about the following proposition.

Proposition 5.10. *Let $\pi : (\widetilde{X}, \tilde{x}) \to (X, x)$ be a finite quasi-étale morphism of klt singularities of degree at least 2. Then we have*

$$\widehat{\mathrm{vol}}(x, X) < \widehat{\mathrm{vol}}(\tilde{x}, \widetilde{X}) \leq \deg(\pi) \cdot \widehat{\mathrm{vol}}(x, X).$$

As mentioned [SS17], one main application of the local volume estimate Theorem 5.8 is to the K-stability question of cubic threefolds.

Theorem 5.11 ([LX17b]). *A cubic threefold is K-(poly/semi)stable if and only if it is GIT (poly/semi)stable. In particular, any smooth cubic threefold is K-stable.*

The general strategy to prove Theorem 5.11 is via the comparison of moduli spaces which has first appeared in [MM93] built on the work of [Tia90]. Later it was also applied in [OSS16, SS17].

First, one can construct a proper algebraic space which is a good quotient moduli space with closed points parametrizing all smoothable K-polystable $\mathbb{Q}$-Fano varieties (see, e.g., [LWX14, Oda15]). Let M be the closed subspace whose closed points parametrize KE cubic threefolds and their K-polystable limits. By [Tia87], we know that at least one cubic threefold, namely the Fermat cubic threefold, admits a KE metric. Hence M is non-empty. By the Zariski openness of K-(semi)stability of smoothable Fano varieties (cf. [Oda15, LWX14]), the K-moduli space M is birational to the GIT moduli space M^{GIT} of cubic threefolds.

Next, we will show that any K-semistable limit X of a family of cubic threefolds $\{X_t\}$ over a punctured curve is necessarily a cubic threefold. The idea is to control the singularity of X use an inequality from [Liu18] (see Theorem 5.12) between the global volume of a K-semistable Fano variety and the local normalized volume. Since the volume of X is the same as the volume of a cubic 3-fold which is 24, Theorem 5.12 immediately implies that $\widehat{\mathrm{vol}}(x, X) \geq \frac{81}{8}$ for any closed point $x \in X$. The limit X carries a $\mathbb{Q}$-Cartier Weil divisor L which is the flat limit of hyperplane sections in the cubic threefolds X_t. It is clear that $-K_X \sim_{\mathbb{Q}} 2L$ and $(L^3) = 3$, thus once we show that L is Cartier, we can claim that X is a cubic threefold using a result of T. Fujita.

Assume to the contrary that L is not Cartier at some point $x \in X$, then we may take the index 1 cover $(\tilde{x} \in \widetilde{X}) \to (x \in X)$ of L. From the finite degree formula Theorem 6.5,

$$\widehat{\mathrm{vol}}(\tilde{x}, \widetilde{X}) = \mathrm{ind}(L) \cdot \widehat{\mathrm{vol}}(x, X) \geq 81/4.$$

Hence $\tilde{x} \in \widetilde{X}$ is a smooth point and $\mathrm{ind}(L) = 2$ by Theorem 5.8. Thus $x \in X$ is a quotient singularity of type $\frac{1}{2}(1,1,0)$ from the smoothable condition. Then using the local Grothendieck–Lefschetz theorem, we can show that L is indeed Cartier at $x \in X$ which is a contradiction.

So far we have shown that any K-polystable point X in M is a cubic threefold. By an argument of Paul and Tian in [Tia94], we know that any K-(poly/semi)stable hypersurface is GIT (poly/semi)stable. Thus we obtain an injective birational morphism $M \to M^{\mathrm{GIT}}$ between proper algebraic spaces. This implies that M is isomorphic to M^{GIT} which finishes the proof.

Theorem 5.12 ([Liu18]). *Let X be an n-dimensional K-semistable Fano variety. Then for any closed point $x \in X$, we have*

$$(-K_X)^n \leq \left(1+\frac{1}{n}\right)^n \widehat{\mathrm{vol}}(x, X).$$

When X is smooth, the above result was first proved in [Fuj18].

6. Questions and future research

6.1. Revisit stable degeneration conjecture

The following two parts of stable degeneration conjecture, proposed in [Li18a], are still missing.

Conjecture 6.1 (Quasi-monomial). *Let $x \in (X, D)$ be a klt singularity. Any minimizer of $\widehat{\mathrm{vol}}_{(X,D),x}$ is quasi-monomial.*

Conjecture 6.2 (Finite generation). *Let $x \in (X = \mathrm{Spec}(R), D)$ be a klt singularity. Any minimizer of $\widehat{\mathrm{vol}}_{(X,D),x}$ has its associated graded ring $\mathrm{gr}_v(R)$ to be finitely generated.*

Due to the fundamental role of the stable degeneration conjecture, it implies many other interesting properties. We discuss a number of special cases or consequences, with the hope that some of them might be solved first.

One interesting consequence of the uniqueness of the minimizer is the following

Conjecture 6.3 (Group action). *If there is a group G acting on the klt singularity $x \in (X, D)$ such that x is a fixed point, then there exists a G-invariant minimizer.*

Applying this conjecture to a cone singularity, it implies that to test the K-semistability of a log Fano (S, B) with a G-action, we only need to test on G-equivariant test configurations, a fact known for a Fano manifold X and G-reductive.

There are two special cases naturally appearing in contexts. The first one is that when G is a torus group T. It follows the argument in [Blu18] and the techniques of degenerating ideals to their initials, that there is a T-equivariant minimizer. This is the philosophy behind Section 5.1. It also follows from [LX17a] that any quasi-monomial minimizer is T-equivariant.

A more challenging case is when G is a finite group. Indeed, Conjecture 6.3 for finite group G implies the following finite degree formula.

Conjecture 6.4 (Finite degree formula). *If $\pi\colon (y \in Y, D') \to (x \in X, D)$ is a dominant finite morphism between klt singularities, such that $K_Y + D' = \pi^*(K_X + D)$, then*

$$\deg(\pi) \cdot \widehat{\mathrm{vol}}(x, X, D) = \widehat{\mathrm{vol}}(y, Y, D').$$

This is useful when we want to bound the klt singularities $x \in (X, D)$ with a large volume.

Theorem 6.5 ([LX17a]). *Conjecture 6.4 is true when (X, x) is on a Gromov–Hausdorff limit of Kähler–Einstein Fano manifolds.*

Proof. Let $\pi = \pi_X : (Y, y) \to (X, x)$ be a quasi-étale morphism, i.e., π_X is étale in codimension one. Then π_X induces a quasi-étale morphism along the 2-step degeneration of X.

$$\begin{tikzcd} Y \arrow[r, rightsquigarrow] \arrow[d, "\pi_X"] & W_Y \arrow[r, rightsquigarrow] \arrow[d, "\pi_W"] & C_Y \arrow[d, "\pi_C"] \\ X \arrow[r, rightsquigarrow] & W \arrow[r, rightsquigarrow] & C \end{tikzcd} \tag{33}$$

We can use the above diagram to prove the degree multiplication formula. Roughly speaking, because C admits a Ricci-flat Kähler cone metric ω_C with radius function r^2 and π_C is quasi-étale, we can pull back it to get $\pi_C^* r^2$ which is also a potential for a weak Ricci-flat Kähler cone metric ω_{C_Y}. By Theorem 3.17, Theorem 4.6 and Theorem 4.14, we know that the Reeb vector field associated to ω_C (resp. ω_{C_Y}) induces minimizing valuations of $\widehat{\mathrm{vol}}_X$ (resp. $\widehat{\mathrm{vol}}_Y$). So we get

$$\widehat{\mathrm{vol}}(y, Y) = \widehat{\mathrm{vol}}(o_{C_Y}, C_Y) = \deg(\pi_C) \cdot \widehat{\mathrm{vol}}(o_C, C) = \deg(\pi_C) \cdot \widehat{\mathrm{vol}}(x, X). \qquad \square$$

Another consequence of the stable degeneration conjecture is the following strengthening of Theorem 2.11.

Conjecture 6.6. *Let $\pi : (\mathcal{X}, \mathcal{D}) \to T$ together with a section $t \in T \mapsto x_t \in \mathcal{X}_t$ be a $\mathbb{Q}$-Gorenstein flat family of klt singularities. Then the function $t \mapsto \widehat{\mathrm{vol}}(x_t, \mathcal{X}_t, \mathcal{D}_t)$ is constructible with respect to the Zariski topology.*

Besides the stable degeneration conjecture, to prove Conjecture 6.6, we also need to know the well-expected speculation that K-semistability is an open condition. It is also natural to consider the volume of non-closed point. However, the following conjecture says after the right scaling, it does not contribute more information.

Conjecture 6.7. *If a klt pair (X, D) has a non-closed point η, and let $Z = \overline{\{\eta\}}$ has dimension d. Pick a general closed point $x \in Z$, then*

$$\widehat{\mathrm{vol}}(x, X, D) = \widehat{\mathrm{vol}}(\eta, X, D) \cdot \frac{n^n}{(n-d)^{n-d}}.$$

In fact, combining the argument in [LZ18], for any valuation $v \in \mathrm{Val}_X$ such that its center $Z = \mathrm{Center}_X(v)$ on X is of dimension d and $x \in Z$, denoted by η is the generic point of Z, one can show that

$$\frac{\widehat{\mathrm{vol}}_{(X,D),\eta}(v) \cdot n^n}{(n-d)^{n-d}} \geq \widehat{\mathrm{vol}}(x, X, D),$$

i.e.,

$$\widehat{\mathrm{vol}}(x,X,D)=\inf_{v}\left\{\frac{n^n\cdot\widehat{\mathrm{vol}}_{(X,D),\eta}(v)}{(n-d)^{n-d}}\mid x\in Z=\overline{\{\eta\}}=\mathrm{Center}_X(v),\dim(Z)=d\right\}.$$

6.2. Birational geometry study

A different invariant attached to a klt singularities, called the minimal log discrepancy has been intensively studied in the minimal model program, though there are still many deep questions unanswered. We can formulate many similar questions for $\widehat{\mathrm{vol}}$.

6.2.1. Inversion of adjunction. One could look for a theory of the change of the volumes when the klt pair is 'close' to a log canonical singularities, using the inversion of adjunction. We have some results along this line.

Proposition 6.8. *Let $x\in(X,\Delta)$ be an n-dimensional klt singularity. Let D be a normal $\mathbb{Q}$-Cartier divisor containing x such that $(X,D+\Delta)$ is plt. Denote by Δ_D the different of Δ on D. Then*

$$\lim_{\epsilon\to 0+}\frac{\widehat{\mathrm{vol}}(x,X,(1-\epsilon)D+\Delta)}{n^n\epsilon}=\frac{\widehat{\mathrm{vol}}(x,D,\Delta_D)}{(n-1)^{n-1}}.$$

Proof. Using the degeneration argument in [LZ18], we know that

$$\epsilon^{-1}\widehat{\mathrm{vol}}(x,X,(1-\epsilon)D+\Delta)\geq\frac{n^n}{(n-1)^{n-1}}\widehat{\mathrm{vol}}(x,D,\Delta_D).$$

Hence it suffices to show the reverse inequality is true after taking limits. Let us pick an arbitrary Kollár component S over $x\in(D,\Delta_D)$ with valuation ideals $\mathfrak{a}_m:=\mathfrak{a}_m(\mathrm{ord}_S)$. Choose m sufficiently divisible so that $\mathfrak{a}_{im}=\mathfrak{a}_m^i$ for any $i\in\mathbb{N}$. Then we know that $\mathrm{lct}(D,\Delta_D;\mathfrak{a}_m)=A_X(\mathrm{ord}_S)/m=:c$. Let $\mathfrak{b}_m$ be the pull-back ideal of $\mathfrak{a}_m$ on X. By inversion of adjunction, we have $\mathrm{lct}(X,D+\Delta;\mathfrak{b}_m)=\mathrm{lct}(D,\Delta_D;\mathfrak{a}_m)=c$.

Let E be an exceptional divisor over X computing $\mathrm{lct}(X,D+\Delta;\mathfrak{b}_m)$. Then E is centered at $x\in X$ since $(X,D+\Delta)$ is plt. For $\epsilon_1>0$ sufficiently small, we have that $(X,\Delta+(1-\epsilon_1)(D+c\cdot\mathfrak{b}_m))$ is a klt pair over which the discrepancy of E is negative. Thus [BCHM10] implies that there exists a proper birational model $\mu:Y\to X$ which only extracts E. Moreover, $\mu:Y\to X$ is a log canonical modification of $(X,\Delta+D+c\cdot\mathfrak{b}_m)$. Let $\widetilde{D}$ be the normalization of $\mu_*^{-1}D$. Then by adjunction, the lifting morphism $\tilde{\mu}:\widetilde{D}\to D$ is a log canonical (in fact plt) modification of $(D,\Delta_D+c\cdot\mathfrak{a}_m)$. Since $\mathrm{Bl}_{\mathfrak{a}_m}D\to D$ provides a model of the Kollár component S, this is the only log canonical modification of $(D,\Delta_D+c\cdot\mathfrak{a}_m)$. Hence $E|_{\widetilde{D}}=S$ and $(\widetilde{D},\tilde{\mu}_*^{-1}\Delta_D+E|_{\widetilde{D}})$ is plt. Then by inversion of adjunction, $(Y,\mu_*^{-1}\Delta+\mu_*^{-1}D+E)$ is qdlt and $\mu_*^{-1}D=\widetilde{D}$ is normal. Note that all the constructions so far are independent of the choice of ϵ.

Over the qdlt model $(Y,\mu_*^{-1}\Delta+\mu_*^{-1}D+E)$, we consider a quasi-monomial valuation v_λ of weights 1 and λ along divisors $\widetilde{D}$ and E respectively. By adjunction, we know that $A_{(X,\Delta)}(\mathrm{ord}_E) = A_{(D,\Delta_D)}(\mathrm{ord}_S) + \mathrm{ord}_E(D)$. Hence computation shows that

$$A_{(X,\Delta+(1-\epsilon)D)}(v_\lambda) = \lambda A_{(D,\Delta_D)}(\mathrm{ord}_S) + \lambda\epsilon \cdot \mathrm{ord}_E(D) + \epsilon.$$

Then using the Okounkov body description of the volume (see [LM09, KK12]), we easily see that $\mathrm{vol}(v_\lambda) \le \lambda^{1-n}\mathrm{vol}(\mathrm{ord}_S)$. Hence

$$\begin{aligned}\widehat{\mathrm{vol}}_{(X,\Delta+(1-\epsilon)D)}(v_\lambda) &\le \lambda^{1-n}((A_{(D,\Delta_D)}(\mathrm{ord}_S) + \epsilon \cdot \mathrm{ord}_E(D))\lambda + \epsilon)^n \mathrm{vol}(\mathrm{ord}_S)\\ &=: \phi(\lambda).\end{aligned}$$

It is easy to see that $\phi(\lambda)$ reaches its minimum at

$$\lambda_0 = \frac{(n-1)\epsilon}{A_{D,\Delta_D}(\mathrm{ord}_S) + \epsilon \cdot \mathrm{ord}_E(D)}.$$

Hence computation shows

$$\epsilon^{-1}\widehat{\mathrm{vol}}_{(X,\Delta+(1-\epsilon)D)}(v_{\lambda_0}) \le \frac{n^n}{(n-1)^{n-1}}(A_{(D,\Delta_D)}(\mathrm{ord}_S) + \epsilon \cdot \mathrm{ord}_E(D))^{n-1}\mathrm{vol}(\mathrm{ord}_S).$$

Thus

$$\limsup_{\epsilon\to 0} \epsilon^{-1}\widehat{\mathrm{vol}}(x, X, \Delta + (1-\epsilon)D) \le \frac{n^n}{(n-1)^{n-1}}\widehat{\mathrm{vol}}_{(D,\Delta_D)}(\mathrm{ord}_S).$$

Since this inequality holds for any Kollár component S over $x \in (D, \Delta_D)$, the proof is finished. □

When the center is zero-dimensional, we also have

Proposition 6.9. *Let $x \in (X, \Delta)$ be a klt singularity. Let $D \ge 0$ be a $\mathbb{Q}$-Cartier divisor such that $(X, \Delta + D)$ is log canonical with $\{x\}$ being the minimal non-klt center. Then there exists $\epsilon_0 > 0$ (depending only on the coefficient of Δ, D and n) and a quasi-monomial valuation $v \in \mathrm{Val}_{X,x}$ such that v computes both $\mathrm{lct}(X, \Delta; D)$ and $\widehat{\mathrm{vol}}(x, X, \Delta + (1-\epsilon)D)$ for any $0 < \epsilon < \epsilon_0$. In particular,*

$$\widehat{\mathrm{vol}}(x, X, \Delta + (1-\epsilon)D) = \widehat{\mathrm{vol}}_{x,(X,\Delta)}(v) \cdot \epsilon^n \text{ for any } 0 < \epsilon < \epsilon_0.$$

Proof. Let $Y^{\mathrm{dlt}} \to X$ be a dlt modification of $(X, \Delta + D)$. Let $K_{Y^{\mathrm{dlt}}} + \Delta^{\mathrm{dlt}}$ be the log pull back of $K_X + \Delta + D$. Then by [dFKX17], the dual complex $\mathcal{DR}(\Delta^{\mathrm{dlt}})$ form a natural subspace of $\mathrm{Val}^{=1}_{X,x}$. Any divisorial valuation ord_E computing $\mathrm{lct}(X, \Delta; D)$ corresponds to a rescaling of a valuation in $\mathcal{DR}(\Delta^{\mathrm{dlt}})$. Consider the function $\mathrm{vol}_X : \mathcal{DR}(\Delta^{\mathrm{dlt}}) \to \mathbb{R}_{>0} \cup \{+\infty\}$. Denote by $\mathcal{DR}^\circ(\Delta^{\mathrm{dlt}})$ the open subset of $\mathcal{DR}(\Delta^{\mathrm{dlt}})$ consisting of valuations centered at x. Since $\{x\}$ is the minimal non-klt center of $(X, \Delta + D)$, we know that $\mathcal{DR}^\circ(\Delta^{\mathrm{dlt}})$ is non-empty. By [BFJ14] the function vol is continuous on $\mathcal{DR}(\Delta^{\mathrm{dlt}})$, so we can take a vol-minimizing valuation $v \in \mathcal{DR}^\circ(\Delta^{\mathrm{dlt}})$. Hence v is a minimizer of $\widehat{\mathrm{vol}}$ restricted to $\mathcal{DR}^\circ(\Delta^{\mathrm{dlt}})$.

Assume S is an arbitrary Kollár component over $(X, \Delta + (1-\epsilon)D)$. Then we have a birational morphism $\mu : Y \to X$ such that $K_Y + \mu_*^{-1}(\Delta + (1-\epsilon)D) + S$ is plt, and μ is an isomorphism away from x with $S = \mu^{-1}(x)$. Then by ACC of lct [HMX14], we know that there exists ϵ_0 such that $K_Y + \mu_*^{-1}(\Delta + D) + S$ is log canonical whenever $0 < \epsilon < \epsilon_0$. Let v' be an arbitrary divisorial valuation in $\mathcal{DR}^\circ(\Delta^{\rm dlt})$. Since $K_Y + \mu_*^{-1}(\Delta + D) + S \sim_{\mathbb{Q}} \mu^*(K_X + \Delta + D) + A_{(X,\Delta+D)}(\mathrm{ord}_S)S$, we have

$$0 \le A_{(Y,\mu_*^{-1}(\Delta+D)+S)}(v') = A_{(X,\Delta+D)}(v') - A_{(X,\Delta+D)}(\mathrm{ord}_S)\cdot v'(S).$$

Since $A_{(X,\Delta+D)}(v') = 0$ and $v'(S) > 0$ since $\{x\}$ is the only lc center, we know that $A_{(X,\Delta+D)}(\mathrm{ord}_S) = 0$. Thus a rescaling of ord_S belongs to $\mathcal{DR}(\Delta^{\rm dlt})$. Then by [LX16] we see that

$$\begin{aligned}\widehat{\mathrm{vol}}(x, X, \Delta + (1-\epsilon)D) &= \min_{v' \in \mathcal{DR}(\Delta^{\rm dlt})} \widehat{\mathrm{vol}}_{(X,\Delta+(1-\epsilon)D)}(v')\\ &= \epsilon^n \min_{v' \in \mathcal{DR}(\Delta^{\rm dlt})} \mathrm{vol}_X(v') = \widehat{\mathrm{vol}}_{x,(X,\Delta)}(v)\cdot \epsilon^n. \qquad \square\end{aligned}$$

One should be able to solve the following question using the above techniques.

Question 6.10. Let $x \in (X, \Delta)$ be an n-dimensional klt singularity. Let D be an effective $\mathbb{Q}$-Cartier $\mathbb{Q}$-Weil divisor through x. Let $c = \mathrm{lct}(X, \Delta; D)$, and let W be the minimal log canonical center of $(X, \Delta + cD)$ containing x. By Kawamata's subadjunction, we have $(K_X + \Delta + cD)|_W = K_W + \Delta_W + J_W$, where $(W, \Delta_W + J_W)$ is a generalized klt pair. Denote by $k := \mathrm{codim}_X W$, then is it true that

$$\lim_{\epsilon\to 0+} \epsilon^{-k}\frac{\widehat{\mathrm{vol}}(x, X, \Delta + (1-\epsilon)cD)}{n^n} \ge \frac{\widehat{\mathrm{vol}}(w, X, \Delta)}{k^k}\cdot\frac{\widehat{\mathrm{vol}}(x, W, \Delta_W + J_W)}{(n-k)^{n-k}}$$

where w is the generic point of W in X and $\widehat{\mathrm{vol}}(x, W, \Delta_W + J_W)$ is similarly defined as for the usual klt pair case in Definition 2.3?

6.2.2. Uniform bound. The following is conjectured in [SS17] (see also [LX17b]).

Conjecture 6.11. *Let $x \in X$ be an n-dimensional singular point, then $\widehat{\mathrm{vol}}(x, X) \le 2(n-1)^n$.*

The constant $2(n-1)^n$ is the volume of a rational double point. When $n = 3$, it is proved in Theorem 5.8. The implication to the K-stability question of cubic hypersurfaces as in the argument of Theorem 5.11 holds in any dimension.

We also ask whether the following strong property of the set of local volumes holds.

Question 6.12. Fix the dimension n, and a finite set $I \subset [0, 1]$. Is it true that the set $\mathrm{Vol}^{\rm loc}_{n,I}$ consisting of all possible local volumes of n-dimensional klt singularities $x \in (X, D)$ with (coefficients of D) $\subset I$ has the only accumulation point 0?

Next we give a comparison between local volumes and minimal log discrepancies.

Theorem 6.13. *Let $x \in (X, \Delta)$ be an n-dimensional complex klt singularity. Then there exists a neighborhood U of $x \in X$ such that $(U, \Delta|_U)$ is $(\widehat{\mathrm{vol}}(x, X, \Delta)/n^n)$-lc. Moreover,* $\mathrm{mld}(x, X, \Delta) > \widehat{\mathrm{vol}}(x, X, \Delta)/n^n$.

Proof. If $x \in X$ is not $\mathbb{Q}$-factorial then we may replace X by its $\mathbb{Q}$-factorial modification under which the local volume will increase by [LX17b, Corollary 2.11]. Let Δ_i be any component of Δ containing x. Then [BL18, Theorem 33] implies that $A_{(X,\Delta)}(\Delta_i) \geq \widehat{\mathrm{vol}}(x, X, \Delta)/n^n$. Let E be any exceptional divisor over X such that x is contained in the Zariski closure of $c_X(E)$ and $a(E; X, \Delta) < 0$. Then by [Kol13, Corollary 1.39], there exists a proper birational morphism $\mu : Y \to X$ such that Y is normal, $\mathbb{Q}$-factorial and $E = \mathrm{Ex}(\mu) \supset \mu^{-1}(x)$. Since $K_Y + \mu_*^{-1}\Delta - a(E; X, \Delta)E = \mu^*(K_X + \Delta)$, we know that $(Y, \mu_*^{-1}\Delta - a(E; X, \Delta)E)$ is klt. Let $y \in \mu^{-1}(x)$ be a point, then y lies on E. Hence by [LX17b, Corollary 2.11] and [BL18, Theorem 33] we have

$$\widehat{\mathrm{vol}}(x, X, \Delta) < \widehat{\mathrm{vol}}(y, Y, \mu_*^{-1}\Delta - a(E; X, \Delta)E) \leq A_{(X,\Delta)}(E)n^n.$$

Thus $A_{(X,\Delta)}(E) > \widehat{\mathrm{vol}}(x, X, \Delta)/n^n$ which finishes the proof. $\square$

Next we will discuss application to boundedness generalizing a result by C. Jiang [Jia17, Theorem 1.6].

Corollary 6.14. *Let n be a natural number and c a positive real number. Then the projective varieties X satisfying the following properties:*

- *(X, Δ) is a klt pair of dimension n for some effective $\mathbb{Q}$-divisor Δ,*
- *$-(K_X + \Delta)$ is nef and big,*
- *$\alpha(X, \Delta)^n(-(K_X + \Delta))^n \geq c$,*

form a bounded family.

Proof. By [BJ17, Theorem A and D] (generalizing [Liu18]), for any closed point $x \in X$ we have

$$c \leq \alpha(X, \Delta)^n(-(K_X + \Delta))^n \leq \delta(X, \Delta)^n(-(K_X + \Delta))^n \leq \left(1 + \frac{1}{n}\right)^n \widehat{\mathrm{vol}}(x, X, \Delta).$$

Hence Theorem 6.13 implies (X, Δ) is $(c/(n+1)^n)$-lc. Therefore, the BAB Conjecture proved by Birkar in [Bir16, Theorem 1.1] implies the boundedness of X. $\square$

Remark 6.15. In the conditions of Corollary 6.14 if we also assume that the coefficients of Δ are at least ϵ for any fixed $\epsilon \in (0, 1)$, then such pairs (X, Δ) are log bounded. This partially generalizes [Che18, Theorem 1.4]. Besides, all results should hold for $\mathbb{R}$-pairs.

Question 6.16. Is it true that for any n-dimensional klt singularity $x \in X$, we have $\mathrm{mld}(x, X) \geq \widehat{\mathrm{vol}}(x, X)/n^{n-1}$?

6.3. Miscellaneous questions

6.3.1. Positive characteristics. In this section, we consider a variety X over an algebraically closed field $\mathbb{k}$ of characteristic $p > 0$. From [Har98, HW02], we know that klt singularities are closely related to strongly F-regular singularities in positive characteristic. Moreover, log canonical thresholds (lct) correspond to F-pure thresholds (fpt) in positive characteristic (see [HW02]). In spirit of Theorem 2.6, we define the F-volume of singularities in characteristic p as follows.

Definition 6.17 ([Liu19]). Let X be an n-dimensional strongly F-regular variety over an algebraically closed field $\Bbbk$ of positive characteristic. Let $x \in X$ be a closed point. We define the *F-volume* of $(x \in X)$ as

$$\mathrm{Fvol}(x, X) := \inf_{\mathfrak{a}\colon \mathfrak{m}_x\text{-primary}} \mathrm{fpt}(X; \mathfrak{a})^n \mathrm{mult}(\mathfrak{a}).$$

Similar to [dFEM03], Takagi and Watanabe [TW04] showed that if $x \in X$ is a smooth point, then $\mathrm{Fvol}(x, X) = n^n$.

Another interesting invariant of a strongly F-regular singularity $x \in X$ is its F-signature $s(x, X)$, see [SVdB97, HL02, Tuc12]. In [Liu19], we establish the following comparison result between the F-volume and the F-signature.

Theorem 6.18 ([Liu19]). *Let $x \in X$ be an n-dimensional strongly F-regular singularity. Then*

$$n! \cdot s(x, X) \le \mathrm{Fvol}(x, X) \le n^n \min\{1, n! \cdot s(x, X)\}.$$

It would be interesting to study the limiting behavior of F-volumes of mod-p reductions of a klt singularity over characteristic zero when p goes to infinity.

Conjecture 6.19. *Let $x \in (X, \Delta)$ be a klt singularity over characteristic* 0. *Let $x_p \in (X_p, \Delta_p)$ be its reduction mod $p \gg 0$, then*

$$\widehat{\mathrm{vol}}(x, X, \Delta) = \lim_{p\to\infty} \mathrm{Fvol}(x_p, X_p, \Delta_p).$$

Remark 6.20. Together with Theorem 6.18, this will imply that for the reductions (X_p, Δ_p), the F-signature $s(x_p, X_p, \Delta_p)$ has a uniform lower bound as $p \to \infty$, as asked in [CRST18, Question 5.9].

6.3.2. Relation to local orbifold Euler numbers. In [Lan03], Langer introduced local orbifold Euler numbers for general log canonical surface singularities and used it to prove a Miyaoka–Yau inequality for any log canonical surface. In an attempt to understand Langer's inequality using the Kähler–Einstein metric on a log canonical surface, Borbon–Spotti conjectured recently in [BS17] that the volume densities of the singular Kähler–Einstein metrics should match Langer's local Euler numbers (at least for log terminal surface singularities). They verified this in special examples by comparing the known values of both sides. On the other hand, from Theorem 5.6, we know that the normalized volume is equal to the volume density up to a factor $(\dim X)^{\dim X}$ for any point (X, x) that lives on a Gromov–Hausdorff limit of smooth Kähler–Einstein manifolds ([HS17, LX17a]). In view of this connection, one can formulate a purely algebraic problem about two algebraic invariants of the singularities. This problem was already posed by in [BS17] at least in the log terminal case. We formulate the following form by including one of Langer's expectations (see [Lan03, p.381]):

Conjecture 6.21 (see [BS17, p.37]**).** *Let (X, D, x) be a germ of log canonical surface singularity with $\mathbb{Q}$-boundary. Then we have*

$$e_{\mathrm{orb}}(x, X, D) = \begin{cases} \frac{1}{4}\widehat{\mathrm{vol}}(x, X, D), & \text{if } (X, D) \text{ is log terminal};\\ 0, & \text{if } (X, D) \text{ is not log terminal.}\end{cases} \tag{34}$$

In [Li18b], it was proved that the above conjecture is true when (X, D, x) is a 2-dimensional log-Fano cone or a log-CY cone. In particular, combined with Langer's calculation, one gets the local orbifold Euler numbers of line arrangements.

Proposition 6.22 ([Lan03, Li18b]). *Let $L_1, \ldots, L_n$ be m distinct lines in $\mathbb{C}^2$ passing through 0. Let $D = \sum_{i=1}^{m} \delta_i L_i$, where $0 \le \delta_1 \le \delta_2 \le \cdots \le \delta_m \le 1$. Denote $\delta = \sum_{i=1}^{m} \delta_i$. Then we have:*

$$e_{\mathrm{orb}}(0, \mathbb{C}^2, D) = \begin{cases} 0 & \text{if} \quad (\mathbb{C}^2, D, 0) \text{ is not klt };\\ (1-\delta+\delta_m)(1-\delta_m) & \text{if} \quad \delta < 2\delta_m;\\ \frac{(2-\delta)^2}{4} & \text{if} \quad 2\delta_m \le \delta \le 2. \end{cases} \tag{35}$$

Here we point out a possible application of Theorem 4.15 (i.e., two-dimensional case conjecture 4.1) for studying Conjecture 6.21 for any log terminal singularity (x, X, D). First, by Theorem 4.15 there exists a unique Kollár component $S \cong \mathbb{P}^1$ which minimizes the normalized volume. Let $\mu : Y \to X$ be the extraction of S and $\Delta = \mathrm{Diff}_S(D)$. By Theorem 4.14 we know that $(S, \Delta) \cong (\mathbb{P}^1, \sum_i \delta_i p_i)$ is indeed K-semistable (see [LX16, Section 6]). Then $\mathscr{F} := \Omega^1(\log(S + D))$ (defined using ramified coverings as in [Lan03]) restricted to S fits into an exact sequence of orbifold sheaves:

$$0 \longrightarrow \Omega^1_S(\log(\Delta)) \to \mathscr{F}|_S \to \mathscr{O}_S \to 0. \tag{36}$$

By [Li18b, Theorem 1.3], we know that $\mathscr{E} := \mathscr{F}|_S$ is slope semistable. Then the generalization of [Wah93, Proposition 3.16] to the logarithmic/orbifold setting combined together with Langer's work should imply that $e_{\mathrm{orb}}(x, X, D) = \frac{c_1(\mathscr{E})^2}{4(-S\cdot S)_Y}$ which is indeed equal to $\frac{\widehat{\mathrm{vol}}(\mathrm{ord}_S)}{4}$.

6.3.3. Normalized volume function. We have mainly concentrated on the minimizer of the normalized volume function. We can also ask questions on the general behavior of the normalized volume function. For example:

Question 6.23 (Convexity). Let $\sigma \subset \mathrm{Val}_{X,x}$ be a simplex of quasi-monomial valuations. Is it true that $\widehat{\mathrm{vol}}(\cdot)$ is always convex on σ? Is there a more general convexity property for $\widehat{\mathrm{vol}}$ on $\mathrm{Val}_{X,x}$?

Question 6.24. Is the normalized volume a lower semicontinuous function on $\mathrm{Val}_{X,x}$? If this is true, then it would directly imply the existence of minimizer of $\widehat{\mathrm{vol}}$ using the properness estimate in Theorem 2.5.

Acknowledgment

We would like to use this chance to express our deep gratitude to Gang Tian, from whom we all learn a large amount of knowledge related to K-stability questions on Fano varieties in these years. We want to thank Harold Blum and Sébastien Boucksom for helpful discussions.

Chi Li is partially supported by NSF (Grant No. DMS-1405936) and an Alfred P. Sloan research fellowship. Chenyang Xu is partially supported by the National Science Fund for Distinguished Young Scholars (11425101). A large part of the work was written while he visited the Institut Henri Poincaré under the program 'Poincaré Chair'. He wants to thank the wonderful environment.

References

[AIP+12] K. Altmann, N.O. Ilten, L. Petersen, H. Süss, and R. Vollmert, *The geometry of T-varieties*, Contributions to algebraic geometry, 2012, pp. 17–69.

[And13] D. Anderson, *Okounkov bodies and toric degenerations*, Math. Ann. **356** (2013), no. 3, 1183–1202.

[Bir16] C. Birkar, *Singularities of linear systems and boundedness of Fano varieties*, arXiv:1609.05543 (2016).

[BCHM10] C. Birkar, P. Cascini, C. Hacon, and J. McKernan, *Existence of minimal models for varieties of log general type*, J. Amer. Math. Soc. **23** (2010), no. 2, 405–468.

[BdFFU15] S. Boucksom, T. de Fernex, C. Favre, and S. Urbinati, *Valuation spaces and multiplier ideals on singular varieties*, Recent advances in algebraic geometry, 2015, pp. 29–51.

[BdFF12] S. Boucksom, T. de Fernex, and C. Favre, *The volume of an isolated singularity*, Duke Math. J. **161** (2012), no. 8, 1455–1520.

[BFJ14] S. Boucksom, C. Favre, and M. Jonsson, *A refinement of Izumi's theorem*, Valuation theory in interaction, 2014, pp. 55–81.

[BHJ17] S. Boucksom, T. Hisamoto, and M. Jonsson, *Uniform K-stability, Duistermaat–Heckman measures and singularities of pairs*, Ann. Inst. Fourier (Grenoble) **67** (2017), no. 2, 743–841.

[Blu18] H. Blum, *Existence of Valuations with Smallest Normalized Volume*, Compos. Math. **154** (2018), no. 4, 820–849.

[BJ17] H. Blum and M. Jonsson, *Thresholds, valuations, and K-stability*, arXiv: 1706.04548 (2017).

[BJ18] S. Boucksom and M. Jonsson, *A non-Archimedean approach to K-stability*, arXiv:1805.11160 (2018).

[BL18] H. Blum and Y. Liu, *The normalized volume of a singularity is lower semicontinuous*, arXiv:1802.09658 (2018).

[BS17] M. Borbon and C. Spotti, *Calabi–Yau metrics with conical singularities along line arrangements*, arXiv:1712.07967 (2017).

[CCT02] J. Cheeger, T. Colding, and G. Tian, *On the singularities of spaces with bounded Ricci curvature*, Geom. Funct. Anal. **12** (2002), no. 5, 873–914.

[Che18] W. Chen, *Boundedness of varieties of Fano type with alpha-invariants and volumes bounded below*, arXiv:1810.04019 (2018).

[CDS15] X. Chen, S. Donaldson, and S. Sun, *Kähler–Einstein metrics on Fano manifolds. I: Approximation of metrics with cone singularities, II: Limits with cone angle less than 2π, III: Limits as cone angle approaches 2π and completion of the main proof*, J. Amer. Math. Soc. **28** (2015), no. 1, 183–197, 199–234, 235–278.

[CRST18] J. Carvajal-Rojas, K. Schwede, and K. Tucker, *Fundamental groups of F-regular singularities via F-signature*, Ann. Sci. Éc. Norm. Supér. (4) **51** (2018), no. 4, 993–1016.

[CS16] G. Codogni and J. Stoppa, *Torus equivariant K-stability*, arXiv:1602.03451v2 (2016).

[CS15] T. Collins and G. Székelyhidi, *Sasaki–Einstein metrics and K-stability*, arXiv: 1512.07213, To appear in Geom. Topol. (2015).

[CS18] ______, *K-semistability for irregular Sasakian manifolds*, J. Differential Geom. **109** (2018), no. 1, 81–109.

[Cut13] S. Cutkosky, *Multiplicities associated to graded families of ideals*, Algebra Number Theory **7** (2013), no. 9, 2059–2083.

[dFKX17] T. de Fernex, J. Kollár, and C. Xu, *The dual complex of singularities*, Higher dimensional algebraic geometry–in honour of Professor Yujiro Kawamata's sixtieth birthday, 2017, pp. 103–130.

[dFEM03] T. de Fernex, L. Ein, and M. Mustaţă, *Bounds for log canonical thresholds with applications to birational rigidity*, Math. Res. Lett. **10** (2003), no. 2-3, 219–236.

[DS16] V. Datar and G. Székelyhidi, *Kähler–Einstein metrics along the smooth continuity method*, Geom. Funct. Anal. **26** (2016), no. 4, 975–1010.

[DS14] S. Donaldson and S. Sun, *Gromov–Hausdorff limits of Kähler manifolds and algebraic geometry*, Acta Math. **213** (2014), no. 1, 63–106.

[DS17] ______, *Gromov–Hausdorff limits of Kähler manifolds and algebraic geometry, II*, J. Differential Geom. **107** (2017), no. 2, 327–371.

[ELS03] L. Ein, R. Lazarsfeld, and K. Smith, *Uniform approximation of Abhyankar valuation ideals in smooth function fields*, Amer. J. Math. **125** (2003), no. 2, 409–440.

[FJ04] C. Favre and M. Jonsson, *The valuative tree*, Lecture Notes in Mathematics, vol. 1853, Springer-Verlag, Berlin, 2004.

[FO16] K. Fujita and Y. Odaka, *On the K-stability of Fano varieties and anticanonical divisors*, arXiv:1602.01305, to appear in Tohoku Math. J. (2016).

[Fuj16] K. Fujita, *A valuative criterion for uniform K-stability of $\mathbb{Q}$-Fano varieties*, arXiv:1602.00901, To appear in J. Reine Angew. Math. (2016).

[Fuj18] ______, *Optimal bounds for the volumes of Kähler–Einstein Fano manifolds*, Amer. J. Math. **140** (2018), no. 2, 391–414.

[Ful13] M. Fulger, *Local volumes of Cartier divisors over normal algebraic varieties*, Ann. Inst. Fourier (Grenoble) **63** (2013), no. 5, 1793–1847.

[Gig78] B. Gigena, *Integral invariants of convex cones*, J. Differential Geometry **13** (1978), 191–222.

[Har98] N. Hara, *Classification of two-dimensional F-regular and F-pure singularities*, Adv. Math. **133** (1998), no. 1, 33–53.

[HL02] C. Huneke and G. Leuschke, *Two theorems about maximal Cohen–Macaulay modules*, Math. Ann. **324** (2002), no. 2, 391–404.

[HMX14] C. Hacon, J. McKernan, and C. Xu, *ACC for log canonical thresholds*, Ann. of Math. **180** (2014), 523–571.

[HS17] H.-J. Hein and S. Sun, *Calabi–Yau manifolds with isolated conical singularities*, Publ. Math. IHES **126** (2017), 73–130.

[HW02] N. Hara and K. Watanabe, *F-regular and F-pure rings vs. log terminal and log canonical singularities*, J. Algebraic Geom. **11** (2002), no. 2, 363–392.

[Jia17] C. Jiang, *Boundedness of $\mathbb{Q}$-Fano varieties with degrees and alpha-invariants bounded from below*, arXiv:1705.02740 (2017).

[JM12] M. Jonsson and M. Mustaţă, *Valuations and asymptotic invariants for sequences of ideals*, Ann. Inst. Fourier (Grenoble) **62** (2012), no. 6, 2145–2209.

[KK12] K. Kaveh and A. Khovanskii, *Newton–Okounkov bodies, semigroups of integral points, graded algebras and intersection theory*, Ann. of Math. (2) **176** (2012), no. 2, 925–978, DOI 10.4007/annals.2012.176.2.5.

[KK14] ———, *Convex bodies and multiplicities of ideals*, Proc. Steklov Inst. Math. **286** (2014), no. 1, 268–284.

[KM98] J. Kollár and S. Mori, *Birational geometry of algebraic varieties*, Cambridge Tracts in Mathematics, vol. 134, Cambridge University Press, Cambridge, 1998. With the collaboration of C.H. Clemens and A. Corti; Translated from the 1998 Japanese original.

[Kol04] J. Kollár, *Seifert G_m-bundles*, arXiv:0404386 (2004).

[Kol13] J. Kollár, *Singularities of the minimal model program*, Cambridge Tracts in Mathematics, vol. 200, Cambridge University Press, Cambridge, 2013. With a collaboration of Sándor Kovács.

[Lan03] A. Langer, *Lograrithmic orbifold Euler numbers of surfaces with applications*, Proc. London Math. Soc. **86** (2003), 358–396.

[Laz04] Robert Lazarsfeld, *Positivity in algebraic geometry. II*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 49, Springer-Verlag, Berlin, 2004.

[Li17] C. Li, *K-semistability is equivariant volume minimization*, Duke Math. J. **166** (2017), no. 16, 3147–3218.

[Li18a] ———, *Minimizing normalized volumes of valuations*, Math. Zeit. **289** (2018), no. 1-2, 491–513.

[Li18b] ———, *On the stability of extensions of tangent sheaves on Kähler–Einstein Fano/Calabi–Yau pairs*, arXiv:1803.01734 (2018).

[Liu18] Y. Liu, *The volume of singular Kähler–Einstein Fano varieties*, Compos. Math. **154** (2018), no. 6, 1131–1158.

[Liu19] ———, *The F-volume of singularities in positive characteristic*, In preparation (2019).

[LL19] C. Li and Y. Liu, *Kähler–Einstein metrics and volume minimization*, Adv. Math. **341** (2019), 440–492.

[LM09] R. Lazarsfeld and M. Mustaţă, *Convex bodies associated to linear series*, Ann. Sci. Éc. Norm. Supér. (4) **42** (2009), no. 5, 783–835.

[LS13] A. Liendo and H. Süss, *Normal singularities with torus actions*, Tohoku Mathematical Journal **65** (2013), no. 1, 105–130.

[LWX14] C. Li, X. Wang, and C. Xu, *On proper moduli space of smoothable Kähler–Einstein Fano varieties*, arXiv:1411.0761v3, to appear in Duke Math. J. (2014).

[LWX18] ———, *Algebraicity of the metric tangent cones and equivariant K-stability*, arXiv:1805.03393 (2018).

[LX14] C. Li and C. Xu, *Special test configuration and K-stability of Fano varieties*, Ann. of Math. (2) **180** (2014), no. 1, 197–232.

[LX16] ———, *Stability of Valuations and Kollár Components*, arXiv:1604.05398, to appear in J. Eur. Math. Soc. (2016).

[LX17a] ———, *Stability of Valuations: Higher Rank*, arXiv:1707.05561, to appear in Peking Math. J. (2017).

[LX17b] Y. Liu and C. Xu, *K-stability of cubic threefolds*, arXiv:1706.01933 (2017).

[LZ18] Y. Liu and Z. Zhuang, *Birational superrigidity and K-stability of singular Fano complete intersections*, arXiv:1803.08871 (2018).

[McL16] M. McLean, *Reeb orbits and the minimal discrepancy of an isolated singularity*, Invent. Math. **204** (2016), no. 2, 505–594.

[MM93] T. Mabuchi and S. Mukai, *Stability and Einstein–Kähler metric of a quartic del Pezzo surface*, Einstein metrics and Yang–Mills connections (Sanda, 1990), 1993, pp. 133–160.

[MSY06] D. Martelli, J. Sparks, and S. Yau, *The geometric dual of a-maximisation for toric Sasaki–Einstein manifolds*, Comm. Math. Phys. **268** (2006), 39–65.

[MSY08] ———, *Sasaki–Einstein manifolds and volume minimisation*, Comm. Math. Phys. **280** (2008), no. 3, 611–673.

[Oda15] Y. Odaka, *Compact moduli spaces of Kähler–Einstein Fano varieties*, Publ. Res. Inst. Math. Sci. **51** (2015), no. 3, 549–565.

[OSS16] Y. Odaka, C. Spotti, and S. Sun, *Compact moduli spaces of del Pezzo surfaces and Kähler–Einstein metrics*, J. Differential Geom. **102** (2016), no. 1, 127–172.

[PS11] L. Petersen and H. Süss, *Torus invariant divisors*, Israel J. Math. **182** (2011), 481–504.

[Rei94] M. Reid, *Nonnormal del Pezzo surfaces*, Publ. Res. Inst. Math. Sci. **30** (1994), no. 5, 695–727.

[SS17] C. Spotti and S. Sun, *Explicit Gromov–Hausdorff compactifications of moduli spaces of Kähler–Einstein Fano manifolds*, Pure Appl. Math. Q. **13** (2017), no. 3, 477–515.

[SVdB97] K. Smith and M. Van den Bergh, *Simplicity of rings of differential operators in prime characteristic*, Proc. London Math. Soc. (3) **75** (1997), no. 1, 32–62.

[Tak06] S. Takagi, *Formulas for multiplier ideals on singular varieties*, Amer. J. Math. **128** (2006), no. 6, 1345–1362.

[Tei03] B. Teissier, *Valuations, deformations, and toric geometry*, Valuation theory and its applications, Vol. II (Saskatoon, SK, 1999), 2003, pp. 361–459.

[Tei14] ———, *Overweight deformations of affine toric varieties and local uniformization*, arXiv:1401.5204 (2014).

[Tia87] G. Tian, *On Kähler–Einstein metrics on certain Kähler manifolds with $C_1(M) > 0$*, Invent. Math. **89** (1987), no. 2, 225–246, DOI 10.1007/BF01389077.

[Tia90] ———, *On Calabi's conjecture for complex surfaces with positive first Chern class*, Invent. Math. **101** (1990), no. 1, 101–172.

[Tia94] ———, *The K-energy on hypersurfaces and stability*, Comm. Anal. Geom. **2** (1994), no. 2, 239–265.

[Tia97] ———, *Kähler–Einstein metrics with positive scalar curvature*, Invent. Math. **130** (1997), no. 1, 1–37.

[Tia13] ———, *Partial C^0-estimate for Kähler–Einstein metrics*, Commun. Math. Stat. **1** (2013), no. 2, 105–113.

[Tia15] ———, *K-stability and Kähler–Einstein metrics*, Comm. Pure Appl. Math. **68** (2015), no. 7, 1085–1156.

[Tuc12] K. Tucker, *F-signature exists*, Invent. Math. **190** (2012), no. 3, 743–765.

[TW04] S. Takagi and K. Watanabe, *On F-pure thresholds*, J. Algebra **282** (2004), no. 1, 278–297.

[Wah93] J. Wahl, *Second Chern class and Riemann–Roch for vector bundles on resolutions of surface singularities*, Math. Ann. **295** (1993), 81–110.

[Xu14] C. Xu, *Finiteness of algebraic fundamental groups*, Compos. Math. **150** (2014), no. 3, 409–414.

Chi Li
Department of Mathematics
Purdue University
West Lafayette, IN 47907-2067, USA
e-mail: `li2285@purdue.edu`

Yuchen Liu
Department of Mathematics
Yale University
New Haven, CT 06511, USA
e-mail: `yuchen.liu@yale.edu`

Chenyang Xu
Beijing International Center
for Mathematical Research
Beijing 100871, China
e-mail: `cyxu@math.pku.edu.cn`

Progress in Mathematics, Vol. 333, 221–244

Towards a Liouville Theorem for Continuous Viscosity Solutions to Fully Nonlinear Elliptic Equations in Conformal Geometry

YanYan Li, Luc Nguyen and Bo Wang

Dedicated to Gang Tian on his 60th birthday with friendship

Abstract. We study entire continuous viscosity solutions to fully nonlinear elliptic equations involving the conformal Hessian. We prove the strong comparison principle and Hopf Lemma for (non-uniformly) elliptic equations when one of the competitors is $C^{1,1}$. We obtain as a consequence a Liouville theorem for entire solutions which are approximable by $C^{1,1}$ solutions on larger and larger compact domains, and, in particular, for entire $C^{1,1}_{\rm loc}$ solutions: they are either constants or standard bubbles.

Mathematics Subject Classification (2010). Primary 35J60; Secondary 35J70.

Keywords. Liouville theorem, strong comparison principle, Hopf Lemma, conformal geometry; viscosity solution.

1. Introduction

It is of interest to prove Liouville theorems for entire continuous viscosity solutions of a fully nonlinear elliptic equation of the form

$$f(\lambda(A^u)) = 1, \qquad \lambda(A^u) \in \Gamma, \qquad u > 0 \text{ on } \mathbb{R}^n, \tag{1.1}$$

where the conformal Hessian A^u of u is defined for $n \geq 3$ by

$$A^u = -\frac{2}{n-2}u^{-\frac{n+2}{n-2}}\nabla^2 u + \frac{2n}{(n-2)^2}u^{-\frac{2n}{n-2}}\nabla u \otimes \nabla u - \frac{2}{(n-2)^2}u^{-\frac{2n}{n-2}}|\nabla u|^2 I,$$

I is the $n \times n$ identity matrix, $\lambda(A^u)$ denotes the eigenvalues of A^u, Γ is an open subset of $\mathbb{R}^n$ and $f \in C^0(\bar{\Gamma})$. (See [30], or Definition 2.2 below with $\psi = -\ln u$, for

YanYan Li is partially supported by NSF Grant DMS-1501004. Bo Wang is partially supported by NNSF (11701027) and Beijing Institute of Technology Research Fund Program for Young Scholars.

the definition of viscosity solutions as well as sub- and super-solutions.) Typically, (f,Γ) is assumed to satisfy the following structural conditions.

(i) (f,Γ) is symmetric, i.e.,

$$\text{if } \lambda \in \Gamma \text{ and } \lambda' \text{ is a permutation of } \lambda, \text{ then } \lambda' \in \Gamma \text{ and } f(\lambda') = f(\lambda). \tag{1.2}$$

(ii) (f,Γ) is elliptic, i.e.,

$$\text{if } \lambda \in \Gamma \text{ and } \mu \in \bar{\Gamma}_n, \text{ then } \lambda + \mu \in \Gamma \text{ and } f(\lambda+\mu) \geq f(\lambda), \tag{1.3}$$

where $\Gamma_n := \{\mu \in \mathbb{R}^n : \mu_i > 0\}$ is the positive cone.

(iii) (f,Γ) is locally strictly elliptic, i.e., for any compact subset K of Γ, there is some constant $\delta(K) > 0$ such that

$$f(\lambda+\mu) - f(\lambda) \geq \delta(K)|\mu| \text{ for all } \lambda \in K, \mu \in \bar{\Gamma}_n. \tag{1.4}$$

(iv) f is locally Lipschitz, i.e., for any compact subset K of Γ, there is some constant $C(K) > 0$ such that

$$|f(\lambda') - f(\lambda)| \leq C(K)|\lambda' - \lambda| \text{ for all } \lambda, \lambda' \in K. \tag{1.5}$$

(v) The 1-superlevel set of f stays in Γ, namely

$$f^{-1}([1,\infty)) \subset \Gamma. \tag{1.6}$$

(vi) Γ satisfies

$$\Gamma \subset \Gamma_1 := \{\mu \in \mathbb{R}^n : \mu_1 + \cdots + \mu_n > 0\}. \tag{1.7}$$

It should be noted that equation (1.1) is not necessarily uniformly elliptic and that we do not assume that Γ be convex nor f be concave.

Standard examples of (f,Γ) satisfying (1.2)–(1.7) are given by $(f,\Gamma) = (\sigma_k^{1/k}, \Gamma_k)$, $1 \leq k \leq n$, where σ_k is the kth elementary symmetric function and Γ_k is the connected component of $\{\lambda \in \mathbb{R}^n : \sigma_k(\lambda) > 0\}$ containing the positive cone Γ_n.

Liouville theorems for (1.1) have been studied extensively. We mention here earlier results of Gidas, Ni and Nirenberg [15], Caffarelli, Gidas and Spruck [10] in the semi-linear case, of Viaclovsky [39, 40] for the σ_k-equations for C^2 solutions which are regular at infinity, of Chang, Gursky and Yang [11] for the σ_2-equation in four dimensions, of Li and Li [26, 27] for C^2 solutions, and of Li and Nguyen [32] for continuous viscosity solutions which are approximable by C^2 solutions on larger and larger compact domains.

The key use of the C^2 regularity in the proof of the Liouville theorem in [32] is the strong comparison principle and the Hopf Lemma for (1.1). In fact, if the strong comparison principle and the Hopf Lemma can be established for $C^{1,\alpha}$ solutions ($0 \leq \alpha \leq 1$), a Liouville theorem is then proved in $C^{1,\alpha}$ regularity by the same arguments.

The present note is an exploration in the above direction. We establish the strong comparison principle and the Hopf Lemma when one competitor is $C^{1,1}$, and obtain as a consequence a Liouville theorem in this regularity.

Theorem 1.1 (Strong comparison principle). *Let Ω be an open, connected subset of $\mathbb{R}^n$, $n \geq 3$, Γ be a non-empty open subset of $\mathbb{R}^n$ and $f \in C^0(\bar{\Gamma})$ satisfy* (1.2)–(1.6). *Assume that*

(i) *$u_1 \in USC(\Omega; [0, \infty))$ and $u_2 \in LSC(\Omega; (0, \infty])$ are a sub-solution and a super-solution to $f(\lambda(A^u)) = 1$ in Ω in the viscosity sense, respectively,*
(ii) *and that $u_1 \leq u_2$ in Ω.*

If one of $\ln u_1$ and $\ln u_2$ belongs to $C^{1,1}_{\mathrm{loc}}(\Omega)$, then either $u_1 \equiv u_2$ in Ω or $u_1 < u_2$ in Ω.

Theorem 1.2 (Hopf Lemma). *Let Ω be an open subset of $\mathbb{R}^n$, $n \geq 3$, such that $\partial\Omega$ is C^2 near some point $\hat{x} \in \partial\Omega$, Γ be a non-empty open subset of $\mathbb{R}^n$ and $f \in C^0(\bar{\Gamma})$ satisfying* (1.2)–(1.6). *Assume that*

(i) *$u_1 \in USC(\Omega \cup \{\hat{x}\}; [0, \infty))$ and $u_2 \in LSC(\Omega \cup \{\hat{x}\}; (0, \infty])$ are a sub-solution and a super-solution to $f(\lambda(A^u)) = 1$ in Ω in the viscosity sense, respectively,*
(ii) *and that $u_1 < u_2$ in Ω, and $u_1(\hat{x}) = u_2(\hat{x})$.*

If one of $\ln u_1$ and $\ln u_2$ belongs to $C^{1,1}(\Omega \cup \{\hat{x}\})$, then

$$\liminf_{s \to 0^+} \frac{(u_2 - u_1)(\hat{x} - s\nu(\hat{x}))}{s} > 0,$$

where $\nu(\hat{x})$ is the outward unit normal to $\partial\Omega$ at $\hat{x}$.

Our proof of the strong comparison principle and the Hopf Lemma uses ideas in Caffarelli, Li and Nirenberg [9] and an earlier work of the authors [33]. In fact we establish them for more general equations of the form

$$F(x, \psi, \nabla\psi, \nabla^2\psi) = 1.$$

See Section 2, Theorem 2.3 and Theorem 2.4.

There has been a lot of studies on the (strong) comparison principle and Hopf Lemma for elliptic equations in related contexts. See for instance [1–9, 12–14, 16–25, 29–31, 33–36, 38, 41] and the references therein.

As mentioned earlier, a combination of the above strong comparison principle and Hopf Lemma and the proof of [32, Theorem 1.1] give the following Liouville theorem.

Theorem 1.3 (Liouville theorem). *Assume that $n \geq 3$ and (f, Γ) satisfies* (1.2)–(1.7). *Suppose that there exist $v_k \in C^{1,1}(B_{R_k}(0))$, $R_k \to \infty$, such that $f(\lambda(A^{v_k})) = 1$, $\lambda(A^{v_k}) \in \Gamma$ in the ball $B_{R_k}(0)$ of radius R_k in the viscosity sense, v_k converges uniformly on compact subsets of $\mathbb{R}^n$ to some function $v > 0$. Then*

either (i) *v is identically constant, $0 \in \Gamma$ and $f(0) = 1$,*

or (ii) *v has the form*

$$v(x) = \left(\frac{a}{1 + b^2|x - x_0|^2}\right)^{\frac{n-2}{2}} \tag{1.8}$$

for some $x_0 \in \mathbb{R}^n$ and some $a, b > 0$ satisfying $f(2b^2a^{-2}, \ldots, 2b^2a^{-2}) = 1$.

It is a fact that if u is $C^{1,1}$ in some open set Ω, u satisfies $f(\lambda(A^u)) = 1$ in the viscosity sense in Ω if and only if it satisfies $f(\lambda(A^u)) = 1$ almost everywhere in Ω. See, e.g., Lemma 2.5.

It should be clear that if $0 \in \Gamma$ and $f(0) = 1$, then, by (1.3) and (1.4), $(t, \dots, t) \in \Gamma$ and $f(t, \dots, t) > 1$ for all $t > 0$. Hence if some constant is a solution of (1.1), then all entire solutions of (1.1) are constant, and likewise if some function of the form (1.8) is a solution of (1.1), then all entire solutions of (1.1) are of the form (1.8).

An immediate consequence is:

Corollary 1.4. *Assume that $n \geq 3$ and (f, Γ) satisfies* (1.2)–(1.7). *If $v \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ is a viscosity solution of* (1.1), *then v is either a constant or of the form* (1.8).

The rest of the paper contains two sections. In Section 2, we state and prove our strong comparison principle and the Hopf Lemma for a class of elliptic equations which is more generalized than $f(\lambda(A^u)) = 1$. In Section 3, we prove the Liouville theorem (Theorem 1.3).

2. The strong comparison principle and the Hopf Lemma

In this section we prove the strong comparison principle and the Hopf Lemma for elliptic equations of the form

$$F(x, \psi, \nabla\psi, \nabla^2\psi) = 1 \text{ in } \Omega \tag{2.1}$$

where Ω is an open subset of $\mathbb{R}^n$, $n \geq 1$, $F \in C(\bar{\mathscr{U}})$, $\mathscr{U}$ is a non-empty open subset of $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$, and $(F, \mathscr{U})$ satisfies the following conditions.

(i) $(F, \mathscr{U})$ is elliptic, i.e., for all $(x, s, p, M) \in \mathscr{U}$, $N \in \mathrm{Sym}_n$, $N \geq 0$,

$$(x, s, p, M + N) \in \mathscr{U} \text{ and } F(x, s, p, M + N) \geq F(x, s, p, M). \tag{2.2}$$

Here and below we write $N \geq 0$ for a non-negative definite matrix N.

(ii) For $x \in \bar{\Omega}$, let $\mathscr{U}_x := \{(s, p, M) \in \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n : (x, s, p, M) \in \mathscr{U}\}$. Then, for $x \in \bar{\Omega}$, the 1-superlevel set of $F(x, \cdot)$ stays in $\mathscr{U}_x$, i.e.,

$$F(x, s, p, M) < 1 \text{ for all } x \in \bar{\Omega} \text{ and } (s, p, M) \in \partial\mathscr{U}_x, \tag{2.3}$$

or, equivalently,

$$\{(s, p, M) \in \bar{\mathscr{U}}_x : F(x, s, p, M) \geq 1\} \subset \mathscr{U}_x.$$

(iii) $(F, \mathscr{U})$ is locally strictly elliptic, i.e., for any compact subset $\mathscr{K}$ of $\mathscr{U}$, there is some constant $\delta = \delta(\mathscr{K}) > 0$ such that, for all $(x, s, p, M) \in \mathscr{K}$, $N \in \mathrm{Sym}_n$, $N \geq 0$,

$$F(x, s, p, M + N) - F(x, s, p, M) \geq \delta(\mathscr{K})|N|. \tag{2.4}$$

(iv) F satisfies a local Lipschitz condition with respect to (s, p, M), namely for every compact subset $\mathscr{K}$ of $\mathscr{U}$, there there exists $C(\mathscr{K}) > 0$ such that, for all $(x, s, p, M), (x, s', p', M') \in \mathscr{K}$,

$$|F(x, s, p, M) - F(x, s', p', M')| \le C(\mathscr{K})(|s - s'| + |p - p'| + |M - M'|). \quad (2.5)$$

To keep the notation compact, we abbreviate

$$J_2[\psi] = (\psi, \nabla\psi, \nabla^2\psi) \in \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n.$$

We note that equation (1.1) can be put in the form (2.1) by writing $\psi = -\ln u$, $F(J_2[\psi]) = f(\lambda(A^u))$.

To dispel confusion, we remark that $\mathscr{U}$ is defined as a subset of $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$ rather than that of $\Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$. In particular, the 'local' properties in (iii)–(iv) are local with respect to the (s, p, M)-variables and not the x-variables.

Let us start with the definition of classical and viscosity (sub-/super-)solutions. For this we only need the ellipticity condition (2.2) and the following condition which is weaker than (2.3):

(ii′) There holds

$$F(x, s, p, M) \le 1 \text{ for all } x \in \bar{\Omega} \text{ and } (s, p, M) \in \partial\mathscr{U}_x. \quad (2.6)$$

or, equivalently,

$$\{(s, p, M) \in \bar{\mathscr{U}}_x : F(x, s, p, M) > 1\} \subset \mathscr{U}_x.$$

Definition 2.1 (Classical (sub-/super-)solutions). Let $\Omega \subset \mathbb{R}^n$, $n \ge 1$, be an open set, and $\mathscr{U}$ be a non-empty open subset of $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$ and $F \in C^0(\bar{\mathscr{U}})$ satisfying (2.2) and (2.6). For a function $\psi \in C^2(\Omega)$, we say that

$$F(x, J_2[\psi]) \le 1 \quad (F(x, J_2[\psi]) \ge 1 \text{ resp.}) \quad \text{classically in } \Omega$$

if there holds

$$\text{either } (x, J_2[\psi](x)) \notin \bar{\mathscr{U}} \text{ or } F(x, J_2[\psi](x)) \le 1 \text{ for all } x \in \Omega$$
$$\big(\, (x, J_2[\psi](x)) \in \bar{\mathscr{U}} \text{ and } F(x, J_2[\psi](x)) \ge 1 \text{ for all } x \in \Omega \text{ resp.} \,\big).$$

We say that a function $\psi \in C^2(\Omega)$ is a classical solution of (2.1) in Ω if we have that $(x, J_2[\psi](x)) \in \bar{\mathscr{U}}$ and $F(x, J_2[\psi](x)) = 1$ for every $x \in \Omega$.

When $F(x, J_2[\psi]) \le 1$ ($F(x, J_2[\psi]) \ge 1$, resp.) in Ω, we also say interchangeably that u is a super-solution (sub-solution) to (2.1) in Ω.

In the above definition, the role of condition (2.6) is manifested in the property that if ψ_k is a sequence of super-solutions which converges in C^2 to some ψ, then ψ is also a super-solution. When discussing only sub-solutions, condition (2.6) can be dropped.

Definition 2.2 (Viscosity (sub-/super-)solutions). Let $\Omega \subset \mathbb{R}^n$, $n \ge 1$, be an open set, and $\mathscr{U}$ be a non-empty open subset of $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$ and $F \in C^0(\bar{\mathscr{U}})$ satisfying (2.2) and (2.6). For a function $\psi \in LSC(\Omega; \mathbb{R} \cup \{\infty\})$ ($\psi \in USC(\Omega; \mathbb{R} \cup \{-\infty\})$ resp.), we say that

$$F(x, J_2[\psi]) \le 1 \quad (F(x, J_2[\psi]) \ge 1 \text{ resp.}) \quad \text{in } \Omega$$

in the viscosity sense if for any $x_0 \in \Omega$, $\varphi \in C^2(\Omega)$, $(\psi - \varphi)(x_0) = 0$ and

$$\psi - \varphi \geq 0 \quad (\psi - \varphi \leq 0 \text{ resp.}) \quad \text{near } x_0,$$

there holds

$$\text{either } (x_0, J_2[\varphi](x_0)) \notin \bar{\mathscr{U}} \text{ or } F(x_0, J_2[\varphi](x_0)) \leq 1$$

$$\Big((x_0, J_2[\varphi](x_0)) \in \bar{\mathscr{U}} \text{ and } F(x_0, J_2[\varphi](x_0)) \geq 1 \text{ resp. } \Big).$$

We say that a function $\psi \in C^0(\Omega)$ satisfies (2.1) in the viscosity sense in Ω if we have both that $F(x, J_2[\psi]) \geq 1$ and $F(x, J_2[\psi]) \leq 1$ in Ω in the viscosity sense.

When $F(x, J_2[\psi]) \leq 1$ ($F(x, J_2[\psi]) \geq 1$, resp.) in Ω in the viscosity sense, we also say interchangeably that u is a viscosity super-solution (sub-solution) to (2.1) in Ω.

The main results in this section are the following.

Theorem 2.3 (Strong comparison principle). *Let Ω be an open, connected subset of $\mathbb{R}^n$, $n \geq 1$, $\mathscr{U}$ be a non-empty open subset of $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$ and $F \in C^0(\bar{\mathscr{U}})$ satisfy* (2.2)–(2.5). *Assume that*

(i) *$\psi_1 \in USC(\Omega; \mathbb{R} \cup \{-\infty\})$ and $\psi_2 \in LSC(\Omega; \mathbb{R} \cup \{\infty\})$ are a sub-solution and a super-solution to* (2.1) *in Ω in the viscosity sense, respectively,*
(ii) *and that $\psi_1 \leq \psi_2$ in Ω.*

If one of ψ_1 and ψ_2 belongs to $C^{1,1}_{\mathrm{loc}}(\Omega)$, then either $\psi_1 \equiv \psi_2$ in Ω or $\psi_1 < \psi_2$ in Ω.

Theorem 2.4 (Hopf Lemma). *Let Ω be an open subset of $\mathbb{R}^n$, $n \geq 1$, such that $\partial\Omega$ is C^2 near some point $\hat{x} \in \partial\Omega$, $\mathscr{U}$ be a non-empty open subset of $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$ and $F \in C^0(\bar{\mathscr{U}})$ satisfy* (2.2)–(2.5). *Assume that*

(i) *$\psi_1 \in USC(\Omega \cup \{\hat{x}\}; \mathbb{R} \cup \{-\infty\})$ and $\psi_2 \in LSC(\Omega \cup \{\hat{x}\}; \mathbb{R} \cup \{\infty\})$ are a sub-solution and a super-solution to* (2.1) *in Ω in the viscosity sense, respectively,*
(ii) *and that $\psi_1 < \psi_2$ in Ω, and $\psi_1(\hat{x}) = \psi_2(\hat{x})$.*

If one of ψ_1 and ψ_2 belongs to $C^{1,1}(\Omega \cup \{\hat{x}\})$, then

$$\liminf_{s \to 0^+} \frac{(\psi_2 - \psi_1)(\hat{x} - s\nu(\hat{x}))}{s} > 0,$$

where $\nu(\hat{x})$ is the outward unit normal to $\partial\Omega$ at $\hat{x}$.

If ψ_1 and ψ_2 are continuous and one of them is C^2, the above theorems were proved in Caffarelli, Li, Nirenberg [9].

Before turning to the proof of the above theorems, we give some simple statements about viscosity solutions.

Lemma 2.5. *Let $\Omega \subset \mathbb{R}^n$, $n \geq 1$, be an open set, and $\mathscr{U}$ be a non-empty open subset of $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$ and $F \in C^0(\bar{\mathscr{U}})$ satisfy* (2.2) *and* (2.6). *Suppose that ψ is semi-concave* (*semi-convex resp.*) *in Ω, then*

$$F(x, J_2[\psi]) \leq 1 \quad (F(x, J_2[\psi]) \geq 1 \textit{ resp.}) \quad \textit{in } \Omega \textit{ in the viscosity sense}$$

if and only if

$$\textit{either } (x, J_2[\psi](x)) \notin \bar{\mathscr{U}} \textit{ or } F(x, J_2[\psi](x)) \le 1 \quad \textit{a.e. in } \Omega$$
$$\big((x, J_2[\psi](x)) \in \bar{\mathscr{U}} \textit{ and } F(x, J_2[\psi](x)) \ge 1 \quad \textit{a.e. in } \Omega \textit{ resp.}\big)\,.$$

Recall that ψ is semi-concave (semi-convex resp.) in Ω if there is some $K > 0$ such that $\psi - \frac{K}{2}|x|^2$ ($\psi + \frac{K}{2}|x|^2$ resp.) is locally concave (convex resp.) in Ω. By a theorem of Alexandrov, Buselman and Feller (see, e.g., [8, Theorem 1.5]), semi-concave (or semi-convex) functions are almost everywhere punctually second-order differentiable.

Proof. (a) Consider the inequality $F(x, J_2[\psi]) \le 1$.

Since ψ is semi-concave, it is almost everywhere punctually second-order differentiable. Suppose that $F(x, J_2[\psi]) \le 1$ in Ω in the viscosity sense and x_0 is a point where ψ is punctually second-order differentiable. Then we can use

$$\varphi(x) = \psi(x_0) + \nabla\psi(x_0) \cdot (x - x_0) + (x - x_0)^T \nabla^2\psi(x_0)(x - x_0) - \delta|x - x_0|^2$$

for any $\delta > 0$ as test functions at x_0 to see that

$$\text{either } (x_0, J_2[\psi](x_0) - (0, 0, 2\delta I)) \notin \bar{\mathscr{U}} \text{ or } F(x_0, J_2[\psi](x_0) - (0, 0, 2\delta I)) \le 1.$$

Sending $\delta \to 0$ and using (2.6), we obtain

$$\text{either } (x_0, J_2[\psi](x_0)) \notin \bar{\mathscr{U}} \text{ or } F(x_0, J_2[\psi](x_0)) \le 1.$$

Conversely, assume that either $(x, J_2[\psi](x)) \notin \bar{\mathscr{U}}$ or $F(x, J_2[\psi](x)) \le 1$ for almost all $x \in \Omega$, and suppose, for some $x_0 \in \Omega$ and $\varphi \in C^2(\Omega)$, that $(\psi - \varphi)(x_0) = 0$ and $\psi - \varphi \ge 0$ near x_0. We need to show that

$$\text{either } (x_0, J_2[\varphi](x_0)) \notin \bar{\mathscr{U}} \text{ or } F(x_0, J_2[\varphi](x_0)) \le 1.$$

If $(x_0, J_2[\varphi](x_0)) \notin \mathscr{U}$, we are done by (2.6). We assume henceforth that

$$(x_0, J_2[\varphi](x_0)) \in \mathscr{U}.$$

Replacing φ by $\varphi - \delta|x - x_0|^2$ for some small $\delta > 0$ and letting $\delta \to 0$ eventually, we may assume without loss of generality that

$$\psi > \varphi \text{ in } B_{2r_0}(x_0) \setminus \{x_0\} \subset \Omega \text{ for some } r_0 > 0.$$

For small $\eta > 0$, let $\xi = \xi_\eta = (\psi - \varphi - \eta)^-$ and let Γ_ξ be the concave envelop of ξ in $B_{2r_0}(x_0)$. We have by [8, Lemma 3.5] that

$$\int_{\{\xi = \Gamma_\xi\}} \det(-\nabla^2\Gamma_\xi) \ge \frac{1}{C}(\sup_{B_{2r_0}(x_0)} \xi)^n > 0.$$

In particular, the set $\{\xi = \Gamma_\xi\}$ has positive measure. Thus, we can find $y_\eta \in \{\xi = \Gamma_\xi\}$ such that ψ is punctually second-order differentiable at y_η, either $(y_\eta, J_2[\psi](y_\eta)) \notin \bar{\mathscr{U}}$ or $F(y_\eta, J_2[\psi](y_\eta)) \le 1$ and

$$0 > \xi(y_\eta) = \psi(y_\eta) - \varphi(y_\eta) - \eta \ge -\eta, \tag{2.7}$$
$$|\nabla\xi(y_\eta)| = |\nabla\psi(y_\eta) - \nabla\varphi(y_\eta)| \le C\eta, \tag{2.8}$$
$$\nabla^2\xi(y_\eta) = \nabla^2\psi(y_\eta) - \nabla^2\varphi(y_\eta) \ge 0. \tag{2.9}$$

Recalling that $(x_0, J_2[\varphi](x_0)) \in \mathscr{U}$ and noting that $y_\eta \to x_0$ as $\eta \to 0$, we deduce from (2.2) and (2.7)–(2.9) that, for all small η, $(y_\eta, J_2[\varphi](y_\eta))$, $(y_\eta, J_2[\psi](y_\eta))$ and $(y_\eta, \psi(y_\eta), \nabla\psi(y_\eta), \nabla^2\varphi(y_\eta))$ belong to $\mathscr{U}$. We then have

$$\begin{aligned} 1 &\geq F(y_\eta, J_2[\psi](y_\eta)) \\ &\overset{(2.2),(2.9)}{\geq} F(y_\eta, \psi(y_\eta), \nabla\psi(y_\eta), \nabla^2\varphi(y_\eta)) \\ &\overset{(2.7),(2.8)}{\geq} F(y_\eta, \varphi(y_\eta), \nabla\varphi(y_\eta), \nabla^2\varphi(y_\eta)) + o_\eta(1), \end{aligned}$$

where $o_\eta(1) \to 0$ as $\eta \to 0$ and where we have used the uniform continuity of F on compact subsets of $\bar{\mathscr{U}}$. Letting $\eta \to 0$, we obtain the assertion.

(b) Consider now the inequality $F(x, J_2[\psi]) \geq 1$. This case is treated similarly, but is slightly easier as we do not have a dichotomy in the almost everywhere sense.

Since ψ is semi-convex, it is almost everywhere punctually second-order differentiable. If $F(x, J_2[\psi]) \geq 1$ is satisfied in the viscosity sense, then, as in the previous case, if x_0 is a point where ψ is punctually second-order differentiable, then

$$(x_0, J_2[\psi](x_0) + (0,0,2\delta)) \in \bar{\mathscr{U}} \text{ and } F(x_0, J_2[\psi](x_0) + (0,0,2\delta)) \geq 1 \text{ for any } \delta > 0,$$

and so, upon sending $\delta \to 0$, we obtain

$$(x_0, J_2[\psi](x_0)) \in \bar{\mathscr{U}} \text{ and } F(x_0, J_2[\psi](x_0)) \geq 1.$$

Suppose that $F(x, J_2[\psi](x)) \geq 1$ holds almost everywhere in Ω and suppose, for some $x_0 \in \Omega$ and $\varphi \in C^2(\Omega)$, that $(\psi - \varphi)(x_0) = 0$ and $\psi - \varphi \leq 0$ near x_0. We need to show that

$$F(x_0, J_2[\varphi](x_0)) \geq 1.$$

Replacing φ by $\varphi + \delta|x - x_0|^2$ for some small $\delta > 0$ and letting $\delta \to 0$ eventually, we may assume without loss of generality that

$$\psi < \varphi \text{ in } B_{2r_0}(x_0) \setminus \{x_0\} \subset \Omega \text{ for some } r_0 > 0.$$

For small $\eta > 0$, let $\xi = \xi_\eta = (\psi - \varphi + \eta)^+$ and let Γ_ξ be the concave envelop of ξ in $B_{2r_0}(x_0)$. We have by [8, Lemma 3.5] that

$$\int_{\{\xi = \Gamma_\xi\}} \det(-\nabla^2\Gamma_\xi) \geq \frac{1}{C}(\sup_{B_{2r_0}(x_0)} \xi)^n > 0.$$

In particular, the set $\{\xi = \Gamma_\xi\}$ has positive measure. Thus, we can find $y_\eta \in \{\xi = \Gamma_\xi\}$ such that ψ is punctually second-order differentiable at y_η, $F(y_\eta, J_2[\psi](y_\eta)) \geq 1$ and

$$0 < \xi(y_\eta) = \psi(y_\eta) - \varphi(y_\eta) + \eta \leq \eta, \tag{2.10}$$

$$|\nabla\xi(y_\eta)| = |\nabla\psi(y_\eta) - \nabla\varphi(y_\eta)| \leq C\eta, \tag{2.11}$$

$$\nabla^2\xi(y_\eta) = \nabla^2\psi(y_\eta) - \nabla^2\varphi(y_\eta) \leq 0. \tag{2.12}$$

It follows that

$$\begin{aligned} 1 &\le F(y_\eta, J_2[\psi](y_\eta)) \\ &\overset{(2.2),(2.12)}{\le} F(y_\eta, \psi(y_\eta), \nabla\psi(y_\eta), \nabla^2\varphi(y_\eta)) \\ &\overset{(2.10),(2.11)}{\le} F(y_\eta, \varphi(y_\eta), \nabla\varphi(y_\eta), \nabla^2\varphi(y_\eta)) + o_\eta(1), \end{aligned}$$

where $o_\eta(1) \to 0$ as $\eta \to 0$ and where we have used the uniform continuity of F on compact subsets of $\bar{\mathscr{U}}$. Letting $\eta \to 0$ and noting that $y_\eta \to x_0$, we conclude the proof. $\square$

2.1. Proof of the strong comparison principle

We first prove the strong comparison principle for subsolutions and $C^{1,1}$ strict super-solutions.

Proposition 2.6. *Let Ω be an open, connected subset of $\mathbb{R}^n$, $n \ge 1$, $\mathscr{U}$ be a non-empty open subset of $\bar\Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$ and $F \in C^0(\bar{\mathscr{U}})$ satisfy* (2.2)–(2.3). *Assume that*

(i) $\psi_1 \in USC(\Omega; \mathbb{R} \cup \{-\infty\})$ *satisfies*

$$F(x, J_2[\psi_1]) \ge 1 \text{ in } \Omega \text{ in the viscosity sense,}$$

(ii) $\psi_2 \in C^{1,1}_{\mathrm{loc}}(\Omega)$ *satisfies for some constant $a < 1$,*

$$\text{either } (x, J_2[\psi_2](x)) \notin \bar{\mathscr{U}} \text{ or } F(x, J_2[\psi_2](x)) \le a \qquad \text{a.e. in } \Omega,$$

(iii) $\psi_1 \le \psi_2$ *in Ω and $\psi_1 < \psi_2$ near $\partial\Omega$.*

Then $\psi_1 < \psi_2$ in Ω.

Proof. We follow [33]. Assume by contradiction that there exists some $\hat{x} \in \Omega$ such that $\psi_1(\hat{x}) = \psi_2(\hat{x})$.

Step 1: We regularize ψ_1 using sup-convolution.

This step is well known, see, e.g., [8, Chapter 5].

Take some bounded domain A containing $\hat{x}$ such that $\bar{A} \subset \Omega$ and $\psi_1 < \psi_2$ on ∂A.

We define, for small $\varepsilon > 0$ and $x \in A$,

$$\hat{\psi}_\varepsilon(x) = \sup_{y \in \Omega} \Big(\psi_1(y) - \frac{1}{\varepsilon}|x - y|^2 \Big).$$

It is well known that $\hat{\psi}_\varepsilon \ge \psi_1$, $\hat{\psi}_\varepsilon$ is semi-convex, $\nabla^2 \hat{\psi}_\varepsilon \ge -\frac{2}{\varepsilon} I$ a.e. in A, and $\hat{\psi}_\varepsilon$ converges monotonically to ψ_1 as $\varepsilon \to 0$. Furthermore, for every $x \in A$, there exists $x^* = x^*(\varepsilon, x)$ such that

$$\hat{\psi}_\varepsilon(x) = \psi_1(x^*) - \frac{1}{\varepsilon}|x - x^*|^2. \tag{2.13}$$

We note that if x is a point where $\hat\psi_\varepsilon$ is punctually second-order differentiable, then ψ_1 'can be touched from above' at x^* by a quadratic polynomial:

$$\psi_1(x^*+z) \le \hat\psi_\varepsilon(x) + \frac{1}{\varepsilon}|x^*-x|^2 + \nabla\hat\psi_\varepsilon(x)\cdot z + \frac{1}{2}z^T\nabla^2\hat\psi_\varepsilon(x)z + o(|z|^2) \quad \text{as } z\to 0, \tag{2.14}$$

which is a consequence of the inequalities

$$\hat\psi_\varepsilon(x+z) \le \hat\psi_\varepsilon(x) + \nabla\hat\psi_\varepsilon(x)\cdot z + \frac{1}{2}z^T\nabla^2\hat\psi_\varepsilon(x)z + o(|z|^2), \quad \text{as } z\to 0,$$
$$\hat\psi_\varepsilon(x+z) \ge \psi_1(x^*+z) - \frac{1}{\varepsilon}|x^*-x|^2.$$

(Here we have used the definition of $\hat\psi_\varepsilon$ in the last inequality.)

An immediate consequence of (2.13)–(2.14) and the fact that ψ_1 is a subsolution of (2.1) is that

$$F(x^*, \hat\psi_\varepsilon(x) + \frac{1}{\varepsilon}|x^*-x|^2, \nabla\hat\psi_\varepsilon(x), \nabla^2\hat\psi_\varepsilon(x)) \ge 1. \tag{2.15}$$

Step 2: We proceed to derive a contradiction as in [33].

For small $\eta > 0$, let $\tau = \tau(\varepsilon,\eta)$ be such that

$$\eta = \sup\nolimits_A(\hat\psi_\varepsilon - \psi_2 + \tau).$$

Then

$$\tau = \psi_1(\hat x) - \psi_2(\hat x) + \tau \le \hat\psi_\varepsilon(\hat x) - \psi_2(\hat x) + \tau \le \eta, \tag{2.16}$$
$$\tau = \eta - \sup_A(\hat\psi_\varepsilon - \psi_2) \ \ \ge \eta - \sup_A(\hat\psi_\varepsilon - \psi_1). \tag{2.17}$$

Suppose that ε and η are sufficiently small so that $\xi := \hat\psi_\varepsilon - \psi_2 + \tau$ is negative on ∂A. Let Γ_{ξ^+} denote the concave envelop of $\xi^+ = \max(\xi, 0)$. Since ξ is semi-convex and $\xi \le 0$ on ∂A, we have by [8, Lemma 3.5] that

$$\int_{\{\xi=\Gamma_{\xi^+}\}} \det(-\nabla^2\Gamma_{\xi^+}) \ge \frac{1}{C(\Omega)}(\sup_\Omega \xi)^n > 0.$$

In particular, the set $\{\xi = \Gamma_{\xi^+}\}$ has positive measure. Recall that $\hat\psi_\varepsilon$ and ψ_2 is almost everywhere punctually second-order differentiable, we can find $y = y_{\varepsilon,\eta} \in \{\xi = \Gamma_{\xi^+}\}$ such that $\hat\psi_\varepsilon$ and ψ_2 are punctually second-order differentiable at y, $|J_2[\psi_2](y)| \le C\|\psi\|_{C^{1,1}(\bar A)}$, either $(y, J_2[\psi_2](y)) \notin \bar{\mathscr U}$ or $F(y, J_2[\psi_2](y)) \le a$, and

$$0 < \xi(y) = \hat\psi_\varepsilon(y) - \psi_2(y) + \tau \quad \le \eta, \tag{2.18}$$
$$|\nabla\xi(y)| = |\nabla\hat\psi_\varepsilon(y) - \nabla\psi_2(y)| \ \le C\eta, \tag{2.19}$$
$$\nabla^2\xi(y) = \nabla^2\hat\psi_\varepsilon(y) - \nabla^2\psi_2(y) \le 0. \tag{2.20}$$

We claim that

$$\liminf_{\varepsilon\to 0} \frac{1}{\varepsilon}|y^*-y|^2 \le \eta, \tag{2.21}$$

where $y^* = x^*(\varepsilon, y)$ and x^* is defined in (2.13).

Let us assume (2.21) for now and go on with the proof. From, (2.2), (2.15), (2.20), we have $(y^*, \hat{\psi}_\varepsilon(y) + \frac{1}{\varepsilon}|y^* - y|^2, \nabla\hat{\psi}_\varepsilon(y), \nabla^2\psi_2(y)) \in \bar{\mathscr{U}}$ and

$$\begin{aligned} 1 &\overset{(2.15)}{\le} F(y^*, \hat{\psi}_\varepsilon(y) + \frac{1}{\varepsilon}|y^* - y|^2, \nabla\hat{\psi}_\varepsilon(y), \nabla^2\hat{\psi}_\varepsilon(y)) \\ &\overset{(2.2),(2.20)}{\le} F(y^*, \hat{\psi}_\varepsilon(y) + \frac{1}{\varepsilon}|y^* - y|^2, \nabla\hat{\psi}_\varepsilon(y), \nabla^2\psi_2(y)). \end{aligned} \tag{2.22}$$

By the boundedness of $J_2[\psi_2](y)$, we may assume that

$$(y, J_2[\psi_2](y)) = (y_{\varepsilon,\eta}, J_2[\psi_2](y_{\varepsilon,\eta})) \to (y_0, p_0) \text{ along a sequence } \varepsilon, \eta \to 0. \tag{2.23}$$

By (2.16), (2.17), (2.18) and (2.19), we then have

$$(y^*, \hat{\psi}_\varepsilon(y) + \frac{1}{\varepsilon}|y^* - y|^2, \nabla\hat{\psi}_\varepsilon(y), \nabla^2\psi_2(y)) \to (y_0, p_0).$$

Thus by (2.3) and (2.22), $(y_0, p_0) \in \mathscr{U}$ and $F(y_0, p_0) \ge 1$. But this implies, in view of (2.23), that $(y, J_2[\psi_2](y)) \in \mathscr{U}$ along a sequence $\varepsilon, \eta \to 0$ and so

$$1 \le F(y_0, p_0) = \lim_{\varepsilon,\eta\to 0} F(y, J_2[\psi_2](y)) \le a,$$

which is a contradiction.

To conclude the proof, it remains to establish (2.21).

Proof of (2.21): Suppose for some η and some sequence $\varepsilon_m \to 0$ that $\frac{1}{\varepsilon_m}|y_m^* - y_m|^2 \to d$ where $y_m := y_{\varepsilon_m,\eta}$ and $y_m^* := y^*_{\varepsilon_m,\eta}$. (Note that $\frac{1}{\varepsilon}|x^* - x|^2 \le C$, so this assumption makes sense.) We need to show that $d \le \eta$.

Let $\tau_m = \tau(\varepsilon_m, \eta)$. Without loss of generality, we assume further that $y_m \to y_0$ and $\tau_m \to \tau_0$. By the convergence of y_m and of $\frac{1}{\varepsilon_m}|y_m^* - y_m|^2$, we have that $y_m^* \to y_0$. Thus, by the upper semi-continuity of ψ_1, we have

$$\limsup_{m\to\infty} \psi_1(y_m^*) \le \psi_1(y_0).$$

Hence, by (2.13), (2.16) and the left half of (2.18), we have

$$\begin{aligned} 0 &\le \limsup_{m\to\infty} \frac{1}{\varepsilon_m}|y_m^* - y_m|^2 \overset{(2.13)}{=} \limsup_{m\to\infty} (\psi_1(y_m^*) - \hat{\psi}_{\varepsilon_m}(y_m)) \\ &\overset{(2.18)}{\le} \limsup_{m\to\infty} (\psi_1(y_m^*) - \psi_2(y_m) + \tau_m) \\ &\overset{(2.16)}{\le} \psi_1(y_0) - \psi_2(y_0) + \eta = \lim_{m\to\infty} (\hat{\psi}_{\varepsilon_m}(y_0) - \psi_2(y_0)) + \eta \\ &\le \lim_{m\to\infty} \sup_A (\hat{\psi}_{\varepsilon_m} - \psi_2) + \eta \le \sup_A (\psi_1 - \psi_2) + \eta = \eta. \end{aligned}$$

This proves (2.21) and concludes the proof. □

By analogous arguments, we have:

Proposition 2.7. *Let Ω be an open, connected subset of $\mathbb{R}^n$, $n \ge 1$, $\mathscr{U}$ be a non-empty open subset of $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}_n$ and $F \in C^0(\bar{\mathscr{U}})$ satisfy* (2.2) *and* (2.6). *Assume that*

(i) $\psi_1 \in C^{1,1}_{\rm loc}(\Omega;\mathbb{R})$ *and* $\psi_2 \in LSC(\Omega \cup \{\infty\})$ *satisfy for some constant* $a' > 1$,

$$F(x, J_2[\psi_1]) \geq a' \text{ and } F(x, J_2[\psi_2]) \leq 1 \text{ in } \Omega \text{ in the viscosity sense,}$$

(ii) $\psi_1 \leq \psi_2$ *in* Ω *and* $\psi_1 < \psi_2$ *near* $\partial\Omega$.

Then $\psi_1 < \psi_2$ *in* Ω.

Proof. We argue as in the proof of Proposition 2.6, exchanging the roles of ψ_1 and ψ_2 and sup-convolution and inf-convolution.

Assume by contradiction that there exists some $\hat{x} \in \Omega$ such that $\psi_1(\hat{x}) = \psi_2(\hat{x})$.

Step 1: We regularize ψ_2 by using inf-convolution.

Take some bounded domain A containing $\hat{x}$ such that $\bar{A} \subset \Omega$ and $\psi_1 < \psi_2$ on ∂A.

We define, for small $\varepsilon > 0$ and $x \in A$,

$$\hat{\psi}^\varepsilon(x) = \inf_{y\in\Omega}\left(\psi_2(y) + \frac{1}{\varepsilon}|x-y|^2\right).$$

It is well known that $\hat{\psi}^\varepsilon \leq \psi_2$, $\hat{\psi}^\varepsilon$ is semi-concave, $\nabla^2\hat{\psi}^\varepsilon \leq \frac{2}{\varepsilon}I$ a.e. in A, and $\hat{\psi}^\varepsilon$ converges monotonically to ψ_2 as $\varepsilon \to 0$. Furthermore, for every $x \in A$, there exists $x_* = x_*(\varepsilon, x)$ such that

$$\hat{\psi}^\varepsilon(x) = \psi_2(x_*) + \frac{1}{\varepsilon}|x - x_*|^2. \tag{2.24}$$

We note that if x is a point where $\hat{\psi}^\varepsilon$ is punctually second-order differentiable, then ψ_2 'can be touched from below' at x_* by a quadratic polynomial:

$$\psi_2(x_*+z) \geq \hat{\psi}^\varepsilon(x) - \frac{1}{\varepsilon}|x_* - x|^2 + \nabla\hat{\psi}^\varepsilon(x)\cdot z + \frac{1}{2}z^T\nabla^2\hat{\psi}^\varepsilon(x)z + o(|z|^2), \quad \text{as } z \to 0, \tag{2.25}$$

which is a consequence of the inequalities

$$\hat{\psi}^\varepsilon(x+z) \geq \hat{\psi}^\varepsilon(x) + \nabla\hat{\psi}^\varepsilon(x)\cdot z + \frac{1}{2}z^T\nabla^2\hat{\psi}^\varepsilon(x)z + o(|z|^2), \quad \text{as } z \to 0,$$
$$\hat{\psi}^\varepsilon(x+z) \leq \psi_2(x_*+z) + \frac{1}{\varepsilon}|x_* - x|^2.$$

(Here we have used the definition of $\hat{\psi}^\varepsilon$ in the last inequality.)

An immediate consequence of (2.24)–(2.25) and the fact that ψ_2 is a supersolution of (9) is that either

$$(x_*, \hat{\psi}^\varepsilon(x) - \frac{1}{\varepsilon}|x_* - x|^2, \nabla\hat{\psi}^\varepsilon(x), \nabla^2\hat{\psi}^\varepsilon(x)) \notin \bar{\mathscr{U}}, \tag{2.26}$$

or

$$F(x_*, \hat{\psi}^\varepsilon(x) - \frac{1}{\varepsilon}|x_* - x|^2, \nabla\hat{\psi}^\varepsilon(x), \nabla^2\hat{\psi}^\varepsilon(x)) \leq 1. \tag{2.27}$$

Step 2: We proceed to derive a contradiction as in [33].

For small $\eta > 0$, let $\tau = \tau(\varepsilon, \eta)$ be such that

$$\eta = \sup_A (\psi_1 - \hat{\psi}^\varepsilon + \tau).$$

Then

$$\tau = \psi_1(\hat{x}) - \psi_2(\hat{x}) + \tau \le \psi_1(\hat{x}) - \hat{\psi}^\varepsilon(\hat{x}) + \tau \le \eta, \tag{2.28}$$

$$\tau = \eta - \sup_A (\psi_1 - \hat{\psi}^\varepsilon) \ \ \ge \eta - \sup_A (\psi_2 - \hat{\psi}^\varepsilon). \tag{2.29}$$

Suppose that ε and η are sufficiently small so that $\xi := \psi_1 - \hat{\psi}^\varepsilon + \tau$ is negative on ∂A. Let Γ_{ξ^+} denote the concave envelop of $\xi^+ = \max\{\xi, 0\}$. Since ξ is semi-convex and $\xi \le 0$ on ∂A, we have by [8, Lemma 3.5] that

$$\int_{\{\xi = \Gamma_{\xi^+}\}} \det(-\nabla^2 \Gamma_{\xi^+}) \ge \frac{1}{C(\Omega)} (\sup_\Omega \xi)^n > 0.$$

In particular, the set $\{\xi = \Gamma_{\xi^+}\}$ has positive measure. Recall that $\hat{\psi}^\varepsilon$ and ψ_1 are almost everywhere punctually second-order differentiable, we can find $y = y_{\varepsilon,\eta} \in \{\xi = \Gamma_{\xi^+}\}$ such that $\hat{\psi}^\varepsilon$ and ψ_1 are punctually second-order differentiable at y, $|J_2[\psi_1](y)| \le C\|\psi_1\|_{C^{1,1}(\bar{A})}$,

$$0 < \xi(y) = \psi_1(y) - \hat{\psi}^\varepsilon(y) + \tau \quad \le \eta, \tag{2.30}$$

$$|\nabla \xi(y)| = |\nabla \psi_1(y) - \nabla \hat{\psi}^\varepsilon(y)| \ \le C\eta, \tag{2.31}$$

$$\nabla^2 \xi(y) = \nabla^2 \psi_1(y) - \nabla^2 \hat{\psi}^\varepsilon(y) \le 0, \tag{2.32}$$

and

$$(y, J_2[\psi_1](y)) \in \bar{\mathscr{U}}, \ \ F(y, J_2[\psi_1](y)) \ge a'. \tag{2.33}$$

We claim that

$$\liminf_{\varepsilon \to 0} \frac{1}{\varepsilon} |y_* - y|^2 \le \eta, \tag{2.34}$$

where $y_* = x_*(\varepsilon, y)$ and x_* is defined in (2.24).

Let us assume (2.34) for now and go on with the proof. As in Case 1, we may assume that $(y, J_2[\psi_1](y)) \to (y_0, p_0)$ as $\varepsilon, \eta \to 0$. By (2.33), $F(y_0, p_0) \ge a'$ and so by (2.6), $(y_0, p_0) \in \mathscr{U}$. Also, by (2.28), (2.29), (2.30) and (2.31),

$$\left(y_*, \hat{\psi}^\varepsilon(y) - \frac{1}{\varepsilon}|y_* - y|^2, \nabla \hat{\psi}^\varepsilon(y), \nabla^2 \psi_1(y)\right) \to (y_0, p_0) \text{ as } \varepsilon, \eta \to 0,$$

and so

$$\left(y_*, \hat{\psi}^\varepsilon(y) - \frac{1}{\varepsilon}|y_* - y|^2, \nabla \hat{\psi}^\varepsilon(y), \nabla^2 \psi_1(y)\right) \in \mathscr{U} \text{ along a sequence } \varepsilon, \eta \to 0,$$

Now, we have by (2.2) and (2.32) that (2.27) holds at $x = y$ and so

$$\begin{aligned} 1 &\overset{(2.27)}{\geq} F(y_*, \hat{\psi}^\varepsilon(y) - \frac{1}{\varepsilon}|y_* - y|^2, \nabla\hat{\psi}^\varepsilon(y), \nabla^2\hat{\psi}^\varepsilon(y)) \\ &\overset{(2.32)}{\geq} F(y_*, \hat{\psi}^\varepsilon(y) - \frac{1}{\varepsilon}|y_* - y|^2, \nabla\hat{\psi}^\varepsilon(y), \nabla^2\psi_1(y)) \\ &= F(y, J_2[\psi_1](y)) + o_{\varepsilon,\eta}(1) \\ &\overset{(2.33)}{\geq} a' + o_{\varepsilon,\eta}(1), \end{aligned}$$

where $\lim_{\varepsilon,\eta\to 0} o_{\varepsilon,\eta}(1) = 0$ and where we have used the (local uniform) continuity of F in the second-to-last equality. This gives a contradiction as $a' > 1$.

To conclude the proof, it remains to establish (2.34).

<u>Proof of (2.34):</u> Suppose for some $\eta > 0$ and some sequence $\varepsilon_m \to 0$ that

$$\frac{1}{\varepsilon_m}|(y_m)_* - y_m|^2 \to d$$

where $y_m := y_{\varepsilon_m,\eta}$ and $(y_m)_* := (y_{\varepsilon_m,\eta})_*$. (Note that $\frac{1}{\varepsilon}|x_* - x|^2 \leq C$, so this assumption makes sense.) We need to show that $d \leq \eta$.

Let $\tau_m = \tau(\varepsilon_m, \eta)$. Without loss of generality, we assume further that $y_m \to y_0$ and $\tau_m \to \tau_0$. By the convergence of y_m and of $\frac{1}{\varepsilon_m}|(y_m)_* - y_m|^2$, we have that $(y_m)_* \to y_0$. Thus, by the lower semi-continuity of ψ_2, we have

$$\liminf_{m\to\infty} \psi_2((y_m)_*) \geq \psi_2(y_0).$$

Hence, by (2.24), (2.28) and the left half of (2.30), we have

$$\begin{aligned} 0 &\leq \liminf_{m\to\infty} \frac{1}{\varepsilon_m}|(y_m)_* - y_m|^2 \overset{(2.24)}{=} \liminf_{m\to\infty} \left(\hat{\psi}^{\varepsilon_m}(y_m) - \psi_2((y_m)_*)\right) \\ &\overset{(2.30)}{\leq} \liminf_{m\to\infty}(\psi_1(y_m) - \psi_2((y_m)_*) + \tau_m) \\ &\overset{(2.28)}{\leq} \psi_1(y_0) - \psi_2(y_0) + \eta = \lim_{m\to\infty}(\psi_1(y_0) - \hat{\psi}^{\varepsilon_m}(y_0)) + \eta \\ &\leq \lim_{m\to\infty} \sup_A(\psi_1 - \hat{\psi}^{\varepsilon_m}) + \eta \leq \sup_A(\psi_1 - \psi_2) + \eta = \eta. \end{aligned}$$

This proves (2.34) and concludes the proof. □

We now give the

Proof of Theorem 2.3. Arguing by contradiction, suppose the conclusion is wrong, then we can find a closed ball $\bar{B} \subset \Omega$ of radius $R > 0$ and a point $\hat{x} \in \partial B$ such that

$$\psi_1 < \psi_2 \text{ in } \bar{B} \setminus \{\hat{x}\} \text{ and } \psi_1(\hat{x}) = \psi_2(\hat{x}).$$

Without loss of generality, we assume the center of B is the origin.

<u>Case 1:</u> Consider first the case ψ_2 is $C^{1,1}$.

In the proof, C denotes some generic constant which may vary from lines to lines but depends only on an upper bound for $\|\psi_2\|_{C^{1,1}(\bar\Omega)}$, Ω and $(F, \mathscr{U})$.

In view of Proposition 2.6, it suffices to deform ψ_2 to a strict super-solution $\tilde\psi_2$ in some open ball A around $\hat x$ such that $\tilde\psi_2 > \psi_1$ on ∂A and $\inf_A(\tilde\psi_2 - \psi_1) = 0$. We adapt the argument in [9], which assumes that ψ_2 is C^2.

Using that ψ_2 is $C^{1,1}$, a theorem of Alexandrov, Buselman and Feller (see, e.g., [8, Theorem 1.5]) and Lemma 2.5, we can find some $\Lambda > 0$ and a set Z of zero measure such that ψ_2 is punctually second-order differentiable in $\Omega \setminus Z$,

$$|J_2[\psi_2]| \le \Lambda \text{ in } \Omega \setminus Z. \tag{2.35}$$

and

$$\text{either } (x, J_2[\psi_2](x)) \notin \bar{\mathscr{U}} \text{ or } F(x, J_2[\psi_2](x)) \le 1 \text{ in } \Omega \setminus Z. \tag{2.36}$$

By (2.3), there is some small constant $\theta_0 > 0$

$$F(x,s,p,M) \le 1 - 2\theta_0 \text{ for all } x \in \bar\Omega, (s,p,M) \in \partial\mathscr{U}_x, |s| + |p| + |M| \le \Lambda + 2.$$

Hence

$$\mathscr{K} := \Big\{(x,s,p,M) \in \mathscr{U} : F(x,s,p,M) \ge 1 - \theta_0, x \in \bar\Omega, |s| + |p| + |M| \le \Lambda + 1\Big\}$$

and

$$\mathscr{K}' := \Big\{(x,s,p,M) \in \mathscr{U} : F(x,s,p,M) \ge 1 - \theta_0/2, x \in \bar\Omega, \\ |s| + |p| + |M| \le \Lambda + 1/2\Big\} \subset \mathscr{K}$$

are compact.

For $\alpha > 1$, $\mu > 0$ and $\tau > 0$ which will be fixed later, let

$$\begin{aligned}
E(x) &= E_\alpha(x) = e^{-\alpha|x|^2},\\
h(x) &= h_\alpha(x) = e^{-\alpha|x|^2} - e^{-\alpha R^2},\\
\zeta(x) &= \zeta_\alpha(x) = \cos(\alpha^{1/2}(x_1 - \hat x_1)),\\
\tilde\psi_{\mu,\tau} &= \psi_2 - \mu\,(h - \tau)\zeta.
\end{aligned} \tag{2.37}$$

Let A be a ball centered at $\hat x$ such that $\zeta > \frac{1}{2}$ in A and $\tau_0 = \sup_A h > 0$.

It is clear that, for $0 \le \tau \le \tau_0$ and all sufficiently small μ,

$$\tilde\psi_{\mu,\tau} > \psi_1 \text{ on } \partial A.$$

We compute

$$\begin{aligned}
\nabla\tilde\psi_{\mu,\tau}(x) &= \nabla\psi_2(x) + 2\mu\,\alpha\,E\,\zeta\,x + \mu\alpha^{1/2}(h-\tau)\sin(\alpha^{1/2}(x_1 - \hat x_1))e_1,\\
\nabla^2\tilde\psi_{\mu,\tau}(x) &= \nabla^2\psi_2(x) - 2\mu\alpha E\,\zeta(2\alpha\,x\otimes x - I)\\
&\quad - 2\mu\,\alpha^{3/2}\,E\,\sin(\alpha^{1/2}(x_1 - \hat x_1)(e_1 \otimes x + x \otimes e_1)\\
&\quad + \mu\alpha(h-\tau)\zeta e_1 \otimes e_1.
\end{aligned}$$

We thus have

$$J_2[\tilde\psi_{\mu,\tau}](x) = J_2[\psi_2](x) - \big(0, 0, 4\mu\,\alpha^2\,E\zeta\,x\otimes x + \mu\tau\alpha\zeta e_1\otimes e_1\big) + O(\mu(\alpha^{3/2}E + \alpha^{1/2}\tau)).$$

Now if $x \in A \setminus Z$ is such that

$$(x, J_2[\tilde{\psi}_{\mu,\tau}](x)), (x, J_2[\psi_2](x))$$

and

$$(x, J_2[\tilde{\psi}_{\mu,\tau}] + (0, 0, 4\mu\,\alpha^2\,E\zeta\,x \otimes x + \mu\tau\alpha\zeta e_1 \otimes e_1)$$

lie in $\mathscr{K}$, then

$$\begin{aligned} &F(x, J_2[\psi_2]) + C\mu\alpha^{3/2}E + C\mu\tau\alpha^{1/2} \\ &\quad\overset{(2.5)}{\geq} F\big(x, J_2[\tilde{\psi}_{\mu,\tau}] + (0, 0, 4\mu\,\alpha^2\,E\zeta\,x \otimes x + \mu\tau\alpha\zeta e_1 \otimes e_1)\big) \\ &\quad\overset{(2.4)}{\geq} F\big(x, J_2[\tilde{\psi}_{\mu,\tau}]\big) + \frac{1}{C}\mu\,\alpha^2\,E + \frac{1}{C}\mu\tau\alpha, \end{aligned}$$

and so, by selecting a sufficiently large α, we thus obtain for some $\beta > 0$ and all sufficiently small μ,

$$F\big(x, J_2[\tilde{\psi}_{\mu,\tau}]\big) \leq F\big(x, J_2[\psi_2]\big) - \beta\mu \overset{(2.36)}{\leq} 1 - \beta\mu. \tag{2.38}$$

Now for every $x \in A \setminus Z$ satisfying

$$J_2[\tilde{\psi}_{\mu,\tau}](x) \in \mathscr{U}_x \quad \text{and} \quad F(x, J_2[\tilde{\psi}_{\mu,\tau}](x)) \geq 1 - \theta_0/2,$$

we have, in view of (2.35), that $|J_2[\tilde{\psi}_{\mu,\tau}](x)| \leq \Lambda + 1/2$ and so $(x, J_2[\tilde{\psi}_{\mu,\tau}](x))$ lies in $\mathscr{K}'$ for all small μ. By squeezing μ further, we then have that $(x, J_2[\psi_2](x))$ and $(x, J_2[\tilde{\psi}_{\mu,\tau}] + (0, 0, 4\mu\,\alpha^2\,E\zeta\,x \otimes x + \mu\tau\alpha\zeta e_1 \otimes e_1)$ lie in $\mathscr{K}$. In particular, (2.38) holds.

Taking $\tilde{\beta} = \min(\beta, \frac{\theta_0}{2\mu})$, we thus obtain that

$$\text{either } J_2[\tilde{\psi}_{\mu,\tau}](x) \notin \bar{\mathscr{U}}_x \text{ or } F(x, J_2[\tilde{\psi}_{\mu,\tau}](x)) \leq 1 - \tilde{\beta}\mu \text{ in } A \setminus Z.$$

Noting that

$$\inf_A(\tilde{\psi}_{\mu,0} - \psi_1) \leq 0 \leq \inf_A(\tilde{\psi}_{\mu,\tau_0} - \psi_1),$$

we can select $\tau_1 \in [0, \tau_0]$ such that

$$\inf_A(\tilde{\psi}_{\mu,\tau_1} - \psi_1) = 0.$$

The desired $\tilde{\psi}_2$ is taken to be $\tilde{\psi}_{\mu,\tau_1}$. The conclusion follows from Proposition 2.6.

Case 2: Consider now the case ψ_1 is $C^{1,1}$.

The proof is similar. C will now denote some generic constant which depends only on an upper bound for $\|\psi_1\|_{C^{1,1}(\bar{\Omega})}$, Ω and $(F, \mathscr{U})$.

In view of Proposition 2.7, it suffices to deform ψ_1 to a strict sub-solution $\tilde{\psi}_1$ in some open ball A around $\hat{x}$ such that $\psi_2 > \tilde{\psi}_1$ on ∂A and $\inf_A(\psi_2 - \tilde{\psi}_1) = 0$.

Using that ψ_1 is $C^{1,1}$, a theorem of Alexandrov, Buselman and Feller and Lemma 2.5, we can find some $\Lambda > 0$ and a set Z of zero measure such that ψ_1 is punctually second-order differentiable in $\Omega \setminus Z$,

$$|J_2[\psi_1]| \leq \Lambda \text{ in } \Omega \setminus Z,$$

and, by (2.3),

$$(x, J_2[\psi_1](x)) \in \mathscr{U} \text{ and } F(x, J_2[\psi_1](x)) \geq 1 \text{ in } \Omega \setminus Z. \tag{2.39}$$

For $\alpha > 1$, $\mu > 0$ and $\tau > 0$ which will be fixed later, let E, h, ζ, A, τ_0 be as in Case 1, and amend the definition of $\tilde{\psi}_{\mu,\tau}$ to

$$\tilde{\psi}_{\mu,\tau} = \psi_1 + \mu(h - \tau)\zeta. \tag{2.40}$$

It is clear that, for $0 \leq \tau \leq \tau_0$ and all sufficiently small μ,

$$\tilde{\psi}_{\mu,\tau} < \psi_2 \text{ on } \partial A.$$

As before, we have

$$J_2[\tilde{\psi}_{\mu,\tau}](x) = J_2[\psi_1](x) + (0, 0, 4\mu\alpha^2 E\zeta x \otimes x + \mu\tau\alpha\zeta e_1 \otimes e_1) + O(\mu(\alpha^{3/2}E + \alpha^{1/2}\tau)).$$

It is clear from (2.39) that $(x, J_2[\psi_1](x))$ belongs to $\mathscr{K}'$ for all $x \in A \setminus Z$. We thus have for all sufficiently small μ and $x \in A \setminus Z$ that

$$\begin{aligned}&(x, J_2[\tilde{\psi}_{\mu,\tau}]),\ (x, J_2[\psi_1])\\ &\qquad \text{and } (x, J_2[\tilde{\psi}_{\mu,\tau}] - (0, 0, 4\mu\alpha^2 E\zeta x \otimes x + \mu\tau\alpha\zeta e_1 \otimes e_1)) \text{ lie in } \mathscr{K}.\end{aligned}$$

Therefore,

$$\begin{aligned}&F(x, J_2[\psi_1]) - C\mu\alpha^{3/2}E - C\mu\tau\alpha^{1/2}\\ &\overset{(2.5)}{\leq} F\big(x, J_2[\tilde{\psi}_{\mu,\tau}] - (0, 0, 4\mu\alpha^2 E\zeta x \otimes x + \mu\tau\alpha\zeta e_1 \otimes e_1)\big)\\ &\overset{(2.4)}{\leq} F(x, J_2[\tilde{\psi}_{\mu,\tau}]) - \frac{1}{C}\mu\alpha^2 E - \frac{1}{C}\mu\tau\alpha,\end{aligned}$$

and so, by selecting a sufficiently large α, we thus obtain for some $\beta > 0$ and all sufficiently small μ,

$$F(x, J_2[\tilde{\psi}_{\mu,\tau}]) \geq F(x, J_2[\psi_1]) + \beta\mu \overset{(2.39)}{\geq} 1 + \beta\mu.$$

Noting that

$$\inf_A(\psi_2 - \tilde{\psi}_{\mu,0}) \leq 0 \leq \inf_A(\psi_2 - \tilde{\psi}_{\mu,\tau_0}),$$

we can select $\tau_1 \in [0, \tau_0]$ such that

$$\inf_A(\psi_2 - \tilde{\psi}_{\mu,\tau_1}) = 0.$$

The desired $\tilde{\psi}_1$ is taken to be $\tilde{\psi}_{\mu,\tau_1}$. The conclusion follows from Proposition 2.7 (and Lemma 2.5). $\square$

2.2. Proof of the Hopf Lemma

Proof of Theorem 2.4. We will only consider the case that ψ_2 is $C^{1,1}$, since the case when ψ_1 is $C^{1,1}$ can be treated similarly.

Since $\partial\Omega$ is C^2 near $\hat{x}$, we can find a ball B such that $\bar{B} \subset \Omega \cup \{\hat{x}\}$ and $\hat{x} \in \partial B$. Thus we may assume without loss of generality that $\Omega = B$ is a ball centered at the origin, u_1 and u_2 are defined on $\bar{B}$ and $u_1 < u_2$ in $\bar{B} \setminus \{\hat{x}\}$.

The function $\tilde{\psi}_{\mu,\tau} = \psi_2 - \mu(h-\tau)\zeta$ defined by (2.37) in the proof of Theorem 2.3 satisfies for some open ball A centered at $\hat{x}$, some constant $\beta > 0$, and all $0 \leq \tau \leq \tau_0 := \sup_{A\cap B} h$ that

$$\text{either } J_2[\tilde{\psi}_{\mu,\tau}](x) \notin \bar{\mathscr{U}}_x \text{ or } F(x, J_2[\tilde{\psi}_{\mu,\tau}](x)) \leq 1 - \beta\mu \text{ a.e. in } A \cap B. \tag{2.41}$$

If $\tilde{\psi}_{\mu,0} \geq \psi_1$ in $A \cap B$ for some $\mu > 0$, we are done by the explicit form of h. Suppose otherwise that

$$\inf_{A\cap B} (\tilde{\psi}_{\mu,0} - \psi_1) < 0.$$

Noting that

$$0 \leq \inf_{A\cap B} (\tilde{\psi}_{\mu,\tau_0} - \psi_1),$$

we can find $\tau_1 \in (0, \tau_0]$ such that

$$\inf_{A\cap B} (\tilde{\psi}_{\mu,\tau_1} - \psi_1) = 0.$$

Recall the definition of h, we have also that

$$\inf_{\partial(A\cap B)} (\tilde{\psi}_{\mu,\tau_1} - \psi_1) > 0.$$

Recalling (2.41), we obtain a contradiction to Proposition 2.6. □

3. Proof of the Liouville theorem

In this section, we prove our Liouville theorem. Let us start with some preliminary. Define

$$U = \{M \in \mathrm{Sym}_n : \lambda(U) \in \Gamma\}$$

and

$$F(M) = f(\lambda(M)).$$

By (1.2)–(1.7), we have

(i) (F, U) is elliptic, i.e.,

$$\text{if } M \in U \text{ and } N \geq 0, \text{ then } M + N \in U \text{ and } F(M+N) \geq F(M). \tag{3.1}$$

(ii) (F, U) is locally strictly elliptic, i.e., for any compact subset K of U, there is some constant $\delta(K) > 0$ such that

$$F(M+N) - F(M) \geq \delta(K)|N| \text{ for all } M \in K, N \geq 0. \tag{3.2}$$

(iii) F is locally Lipschitz, i.e., for any compact subset K of U, there is some constant $C(K) > 0$ such that

$$|F(M') - F(M)| \leq C(K)|M' - M| \text{ for all } M, M' \in K. \tag{3.3}$$

(iv) The 1-superlevel set of F stays in U, namely

$$F^{-1}([1, \infty)) \subset U. \tag{3.4}$$

(v) (F, U) is invariant under the orthogonal group $O(n)$, i.e.,

$$\text{if } M \in U \text{ and } R \in O(n), \text{ then } R^t M R \in U \text{ and } F(R^t M R) = F(M). \tag{3.5}$$

(vi) U satisfies

$$\mathrm{tr}(M) \geq 0 \text{ for all } M \in U. \tag{3.6}$$

From (3.1)–(3.4), we see that the strong comparison principle (Theorem 2.3) and the Hopf Lemma (Theorem 2.4) are applicable to the equation $F(A^u) = 1$ by setting $\psi = -\ln u$.

An essential ingredient for our proof is a conformal property of the conformal Hessian A^w, inherited from the conformal structure of $\mathbb{R}^n$. Recall that a map $\varphi : \mathbb{R}^n \cup \{\infty\} \to \mathbb{R}^n \cup \{\infty\}$ is called a Möbius transformation if it is the composition of finitely many translations, dilations and inversions. Now if φ is a Möbius transformation and if we set $w_\varphi = |J_\varphi|^{\frac{n-2}{2n}} w \circ \varphi$ where J_φ is the Jacobian of φ, then

$$A^{w_\varphi}(x) = O_\varphi(x)^t A^w(\varphi(x)) O_\varphi(x)$$

for some orthogonal $n \times n$ matrix $O_\varphi(x)$. In particular, by (3.5),

$$F(A^{w_\varphi}(x)) = F(A^w(\varphi(x))). \tag{3.7}$$

Proof of Theorem 1.3. Having established the Hopf Lemma and the strong comparison principle, we can follow the proof of [32, Theorem 1.1], which draws on ideas from [27], to reach the conclusion. We give a sketch here for readers' convenience. For details, see [32, Section 2].

We use the method of moving spheres. For a function w defined on a subset of $\mathbb{R}^n$, we define

$$w_{x,\lambda}(y) = \frac{\lambda^{n-2}}{|y - x|^{n-2}} w\Big(x + \frac{\lambda^2 (y - x)}{|y - x|^2}\Big)$$

wherever the expression makes sense.

Step 1: We set up the moving sphere method.

Since v_k is locally uniformly bounded, local gradient estimates (see, e.g., [32, Theorem 2.1], [30, Theorem 1.10]), imply that $|\nabla v_k|$ is locally uniformly bounded and so v_k converges to v in $C^{0,\alpha}_{loc}(\mathbb{R}^n)$ and $v \in C^{0,1}_{loc}(\mathbb{R}^n)$.

We note that, by (3.6), v is super-harmonic. Thus, by the positivity of v and the maximum principle, we have

$$v(y) \geq \frac{1}{C}(1 + |y|)^{2-n} \text{ for all } y \in \mathbb{R}^n, \tag{3.8}$$

and so we may also assume without loss of generality that

$$\|v_k - v\|_{C^0(B_{R_k}(0))} \le R_k^{-n} \text{ and } v_k(y) \ge \frac{1}{C}(1+|y|)^{2-n} \text{ for all } y \in B_{R_k}(0). \quad (3.9)$$

Using (3.9) and the local uniform boundedness of $|\nabla v_k|$, one can show that there is a function $\lambda^{(0)} : \mathbb{R}^n \to (0,\infty)$ such that for all k,

$$(v_k)_{x,\lambda} \le v_k \text{ in } B_{R_k}(0) \setminus B_\lambda(x), \forall\ 0 < \lambda < \lambda^{(0)}(x), |x| < R_k/5. \quad (3.10)$$

See [32, Lemma 2.2].

Define, for $|x| < R_k/5$,

$$\bar\lambda_k(x) = \sup\Big\{0 < \mu < R_k/5 : u_{x,\lambda} \le u \text{ in } B_{R_k}(0) \setminus B_\lambda(x), \forall 0 < \lambda < \mu\Big\}.$$

By (3.10), $\bar\lambda_k(x) \in [\lambda^{(0)}(x), R_k/5]$. Set

$$\bar\lambda(x) = \liminf_{k\to\infty} \bar\lambda_k(x) \in [\lambda^{(0)}(x), \infty].$$

$\bar\lambda(x)$ is sometimes referred to as the moving sphere radius of v at x,

Step 2: We show that if $\bar\lambda(x) < \infty$ for some $x \in \mathbb{R}^n$, then

$$\alpha := \liminf_{|y|\to\infty} |y|^{n-2} u(y) = \lim_{|y|\to\infty} |y|^{n-2} v_{x,\bar\lambda(x)}(y) = \bar\lambda(x)^{n-2} v(x) < \infty. \quad (3.11)$$

(Note that $\alpha > 0$ by (3.8).)

We have

$$(v_k)_{x,\bar\lambda_k(x)} \le v_k \text{ in } \mathbb{R}^n \setminus B_{\bar\lambda_k(x)}(x),$$

By the conformal invariance of the conformal Hessian (3.7), $(v_k)_{x,\bar\lambda_k(x)}$ satisfies

$$F(A^{(v_k)_{x,\bar\lambda_k(x)}}) = 1 \quad \text{in } \mathbb{R}^n \setminus \overline{B_{\bar\lambda_k(x)}(x)}.$$

We can now apply the strong comparison principle (Theorem 2.3) and the Hopf Lemma (Theorem 2.4) to conclude that there exists $y_k \in \partial B_{R_k}(0)$ such that $(v_k)_{x,\bar\lambda_k(x)} = v_k(y_k)$. (See the proof of [27, Lemma 4.5].)

It follows that

$$\begin{aligned}\alpha &\le \liminf_{k\to\infty} |y_k|^{n-2} v(y_k) = \liminf_{k\to\infty} |y_k|^{n-2} v_k(y_k)\\ &= \liminf_{k\to\infty} |y_k|^{n-2} (v_k)_{x,\bar\lambda_k(x)}(y_k) = (\bar\lambda(x))^{n-2} v(x) < \infty.\end{aligned}$$

The opposite inequality that $\alpha \ge (\bar\lambda(x))^{n-2} v(x)$ is an easy consequence of the inequality $v_{x,\bar\lambda(x)} \le v$ in $\mathbb{R}^n \setminus B_{\bar\lambda(x)}(x)$. This proves (3.11).

Step 3: We show that either v is constant or $\bar\lambda(x) < \infty$ for all $x \in \mathbb{R}^n$.

Suppose that $\bar\lambda(x_0) = \infty$ for some x_0. Then we have

$$v_{x_0,\lambda} \le v \text{ in } \mathbb{R}^n \setminus B_\lambda(x_0) \text{ for all } \lambda > 0.$$

It follows that, for every unit vector e, the function $r \mapsto r^{\frac{n-2}{2}} v(x_0 + re)$ is non-decreasing. It follows that

$$r^{n-2} \inf_{\partial B_r(x_0)} v \ge r^{\frac{n-2}{2}} \inf_{\partial B_1(x_0)} v$$

and so

$$\alpha = \liminf_{|y|\to\infty} |y|^{n-2} v(y) = \infty.$$

Thus, by Step 2 above, we have $\bar{\lambda}(x) = \infty$ for all $x \in \mathbb{R}^n$. This implies that v is constant; see, e.g., [28], [32, Lemma C.1]. This implies that $0 \in \Gamma$ and $f(0) = 1$.

Step 4: By Steps 2 and 3, it remains to consider the case where, for every $x \in \mathbb{R}^n$, there exists $0 < \bar{\lambda}(x) < \infty$ such that

(i) $v_{x,\bar{\lambda}(x)} \le v$ in $\mathbb{R}^n \setminus B_{\bar{\lambda}(x)}(x)$,

(ii) and

$$\alpha = \lim_{|y|\to\infty} |y|^{n-2} v(y) = \lim_{|y|\to\infty} |y|^{n-2} v_{x,\bar{\lambda}(x)}(y).$$

In some sense, we have a strong comparison principle situation where touching occurs at infinity. If v was $C^{1,1}$, this would imply that $v_{x,\bar{\lambda}(x)} \equiv v$ and so a calculus argument would then show that v has the desired form (see [37, Lemma 11.1]).

Since we have not established the strong comparison principle in $C^{0,1}$ regularity, we resort to a different argument, which was first observed in [27] for C^2 solution and [29] for $C^{0,1}$ solutions. It turns out that, (i) and (ii) together with the super-harmonicity of v imply directly that there exist $a, b > 0$ and $x_0 \in \mathbb{R}^n$ such that

$$u(x) = \left(\frac{a}{1 + b^2|x - x_0|^2}\right)^{\frac{n-2}{2}}.$$

See [29, 32]. This concludes the proof. □

References

[1] M.E. Amendola, G. Galise, and A. Vitolo, *Riesz capacity, maximum principle, and removable sets of fully nonlinear second-order elliptic operators*, Differential Integral Equations, 26 (2013), pp. 845–866.

[2] M. Bardi and F. Da Lio, *On the strong maximum principle for fully nonlinear degenerate elliptic equations*, Arch. Math. (Basel), 73 (1999), pp. 276–285.

[3] H. Berestycki, I. Capuzzo Dolcetta, A. Porretta, and L. Rossi, *Maximum principle and generalized principal eigenvalue for degenerate elliptic operators*, J. Math. Pures Appl. (9), 103 (2015), pp. 1276–1293.

[4] I. Birindelli and F. Demengel, *Comparison principle and Liouville type results for singular fully nonlinear operators*, Ann. Fac. Sci. Toulouse Math. (6), 13 (2004), pp. 261–287.

[5] ———, *Eigenvalue, maximum principle and regularity for fully non linear homogeneous operators*, Commun. Pure Appl. Anal., 6 (2007), pp. 335–366.

[6] I. Birindelli, G. Galise, and H. Ishii, *A family of degenerate elliptic operators: maximum principle and its consequences*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 35 (2018), pp. 417–441.

[7] ———, *Towards a reversed Faber–Krahn inequality for the truncated Laplacian*, preprint, (2018).

[8] L. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, vol. 43 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 1995.

[9] L. Caffarelli, Y.Y. Li, and L. Nirenberg, *Some remarks on singular solutions of nonlinear elliptic equations III: viscosity solutions including parabolic operators*, Comm. Pure Appl. Math., 66 (2013), pp. 109–143.

[10] L.A. Caffarelli, B. Gidas, and J. Spruck, *Asymptotic symmetry and local behavior of semilinear elliptic equations with critical Sobolev growth*, Comm. Pure Appl. Math., 42 (1989), pp. 271–297.

[11] S.-Y.A. Chang, M.J. Gursky, and P.C. Yang, *Entire solutions of a fully nonlinear equation*, in Lectures on partial differential equations, vol. 2 of New Stud. Adv. Math., Int. Press, Somerville, MA, 2003, pp. 43–60.

[12] M.G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.), 27 (1992), pp. 1–67.

[13] I.C. Dolcetta and A. Vitolo, *On the maximum principle for viscosity solutions of fully nonlinear elliptic equations in general domains*, Matematiche (Catania), 62 (2007), pp. 69–91.

[14] ———, *The weak maximum principle for degenerate elliptic operators in unbounded domains*, Int. Math. Res. Not. IMRN, (2018), pp. 412–431.

[15] B. Gidas, W.M. Ni, and L. Nirenberg, *Symmetry of positive solutions of nonlinear elliptic equations in* $\mathbf{R}^n$, in Mathematical analysis and applications, Part A, vol. 7 of Adv. in Math. Suppl. Stud., Academic Press, New York-London, 1981, pp. 369–402.

[16] P. Hartman and L. Nirenberg, *On spherical image maps whose Jacobians do not change sign*, Amer. J. Math., 81 (1959), pp. 901–920.

[17] F.R. Harvey and H.B. Lawson, Jr., *Existence, uniqueness and removable singularities for nonlinear partial differential equations in geometry*, in Surveys in differential geometry. Geometry and topology, vol. 18 of Surv. Differ. Geom., Int. Press, Somerville, MA, 2013, pp. 103–156.

[18] H. Ishii, *On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDEs*, Comm. Pure Appl. Math., 42 (1989), pp. 15–45.

[19] H. Ishii and P.-L. Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations*, J. Differential Equations, 83 (1990), pp. 26–78.

[20] R. Jensen, *The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations*, Arch. Rational Mech. Anal., 101 (1988), pp. 1–27.

[21] B. Kawohl and N. Kutev, *Strong maximum principle for semicontinuous viscosity solutions of nonlinear partial differential equations*, Arch. Math. (Basel), 70 (1998), pp. 470–478.

[22] ———, *Comparison principle and Lipschitz regularity for viscosity solutions of some classes of nonlinear partial differential equations*, Funkcial. Ekvac., 43 (2000), pp. 241–253.

[23] ———, *Comparison principle for viscosity solutions of fully nonlinear, degenerate elliptic equations*, Comm. Partial Differential Equations, 32 (2007), pp. 1209–1224.

[24] S. Koike and T. Kosugi, *Remarks on the comparison principle for quasilinear PDE with no zeroth order terms*, Commun. Pure Appl. Anal., 14 (2015), pp. 133–142.

[25] S. Koike and O. Ley, *Comparison principle for unbounded viscosity solutions of degenerate elliptic PDEs with gradient superlinear terms*, J. Math. Anal. Appl., 381 (2011), pp. 110–120.

[26] A. Li and Y.Y. Li, *On some conformally invariant fully nonlinear equations*, Comm. Pure Appl. Math., 56 (2003), pp. 1416–1464.

[27] ———, *On some conformally invariant fully nonlinear equations. II. Liouville, Harnack and Yamabe*, Acta Math., 195 (2005), pp. 117–154.

[28] Y.Y. Li, *Conformally invariant fully nonlinear elliptic equations and isolated singularities*, J. Funct. Anal., 233 (2006), pp. 380–425.

[29] ———, *Degenerate conformally invariant fully nonlinear elliptic equations*, Arch. Ration. Mech. Anal., 186 (2007), pp. 25–51.

[30] ———, *Local gradient estimates of solutions to some conformally invariant fully nonlinear equations*, Comm. Pure Appl. Math., 62 (2009), pp. 1293–1326. (C. R. Math. Acad. Sci. Paris 343 (2006), no. 4, 249–252).

[31] Y.Y. Li and D.D. Monticelli, *On fully nonlinear CR invariant equations on the Heisenberg group*, J. Differential Equations, 252 (2012), pp. 1309–1349.

[32] Y.Y. Li and L. Nguyen, *Symmetry, quantitative Liouville theorems and analysis of large solutions of conformally invariant fully nonlinear elliptic equations*, Calc. Var. Partial Differential Equations, 56 (2017), pp. Art. 99, 35.

[33] Y.Y. Li, L. Nguyen, and B. Wang, *Comparison principles and Lipschitz regularity for some nonlinear degenerate elliptic equations*, Calc. Var. Partial Differential Equations, 57 (2018), pp. Art. 96, 29.

[34] Y.Y. Li and L. Nirenberg, *A miscellany*, in Percorsi incrociati (in ricordo di Vittorio Cafagna), Collana Scientifica di Ateneo, Universita di Salerno, 2010, pp. 193–208. http://arxiv.org/abs/0910.0323.

[35] Y.Y. Li and B. Wang, *Strong comparison principles for some nonlinear degenerate elliptic equations*, Acta Math. Sci. Ser. B (Engl. Ed.), 38 (2018), pp. 1583–1590.

[36] ———, *Comparison principles for some fully nonlinear subelliptic equations on the Heisenberg group*, Analysis in Theory and Applications, 35 (2019), no. 3, pp. 312–334.

[37] Y.Y. Li and L. Zhang, *Liouville-type theorems and Harnack-type inequalities for semilinear elliptic equations*, J. Anal. Math., 90 (2003), pp. 27–87.

[38] N.S. Trudinger, *Comparison principles and pointwise estimates for viscosity solutions of nonlinear elliptic equations*, Rev. Mat. Iberoamericana, 4 (1988), pp. 453–468.

[39] J.A. Viaclovsky, *Conformal geometry, contact geometry, and the calculus of variations*, Duke Math. J., 101 (2000), pp. 283–316.

[40] ———, *Conformally invariant Monge–Ampère equations: global solutions*, Trans. Amer. Math. Soc., 352 (2000), pp. 4371–4379.

[41] B. Wang, *Removable singularities for degenerate elliptic Pucci operator on the Heisenberg group*, Nonlinear Anal., 160 (2017), pp. 177–190.

YanYan Li
Department of Mathematics
Rutgers University
Hill Center
Busch Campus
110 Frelinghuysen Road
Piscataway, NJ 08854, USA
e-mail: yyli@math.rutgers.edu

Luc Nguyen
Mathematical Institute and St Edmund Hall
University of Oxford
Andrew Wiles Building
Radcliffe Observatory Quarter
Woodstock Road
Oxford OX2 6GG, UK
e-mail: luc.nguyen@maths.ox.ac.uk

Bo Wang
School of Mathematics and Statistics
Beijing Institute of Technology
Beijing 100081, China
e-mail: wangbo89630@bit.edu.cn

Progress in Mathematics, Vol. 333, 245–263

Arsove–Huber Theorem in Higher Dimensions

Shiguang Ma and Jie Qing

Dedicated to 60th birthday of Professor Gang Tian

Abstract. In this note we briefly present the progress in the research project to extend Huber's theory of surfaces to general dimensions. The full paper [42] is in progress. We discuss n-Laplace equations and n-subharmonic functions using nonlinear potential theory. Particularly we build the Brezis–Merle type sharp inequality for Wolff potential and establish Taliaferro's estimates in higher dimensions. We then apply the theory of n-subharmonic functions developed here to study hypersurfaces in hyperbolic space with nonnegative Ricci curvature as well as locally conformal flat manifolds with nonnegative Ricci.

Mathematics Subject Classification (2010). Primary 53A30; Secondary 53C45; 31B15; 35J60; 58J05; 31B05.

Keywords. Arsove–Huber Theorem, n-superharmonic functions, nonlinear potential theory, hypersurfaces in hyperbolic space, locally conformally flat manifolds.

1. Introduction: the story in two dimensions

Thanks to the seminal paper [35] of Huber in 1957 (see also [9, 20, 26, 33]), to explore the connection between geometric properties of surfaces and potential theory based on Gauss curvature equations has been the major part of the theory of surfaces. The Gauss curvature equation in isothermal coordinates on a surface is

$$-\Delta u = Ke^{2u},$$

where K is the Gauss curvature of the surface metric $e^{2u}|dx|^2$. Let us focus on one thread of developments on this subject: local behavior of superharmonic functions near an isolated singular point or asymptotic behavior at infinity of superharmonic functions on the entire plane.

A function that is subharmonic on the entire plane is representable as a function of potential type

$$v(z) = \int_{\mathbb{C}} \log|1 - \frac{z}{\xi}| d\mu(\xi)$$

for $z, \xi \in \mathbb{C}$ the complex plane, where μ is a positive mass distribution and vanishes in a neighborhood of the origin for our purposes. To describe the asymptotic behavior of the function v at infinity one aims to understand the limit

$$\lim_{z\to\infty} \frac{v(z)}{\log|z|}.$$

The first major breakthrough is achieved in [5] in 1973 based on previous works in [4, 29, 32, 34] (cf. [30]). In [5], it was shown that, there is a set E that is thin at infinity and

$$\lim_{z\notin E \text{ and } z\to\infty} \frac{v(z)}{\log|z|} = \limsup_{z\to\infty} \frac{v(z)}{\log|z|}. \tag{1.1}$$

Consequently, this asymptotic result [5, 30] enables Alexander and Currier to study the asymptotic behavior at infinity of complete convex embedded hypersurfaces in hyperbolic space in [2, 3] around 1990. The notion of being thin at infinity is first defined by Brelot in 1940 in [12]. In [5, (1.8)], a subset $E \subset \mathbb{C}$ is said to be thin at infinity if either it is bounded or there exists a function that is subharmonic on the entire plane such that

$$\limsup_{z\in E \text{ and } z\to\infty} \frac{v(z)}{\log|z|} < \limsup_{z\to\infty} \frac{v(z)}{\log|z|}.$$

Most importantly, in [5], the Wiener type criterion for a set to be thin at infinity was established.

Theorem ([5, Theorem 1.3]). *Let E be a Borel subset set in the plane and γ_n be the logarithmic capacity of the part of E lying in the annulus $\{z \in \mathbb{C} : r^n < |z| \le r^{n+1}\}$ for a fixed number $r > 1$ and $n = 1, 2, 3, \ldots$. Then E is thin at infinity if and only if $\gamma_n \to 0$ as $n \to \infty$ and*

$$\sum_{n=1}^{\infty} \frac{n}{\log \frac{1}{\gamma_n}} < \infty. \tag{1.2}$$

Naturally, one asks what is the condition for a function of potential type to have no such set E thin at infinity in (1.1)? It is until very recent this question was solved analytically in [51, Theorem 2.1] in 2006 and geometrically in [10, Lemma 4.2] in 2016. Namely,

Theorem ([10, Lemma 4.2]). *Suppose that $(\mathbb{C}, e^{2u}|dz|^2)$ is complete with nonnegative and bounded Gauss curvature. Then*

$$u(z) = m \log \frac{1}{|z|} + o(\log|z|) \text{ as } |z| \to \infty \tag{1.3}$$

for $m \in [0, 1]$.

It is known that $m = \frac{1}{2\pi}\int_{\mathbb{C}} Ke^{2u}dz$ and $m \in [0,1]$ due to [20, 35], where $m = 0$ implies u is a constant. The proof of the above result in [10] relies on two important ingredients that are deep in geometric analysis and partial differential equation. One is the non-collapsing result of Croke–Karcher [21, Theorem A] in 1988 for complete surfaces with nonnegative Gauss curvature; the other is asymptotic estimates for nonnegative solutions to Gauss curvature type equations of Taliaferro in [51, Theorem 2.1] in 2006 (see also his previous work [49, 50]). One of the key analytic ingredients in [49–51] is based on the Brezis–Merle inequality of Moser–Trudinger type

$$\int_\Omega e^{\frac{(4\pi-\delta)|u(x)|}{\|\Delta u\|_{L^1(\Omega)}}}dx \le (\mathrm{diam}(\Omega))^2\frac{4\pi^2}{\delta} \tag{1.4}$$

for $u|_{\partial\Omega} = 0$ and $\delta \in (0, 4\pi)$, established in [13, Theorem 1] in 1991.

Taliaferro's estimates in [49–51] are the major work in the theory of local behavior of a class of subharmonic functions near an isolated singular point in two dimensions. And, in the spirit of Huber that was reflected in [35], on geometric side, it was a very successful story that the above theorem of sharp local behavior (cf. [10, Lemma 4.2]) turns out to be essential to the proof of [10, Main Theorem] in two dimensions that a complete, nonnegatively curved, immersed surface in hyperbolic 3-space is necessarily properly embedded, except coverings of equidistant surfaces, which was conjectured by Epstein and Alexander–Currier in [2, 3, 23–25] around 1990.

2. n-Laplace equations as higher-dimensional analogues

What can we do in higher dimensions following the approach in [35] by Huber? We have seen successful efforts in [14, 15, 18, 48, 54] to explored higher-dimensional counterparts of Gauss curvature equations such as the scalar curvature equations

$$-\frac{4(n-1)}{n-2}\Delta u + Ru = \bar{R}u^{\frac{n+2}{n-2}},$$

where R and $\bar{R}$ are scalar curvature of the metrics g and $\bar{g} = u^{\frac{4}{n-2}}g$ respectively in dimensions $n \ge 3$, and the higher-order analogue: Q-curvature equations,

$$P_n w + Q_n = \bar{Q}_n e^{2nw},$$

where $P_n = (-\Delta)^n +$ lower order is the so-called Paneitz type operator and $Q_n, \bar{Q}_n$ are so-called Q-curvature of the metrics g and $\bar{g} = e^{2w}g$ respectively in dimensions $2n \ge 2$. We have also seen remarkable successes in using fully nonlinear equations of Weyl–Schouten curvature, as replacements of Gauss curvature equations, in [16, 17, 27, 28]. The above-mentioned seem to represent major developments in conformal geometry and conformally invariant partial differential equations following the approach in [35] by Huber.

2.1. Introduction of n-Laplace equations in conformal geometry

Recall the change of Ricci curvature under conformal change of metrics is

$$\bar{R}_{ij} = R_{ij} - \Delta\phi g_{ij} + (2-n)\phi_{i,j} + (n-2)\phi_i\phi_j + (2-n)|\nabla\phi|^2 g_{ij},$$

where $R_{ij}, \bar{R}_{ij}$ are Ricci curvature tensors for the metrics g and $\bar{g} = e^{2\phi}g$ respectively in n dimensions. Contracting with ϕ^i and ϕ^j on both sides of the above equation, one gets that

$$\phi^i\phi^j\bar{R}_{ij} = \phi^i\phi^j R_{ij} - |\nabla\phi|^{4-n}\mathrm{div}(|\nabla\phi|^{n-2}\nabla\phi).$$

Therefore one arrives at another generalization of Gauss curvature equations in higher dimensions,

$$-|\nabla\phi|^{2-n}\mathrm{div}(|d\phi|^{n-2}\nabla\phi) + \mathrm{Ric}\left(\frac{\nabla\phi}{|\nabla\phi|}, \frac{\nabla\phi}{|\nabla\phi|}\right) = \bar{\mathrm{Ric}}\left(\frac{\bar{\nabla}\phi}{|\bar{\nabla}\phi|}, \frac{\bar{\nabla}\phi}{|\bar{\nabla}\phi|}\right)e^{2\phi}. \quad (2.1)$$

Particularly, when g is Ricci-flat, we have

$$-\Delta_n\phi = \bar{\mathrm{Ric}}\left(\frac{\bar{\nabla}\phi}{|\bar{\nabla}\phi|}, \frac{\bar{\nabla}\phi}{|\bar{\nabla}\phi|}\right)e^{2\phi}|\nabla\phi|^{n-2}, \quad (2.2)$$

where $\Delta_n\phi = \mathrm{div}(|\nabla\phi|^{n-2}\nabla\phi)$ is the so-called n-Laplace operator. In this paper we want to explore properties of n-superharmonic functions and the geometric consequences. Following the approach in [35] by Huber we want to extend the success in two dimensions to higher dimensions and complement contemporary developments in conformal geometry and conformally invariant partial differential equations.

2.2. Non-linear potential theory for n-Laplace equations

To study n-Poisson equations

$$-\Delta_n w = \mu, \quad (2.3)$$

following the roadmap in two dimensions, our first goal is to understand local behavior of n-superharmonic functions near an isolated singular point analogous to (1.1). The theory of n-Laplace equations is as fundamental as that of classic Laplace equations since it is also in the center of the interplay of several important fields of mathematics including calculus of variations, partial differential equations, (nonlinear) potential theory, and mathematical physics, except the principle of superposition is no longer available. We would like to develop the geometric aspects similar to what have been developed for the theory of subharmonic functions in [5, 10, 51] regarding local or asymptotic behavior. The geometric problems that we will use the theory of n-superharmonic functions to study in this paper will be asymptotic behaviors at end of complete embedded convex hypersurfaces in hyperbolic space as well as complete locally conformally flat Ricci nonnegative manifolds.

There is the nonlinear potential theory developed to deal with the lack of the principle of superposition for some general quasilinear and fully nonlinear equations, particularly n-Laplace equations (cf. [31, 43]). The fundamental tool is the Wolff potential

$$W^{\mu}_{1,n}(x_0, r) = \int_0^r \mu(B(x_0,t))^{\frac{1}{n-1}} \frac{dt}{t}, \tag{2.4}$$

where μ is a Radon measure representing the mass distribution. And the foundational estimates in nonlinear potential for the equation (2.3) is as follows:

Theorem ([38, Theorem 1.6]). *Suppose that w is a nonnegative n-superharmonic function satisfying* (2.3) *for a Radon measure μ in $B(x_0, 3r)$. Then*

$$C_1 W^{\mu}_{1,n}(x_0, r) \le w(x_0) \le C_2 \inf_{B(x_0,r)} w + C_3 W^{\mu}_{1,n}(x_0, 2r) \tag{2.5}$$

for some constants $C_1, C_2, C_3 > 0$ which depend only on dimension n.

2.3. Isolated singularity for nonnegative n-superharmonic functions

On the other hand, there are significant developments of the study on local and global behaviors for solutions to (degenerate) quasilinear elliptic equations that include the study of n-Laplace equations, for example, [8, 39, 53] and references therein. The following result on the isolated singularities of nonnegative n-superharmonic functions is particularly useful to us.

Theorem ([8, Proposition 1.1]). *Suppose that w is a nonnegative n-superharmonic function on the punctured ball $B(0,r) \setminus \{0\}$ for some $r > 0$. Assume that w is continuous and $|\nabla w|^n$ is locally integrable in the punctured ball. Furthermore, assume that $-\Delta_n w$ is also locally integrable in the punctured ball. Then, if $\lim_{x\to 0} w(x) = \infty$, then there are a function $g \in L^1(B(0,r))$ and a number $\beta \ge 0$ such that*

$$-\Delta_n w = g + \beta \delta_0 \tag{2.6}$$

in the distributional sense, where δ_0 is the Dirac function at the origin.

Our approach combines that in [39] and the use of the nonlinear potential theory [31, 38] with relevant ideas from [5, 13, 49–51]. In [39, 40], based on [46, 47, 52], it was established that

Theorem ([39, Theorem 1.1]). *Suppose that u is a nonnegative n-harmonic function on the punctured ball $B(0, r_0) \setminus \{0\}$. Then, there is a number γ such that*

$$u(x) - \gamma \log \frac{1}{|x|} \in L^{\infty}_{\mathrm{loc}}(B(0, r_0)). \tag{2.7}$$

2.4. Higher-dimensional analogue of Arsove–Huber estimates

To study the local behavior for a nonnegative n-superharmonic function w on the punctured ball, we follow the idea from [39] to consider the blow-down

$$w_r(\xi) = \frac{w(r\xi)}{\log \frac{1}{r}} \text{ for } \xi \in B(0, \frac{r_0}{r}) \setminus \{0\} \text{ as } r \to 0. \tag{2.8}$$

The first is to observe that, the quotient $\frac{w(x)}{\log\frac{1}{|x|}}$ is mostly uniformly bounded (cf. Lemma 2.1). To be precise we need to use the n-capacity. We refer to [38, Section 3] for the definition. For $x_0 \in \mathbb{R}^n$, we let

$$\omega(x_0, i) = \{x \in \mathbb{R}^n : \ 2^{-i-1} \leq |x - x_0| \leq 2^{-i}\} \text{ and}$$
$$\Omega(x_0, i) = \{x \in \mathbb{R}^n : \ 2^{-i-2} < |x - x_0| < 2^{-i+1}\}.$$

and let

$$\omega(\infty, i) = \{x \in \mathbb{R}^n : \ 2^{i} \leq |x| \leq 2^{i+1}\} \text{ and}$$
$$\Omega(\infty, i) = \{x \in \mathbb{R}^n : \ 2^{i-1} < |x| < 2^{i+2}\}.$$

Definition 2.1. Let $E \subset \mathbb{R}^n$ and $x_0 \in \mathbb{R}^n$. We say E is measure theoretically n-thin at x_0, if

$$\sum_{i=1}^{\infty} i^{n-1} \operatorname{cap}_n(E \cap \omega(x_0, i), \Omega(x_0, i)) < +\infty.$$

We say E is measure theoretically n-thin at ∞, if

$$\sum_{i=1}^{\infty} i^{n-1} \operatorname{cap}_n(E \cap \omega(\infty, i), \Omega(\infty, i)) < +\infty.$$

We remark here that our definition of measure theoretical n-thinness is different from the definition of n-thinness in [1], where it defines that a subset $E \subset \mathbb{R}^n$ is n-thin at $x_0 \in \bar{E} \setminus E$ if

$$W_n(E, x_0) = \int_0^1 \operatorname{Cap}_n(B(x_0, t) \cap E, B(x_0, 2t))^{\frac{1}{n-1}} \frac{dt}{t} < \infty.$$

Later, in [38, Theorem 1.3], it shows that the above definition from [1] is equivalent to the existence of an $\mathcal{A}$-superharmonic function u in the neighborhood of x_0 such that

$$\liminf_{x \to x_0 \text{ and } x \in E} u(x) > u(x_0).$$

We refer these definitions as function theoretic ones. We like to mention that, in [5, Theorem 1.3], Arsove and Huber indeed were able to show that a similar definition of measure theoretical thinness in two dimensions is equivalent to the function theoretic definitions in [12, 38]. We are content to work with measure theoretic thinness from geometric perspective in this paper.

The first major step is the following higher-dimensional analogue of [5, Theorem 1.3] for subharmonic functions in two dimensions.

Theorem 2.1. *Let $1 \leq w \in C^2(B(0,1)\backslash\{0\})$ satisfy*

$$-\Delta_n w \geq 0 \quad \textit{and} \quad \lim_{|x| \to 0} w(x) = +\infty.$$

Then there is a set $E \subset \mathbb{R}^n$, which is measure theoretically n-thin at the origin such that

$$\lim_{|x|\to 0, x\notin E} \frac{w(y)}{\log\frac{1}{|x|}} = \liminf_{|x|\to 0} \frac{w(x)}{\log\frac{1}{|x|}} = m \geq 0$$

and

$$w(x) \geq m \log\frac{1}{|x|} \text{ for } x \in B(0,1)\setminus\{0\}.$$

Moreover, if $(B(0,1)\setminus\{0\}, e^{2w}|dx|^2)$ is complete at the origin, then $m \geq 1$.

The first key ingredient to prove Theorem 2.1, based on the nonlinear potential theory in [31, 38, 41, 43] and the earlier work of Arsove and Huber [5], is as follows:

Lemma 2.1. *Assume the same assumptions as in Theorem* 2.1*. Then, there is a set $\hat{E}$, which is measure theoretically n-thin at the origin, and a constant $\hat{C}$ such that the quotient*

$$0 \leq \frac{w(x)}{\log\frac{1}{|x|}} \leq \hat{C} \tag{2.9}$$

for $x \in (B(0,1)\setminus\{0\})\setminus\hat{E}$.

The proof of the above lemma starts with the following simple fact observed in [38, Lemma 3.9].

Lemma ([38, Lemma 3.9]). *Suppose that $u \in W_0^{1,n}(\Omega)$ is an n-superharmonic function satisfying*

$$-\Delta_n u = \mu$$

for a Radon measure μ. Then, for $\lambda > 0$,

$$\lambda^{n-1}\operatorname{Cap}_n(\{x\in\Omega : u(x) > \lambda\}, \Omega) \leq \mu(\Omega). \tag{2.10}$$

The proof is to use $\frac{\min\{u,\lambda\}}{\lambda}$ as a test function and immediately gets

$$\int_\Omega |\nabla \min\{u,\lambda\}|^n \leq \frac{\mu(\Omega)}{\lambda^{n-1}}$$

which is easily seen to imply the above n-capacity estimate (2.10). For the purpose of working with all scales, we consider

$$\omega_i = \omega(0,i) = \{x\in\mathbb{R}^n : 2^{-1-i} \leq |x| \leq 2^{-i}\} = 2^{-i}\omega(0,0)$$
$$\Omega_i = \Omega(0,i) = \{x\in\mathbb{R}^n : 2^{-2-i} \leq |x| \leq 2^{-i+1}\} = 2^{-i}\Omega(0,0).$$

Hence

$$\frac{i}{2}\log 2 \leq (i-1)\log 2 \leq \log\frac{1}{|x|} \leq (i+2)\log 2 \leq 2i\log 2 \text{ for all } x\in\Omega_i.$$

In order to use the above-mentioned [38, Lemma 3.9] to prove Lemma 2.1, we will rely on the fundamental estimates in non-linear potential theory (2.5) of [38, Theorem 1.6], Lemma 2.3 in the next subsection, and the following basic existence result (cf. [37, Theorem 2.4])

Lemma 2.2. *For a bounded domain $\Omega \subset \mathbb{R}^n$ and a Radon measure*

$$\mu_f(E) = \int_{E\cap\Omega} f(x)dx$$

of a function $f \in C(\bar{\Omega})$, there always exists a solution $u(x) > 0$ to the equation

$$\begin{cases} -\Delta_n u = \mu_f \ \text{ in } \Omega \\ \quad\quad u = 0 \ \text{ on } \partial\Omega. \end{cases}$$

We also need the following estimates.

Lemma 2.3. *Suppose that w is nonnegative and satisfies*

$$-\Delta_n w = g + \beta\delta_0 \ \text{ in } B(0,2)$$

for some $g \in L^1(B(0,2))$ and $\beta \geq 0$. Then there is a constant C such that

$$\inf_{B(x_0,\frac{|x_0|}{2})} w(x) \leq C \log\frac{1}{|x_0|} \ \text{ for all } B(x_0, \frac{|x_0|}{2}) \subset B(0,2). \tag{2.11}$$

Lemma 2.3 is derived from [22, Lemma 14 and Lemma 15] and [36, Lemma 1].

The second key ingredient in the proof of Theorem 2.1, for the sake of the blow-down argument as the one used in [39], is to modify the function $\frac{w(r\xi)}{\log\frac{1}{r}}$ to accommodate the lack of boundedness. We use the trick from [22] and consider the cut-off function

$$a_\alpha(s) = \begin{cases} s & \text{when } 0 \leq s \leq \alpha \\ \alpha + \int_\alpha^s (\frac{\alpha}{t})^{\frac{n}{n-1}} dt & \text{when } s > \alpha, \end{cases}$$

where α is to be fixed as $\hat{C}+1$ throughout this paper, where $\hat{C}$ is the one in (2.9). One may calculate that

$$\begin{aligned} a_\alpha(s) &\leq n\alpha \\ \min\{1, (\frac{\alpha}{s})^{\frac{n}{n-1}}\} &\leq a'_\alpha(s) \leq \frac{a_\alpha(s)}{s} \leq 1 \\ -\frac{4n\alpha}{n-1}(1+s)^{-2} &\leq a''_\alpha(s) \leq 0 \\ -\Delta_n a_\alpha(w) &= -(a'_\alpha)^{n-1}\Delta_n w - (n-1)(a'_\alpha)^{n-2}a''_\alpha|\nabla w|^n. \end{aligned} \tag{2.12}$$

Now we are to carry out the blow-down argument as in [39]. For each $r > 0$ and small, we consider the modified blow-down

$$\hat{w}_r(\xi) = a_\alpha(w_r(\xi)) = a_\alpha(\frac{w(r\xi)}{\log\frac{1}{r}}). \tag{2.13}$$

Clearly, we have

$$0 \leq \hat{w}_r(\xi) \leq n\alpha = n(\hat{C}+1) \tag{2.14}$$

for $\xi \in A_{0,\frac{1}{r}} = \{\xi \in \mathbb{R}^n : |\xi| \in (0, \frac{1}{r})\}$ and

$$-\Delta_n^\xi \hat{w}_r(\xi) = \frac{r^n}{(\log \frac{1}{r})^{n-1}} (a'_\alpha)^{n-1} \mu(r\xi) - (n-1)(\frac{r}{\log \frac{1}{r}})^n (a'_\alpha)^{n-2} a''_\alpha |\nabla w|^n (r\xi) \tag{2.15}$$

for $\xi \in A_{0,\frac{1}{r}}$.

Lemma 2.4. *Assume the same assumptions as in Theorem* 2.1*. Then the modified blow-down $\hat{w}_r(\xi)$ is a nonnegative n-superharmonic function satisfying*

$$-\Delta_n^\xi \hat{w}_r = \hat{g}_r \geq 0 \text{ in } A_{0,\frac{1}{r}}$$

for a function $\hat{g}_r \in L^1(B(0, \frac{1}{r}))$ and $\hat{w}_r(\xi) \leq n\hat{C} + n$ for all $x \in A_{0,\frac{1}{r}}$. More importantly, for any fixed $R > 0$,

$$\int_{B(0,R)} \hat{g}_r d\xi \leq \int_{A_{0,rR}} (-\Delta_n^x w) dx + C \int_{B(0,rR)} \frac{|\nabla w|^n}{(1+w)^2} dx \to 0 \text{ as } r \to 0. \tag{2.16}$$

Lemma 2.4 implies that there is no concentration other than that possibly at the origin. Therefore, at least, for sequences $r_k \to 0$, one can manage to show that $\hat{w}_{r_k}(\xi)$ converges to a bounded n-harmonic function on the entire space $\mathbb{R}^n$ except possibly the origin, which can only be a constant due to [44] because the origin and the infinity are removable singularities by [46].

The third key ingredient in the proof of Theorem 2.1 is to show the uniqueness of all possible limits of blow-down. In the other words, the issue now is what are the possible limit constants? We continue to use the approach used as in [39]. One of the key tool is the following weak comparison principle as a consequence of [52, Lemma 3.1] (please also see [39, Corollary 1.1] and the comment in [40]).

Lemma 2.5. *Assume Ω is a connected open subset of $\mathbb{R}^n \setminus \{0\}$ and u is n-superharmonic in Ω. Then*

$$\inf_{\partial\Omega} \frac{u(x)}{\log \frac{1}{|x|}} \leq \inf_{\Omega} \frac{u(x)}{\log \frac{1}{|x|}}. \tag{2.17}$$

For any blow-down sequence $\hat{w}_{r_k}(\xi)$ with $r_k \to 0$, there is ξ_k with $|\xi_k| = 1$ and

$$\hat{w}_{r_k}(\xi_k) = w_{r_k}(\xi_k) = \frac{w(r_k\xi_k)}{\log \frac{1}{r_k}} = \min_{|x|=r_k} \frac{w(x)}{\log \frac{1}{|x|}} \to \liminf_{x\to 0} \frac{w(x)}{\log \frac{1}{|x|}} < \hat{C}.$$

Because, in the light of the weak comparison principle Lemma 2.5, the quotient $\min_{|x|=r} \frac{w(x)}{\log \frac{1}{|x|}}$ is non-increasing as $r \to 0$.

2.5. Higher-dimensional analogue of Taliaferro's estimates

Let us start with Taliaferro's estimates in two dimensions.

Theorem ([51, Theorem 2.1]). *Suppose that u is C^2 positive solution to*

$$0 \leq -\Delta u \leq f(u)$$

in a punctured neighborhood of the origin in $\mathbb{R}^2$, where $f : (0,\infty) \to (0,\infty)$ is a continuous function such that

$$\log f = O(t) \text{ as } t \to \infty.$$

Then, either u has a C^1 extension to the origin or

$$\lim_{x\to 0} \frac{u(x)}{\log \frac{1}{|x|}} = m \tag{2.18}$$

for some finite positive number m.

This can be viewed as the improvement of [5, Theorem 1.3], having eliminated the possible subset that is thin at the origin and where the local behavior may differ from (2.18). Our next goal is to establish the higher-dimensional analogue of [51, Theorem 2.1] as follows:

Theorem 2.2. *Let $1 \leq w \in C^2(B(0,1) \setminus \{0\})$ satisfies*

$$0 \leq -\Delta_n w \leq g(x, w, \nabla w) \tag{2.19}$$

on a punctured neighborhood of the origin in $\mathbb{R}^n$ and that

$$\lim_{x\to 0} w(x) = +\infty,$$

where g is a nonnegative function satisfying

$$g(x, w, \nabla w) \leq C|\nabla w|^{n-2} e^{2w} \tag{2.20}$$

for some fixed constant C. Then

$$\lim_{|x|\to 0} \frac{w(x)}{\log \frac{1}{|x|}} = m \geq 0$$

and

$$w(x) \geq m \log \frac{1}{|x|} \text{ for } x \in B(0,1) \setminus \{0\}.$$

Moreover, if $(B(0,1) \setminus \{0\}, e^{2w}|dx|^2)$ is complete at the origin, then $m \geq 1$.

The key analytic tool to remove the possibility of concentrating for solutions to n-Laplace equations like (2.2) and (2.19) with growth condition (2.20) is the higher-dimensional analogue of the borderline Sobolev inequality established by Brezis and Merle in two dimensions in [13, Theorem 1] in 1991. We extend that to higher dimensions.

Proposition 2.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with the diameter D. And let $f \in L^1(\Omega)$ be nonnegative. Then*

1. $$\int_\Omega e^{\frac{(n-\delta)W_{1,n}^{\mu_f}(x,D)}{\|f\|_{L^1(\Omega)}}}dx \le C(\delta, D), \tag{2.21}$$

where $\delta \in (0,n)$*, the measure is defined as*

$$\mu(U) = \mu_f(U) = \int_{U\cap\Omega} f dx$$

and

$$W_{1,n}^{\mu}(x,D) = \int_0^D \mu(B(x,t))^{\frac{1}{n-1}}\frac{dt}{t}$$

is the associated Wolff potential.

2. *If* $p > n$ *and* $\|f\|_{L^1(\Omega)} \le (\frac{n}{p})^p$*, then for* $\delta \in (0,p)$

$$\int_\Omega \exp((p-\delta)W_{1,n}^{\mu}(x,R))dx \le \frac{p\cdot 2^q}{\delta}(R^n\|f\|_{L^1(\Omega)} + |\Omega|). \tag{2.22}$$

The proof is a nice use of the Hardy–Littlewood maximal function and its weak L^1 estimates with the assistance of Jensen's inequality. In contrast to the proof of Theorem 2.1 in the previous subsection, we will be able to show, based on the growth condition (2.20) and Proposition 2.1, the quotient $\frac{w(x)}{\log\frac{1}{|x|}}$ is bounded: the analogue of [51, Theorem 2.3].

Lemma 2.6. *Assume the same assumptions as in Theorem* 2.2*. Then the quotient*

$$\frac{w(x)}{\log\frac{1}{|x|}}$$

is uniformly bounded in the punctured ball $B(0,1)\setminus\{0\}$*.*

We prove Lemma 2.6 by contradiction. Assume otherwise, there is a sequence $\{x_k\}$ inside the punctured ball such that

$$\frac{w(x_k)}{\log\frac{1}{|x_k|}} \to \infty \text{ as } |x_k| \to 0.$$

One may consider the blow-up sequence

$$v_k(\xi) = w(x_k + \frac{|x_k|}{4}\xi) \text{ for } \xi \in B(0,2)$$

and calculate

$$\begin{aligned}-\Delta_n^\xi v_k &= -\left(\frac{|x_k|}{4}\right)^n \Delta_n^x w\left(x_k + \frac{|x_k|}{4}\xi\right)\\ &= g_k(\xi) \le C|x_k|^2|\nabla^\xi v_k|^{n-2}e^{2v_k} \text{ for } \xi\in B(0,2)\end{aligned}$$

$$\int_{B(0,2)} g_k(\xi)d\xi = \int_{B(x_k,\frac{|x_k|}{2})} g(x)dx \to 0 \text{ as } k\to\infty$$

where $-\Delta_n w = g + \beta\delta_0$ and $g \in L^1_{\text{loc}}(B(0,2))$ according to [8, Proposition 1.1]. We will argue in the similar way to that in [51]. For convenience, let us denote

$$\lambda_k = \log \frac{1}{|x_k|} \to \infty \text{ as } k \to \infty.$$

Then it is implied from [38, Theorem 1.6] and Lemma 2.3 that

$$\frac{1}{\lambda_k} W^{\mu_{g_k}}_{1,n}(0,2) \to \infty \tag{2.23}$$

$$g_k(\xi) \le C|x_k|^2 |\nabla^\xi v_k|^{n-2} e^{C_1\lambda_k + C_2 W^{\mu_{g_k}}_{1,n}(\xi,2)} \text{ for } \xi \in B(0,1). \tag{2.24}$$

A very important observation is that, when dealing with competing terms like λ_k and $W^{\mu_{g_k}}_{1,n}(0,2)$, for

$$\Omega_k = \{\xi \in B(0,1) : W^{\mu_{g_k}}_{1,n}(\xi,2) \ge \lambda_k\}$$

we have

$$\begin{aligned}
\int_{\Omega_k} |g_k|^{\frac{n-1}{n-2}} d\xi &\le C|x_k|^{\frac{2(n-1)}{n-2}} \int_{\Omega_k} |\nabla^\xi v_k|^{n-1} e^{\frac{2(n-1)}{n-2} v_k} d\xi \\
&\le C|x_k|^{\frac{2(n-1)}{n-2}} \int_{\Omega_k} |\nabla^\xi v_k|^{n-1} e^{\frac{2(n-1)}{n-2}(C_1 \inf_{B(0,1)} v_k + C_2 W^{\mu_{g_k}}_{1,n}(\xi,2))} d\xi \\
&\le C|x_k|^{\frac{2(n-1)}{n-2}} \int_{B(0,1)} |\nabla^\xi v_k|^{n-1} e^{C_3 W^{\mu_{g_k}}_{1,n}(\xi,2))} d\xi \\
&\le C.
\end{aligned} \tag{2.25}$$

Make a note that $\frac{n-1}{n-2} > 1$. The last step in the above inequalities relies on Proposition 2.1 and the simple L^p-gradient estimates for n-superharmonic functions for any $p < n$. This implies that

$$\mu_{g_k}(B(0,t) \cap \Omega_k) \le C t^{\frac{1}{n-1}}$$

for some positive constant $C > 0$. Observe that

$$\mu_{g_k}(B(0,t))^{\frac{1}{n-1}} \le \mu_{g_k}(B(0,t) \cap \Omega_k)^{\frac{1}{n-1}} + \mu_{g_k}(B(0,t) \setminus \Omega_k)^{\frac{1}{n-1}}$$

which implies

$$W^{\mu_{g_k}}_{1,n}(0,2) \le C + C \int_0^2 \mu_{g_k}(B(0,t) \setminus \Omega_k)^{\frac{1}{n-1}} \frac{dt}{t}. \tag{2.26}$$

To estimate the second term on the right side the above equation, one notices that, for $\xi \in B(0,1) \setminus \Omega_k$,

$$g_k(\xi) \le C|x_k|^2 |\nabla^\xi v_k|^{n-2} e^{C_2\lambda_k + C_3 W^{\mu_{g_k}}_{1,n}(\xi,2)} \le C|x_k|^2 |\nabla^\xi v_k|^{n-2} e^{C_4\lambda_k}$$

from (2.24). Therefore

$$\begin{aligned}\int_{B(0,t)\setminus\Omega_k} g_k(\xi)d\xi &\le C\int_{B(0,t)\setminus\Omega_k} |x_k|^2|\nabla^\xi v_k|^{n-2}e^{C_4\lambda_k}d\xi\\ &\le C|x_k|^{2-\frac{n-2}{n-1}}e^{C_4\lambda_k}\int_{B(0,t)\setminus\Omega_k}|x_k|^{\frac{n-2}{n-1}}|\nabla^\xi v_k|^{n-2}d\xi\\ &\le C|x_k|^{2-\frac{n-2}{n-1}}e^{C_4\lambda_k}t^{\frac{n}{n-1}}\left(\int_{B(0,t)\setminus\Omega_k}\left(|x_k|^{\frac{n-2}{n-1}}|\nabla^\xi v_k|^{n-2}\right)^{\frac{n-1}{n-2}}d\xi\right)^{\frac{n-2}{n-1}}\\ &\le C|x_k|^{2-\frac{n-2}{n-1}}e^{C_4\lambda_k}t^{\frac{n}{n-1}}\left(\int_{B^x(0,1)}|\nabla w|^{n-1}dx\right)^{\frac{n-2}{n-1}},\end{aligned}$$

where the last inequality follows because $|\nabla^\xi v_k| = |x_k||\nabla^x w|$ and $d\xi = |x_k|^{-n}dx$. We now calculate separately, for ρ_k to be fixed next,

$$\begin{aligned}\int_0^2 \mu_{g_k}(B(0,t)\setminus\Omega_k)^{\frac{1}{n-1}}\frac{dt}{t} &= \int_0^{\rho_k}\mu_{g_k}(B(0,t)\setminus\Omega_k)^{\frac{1}{n-1}}\frac{dt}{t}\\ &\quad+\int_{\rho_k}^2\mu_{g_k}(B(0,t)\setminus\Omega_k)^{\frac{1}{n-1}}\frac{dt}{t}\\ &\le (n-1)^2C|x_k|^2e^{C_4\lambda_k}\rho_k^{\frac{n}{(n-1)^2}} + C\log\frac{1}{\rho_k}+C.\end{aligned}$$

Let us fix

$$\rho_k = e^{-\frac{(n-1)^2}{n}C_4\lambda_k}\in(0,2).$$

We thus get

$$\int_0^2\mu_{g_k}(B(0,t)\setminus\Omega_k)^{\frac{1}{n-1}}\frac{dt}{t}\le C+C\lambda_k,$$

which contradicts with (2.23) in the light of (2.26). So Lemma 2.6 is proved.

Next we proceed with the blow-down argument as used in the proof Theorem 2.1 in the previous subsection. But this time it is easier because of Lemma 2.6. We consider the blow-down

$$w_r(\xi) = \frac{w(r\xi)}{\log\frac{1}{r}}$$

and calculate that, from Lemma 2.6,

$$|w_r(\xi)|\le C\frac{\log\frac{1}{r}+|\log\frac{1}{|\xi|}|}{\log\frac{1}{r}}\le 2C \text{ for all } \xi\in A_{r,\frac{1}{r}} = \{\xi\in\mathbb{R}^n : |\xi|\in(r,\frac{1}{r})\}.$$

From here, similarly as in the proof of Theorem 2.1 in the previous subsection, one may complete the proof of Theorem 2.2.

3. n-Laplace equations in conformal geometry

In this section we are going to use the local property of n-superharmonic functions to study the asymptotical behavior at the end of a complete locally conformally flat manifold (M^n, g). Based on the injectivity of the development maps of [48, Theorem 4.5], in [54, Theorem 1], and later in [15] with the appendix where the version of positive mass theorem that was used in the proof of [48, Theorem 4.5] was proved under nonnegative Ricci assumption, concluded the following classification result.

Theorem ([54, Theorem 1], [15]). *Let (M^n, g) be a complete conformally flat manifold of dimension $n \geq 3$ with nonnegative Ricci curvature. Then, exactly one of the following holds:*

- *The universal cover of (M^n, g) is globally conformally equivalent to the flat Euclidean space;*
- *(M^n, g) is globally conformally equivalent to the Euclidean space $(\mathbb{R}^n, |dx|^2)$ with a non-flat conformal metric $g = e^{2\phi}|dx|^2$ with nonnegative Ricci curvature for a smooth function on $\mathbb{R}^n$;*
- *The universal cover of (M^n, g) is globally conformally equivalent to a round sphere $(\mathbb{S}^n, g_{\mathbb{S}^n})$;*
- *(M^n, g) is locally isometric to the standard cylinder $R \times \mathbb{S}^{n-1}$.*

Recall that, on $(\mathbb{R}^n, e^{2\phi}|dx|^2)$, in the light of (2.2),

$$-\Delta_n \phi = \mathrm{Ric}_g(\nabla^g \phi)|\nabla \phi|^{n-2} e^{2\phi},$$

where $\mathrm{Ric}_g(\nabla^g \phi)$ is the Ricci curvature of the conformal metric $g = e^{2\phi}|dx|^2$ in the $\nabla^g \phi$ direction. As a consequence of Theorem 2.1 and Theorem 2.2, for a globally conformally flat manifold $(\mathbb{R}^n, e^{2\phi}|dx|^2)$, we therefore are able to deduce the following:

Theorem 3.1. *Suppose that $(\mathbb{R}^n, e^{2\phi}|dx|^2)$ is complete with nonnegative Ricci $(n \geq 3)$. Then there is a subset $E \subset \mathbb{R}^n$, which is measure theoretically n-thin at infinity, such that*

$$\lim_{x \notin E \to \infty} \frac{\phi(x)}{\log \frac{1}{|x|}} = \liminf_{x \to \infty} \frac{\phi(x)}{\log \frac{1}{|x|}} = m \tag{3.1}$$

and

$$\phi(x) \geq m \log \frac{1}{|x|} - C$$

for $|x| \geq R_0$, where C and R_0 are some constants and

$$m^{n-1} vol(\mathbb{S}^{n-1}) = \int_{\mathbb{R}^n} \mathrm{Ric}_g(\nabla^g \phi)|\nabla \phi|^{n-2} e^{2\phi} dx.$$

Moreover,

- *$m \in [0, 1)$ and $m = 0$ if and only if g is flat, i.e., $\phi(x)$ is a constant function.*

- *if* Ric_g *is bounded in addition, then*

$$\lim_{x\to\infty} \frac{\phi(x)}{\log \frac{1}{|x|}} = \liminf_{x\to\infty} \frac{\phi(x)}{\log \frac{1}{|x|}} = m; \tag{3.2}$$

We remark that Theorem 3.1 should be compared with [6, 19]. In [6] it was proved that, a complete noncompact manifold $(M^n,\ g)$ satisfying

$$\mathrm{Ric} \geq 0$$
$$\mathrm{vol}(B(0,r)) \geq \gamma r^n \text{ for some } \gamma > \frac{1}{2}$$
$$|\mathrm{Rm}| \leq Cr^{-2}$$

and in addition,

$$\text{either } |\mathrm{Rm}| = o(r^{-2}) \text{ or } \int_M |\mathrm{Rm}|^{\frac{n}{2}} \,\mathrm{dvol} < \infty,$$

is actually isometric to the Euclidean space. The assumption of $\gamma > \frac{1}{2}$ is essential, in the light of Eguchi–Hanson metrics. In [19], on the other hand, it was proved, a complete noncompact conformally flat manifold with nonnegative Ricci and satisfying

$$\frac{1}{\mathrm{vol}(B(x_0,r))} \int_{B(x_0,r)} R \,\mathrm{dvol} = o(r^{-2})$$

where the scalar curvature R is bounded, is actually isometric to the Euclidean space. The comparison of Theorem 3.1 to the main theorem in [19] would be more direct if one is able to compare the intrinsic distance function r on the conformally flat manifold with the one $|x|$ in Euclidean space as the background metric, which seems require something stronger than (3.2).

4. Hypersurfaces in hyperbolic space

In this section we discuss applications of Theorem 2.1 and Theorem 2.2 to study the asymptotic end structure of embedded hypersurfaces with nonnegative Ricci in hyperbolic space H^{n+1}, $n \geq 3$. The understanding of end structure in [2, 3] are significantly improved.

We recall that the main observation in [11] is that

Lemma ([11, Theorem 3.1]). *For a complete vertical graph in hyperbolic space with nonnegative Ricci, the height function f in Busemann coordinates is n-subharmonic.*

Consequently, as the main theorem in [11], one gets

Theorem ([11, Main Theorem]). *A properly embedded complete hypersurface with nonnegative Ricci has at most two single-point ends. Having two ends is a rigidity condition that forces the hypersurface to be an equidistant hypersurface.*

Therefore one may focus on the study of end structure for such hypersurfaces that are graphs of the height function over a convex domain in $\mathbb{R}^n$ in Busemann coordinates, using the theory of n-subharmonic functions developed in Section 2.

Before doing so, for a vertical graph in Busemann coordinates of hyperbolic space, one considers the inscribed radially symmetric graph (which is called inner rotation hypersurface in [2, 3]). More precisely, let

$$\hat{f}(r) = \sup_{|x|=r} f(x)$$

for a vertical graph $y = f(x)$ in Busemann coordinates, where the hyperbolic metric is given by

$$g_{\mathbb{H}} = dy^2 + e^{-2y}|dx|^2$$

for $(x, y) \in \mathbb{R}^n \times \mathbb{R}$. It is worth to mention that, in such coordinates, any equidistant hypersurface is represented by

$$y = \log|x| + C$$

for some constant C. We observe that, from the work in [2, 3],

Lemma 4.1. *Suppose that the graph $y = f(x)$ over a convex domain Ω in $\mathbb{R}^n$ in Busemann coordinates in hyperbolic space is complete and with nonnegative Ricci. Then, there is an equidistant hypersurface $y = \log|x| + C$ such that*

$$f(x) \leq \hat{f}(|x|) \leq \log|x| + C$$

for all $|x| \geq r_0$ for some $r_0 > 0$. Therefore such graph is always a global graph, i.e., $\Omega = \mathbb{R}^n$.

Thus, based on Theorem 2.1 and its proof in Section 2, we obtain

Theorem 4.1. *Suppose that Σ is a properly embedded, complete hypersurface with nonnegative Ricci and one end. Then it is a global graph of $y = f(x)$ in Busemann coordinates and it is asymptotically rotationally symmetric in the sense that there is a number $m \in [0, 1]$ such that*

$$m \log|x| + o(\log|x|) \leq f(x) \leq m \log|x| + C.$$

Moreover, $m = 0$ implies that the hypersurface is a horosphere. Therefore such hypersurface at end supports equidistant hypersurfaces and is supported by horospheres when $m > 0$.

One may simply apply Theorem 2.1 and realize that deeper geometric argument allows us to reach the conclusion of Theorem 2.1 with no exception thin set E. Theorem 4.1 extends [2, Theorem 1.2] completely.

Acknowledgment

Shiguang Ma is partially supported by NSFC 11571185, 11871283 and Jie Qing by NSF DMS-1608782.

References

[1] D. Adams and N.G. Meyers, *Thinness and Wiener criteria for non-linear potentials*, Indiana Univ. Math. J. **22** (1972), 169–197.

[2] S. Alexander and R.J. Currier, *Non-negatively curved hypersurfaces of hyperbolic space and subharmonic functions*, J. London Math. Soc. **41** (2) (1990), 347–360.

[3] S. Alexander and R.J. Currier, *Hypersurfaces and nonnegative curvature*, Proceedings of Symposia in Pure Mathematics **54** (3) (1993), 37–44.

[4] M. Arsove, *Functions of potential type*, Trans. AMS **75** (1953), 526–551.

[5] M. Arsove and A. Huber, *Local behavior of subharmonic functions*, Indiana Univ. Math. J. **22** (1973), 1191–1199.

[6] S. Bando, A. Kasue, and H. Nakajima, *On a construction of coordinates at infinity on manifolds with fast curvature decay and maximal volume growth*, Invent. Math., **97** (1989), 313–349.

[7] L. Boccardo and T. Gallouët *Nonlinear elliptic equations with right-hand side measures*, Comm. Partial Differential Equations **17** (1992), 641–655.

[8] M.F. Bidaut-Veron. *Local and global behavior of solutions of quasi- linear equations of Emden–Fowler type*, Archive for Rational Mechanics and Analysis, **107** (4) (1989), 293–324.

[9] Ch. Blanc and F. Fiala, *le type d'une surface et sa courbure totale*, Comment. Math. Helv. **14** (1941–42), 230–233.

[10] V. Bonini, S. Ma, and J. Qing, *On non-negatively curved hypersurfaces in hyperbolic space*, to appear in Mathematische Annalen, arXiv:1603.03862.

[11] V. Bonini, S. Ma, and J. Qing, *Hypersurfaces with nonnegative Ricci curvature in hyperbolic space*, preprint 2017, arXiv:1709.00091.

[12] M. Brelot, *Points irreguliers et transformations continues en théorie du potentiel*, J. de Math. **19** (1040), 319–337.

[13] H. Brezis and F. Merle, *Uniform estimates and blow-up behavior for solutions of* $-\Delta u = V(x)e^u$ *in two dimensions*, Comm. in PDE **16** (1991), 1223–1253.

[14] G. Carron and M. Herzlich, *The Huber theorem for non-compact conformally flat manifolds*, Comment. Math. Helv. **77** (2002), 192–220.

[15] G. Carron and M. Herzlich, *Conformally flat manifolds with nonnegative Ricci curvature*, Compos. Math. **142** (2006), no. 3, 798–810.

[16] S.-Y.A. Chang, F.B. Hang, P. Yang, *On a class of locally conformally flat manifolds*, IMRN **4** (2004), 185–209.

[17] S.-Y.A. Chang, M. Gursky, P. Yang, *An equation of Monge–Ampère type in conformal geometry, and four-manifolds of positive Ricci curvature*, Ann. of Math. (2) **155** (2002), no. 3, 709–787.

[18] S.-Y.A. Chang, J. Qing, P. Yang, *Compactification of a class of conformally flat 4-manifold*, Invent. Math. **142** (2000), 65–93.

[19] B. Chen and X. Zhu, *A gap theorem for complete noncompact manifolds with nonnegative Ricci curvature*, Comm. in Geom. and Anal. **10** (2002), no. 1, 217–239.

[20] S. Cohn-Vossen, *Kürzeste Wege und Totalkrümmung auf Flächen*, Compositio Math. **2** (1935), 69–133.

[21] C.B. Croke and H. Karcher, *Volume of small balls on open manifolds:lower bounds and examples*, Trans. Amer. Math. Soc. **309** (2) (1988), 753–762.

[22] G. Dolzmann, N. Hungerbühler and S. Müller, *Non-linear elliptic systems with measure-valued right-hand side*, Math. Z. **226** (1997), 545–574.

[23] C.L. Epstein, *The hyperbolic Gauss map and quasi-conformal reflections*, J. Reine Angew. Math. **372** (1986), 96–135.

[24] C.L. Epstein, *Envelopes of horospheres and Weingarten surfaces in hyperbolic 3-space*, Unpublished (1986). http://www.math.upenn.edu/ cle/papers/index.html.

[25] C.L. Epstein, *The asymptotic boundary of a surface imbedded in* $\mathbb{H}^3$ *with nonnegative curvature*, Michigan Math. J. **34** (1987), 227–239.

[26] F. Fiala, *Le problème des isopérimètres sur les surfaces ouvertes à courbure positive*, Comment. Math. Helv. **13** (1940–41), 293–346.

[27] M. González, *Singular sets of a class of locally conformally flat manifolds*, Duke Math. J. **129** (2005), no. 3, 551–572.

[28] P. Guan, C. Lin, G. Wang, *Schouten tensor and some topological properties*, Comm. Anal. Geom. **13** (2005), no. 5, 887–902.

[29] W.K. Hayman, *Slowly growing integral and subharmonic functions*, Comment. Math. Helv. **34** (1960), 75–84.

[30] W.K. Hayman and P.B. Kennedy, *Subharmonic functions*, vol. 1, Academic Press, London, New York, San Francisco, 1976.

[31] J. Heinonen, T. Kilpeläinen, and O. Martio, *Nonlinear potential theory of degenerate elliptic equations*, Oxford Univ. Press, Oxford, 1993.

[32] M.H. Heins, *Entire functions with bounded minimum modulus; subharmonic function analogous*, Ann. of Math. **49** (1948), 200–213.

[33] H. Hopf, *Sulla geometria riemanniana globale della superficie*, Rendiconti del Seminario Matematico e Fisico di Milano **23** (1953), 48–63.

[34] A. Huber, *Über Wachstumseigenschaften gewisser Klassen von subharmonischen Funktionen*, Comment. Math. Helv. **26** (1952), 81–116.

[35] A. Huber, *On subharmonic functions and differential geometry in the large*, Comment. Math. Helv. **32** (1957), 13–72.

[36] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. Pure. Appl. Math. **14** (1961), 415–426.

[37] T. Kilpeläinen and J. Malý, *Degenerate elliptic equations with measure data and nonlinear potentials*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **19** (1992), no. 4, 591–613.

[38] T. Kilpeläinen and J. Malý, *The wiener test and potential estimates for quasilinear elliptic equations*, Acta. Math. **172** (1994), 137–161.

[39] S. Kichenassamy and L. Veron, *Singular solutions of the p-Laplace equation*, Math. Ann. **275** (1986), 599–615.

[40] S. Kichenassamy and L. Veron, *Erratum: "Singular solutions of the p-Laplace equation"* [Math. Ann. **275** (1986), no. 4, 599–615], Math. Ann. **277** (1987), no. 2, 352.

[41] P. Lindqvist, *Notes on the p-Laplace equation*, University of Jyvaskyla Lecture Notes 2006.

[42] S. Ma and J. Qing, *On n-subharmonic functions and their applications in differential geometry*, in progress 2018.

[43] N. Phuc and I. Verbitsky, *Quasilinear and Hessian equations of Lane–Emden type*, Ann. Math. **168** (2008), 859–914.

[44] Y.G. Reshetniak, *Mappings with bounded deformation as extremals of Dirichlet type integrals*, Sibirsk. Mat. Zh. **9** (1966), 652–666.

[45] J. Rakotoson, *Quasilinear elliptic problems with measures as data*, Differential Integral Equations **4** (1991), no. 3, 449–457.

[46] J. Serrin, *Local behavior of solutions of quasilinear equations*, Acta Math. **111** (1964), 247–302.

[47] J. Serrin, *Isolated singularities of solutions of quasilinear equations*, Acta Math. **113** (1965), 219–240.

[48] R. Schoen, S.T. Yau, *Conformally flat manifolds, Kleinian groups and scalar curvature*, Invent. Math. **92** (1988), 47–71.

[49] S. Taliaferro, *On the growth of superharmonic functions near an isolated singularity I*, J. Differential Equations **158** (1999), 28–47.

[50] S. Taliaferro, *On the growth of superharmonic functions near an isolated singularity I*, Indiana Univ. Math. J. **50** (2001), no. 4, 1885–1897.

[51] S. Taliaferro, *Isolated Singularities of Nonlinear Elliptic Inequalities. II. Asymptotic Behavior of Solutions*, Indiana Univ. Math. J. **55** (2006), no. 6, 1791–1811.

[52] P. Tolksdorf, *On the Dirichlet problem for quasilinear equations in domain with conical boundary points*, Commun. Partial. Differ. Equations, bf 8 (1983), 773–817.

[53] L. Veron, *Local and Global Aspects of Quasilinear Degenerate Elliptic Equations*, New Jersey, World Scientific, 2017.

[54] S. Zhu, *The classification of complete locally conformally flat manifolds of nonnegative Ricci curvature, Pac. J. Math.* **163** (1) (1994), 189–199.

Shiguang Ma
Department of Mathematics
Nankai University
Tianjin, China
e-mail: `msgdyx8741@nankai.edu.cn`

Jie Qing
Department of Mathematics
University of California
Santa Cruz, CA 95064, USA
e-mail: `qing@ucsc.edu`

Progress in Mathematics, Vol. 333, 265–286

From Local Index Theory to Bergman Kernel: A Heat Kernel Approach

Xiaonan Ma

Dedicated to Gang Tian for his 60th birthday.

Abstract. The aim of this note is to explain a uniform approach of three different topics: Atiyah–Singer index theorem, holomorphic Morse inequalities and asymptotic expansion of Bergman kernel, by using heat kernels.

Mathematics Subject Classification (2010). Primary 58J20; Secondary 32L10.

Keywords. Dirac operator, index theorem, holomorphic Morse inequalities, Bergman kernel.

0. Introduction

In this note, we explain how to give a uniform approach of three different topics: Atiyah–Singer index theorem, holomorphic Morse inequalities and asymptotic expansion of Bergman kernel, by using heat kernels.

Roughly the Atiyah–Singer index theorem announced in 1963, as one of the most important theorems in mathematics of 20 century, computes the index of an elliptic operator by using the characteristic classes, i.e., the topological way. Its heat kernel approach (as a solution of the McKean–Singer conjecture or as the local index theorem) was developed by Gilkey in his thesis in 1973 and also by Atiyah–Bott–Patodi, which needs to use Weyl's invariant theory and computes infinite examples to detect the final formula. In the 1980s, influenced by the supersymmetry in physics, Bismut and Getzler independently developed direct heat kernel proofs of the Atiyah–Singer index theorem. In modern index theory, the local index techniques plays a central role which allows us to study the more refined spectral invariants such as the analytic torsion and the eta invariant.

In complex geometry, the Atiyah–Singer index theorem reduces to the classical Riemann–Roch–Hirzebruch theorem, which computes the alternating sum of dimensions of the Dolbeault cohomology groups of a holomorphic vector bundle.

The holomorphic Morse inequalities give an asymptotic estimate of the dimension of each Dolbeault cohomology group of a pth tensor power of a line bundle when p goes to infinity. This was first established by Demailly in 1985 [7] answering a question of Siu after Siu's solution of Grauert–Riemenschneider conjecture, and Bismut [2] gave a heat kernel approach. If the line bundle is positive, then by the Kodaira vanishing theorem, for large p the associated Dolbeault cohomology group of positive degree is zero and the dimension of its zero degree part, i.e., the space of holomorphic sections of its pth tensor power, is given by the Riemann–Roch–Hirzebruch theorem. Its analytic refinement is the smooth kernel of the orthogonal projection from the space of smooth sections onto the space of holomorphic sections: the Bergman kernel. In his thesis [16] in 1990, Tian initiated the study of the asymptotic of Bergman kernels. Since then, it is a very active research direction.

In this note, we explain a uniform approach of the above three topics by using heat kernels, which is inspired a lot from the analytic localization techniques of Bismut–Lebeau in local index theory. The basic references of this note are [1, Chap. 4] on the local index theorem, and [6], [11, §1.6, §4.1], [10], where the readers can also find a complete list of references. In particular, based on our contributions with Dai, Liu and Marinescu, [11] gives a comprehensive study on holomorphic Morse inequalities and Bergman kernels and their applications. To keep this note in a reasonable size, we omit many technical details, and hope that this note can be served as an introduction of the subject and motivation to read the book [11] and recent developments.

This note is organized as follows: In Section 1, we explain the Atiyah–Singer index theorem and the basic ideas on its local version: the local index theorem. In Sections 2, 3, we show how to apply the ideas from the local index theory to give a heat kernel approach of the holomorphic Morse inequalities and Berman kernels.

This note is based on the three lectures I gave in January 2018 in the workshop 'International workshop on differential geometry' at Sydney in celebration of Professor Gang Tian's 60th birthday.

Notations: we denote by dim or $\dim_{\mathbb{C}}$ the complex dimension of a complex vector space. Denote also $\dim_{\mathbb{R}}$ the real dimension of a space. $\mathrm{supp}(f)$ means the support of a function f.

1. Local index theorem

In this section, we review briefly the Chern–Weil theory, the Atiyah–Singer index theorem for Dirac operators and the heat kernel proof of the local index theorem.

1.1. Chern–Weil Theory

Let X be a smooth manifold of dimension n. Let TX be its tangent bundle and T^*X its cotangent bundle. Let $\Omega^k(X) = \mathscr{C}^\infty(X, \Lambda^k(T^*X))$ be the space of smooth k-forms on X and $\Omega^\bullet(X) = \oplus_k \Omega^k(X)$, and $d : \Omega^k(X) \to \Omega^{k+1}(X)$ be the exterior differential operator.

Definition 1.1. Let E be a smooth manifold, and let $\pi : E \to X$ be a smooth map. Then E is called a complex vector bundle over X if there exist a covering $\{U_i\}_{i=1}^l$ of

X and a family of diffeomorphisms $\phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{C}^m, \phi_i(v) = (\pi(v), \psi_i(v))$ such that if $U_i \cap U_j \neq \emptyset$, then for $x \in U_i \cap U_j$, $\psi_{ji}(x,\cdot) \in \mathrm{GL}(m,\mathbb{C})$, i.e., an invertible $\mathbb{C}$-linear map on $\mathbb{C}^m$, and is smooth on x, where $\psi_{ji}(x,\cdot)$ is given by

$$\begin{aligned} \phi_j \circ \phi_i^{-1} &: (U_i \cap U_j) \times \mathbb{C}^m \to (U_i \cap U_j) \times \mathbb{C}^m, \\ &\phi_j \circ \phi_i^{-1}(x,w) = (x, \psi_{ji}(x,w)). \end{aligned} \tag{1.1}$$

That is, if we write $\psi_{ji}(x,w) = \psi_{ji}(x)w$, then $\psi_{ji}(x) \in \mathrm{GL}(m,\mathbb{C})$. For $x \in X$, $E_x := \pi^{-1}(x)$ is called the fiber of E at x. The integer m is called the rank of E and is denoted by $\mathrm{rk}(E)$. If $\mathrm{rk}(E) = 1$, then E is called a line bundle.

Denote by $\mathscr{C}^\infty(X,E)$ the space of smooth sections of E on X, i.e., the space of smooth maps from X to E such that its composition with π is the identity map on X. Denote by $\Omega^\bullet(X,E) := \mathscr{C}^\infty(X, \Lambda(T^*X) \otimes E)$ the space of smooth forms on X with values in E. We denote by $\mathscr{C}^\infty(X,\mathbb{C})$ the space of smooth $\mathbb{C}$-valued functions on X.

Definition 1.2. A map $\nabla^E : \mathscr{C}^\infty(X,E) \to \mathscr{C}^\infty(X, T^*X \otimes E)$ is called a connection if

1) ∇^E is $\mathbb{C}$-linear,
2) For any $s \in \mathscr{C}^\infty(X,E)$ and $\varphi \in \mathscr{C}^\infty(X,\mathbb{C})$,

$$\nabla^E(\varphi s) = d\varphi \otimes s + \varphi \nabla^E s. \tag{1.2}$$

A Hermitian metric h^E on E is a family of Hermitian products h^{E_x} on E_x which is smooth on $x \in X$. In this case, we call (E, h^E) a Hermitian vector bundle and as usual, we also denote h^E by $\langle \ \ \rangle$. A connection ∇^E is a Hermitian connection on (E,h^E) if for any $s_1, s_2 \in \mathscr{C}^\infty(X,E)$,

$$\left\langle \nabla^E s_1, s_2 \right\rangle + \left\langle s_1, \nabla^E s_2 \right\rangle = d \left\langle s_1, s_2 \right\rangle. \tag{1.3}$$

Let $\nabla^E : \mathscr{C}^\infty(X,E) \to \mathscr{C}^\infty(X, T^*X \otimes E)$ be a connection on E.

Definition 1.3. Let $\nabla^E : \Omega^k(X,E) \to \Omega^{k+1}(X,E)$ be the operator induced by ∇^E such that for any $\alpha \in \Omega^k(X)$ and $s \in \mathscr{C}^\infty(X,E)$,

$$\nabla^E(\alpha \wedge s) = d\alpha \wedge s + (-1)^k \alpha \wedge \nabla^E s. \tag{1.4}$$

The operator $(\nabla^E)^2$ defines a homomorphism $R^E := (\nabla^E)^2 : E \to \Lambda^2(T^*X) \otimes E$. The $R^E \in \Omega^2(X, \mathrm{End}(E))$ is called the curvature operator of ∇^E.

Example. For $E = \mathbb{C}$, the exterior differential $d : \Omega^k(X,\mathbb{C}) \to \Omega^{k+1}(X,\mathbb{C})$ is a connection on the trivial line bundle $\mathbb{C}$ and $d^2 = 0$. The de Rham cohomology of X is defined by

$$H^k(X,\mathbb{C}) := \frac{\mathrm{Ker}\left(d|_{\Omega^k(X,\mathbb{C})}\right)}{\mathrm{Im}\left(d|_{\Omega^{k-1}(X,\mathbb{C})}\right)}, \quad H^\bullet(X,\mathbb{C}) = \bigoplus_{k=0}^n H^k(X,\mathbb{C}). \tag{1.5}$$

Theorem 1.4 (Chern–Weil). *For $f \in \mathbb{R}[z]$, i.e., f is a real polynomial on z, set*

$$F(R^E) = \mathrm{Tr}\left[f\left(\frac{i}{2\pi} R^E \right) \right] \in \Omega^{2\bullet}(X,\mathbb{C}). \tag{1.6}$$

Then $F(R^E)$ is closed. Moreover, its cohomology class $[F(R^E)] \in H^{2\bullet}(X,\mathbb{R})$ and it does not depend on the choice of the connection ∇^E.

Proof. By the definition of R^E, we have the Bianchi identity:

$$\left[\nabla^E, R^E\right] = \left[\nabla^E, (\nabla^E)^2\right] = 0. \tag{1.7}$$

Then

$$dF(R^E) = d\,\mathrm{Tr}\left[f\left(\frac{i}{2\pi}R^E\right)\right] = \mathrm{Tr}\left[\left[\nabla^E, f\left(\frac{i}{2\pi}R^E\right)\right]\right] = 0. \tag{1.8}$$

That is, $F(R^E)$ is closed.

Denote by $\pi : X \times \mathbb{R} \to X$ the natural projection. Let ∇_0^E, ∇_1^E be two connections on E. Then

$$\nabla^{\pi^*E} = (1-t)\nabla_0^E + t\nabla_1^E + dt \wedge \frac{\partial}{\partial t} \tag{1.9}$$

is a connection on π^*E, the pullback of E over $X \times \mathbb{R}$. Set $\nabla_t^E = (1-t)\nabla_0^E + t\nabla_1^E$. Its curvature $R_t^E = (\nabla_t^E)^2 \in \Omega^2\big(X, \mathrm{End}(E)\big)$ and $R^{\pi^*E} = (\nabla^{\pi^*E})^2 = R_t^E + dt \wedge \cdot$, thus there exists $Q_t \in \Omega^\bullet(X)$ such that

$$F(R^{\pi^*E}) = F(R_t^E) + dt \wedge Q_t. \tag{1.10}$$

Applying (1.8) for π^*E, we get $d^{X\times\mathbb{R}}F(R^{\pi^*E}) = 0$. By (1.10) and comparing the coefficient of dt in $d^{X\times\mathbb{R}}F(R^{\pi^*E}) = 0$, we obtain

$$\frac{\partial}{\partial t}F(R_t^E) = dQ_t. \tag{1.11}$$

Thus

$$F(R_1^E) - F(R_0^E) = d\int_0^1 Q_t dt, \tag{1.12}$$

which implies

$$\left[F(R_1^E)\right] = \left[F(R_0^E)\right] \in H^{2\bullet}(X,\mathbb{C}). \tag{1.13}$$

Finally, we can choose a Hermitian metric h^E on E and a Hermitian connection ∇^E on (E,h^E), then $\frac{i}{2\pi}R^E$ is self-adjoint with respect to h^E, thus $\mathrm{Tr}\Big[f\big(\frac{i}{2\pi}R^E\big)\Big]$ is a real differential form, which implies that $[F(R^E)] \in H^{2\bullet}(X,\mathbb{R})$. The proof of Theorem 1.4 is completed. □

Example. 1). For $f(z) = e^z$, the Chern character form of (E,∇^E) is

$$\mathrm{ch}(E,\nabla^E) = \mathrm{Tr}\left[\exp\left(\frac{i}{2\pi}R^E\right)\right]. \tag{1.14}$$

The Chern character of E is

$$\mathrm{ch}(E) := \left[\,\mathrm{ch}(E,\nabla^E)\right] \in H^{2\bullet}(X,\mathbb{R}). \tag{1.15}$$

The first Chern form of (E,∇^E) is $c_1(E,\nabla^E) = \mathrm{Tr}\left[\frac{i}{2\pi}R^E\right]$. Its cohomology class is the first Chern class $c_1(E)$.

2). For $f(z) = \log(\frac{z}{1-e^{-z}})$, the Todd form of (E, ∇^E) is

$$\mathrm{Td}(E, \nabla^E) = \det\left[\frac{\frac{i}{2\pi}R^E}{1 - e^{-\frac{i}{2\pi}R^E}}\right] = \exp\left\{\mathrm{Tr}\left[\log\left(\frac{\frac{i}{2\pi}R^E}{1 - e^{-\frac{i}{2\pi}R^E}}\right)\right]\right\}. \tag{1.16}$$

The Todd class of E is

$$\mathrm{Td}(E) = \left[\mathrm{Td}(E, \nabla^E)\right] \in H^{2\bullet}(X, \mathbb{R}). \tag{1.17}$$

3). Let g^{TX} be a Riemannian metric on TX and ∇^{TX} be the Levi-Civita connection on (X, g^{TX}). The $\widehat{A}$-form of (TX, ∇^{TX}) is

$$\widehat{A}(TX, \nabla^{TX}) = \det{}^{1/2}\left[\frac{\frac{i}{4\pi}R^{TX}}{\sinh\left(\frac{i}{4\pi}R^{TX}\right)}\right]. \tag{1.18}$$

The $\widehat{A}$-genus of TX is

$$\widehat{A}(TX) = \left[\widehat{A}(TX, \nabla^{TX})\right] \in H^{4\bullet}(X, \mathbb{R}). \tag{1.19}$$

1.2. Atiyah–Singer index theorem

Let X be an n-dimensional compact spin manifold with n even (in particular, X is orientable) and g^{TX} be a Riemannian metric on X. Let $S(TX)$ be the spinor bundle of (TX, g^{TX}). Then $S(TX)$ is a $\mathbb{Z}_2$-graded vector bundle on X:

$$S(TX) = S^+(TX) \oplus S^-(TX). \tag{1.20}$$

For $U \in TX$, let $c(U) \in \mathrm{End}(S(TX))$ be the Clifford action of U on $S(TX)$. We will not explain in detail the Clifford action, but only recall that $c(U)$ exchange $S^+(TX)$ and $S^-(TX)$ and $c(U)^2 = -|U|^2_{g^{TX}}$.

The Levi-Civita connection ∇^{TX} on (X, g^{TX}) induces canonically the Clifford connection $\nabla^{S(TX)} = \nabla^{S^+(TX)} \oplus \nabla^{S^-(TX)}$ on $S(TX)$, i.e., the connection preserves the splitting (1.20) and compatible with the Clifford action:

$$\left[\nabla^{S(TX)}_V, c(U)\right] = c\left(\nabla^{TX}_V U\right) \quad \text{for } U, V \in \mathscr{C}^\infty(X, TX). \tag{1.21}$$

Let (E, h^E) be a Hermitian vector bundle on X. Let ∇^E be a Hermitian connection on (E, h^E). Denote by $\nabla^{S(TX)\otimes E}$ the connection on $S(TX)\otimes E$ induced by $\nabla^{S(TX)}$ and ∇^E.

Definition 1.5. The Dirac operator is defined by

$$D = \sum_{j=1}^{n} c(e_j)\nabla^{S(TX)\otimes E}_{e_j} : \mathscr{C}^\infty(X, S^\pm(TX) \otimes E) \to \mathscr{C}^\infty(X, S^\mp(TX) \otimes E), \tag{1.22}$$

where $\{e_j\}$ is an orthonormal frame of (TX, g^{TX}).

Let dv_X be the Riemannian volume form on (X, g^{TX}). For

$$s_1, s_2 \in \mathscr{C}^\infty(X, S(TX) \otimes E),$$

their Hermitian product is defined as

$$\langle s_1, s_2 \rangle = \int_X \langle s_1, s_2 \rangle (x) dv_X(x).$$

The Dirac operator D is a first-order self-adjoint elliptic differential operator. As X is compact, D is a Fredholm operator, in particular, its kernel $\mathrm{Ker}(D)$ is a finite-dimensional complex vector space.

Set

$$D_\pm = D|_{\mathscr{C}^\infty(X, S^\pm(TX)\otimes E)}.$$

Then under the decomposition (1.20),

$$D = \begin{pmatrix} 0 & D_- \\ D_+ & 0 \end{pmatrix}, \tag{1.23}$$

and D^2 preserves $\mathscr{C}^\infty(X, S^\pm(TX) \otimes E)$. As D is self-adjoint, $\mathrm{CoKer}(D_+)$, the cokernel of D_+, is $\mathrm{Ker}(D_-)$. Thus the index of D_+ is defined by

$$\mathrm{Ind}(D_+) = \dim \mathrm{Ker}(D_+) - \dim \mathrm{Ker}(D_-) \in \mathbb{Z}. \tag{1.24}$$

Theorem 1.6 (Atiyah–Singer index theorem (1963)).

$$\mathrm{Ind}(D_+) = \int_X \widehat{A}(TX)\mathrm{ch}(E). \tag{1.25}$$

1.3. Heat kernel and McKean–Singer formula

The heat kernel $e^{-tD^2}(x, y)$ is the smooth kernel of the heat operator e^{-tD^2} with respect to the Riemannian volume form $dv_X(y)$. The following result is well known.

Theorem 1.7. *For any $t > 0$ and $x, y \in X$,*

$$e^{-tD^2}(x, y) = \sum_{j=1}^\infty e^{-t\lambda_j} \psi_j(x) \otimes \psi_j(y)^*, \tag{1.26}$$

where ψ_j is a unit eigenfunction of D^2 corresponding to the eigenvalue λ_j with $0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \cdots$, $\lambda_j \to +\infty$ such that $\{\psi_j\}_j$ form a complete orthonormal basis of the space of L^2-integrable sections, $L^2(X, S(TX)\otimes E)$, and $\psi_j(y)^* \in (S(TX)\otimes E)^*_y$ is the metric dual of $\psi_j(y)$, i.e.,

$$\psi_j(y)^*(v) = \big\langle v, \psi_j(y) \big\rangle \quad \text{for } v \in (S(TX) \otimes E)_y. \tag{1.27}$$

Theorem 1.8 (McKean–Singer (1967)). *For any $t > 0$, we have*

$$\mathrm{Ind}(D_+) = \mathrm{Tr}_s\left[e^{-tD^2}\right] = \int_X \mathrm{Tr}_s\left[e^{-tD^2}(x, x)\right] dv_X(x), \tag{1.28}$$

where the supertrace Tr_s is given by

$$\mathrm{Tr}_s = \mathrm{Tr}\left|_{\mathscr{C}^\infty(X, S^+(TX)\otimes E)} - \mathrm{Tr}\right|_{\mathscr{C}^\infty(X, S^-(TX)\otimes E)}. \tag{1.29}$$

Proof. By (1.26),

$$\lim_{t\to+\infty} e^{-tD^2}(x,x) = \sum_{\lambda_j=0} \psi_j(x)\otimes\psi_j(x)^*,$$

which implies

$$\lim_{t\to+\infty} \mathrm{Tr}_s\left[e^{-tD^2}\right] = \mathrm{Ind}(D_+). \tag{1.30}$$

Then it suffices to prove that $\mathrm{Tr}_s\left[e^{-tD^2}\right]$ is independent of $t>0$. In fact,

$$\begin{aligned}\frac{\partial}{\partial t}\mathrm{Tr}_s\left[e^{-tD^2}\right] &= -\,\mathrm{Tr}_s\left[D^2e^{-tD^2}\right]\\ &= -\,\frac{1}{2}\mathrm{Tr}_s\left[\left[De^{-tD^2/2}, De^{-tD^2/2}\right]\right] = 0.\end{aligned} \tag{1.31}$$

Here for a $\mathbb{Z}_2$-graded vector space $\mathcal{E}=\mathcal{E}^+\oplus\mathcal{E}^-$, the $\mathbb{Z}_2$-grading on $\mathrm{End}(\mathcal{E})$ is given by

$$\begin{aligned}\mathrm{End}(\mathcal{E})^+ &= \mathrm{Hom}(\mathcal{E}^+,\mathcal{E}^+)\oplus\mathrm{Hom}(\mathcal{E}^-,\mathcal{E}^-),\\ \mathrm{End}(\mathcal{E})^- &= \mathrm{Hom}(\mathcal{E}^+,\mathcal{E}^-)\oplus\mathrm{Hom}(\mathcal{E}^-,\mathcal{E}^+),\end{aligned} \tag{1.32}$$

and $[\cdot,\cdot]$ is the supercommutator of $\mathrm{End}(\mathcal{E})$, i.e.,

$$[A,B] = \begin{cases} AB-BA & \text{if } A \text{ or } B\in\mathrm{End}(\mathcal{E})^+,\\ AB+BA & \text{if } A,B\in\mathrm{End}(\mathcal{E})^-.\end{cases} \tag{1.33}$$

Then we verify easily as for matrices that $\mathrm{Tr}_s\big[[A,B]\big]=0$. This completes the proof of Theorem 1.8. □

When $t\to 0$, classically the following asymptotic expansion of the heat kernel holds for any $k\in\mathbb{N}$:

$$e^{-tD^2}(x,x) = \sum_{j=-l}^{k} a_j(x)t^j + \mathscr{O}(t^{k+1}) \ \text{ uniformly on } X, \tag{1.34}$$

where $l=n/2$ and the coefficients $a_j(x)$ depend only on the restriction of D^2 on $B^X(x,\varepsilon)$, the ball in X of center x and radius ε for any $\varepsilon>0$. Then the McKean–Singer formula implies that

$$\int_X \mathrm{Tr}_s\left[a_j(x)\right]dv_X(x) = \begin{cases}0 & \text{for } j<0,\\ \mathrm{Ind}(D_+) & \text{for } j=0.\end{cases} \tag{1.35}$$

McKean–Singer conjectured that in fact a pointwise version of (1.35) holds, which they called the “miraculous cancellation”. The solution of this conjecture is called the local index theorem stated as follows. For $\alpha\in\Omega(X)$, we denote $\alpha^{\max}$ the degree n component of the differential form α.

Theorem 1.9 (Local index theorem).

$$\operatorname{Tr}_s\left[a_j(x)\right]dv_X(x)=\begin{cases}0 & \textit{for } j<0,\\ \left\{\widehat{A}(TX,\nabla^{TX})\operatorname{ch}(E,\nabla^E)\right\}_x^{\max} & \textit{for } j=0.\end{cases}\tag{1.36}$$

Equivalently,

$$\lim_{t\to 0}\operatorname{Tr}_s\left[e^{-tD^2}(x,x)\right]dv_X(x)=\left\{\widehat{A}(TX,\nabla^{TX})\operatorname{ch}(E,\nabla^E)\right\}_x^{\max}.\tag{1.37}$$

By Theorems 1.8 and 1.9, we get the Atiyah–Singer index theorem, Theorem 1.6.

Remark 1.10. Using the Bott periodicity theorem in K-theory, we obtain the index theorem for any elliptic operator P on X:

$$\operatorname{Ind}(P)=\int_{T^*X}\widehat{A}(TX)^2\operatorname{ch}\big(\sigma(P)\big).\tag{1.38}$$

Here $\sigma(P)$ is the principal symbol of P, which can be understood as an element in $K(T^*X)$, the K-group of T^*X.

1.4. Proof of the local index theorem

The proof presented here consists of Bismut–Lebeau's analytic localization techniques [3, §11] and Getzler rescaling trick. We need to compute the limit as $t\to 0$,

$$\lim_{t\to 0}\operatorname{Tr}_s\left[e^{-tD^2}(x,x)\right]dv_X(x).$$

Step 1. The asymptotic of $e^{-tD^2}(x,x)$ is local, i.e., only depends on the restriction of D^2 on any neighborhood of x.

Recall

$$e^{-a^2/2}=\int_{-\infty}^{+\infty}\cos(va)e^{-v^2/2}\frac{dv}{\sqrt{2\pi}},\quad \text{for any } a\in\mathbb{R}.\tag{1.39}$$

Thus for the heat operator e^{-tD^2}, we have

$$e^{-tD^2}=\int_{-\infty}^{+\infty}\cos\left(v\sqrt{2t}D\right)e^{-v^2/2}\frac{dv}{\sqrt{2\pi}}.\tag{1.40}$$

We can formally verify (1.40) from (1.39) as D^2 is an infinite-dimensional diagonal matrix. Rigorously, the wave operator $\cos\big(v\sqrt{2t}D\big)$ is given by

$$\cos\big(vD\big)(x,y)=\sum\nolimits_j\cos(v\sqrt{\lambda_j})\psi_j(x)\otimes\psi_j(y)^*.\tag{1.41}$$

In fact, $w_t(x)=\cos\big(tD\big)(x,\cdot)$ is the fundamental solution of the equation

$$\left(\frac{\partial^2}{\partial t^2}+D^2\right)w_t(x)=0\tag{1.42}$$

with the initial conditions

$$\lim_{t\to 0}w_t\phi=\phi,\quad\text{for any } \phi\in L^2(X,S(TX)\otimes E).\tag{1.43}$$

Using the energy estimates, we obtain the property of the finite propagation speed for the wave operator $\cos\big(tD\big)$:

$$\operatorname{supp}\cos\big(tD\big)(x,\cdot)\subset B^X(x,t) \tag{1.44}$$

and $\cos\big(tD\big)(x,\cdot)$ depends only on $D^2|_{B^X(x,t)}$.

Let $f:\mathbb{R}\to[0,1]$ be a smooth even function such that $f(v)=1$ for $|v|\leqslant\varepsilon/2$ and that $f(v)=0$ for $|v|\geqslant\varepsilon$. For $u>0$, set $F_u(a),G_u(a)$ even functions on $\mathbb{R}$ defined by

$$\begin{aligned} F_u(a)&=\int_{-\infty}^{+\infty}\cos(va)e^{-v^2/2}f(\sqrt{u}v)dv/\sqrt{2\pi},\\ G_u(a)&=\int_{-\infty}^{+\infty}\cos(va)e^{-v^2/2}\Big(1-f(\sqrt{u}v)\Big)dv/\sqrt{2\pi}. \end{aligned} \tag{1.45}$$

Clearly, from (1.39) and (1.45),

$$F_u(a)+G_u(a)=e^{-a^2/2}. \tag{1.46}$$

From (1.44),

$$\operatorname{supp}F_u\big(\sqrt{u}D\big)(x,\cdot)\subset B^X(x,\varepsilon). \tag{1.47}$$

Clearly, from (1.45),

$$G_u\big(\sqrt{u}a\big)=\int_{|v|\geqslant\varepsilon/2}e^{iva}\exp\left(-\frac{v^2}{2u}\right)\big(1-f(v)\big)\frac{dv}{\sqrt{2\pi u}}. \tag{1.48}$$

Then

$$\begin{aligned} a^mG_u\big(\sqrt{u}a\big)&=i^m\int_{|v|\geqslant\varepsilon/2}e^{iva}\frac{\partial^m}{\partial v^m}\left[\exp\left(-\frac{v^2}{2u}\right)\big(1-f(v)\big)\right]\frac{dv}{\sqrt{2\pi u}}\\ &=\int_{|v|\geqslant\varepsilon/2}e^{iva}\exp\left(-\frac{v^2}{2u}\right)\sum_{j=0}^m\frac{\partial^j}{\partial v^j}\big(1-f(v)\big)P_j\left(\frac{1}{u},v\right)\frac{dv}{\sqrt{2\pi u}}, \end{aligned} \tag{1.49}$$

where $P_j(\frac{1}{u},v)$ are polynomials on $\frac{1}{u}$ and v. Thus, there exists $C>0$ such that for $u\in(0,1]$,

$$\Big|a^mG_u\big(\sqrt{u}a\big)\Big|\leqslant C\,e^{-\frac{\varepsilon^2}{16u}},\ \text{for any } a\in\mathbb{R}. \tag{1.50}$$

Again from (1.50) in view of D^2 as a diagonal matrix, we get the estimate of operator norm $\|\cdot\|^{0,0}$ from L^2 to L^2 as

$$\Big\|D^mG_u\big(\sqrt{u}D\big)\Big\|^{0,0}\leqslant C\,e^{-\frac{\varepsilon^2}{16u}}. \tag{1.51}$$

Using Sobolev inequalities and (1.51), we obtain uniformly on $x,y\in X$,

$$\Big|G_u\big(\sqrt{u}D\big)(x,y)\Big|\leqslant c_1e^{-c_2/u}. \tag{1.52}$$

This concludes that the asymptotic expansion of $e^{-tD^2}(x,x)$ is local! Note that the heat operator e^{-tD^2} is defined globally by means of eigenvalues and eigenfunctions of D^2.

Step 2. Replace X by $\mathbb{R}^n$, we work on $\mathbb{R}^n$. Fix $x_0 \in X$. We identify $B^{T_{x_0}X}(0,4\varepsilon)$ to $B^X(x_0,4\varepsilon)$ by the exponential map: $v \to \exp_{x_0}(v)$. For $Z \in B^{T_{x_0}X}(0,4\varepsilon) \subset T_{x_0}X$, we identify $S(TX)_Z$, E_Z to $S(TX)_{x_0}$, E_{x_0} by parallel transport with respect to $\nabla^{S(TX)}$, ∇^E along the path $\gamma : [0,1] \to X, \gamma(s) = sZ$. Then we extend $D^2|_{B^{T_{x_0}X}(0,2\varepsilon)}$ to an operator on $\mathbb{R}^n$ which is the canonical (positive) Laplacian outside $B^{T_{x_0}X}(0,4\varepsilon)$.

Step 3. Rescaling. Set $\mathbf{E}_{x_0} = (S(TX)\otimes E)_{x_0}$. For $s \in \mathscr{C}_0^\infty(\mathbb{R}^n, \mathbf{E}_{x_0})$, $Z \in \mathbb{R}^n$, set

$$\big(S_t s\big)(Z) = s\left(\frac{Z}{\sqrt{t}}\right), \quad L_2^t = S_t^{-1} t D^2 S_t. \tag{1.53}$$

Let $\{e_j\}_{j=1}^n$ be an oriented orthonormal basis of $T_{x_0}X$. For $1 \leqslant j \leqslant n$, $t \in (0,1]$, set

$$c_t(e_j) = \frac{1}{\sqrt{t}} e^j \wedge -\sqrt{t}\, i_{e_j} \in \mathrm{End}\big(\Lambda(T^*_{x_0}X)\big). \tag{1.54}$$

Let L_3^t be the operator obtained from L_2^t by replacing $c(e_j)$ by $c_t(e_j)$ in the explicit formula of the operator L_2^t. Then L_3^t acts on $\mathscr{C}^\infty(\mathbb{R}^n, (\Lambda(T^*X)\otimes E)_{x_0})$. We claim that as $t \to 0$,

$$\mathrm{Tr}_s\left[e^{-tD^2}(x_0,x_0)\right] = (-2i)^{n/2}\, \mathrm{Tr}\Big|_E \left[e^{-L_3^t}(0,0)\right]^{\max} + O(e^{-c/t}), \tag{1.55}$$

which follows from the simple linear algebra identity: for $1 \leqslant i_1 < \cdots < i_j \leqslant n$,

$$\mathrm{Tr}_s\Big|_{S(TX)}\big[c(e_{i_1})\cdots c(e_{i_j})\big] = \begin{cases} 0 & \text{if } j < n = 2l, \\ (-2i)^{n/2} & \text{if } j = n, \end{cases} \tag{1.56}$$

and thus

$$\mathrm{Tr}_s\left[e^{-L_2^t}(0,0)\right] = (-2i)^{n/2} t^{n/2}\, \mathrm{Tr}\Big|_E \left[e^{-L_3^t}(0,0)\right]^{\max}. \tag{1.57}$$

Theorem 1.11. *As $t \to 0$,*

$$L_3^t \to L_3^0 = -\sum_{j=1}^n \left[\frac{\partial}{\partial Z_j} + \frac{1}{4}\left\langle R_{x_0}^{TX} Z, \frac{\partial}{\partial Z_j}\right\rangle\right]^2 + R_{x_0}^E. \tag{1.58}$$

The following Lichnerowicz formula allows us to obtain (1.58):

$$D^2 = \Delta + \frac{1}{4} r^X + {}^c R^E, \tag{1.59}$$

where Δ is the (positive) Bochner Laplacian on $S(TX)\otimes E$ associated with the connection $\nabla^{S(TX)\otimes E}$, and r^X is the scale curvature of (X, g^{TX}) and for $\{e_j\}_{j=1}^n$

an orthonormal frame of (X, g^{TX}),

$$^cR^E = \frac{1}{2}\sum_{i,j=1}^{n} R^E(e_i,e_j)c(e_i)c(e_j). \tag{1.60}$$

By using weighted Sobolev norm adapted from the structure of the operator L_3^t, we can obtain as $t\to 0$,

$$e^{-L_3^t}(0,0) \to e^{-L_3^0}(0,0). \tag{1.61}$$

By Mehler's formula, we get

$$\begin{aligned} e^{-tL_3^0}(Z,Z') = (4\pi)^{-n/2}\exp(-tR^E_{x_0})\det{}^{1/2}\left[\frac{R^{TX}_{x_0}}{e^{tR^{TX}_{x_0}/2}-e^{-tR^{TX}_{x_0}/2}}\right] \\ \times\exp\left\{\left\langle -\frac{R^{TX}_{x_0}/4}{2\tanh(tR^{TX}_{x_0}/2)}Z,Z\right\rangle - \left\langle \frac{R^{TX}_{x_0}/4}{2\tanh(tR^{TX}_{x_0}/2)}Z',Z'\right\rangle \right. \\ \left. + \left\langle \frac{e^{tR^{TX}_{x_0}/4}R^{TX}_{x_0}/4}{2\sinh(tR^{TX}_{x_0}/2)}Z,Z'\right\rangle\right\}. \end{aligned} \tag{1.62}$$

In particular,

$$e^{-L_3^0}(0,0) = (4\pi)^{-n/2}\det{}^{1/2}\left[\frac{R^{TX}_{x_0}}{e^{R^{TX}_{x_0}/2}-e^{-R^{TX}_{x_0}/2}}\right]\exp(-R^E_{x_0}). \tag{1.63}$$

Combining (1.55), (1.61) and (1.63), we obtain

$$\begin{aligned} &\lim_{t\to 0}\operatorname{Tr}_s\left[e^{-tD^2}(x_0,x_0)\right]dv_X(x_0) \\ &\quad = (-2i)^{n/2}\operatorname{Tr}\Big|_E\left[e^{-L_3^0}(0,0)\right]^{\max} \\ &\quad = (-2i)^{n/2}(4\pi)^{-n/2}\left\{\det{}^{1/2}\left[\frac{R^{TX}_{x_0}}{e^{R^{TX}_{x_0}/2}-e^{-R^{TX}_{x_0}/2}}\right]\operatorname{Tr}\left[e^{-R^E_{x_0}}\right]\right\}^{\max} \\ &\quad = \left\{\det{}^{1/2}\left[\frac{R^{TX}_{x_0}/(2\pi i)}{e^{R^{TX}_{x_0}/(4\pi i)}-e^{-R^{TX}_{x_0}/(4\pi i)}}\right]\operatorname{Tr}\left[\exp\left(-\frac{R^E_{x_0}}{2\pi i}\right)\right]\right\}^{\max} \\ &\quad = \left\{\widehat{A}(TX,\nabla^{TX})\operatorname{ch}(E,\nabla^E)\right\}^{\max}. \end{aligned} \tag{1.64}$$

This completes the proof of (1.37). Then we finish the proof of the Atiyah–Singer index theorem. □

2. Holomorphic Morse inequalities

Let (X,J) be a compact complex manifold with complex structure J and $\dim_{\mathbb{C}} X = n$. Then we can identify the holomorphic tangent bundle $T^{(1,0)}X$ (resp. anti-holomorphic tangent bundle $T^{(0,1)}X$) as the eigenspace of J with eigenvalue i

(resp. $-i$) on $TX \otimes_{\mathbb{R}} \mathbb{C}$. Let $T^{*(0,1)}X$ be the anti-holomorphic cotangent bundle of X. Then formally,

$$\Lambda(T^{*(0,1)}X) = S(TX) \otimes (\det T^{(1,0)}X)^{1/2}, \tag{2.1}$$

here $\det F = \Lambda^{\mathrm{rk}(F)}F$ as the determinant line bundle of a vector bundle F.

If E is a holomorphic vector bundle on X. Let

$$\Omega^{0,\bullet}(X,E) = \mathscr{C}^\infty(X, \Lambda^\bullet(T^{*(0,1)}X) \otimes E)$$

be the space of anti-holomorphic differential forms on X with values in E. Then as in (1.4), we can define the Dolbeault operator

$$\overline{\partial}^E : \Omega^{0,k}(X,E) \to \Omega^{0,k+1}(X,E)$$

by using $\overline{\partial}^E$ on $\mathscr{C}^\infty(X,E)$ induced by the holomorphic structure on E. Moreover, $\left(\overline{\partial}^E\right)^2 = 0$. Denote by $H^\bullet(X,E)$ the Dolbeault cohomology of X with values in E, i.e.,

$$H^q(X,E) = \frac{\mathrm{Ker}(\overline{\partial}^E|_{\Omega^{0,q}(X,E)})}{\mathrm{Im}(\overline{\partial}^E|_{\Omega^{0,q-1}(X,E)})}. \tag{2.2}$$

Let g^{TX} be a J-invariant metric on TX and h^E be a Hermitian metric on E. Then they induce naturally an L^2-Hermitian product on $\Omega^{0,\bullet}(X,E)$ via

$$\langle s_1, s_2 \rangle = \int_X \langle s_1, s_2 \rangle(x) dv_X(x). \tag{2.3}$$

Let $\overline{\partial}^{E,*}$ be the formal adjoint of $\overline{\partial}^E$, and

$$D = \sqrt{2}\big(\overline{\partial}^E + \overline{\partial}^{E,*}\big). \tag{2.4}$$

Then

$$D^2 = 2\big(\overline{\partial}^E\overline{\partial}^{E,*} + \overline{\partial}^{E,*}\overline{\partial}^E\big). \tag{2.5}$$

Thus D^2 preserves the $\mathbb{Z}$-grading on $\Omega^{0,\bullet}(X,E)$. By Hodge theory, we have

$$\mathrm{Ker}\left(D^2|_{\Omega^{0,q}(X,E)}\right) \simeq H^q(X,E) \quad \text{for any } q. \tag{2.6}$$

Remark 2.1. If (X, g^{TX}) is Kähler and (E, h^E) is a holomorphic Hermitian vector bundle on X with Chern connection ∇^E, i.e., the $(0,1)$-part of ∇^E is $\overline{\partial}^E$ and ∇^E is Hermitian, then D in (2.4) is the Dirac operator in (1.22) acting on $\Lambda^\bullet(T^{*(0,1)}X) \otimes E$.

Theorem 2.2 (Riemann–Roch–Hirzebruch Theorem).

$$\sum_{j=0}^n (-1)^j \dim H^j(X,E) = \int_X \mathrm{Td}(T^{(1,0)}X)\,\mathrm{ch}(E). \tag{2.7}$$

If X is projective, then Theorem 2.2 is the original Riemann–Roch–Hirzebruch theorem. If X is only a compact complex manifold, then (2.7) is a consequence of the Atiyah–Singer index theorem for the Spin^c Dirac operator and (2.6).

Question: How to estimate $\dim H^q(X,E)$ in geometric way? If it is not possible, then at least asymptotically?

The following Theorem 2.3 gives a positive answer to the above question. It is an analogue of the classical Morse inequalities: For a Morse function f on a compact manifold M, let $C_j(f)$ be the number of critical points of f with index j, then

$$\sum_{j=0}^{q}(-1)^{q-j}\dim H^j(M,\mathbb{C}) \leqslant \sum_{j=0}^{q}(-1)^{q-j}C_j(f) \quad \text{for any } 0\le q\le \dim_{\mathbb{R}} M. \tag{2.8}$$

Let L be a holomorphic Hermitian line bundle on X. Set $L^p = L^{\otimes p}$, the pth tensor power of L. For $0\leqslant j\leqslant n$, set

$$B_j^p = \dim H^j(X, L^p\otimes E). \tag{2.9}$$

Let h^L be a Hermitian metric on L and ∇^L be the Chern connection on (L,h^L) with curvature $R^L = (\nabla^L)^2$. We define $\dot{R}^L_x \in \mathrm{End}(T_x^{(1,0)}X)$ by

$$\langle \dot{R}^L u, \overline{v}\rangle = R^L(u,\overline{v}). \tag{2.10}$$

Set

$$X(q) = \Big\{x\in X : iR^L_x \text{ non-degenerate}, \dot{R}^L_x \text{ has exactly } q \text{ negative eigenvalues}\Big\},$$
$$X(\leqslant q) = \cup_{k\leqslant q} X(k). \tag{2.11}$$

Theorem 2.3 (Demailly). *As $p\to+\infty$, the following strong Morse inequalities hold for every $q=0,1,\dots,n$:*

$$\sum_{j=0}^{q}(-1)^{q-j}B_j^p \leqslant \mathrm{rk}(E)\frac{p^n}{n!}\int_{X(\leqslant q)}(-1)^q\left(\frac{i}{2\pi}R^L\right)^n + o(p^n), \tag{2.12}$$

with equality for $q=n$. In particular, we get the weak Morse inequalities

$$B_j^p \leqslant \mathrm{rk}(E)\frac{p^n}{n!}\int_{X(q)}(-1)^q\left(\frac{i}{2\pi}R^L\right)^n + o(p^n). \tag{2.13}$$

In 1987, Bismut gave a heat kernel proof of Demailly's holomorphic Morse inequalities by using probability theory. Here we gave a heat kernel proof by using Bismut–Lebeau's analytic localization techniques in local index theory [3, §11]. The starting point is the following analogue of the McKean–Singer formula in current context obtained first by Bismut [2]. As in (2.4), set

$$D_p = \sqrt{2}\big(\overline{\partial}^{L^p\otimes E} + \overline{\partial}^{L^p\otimes E,*}\big). \tag{2.14}$$

Theorem 2.4. *For any $u > 0$, $0 \leqslant q \leqslant n$, we have*

$$\sum_{j=0}^{q}(-1)^{q-j} B_j^p \leqslant \sum_{j=0}^{q}(-1)^{q-j} \operatorname{Tr}_j \left[e^{-\frac{u}{p}D_p^2}\right], \tag{2.15}$$

with equality for $q = n$. Again $\operatorname{Tr}_j \left[e^{-\frac{u}{p}D_p^2}\right]$ is the trace of $e^{-\frac{u}{p}D_p^2}$ on $\Omega^j(X, L^p \otimes E)$ which is given by

$$\operatorname{Tr}_j[e^{-\frac{u}{p}D_p^2}] = \int_X \operatorname{Tr}|_{\Lambda^j(T^{*(0,1)}X)\otimes L^p\otimes E}\left[e^{-\frac{u}{p}D_p^2}(x,x)\right] dv_X(x). \tag{2.16}$$

Note that

$$e^{-\frac{u}{p}D_p^2}(x,y) \in \bigoplus_{j=0}^{n} E^j_{p,x} \otimes E^{j,*}_{p,y}, \text{ with } E_p^j = \Lambda^j(T^{*(0,1)}X) \otimes L^p \otimes E. \tag{2.17}$$

As $\operatorname{End}(L) = \mathbb{C}$, thus

$$e^{-\frac{u}{p}D_p^2}(x,x) \in \bigoplus_{j=0}^{n} \operatorname{End}\left(\Lambda^j(T^{*(0,1)}X) \otimes E\right)_x.$$

Theorem 2.5 (Bismut). *For $u > 0$ fixed, as $p \to +\infty$, we have*

$$\begin{aligned}\exp\left(-\frac{u}{p}D_p^2\right)(x,x) &= (2\pi)^{-n}\frac{\det(\dot{R}^L)\exp(2u\omega_d)}{\det(1-\exp(-2u\dot{R}^L))} \otimes \operatorname{Id}_E p^n + o(p^n)\\ &= \prod_{j=1}^{n} \frac{a_j(x)\left(1+(e^{-2ua_j(x)}-1)\overline{w}^j \wedge i_{\overline{w}_j}\right)}{2\pi(1-e^{-2ua_j(x)})} \otimes \operatorname{Id}_E p^n + o(p^n),\end{aligned} \tag{2.18}$$

where we choose an orthonormal basis w_j of $T^{(1,0)}X$ such that

$$\dot{R}^L(x) = \operatorname{diag}(a_1(x), \dots, a_n(x)) \in \operatorname{End}(T_x^{(1,0)}X), \tag{2.19}$$

and

$$\omega_d = -\sum_{j=1}^{n} a_j(x)\overline{w}^j \wedge i_{\overline{w}_j}. \tag{2.20}$$

If $\omega(\cdot, J\cdot) = g^{TX}(\cdot,\cdot)$, then $a_j(x) = 2\pi$.

Proof. Bismut used probability theory to prove the result. Our proof is based on the analytic localization techniques of Bismut–Lebeau.

Step 1. The problem is local! Recall that from (1.50) there exists $C > 0$ such that

$$\left|a^k G_u(\sqrt{u}a)\right| \leqslant Ce^{-\frac{\varepsilon^2}{16u}}, \quad \text{for any } u \in (0,1],\ a \in \mathbb{R}. \tag{2.21}$$

Thus for $u > 0$ fixed, there exists $C_k > 0$ such that for $p \in \mathbb{N}$,

$$\left\| D_p^k G_{\frac{u}{p}}\left(\sqrt{\frac{u}{p}}D_p\right)\right\|^{0,0} \leqslant C_k e^{-\frac{\varepsilon^2 p}{32u}}. \tag{2.22}$$

Once we study carefully the Sobolev embedding theorem with parameter p, from (2.22) we know that there exist $c_1 > 0, c_2 > 0$ such that for any $x, y \in X$,

$$\left| G_{\frac{u}{p}} \left(\sqrt{\frac{u}{p}} D_p \right) (x, y) \right| \leqslant c_1 e^{-c_2 p}. \tag{2.23}$$

But supp $F_{\frac{u}{p}} \left(\sqrt{\frac{u}{p}} D_p \right) (x, \cdot) \subset B(x, \varepsilon)$ and $F_{\frac{u}{p}} \left(\sqrt{\frac{u}{p}} D_p \right) (x, \cdot)$ only depends on the restriction of D_p to $B(x, \varepsilon)$.

Step 2. Replace X by $\mathbb{R}^{2n} \simeq \mathbb{C}^n$, we work on $\mathbb{C}^n$.

Fix $x_0 \in X$. We identify $B^{T_{x_0}X}(0, 4\varepsilon)$ to $B^X(x_0, 4\varepsilon)$ by the exponential map: $v \to \exp_{x_0}(v)$. For $Z \in B^{T_{x_0}X}(0, 4\varepsilon) \subset T_{x_0}X$, we identify $\Lambda^\bullet(T_Z^{*(0,1)}X)$, L_Z and E_Z to $\Lambda^\bullet(T_{x_0}^{*(0,1)}X)$, L_{x_0} and E_{x_0} by parallel transport with respect to $\nabla^{B,\Lambda^{0,*}}$, ∇^L and ∇^E along the path $\gamma : [0,1] \to X, \gamma(s) = sZ$, where $\nabla^{B,\Lambda^{0,*}}$ is the connection on $\Lambda^\bullet(T^{*(0,1)}X)$ induced by the Bismut connection ∇^B on $T^{(1,0)}X$, in particular, it preserves the $\mathbb{Z}$-grading on $\Lambda^\bullet(T^{*(0,1)}X)$.

Step 3. Rescaling. Once we trivialized L we can consider that D_p^2 acts on $\mathscr{C}^\infty(\mathbb{R}^{2n}, \mathbf{E}_{x_0})$ with $\mathbf{E}_{x_0} = (\Lambda^\bullet(T^{*(0,1)}X) \otimes E)_{x_0}$. For $s \in \mathscr{C}^\infty(\mathbb{R}^{2n}, \mathbf{E}_{x_0})$, $Z \in \mathbb{R}^{2n}$ and $t = \frac{1}{\sqrt{p}}$, set

$$(S_t s)(Z) = s\left(\frac{Z}{t}\right), \qquad L_2^t = S_t^{-1} \frac{1}{p} D_p^2 S_t. \tag{2.24}$$

Then as $t \to 0$, with $\tau_{x_0} = \sum_{j=1}^n a_j(x_0)$,

$$L_2^t \to L_2^0 = -\sum_{j=1}^{2n} \left[\frac{\partial}{\partial Z_j} + \frac{1}{2} R_{x_0}^L \left(Z, \frac{\partial}{\partial Z_j} \right) \right]^2 - 2\omega_{d,x_0} - \tau_{x_0}. \tag{2.25}$$

Again (2.25) is obtained from the Lichnerowicz formula for D_p^2 obtained by Bismut:

$$D_p^2 = \Delta^{B,\Lambda^{0,*}} + p\ {}^c R^L + 0\text{-order term independent of } p, \tag{2.26}$$

and $\Delta^{B,\Lambda^{0,*}}$ is the Bochner Laplacian acting on $\mathscr{C}^\infty(X, E_p)$ associated with $\nabla^{B,\Lambda^{0,*}}$, ∇^L and ∇^E.

From the finite propagation speed for the wave operator, for $u > 0$ fixed, we obtain as $p \to +\infty$,

$$e^{-\frac{u}{p} D_p^2}(x_0, x_0) = p^n e^{-uL_2^t}(0,0) + \mathscr{O}(e^{-cp}). \tag{2.27}$$

By using weighted Sobolev norms adapted from the structure of the operator L_2^t, we get

Theorem 2.6. *As $t \to 0$,*

$$e^{-uL_2^t}(0,0) \to e^{-uL_2^0}(0,0). \tag{2.28}$$

Finally, from (1.62) and (2.25), we get

$$e^{-uL_2^0}(0,0) = (2\pi)^{-n} \frac{\det(\dot{R}^L_{x_0})e^{2u\omega_{d,x_0}}}{\det(1-e^{-2u\dot{R}^L_{x_0}})}. \tag{2.29}$$

From Theorem 2.4, (2.27)–(2.29), as $p \to +\infty$,

$$\begin{aligned} &\sum_{j=0}^{q}(-1)^{q-j}B_j^p \\ &\leqslant \sum_{j=0}^{q}(-1)^{q-j}\int_X \operatorname{Tr}|_{\Lambda^j(T^{*(0,1)}X)\otimes E}\left[(2\pi)^{-n}\frac{\det(\dot{R}^L_{x_0})e^{2u\omega_{d,x_0}}}{\det(1-e^{-2u\dot{R}^L_{x_0}})}\otimes \mathrm{Id}_E\right]dv_X(x)\cdot p^n \\ &\qquad + o(p^n). \end{aligned} \tag{2.30}$$

One can verify directly that

$$\begin{aligned} \lim_{u\to+\infty}\int_X \operatorname{Tr}|_{\Lambda^j(T^{*(0,1)}X)}\left[(2\pi)^{-n}\frac{\det(\dot{R}^L_{x_0})e^{2u\omega_{d,x_0}}}{\det(1-e^{-2u\dot{R}^L_{x_0}})}\right]dv_X(x) \\ = \int_{X(j)}(-1)^j\frac{1}{n!}\left(\frac{iR^L}{2\pi}\right)^n. \end{aligned} \tag{2.31}$$

Combining (2.30) and (2.31) yields (2.12). □

3. Bergman kernels

3.1. Asymptotic expansion of Bergman kernels

Let (X,J) be a compact complex manifold with complex structure J and $\dim_{\mathbb{C}} X = n$. Let (L,h^L), (E,h^E) be holomorphic Hermitian vector bundles on X and $\mathrm{rk}(L) = 1$. Let ∇^L be the Chern connection on (L,h^L) with curvature

$$R^L = (\nabla^L)^2 \in \Omega^{1,1}(X,\mathrm{End}(L)) = \Omega^{1,1}(X,\mathbb{C}). \tag{3.1}$$

Assumption: $\omega = \frac{i}{2\pi}R^L$ is positive (equivalently, $w(\cdot,J\cdot)$ defines a metric on TX).

By the Kodaira vanishing theorem, we have for any $q>0$,

$$H^q(X,L^p\otimes E) = 0 \qquad \text{for } p \gg 1. \tag{3.2}$$

Let g^{TX} be any J-invariant Riemannian metric on TX. Let P_p be the orthogonal projection from $\mathscr{C}^\infty(X,L^p\otimes E)$ onto $H^0(X,L^p\otimes E)$. Its smooth kernel is

$$P_p(x,y) = \sum_{i=1}^{d_p} S_i^p(x)\otimes(S_i^p(y))^* \in (L^p\otimes E)_x\otimes(L^p\otimes E)_y^*, \tag{3.3}$$

where $\{S_i^p\}_{i=1}^{d_p}$ ($d_p := \dim H^0(X,L^p\otimes E)$) is an orthonormal basis of $H^0(X,L^p\otimes E)$. In particular,

$$P_p(x,x) \in \mathrm{End}(L^p\otimes E)_x = \mathrm{End}(E)_x. \tag{3.4}$$

If $E = \mathbb{C}$, then

$$P_p(x,x) = \sum_{i=1}^{d_p} |S_i^p(x)|^2 : \ X \to [0,+\infty). \tag{3.5}$$

By the Riemann–Roch–Hirzebruch theorem and (3.2), we have for p large enough,

$$\begin{aligned}\int_X \operatorname{Tr}|_E[P_p(x,x)]dv_X(x) &= \dim H^0(X, L^p\otimes E) = \int_X \operatorname{Td}(T^{(1,0)}X)\operatorname{ch}(L^p\otimes E)\\ &= \int_X \operatorname{Td}(T^{(1,0)}X)\operatorname{ch}(E)e^{p\omega} = \operatorname{rk}(E)\int_X \frac{c_1(L)^n}{n!}p^n \\ &\quad + \int_X \left(c_1(E) + \frac{\operatorname{rk}(E)}{2}c_1(T^{(1,0)}X)\right)\frac{c_1(L)^{n-1}}{(n-1)!}p^{n-1} + \mathscr{O}(p^{n-2}).\end{aligned} \tag{3.6}$$

Question: Whether as $p \to +\infty$,

$$\operatorname{Tr}|_E[P_p(x,x)]dv_X(x) = \operatorname{Td}(T^{(1,0)}X, \nabla^{T^{(1,0)}X})\operatorname{ch}(E,\nabla^E)_x e^{p\omega_x} + \mathscr{O}(p^{-\infty}), \tag{3.7}$$

where $\nabla^{T^{(1,0)}X}$ is the Chern connection on $(T^{(1,0)}X, g^{TX})$.

The following is a local version of the expansion.

Theorem 3.1 (Tian, Ruan, Catlin, Zelditch, Boutet de Monvel–Sjöstrand, Dai–Liu–Ma, Ma–Marinescu, ...). *There exist* $\boldsymbol{b}_j \in \mathscr{C}^\infty(X, \operatorname{End}(E))$ *such that for any* k*, as* $p \to +\infty$*, we have uniformly on* X*,*

$$p^{-n}P_p(x,x) = \sum_{j=0}^{k} \boldsymbol{b}_j(x)p^{-j} + \mathscr{O}(p^{-k-1}), \tag{3.8}$$

with

$$\boldsymbol{b}_0 = \det(\dot{R}^L/(2\pi))\operatorname{Id}_E. \tag{3.9}$$

The Kodaira embedding theorem shows that for $p \gg 1$, L^p give rise to holomorphic embeddings $\Phi_p : X \to \mathbb{P}(H^0(X,L^p)^*)$. Moreover, $L^p = \Phi_p^*\mathcal{O}(1)$ and $h^{L^p}(x) = P_p(x,x)h^{\Phi_p^*\mathcal{O}(1)}(x)$ (cf. [11, Theorem 5.1.3]). Here $\mathcal{O}(1)$ is the hyperplane line bundle on $\mathbb{P}(H^0(X,L^p)^*)$ with the metric $h^{\mathcal{O}(1)}$ induced naturally from the Hermitian product on $H^0(X,L^p)$. Thus

$$\frac{1}{p}\Phi_p^*\omega_{FS} - \omega = -\frac{i}{2\pi p}\overline{\partial}\partial \log P_p(x,x). \tag{3.10}$$

Where ω_{FS} is the Fubini–Study form on the projective space $\mathbb{P}(H^0(X,L^p)^*)$.

From (3.8) and (3.10), we know that the induced Fubini–Study forms via Kodaira embedding maps Φ_p is dense in the space of Kähler form in the Kähler class $c_1(L)$. More precisely,

Corollary 3.2 (Tian, Ruan). *For* $k > 0$*, there exists* $C > 0$ *such that*

$$\left|\frac{1}{p}\Phi_p^*(\omega_{\mathrm{FS}}) - \omega\right|_{\mathscr{C}^k} \leqslant \frac{C}{p}\cdot \tag{3.11}$$

Ruan improved Tian's asserting convergence in $\mathscr{C}^2$ topology with speed rate $p^{-1/2}$. The optimal convergence speed of the induced Fubini–Study forms in the symplectic case was obtained in [9].

3.2. Proof of the asymptotic expansion of Bergman kernels

In this subsection, we obtain the asymptotic behavior of $P_p(x,y)$ as $p \to +\infty$ via the analytic localization techniques of Bismut–Lebeau [3, §11]. The method also works in the symplectic case by Dai–Liu–Ma [6], also Ma–Marinescu [12]. The starting point of the approach is the following spectral gap result.

Theorem 3.3 (Bismut–Vasserot (1989); Ma–Marinescu, symplectic version (2002)). *There exists $C > 0$ such that for any $p \in \mathbb{N}^*$,*

$$\operatorname{Spec}(D_p^2) \subset \{0\} \cup [2p\mu_0 - C, +\infty), \tag{3.12}$$

where

$$\mu_0 = \inf_{\substack{x\in X,\\ 0\neq u\in T_x^{(1,0)}X}} \frac{R^L(u,\overline{u})}{|u|^2}. \tag{3.13}$$

If $w(\cdot, J\cdot) = g^{TX}$, then $\mu_0 = 2\pi$.

Proof of Theorem 3.1. We divide the proof into three steps.

Step 1. The problem is local, i.e., module $\mathscr{O}(p^{-\infty})$, $P_p(x_0,\cdot)$ depends only on $D_p|_{B^X(x_0,\varepsilon)}$. Let $f : \mathbb{R} \to [0,1]$ be a smooth even function such that $f(v) = 1$ for $|v| \leqslant \varepsilon/2$ and that $f(v) = 0$ for $|v| \geqslant \varepsilon$. Take

$$F(a) = \left(\int_{-\infty}^{+\infty} f(v)dv\right)^{-1} \int_{-\infty}^{+\infty} e^{iva} f(v)dv. \tag{3.14}$$

Then $F(0) = 1$ and for $p > C/\mu_0$,

$$P_p = F(D_p) - 1_{[\sqrt{p\mu_0},+\infty)}(|D_p|)F(D_p). \tag{3.15}$$

On one hand, by the finite propagation speed of solutions of wave equations, we have $\operatorname{supp} F(D_p)(x_0,\cdot) \subset B^X(x_0,\varepsilon)$ and $F(D_p)(x_0,\cdot)$ depends only on $D_p|_{B^X(x_0,\varepsilon)}$. On the other hand, as

$$\sup_{a\in\mathbb{R}} |a|^m |F(a)| \leqslant C_m, \tag{3.16}$$

which implies that the smooth kernel of the operator $1_{[\sqrt{p\mu_0},+\infty)}(|D_p|)F(D_p)$ has the following property: as $p \to +\infty$,

$$1_{[\sqrt{p\mu_0},+\infty)}(|D_p|)F(D_p)(x,y) = \mathscr{O}(p^{-\infty}). \tag{3.17}$$

As $F(D_p)(x,y) = 0$ if $d(x,y) > \varepsilon$, where $d(\,,\,)$ is the Riemannian distance on (X, g^{TX}). Thus we know that if $d(x,y) > \varepsilon$, then

$$P_p(x,y) = \mathscr{O}(p^{-\infty}). \tag{3.18}$$

Step 2. We replace X by $\mathbb{R}^{2n} =: X_0$. We identify $B^{T_{x_0}X}(0,4\varepsilon)$ in $T_{x_0}X$ to $B^X(x_0,4\varepsilon)$ by the exponential map: $v \to \exp_{x_0}(v)$. For $Z \in B^{T_{x_0}X}(0,4\varepsilon) \subset T_{x_0}X$, we identify $\Lambda(T_Z^{*(0,1)}X)$, L_Z and E_Z to $\Lambda(T_{x_0}^{*(0,1)}X)$, L_{x_0} and E_{x_0} by parallel transport with respect to $\nabla^{B,\Lambda^{0,*}}$, ∇^L and ∇^E along the path $\gamma : [0,1] \to X, \gamma(s) = sZ$.

Step 3. Rescaling. Let $\rho : \mathbb{R} \to [0,1]$ be a smooth even function such that

$$\rho(v) = 1 \quad \text{if} \quad |v| < 2; \qquad \rho(v) = 0 \quad \text{if} \quad |v| > 4. \tag{3.19}$$

Set $\varphi_\varepsilon(Z) = \rho(|Z|/\varepsilon)Z$. For the trivial vector bundle $L_0 := (L_{x_0}, h^{L_{x_0}})$, we defined a Hermitian connection on $X_0 := T_{x_0}X$ by

$$\nabla^{L_0}|_Z = \varphi_\varepsilon^* \nabla^L + \frac{1}{2}(1 - \rho^2(|Z|/\varepsilon))R^L_{x_0}(Z, \cdot). \tag{3.20}$$

The important observation is that the curvature $(\nabla^{L_0})^2$ of ∇^{L_0} is uniformly positive on $\mathbb{R}^{2n}$ and its small eigenvalues in the sense of (3.13) is bigger than $\frac{4}{5}\mu_0$ for ε small enough. We obtain a modified Dirac operator $D_{0,p}$ on X_0 with

$$\operatorname{Spec}(D_{0,p}^2) \subset \{0\} \cup \left[\frac{8}{5}p\mu_0 - C, +\infty\right). \tag{3.21}$$

Denote by $P_{0,p}$ the orthogonal projection from $L^2(X_0, E_{0,p})$ onto $\operatorname{Ker}(D_{0,p}^2)$. Then

$$P_p = P_{0,p} + \mathscr{O}(p^{-\infty}). \tag{3.22}$$

For large p, we have

$$\begin{aligned} P_{0,p} &= e^{-\frac{u}{p}D_{0,p}^2} - e^{-\frac{u}{p}D_{0,p}^2} 1_{(p\mu_0,+\infty)}(D_{0,p}^2) \\ &= e^{-\frac{u}{p}D_{0,p}^2} - \int_u^\infty \frac{1}{p} D_{0,p}^2 e^{-\frac{v}{p}D_{0,p}^2} dv. \end{aligned} \tag{3.23}$$

Then for u fixed we have the asymptotic expansion of $e^{-\frac{u}{p}D_{0,p}^2}$ and

$$\frac{1}{p} D_{0,p}^2 e^{-\frac{u}{p}D_{0,p}^2} = \mathscr{O}(e^{-cu}).$$

This indicates that we can approximate the Bergman kernel by using heat kernels. The detail of this approach was first realized by Dai–Liu–Ma in [6] by using the analytic localization techniques of Bismut–Lebeau. In fact they obtain the full asymptotics of $P_p(x,y)$ as $p \to +\infty$. This approach works for the symplectic case, also the singular case with orbifold singularities. Ma–Marinescu [8, 11, 13, 14] use this kind of expansion to establish the Berezin–Toeplitz geometric quantization theory in symplectic case. The Berezin–Toeplitz theory has played an important role in the recent works on the asymptotics of analytic torsions [4], [15]. □

3.3. Coefficients of the asymptotic expansion of Bergman kernels

In the last part, we explain how to compute the coefficient in the expansion.

For $t = \frac{1}{\sqrt{p}}$, set

$$(S_t s)(Z) = s\left(\frac{Z}{t}\right), \ \mathscr{L}_t = S_t^{-1} \frac{1}{p} D_{0,p}^2 S_t. \tag{3.24}$$

Then the Taylor expansion of $\mathscr{L}_t$ gives

$$\mathscr{L}_t = \mathscr{L}_0 + \sum_{r=1}^{k} t^r \mathcal{O}_r + \mathscr{O}(t^{k+1}), \tag{3.25}$$

where under the notation of (2.20),

$$\mathscr{L}_0 = \sum_j \left(-2\tfrac{\partial}{\partial z_j} + \frac{1}{2} a_j \overline{z}_j\right)\left(2\tfrac{\partial}{\partial \overline{z}_j} + \frac{1}{2} a_j z_j\right) + 2a_j \overline{w}^j \wedge i_{\overline{w}_j}. \tag{3.26}$$

Set

$$b_j = -2\tfrac{\partial}{\partial z_j} + \frac{1}{2} a_j \overline{z}_j\,, \quad b_j^+ = 2\tfrac{\partial}{\partial \overline{z}_j} + \frac{1}{2} a_j z_j\,, \quad \mathscr{L} = \sum_j b_j b_j^+. \tag{3.27}$$

Then

$$\mathscr{L}_0 = \mathscr{L} - 2\omega_{d,x_0} = \mathscr{L} + 2\sum_j a_j \overline{w}^j \wedge i_{\overline{w}_j}. \tag{3.28}$$

One verifies directly that for the spectrum of $\mathscr{L}$,

$$\mathrm{Spec}(\mathscr{L}|_{L^2(\mathbb{R}^{2n})}) = \left\{2\sum\nolimits_{j=1}^{n} \alpha_j a_j \,:\, \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n\right\}, \tag{3.29}$$

and that an orthogonal basis of the eigenspace of $2\sum_{j=1}^n \alpha_j a_j$ is given by

$$b^\alpha \left(z^\beta \exp\left(-\frac{1}{4}\sum\nolimits_j a_j |z_j|^2\right)\right), \quad \text{with } \beta \in \mathbb{N}^n\,. \tag{3.30}$$

Thus

$$\mathrm{Ker}(\mathscr{L}) = \left\{z^\beta \exp\left(-\frac{1}{4}\sum\nolimits_j a_j |z_j|^2\right), \quad \text{with } \beta \in \mathbb{N}^n\right\}. \tag{3.31}$$

The orthogonal projection from $L^2(\mathbb{R}^{2n}, \mathbb{C})$ onto $\mathrm{Ker}(\mathscr{L})$ is the classical Bergman kernel on $\mathbb{C}^n$ associated with the trivial line bundle with metric

$$|1|_{h^L}(Z) = \exp\left(-\frac{1}{4}\sum\nolimits_j a_j |z_j|^2\right). \tag{3.32}$$

The classical Bergman kernel is given by

$$\mathcal{P}(Z, Z') = \prod_{j=1}^{n} \frac{a_j}{2\pi} \exp\left(-\frac{1}{4}\sum\nolimits_j a_j \left(|z_j|^2 + |z_j'|^2 - 2z_j \overline{z}_j'\right)\right). \tag{3.33}$$

As all our operators preserve $\mathbb{Z}$-grading of $\mathbf{E}_{x_0}$ and the degree $\geqslant 1$ part is zero. We can restrict all the following computation on 0-degree part, i.e., on $\mathscr{C}^\infty(X_0, E_{x_0})$.

Let $P_{0,t}$ be the spectral projection

$$P_{0,t} : L^2(X_0, E_{x_0}) \to \mathrm{Ker}(\mathscr{L}_t)$$

and $P_{0,t}(Z,Z)$ the smooth kernel of $P_{0,t}$. From (3.25), by the formal expansion of the resolvent for $|\lambda|=\mu_0/4$

$$(\lambda-\mathscr{L}_t)^{-1}=\sum_{r=0}^{\infty}t^r f_r(\lambda), \tag{3.34}$$

we obtain as $t\to 0$,

$$P_{0,t}=\frac{1}{2\pi i}\int_{|\lambda|=\mu_0/4}(\lambda-\mathscr{L}_t)^{-1}d\lambda=P^N+\frac{1}{2\pi i}\sum_{r=1}^{k}t^r\int_{|\lambda|=\mu_0/4}f_r(\lambda)d\lambda+\mathscr{O}(t^{k+1}), \tag{3.35}$$

with $P^N=\mathcal{P}\ \mathrm{Id}_E$. Then from (3.24)

$$P_{0,p}(Z,Z')=t^{-2n}P_{0,t}\left(\frac{Z}{t},\frac{Z'}{t}\right). \tag{3.36}$$

From (3.22) and (3.36), the kernel of the coefficient of t^r in (3.35) gives the coefficient of $p^{-r/2}$ in the off-diagonal expansion of $p^{-n}P_p(Z,Z')$ by Dai–Liu–Ma, Ma–Marinescu. In particular, $\boldsymbol{b}_j$ in (3.8) is given by the evaluation of the kernel of the coefficient of t^{2j} in (3.35) at $(0,0)$.

Remark 3.4. In the Kähler case, i.e., $\omega(\cdot,J\cdot)=g^{TX}(\cdot,\cdot)$, then all $a_j=2\pi$,

$$\mathcal{O}_1=0,$$

and

$$\boldsymbol{b}_1(x)=\big(-\mathscr{L}^{-1}\mathcal{O}_2P^N-P^N\mathcal{O}_2\mathscr{L}^{-1}\big)(0,0)=\frac{1}{8\pi}\big[r^X+4R^E(w_j,\overline{w}_j)\big], \tag{3.37}$$

here $\{w_j\}$ is an orthonormal basis of $T^{(1,0)}_{x_0}X$ and r^X is the scalar curvature of (X,g^{TX}). Note that in the Kähler case, $\boldsymbol{b}_1$ was obtained first by Lu and Wang by using the pick section trick in complex analysis as in [16].

Acknowledgment

We thank Wen Lu for his help on this note. This work was partially supported by ANR contract ANR-14-CE25-0012-01, NNSFC No. 11829102.

References

[1] N. Berline, E. Getzler, and M. Vergne, *Heat kernels and Dirac operators*, Grundlehren Text Editions, Springer-Verlag, Berlin, 2004, Corrected reprint of the 1992 original.

[2] J.-M. Bismut, *Demailly's asymptotic inequalities: a heat kernel proof*, J. Funct. Anal. 72 (1987), 263–278.

[3] J.-M. Bismut and G. Lebeau, *Complex immersion and Quillen metrics*, Inst. Hautes Étud. Sci. Publ. Math. 74 (1991), 1–297.

[4] J.-M. Bismut, X. Ma and W. Zhang, *Asymptotic torsion and Toeplitz operators*, Journal of the Institute of Mathematics of Jussieu, **16** (2017), 223–349.

[5] J.-M. Bismut and E. Vasserot, *The asymptotics of the Ray–Singer analytic torsion associated with high powers of a positive line bundle*, Comm. Math. Phys. 125 (1989), 355–367.

[6] X. Dai, K. Liu and X. Ma, *On the asymptotic expansion of Bergman kernel*, J. Differential Geom. 72 (2006), 1–41.

[7] J.-P. Demailly, *Champs magnétiques et inequalités de Morse pour la d-cohomologie*, Ann. Inst. Fourier (Grenoble) 35 (1985), 189–229.

[8] L. Ioos, W. Lu, X. Ma and G. Marinescu, *Berezin–Toeplitz quantization for eigenstates of the Bochner-Laplacian on symplectic manifolds*, arXiv:1703.06420, J. Geom. Anal.

[9] W. Lu, X. Ma and G. Marinescu, *Optimal convergence speed of Bergman metrics on symplectic manifolds*, 23pp. arXiv:1702.00974, Journal of Symplectic Geometry.

[10] X. Ma, *Geometric quantization on Kähler and symplectic manifolds*, Proceedings of the International Congress of Mathematicians. Volume II, 785–810, Hindustan Book Agency, New Delhi, 2010.

[11] X. Ma and G. Marinescu, *Holomorphic Morse inequalities and Bergman kernels*, Progr. Math. 254, Birkhäuser Verlag, Basel, 2007.

[12] X. Ma and G. Marinescu, *Generalized Bergman kernels on symplectic manifolds*, Adv. in Math. 217 (2008), 1756–1815.

[13] X. Ma and G. Marinescu, *Toeplitz operators on symplectic manifolds*, J. Geom. Anal., **18** (2008), 565–611.

[14] X. Ma and G. Marinescu, *Berezin–Toeplitz quantization on Kähler manifolds*, J. Reine Angew. Math. 662 (2012), 1–58.

[15] M. Puchol, *The asymptotics of the holomorphic analytic torsion forms*, arXiv: 1511.04694.

[16] G. Tian, *On a set of polarized Kähler metrics on algebraic manifolds*, J. Differential Geom. 32 (1990), 99–130.

Xiaonan Ma
Université de Paris & Université Paris Diderot – Paris 7
UFR de Mathématiques
Case 7012
F-75205 Paris Cedex 13, France
e-mail: `xiaonan.ma@imj-prg.fr`

Progress in Mathematics, Vol. 333, 287–303

Fourier–Mukai Transforms, Euler–Green Currents, and K-Stability

Sean Timothy Paul and Kyriakos Sergiou

Abstract. Inspired by Gang Tian's work in [4, 10, 11], and [12] we exhibit a wide range of energy functionals in Kähler geometry as Fourier–Mukai transforms. Consequently these energies are completely determined by dual type varieties and therefore have logarithmic singularities when restricted to the space of algebraic potentials. This paper is dedicated to Gang Tian on the occasion of his 60th birthday.

Mathematics Subject Classification (2010). 53C55.

Keywords. Discriminants, resultants, K-energy maps, Bott–Chern forms, csc Kähler metrics, K-stability.

1. Introduction and statement of results

Let $\mathbb{X} \xrightarrow{\pi} B$ be a flat (relative dimension n) family of smooth polarized, linearly normal, complex subvarieties of some fixed $\mathbb{P}^N$ parametrized by a reduced and irreducible (quasi) projective base B. We do not assume that B is smooth. Let Q be a locally free sheaf of rank $n+1$ over $\mathbb{X}$. We assume that

$$\begin{array}{c}\text{There is a finite-dimensional subspace}\\ W \subset \Gamma(\mathbb{X}\ \ Q) \text{ generating each fiber.}\end{array} \tag{$**$}$$

This is equivalent to an exact sequence of vector bundles over $\mathbb{X}$

$$0 \longrightarrow \mathcal{S} \longrightarrow \mathbb{X} \times W \longrightarrow Q \longrightarrow 0$$

where the sub bundle $\mathcal{S}$ is the kernel of the evaluation map $\mathbb{X} \times W \longrightarrow Q$.
This defines

- $\mathcal{E} := Q \otimes \mathcal{O}_{\mathbb{P}(W)}(1)$.
- $\Psi \in \Gamma(\mathbb{X} \times \mathbb{P}(W), Q \otimes \mathcal{O}_{\mathbb{P}(W)}(1))$ (the evaluation section).
- $I := \mathbb{P}(\mathcal{S}) = (\Psi = 0) \subset \mathbb{X} \times \mathbb{P}(W)$.
- $Z := \pi(I)$.

By abuse of notation we also denote by π the projection

$$\pi : \mathbb{X} \times \mathbb{P}(W) \longrightarrow B \times \mathbb{P}(W).$$

The situation may be visualized as follows

$$\begin{tikzcd} & \mathcal{E} \arrow[d] \\ I \arrow[r, hook, "\iota"] \arrow[d] & \mathbb{X} \times \mathbb{P}(W) \arrow[u, bend left, "\Psi"] \arrow[d, "\pi"] \\ Z \arrow[r, hook] \arrow[d] & B \times \mathbb{P}(W) \arrow[d] \\ B \arrow[r] & B\,. \end{tikzcd} \tag{1.1}$$

Observe that the codimension of I in $\mathbb{X} \times \mathbb{P}(W)$ is equal to $n+1$.

The **basic assumption** in this paper is that $I \longrightarrow^{\pi} Z$ is a *simultaneous resolution of singularities*. This means:

i) $I \xrightarrow{\pi} Z$ is birational.

ii) For each $b \in B$, $\quad I_b := \mathbb{P}(\mathcal{S}|_{\mathbb{X}_b}) \xrightarrow{\pi} Z_b := Z \cap (\{b\} \times \mathbb{P}(W))$ is also birational.

Our basic assumption implies that the codimension of Z_b in $\mathbb{P}(W)$ is one. All of this data will be called a *basic set up* for the family $\mathbb{X} \longrightarrow B$.

Next we observe that the degrees of the divisors Z_b are constant in $b \in B$. This follows at once from the identity

$$\deg(Z_b) = \int_{\mathbb{X}_b} c_n(Q|_{\mathbb{X}_b}) > 0\,.$$

We will denote this common degree by $\widehat{d}$

$$\deg(Z_b) = \widehat{d} \quad \text{for all } b \in B\,.$$

Therefore, for each $b \in B$ there is a unique (up to scale) irreducible polynomial $f_b \in \mathbb{C}_{\widehat{d}}\,[W] \setminus \{0\}$ such that

$$(f_b = 0) = Z_b\,.$$

We may therefore define a map

$$\Delta : B \longrightarrow |\mathcal{O}(\widehat{d})| := \mathbb{P}(\mathbb{C}_{\widehat{d}}\,[W]) \quad \Delta(b) := [f_b]\,.$$

The main result of this paper is the following

Theorem 1.1. Δ *is a morphism of quasi-projective varieties. In particular,* Δ *is holomorphic.*

1.1. Hermitean metrics and base change

Now we introduce Hermitean metrics on everything via "base change". Let S be a smooth complex algebraic variety together with a morphism to the base of our family $\mathbb{X} \xrightarrow{\pi} B$

$$S \longrightarrow B\,.$$

The basic set up over B may be pulled back to a basic set up over S. The corresponding morphism will still be denoted by Δ. Our first assumption on S is that

$$\Delta^*\mathcal{O}(1) \cong \mathcal{O}_S \qquad \mathcal{O}(1) := \text{the hyperplane over } |\mathcal{O}(\widehat{d})|\,.$$

Equivalently we may "lift" the map Δ the cone over $|\mathcal{O}(\widehat{d})|$

$$\begin{tikzcd} & \mathbb{C}_{\widehat{d}}\,[W] \setminus \{0\} \arrow[d] \\ S \arrow[ur] \arrow[r, "\Delta"] & |\mathcal{O}(\widehat{d})| \end{tikzcd}$$

By abuse of notation the lifted map will also be denoted by Δ. Next we fix a positive definite Hermitean form $\langle\cdot\,,\,\cdot\rangle$ on W. We let $\mathcal{H}^+(W, \langle\cdot\,,\,\cdot\rangle)$ denote self adjoint positive linear maps on W. This parametrizes all positive Hermitian forms on W. Let h be a smooth map

$$h : S \longrightarrow \mathcal{H}^+(W, \langle\cdot\,,\,\cdot\rangle)\,.$$

Using this map we define a "dynamic" metric H_S on the trivial bundle $\mathbb{X}_S \times W$

$$H_S(p)\,\langle w_1, w_2\rangle := \langle h(\pi(p))w_1, w_2\rangle\ .$$

Similarly we have the "static" metric

$$H_0(p)\,\langle w_1, w_2\rangle := \langle w_1, w_2\rangle\ .$$

These metrics descend to metrics H_S^Q and H_0^Q on Q. The Hermitean inner product on W induces a Fubini Study metric h^m_{FS} on $\mathcal{O}_{\mathbb{P}(W)}(m)$ for any $m \in \mathbb{Z}$ as well as a Kähler form ω_{FS} on $\mathbb{P}(W)$. This will be fixed throughout the paper. Tensoring H_S^Q and H_0^Q with h_{FS} gives two metrics $H_S^{\mathcal{E}}$ and $H_0^{\mathcal{E}}$on $\mathcal{E}$.

We need the following result from [1].

Theorem 1.2 (Bismut, Gillet, Soulé)). *Let $\mathcal{E} \xrightarrow{\pi} M$ be a rank r holomorphic vector bundle over a complex manifold M. Given any Hermitean metric $H^{\mathcal{E}}$ on $\mathcal{E}$ there exists a current $\hat{\mathbf{e}}(H^{\mathcal{E}}) \in \mathcal{D}^{'}(\mathcal{E})$ whose wave front set is included in $\mathcal{E}^*_{\mathbb{R}}$ and which satisfies the following equation of currents on $\mathcal{E}$*

$$\frac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\hat{\mathbf{e}}(H^{\mathcal{E}}) = \delta_M - \pi^*\mathbf{e}(H^{\mathcal{E}})\,.$$

The current $\hat{\mathbf{e}}(H^{\mathcal{E}})$ can be pulled back to M by any section $\mathbf{s}$ and satisfies

$$\frac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\mathbf{s}^*\hat{\mathbf{e}}(H^{\mathcal{E}}) = \delta_{Z(\mathbf{s})} - \mathbf{e}(H^{\mathcal{E}})$$

where $\mathbf{e}(H^{\mathcal{E}}) \in C^\infty(\Lambda_M^{r,r})$ is the Euler form of $\mathcal{E}$ in Chern–Weil theory. Moreover, given two Hermitean metrics $H_0^{\mathcal{E}}$ and $H_1^{\mathcal{E}}$ the difference of the corresponding currents is smooth and up to ∂ and $\overline{\partial}$ terms is given by

$$\mathbf{s}^*\hat{\mathbf{e}}(H_0^{\mathcal{E}}) - \mathbf{s}^*\hat{\mathbf{e}}(H_1^{\mathcal{E}}) = -\mathbf{e}(H_1^{\mathcal{E}}, H_0^{\mathcal{E}})$$

where $\mathbf{e}(H_1^{\mathcal{E}}, H_0^{\mathcal{E}})$ denotes the Bott–Chern double transgression of the Euler form with respect the given metrics.

We apply this result to our situation. $M = \mathbb{X} \times \mathbb{P}(W)$, $\mathcal{E} = \mathcal{Q} \otimes \mathcal{O}(1)$, $\mathbf{s} = \Psi$ so that $Z(\mathbf{s}) = I$.

We assume that our two metrics $H_0^{\mathcal{E}}$ and $H_S^{\mathcal{E}}$ satisfy the following conditions[1]:

A1. $\dfrac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\Psi^*\hat{\mathbf{e}}(H_0^{\mathcal{E}}) \wedge \omega_{FS}^l = \delta_I \wedge \omega_{FS}^l \quad l+1 := \dim(W)$

A2. The function $S \ni s \mapsto \displaystyle\int_{\mathbb{X}_s \times \mathbb{P}(W)} \Psi^*\hat{\mathbf{e}}(H_S^{\mathcal{E}}) \wedge \omega_{FS}^l$ is pluriharmonic.

We need one more ingredient to state our next result. Let $d \in \mathbb{Z}_{>0}$. Recall that the Mahler measure is defined by

$$\Theta : |\mathcal{O}(d)| \longrightarrow \mathbb{R} \qquad \Theta([f]) := \int_{\mathbb{P}(W)} \log \frac{|f|^2_{h^d_{FS}}(\cdot)}{||f||_2^2} \omega_{FS}^l \,.$$

Tian has shown [11] that

Proposition 1.1. Θ *is Hölder continuous, in particular it is bounded.*

Finally define $\theta(s) := \Theta(\Delta(s))$, and let o be a basepoint in S such that $H_S^{\mathcal{E}}|_{\mathbb{X}_o} = H_0^{\mathcal{E}}|_{\mathbb{X}_o}$[2].

The following corollary of Theorem 1.1 extends several ideas of Gang Tian (especially [10], [12], and [11]) as well as the first author [9].

Corollary 1.1. *The function*

$$S \ni s \mapsto \log\left(e^{\theta(s)}\frac{||\Delta(s)||_2^2}{||\Delta(o)||_2^2}\right) - \int_{\mathbb{X}_s \times \mathbb{P}(W)} \mathbf{e}(H_0^{\mathcal{E}}, H_S^{\mathcal{E}}) \wedge \omega_{FS}^l \tag{1.2}$$

is pluriharmonic.

[1] Condition A2 comes from Lemma 4.1 on p. 278 of [9].
[2] In our case this comes from assuming that $h(o) = Id_W \in \mathcal{H}^+(W, \langle\cdot\ ,\ \cdot\rangle)$.

In all applications this pluriharmonic function is identically $\theta(o)$ (a constant).

To summarize, $\Delta(s)$ is for each $s \in S$ a polynomial on W and $\mathbf{e}(H_0^{\mathcal{E}}, H_S^{\mathcal{E}})$ is essentially a generalization of the Mabuchi K-energy map. For special choices of Q it *is* the Mabuchi energy up to lower-order terms. Our work suggests that most "action functionals" of interest on a polarized manifold are obtained through the basic set up and in particular have logarithmic singularities in the parameter $s \in S$. In applications one changes base to $S = G$ a reductive algebraic group (usually $SL(N+1, \mathbb{C})$). Then the right-hand side of 1.2 is exhibited as an action restricted to the usual space of Bergman potentials. Therefore, along algebraic one parameter subgroups $\lambda(t)$ of G the asymptotic expansion of the integral of the Bott–Chern form as $t \longrightarrow 0$ is determined by the limiting behavior of the *coefficients* of $\Delta(\lambda(t))$ as $t \longrightarrow 0$ [3].

Organization. Δ is a (homogeneous) polynomial with zero set $Z \subset \mathbb{P}(W)$ say. In the second section we discuss the classical setting of how such Z's arise. They are objects of elimination theory. In the third section we recall in detail an essential and rather ingenious idea of Cayley that allows us to *construct* Δ from Z. Usually one can easily detect whether or not Z has codimension one but it is very hard to find an explicit defining polynomial. This is the main problem of classical elimination theory. In the fourth section we use a vast generalization of Cayley's idea due to Grothendieck, Knudsen, and Mumford [8] which extends Cayley's construction from a *fixed* variety to the more natural setting of a *family* of varieties over some base S, this is where the Fourier–Mukai Transform makes an appearance. In the final section we prove Corollary 1.1 by introducing metrics on everything and invoking the currents of Bismut, Gillet, and Soulé on the one hand, and the Poincaré Lelong formula on the other.

2. Classical elimination theory

Let $X^n \hookrightarrow \mathbb{P}^N$ be a linearly normal subvariety. We consider a parameter space $\mathcal{F}$ of "linear sub-objects" f of $\mathbb{P}^N$. That is, $\mathcal{F}$ may parametrize points, lines, planes, flags, etc. of $\mathbb{P}^N$. Consider the admittedly vague statement

"*The general member f of $\mathcal{F}$ has a certain order of contact along X*" $\qquad (**)$

Let Z be the (proper) subvariety of those $f \in \mathcal{S}$ *violating* $(**)$.

Example 1 (Cayley–Chow Forms). We take $\mathcal{S}$ to be the Grassmannian of $N-n-1$-dimensional linear subspaces of $\mathbb{P}^N$

$$\mathcal{F} = \mathbb{G}(N-n-1, \mathbb{P}^N) := \{L \subset \mathbb{C}^{N+1} \mid \dim(L) = N-n-1\}\,.$$

In this case the general member $f = L$ of $\mathcal{S}$ *fails* to meet X. Therefore we take

$$Z = \{L \in \mathbb{G}(N-n-1, \mathbb{P}^N) \;|L \cap X \neq \emptyset\}\,.$$

[3]For a concrete example of this see Theorem B of [9].

Example 2. (*Dual Varieties*) Let $\mathcal{F} = \check{\mathbb{P}}^N$ the dual projective space parametrizing hyperplanes H inside $\mathbb{P}^N$. Bertini's theorem provides us with our "order of contact" condition. Namely, for generic $H \in \check{\mathbb{P}}^N$ $H \cap X$ is *smooth*. Therefore we define

$$Z = \{H \in \check{\mathbb{P}}^N \mid H \cap X \text{ is singular } \}.$$

Observe that

$$Z = \{H \in \check{\mathbb{P}}^N \mid \text{there exists } p \in X \text{ such that } \mathbb{T}_p(X) \subset H\}.$$

The next example includes both **1** and **2** as extreme cases.

Example 3. (Higher associated hypersurfaces, see [5] p. 104). Fix $L \subset \mathbb{C}^{N+1}$, $\dim(L) = n + 1 < N + 1$. Consider the subset $\mathcal{U}$ of the Grassmannian defined by

$$\mathcal{U} := \{E \in G(r;\ \mathbb{C}^{N+1}) |\ H^{\bullet} \left(0 \longrightarrow E \cap L \longrightarrow E \xrightarrow{\pi_L} \mathbb{C}^{N+1}/L \longrightarrow 0\right) = 0\}.$$

In order that $\mathcal{U}$ be open and dense in $G(r;\ \mathbb{C}^{N+1})$ it is enough to have

$$r = \dim(E) \geq N + 1 - (n + 1) = N - n.$$

Therefore we set $r = N - l$ where $0 \leq l \leq n$. By the rank plus nullity theorem of linear algebra we have

$$\dim(E \cap L) + \dim(\pi_L(E) = \dim(E) = N - l\,.$$

Therefore $E \in Z := G(r;\ \mathbb{C}^{N+1}) \setminus \mathcal{U}$ if and only if

$$\dim(E \cap L) \geq n - l + 1\,.$$

This motivates the following. Let $X^n \longrightarrow \mathbb{P}^N$. Fix $0 \leq l \leq n$. We define

$$Z := \{E \in \mathbb{G}(N - (l + 1), \mathbb{P}^N) \mid \exists\, p \in X$$

such that

$$p \in E \ \textit{and}\ \dim(E \cap \mathbb{T}_p(X)) \geq n - l\}\,.$$

In this situation we denote Z by $Z_{l+1}(X)$. The reader should observe that

$$Z_{n+1}(X) = \{R_X = 0\} \text{ and } Z_1(X) = \{\Delta_X = 0\}\,.$$

$Z_n(X)$ plays a fundamental role in our study of the K-energy (see [9]). In fact, we have that

$$Z_n(X) = \{\Delta_{X\times\mathbb{P}^{n-1}} = 0\}\ ,$$

the dual variety of the Segre image of $X \times \mathbb{P}^{n-1}$.

In the situations of interest to us we may assume that $\mathcal{F} \cong W$ for an appropriate finite-dimensional vector space W. For example in **1** we may replace $\mathbb{G}(N - n - 1, \mathbb{P}^N)$ with $W = M_{(n+1)\times(N+1)}(\mathbb{C})$.

In our applications Z has *codimension one*. Then $Z \subset W$ is an irreducible algebraic hypersurface defined by a single polynomial R_Z

$$Z = \{f \in W \mid R_Z(f) = 0\}\,.$$

Z is naturally dominated by the variety of zeros of a larger system $\{s_j(p,f)=0\}$ in more variables $p \in X$. We define the *incidence variety* I_X by

$$I_X := \{(p,\ f) \in X \times W \mid s_j(p\ ,\ f) = 0\} \subset X \times W.$$

In geometric terms $(p,f) \in I_X$ if and only if f fails to meet X generically at p (and possibly at some other point q). Therefore, Z is the *resultant system* obtained by eliminating the variable p

$R_Z(f) = 0$ iff $f \in Z$ iff $\{s_j(\cdot\ ,\ f) = 0\}$ has a solution p in X for the fixed $f \in W$.

3. Linear algebra of complexes and the torsion of a exact complex

To begin let $(E^\bullet, \partial^\bullet)$ be a bounded complex of finite-dimensional $\mathbb{C}$ vector spaces.

$$0 \longrightarrow E^0 \xrightarrow{\partial_0} \cdots \longrightarrow E^i \xrightarrow{\partial_i} E^{i+1} \xrightarrow{\partial_{i+1}}$$
$$\cdots \xrightarrow{\partial_n} E^{n+1} \longrightarrow 0\,.$$

Recall that the *determinant of the complex* $(E^\bullet, \partial^\bullet)$ is defined to be the one-dimensional vector space

$$\mathbf{Det}(E^\bullet)^{(-1)^n} := \bigotimes_{i=0}^{n+1} \left(\bigwedge^{r_i} E^i\right)^{(-1)^{i+1}} \qquad r_i := \dim(E^i)\,.$$

Remark 1. $\mathbf{Det}(E^\bullet)$ does not depend the boundary operators.

As usual, for any vector space V we set $V^{-1} := \mathrm{Hom}_{\mathbb{C}}(V, \mathbb{C})$, the dual space to V. Let $H^i(E^\bullet, \partial^\bullet)$ denote the ith cohomology group of this complex. When $V = \mathbf{0}$, the zero vector space, we set $\mathbf{det}(V) := \mathbb{C}$. The determinant of the cohomology is defined in exactly the same way

$$\mathbf{Det}(H^\bullet(E^\bullet, \partial^\bullet))^{(-1)^n} := \bigotimes_{i=0}^{n+1} \left(\bigwedge^{b_i} H^i\left(E^\bullet, \partial^\bullet\right)\right)^{(-1)^{i+1}} \qquad b_i := \dim(H^i(E^\bullet, \partial^\bullet))\,.$$

We have the following well-known facts ([8], [2]).

D1 *Assume that the complex* $(E^\bullet, \partial^\bullet)$ *is acyclic, then* $\mathbf{Det}(E^\bullet)$ *is canonically trivial*

$$\tau(\partial^\bullet) : \mathbf{Det}(E^\bullet) \cong \underline{\mathbb{C}}\,.$$

As a corollary of this we have,

D2 *There is a canonical isomorphism*[4] *between the determinant of the complex and the determinant of its cohomology*:

$$\tau(\partial^\bullet) : \mathbf{Det}(E^\bullet) \cong \mathbf{Det}(H^\bullet(E^\bullet, \partial^\bullet))\,.$$

---

[4] A "canonical isomorphism" is one that only depends on the boundary operators, not on any choice of basis.

It is **D1** which is relevant for our purpose. It says is that *there is a canonically given* **nonzero** *element of* $\mathbf{Det}(E^\bullet)$, called the *torsion* of the complex, provided this complex is exact. The torsion is the essential ingredient in the construction of X-resultants (Cayley–Chow forms) and X-discriminants (dual varieties). We recall the construction for more information see [2].

Define $\kappa_i := \dim(\partial_i E^i)$, now choose $S_i \in \wedge^{\kappa_i}(E^i)$ with $\partial_i S_i \neq 0$, then $\partial_i S_i \wedge S_{i+1}$ spans $\bigwedge^{r_{i+1}} E^{i+1}$ (since the complex is exact), that is

$$\bigwedge^{r_{i+1}} E^{i+1} = \mathbb{C}\partial_i S_i \wedge S_{i+1}.$$

With this said we define[5]

$$\mathbf{Tor}\,(E^\bullet, \partial^\bullet)^{(-1)^n} := (S_0)^{-1} \otimes (\partial_0 S_0 \wedge S_1) \otimes (\partial_1 S_1 \wedge S_2)^{-1} \otimes \cdots \otimes (\partial_n S_n)^{(-1)^n}.$$

Then we have the following reformulation of **D1**.

D3
$$\mathbf{Tor}\,(E^\bullet, \partial^\bullet) \tag{3.1}$$

is independent of the choices S_i.

By fixing a basis $\{e^i_1, e^i_2, \ldots, e^i_{r_i}\}$ in each of the terms E_i we may associate to this *based* exact complex a *scalar*.

$$\mathbf{Tor}\left(E^\bullet, \partial^\bullet; \{e^i_1, e^i_2, \ldots, e^i_{r_i}\}\right) \in \mathbb{C}^*.$$

Which is defined through the identity:

$$\mathbf{Tor}\,(E^\bullet, \partial^\bullet) = \mathbf{Tor}\left(E^\bullet, \partial^\bullet; \{e^i_1, e^i_2, \ldots, e^i_{r_i}\}\right) \mathbf{det}(\ldots e^i_1, e^i_2, \ldots, e^i_{r_i}, \ldots).$$

Where we have set

$$\begin{aligned}&\mathbf{det}(\ldots e^i_1, e^i_2, \ldots, e^i_{r_i} \ldots)^{(-1)^n}\\ &\quad := (e^0_1 \wedge \cdots \wedge e^0_{r_0})^{-1} \otimes \cdots \otimes (e^{n+1}_1 \wedge \cdots \wedge e^{n+1}_{r_{n+1}})^{(-1)^n}.\end{aligned}$$

When we have fixed a basis of our exact complex (that is, a basis of each term in the complex) we will call $\mathbf{Tor}\left(E^\bullet, \partial^\bullet; \{e^i_1, e^i_2, \ldots, e^i_{r_i}\}\right)$ the Torsion of the *based exact* complex. It is, as we have said, a *scalar quantity*.

Remark 2. In the following sections we often base the complex without mentioning it explicitly and in such cases we write (incorrectly) $\mathbf{Tor}\,(E^\bullet, \partial^\bullet)$ instead of

$$\mathbf{Tor}\left(E^\bullet, \partial^\bullet; \{e^i_1, e^i_2, \ldots, e^i_{r_i}\}\right).$$

We have the following well-known *scaling behavior* of the Torsion, which we state in the next proposition. Since it is so important for us, we provide the proof, which is yet another application of the rank plus nullity theorem.

Proposition 3.1 (The degree of the Torsion as a polynomial in the boundary maps).

$$\deg\mathbf{Tor}\,(E^\bullet, \partial^\bullet) = (-1)^{n+1}\sum_{i=0}^{n+1}(-1)^i i \dim(E^i)\,. \tag{3.2}$$

[5] The purpose of the exponent $(-1)^n$ will be revealed in the next section.

Proof. Let $\mu \in \mathbb{C}^*$ be a parameter. Then

$$\begin{aligned}\mathbf{Tor}\,(E^\bullet, \mu\partial^\bullet)^{(-1)^n} &= (S_0)^{-1} \otimes (\mu\partial_0 S_0 \wedge S_1) \otimes (\mu\partial_1 S_1 \wedge S_2)^{-1} \otimes \cdots \\ &\quad \cdots \otimes (\mu\partial_n S_n)^{(-1)^n} \\ &= \mu^{\kappa_0 - \kappa_1 + \kappa_2 - \cdots + (-1)^n \kappa_n}\,\mathbf{Tor}\,(E^\bullet, \partial^\bullet)^{(-1)^n} \,.\end{aligned}$$

It is clear that

$$\kappa_0 - \kappa_1 + \kappa_2 - \cdots + (-1)^n \kappa_n = \sum_{i=0}^{n+1} (-1)^{i+1} i(\kappa_i + \kappa_{i-1}) \quad (\kappa_{n+1} = \kappa_{-1} := 0)\,.$$

Exactness of the complex implies that we have the short exact sequence

$$0 \longrightarrow \partial_{i-1} E^{i-1} \xrightarrow{\ \iota\ } E^i \xrightarrow{\ \partial_i\ } \partial_i E^i \longrightarrow 0\,.$$

Therefore $\kappa_i + \kappa_{i-1} = \dim(E_i)$. □

4. Fourier–Mukai transforms and the geometric technique

In this section we prove Theorem 1.1. The method of proof is follows the *Geometric Technique*. The author's understanding is that the method is due to many mathematicians (see [3, 5–8]). The author has also learned a great deal about the method from the monograph of J. Weyman [13], especially the "basic set-up" of Chapter 5.

Let X a complex variety. Let $(\mathcal{E}^\bullet\ ;\ \delta^\bullet)$ be an exact (bounded) complex of locally free sheaves over X. The discussion in the previous section implies the following

Proposition 4.1. *The determinant line bundle of the complex admits a canonical nowhere vanishing section* Δ

$$\mathbf{Det}(\mathcal{E}^\bullet\ ;\ \delta^\bullet) \overset{\Delta}{\cong} \mathbb{C}\,.$$

We return to the situation described in the introduction. The object of study is a smooth, linearly normal, family $\mathbb{X} \longrightarrow B$ of relative dimension n together with a rank $n+1$ locally free sheaf $\mathcal{Q}$ satisfying the requirements of a basic set up. Recall the visualization in (1.1), p. 288.

The basic set up exchanges a high codimension family $\mathbb{X} \longrightarrow B$ for a *family of divisors* $Z_b \subset \mathbb{P}(W)$ parametrized by the same base

$$Z_b := Z \cap (\{b\} \times \mathbb{P}(W))\ ,\ b \in B\,.$$

Recall that each Z_b is irreducible and moreover that the degree of Z_b in $\mathbb{P}(W)$ is given by

$$\deg(Z_b) = \int_{\mathbb{X}_b} c_n(\mathcal{Q})\ ,$$

and therefore constant in $b \in B$.

To facilitate the study of Z we pass to the birationally equivalent I which is much easier to deal with. More precisely we study the direct image of the structure sheaf of I viewed as a coherent sheaf on $\mathbb{X}\times\mathbb{P}(W)$.

Recall that the Koszul complex $(K^\bullet\mathcal{E}\ ;\ \partial^\bullet :=\lrcorner\, \mathbf{s})$ associated to $(\mathcal{E},\mathbf{s})$

$$\xrightarrow{\lrcorner\,\mathbf{s}} \Lambda^{j+1}(\mathcal{E}^\vee) \xrightarrow{\lrcorner\,\mathbf{s}} \Lambda^{j}(\mathcal{E}^\vee) \xrightarrow{\lrcorner\,\mathbf{s}} \cdots \xrightarrow{\lrcorner\,\mathbf{s}} \mathcal{O}_{\mathbb{X}\times\mathbb{P}(W)}$$

resolves $\iota_*(\mathcal{O}_I)$. The main player in this paper is the following Fourier–Mukai transform between the bounded derived categories[6]

$$\Phi^{K^\bullet\mathcal{E}}_{\mathbb{X}\mapsto B\times\mathbb{P}(W)}(\cdot) : D^b(\mathbb{X}) \longrightarrow D^b(B\times\mathbb{P}(W))$$

associated to the projections

$$\begin{tikzcd} & \mathbb{X}\times\mathbb{P}(W) \arrow[dl, "p"'] \arrow[dr, "q"] & \\ \mathbb{X} & & B\times\mathbb{P}(W)\,. \end{tikzcd}$$

Recall that this transform is defined by the formula

$$\Phi^{K^\bullet\mathcal{E}}_{\mathbb{X}\mapsto B\times\mathbb{P}(W)}(F^\bullet) := \mathbf{R}^\bullet q_*\left(\mathbf{L}^\bullet p^*(F^\bullet)\otimes K^\bullet\mathcal{E}\right)$$

where $\mathbf{R}^\bullet q_*$ and $\mathbf{L}^\bullet p^*$ denote the usual derived functors.

We remind the reader that by Grauert's Theorem of coherence of higher direct images that the complex

$$\Phi^{K^\bullet\mathcal{E}}_{\mathbb{X}\mapsto B\times\mathbb{P}(W)}(F^\bullet)$$

is represented by a bounded complex of (quasi) coherent $\mathcal{O}_{B\times\mathbb{P}(W)}$-modules with coherent cohomology sheaves.

We are interested in the value of this transform at a particular point of $D^b(\mathbb{X})$

$$\Phi^{K^\bullet\mathcal{E}}_{\mathbb{X}\mapsto B\times\mathbb{P}(W)}(\mathcal{O}_{\mathbb{X}}(m)) \in D^b\left(B\times\mathbb{P}(W)\right) \quad m >> 0\,.$$

Since the (twisted) Koszul complex resolves $\iota_*\mathcal{O}_I\otimes p^*\mathcal{O}_{\mathbb{X}}(m)$ we have the following

Proposition 4.2. *The cohomology sheaves of the complex* $\Phi^{K^\bullet\mathcal{E}}_{\mathbb{X}\mapsto B\times\mathbb{P}(W)}(\mathcal{O}_{\mathbb{X}}(m))$ *are all supported on* Z.

We assume that m is large enough to force term wise vanishing of all higher direct image sheaves

$$\mathbf{R}^i q_*\left(\Lambda^j(\mathcal{E}^\vee)\otimes p^*\mathcal{O}_{\mathbb{X}}(m)\right) = 0 \qquad i>0 \ \text{ all } j\,.$$

This has the crucial implication that the natural map

$$q_*\left(\Lambda^\bullet\mathcal{E}^\vee\otimes p^*\mathcal{O}_{\mathbb{X}}(m)\ ;\ \delta^\bullet\right) \longrightarrow \Phi^{K^\bullet\mathcal{E}}_{\mathbb{X}\mapsto B\times\mathbb{P}(W)}(\mathcal{O}_{\mathbb{X}}(m))$$

[6]In this paper $D^b(X)$ is just notation for the set of bounded complexes of coherent sheaves on X.

is the identity in $D^b(B \times \mathbb{P}(W))$, in other words, the two complexes are quasi-isomorphic. This is an application of the "Cartan–Eilenberg spectral sequence"

$$E_2^{ij} := \mathbf{R}^i q_* \left(\Lambda^j \mathcal{E}^\vee(m)\right) \implies \mathbf{R}^{i+j} q_* \left(\Lambda^\bullet \mathcal{E}^\vee(m)\right) .$$

Therefore we may replace the a priori quite complicated object

$$\Phi^{K^\bullet \mathcal{E}}_{\mathbb{X} \mapsto B \times \mathbb{P}(W)} \left(\mathcal{O}_{\mathbb{X}}(m)\right)$$

with the much simpler (termwise) direct image complex.

The purpose of the foregoing discussion was to put us in the following situation

Proposition 4.3. *The complex of locally free sheaves over $B \times \mathbb{P}(W)$*

$$q_* \left(\Lambda^\bullet \mathcal{E}^\vee \otimes p^* \mathcal{O}_{\mathbb{X}}(m) \ ; \ \partial^\bullet\right)$$

is exact away from Z.

Therefore away from Z the determinant of the direct image complex has a nowhere vanishing section Δ

$$\begin{tikzcd} \mathbf{Det}\ q_* \left(\Lambda^\bullet \mathcal{E}^\vee \otimes p^* \mathcal{O}_{\mathbb{X}}(m) \ ; \ \partial^\bullet\right) \arrow[d] \\ B \times \mathbb{P}(W) \setminus Z \arrow[u, bend left, "\Delta"] \end{tikzcd}$$

$$\Delta := \mathbf{Tor}\ q_* \left(\Lambda^\bullet \mathcal{E}^\vee \otimes p^* \mathcal{O}_{\mathbb{X}}(m) \ ; \ \partial^\bullet\right) .$$

Proposition 4.4. *There is an invertible sheaf $\mathcal{A}$ over B such that*

$$\mathbf{Det}\ q_* \left(\Lambda^\bullet \mathcal{E}^\vee \otimes p^* \mathcal{O}_{\mathbb{X}}(m) \ ; \ \partial^\bullet\right) \cong p_1^* \mathcal{A} \otimes p_2^* \mathcal{O}_W(\widehat{d}) .$$

Proof. The proof is quite easy. Observe that by the projection formula, the stalk of the direct image of our Koszul complex at $b \in B$ is given by

$$q_* \left(\Lambda^j \mathcal{E}^\vee \otimes p^* \mathcal{O}_{\mathbb{X}}(m)\right)|_b \cong H^0\left(\mathbb{X}_b \ , \ \Lambda^j Q^\vee(m)\right) \otimes \mathcal{O}_W(-j) .$$

We define

$$r_j(m) := \dim H^0\left(\mathbb{X}_b \ , \ \Lambda^j Q^\vee(m)\right) .$$

Then we define $\mathcal{A}$ as follows

$$\mathcal{A}|_b := \bigotimes_{0 \le j \le n+1} \Lambda^{r_j(m)} H^0\left(\mathbb{X}_b \ , \ \Lambda^j Q^\vee(m)\right)^{(-1)^{j+1}} .$$

Therefore the determinant is given by

$$\mathbf{Det}\ q_* \left(\Lambda^\bullet \mathcal{E}^\vee \otimes p^* \mathcal{O}_{\mathbb{X}}(m) \ ; \ \partial^\bullet\right) \cong p_1^* \mathcal{A} \otimes \mathcal{O}_W(\chi) ,$$

where we have defined $\chi \in \mathbb{Z}$ by

$$\chi := \sum_{0 \le j \le n+1} (-1)^{j+1} j r_j(m) \in \mathbb{Z} .$$

So the argument comes down to showing that $\chi = \widehat{d}$ which a straightforward but tedious calculation with the Hirzebruch–Riemann–Roch Theorem will verify. $\square$

Fix $b \in B$. Let f_b denote a defining polynomial of Z_b. Since $\Delta|_{\{b\}\times\mathbb{P}(W)}$ is without zeros or poles away from Z_b there is an integer

$$\mathrm{ord}_{Z_b}(\Delta|_{\{b\}\times\mathbb{P}(W)})$$

satisfying

$$\Delta|_{\{b\}\times\mathbb{P}(W)} = f_b^{\mathrm{ord}_{Z_b}(\Delta|_{\{b\}\times\mathbb{P}(W)})}\,.$$

Our computation of the degree of the torsion shows that

$$\deg(\Delta|_{\{b\}\times\mathbb{P}(W)}) = \chi\,.$$

Therefore $\mathrm{ord}_{Z_b}(\Delta|_{\{b\}\times\mathbb{P}(W)}) = 1$, and we have shown

$$\mathbb{C}^*\Delta|_{\{b\}\times\mathbb{P}(W)} = \mathbb{C}^* f_b\,.$$

Therefore we have the following

Proposition 4.5. *Δ vanishes on Z and in particular extends to a global section of the determinant line.*

Therefore, to any basic set up for the family $\mathbb{X} \longrightarrow B$ we may associate the following

- An invertible sheaf $\mathcal{A} \in \mathrm{Pic}(B)$.
- An algebraic section
$$\Delta \ \in H^0\Big(B \times \mathbb{P}(W)\ ,\ p_1^*\mathcal{A} \otimes p_2^*\mathcal{O}_W(\widehat{d})\Big)\ .$$
- A relative Cartier divisor Z over B
$$Z := \mathrm{Div}(\Delta) \subset B \times \mathbb{P}(W)\,.$$
- Moreover, the direct image T of Δ
$$\begin{tikzcd} \mathcal{A} \otimes H^0(\mathbb{P}(W)\,, \mathcal{O}_W(\widehat{d})) \arrow[d] \\ B \arrow[u, bend left, "T:={p_1}_*(\Delta)"] \end{tikzcd}$$
never vanishes on B.

This package induces a morphism (also denoted by Δ) from B to the complete linear system on $\mathbb{P}(W)$

$$\Delta : B \longrightarrow |\mathcal{O}_W(\widehat{d})|$$

as follows. Since $T \neq 0$, T gives an injection

$$0 \longrightarrow \mathcal{O}_B \xrightarrow{\times T} \mathcal{A} \otimes H^0(\mathbb{P}(W), \mathcal{O}(\widehat{d}))\,.$$

Dualizing and tensoring with $\mathcal{A}$ gives a surjection

$$H^0(\mathbb{P}(W), \mathcal{O}(\widehat{d}))^\vee \times B \longrightarrow \mathcal{A} \longrightarrow 0\ ,$$

hence $\mathcal{A}$ is globally generated.

Therefore we obtain a morphism as required. This completes the proof of Theorem 1.1. Moreover we see that

$$\Delta^*\mathcal{O}_{|\mathcal{O}(\widehat{d})|}(1) \cong \mathcal{A}\ ,$$

and the natural map

$$H^0(\mathbb{P}(W), \mathcal{O}(\widehat{d}))^\vee \longrightarrow H^0(B, \mathcal{A})$$

is an injection. The map exhibits a large (generating) finite-dimensional subspace of the space of sections of $\mathcal{A}$ over B. Conversely, given such a map Δ, we define

$$\mathcal{A} := \Delta^*\mathcal{O}_{|\mathcal{O}(\widehat{d})|}(1)$$

$$\Delta(b, [w]) := \Delta(b)([w]) \in \mathcal{A}_b \otimes \mathcal{O}_W(\widehat{d})\,.$$

Next we give several examples of basic set ups for a given family $\mathbb{X} \longrightarrow B$.

4.1. The basic set up for resultants

Let $\mathbb{X} \to B$ be a flat family of polarized subvarieties of $\mathbb{P}^N$. We can arrange a basic set up

$$\begin{tikzcd}
 & \mathcal{E} \arrow[r] \arrow[d] & p_1^*\mathcal{O}_{\mathbb{P}^N}(1)\otimes p_2^*\mathcal{Q} \arrow[d] \\
I \arrow[r, hook, "\iota"] \arrow[d] & \mathbb{X}\times\mathbb{G} \arrow[u, bend left, "\Psi"] \arrow[r, "\pi_2\times 1"] \arrow[d, "\pi"] & \mathbb{P}^N\times\mathbb{G} \arrow[u, bend left, "\Pi"] \\
Z \arrow[r, hook] & B\times\mathbb{G} \arrow[d, "p_1"] & \\
 & B &
\end{tikzcd}$$

for this family if we define

- $\Pi|_{([v],L)} : \mathbb{C}v \longrightarrow \mathbb{C}^{N+1}/L\ ,\qquad \Pi(zv) = \pi_L(zv)$
  $\pi_L : \mathbb{C}^{N+1} \longrightarrow \mathbb{C}^{N+1}/L$ denotes the projection.
  $\Psi := (\pi_2 \times 1)^*\Pi\,.$
- $\mathbb{G} := \mathbb{G}(N-n-1, N)\,.$
- $\mathcal{E} := (\pi_2\times 1)^*\left(p_1^*\mathcal{O}_{\mathbb{P}^N}(1)\otimes p_2^*\mathcal{Q}\right)$
  $I = (\Psi = 0)\,.$
- $\pi : \mathbb{X}\times\mathbb{G} \longrightarrow B\times\mathbb{G}$ is defined by[7]Abuse of notation. $\pi(x, L) := (\pi(x), L)$
  $Z := \pi(I)\,.$
- Z_b is the **Cayley form** of X_b and Δ_b is the X_b-**resultant**.

[7]Abuse of notation.

4.2. The basic set up for discriminants

In this section we consider a flat family $\mathbb{X} \xrightarrow{\pi} B$ of polarized manifolds. The basic set up that we consider in this case has the form

$$
\begin{tikzcd}
 & \mathcal{E} \arrow[r] \arrow[d] & p_1^*\mathscr{U}^\vee \otimes p_2^*\mathcal{O}_{\check{\mathbb{P}}^N}(1) \arrow[d] \\
\Gamma_{\mathbb{X}} \arrow[r, hook, "\iota"] \arrow[d] & \mathbb{X} \times \check{\mathbb{P}}^N \arrow[r, "\rho\times 1"] \arrow[u, bend left, "\mathbf{s}"] \arrow[d, "p"] & \mathbb{G}(n,N) \times \check{\mathbb{P}}^N \arrow[u, bend left, "\Lambda"] \\
Z \arrow[r, hook] & B \times \check{\mathbb{P}}^N \arrow[d, "p_1"] & \\
 & B &
\end{tikzcd}
$$

In the diagram above we have defined

- $\Lambda|_{(L,[f])} : L \longrightarrow \mathbb{C}^{N+1}/\ker(f)\ , \qquad \Lambda(u) = \pi_{\ker(f)}(u)$
  $\pi_{\ker(f)} : \mathbb{C}^{N+1} \longrightarrow \mathbb{C}^{N+1}/\ker(f)$ denotes the projection.
  Observe that $\Lambda|_{(L,[f])} = 0$ if and only if $L \subset \ker(f)$.
- $\rho = \rho_{\mathbb{X}} : \mathbb{X} \longrightarrow \mathbb{G}(n,N)$ is the fiber wise Gauss map.
  $\mathscr{U}$ is the tautological bundle.
- $\mathcal{E} := (\rho \times 1)^* \left(p_1^*\mathscr{U}^\vee \otimes p_2^*\mathcal{O}_{\check{\mathbb{P}}^N}(1)\right)$
  $\Gamma_{\mathbb{X}} := (\mathbf{s} = 0)\ ;\ Z := q(\Gamma_{\mathbb{X}})$.
- Z_b is the **dual variety** of X_b and Δ_b is the X_b-**discriminant**.

Remark 3. An interesting generalization of these two examples is constructed as follows. Given a family $\mathbb{X} \longrightarrow B$ let $\mathscr{E}_k$ denote the rank $n-k+1$ trivial bundle over $\mathbb{X}$. Then we obtain a new family

$$\mathbb{P}(\mathscr{E}_k) \longrightarrow B\,.$$

The fiberwise Segre embedding exhibits this family as a family of subvarieties of the projective space of matrices of size $(N+1)\times(n-k+1)$. If the original family is smooth one may apply the set up for discriminants to this new family. When $k=1$ the corresponding polynomial Δ_b is the X_b-**hyperdiscriminant**.

5. Comparing the currents δ_Z and δ_I over S

Let $S \longrightarrow B$ be a morphism from a *smooth*[8] variety S. As stated in the introduction we assume

$$\mathcal{A} \cong \mathcal{O}_S$$

[8]Smoothness is required to apply the Poincaré Lelong formula.

and there exists a smooth map

$$h : S \longrightarrow \mathcal{H}^+(W, \langle \cdot\ ,\ \cdot \rangle)$$

satisfying conditions A_1 and A_2. There is an induced Hermitean metric on h on the determinant line bundle

$$p_1^*\mathcal{A} \otimes p_2^*\mathcal{O}_W(\widehat{d})$$

and it is not hard to see that the square of the length of our section Δ is given by

$$\frac{|\Delta(s)([w])|^2_{h^{\widehat{d}}}}{||\Delta(s)||_2^2}$$

The denominator being the usual L^2 norm

$$||\Delta(s)||_2^2 := \int_{\mathbb{P}(W)} |\Delta(s)(\cdot)|^2_{h^{\widehat{d}}}\ \omega_{FS}^l \quad ;\ l+1 = \dim(W)\,.$$

The Poincaré Lelong formula gives

$$\frac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\log\frac{|\Delta(b)([w])|^2_{h^{\widehat{d}}}}{||\Delta(b)||_2^2} = \delta_Z - c_1\left(\mathcal{A}\otimes\mathcal{O}_W(\widehat{d})\ ;\ h\right).$$

Recall that the triviality of $\mathcal{A}$ over S is equivalent to having a lift of[9] the map Δ to the affine cone

$$\Delta : S \longrightarrow H^0(\mathbb{P}(W), \mathcal{O}(\widehat{d}))\setminus\{0\}\,.$$

Next fix some base point $o \in S$. Then

$$c_1\left(\mathcal{A}\otimes\mathcal{O}_W(\widehat{d})\ ;\ h\right)|_{S\times\mathbb{P}(W)} = \widehat{d}\omega_{FS} + \frac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\log\frac{||\Delta(s)||_2^2}{||\Delta(o)||_2^2}\,.$$

With this said we have the following proposition concerning the direct image of this current under π.

Proposition 5.1. *Let η be a smooth compactly supported form on S*

$$\eta \in C^\infty\Lambda_0^{\dim(S)-1,\dim(S)-1}(S).$$

Then

$$\int_Z \pi^*(\eta)\wedge\omega_{FS}^l = \int_S \eta\wedge\frac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\log\left(e^{\theta(s)}\frac{||\Delta(s)||_2^2}{||\Delta(o)||_2^2}\right)$$

where θ is defined by

$$\theta(s) := \int_{\mathbb{P}(W)} \log\frac{|\Delta(s)([w])|^2_{h^{\widehat{d}}}}{||\Delta(s)||_2^2}\omega_{FS}^l\,.$$

[9] We will also denote the lifted map by Δ as well.

The birationality of I and Z imply that we have the identity

$$\int_Z \pi^*(\eta)\wedge\omega_{FS}^l = \int_I \pi^*(\eta)\wedge\omega_{FS}^l .$$

for all compactly supported forms η on S. Therefore, Proposition 5.1 and property A1 of the current $\widehat{\mathbf{e}}(H_0^{\mathcal{E}})$ imply that

$$\int_Z \pi^*(\eta)\wedge\omega_{FS}^l = \int_{\mathbb{X}\times\mathbb{P}(W)} \frac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\Psi^*\widehat{\mathbf{e}}(H_0^{\mathcal{E}})\wedge\omega^l\wedge\pi^*(\eta)\,.$$

Property A2 and the variation formula show that

$$\begin{aligned}\int_{\mathbb{X}_s\times\mathbb{P}(W)} \Psi^*\widehat{\mathbf{e}}(H_0^{\mathcal{E}})\wedge\omega^l &= \int_{\mathbb{X}_s\times\mathbb{P}(W)} \mathbf{e}(H_0^{\mathcal{E}};H_S^{\mathcal{E}})\wedge\omega^l \\ &\quad + \text{a pluriharmonic function of } s\,.\end{aligned}$$

Therefore we see that for all compactly supported forms η we have

$$\int_S \eta\wedge\frac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\int_{\mathbb{X}_s\times\mathbb{P}(W)} \widehat{\mathbf{e}}(H_0^{\mathcal{E}};H_S^{\mathcal{E}})\wedge\omega^l = \int_S \eta\wedge\frac{\sqrt{-1}}{2\pi}\partial\overline{\partial}\log\left(e^{\theta(s)}\frac{||\Delta(s)||_2^2}{||\Delta(o)||_2^2}\right)\,.$$

Therefore we have proved Corollary 1.1.

Corollary 1.1 (see the "main lemma" from [9]). *The function*

$$S\ni s\mapsto \log\left(e^{\theta(s)}\frac{||\Delta(s)||_2^2}{||\Delta(o)||_2^2}\right) - \int_{\mathbb{X}_s\times\mathbb{P}(W)} \mathbf{e}(H_0^{\mathcal{E}};H_S^{\mathcal{E}})\wedge\omega^l$$

is pluriharmonic.

Given a fixed projective variety $X\subset\mathbb{P}^N$ with a non-degenerate dual variety $\check{X}$ we construct, following [11], a "tautological family" $\mathbb{X}\longrightarrow G$ where $G=SL(N+1,\mathbb{C})$ with fiber σX. Then, for each $0\le k\le\dim(X)$, we consider the new family $\mathbb{P}(\mathscr{E}_k)\longrightarrow G$. Let $\mathcal{Q}$ denote the vector bundle associated to the set up for discriminants for this family. Then for the obvious choice of Hermitean inner product on W, there is a natural map

$$h: G\longrightarrow \mathcal{H}^+(W,\langle\cdot,\cdot\rangle)$$

satisfying A1 and A2 with basepoint $o=e$. Therefore the previous theorem applies. In these cases we can show[10] that the function is actually *constant.*

References

[1] Gilllet H. Bismut, J.M. and Soulé C. Complex immersions and Arakelov geometry. *Grothendieck Festschrift*, 76(1):250–326, 2007.

[2] J.-M. Bismut, H. Gillet, and C. Soulé. Analytic torsion and holomorphic determinant bundles. I. Bott–Chern forms and analytic torsion. *Comm. Math. Phys.*, 115(1):49–78, 1988.

[10] This is forthcoming work of the first author and his student Q. Westrich.

[3] Arthur Cayley. On the theory of elimination. *Cambridge and Dublin Math Journal*, 3, 1848.

[4] Wei Yue Ding and Gang Tian. Kähler–Einstein metrics and the generalized Futaki invariant. *Invent. Math.*, 110(2):315–335, 1992.

[5] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. *Discriminants, resultants, and multidimensional determinants*. Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, 1994.

[6] M.M. Kapranov, B. Sturmfels, and A.V. Zelevinsky. Chow polytopes and general resultants. *Duke Math. J.*, 67(1):189–218, 1992.

[7] George R. Kempf. On the collapsing of homogeneous bundles. *Invent. Math.*, 37(3):229–239, 1976.

[8] Finn Faye Knudsen and David Mumford. The projectivity of the moduli space of stable curves I: Preliminaries on “det” and “Div”. *Math. Scand.*, 39(1):19–55, 1976.

[9] Sean Timothy Paul. Hyperdiscriminant polytopes, Chow polytopes, and Mabuchi energy asymptotics. *Annals of Math.*, (175), 2012.

[10] Gang Tian. The K-energy on hypersurfaces and stability. *Comm. Anal. Geom.*, 2(2):239–265, 1994.

[11] Gang Tian. Kähler–Einstein metrics with positive scalar curvature. *Invent. Math.*, 130(1):1–37, 1997.

[12] Gang Tian. Bott–Chern forms and geometric stability. *Discrete Contin. Dynam. Systems*, 6(1):211–220, 2000.

[13] Jerzy Weyman. *Cohomology of vector bundles and syzygies*, volume 149 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2003.

Sean Timothy Paul and Kyriakos Sergiou
Mathematics Department
University of Wisconsin
Madison, WI 53706, USA
e-mail: `stpaul@wisc.edu`
`sergiou@wisc.edu`

Progress in Mathematics, Vol. 333, 305–379

The Variations of Yang–Mills Lagrangian

Tristan Rivière

Dedicated to Gang Tian on his sixtieth birthday

Abstract. We are giving a survey on some of the analysis methods from gauge theory developed in the last decades. We first cover Uhlenbeck's compensated compactness theory in critical 4 dimension for the Yang–Mills functional. As an application we present the resolution of minimization processes of Yang–Mills in this critical dimension. In the second part of the survey we present the resolution of similar variational questions in super-critical dimensions and we end up the survey by stating some open problems raised by Tian relative to the regularity of Ω-anti-self-dual instantons in high dimensions.

Mathematics Subject Classification (2010). 58E15, 35J25, 35J47, 35J50.

Keywords. Yang–Mills, gauge theory, concentration compactness, critical elliptic PDE, conformally invariant variational problems, instantons, calibrated currents, calibrated geometric analysis.

I. Introduction

Yang–Mills theory is growing at the interface between high energy physics and mathematics It is well known that Yang–Mills theory and gauge theory in general had a profound impact on the development of modern differential and algebraic geometry. One could quote Donaldson invariants in four-dimensional differential topology, Hitchin Kobayashi conjecture relating the existence of Hermitian–Einstein metric on holomorphic bundles over Kähler manifolds and Mumford stability in complex geometry or also Gromov–Witten invariants in symplectic geometry... etc. While the influence of gauge theory in geometry is quite notorious, one tends sometimes to forget that Yang–Mills theory has been also at the heart of fundamental progresses in the non-linear analysis of Partial Differential Equations in the last decades. The purpose of this survey is to present the variations of this important Lagrangian. We shall raise analysis question such as existence and regularity of Yang–Mills minimizers or such as the compactification of the "moduli space" of critical points to Yang–Mills Lagrangian in general.

II. The Plateau problem

Before to move to the Yang–Mills minimization problem we will first recall some fundamental facts regarding the minimization of the area in the parametric approach and some elements of the resolution of the so-called *Plateau problem.*

Let Γ be a simple closed Jordan Curve in $\mathbb{R}^3$: there exists $\gamma \in C^0(S^1, \mathbb{R}^3)$ such that $\Gamma = \gamma(S^1)$.

Plateau problem: *Find a C^1 immersion u of the two-dimensional disc D^2 which is continuous up to the boundary, whose restriction to ∂D^2 is a homeomorphism and which minimizes the area*

$$\mathrm{Area}(u) = \int_{D^2} |\partial_{x_1} u \times \partial_{x_2} u| \; dx_1 \, dx_2 \, .$$

The area is a fairly degenerated functional:

i) It has a huge invariance group: $\mathrm{Diff}(D^2)$, the group of diffeomorphism of the disc. Let u_n be a minimizing sequence of the *Plateau problem* above, then the composition of u_n with any sequence of diffeomorphism Ψ_n of D^2 is still a minimizing sequence. The sequence Ψ_n can for instance degenerate so that $u_n \circ \Psi_n$ converges to a point!

ii) The area of u does not control the image $u(D^2)$ which could be uniformly bounded while $u(D^2)$ becomes dense in $\mathbb{R}^3$!

In order to solve the Plateau problem J. Douglas and independently Radó minimize instead the *Dirichlet energy*

$$\mathrm{Area}(u) \le E(u) = \frac{1}{2} \int_{D^2} [|\partial_{x_1} u|^2 + |\partial_{x_2} u|^2] \; dx_1 \, dx_2 \, ,$$

the inequality comes from the pointwise inequality

$$2 \, |\partial_{x_1} u \times \partial_{x_2} u| \le |\partial_{x_1} u|^2 + |\partial_{x_2} u|^2 \, ,$$

and equality holds if and only if

$$H(u) := |\partial_{x_1} u|^2 - |\partial_{x_2} u|^2 - 2 \, i \; \partial_{x_1} u \cdot \partial_{x_2} u = 0 \, .$$

This condition is satisfied if and only if the differential du preserves angle that is to say u is *conformal.*

The *Dirichlet energy* E has much better properties than the area: it is

i) it is *coercive*:

$$\forall \, u_n \quad \text{s.t.} \quad \limsup_{n \to +\infty} E(u_n) < +\infty \quad \exists \, u_{n'} \rightharpoonup u_\infty \quad \text{in} \quad W^{1,2} \, ,$$

ii) *lower semi-continuous*

$$E(u_\infty) \le \liminf_{n \to +\infty} E(u_n) \, ,$$

iii) it is *invariance group* in the domain is reduced to the three-dimensional Möbius group $\mathcal{M}(D^2)$

$$\mathcal{M}(D^2) := \left\{ \Psi(z) = e^{i\theta}\, \frac{z-a}{1-\overline{a}z} \quad \text{s.t. } \theta \in \mathbb{R} \quad |a| < 1 \right\} .$$

Why minimizing E instead of the area has any chance to give a solution to the Plateau problem?

Let u be an immersion of the disc D^2 and denote g_u the pull back of the standard flat metric $g_{\mathbb{R}^3}$ in $\mathbb{R}^3$: $g_u := u^* g_{\mathbb{R}^3}$. The *uniformization theorem* on D^2 gives the existence of a diffeomorphism Ψ of the disc such that

$$\Psi^* g_u = e^{2\lambda}\ [dx_1^2 + dx_2^2] \quad \text{(i.e.,} \quad u \circ \Psi \text{ is conformal)}$$

for some function λ from D^2 into $\mathbb{R}$. So if u minimizes E in the desired class we can replace it by $v := u \circ \Psi$ which is conformal and for which

$$E(v) = \text{Area}(v) = \text{Area}(u) \leq E(u)\, .$$

Hence u is conformal. Assuming now it is not a minimizer of the area, we would then find w in the desired class such that

$$A(w) < A(u) = E(u)\, ,$$

and taking again Ψ' s.t. $E(w \circ \Psi') = A(w)$ we would contradict that u minimizes E.

Hence the difficulty is to find a minimizer of E in the class of C^1 immersions sending continuously and monotonically ∂D^2 into Γ.

One introduces

$$\mathcal{F} := \left\{ \begin{array}{c} u \in W^{1,2}(D^2) \quad ; \quad u \in C^0(\partial D^2, \Gamma) \\ \text{and } u \text{ is monotone from } \partial D^2 \simeq S^1 \text{ into } \Gamma \end{array} \right\} .$$

Fixing the images of three distinct points in ∂D^2 in order to kill the action of the remaining *gauge group* $\mathcal{M}(D^2)$ one proves the existence of a minimizer of E in $\mathcal{F}$ which happens to be a solution to the Plateau problem (a thorough analysis is required to prove that the minimizer is indeed a C^1 immersion).

II.1. The conformal parametrization choice as a Coulomb gauge

As we have seen the conformal parametrizations of immersed discs play a central role in the resolution of the Plateau problem. In the present subsection we establish a one to one correspondence between this choice of conformal parametrization and a *Coulomb gauge* choice.

Let u be a conformal immersion of the disc D^2 (i.e., $H(u) \equiv 0$ on D^2). Let $\lambda \in \mathbb{R}$ such that

$$e^{\lambda} := |\partial_{x_1} u| = |\partial_{x_2} u|\, .$$

Introduce the *moving frame* associated to this parametrization

$$\vec{e}_j := e^{-\lambda}\, \partial_{x_j} u \qquad \text{for } j = 1, 2\, .$$

The family $(\vec{e}_1, \vec{e}_2)$ realizes an orthonormal basis of the tangent space of $u(D^2)$ at $u(x_1, x_2)$. This can be also interpreted as a section of the *frame bundle* of $u(D^2)$ equipped with the induced metric $g_{\mathbb{R}^3}$.

A simple computation gives

$$\begin{aligned}
\operatorname{div}(\vec{e}_1, \nabla\vec{e}_2) &= \partial_{x_1}\left[e^{-2\lambda}\,\partial_{x_1}u\cdot\partial^2_{x_1x_2}u\right] + \partial_{x_2}\left[e^{-2\lambda}\,\partial_{x_1}u\cdot\partial^2_{x_2^2}u\right]\\
&= 2^{-1}\partial_{x_1}\left[e^{-2\lambda}\,\partial_{x_2}|\partial_{x_1}u|^2\right] - 2^{-1}\partial_{x_2}\left[e^{-2\lambda}\,\partial_{x_1}|\partial_{x_2}u|^2\right]\\
&= \partial^2_{x_1x_2}\lambda - \partial^2_{x_2x_1}\lambda = 0\,.
\end{aligned} \tag{II.1}$$

In other words, introducing the 1-form on D^2 given by $A := \vec{e}_1 \cdot d\vec{e}_2$, which is nothing but the *connection form* associated to the *Levi-Civita connection* induced by $g_u = u^* g_{\mathbb{R}^3}$ on the corresponding frame bundle for the trivialization given by $(\vec{e}_1, \vec{e}_2)$, equation (II.1) becomes

$$d^{*_g} A = d^{*_g}\left(\vec{e}_1 \cdot d\vec{e}_2\right) = 0\,, \tag{II.2}$$

where $*_g$ is the *Hodge operator* associated to the induced metric g_u. The equation (II.2) is known as being the *Coulomb condition.* We will see again this condition in the following sections and it is playing a central role in the survey.

Vice versa one proves, see for instance [He] or [40], that for any immersion u, non necessarily conformal, any frame $(\vec{e}_1, \vec{e}_2)$ satisfying the *Coulomb condition* (II.2) corresponds to a conformal parametrization (i.e., $\exists\,\Psi \in \mathrm{Diff}(D^2)$ s.t. $v := u\circ\Psi$ is conformal and $\vec{e}_j = |\partial_{x_j}v|^{-1}\,\partial_{x_j}v$. This observation is the basis of the *Hélein method* for constructing isothermal coordinates.

III. A Plateau type problem on the lack of integrability

In the rest of the class G denotes an arbitrary compact Lie group. We will sometime restrict to the case where G is a special unitary group $SU(n)$ and we will mention it explicitly. The corresponding Lie algebra will be denoted by $\mathcal{G}$ and the neutral element of G is denoted e.

III.1. Horizontal equivariant plane distributions

III.1.1. The definition. Consider the simplest principal fiber structure $P := B^m \times G$ where B^m is the unit m-dimensional ball of $\mathbb{R}^m$. Denote by π the projection map which to $\xi = (x, h) \in P$ assigns the *base point* x. Denote by $Gr_m(TP)$ to be the *Grassmanian* of m-dimensional subspaces of the tangent bundle to P.

We define the notion of *equivariant horizontal distribution of plane* to be a map

$$\begin{aligned}
H\ :\ P = B^m \times G\ &\longrightarrow Gr_m(TP)\\
\xi = (x,h)\ &\longrightarrow\ H_\xi
\end{aligned}$$

satisfying the following 3 conditions

i) the *bundle condition*:

$$\forall\,\xi \in P \qquad H_\xi \in T_\xi P\,,$$

ii) the *horizontality condition*

$$\forall\, \xi \in P \qquad \pi_* H_\xi = T_{\pi(\xi)} B^m \,,$$

iii) the *equivariance condition*

$$\forall\, \xi \in P \quad \forall\, g \in G \qquad (R_g)_* H_\xi = H_{R_g(\xi)}$$

where R_g is the *right multiplication map* by g on P which to any $\xi = (x,h)$ assigns $R_g(\xi) := (x, h\, g)$.

III.1.2. Characterizations of equivariant horizontal distribution of plane by 1-forms on B^m taking values into $\mathcal{G}$. Let H be an equivariant horizontal distribution of plane in $P = B^m \times G$. Clearly the following holds

$$\forall\, \xi = (x,h) \in P \quad \forall X \in T_x B^m \quad \exists!\, X^H(\xi) \in T_\xi P \quad \text{ s.t. } \pi_* X^H = X \,.$$

The vector $X^H(\xi)$ is called the *horizontal lifting* of X at ξ.

At the point (x,e) (recall that e denotes the neutral element of G) we identify $T_{(x,e)}P \simeq T_x M \oplus \mathcal{G}$. Using this identification we deduce the existence of $A_x \cdot X$ such that

$$X^H(x,e) = (X, -\, A_x \cdot X) \,.$$

The 1-form A is called *connection* 1*-form* associated to H. The linearity of A_x with respect to X is a straightforward consequence of the definition of X^H and therefore A defines a 1-form on B^m taking values into $\mathcal{G}$.

For any element $B \in \mathcal{G} \simeq T_e G$ we denote by B^* the unique vector field on G satisfying

$$B^*(e) = B \quad \text{and} \quad \forall g \in G \quad B^*(g) := (R_g)_* B \,,$$

and by an abuse of notation $B^*(g)$ is simply denoted $B\, g$. Using this notation we have

$$\forall\, \xi = (x,h) \in P \quad \forall X \in T_x B^m$$
$$X^H(\xi) = (X, -(A_x \cdot X)^*(h)) = (R_h)_* X^H(x,e) \,.$$

At any point $\xi \in P$ Any vector $Z \in T_\xi P$ admits a decomposition according to H: we denote by Z^V the projection parallel to H_ξ onto the tangent plane to the vertical fiber given by the kernel of π_*:

$$Z^V := Z - (\pi_* Z)^H \,.$$

III.2. The lack of integrability of equivariant horizontal distribution of planes

A m-dimensional plane distribution H is said to be *integrable* if it identifies at every point with the tangent space to a m-dimensional *foliation.*

We aim to “measure” the *lack of integrability* of an equivariant horizontal distribution of planes. To that aim we recall the following classical result

Theorem III.1 (Frobenius). *An m-dimensional plane distribution H is integrable if and only if for any pair of vector fields Y and Z contained in H at every point the bracket $[Y,Z]$ is still included in H at every point.* □

In the particular case of equivariant horizontal distribution of planes in $P = B^m \times G$ we have that H is integrable if and only if

$$\forall\, X, Y \text{ vector fields on } B^m \qquad [X^H, Y^H]^V \equiv 0\,.$$

We shall now compute $[X^H, Y^H]^V$ in terms of the 1-form A.

We write

$$\begin{aligned}{}_{(x,e)} &= [(X, -(A\cdot X)^*), (Y, -(A\cdot Y)^*)]_{(x,e)}\\ &= [(X,0),(Y,0)]_{(x,e)} + [(X,0),(0,-(A\cdot Y)^*)]_{(x,e)}\\ &\quad + [(0,-(A\cdot X)^*),(Y,0)]_{(x,e)} + [(0,(A\cdot X)^*),(0,(A\cdot Y)^*)]_{(x,e)}\,.\end{aligned} \tag{III.1}$$

The definition of the Bracket operation on the Lie algebra $\mathcal{G}$ together with the commutation of the vector field bracket operation with the push-forward operation of the right multiplication map gives that

$$[(A\cdot X)^*),(A\cdot Y)^*] = ([A\cdot X, A\cdot Y])^*\,, \tag{III.2}$$

where the brackets in the r.h.s. of the identity is the Lie algebra bracket operation. The definition of the Lie bracket of vector fields gives

$$[(X,0),(0,-(A\cdot Y)^*)]_{(x,e)} = (0, -\,d(A\cdot Y)\cdot X)\,. \tag{III.3}$$

Finally we write

$$[(X,0),(Y,0)] = ([X,Y],0) = ([X,Y], -A\cdot[X,Y]) + (0, A\cdot[X,Y])\,. \tag{III.4}$$

Combining (III.1), (III.2), (III.3) and (III.4) gives

$$[X^H, Y^H]^V = d(A\cdot X)\cdot Y - d(A\cdot Y)\cdot X + A([X,Y]) + [A\cdot X, A\cdot Y]\,,$$

and using the Cartan formula on the expression of the exterior derivative of a 1-form we obtain

$$[X^H, Y^H]^V = dA(X,Y) + [A\cdot X, A\cdot Y]\,. \tag{III.5}$$

The 2-form we obtained

$$F_A(X,Y) := dA(X,Y) + [A\cdot X, A\cdot Y] \tag{III.6}$$

is the so-called *curvature* of the plane distribution H and "measures" the lack of integrability of H. It will be conventionally denoted

$$F_A = dA + \tfrac{1}{2}[A\wedge A] \quad \text{ or simply } \quad F_A = dA + A\wedge A\,.$$

The Lie algebra, and hence the compact Lie group, is equipped with the Killing form associated to a finite-dimensional representation, hence unitary, for which the form defines a scalar product invariant under adjoint action. For instance $\mathcal{G} = o(n)$ or $\mathcal{G} = u(n)$

$$\langle B, C\rangle = -\,\mathrm{Tr}(B\,C)\,.$$

If the Lie algebra $\mathcal{G}$ is semi-simple: it is a direct sum of Lie algebras with no non trivial ideal, the Lie algebra is equipped with the *Killing scalar product*:

$$\langle B, C\rangle := -\mathrm{Tr}(\mathrm{ad}(B)\,\mathrm{ad}(C))\,,$$

where $\mathrm{ad}(B)$ is the following endomorphism of $\mathcal{G}$: $\mathrm{ad}(B)(D) := [B, D]$.

The Lagrangian we are considering for measuring the lack of integrability of the plane distribution H is just the L^2 norm of the curvature

$$\int_{B^m} \sum_{i<j} \left| [\partial^H_{x_i}, \partial^H_{x_j}]^V \right|^2 \, dx^m = \int_{B^m} |F_A|^2 \, dx^m \, ,$$

where dx^m is the canonical volume form on B^m. The L^2 norm of the curvature is also known as being the *Yang–Mills energy* of the connection form A and is denoted $YM(A)$. We can now state one of the main problems these notes are addressing.

Yang–Mills Plateau Problem: *Let η be a 1-form on ∂B^m taking values into a Lie algebra $\mathcal{G}$ of a compact Lie group G does there exists a 1-form A into $\mathcal{G}$ realizing*

$$\inf \left\{ YM(A) = \int_{B^m} |dA + A \wedge A|^2 \, dx^m \quad ; \quad \iota^*_{\partial B^m} A = \eta \right\} ,$$

where $\iota_{\partial B^m}$ is the canonical inclusion of the boundary ∂B^m into $\mathbb{R}^m$.

In other words we are asking the following question:

> *Given an equivariant horizontal plane distribution over the boundary of the unit ball in $\mathbb{R}^m$, can one extend it inside the ball in an optimal way with respect to the L^2 norm of the integrability defect?*

In order to study this variational problem we first have to identify its invariance group.

III.3. The gauge invariance

In this subsection we identify the group of the Yang–Mill Plateau problem corresponding to the diffeomorphism group of the disc for the area in the classical Plateau problem.

Let g be a map from B^m into G. We denote by $L_{g^{-1}}$ the left multiplication by g^{-1} defined as follows

$$\begin{aligned} L_{g^{-1}} \ ; \ P = B^m \times G \ &\longrightarrow P \\ \xi = (x,h) \ &\longrightarrow \ (x, g^{-1}\, h) \, . \end{aligned}$$

Let H be an *equivariant horizontal distribution of planes* on P we observe that the push-forward by $L_{g^{-1}}$ of H, $(L_{g^{-1}})_* H$, is still an *equivariant horizontal distribution of planes*. We now compute the connection 1-form associated to this new distribution.

Let X be a vector of $T_x B^m$ and $x(t)$ a path in B^m such that $\dot{x}(0) = X$. Let $h(t) \in G$ such that $\xi(t) := (x(t), h(t))$ is the horizontal lifting of $x(t)$ starting at the neutral element e of G (i.e., $\dot{\xi} = (\dot{x})^H(\xi(t))$ and $\xi(0) = (x,e)$). Since $\dot{xi}^V = 0$ we have in particular

$$\frac{dh}{dt}(0) = -A \cdot X \, .$$

The push-forward by $L_{g^{-1}}$ of the horizontal vector $X^H(x,e)$ is the horizontal lifting of X at (x,g^{-1}) for the distribution $(L_{g^{-1}})_*H$: $X^{(L_{g^{-1}})_*H}(x,g^{-1})$. Hence we have

$$\begin{aligned}
X^{(L_{g^{-1}})_*H}(x,g^{-1}) &= (L_{g^{-1}})_*X^H(x,e) = (L_{g^{-1}})_*(X,-A\cdot X)\\
&= \frac{d}{dt}\left(x(t),g^{-1}h(t)\right) = (X, dg^{-1}\cdot X - g^{-1}\,A\cdot X)\\
&= \left(X, -\left(g^{-1}dg\cdot X + g^{-1}\,A\cdot X\,g\right)g^{-1}\right)\\
&= \left(X, -\left(g^{-1}dg\cdot X + g^{-1}\,A\cdot X\,g\right)^*(g^{-1})\right).
\end{aligned}$$

Hence we have proved that the horizontal lift at (x,e) for the new plane distribution $(L_{g^{-1}})_*H$ is $\left(X,-\left(g^{-1}dg\cdot X+g^{-1}\,A\cdot X\,g\right)\right)$ and the associated connection 1-form associated to the distribution $(L_{g^{-1}})_*H$ is

$$A^g = g^{-1}dg\cdot X + g^{-1}\,A\cdot X\,g\,.$$

The curvature associated to this new distribution is given by

$$F_{A^g}(X,Y) = dA^g(X,Y) + [A^g(X), A^g(Y)]\,.$$

We have on one hand

$$dA^g(X,Y) = dg^{-1}\wedge dg(X,Y) + [dg^{-1}\,g\wedge g^{-1}Ag](X,Y)\,,$$

and on the other hand

$$[A^g(X),A^g(Y)] = [g^{-1}dg\wedge g^{-1}Ag](X,Y) + g^{-1}[A(X),A(Y)]g\,.$$

Summing the two last identities and using the fact that $g^{-1}dg + dg^{-1}g = 0$ gives finally

$$F_{A^g} = g^{-1}\,F_A\,g\,.$$

Since the Killing scalar product on $\mathcal{G}$ is invariant under the adjoint action of G we have

$$YM(A^g) = YM(A)\,.$$

The action of $L_{g^{-1}}$ on the plane distribution H leaves invariant its Yang–Mills energy and realizes therefore a "huge" invariance group of the *Yang–Mills Plateau Problem* which is called the *gauge group* of the problem.

Exactly as for the classical Plateau problem we discussed in the first part of the survey the task for solving the Yang–Mills Plateau problem will be to "kill" this gauge invariance and, here again, the Coulomb gauge choices will be of great help.

III.4. The Coulomb gauges

We first start with the simplest group, the Abelian group $G = S^1$. The Yang–Mills Plateau problem in this case becomes:

Find a minimizer of

$$\inf\left\{YM(A) = \int_{B^m}|dA|^2\;dx^m\quad;\quad \iota^*_{\partial B^m}A = \eta\right\}\,,$$

where η is some given 1-form on the boundary ∂B^m.

In a reminiscent way to the classical Plateau problem our starting functional is degenerate and we shall replace it by a more coercive one

$$YM(A) \leq E(A) = \int_{B^m} |dA|^2 + |d^*A|^2 \ dx^m \,,$$

with equality if and only if $d^*A = 0$ (i.e., A satisfies the Coulomb condition).

The following coercivity inequality holds

$$\forall\, A \in W^{1,2}(B^m, \mathcal{G}) \qquad \text{s.t. } \iota^*_{\partial B^m} A = \eta$$
$$\int_{B^m} |A|^2 + \sum_{i,j=1}^{m} |\partial_{x_i} A_j|^2 \ dx^m \leq C \left[E(A) + \|\eta\|^2_{H^{1/2}(\partial B^m)} \right] \,, \tag{III.7}$$

for some fixed constant independent of η and A, where $H^{1/2}(\partial B^m)$ is the fractional trace space of $W^{1,2}(B^m)$. The convexity of E in $W^{1,2}_\eta(\wedge^1 B^m, \mathcal{G})$ together with the previous coercivity inequality implies that the following problem admits a unique minimizer A_0 of

$$\min_{A \in W^{1,2}_\eta(\wedge^1 B^m, \mathcal{G})} E(A)\,.$$

The Euler Lagrange equation reads

$$\forall\ \phi \in W^{1,2}_0(\wedge^1 B^m, \mathcal{G}) \qquad \int_{B^m} d\phi \wedge * dA + (-1)^m\, d^*A \wedge d * \phi = 0\,. \tag{III.8}$$

This gives in particular that each of the components of A are harmonic. We choose f to be an arbitrary function in $L^2(B^m, \mathcal{G})$ and we solve

$$\begin{cases} -\Delta u = f & \text{in} \quad B^m \\ \quad u = 0 & \text{on} \quad \partial B^m\,. \end{cases}$$

Classical elliptic theory gives that the 1-form $\phi := du$ is in $W^{1,2}_0(\wedge^1 B^m, \mathcal{G})$. By substituting $\phi = du$ in (III.8) we obtain

$$\forall\ f \in L^2(B^m, \mathcal{G}) \quad \int_{B^m} d^*A \cdot f \ dx^m = 0\,.$$

Hence we deduce $d^*A = 0$ in B^m and there exists a unique solution to (III.8) which is also the unique solution of the following system

$$\begin{cases} d^* dA_0 = 0 & \text{in } \mathcal{D}'(B^m) \\ d^* A_0 = 0 & \text{in } \mathcal{D}'(B^m) \\ \iota^*_{\partial B^m} A_0 = \eta\,. \end{cases}$$

The components of A are harmonic in B^m and are therefore smooth moreover we have

$$YM(A_0) = E(A_0)\,.$$

Let now B in $W^{1,2}_\eta(\wedge B^m,\mathcal{G})$ we claim that there exists a gauge change g such that $YM(B)=YM(B^g)=E(B^g)$. This can be seen as follows. Let φ be the solution of

$$\begin{cases} -\Delta\varphi = d^*B & \text{in } \mathcal{D}'(B^m) \\ \varphi = 0 & \text{on } \partial B^m\,. \end{cases}$$

Hence we have

$$\begin{cases} d(B+d\varphi) = dB & \text{in } \mathcal{D}'(B^m) \\ d^*(B+d\varphi) = 0 & \text{in } \mathcal{D}'(B^m) \\ \iota_{\partial B^m}(B+d\varphi) = \eta\,. \end{cases}$$

Taking $g := \exp(i\varphi)$ we have $YM(B)=YM(B^g)=E(B^g)$.

Hence A_0 realizes

$$\min_{A\in W^{1,2}_\eta(\wedge^1 B^m,\mathcal{G})} YM(A)\,.$$

Indeed if there would be $B\in W^{1,2}_\eta(\wedge^1 B^m,\mathcal{G})$ such that $YM(B)<YM(A_0)$ we choose g such that $YM(B^g)=E(B^g)$ and we would contradict the fact that A_0 minimizes E.

Taking now a general compact Lie group G we would also like to propose to minimize E instead of YM but we need first ensure that a Coulomb gauge always exists. We have the following lemma which answers positively to this last question

Lemma III.1. *Let $A\in L^2(\wedge^1 B^m,\mathcal{G})$. The following variational problem*

$$\inf_{g\in W^{1,2}_e(B^m,G)} \int_{B^m} |g^{-1}dg + g^{-1}Ag|^2\ dx^m$$

is achieved and each minimizer satisfies the Coulomb condition

$$d^*(g^{-1}dg + g^{-1}A\,g) = 0\,. \qquad \square$$

Proof of Lemma III.1. Let g_k be a minimizing sequence. Since the group is compact we have that

$$\limsup_{k\to+\infty}\int_{B^m} |dg_k|^2\ dx^m < +\infty\,.$$

Hence, using the Rellich–Kondrachov theorem there exists a subsequence $g_{k'}$ converging weakly in $W^{1,2}(B^m,G)$ to g_∞ and strongly in every $L^p(B^m,G)$ space for any $p<+\infty$ hence $g_\infty\in W^{1,2}_e(B^m,G)$. The same holds for g_k^{-1} and its weak limit in $W^{1,2}(B^m,G)$ is the inverse of g_∞. Hence we have

$$g_k^{-1}dg_k + g_k^{-1}A\,g_k \rightharpoonup g_\infty^{-1}dg_\infty + g_\infty^{-1}Ag_\infty \qquad \text{in } \mathcal{D}'(B^m)\,.$$

The lower semi continuity of the L^2 norm implies that g_∞ is a minimizer of (III.6).

For any $U\in C^\infty_0(B^m,\mathcal{G})$ we introduce

$$g_\infty(t) := g_\infty \exp(t\,U)\,.$$

We have

$$\begin{aligned} &g_\infty^{-1}(t)dg_\infty^{-1}(t) + g_\infty^{-1}(t)Ag_\infty(t) \\ &\qquad = \exp(-t\,U)\,d\exp(t\,U) + \exp(-t\,U)\left[g_\infty^{-1}dg_\infty + g_\infty^{-1}Ag_\infty\right]\exp(t\,U)\,. \end{aligned}$$

Hence

$$\frac{d}{dt}\left[g_\infty^{-1}(t)dg_\infty^{-1}(t)+g_\infty^{-1}(t)Ag_\infty(t)\right]=dU-[U,A^{g_\infty}]\,.$$

Since g_∞ is a minimizer we have

$$0=\frac{d}{dt}\int_{B^m}|A^{g_\infty(t)}|^2\ dx^m=2\int_{B^m}\langle(dU-[U,A^{g_\infty}])\cdot A^{g_\infty}\rangle\ dx^m\,.$$

We use the identity $\langle[U,V],W\rangle=\langle U,[V,W]\rangle$ to deduce that

$$[U,A^{g_\infty}]\cdot A^{g_\infty}=0$$

and then we have proved that for any $U\in C_0^\infty(B^m,\mathcal{G})$ we have

$$0=\int_{B^m}dU\cdot A^{g_\infty}\ dx^m=\int_{B^m}\langle U,d^*A^{g_\infty}\rangle\ dx^m\,.$$

This finishes the proof of the lemma. $\square$

Since every connection form possesses a Coulomb gauge representative it is then tempting to minimize E instead of YM following the main lines of the Abelian case. However due to the non-linearity $A\wedge A$ in F_A it is not clear whether the E energy controls the $W^{1,2}$ norm of A in a similar way of (III.7) in the general case.

In fact the answer to that question is "no" as we can see in the following example. We take $G=SU(2)$ and we identify $su(2)$ with the imaginary quaternions. On $\mathbb{R}^4$ we identify canonically the point of coordinates (x_0,x_1,x_2,x_3) with the quaternion $\mathbf{x}:=x_0+x_1\,\mathbf{i}+x_2\,\mathbf{j}+x_3\,\mathbf{k}$. For a quaternion $\mathbf{y}=y_0+y_1\,\mathbf{i}+y_2\,\mathbf{j}+y_3\,\mathbf{k}$ we denote by $\Im(\mathbf{y})$ the element in $su(2)$ given by $y_1\,\sigma_1+y_2\,\sigma_2+y_3\,\sigma_3$ where σ_i are the Pauli matrices to which we identify $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$

$$\mathbf{i}\leftrightarrow\sigma_1=\begin{pmatrix}i&0\\0&-i\end{pmatrix}\qquad\mathbf{j}\leftrightarrow\sigma_2=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\qquad\mathbf{k}\leftrightarrow\sigma_3=\begin{pmatrix}0&i\\i&0\end{pmatrix}$$

forming an orthogonal basis of $su(2)$ with norms $\sqrt{2}$ for each vector of the basis.

On B^4, for $\lambda\in\mathbb{R}_+^*$, we consider the family of 1-forms into $su(2)$ given by

$$A_\lambda:=\lambda^2\ \frac{\Im\left(\mathbf{x}\,d\overline{\mathbf{x}}\right)}{1+\lambda^2\,|\mathbf{x}|^2}\,.$$

The corresponding curvature is given by

$$F_{A_\lambda}=\lambda^2\,\frac{d\mathbf{x}\wedge d\overline{\mathbf{x}}}{(1+\lambda^2\,|\mathbf{x}|^2)^2}\,.$$

One easily verifies that[1]

$$\lim_{\lambda\to+\infty}\int_{B^4}|F_{A_\lambda}|^2\ dx^4=\int_{\mathbb{R}^4}|F_{A_1}|^2\ dx^4=\int_{\mathbb{R}^4}48\ \frac{dx^4}{(1+|x|^2)^4}=8\pi^2<+\infty$$

[1]The somehow surprising factor 48 comes from the fact that there are 6 curvature coordinates and each curvature coordinate has the form

$$|(F_{A_1})_{ij}|^2=\frac{8}{(1+|x|^2)^2}$$

where we have used that the square of the norm of each Pauli matrix is 2.

but one verifies also that

$$\lim_{\lambda\to+\infty}\int_{B^4}\sum_{i,j=1^4}|\partial_{x_i}(A_\lambda)_j|^2\ dx^4=+\infty$$

we might then think that by changing the gauge we can avoid this blow up of the $W^{1,2}$ norm of the connection form but, as we see now, this cannot be the case. Consider the *second Chern form* $Tr(F_{A_\lambda}\wedge F_{A_\lambda})$, it satisfies

$$\lim_{\lambda\to+\infty}\mathrm{Tr}(F_{A_\lambda}\wedge F_{A_\lambda})=8\pi^2\,\delta_0\ dx^4\,. \tag{III.9}$$

The *second Chern form* is invariant under gauge transformation and for any choice of gauge g this 4-form, which has to be closed, is on B^4 the exterior derivative of the *transgression form* known as the *Chern–Simon* 3*-form*:

$$\forall\, g\ :B^4\ \to SU(2)$$

$$\mathrm{Tr}(F_{A_\lambda}\wedge F_{A_\lambda})=d\left[\mathrm{Tr}\left(A^g_\lambda\wedge dA^g_\lambda+\frac{1}{3}A^g_\lambda\wedge[A^g_\lambda,A^g_\lambda]\right)\right]\,.$$

Assume now there would have been a gauge g_λ s.t.

$$\liminf_{\lambda\to+\infty}\|A^{g_\lambda}_\lambda\|_{W^{1,2}(B^4)}<+\infty\,.$$

Then for some sequence $\lambda_k\to+\infty$, using the Rellich–Kondrachov theorem $A^k:=A^{g_k}_{\lambda_k}$ would weakly converge to some limit A^∞ in $W^{1,2}(\wedge^1B^4,su(2))$ and strongly in $L^p(\wedge^1B^4,su(2))$ for any $p<4$. Hence

$$\mathrm{Tr}\left(A^k\wedge dA^k+\frac{1}{3}A^k\wedge[A^k,A^k]\right)\rightharpoonup\mathrm{Tr}\left(A^\infty\wedge dA^\infty+\frac{1}{3}A^\infty\wedge[A^\infty,A^\infty]\right)$$

in $\mathcal{D}'(B^4)$. Taking now the exterior derivative and using again the gauge invariance of the second Chern form we obtain

$$\mathrm{Tr}(F_{A_{\lambda_k}}\wedge F_{A_{\lambda_k}})\rightharpoonup\mathrm{Tr}(F_{A^\infty}\wedge F_{A^\infty})\qquad\text{in }\mathcal{D}'(B^4)\,.$$

Since A^∞ is in $W^{1,2}$ the 4-form $\mathrm{Tr}(F_{A^\infty}\wedge F_{A^\infty})$ is an L^1 function, however, from (III.9) it is equal to the Dirac mass. This gives a contradiction and we have proved the following proposition

Proposition III.1. *There exists* $A^k\in W^{1,2}(B^4,su(2))$ *such that*

$$\limsup_{k\to+\infty}\int_{B^4}|F_{A^k}|^2\ dx^4<+\infty\,,$$

but

$$\liminf_{k\to+\infty}\inf\left\{\int_{B^4}\sum_{i,j=1}^4|\partial_{x_i}(A^k)^g_j|^2\ dx^4\ ;\ g\in W^{2,2}(B^4,SU(2))\right\}=+\infty\,.\qquad\square$$

Hence by minimizing E instead of YM we don't get enough control on the minimizing sequence A^k in order to extract a converging subsequence to a solution to the Yang–Mills Plateau problem.

The situation would have been much better in dimension less than four where a $W^{1,2}$ control of A in terms of $E(A)$ do exist. In dimension equal to four, despite Proposition III.1, there is still a positive result in that line which says roughly that such a control does exist for some gauge provided the *Yang–Mills energy* stays below some positive threshold. The following section is devoted to the proof of this result by K. Uhlenbeck.

IV. Uhlenbeck's Coulomb gauge extraction method

IV.1. Uhlenbeck's construction

We have seen that in dimension four – and higher of course – there is no hope to control the $W^{1,2}$ norm of sequences of connection forms from the E energy. The fact that the dimension four is critical for this phenomenon comes form the optimal *Sobolev embedding*

$$W^{1,2}(B^4) \hookrightarrow L^4(B^4)\,,$$

which does not hold in higher dimension.

Theorem IV.1. *Let $m \leq 4$ and G be a compact Lie group. There exists $\varepsilon_G > 0$ and $C_G > 0$ such that for any $A \in W^{1,2}(B^m, \mathcal{G})$ satisfying*

$$\int_{B^m} |dA + A \wedge A|^2 \; dx^m < \varepsilon_G\,,$$

there exists $g \in W^{2,2}(B^m, G)$ such that

$$\begin{cases} \displaystyle\int_{B^m} |A^g|^2 + \sum_{i,j=1}^{4} |\partial_{x_i} A^g_j|^2 \, dx^m \leq C_G \int_{B^m} |dA + A \wedge A|^2 \, dx^m \\ d^* A^g = 0 \qquad \text{in } B^m \\ \iota^*_{\partial B^m}(*A^g) = 0\,, \end{cases} \tag{IV.1}$$

where $A^g = g^{-1}dg + g^{-1}Ag$ and $\iota_{\partial B^m}$ is the canonical inclusion map of the boundary of the unit ball into $\mathbb{R}^m$. □

Remark IV.1. The same result holds in arbitrary dimension replacing the L^2-norm for the curvature by the $L^{m/2}$ norm and the $W^{1,2}$ norm of the connection by the $W^{1,m/2}$-norm (see [59]). The proof we are giving below can be transposed word by word in this more general setting. The adaptation requires just a shift of the exponents $4 \to m$, $2 \to m/2$. □

For $m < 4$ the non-linearity $A \wedge A$ is a compact perturbation of dA and the problem is a perturbation to a simple linear one that we solved in the Abelian case. We will then restrict the presentation to the case $m = 4$. We assume that the compact Lie group is represented by a subgroup of invertible matrices in R^n for some $n \in \mathbb{N}^*$ which gives an isometric embedding of G in an Euclidian space.

We aim to solve the Coulomb equation $d^* A^g = 0$ keeping this time a control of A^g in $W^{1,2}$ (lemma III.1 was only giving an L^2 control). Since we have little energy the hope is to use a fixed point argument close to the zero connection and for g close to the identity. The linearization of the *Coulomb non-linear elliptic PDE*

$$\begin{cases} d^*\left[g^{-1}\,dg\right] = -\,d^*\left[g^{-1}A\,g\right] & \text{in } B^4 \\ \partial_r g\,g^{-1} = -\,\langle A,\partial_r\rangle & \text{on } \partial B^4\,, \end{cases}$$

for $A^t = t\,\omega$ and $g^t = \exp(t\,U)$ gives

$$\begin{cases} \Delta U = -d^*dU = d^*\omega & \text{in } B^4 \\ \partial_r U = -\,\langle\omega,\partial_r\rangle & \text{on } \partial B^4\,. \end{cases}$$

This linearized problem is solvable for any map $\omega \in W^{1,2}(B^4,\mathcal{G})$ and we get a unique $U \in W^{2,2}(B^4,\mathcal{G})$ solving the previous linear equation.

However, in order to be able to apply the implicit function theorem we need the following non-linear mapping to be <u>smooth</u>

$$\begin{aligned} \mathcal{N}^0\ :\ W^{1,2}(B^4,\mathcal{G})\times W^{2,2}(B^4,\mathcal{G}) &\longrightarrow W^{1,2}(B^4,\mathcal{G})\times H^{1/2}(\partial B^4,\mathcal{G}) \\ (\omega,U) &\longrightarrow \left(d^*\left[g_U^{-1}\,dg_U + g_U^{-1}\omega\,g_U\right], \partial_r g_U\,g_U^{-1} - \langle\omega,\partial_r\rangle\right)\,, \end{aligned}$$

where $g_U := \exp(U)$. This is however <u>not the case</u> in dimension four. This is due to the fact that $W^{2,2}$ does not embed in C^0 in four dimensions but only in the *Vanishing Mean Oscillation* space $VMO(B^4)$ hence simple algebraic operations such as the multiplication of two G valued $W^{2,2}$ maps is not continuous in four dimensions.

If one replaces however $W^{1,2}\times W^{2,2}$ by a "slightly smaller space" $W^{1,p}\times W^{2,p}$ for any $p>2$ (p being as close as we want from 2) then the space $W^{2,p}$ embeds continuously (and compactly) in C^0 and the map $\mathcal{N}^0$ becomes suddenly <u>smooth</u>! and a fixed point argument is conceivable in this smaller space.

Uhlenbeck's strategy consists in combining a fixed point argument in smaller spaces – in which the problem is invertible – together with a continuity argument.

This method is rather generic in the sense that it can be applied to critical *extensions* or *lifting problems* of maps in the Sobolev space $W^{1,m}(M^m)$ which misses to embed in C^0 but for which however the notion of homotopy class is well defined (see [47] and [61]) and prevents to find global extensions or liftings when the norm of the map is too high. As an illustration we shall give two results.

Theorem IV.2. *For any $m\ge 1$, and any compact Lie group G there exist $\varepsilon_{m,G}>0$ and $C_{m,G}>0$ such that for any map $g\in W^{1,m}(S^m,G)$ satisfying*

$$\int_{S^m} |dg|^m\ \ \mathrm{dvol}_{S^m} < \varepsilon_m\,,$$

there exists an extension $\tilde g \in W^{1,m+1}(B^{m+1},G)$ equal to g on ∂B^{m+1} such that

$$\int_{B^{m+1}} |d\tilde g|^{m+1}\ dx^{m+1} \le C_m \int_{S^m} |dg|^m\ \ \mathrm{dvol}_{S^m}\ . \qquad \square$$

Remark IV.2. The existence of such an extension $\tilde{g} \in W^{1,m+1}(B^{m+1}, G)$ is clearly not true for general $g \in W^{1,m}(S^m, S^m)$. Indeed, consider for instance $m = 3$ and $G = SU(2) \simeq S^3$, if such an extension would exists one would have to use the Stokes theorem.

$$0 = \int_{B^4} \tilde{g}^* dx^4 = \frac{1}{4} \int_{S^3} g^* \, \mathrm{dvol}_{SU(2)} = \frac{|S^3|}{4} \deg(g) \, ,$$

where $\deg(g)$ is the topological degree of the map g which is not necessarily zero. However there always exists an extension in the "slightly" larger space $W^{1,(m+1,\infty)}(B^{m+1}, G)$ of maps from B^m into G with one derivative in the Marcinkiewicz weak $L^{m+1}(B^{m+1})$ space (see Subsection IV.3 below). □

Theorem IV.2 is proved in [33] using Uhlenbeck's method. The second example is the following one

Theorem IV.3. *Let P be a G principal bundle over S^m where G is a compact Lie group and where π is the projection associated to this bundle. There exist $\varepsilon_{m,G} > 0$ and $C_{m,G} > 0$ such that for any $g \in W^{1,m}(S^m, S^m)$ satisfying*

$$\int_{S^m} |dg|^m \; \mathrm{dvol}_{S^m} < \varepsilon_{m,G} \, ,$$

there exists $v \in W^{1,m}(S^m, P)$ such that

$$\int_{S^m} |dv|^m \; \mathrm{dvol}_{S^m} \le C_{m,G} \int_{S^m} |dg|^m \; \mathrm{dvol}_{S^m} \, , \quad \textit{and} \quad \pi \circ v = g \, . \qquad \square$$

Proof of Theorem IV.1. Fix some $2 < p < 4$. For any $\varepsilon > 0$ we introduce

$$\mathcal{U}^\varepsilon := \left\{ A \in W^{1,p}(B^4, \mathcal{G}) \quad \text{s.t.} \quad \int_{B^4} |F_A|^2 \; dx^4 < \varepsilon \right\} ,$$

and for any $\varepsilon > 0$ and $C > 0$ and we consider

$$\mathcal{V}_C^\varepsilon := \left\{ \begin{array}{l} A \in \mathcal{U}^\varepsilon \quad \text{s.t.} \quad \exists\, g \in W^{2,p}(B^4, G) \\ \displaystyle\int_{B^4} |dA^g|^p_{g_{S^4}} \; dx^4 \le C \int_{B^4} |F_A|^p \; dx^4 \\ \displaystyle\int_{B^4} |dA^g|^2_{g_{S^4}} \; dx^4 \le C \int_{B^4} |F_A|^2 \; dx^4 \\ \quad d^* A^g = 0 \quad \text{and} \quad \iota_{\partial B^4} * A^g = 0 \, . \end{array} \right\}$$

The first goal is to show the following

Claim 1 $\qquad \exists\, \varepsilon > 0 \quad C > 0 \quad \text{s.t.} \quad \mathcal{V}_C^\varepsilon = \mathcal{U}^\varepsilon \, .$

In order to prove the claim we shall establish successively

1) The set $\mathcal{U}^\varepsilon$ is path connected.
2) The set $\mathcal{V}_C^\varepsilon$ is closed in $\mathcal{U}^\varepsilon$ for the $W^{1,p}$-topology
3) For $\varepsilon > 0$ chosen small enough and $C > 0$ large enough the set $\mathcal{V}_C^\varepsilon$ is open for the $W^{1,p}$-topology

Since $\mathcal{V}_C^\varepsilon$ is non-empty, this will imply the claim 1 for this choice of ε and C.

Proof of the path connectedness of $\mathcal{U}^\varepsilon$. For A in $\mathcal{U}^\varepsilon$ and $t \in [0,1]$ we define the connection form A^t to be the image of A by the dilation of rate t^{-1}: $A^t = t \sum_{j=1} A_j(t\,x)\; dx_j$. We have in particular

$$F_{A^t} = t^2 \sum_{i,j=1}^{4} (F_A)_{ij}(x)\; dx_i\, dx_j\,,$$

hence

$$\int_{B^4} |F_{A^t}|^2\; dx^4 = \int_{B^4_t} |F_A|^2\; dx^4 < \varepsilon\,,$$

and

$$\sum_{i,j=1}^{4} \int_{B^4} |\partial_{x_i}(A_t)_{x_j}|^p\; dx^4 = O(t^{2p-4})\,. \tag{IV.2}$$

This shows that on one hand $A^t \in \mathcal{U}^\varepsilon$ for any $t \in [0,1]$ and that on the other hand $A^t \to 0$ strongly in $W^{1,p}(B^4)$. Hence A^t is a continuous path contained in $\mathcal{U}^\varepsilon$ connecting A and 0 which prove the path connectedness of $\mathcal{U}^\varepsilon$.

Proof of the closeness of $\mathcal{V}_C^\varepsilon$ in $\mathcal{U}^\varepsilon$. Let $A^k \in \mathcal{V}_C^\varepsilon$ and assume A^k converges strongly in $W^{1,p}$ to some limit $A^\infty \in \mathcal{U}^\varepsilon$. We claim that $A^\infty \in \mathcal{V}_C^\varepsilon$.

Since $A^k \to A^\infty$ strongly in $W^{1,p}$, $dA^k \to dA^\infty$ strongly in L^p and, using the Sobolev embedding, $A^k \to A^\infty$ in $L^{4p/4-p}$. Hence, due to the later,

$$A^k \wedge A^k \longrightarrow A^\infty \wedge A^\infty \text{ strongly in } L^{2p/4-p}(B^4)\,.$$

We have chosen $2 < p < 4$ in such a way that $p < 2p/4 - p$. Hence we deduce that

$$F_{A^k} \longrightarrow F_{A^\infty} \qquad \text{strongly in } L^p(B^4)\,. \tag{IV.3}$$

Let g^k be a sequence such that

$$\int_{B^4} |d(A^k)^{g^k}|^q_{g_{B^4}}\; \mathrm{dvol}_{S^4} \le C \int_{B^4} |F_{A^k}|^q\; \mathrm{dvol}_{S^4}\,, \tag{IV.4}$$

for $q = 2, p$ and

$$d^*(A^k)^{g^k} = 0 \quad \text{ and } \quad \iota^*_{\partial B^4} * (A^k)^{g^k} = 0\,. \tag{IV.5}$$

Since both $d(A^k)^{g^k}$ and $d^*(A^k)^{g^k}$ are uniformly bounded in L^p and since there is no harmonic 3-form[2] on B^4 whose restriction on the boundary is zero, classical L^p-Hodge theory (see for instance [21]) infers that $(A^k)^{g^k}$ is uniformly bounded in $W^{1,p}$. So, using the Sobolev embedding, it is in particular bounded in $L^{4p/4-p}$ and since $(A^k)^{g^k} = (g^k)^{-1}dg^k + (g^k)^{-1}A^k\, g^k$, using that A^k is bounded in $W^{1,p}$

[2] Assume B is a 3 form satisfying $dB = 0$ and $d^*B = 0$ on B^4 and $\iota^*_{\partial B^4}B = 0$ then $*B = d\varphi$, $B = d\beta$ and $\iota^*_{\partial B^4}\beta = dC$. With these notations we have

$$\int_{B^4} |B|^2 = \int_{B^4} d\beta \wedge d\varphi = \int_{\partial B^4} dC \wedge d\varphi = 0$$

hence $B = 0$.

and hence in $L^{4p/4-p}$, we deduce that dg^k is bounded also in $L^{4p/4-p}$. Since $4 < 4p/4-p$, there exists a subsequence $g^{k'}$ converging strongly to some limit g^∞ in C^0. Going back to the weak convergence of $(A^k)^{g^k}$ in $W^{1,p}$, the strong convergence of A^k in $W^{1,p}$ and the weak convergence of g^k in $W^{1,4p/4-p}$ we deduce that

$$dg^k = g^k\,(A^k)^{g^k} - A^k\,g^k$$

is uniformly bounded in $W^{1,p}$ and therefore $g^{k'} \rightharpoonup g^\infty$ weakly in $W^{2,p}(B^4,G)$. Thus we deduce that the following weak convergence in $W^{1,p}(B^4)$ holds

$$(g^{k'})^{-1}dg^{k'} + (g^{k'})^{-1}A^{k'}\,g^{k'} \rightharpoonup (g^\infty)^{-1}dg^\infty + (g^\infty)^{-1}A^\infty\,g^\infty\,.$$

Combining (IV.3) and the latest weak convergence we deduce that

$$\int_{B^4} |d(A^\infty)^{g^\infty}|^q\,dx^4 \le C\int_{B^4} |F_{A^\infty}|^q\ dx^4\,, \tag{IV.6}$$

for $q = 2, p$ and, using the following continuous embedding of

$$W^{1,p}(B^4) \hookrightarrow W^{1-1/p,p}(\partial B^4)$$

we have

$$d^*(A^\infty)^{g^\infty} = 0 \quad \text{ and } \quad \iota^*_{\partial B^4}(A^\infty)^{g^\infty}\,. \tag{IV.7}$$

So we have proved that A^∞ fulfill all the conditions for being in $\mathcal{V}^\varepsilon_C$.

Proof of the openness of $\mathcal{V}^\varepsilon_C$. Let A be an element of $\mathcal{V}^\varepsilon_C$. It is clear that if we find in $\mathcal{V}^\varepsilon_C$ an open neighborhood for the $W^{1,p}$-topology of the $W^{1,p}$ Coulomb gauge A^g, then A possesses also such a neighborhood. So we can assume right away that $d^*A = 0$ and $\iota_{\partial B^4} * A = 0$.

We are looking for the existence of δ sufficiently small – possibly depending on A – such that for any ω satisfying $\|\omega\|_{W^{1,p}} < \delta$ there exists g close to the identity in $W^{2,p}$-norm such that

$$d^*\left[g^{-1}dg + g^{-1}(A+\omega)\,g\right] = 0 \quad \text{ and } \quad \iota_{\partial B^4} * (A+\omega)^g = 0\,.$$

To that purpose we introduce the map

$$\begin{aligned} \mathcal{N}^A\ &:\ W^{1,p}(B^4,\mathcal{G})\times W^{2,p}(B^4,\mathcal{G}) \longrightarrow L^p(B^4,\mathcal{G})\times W^{1-1/p,p}(\partial B^4,\mathcal{G}) \\ &(\omega,U) \ \longrightarrow\ \left(d^*\left[g_U^{-1}\,dg_U + g_U^{-1}(A+\omega)\,g_U\right],\ \iota_{\partial B^4} * (A+\omega)^{g_U}\right)\,, \end{aligned} \tag{IV.8}$$

where $g_U := \exp(U)$. We have seen that this map is smooth.

The derivative of $\mathcal{N}^A$ along the U direction at $(0,0)$ gives

$$\partial_U\mathcal{N}^A(0,0)\cdot V = (-\Delta V + [A,dV], \partial_r V)\,,$$

where $\Delta = \sum_{k=1}^4 \partial^2_{x_k^2}$.

Using Calderòn–Zygmund L^p theory (see for instance [52] or [16]) we have the following a priori estimate for any V satisfying $\int_{B^4} V\ dx^4 = 0$

$$\begin{aligned} \|V\|_{W^{2,p}(B^4)} &\le c\,\left[\|\Delta V\|_{L^p(B^4)} + \|\partial_r V\|_{W^{1-1/p,p}(\partial B^4)}\right] \\ &\le c\,\left[\|\partial_U\mathcal{N}^A(0,0)\cdot V\|_{\mathcal{F}} + \|[A,dV]\|_{L^p(B^4)}\right] \\ &\le c\,\left[\|\partial_U\mathcal{N}^A(0,0)\cdot V\|_{\mathcal{F}} + c\,\|A\|_{L^4(B^4)}\ \|dV\|_{L^{4p/4-p}(B^4)}\right]\,, \end{aligned}$$

where $\mathcal{F}$ is the hyperplane of $L^p(B^4,\mathcal{G})\times W^{1-1/p,p}(\partial B^4,\mathcal{G})$ made of couples (f,g) such that

$$\int_{B^4} f(x)\ dx^4 = -\int_{\partial B^4} g(y)\ \mathrm{dvol}_{\partial B^4}\ .$$

From the fact that $A\in\mathcal{V}^\varepsilon_C$ we deduce that $\|A\|_{L^4}\le C_4\sqrt{C\ \varepsilon}$ where C_4 is the Sobolev constant coming from the embedding into $L^4(B^4)$ of closed 3-forms on B^4 with adjoint exterior derivative in L^2 and whose restriction to ∂B^4 is zero. Hence for any V with average zero on B^4 we have

$$\|V\|_{W^{2,p}(B^4)}\le c\,[\|\partial_U\mathcal{N}^A(0,0)\cdot V\|_{\mathcal{F}}+c\,C_4\sqrt{C\ \varepsilon}\ \|dV\|_{L^{4p/4-p}}\ .$$

Using again the embedding of $W^{1,p}(\wedge^1B^4,\mathcal{G})$ into $L^{4p/4-p}(\wedge^1B^4,\mathcal{G})$ and denoting C_p the corresponding constant, we have then

$$[1-c\,C_4\sqrt{C\ \varepsilon}\ C_p]\ \|V\|_{W^{2,p}(B^4)}\le c\,\|\partial_U\mathcal{N}^A(0,0)\cdot V\|_{\mathcal{F}}\ .$$

Having chosen ε such that $c\,C_4\sqrt{C\ \varepsilon}\ C_p<1/2$ we have that $\partial_U\mathcal{N}^A(0,0)$ has zero kernel. A classical result from Calderòn–Zymund theory (see [16]) asserts that

$$\begin{aligned}\mathcal{L}\ :\ W^{2,p}(B^4,\mathcal{G})&\longrightarrow\mathcal{F}\\ V&\longrightarrow(-\Delta V,\partial_rV)\end{aligned}$$

is invertible and hence has zero index. By continuity of the index the maps

$$\begin{aligned}\mathcal{L}_t\ :\ W^{2,p}(B^4,\mathcal{G})&\longrightarrow\mathcal{F}\\ V&\longrightarrow(-\Delta V+t[A,dV],\partial_rV)\end{aligned}$$

have also zero index and since $\mathcal{L}_1=\partial_U\mathcal{N}^A(0,0)$ has trivial kernel it is invertible. So we can apply the implicit function theorem and there exists $\delta>0$ together with an open neighborhood $\mathcal{O}$ of 0 in the subspace of $W^{2,p}(B^4,\mathcal{G})$ with average 0 on B^4 such that

$$\forall\,\omega\in W^{1,p}(B^4,\mathcal{G})\quad\text{satisfying}\quad\|\omega\|_{W^{1,p}(S^4,\mathcal{G})}<\delta$$

$$\exists\ !\ V_\omega\in\mathcal{O}\quad\text{s.t.}\quad\mathcal{N}^A(V_\omega,\omega)=0\quad\text{and}\quad\int_{B^4}V_\omega=0\,,$$

and $\mathcal{O}$ can be taken smaller and smaller as δ tends to zero. We denote $g_\omega:=\exp(V_\omega)$.

It remains to establish the control of the L^p norm (resp. L^2 norm) of $d(A+\omega)^{g_\omega}$ in terms of the L^p norm (resp. L^2 norm) of $F_{A+\omega}$.

The Coulomb gauge $(A+\omega)^{g_\omega}$ satisfies for $q=2$ and $q=p$

$$\|d(A+\omega)^{g_\omega}\|_{L^q}\le\|F_{A+\omega}\|_{L^p}+\|(A+\omega)^{g_\omega}\wedge(A+\omega)^{g_\omega}\|_{L^p}\ .$$

We have

$$\|(A+\omega)^{g_\omega}\|_{L^4}\le C_4\ \|dA\|_{L^2}+\|\omega\|_{L^4}+\|dg^\omega\|_{L^4}\ .$$

Using the fact that A is the Coulomb gauge whose $W^{1,2}$ norm is controlled by the L^2 norm of F_A – which is assumed itself to be less than ε – by taking ε small enough – independently of A – by taking δ small enough – depending possibly on A – which ensures in particular that $\|dg_\omega\|_{L^4}$ is sufficiently small, and by using the embedding of closed forms with L^q exterior co-derivative and whose restriction to ∂B^4 is zero into $L^{4q/4-q}$ since $d^*(A+\omega)^{g_\omega}=0$ and $\iota^*_{\partial B^4}(A+\omega)^g=0$ we have established that

$$\|d(A+\omega)^{g_\omega}\|_{L^q}\le\|F_{A+\omega}\|_{L^q}+2^{-1}\|d(A+\omega)^{g_\omega}\|_{L^q}\,.$$

This implies that $A+\omega$ fulfills the conditions for being in $\mathcal{V}_C^\varepsilon$ for and ω satisfying $\|\omega\|_{W^{1,p}}<\delta$ where δ has also been taken small enough in such a way that $\|F_{A+\omega}\|<\varepsilon$. This concludes the proof of the openness of $\mathcal{V}_C^\varepsilon$ with respect to the $W^{1,p}$-topology for well-chosen constants $\varepsilon>0$ and $C>0$ and this concludes the proof of the **claim 1**.

End of the proof of Theorem IV.1. With the claim 1 at hand now, we are going to conclude the proof of Theorem IV.1

Let $A\in W^{1,2}(\wedge^1B^4,\mathcal{G})$ such that $\int_{B^4}|F_A|^2\,dvol_{B^4}<\varepsilon$. Since C^∞ is dense in $W^{1,2}$ there exists A^t a family of smooth 1-form on B^4 into $\mathcal{G}$ converging strongly to A in $W^{1,2}$ as t goes to zero. Using again the embedding of $W^{1,2}$ into L^4 we have the existence of $t_0>0$ such that

$$\forall\ t<t_0\qquad\int_{B^4}|F_{A^t}|^2\ \mathrm{dvol}_{B^4}<\varepsilon\,.$$

Thus A^t is in $\mathcal{U}^\varepsilon$ and, due to the claim 1, it is also in $\mathcal{V}_C^\varepsilon$. Let g^t such that $d^*(A^t)^{g^t}=0$ with

$$\int_{B^4}\left|d\left((A^t)^{g^t}\right)\right|^2\,dx^4\le C\int_{B^4}|F_{A^t}|^2\,dx^4\,.$$

Again, since there is no non-trivial closed and co-closed 3-form on B^4 the previous identity implies that $(A^t)^{g^t}$ is bounded in $W^{1,2}$ and then in L^4 too. The approximating connection 1-forms A^t are converging strongly to A in $W^{1,2}$ and hence in L^4 thus $d(g^t)$ is bounded in L^4. We then deduce the existence of a sequence $t_k\to0$ such that g^{t_k} converges weakly in $W^{1,4}(B^4,G)$ to some limit g^0. Using the Rellich–Kondrachov compactness theorem g^{t_k} converges strongly to g^0 in $L^p(B^4)$ for any $p<+\infty$ and hence g^0 is also taking values in G since we have almost everywhere convergence of the sequence. Using the previous convergences we deduce first that

$$(g^{t_k})^{-1}\,dg^{t_k}+(g^{t_k})^{-1}\,A\,g^{t_k}\rightharpoonup(g^0)^{-1}\,dg^0+(g^0)^{-1}\,A\,g^0\quad\text{ in }\mathcal{D}'(B^4)\,.$$

This implies that $d^*((A)^{g^0})=0$. Since both A^t and $(A^t)^{g^t}$ are bounded in $W^{1,2}$ and since g^t is bounded in L^4, using

$$dg^t=g^t\,(A^t)^{g^t}-A^t\,g^t\,,$$

we deduce that g^t is bounded in $W^{2,2}$ and hence the trace of g^{t_k} weakly converges to the trace of g^0 in $H^{3/2}(\partial B^4,G)$. So we can pass to the limit in the equation $\iota^*_{\partial B^4}*(A^{t_k})^{g^{t_k}}=0$ and we obtain

$$\iota^*_{\partial B^4}*(A^0)^{g^0}=0\,.$$

Finally since F_{A^t} is strongly converging to F_A in L^2, using also the lower semi-continuity of the L^2 norm together with the weak convergence of $(A^{t_k})^{g_{t_k}}$ towards A^{g^0} we have

$$\int_{B^4} |d(A^{g^0})|\, dx^4 \le C \int_{B^4} |F_A|^2 \, dx^4\,.$$

This concludes the proof of theorem IV.1. □

IV.2. A refinement of Uhlenbeck's Coulomb gauge extraction theorem

We have seen that Uhlenbeck's result is optimal in the sense that without assuming anything about the smallness of the Yang–Mills energy there is no hope to obtain a gauge of $W^{1,2}$ controlled energy. One might wonder however if the smallness of the L^2 norm of the curvature is the ultimate criterium for ensuring the existence of controlled Coulomb gauges. The answer is "no" and one can very slightly reduce this requirement. Recall the notion of weak L^2 quasi-norm. We say that a measurable function f on B^m is in the *weak L^2 space* if

$$|f|_{2,\infty} := \left[\sup_{\alpha>0} \alpha^2 \ |\{x\in B^m\ ;\ |f(x)|>\alpha\}|\right]^{1/2} < +\infty\,,$$

where $|\cdot|$ denotes the Lebesgue measure on B^m. This quantity defines a quasi-norm which is equivalent to a norm (see for instance [17]) that we denote $\|\cdot\|_{2,\infty}$. The *weak L^2 space* equipped with $\|\cdot\|_{2,\infty}$ is complete and define then a *Banach space* denoted $L^{2,\infty}$ called also *Marcinkiewicz weak L^2 space* or also *Lorentz weak L^2 space*. It is larger than L^2. Indeed, for any function $f\in L^2$ we have

$$\|f\|_{2,\infty} \le \sup_{\alpha>0} \int_{x\ ;|f|(x)>\alpha} |f|^2(x)dx \le \int |f|^2(x)\ dx = \|f\|_2^2\,.$$

It is strictly larger than L^2: the function $f(x) := |x|^{-m/2}$ is in $L^{2,\infty}(B^m)$ but not in $L^2(B^m)$. It is also not difficult to see that $L^{2,\infty}(B^m) \hookrightarrow L^p(B^m)$ for any $1\le p<2$. More generally we define the $L^{q,\infty}$ space of measurable functions f satisfying

$$|f|_{q,\infty} := \left[\sup_{\alpha>0} \alpha^q \ |\{x\in B^m\ ;\ |f(x)|>\alpha\}|\right]^{1/q} < +\infty\,.$$

This defines again a quasi-norm equivalent to a norm[3] if $q>1$. So it is a space which "sits" between $L^q(B^m)$ and all the $L^p(B^m)$ spaces for any $p<q$. This is a space which has the same scaling properties as L^q but has however the big advantage of containing the Riesz functions $|x|^{-m/q}$ which play a central role in the theory of elliptic PDE. As we will see later the space $L^{q,\infty}(B^m)$ has also the

[3] This is not true for $q=1$, the space L^1-weak cannot be made equivalent to a normed space – unfortunately, otherwise the analysis could make the economy of Calderòn–Zygmund theory and a major part of harmonic analysis that would suddenly become trivial. . . !

advantage of being the dual of a Banach space, the *Lorentz space* $L^{q',1}(B^m)$ of measurable functions f satisfying

$$\int_0^{+\infty} |\{x\ ;\ |f|(x) > \alpha\}|^{1/q'}\ d\alpha < +\infty\,, \tag{IV.9}$$

where $1/q' = 1 - 1/q$ (see [17]). This later space has very interesting "geometric" properties that will be useful for the analysis of the Yang–Mills Lagrangian as we will see below.

We have the following theorem

Theorem IV.4. *Let $m \leq 4$ and G be a compact Lie group. There exists $\varepsilon_G > 0$ and $C_G > 0$ such that for any $A \in W^{1,2}(B^m, \mathcal{G})$ satisfying*

$$\sup_{\alpha>0}\ \alpha^2\ |\{x \in B^m\ ;\ |F_A(x)| > \alpha\}| < \varepsilon_G\,, \tag{IV.10}$$

there exists $g \in W^{2,2}(B^m, G)$ such that

$$\left\{\begin{array}{l} \displaystyle\int_{B^m} |A^g|^2 + \sum_{i,j=1}^4 |\partial_{x_i} A^g_j|^2\, dx^m \leq C_G \int_{B^m} |dA + A\wedge A|^2\, dx^m \\ d^*A^g = 0 \qquad\quad in\ B^m \\ \iota^*_{\partial B^m}(*A^g) = 0\,, \end{array}\right. \tag{IV.11}$$

where $A^g = g^{-1}dg + g^{-1}Ag$ and $\iota_{\partial B^m}$ is the canonical inclusion map of the boundary of the unit ball into $\mathbb{R}^m$. Moreover we have also

$$\sum_{i,j=1}^4 \|\partial_{x_i} A^g_j\|^2_{2,\infty} \leq C_G\ \|F_A\|^2_{2,\infty}\,. \qquad \Box$$

The weakening of the smallness criterium by replacing *small* L^2 by the less restrictive *small* $L^{2,\infty}$ condition for the existence of a controlled Coulomb gauge has been first observed in [5]. This was a very precious observation for the control of the loss of energies in so-called *neck annular regions* in the study of conformally invariant problems such as *Willmore surfaces* or also *Yang–Mills Fields* as we will see below. The estimate (IV.10) comes naturally from the ϵ*-regularity property* which holds in neck regions.

Proof of Theorem IV.4. It follows exactly the same scheme as the proof of Theorem IV.4 but we will need to use interpolation spaces between $L^{q,\infty}$ and $L^{q,1}$, the *Lorentz spaces* $L^{q,s}$ and some of their properties.

Let $2 < p < 4$ and

$$\hat{\mathcal{U}}^\varepsilon := \left\{A \in W^{1,p}(B^4, \mathcal{G}) \quad \text{s.t.} \quad |F_A|^2_{L^{2,\infty}(B^4)} < \varepsilon\right\}\,,$$

and for any $\varepsilon > 0$ and $C > 0$ and we consider

$$\hat{\mathcal{V}}_C^\varepsilon := \left\{ \begin{array}{l} A \in \hat{\mathcal{U}}^\varepsilon \quad \text{s.t.} \quad \exists\, g \in W^{2,p}(B^4, G) \\ \displaystyle\int_{B^4} |dA^g|^p\, dx^4 \le C \int_{B^4} |F_A|^p\ dx^4 \\ \displaystyle\int_{B^4} |dA^g|^2\, dx^4 \le C \int_{B^4} |F_A|^2\ dx^4 \\ \|dA^g\|^2_{L^{2,\infty}(B^4)} \le C\ \|F_A\|^2_{L^{2,\infty}(B^4)} \\ \quad d^*A^g = 0 \quad \text{and} \quad \iota_{\partial B^4} * A^g = 0\,. \end{array} \right\}$$

The first goal is to show the following

Claim $\qquad \exists\, \varepsilon > 0 \quad C > 0 \quad \text{s.t.} \quad \hat{\mathcal{V}}_C^\varepsilon = \hat{\mathcal{U}}^\varepsilon\,.$

The proof of the claim is again divided in 3 steps.

Proof of the path connectedness of $\hat{\mathcal{U}}^\varepsilon$. For A in $\mathcal{U}^\varepsilon$ and $t \in [0,1]$ we define the connection form A^t to be the image of A by the dilation of rate t^{-1}: $A^t = t\sum_{j=1} A_j(t\,x)\ dx_j$. We have in particular

$$F_{A^t} = t^2 \sum_{i,j=1}^4 (F_A)_{ij}(x)\ dx_i\, dx_j\ ,$$

hence $|F_{A^t}|(x) = t^2\ |F_A|(t\,x)$ and

$$|F_{A_t}|_{L^{2,\infty}(B^4)} = |F_A|_{L^{2,\infty}(B^4_t)} \le |F_A|_{L^{2,\infty}(B^4)} < \varepsilon$$

and[4] this path connects A to 0 in the $W^{1,p}$ topology due to (IV.2). Hence this concludes the proof of the path connectedness of $\hat{\mathcal{U}}^\varepsilon$.

The *proof of the closeness of $\hat{\mathcal{V}}_C^\varepsilon$ in $\hat{\mathcal{U}}^\varepsilon$* is identical to the proof of the closeness of $\mathcal{V}_C^\varepsilon$ in $\mathcal{U}^\varepsilon$.

Proof of the openness of $\hat{\mathcal{V}}_C^\varepsilon$ in $\hat{\mathcal{U}}^\varepsilon$. We consider the map $\mathcal{N}^A$ defined by (IV.8). We recall[5] the definition of the space $L^{q,s}(B^m)$ where $1 < q < \infty$ and $1 \le s < +\infty$. A measurable function f on B^m belongs to $L^{q,s}(B^m)$ if

$$|f|_{q,s} := \left[\int_0^\infty t^{\frac{s}{q}}\ f^*(t)\ \frac{dt}{t}\right]^{1/s} < +\infty\,, \tag{IV.12}$$

where $f^*(t)$ is the decreasing rearrangement function associated to f, defined on $\mathbb{R}_+$, and satisfying

$$\forall \alpha > 0 \qquad |\{t > 0\ ;\ f^*(t) > \alpha\}| = |\{x \in B^m\ ;\ |f|(x) > \alpha\}|\ .$$

[4]This last inequality illustrates what we meant at the beginning of this subsection by $L^{2,\infty}$ *has the same scaling properties as* L^2.

[5]For a more thorough presentation of the Lorentz spaces and its interaction with Calderòn–Zygmund theory in particular the reader is invited to consult the first chapter of [17] as well as [53] or [55].

This defines again a quasi-norm equivalent to a norm for which the space is complete (see [17]). One verifies that the space $L^{q,1}(B^m)$ defined by (IV.9) coincides with the space given by (IV.12) for $s = 1$. One verifies also that $L^{q,q}(B^m) = L^q(B^m)$ and that for any $q \in (1, +\infty)$ and any $1 \leq s < \sigma \leq +\infty$ we have $L^{q,s}(B^m) \hookrightarrow L^{q,\sigma}(B^m)$. We have also that $\forall q < r$ and $\forall t, s \in [1, \infty]$ the following continuous embedding holds $L^{p,s}(B^m) \hookrightarrow L^{q,t}(B^m)$. The following multiplication rules holds and are continuous bilinear mappings in the corresponding spaces with the corresponding estimates

$$L^{p,s} \cdot L^{q,t} \hookrightarrow L^{r,\sigma}, \tag{IV.13}$$

where $r^{-1} = p^{-1} + q^{-1}$ and $s^{-1} + t^{-1} = \sigma^{-1}$ and where $1 < p, q < +\infty$ such that $r \geq 1$ and $1 \leq s, t \leq \infty$ such that $1 \leq \sigma \leq +\infty$. In particular we have for any $2 \leq p < 4$

$$L^{4,\infty} \cdot L^{\frac{4p}{4-p},p} \hookrightarrow L^p. \tag{IV.14}$$

Before to move on with the proof of Theorem IV.4 we shall need a last tool from function theory: the *improved Sobolev embeddings* (see [55]). For $1 \leq p < m$ the following embedding is continuous

$$W^{1,p}(B^m) \hookrightarrow L^{\frac{m\,p}{m-p},p}(B^m), \tag{IV.15}$$

and more generally for any $t \in [1, +\infty]$

$$W^{1,(p,t)}(B^m) \hookrightarrow L^{\frac{m\,p}{m-p},t}(B^m), \tag{IV.16}$$

where $W^{1,(p,t)}(B^m)$ denotes the space of measurable functions on B^m with distributional derivative in the Lorentz space $L^{p,t}(B^m)$.

Proof of openness of $\hat{\mathcal{V}}^\varepsilon_C$ continued. Using Calderòn–Zygmund $L^{q,t}$ theory we have the following bound

$$\begin{aligned}
\|V\|_{W^{2,p}(B^4)} &\leq c\, \left[\|\Delta V\|_{L^p(B^4)} + \|\partial_r V\|_{W^{1-1/p,p}(\partial B^4)}\right] \\
&\leq c\, \left[\|\partial_U \mathcal{N}^A(0,0) \cdot V\|_{\mathcal{F}} + \|[A, dV]\|_{L^p(B^4)}\right] \\
&\leq c\, \left[\|\partial_U \mathcal{N}^A(0,0) \cdot V\|_{\mathcal{F}} + c\, \|A\|_{L^{4,\infty}(B^4)}\, \|dV\|_{L^{4p/4-p,p}(B^4)}\right],
\end{aligned}$$

where $\mathcal{F} := W^{1,p}(B^4, \mathcal{G}) \times W^{1-1/p,p}(\partial B^4, \mathcal{G})$. From the fact that $A \in \mathcal{V}^\varepsilon_C$ we deduce that $\|A\|_{L^{4,\infty}} \leq C_4 \sqrt{C\, \varepsilon}$ where C_4 is the Sobolev constant coming from the embedding into $L^4(B^4)$ of closed 3 forms on B^4 with adjoint exterior derivative in L^2 and whose restriction to ∂B^4 is zero. Hence for any V with average zero on B^4 we have

$$\|V\|_{W^{2,p}(B^4)} \leq c\, [\|\partial_U \mathcal{N}^A(0,0) \cdot V\|_{\mathcal{F}} + c\, C_4 \sqrt{C\, \varepsilon}\, \|dV\|_{L^{4p/4-p,p}}.$$

Using again the embedding (IV.15) and denoting C_p the corresponding constant, we have then

$$[1 - c\, C_4 \sqrt{C\, \varepsilon}\, C_p]\, \|V\|_{W^{2,p}(B^4)} \leq c\, \|\partial_U \mathcal{N}^A(0,0) \cdot V\|_{\mathcal{F}}.$$

Having chosen ε such that $c\, C_4 \sqrt{C\, \varepsilon}\, C_p < 1/2$ we have that $\partial_U \mathcal{N}^A(0,0)$, which is again a Fredholm operator of index zero, has a trivial kernel and is hence invertible.

The rest of the proof is completed by easily transposing to our present setting each argument of the case of small L^2 Yang–Mills energy which was detailed in the previous subsection.

IV.3. Controlled gauges without small energy assumption

One might wonder why a $W^{1,2}$ control is wished and why one could not give up a bit our requirements and look for some control of a "weaker norm". This is indeed possible, together with Mircea Petrache [33], the author proved the existence of global gauges A^g whose $L^{4,\infty}$ norm is controlled by the Yang–Mills energy which is not necessarily small. Precisely we have.

Theorem IV.5. *Let (M^4, g) be a Riemannian 4-manifold. There exists a function $f : \mathbb{R}^+ \to \mathbb{R}^+$ with the following properties.*
Let ∇ be a $W^{1,2}$ connection over an $SU(2)$-bundle over M. Then there exists a global $W^{1,(4,\infty)}$ section of the bundle (possibly allowing singularities) over the whole M^4 such that in the corresponding trivialization ∇ is given by $d + A$ with the following bound.

$$\|A\|_{L^{(4,\infty)}(M^4)} \le f\left(\|F_\nabla\|_{L^2(M^4)}\right) , \tag{IV.17}$$

where F_∇ is the curvature form of ∇. □

The following question is still unsolved.

Open Problem: *Under the same assumptions as the ones of Theorem* IV.5, *find $A \in L^{4,\infty}(M^4)$ such that* (IV.17) *holds and*

$$d^*A \equiv 0 \,.$$ □

Observe that $F_A \in L^2(M^4)$, $A \in L^{4,\infty}(M^4)$ together with $d^*A = 0$ imply $A \in W^{1,(2\infty)}(M^4)$.

Having the existence a global representative with controlled $W^{1,(2,\infty)}$-norm and satisfying the Coulomb condition could be useful for studying the Yang–Mills flow in four dimensions. A partial positive answer to the above open problem is given by Yu Wang in [60].

The proof of Theorem IV.5 is deduced from the following existence result which is a global counterpart of Theorem IV.2.

Theorem IV.6 ([33]). *Let G be a compact lie group and u be a map in $W^{1,n}(S^n, G)$ then there exists an extension $\tilde{u}$ of u in $W^{1,(n+1,\infty)}(B^{n+1}, G)$ and satisfying*

$$\sup_{\alpha>0} \alpha^{n+1} \left|\left\{x \in B^{n+1} \quad ; \quad |\nabla\tilde{u}| > \alpha\right\}\right| \le \gamma_n\left(\int_{S^n} |\nabla u|^n \ d\mathrm{vol}_{S^n}\right) ,$$

where γ_n is a universal function. □

Here again, there are counterexample to the existence of an extension in the "slightly" smaller space $W^{1,n+1}(B^{n+1}, G)$. In fact it is proved in [36] that the previous result does not require a Lie group structure in the target to be true: it extends to general closed target Riemannian manifold in general.

V. The resolution of the Yang–Mills Plateau problem in the critical dimension

V.1. The small energy case

We first present the resolution of the *Yang–Mills Plateau* problem in the case where the given connection at the boundary has a small trace norm. Precisely we shall prove the following result.

Theorem V.1. *Let G be a compact Lie group and $m \le 4$. There exists $\delta_G > 0$ such that for any 1-form $\eta \in H^{1/2}(\wedge^1 \partial B^m, \mathcal{G})$ satisfying*

$$\|\eta\|_{H^{1/2}(\partial B^m)} < \delta_G\,, \tag{V.1}$$

then the minimization problem

$$\inf\left\{YM(A) = \int_{B^m} |dA + A\wedge A|^2\; dx^m \quad ; \quad \iota^*_{\partial B^m} A = \eta\right\}\,.$$

is achieved by a 1-form $A^0 \in W^{1,2}(\wedge^1 B^m, \mathcal{G})$ □

The previous theorem is a corollary of the following weak closure theorem

Theorem V.2. *Let G be a compact Lie group and $m \le 4$. There exists $\delta_G > 0$ such that for any 1-form $\eta \in H^{1/2}(\wedge^1 \partial B^m, \mathcal{G})$ satisfying*

$$\|\eta\|_{H^{1/2}(\partial B^m)} < \delta_G\,, \tag{V.2}$$

then for any $A^k \in W^{1,2}(\wedge^1 B^m, \mathcal{G})$ satisfying

$$\limsup_{k\rightarrow+\infty} YM(A^k) = \int_{B^m} |dA^k + A^k\wedge A^k|^2\; dx^m < +\infty \quad \textit{and} \quad \iota^*_{\partial B^m} A^k = \eta\,,$$

there exists a subsequence $A^{k'}$ and a Sobolev connection $A^\infty \in W^{1,2}(\wedge^1 B^m, \mathcal{G})$ such that

$$D(A^{k'}, A^\infty) := \inf_{g\in W^{2,2}(B^4,G)} \int_{B^m} |A^{k'} - (A^\infty)^g)|^2\; dx^m \longrightarrow 0\,,$$

moreover

$$YM(A^\infty) \le \liminf_{k'\rightarrow 0} YM(A^{k'}) \quad \textit{and} \quad \iota^*_{\partial B^m} A^\infty = \eta\,.$$ □

Proof of Theorem V.2. We present the proof in the critical case $m = 4$. The case $m < 4$ being almost like the Abelian linear case treated. Let B be the minimizer of E in $W^{1,2}_\eta(\wedge^1 B^4, \mathcal{G})$ and using (III.7) and the Sobolev embedding $W^{1,2}(B^4)$ into $L^4(B^4)$ we have

$$\left[\int_{B^4} |B|^4\; dx^4\right]^{\frac{1}{2}} + \sum_{i,j=1}^m \int_{B^4} |\partial_{x_i} B_j|^2\; dx^4 \le C\left[E(B) + \|\eta\|^2_{H^{1/2}}\right]\,. \tag{V.3}$$

The 1-form B is the harmonic extension of η and classical elliptic estimate gives

$$E(B) \le C\,\|\eta\|^2_{H^{1/2}(\partial B^4)}\,. \tag{V.4}$$

Combining (V.3) and (V.4) we obtain the existence of a constant C independent of η such that

$$\int_{B^4} |F_B|^2 \, dx^4 \leq C \, \left[\|\eta\|^2_{H^{1/2}(\partial B^4)} + \|\eta\|^4_{H^{1/2}(\partial B^4)}\right] .$$

We choose first $\delta_G > 0$ such that $C\ [\delta_G^2 + \delta_G^4] < \varepsilon_G$ in such a way that we can apply Theorem IV.1 and we have the existence of a minimizing sequence A^k of YM in $W^{1,2}_\eta(\wedge^1 B^4, \mathcal{G})$ with a Coulomb gauge $(A^k)^{g^k}$ controlled in $W^{1,2}$:

$$\|(A^k)^{g^k}\|_{W^{1,2}(B^4)} \leq \ C \ \|F_{A^k}\|_{L^2(B^4)} \leq C\ [\delta_G + \delta_G^2] . \tag{V.5}$$

Without loss of generality we can assume that

$$(A^k)^{g^k} \rightharpoonup \hat{A}^\infty \quad \text{weakly in } W^{1,2}(\wedge^1 B^4, \mathcal{G}) ,$$

for some 1-form $\hat{A}^\infty$ which satisfies the Coulomb condition $d^*\hat{A}^\infty = 0$ and for which

$$\int_{B^4} |F_{\hat{A}^\infty}|^2 \, dx^4 \leq \liminf_{k\to+\infty} \int_{B^4} |F_{(A^k)^{g^k}}|^2 \, dx^4 = \liminf_{k\to+\infty} \int_{B^4} |F_{A^k}|^2 \, dx^4 . \tag{V.6}$$

We claim that the restriction of $\hat{A}^\infty$ to ∂B^4 is gauge equivalent to η. Because of the weak convergence of $(A^k)^{g^k}$ to $\hat{A}^\infty$ weakly in $W^{1,2}$, by continuity of the trace operation from $W^{1,2}$ into $H^{1/2}$ we have

$$\iota^*_{\partial B^4}(A^k)^{g^k} = (g^k)^{-1}\, \iota^*_{\partial B^4} dg^k + (g^k)^{-1}\, \eta\, g^k \rightharpoonup \iota^*_{\partial B^4}\hat{A}^\infty \tag{V.7}$$

weakly in $H^{1/2}(\wedge^1 \partial B^4, \mathcal{G})$. Using the continuous embedding

$$H^{1/2}(\partial B^4) \hookrightarrow L^3(\partial B^4) ,$$

we have that the restriction of g^k to ∂B^4 converges weakly to some limit g^∞ in $W^{1,3}(\partial B^4)$ and we have, using (V.5),

$$\begin{aligned} \|dg^\infty\|_{L^3(\partial B^4)} &\leq \liminf_{k\to+\infty} \|dg^k\|_{L^3(\partial B^4)} \\ &\leq C \ \left[\|\eta\|_{H^{1/2}(\partial B^4)} + \liminf_{k\to+\infty} \|(A^k)^{g^k}\|_{H^{1/2}(\partial B^4)}\right] \\ &\leq C\ [\delta_G + \delta_G^2] . \end{aligned} \tag{V.8}$$

Using now the Rellich–Kondrachov theorem (see, for instance, [7]), this convergence is strong in L^q for any $q < +\infty$ which implies that g^∞ takes values almost everywhere in G and $g^\infty \in W^{1,3}(\partial B^4, G)$. We have moreover

$$(g^\infty)^{-1}\, dg^\infty + (g^\infty)^{-1}\, \eta\, g^\infty = \iota^*_{\partial B^4}\hat{A}^\infty .$$

Using the continuous embedding

$$L^\infty \cap W^{1,3}(\partial B^4) \cdot H^{1/2}(\partial B^4) \hookrightarrow H^{1/2}(\partial B^4)$$

(the proof of this continuous embedding is also similar to the one of Lemma B1 in [33]), we have that

$$\begin{aligned}
&\|dg^\infty\|_{H^{1/2}(\partial B^4)}\\
&\quad\le C\ \left[\|g^\infty\|_\infty + \|g^\infty\|_{W^{1,3}(\partial B^4)}\right]\left[\|\eta\|_{H^{1/2}(\partial B^4)} + \|\iota^*_{\partial B^4}\hat{A}^\infty\|_{H^{1/2}(\partial B^4)}\right] \qquad \text{(V.9)}\\
&\quad\le C\ [\delta_G + \delta_G^2]\,.
\end{aligned}$$

We shall now make use of the following theorem which, as for Theorem IV.2 can be proved following Uhlenbeck's Coulomb gauge extraction method.

Theorem V.3. *Let G be a compact Lie group. There exists $\varepsilon_G > 0$ such that for any $g \in H^{3/2}(\partial B^3, G)$ satisfying*

$$\|g\|_{H^{3/2}(\partial B^4,G)} < \varepsilon_G\,,$$

there exists an extension $\tilde{g} \in W^{2,2}(B^4, G)$ of g satisfying

$$\|\tilde{g}\|_{W^{2,2}(B^4,G)} \le C\ \|g\|_{H^{3/2}(\partial B^4,G)}\,. \qquad \square$$

End of the proof of Theorem V.2. We choose δ_G small enough such that the r.h.s. of (V.9) $C\ [\delta_G + \delta_G^2]$ is smaller than ε_G given by the previous theorem. Let $\tilde{g}^\infty \in W^{2,2}(B^4, G)$ be an extension of g^∞ given by Theorem V.3. Then

$$A^\infty := (\hat{A}^\infty)^{(\tilde{g}^\infty)^{-1}} \in W^{1,2}_\eta(\wedge^1 B^4, \mathcal{G})$$

and we have using (V.6)

$$\int_{B^4} |F_{A^\infty}|^2\ dx^4 \le \liminf_{k\to+\infty} \int_{B^4} |F_{A^k}|^2\ dx^4\,.$$

Since A^k is a minimizing sequence of the Yang–Mills Plateau problem in $W^{1,2}_\eta(\wedge^1 B^4, \mathcal{G})$, the connection form A^∞ is a solution to this problem and Theorem V.2 is proved. $\square$

V.2. The general case and the point removability result for $W^{1,2}$ Sobolev connections

Theorem V.2 as it is stated does not hold without the small norm assumption (V.1) this is due to the fact that Theorem V.3 and similar results such as Theorem IV.2 do not hold for general data without smallness assumption (see again Remark IV.2). We shall instead prove the following result where the boundary condition is relaxed to a constrained trace modulo gauge action.

Theorem V.4. *Let G be a compact Lie group and $m \le 4$. For any 1-form $\eta \in H^{1/2}(\wedge^1 \partial B^m, \mathcal{G})$ the following minimization problem*

$$\inf\left\{\int_{B^m} |F_A|^2\ dx^m\ ;\ \iota^*_{\partial B^m} A = \eta^g \quad \textit{for some } g \in H^{3/2}(\partial B^4, G)\right\} \qquad \text{(V.10)}$$

is achieved by a 1-form $A^0 \in W^{1,2}(\wedge^1 B^m, \mathcal{G})$. $\square$

In fact Theorem V.4 is a corollary of a general closure result.

Theorem V.5. *For any compact Lie group G and any dimension $m \leq 4$, the space of Sobolev connections*

$$\mathfrak{A}_\eta(B^m) := \left\{ A \in W^{1,2}(B^4, \mathcal{G}) \ ; \ \iota^*_{\partial B^m} A = \eta^g \quad \textit{for some } g \in H^{3/2}(\partial B^4, G) \right\}$$

is weakly sequentially closed for sequences of controlled Yang–Mills energy. Precisely, for any $A^k \in \mathfrak{A}_\eta(B^m)$ satisfying

$$\limsup_{k \to +\infty} YM(A^k) = \int_{B^m} |dA^k + A^k \wedge A^k|^2 \ dx^m < +\infty \, ,$$

there exists a subsequence $A^{k'}$ and a Sobolev connection $A^\infty \in \mathfrak{A}_\eta(B^m)$ such that

$$d(A^{k'}, A^\infty) := \inf_{g \in W^{1,2}(B^4, G)} \int_{B^m} |A^{k'} - (A^\infty)^g|^2 \ dx^m \longrightarrow 0 \, ,$$

moreover

$$YM(A^\infty) \leq \liminf_{k' \to 0} YM(A^{k'}) \, . \qquad \square$$

Proof of Theorem V.5. Here again we restrict to the most delicate case: $m = 4$.

Let A^k be a sequence of $\mathcal{G}$-valued 1-forms and denote by ε_G the positive constant in Uhlenbeck's theorem IV.1[6]. A straightforward covering argument combined by some induction procedure gives the existence of a subsequence that we still denote A^k and N points $p_1 \cdots p_N$ in $\overline{B^4}$ such that

$$\forall \, \delta > 0 \quad \exists \, \rho_\delta > 0$$

$$\sup_{k \in \mathbb{N}} \sup \left\{ \int_{B_{\rho_\delta}(y) \cap B^4} |F_{A^k}|^2 \ dx^4 \ ; \ y \in B^4 \setminus \cup_{l=1}^N B_\delta(p_l) \right\} < \varepsilon_G \, .$$

The case without concentration: $\{p_1 \cdots p_N\} = \emptyset$.

Let $\rho > 0$ such that

$$\sup_{k \in \mathbb{N}} \sup_{y \in B^4} \left\{ \int_{B_\rho(y)} |F_{A^k}|^2 \ dx^4 \right\} < \varepsilon_G \, .$$

We fix a finite *good covering*[7] of B^4 by balls of radius $\rho/2$. Denote $\{B_{\rho/2}(x_i)\}_{i \in I}$ this covering. On each of the rice larger ball $B_\rho(x_i)$ for any $k \in \mathbb{N}$ we take a controlled Coulomb gauge $(A^k)^{g_i^k}$ such that

$$\left[\int_{B^4_\rho(x_i)} |(A^k)^{g_i^k}|^4 \ dx^4 \right]^{\frac{1}{2}} + \sum_{l,j=1}^m \int_{B^4_\rho(x_i)} |\partial_{x_l}((A)^{g_i^k})_j|^2 \ dx^4 \leq \ C \int_{B_\rho(x_i)} |F_{A^k}|^2 \ dx^4 \, , \tag{V.11}$$

and

$$d^*(A^k)^{g_i^k} = 0 \, . \tag{V.12}$$

[6]We choose in fact ε_G small enough for the controlled gauge Uhlenbeck theorem to be valid for this constant when the domain is any intersection of B^4 with a ball $B_\rho(y)$ for $y \in \overline{B^4}$ and $0 < \rho < 1$

[7]The word "good" means that any intersections of elements of the covering is either empty or diffeomorphic to B^4 (see [6]).

For any pair $i \neq j$ in I such that $B_\rho(x_i) \cap B_\rho(x_i) \neq \emptyset$ we denote

$$g_{ij}^k := g_i^k (g_j^k)^{-1} \in W^{2,2}(B_\rho(x_i) \cap B_\rho(x_i), G)\,,$$

and we have in particular

$$(A^k)^{g_j^k} = (g_{ij}^k)^{-1} dg_{ij}^k + (g_{ij}^k)^{-1}\,(A^k)^{g_i^k}\, g_{ij}^k\,. \tag{V.13}$$

Observe that for any triplet $i \neq j$, $j \neq l$ and $i \neq l$ such that $B_\rho(x_i) \cap B_\rho(x_i) \cap B_\rho(x_l) \neq \emptyset$ we have the co-cycle condition

$$\forall\, k \in \mathbb{N} \qquad g_{ij}^k\; g_{jl}^k = g_{il}^k\,. \tag{V.14}$$

Combining (V.11) and (V.13) together with the improved Sobolev embedding $W^{1,2}(B^4) \hookrightarrow L^{4,2}(B^4)$ where $L^{4,2}$ is the Lorentz interpolation space given by (IV.12) we obtain that for any pair $i \neq j$ such that $B_\rho(x_i) \cap B_\rho(x_j) \neq \emptyset$

$$\|dg_{ij}^k\|^2_{L^{4,2}(B_\rho(x_i) \cap B_\rho(x_j))} \leq C \int_{B_\rho(x_i) \cup B_\rho(x_j)} |F_{A^k}|^2\; dx^4\,. \tag{V.15}$$

From (V.13) we have

$$-\Delta g_{ij}^k = (A^k)^{g_i^k} \cdot dg_{ij}^k - dg_{ij}^k \cdot (A^k)^{g_j^k}\,. \tag{V.16}$$

Using again the improved Sobolev embedding $W^{1,2}(B^4) \hookrightarrow L^{4,2}(B^4)$, inequalities (V.11) and (V.15) together with the continuous embedding

$$L^{4,2} \cdot L^{4,2} \hookrightarrow L^{2,1}\,,$$

we obtain

$$\|\Delta g_{ij}^k\|_{L^{2,1}(B_\rho(x_i) \cap B_\rho(x_i))} \leq C \int_{B_\rho(x_i) \cup B_\rho(x_j)} |F_{A^k}|^2\; dx^4\,. \tag{V.17}$$

Using Calderòn–Zygmund theory in Lorentz interpolation spaces (see [53]) we obtain that $g_{ij}^k \in W^{2,(2,1)}_{\text{loc}}(B_\rho(x_i) \cap B_\rho(x_i))$ where $W^{2,(2,1)}$ denotes the space of functions with two derivatives in $L^{2,1}$ and using (V.15) together with (V.17) we obtain the following estimate

$$\|\nabla^2 g_{ij}^k\|_{L^{2,1}(B_{3\rho/4}(x_i) \cap B_{3\rho/4}(x_i))} \leq C \left[\int_{B_\rho(x_i) \cup B_\rho(x_j)} |F_{A^k}|^2\; dx^4\right]^{1/2}\,. \tag{V.18}$$

We can then extract a subsequence such that

$$\begin{cases} \forall i \in I \quad (A^k)^{g_i^k} \rightharpoonup A^{i,\infty} & \text{weakly in } W^{1,2}(B_\rho(x_i)) \\ \forall i \neq j \qquad g_{ij}^k \rightharpoonup g_{ij}^\infty & \text{weakly in } W^{2,(2,1)}(B_{3\rho/4}(x_i) \cap B_{3\rho/4}(x_i)))\,, \end{cases}$$

moreover $A^{i,\infty}$ and g_{ij}^∞ satisfy the following identities

$$\begin{cases} \forall i \neq j \qquad A^{j,\infty} = (g_{ij}^\infty)^{-1} dg_{ij}^\infty + (g_{ij}^\infty)^{-1}\, A^{i,\infty}\, g_{ij}^\infty \\ \forall i,j,l \quad g_{ij}^\infty\; g_{jl}^\infty = g_{il}^\infty\,, \end{cases} \tag{V.19}$$

and we have the following estimate

$$\|\nabla^2 g_{ij}^\infty\|^2_{L^{2,1}(B_{3\rho/4}(x_i)\cap B_{3\rho/4}(x_j))} \le C \liminf_{k\to+\infty} \int_{B_\rho(x_i)\cup B_\rho(x_j)} |F_{A^k}|^2 \, dx^4 \,. \tag{V.20}$$

It is proved[8] in [39] that

$$W^{2,(2,1)}(B^4) \hookrightarrow C^0(B^4)\,,$$

hence we deduce that $g_{ij}^\infty \in C^0 \cap W^{2,(2,1)}(B_{3\rho/4}(x_i)\cap B_{3\rho/4}(x_i))$ and for any $i \neq j$ there exists $\overline{g_{ij}^\infty} \in G$ such that

$$\begin{aligned}\|g_{ij}^\infty - \overline{g_{ij}^\infty}\|^2_{L^\infty(B_{3\rho/4}(x_i)\cap B_{3\rho/4}(x_j))} &\le C \liminf_{k\to+\infty} \int_{B_\rho(x_i)\cup B_\rho(x_j)} |F_{A^k}|^2 \, dx^4 \\ &< 2\, C\, \varepsilon_G \,.\end{aligned} \tag{V.21}$$

Taking ε_G small enough there exists a unique lifting

$$U_{ij}^\infty \in W^{2,(2,1)}(B_{3\rho/4}(x_i)\cap B_{3\rho/4}(x_j))$$

such that

$$\forall i \neq j \qquad g_{ij}^\infty = \overline{g_{ij}^\infty}\, \exp(U_{ij}^\infty)\,,$$

and

$$\|U_{ij}^\infty\|_\infty \le C\, \varepsilon_G\,,$$

for some constant C depending only on G. Following an induction argument[9] such as the one followed in [28] for the proof of Theorem II.11, we can smooth the U_{ij}^∞ in order to produce a sequence

$$g_{ij}^\infty(t) \in C^\infty(B_{3\rho/4}(x_i)\cap B_{3\rho/4}(x_j), G)$$

satisfying

$$g_{ij}^\infty(t) \longrightarrow g_{ij}^\infty \qquad \text{strongly in } W^{2,(2,1)}(B_{3\rho/4}(x_i)\cap B_{3\rho/4}(x_j)) \quad \text{as } t\to 0\,,$$

and

$$\forall\, t \quad \forall i,j,l \qquad g_{ij}^\infty(t)\, g_{jl}^\infty(t) = g_{il}^\infty(t)\,.$$

Since the ball B^4 is topologically trivial, the previous *cocycle condition* defines a trivial *Čech smooth co-chain* for the *presheaf* of G-valued smooth functions (see

[8]This embedding can be proved on $\mathbb{R}^4$ as follows. Let u be a *Schwartz function* in $\mathcal{S}(\mathbb{R}^4)$. Then the following representation formula holds

$$u(x) = C\,|x|^{-2} \star \Delta u\,,$$

where $C\,|x|^{-2}$ is the Green Kernel of the *Laplace operator*. From this identity one deduce the following inequality

$$\|u\|_{L^\infty} \le C\, \||x|^{-2}\|_{L^{2,\infty}}\, \|\Delta u\|_{L^{2,1}}\,,$$

and the embedding $W^{2,(2,1)}(\mathbb{R}^4) \hookrightarrow C^0(\mathbb{R}^4)$ follows from a density argument.

[9]A co-cycle smoothing argument by induction argument is also proposed in [19] under the weaker hypothesis that the co-cycles g_{ij}^∞ are $W^{1,4}$ in four dimension. This is made possible due to the fact that $C^\infty(B^4, G)$ is dense in $W^{1,4}(B^4, G)$ (see [47]). The works of Takeshi Isobe [19], [20] are proposing a framework for studying the analysis of gauge theory in conformal and super-critical dimension.

for instance [6] Section 10 Chapter II) and for any $i \in I$ and any $t > 0$ there exists $\rho_i(t) \in C^\infty(B_{3\rho/4}(x_i), G)$ such that

$$g_{ij}^\infty(t) = \rho_i(t)\,\rho_j(t)^{-1}\,. \tag{V.22}$$

We shall now make use of the following technical lemma which is proved in [59].

Lemma V.1. *Let G be a compact Lie group and $\{U_i\}_{i\in I}$ be a good covering of B^4. There exists $\delta > 0$ such that for any pair of co-chains*

$$\forall i \neq j \qquad h_{ij}, g_{ij} \in W^{2,2} \cap C^0(U_i \cap U_j, G)\,,$$

satisfying

$$\forall i,j,l \qquad g_{ij}\ g_{jl} = g_{il} \quad \text{ and } \quad h_{ij}\ h_{jl} = h_{il} \quad \text{ in } U_i \cap U_j \cap U_l\,.$$

Assume

$$\forall i \neq j \qquad \|g_{ij}^{-1}\, h_{ij} - e\|_{L^\infty(U_i\cap U_j)} < \delta\,,$$

where e is the constant map equal to the neutral element of G, then, for any strictly smaller good covering of B^4 $\{V_i\}_{i\in I}$ satisfying $\overline{V_i} \subset U_i$, there exists a family of maps $\sigma_i \in W^{2,2} \cap C^0(U_i \cap U_j, G)$ such that

$$\forall i \neq j \qquad h_{ij} = (\sigma_i)^{-1} g_{ij}\, \sigma_j \qquad \text{ in } V_i \cap V_j\,. \qquad \square$$

We apply the previous lemma to $h_{ij} := g_{ij}^\infty(t)$ and $g_ij := g_{ij}^\infty$ for t small enough and we deduce the existence of

$$\sigma_i(t) \in W^{2,2} \cap C^0(B_{\rho/2}(x_i) \cap B_{\rho/2}(x_j))\,,$$

such that

$$\forall i \neq j \qquad g_{ij}^\infty(t) = \sigma_i(t)^{-1}\, g_{ij}^\infty\, \sigma_j(t) \qquad \text{ in } B_{\rho/2}(x_i) \cap B_{\rho/2}(x_j)\,. \tag{V.23}$$

Combining (V.22) and (V.23) we have

$$\forall i \neq j \qquad g_{ij}^\infty = \sigma_i \rho_i\, (\sigma_j \rho_j)^{-1} \qquad \text{ in } B_{\rho/2}(x_i) \cap B_{\rho/2}(x_j)\,.$$

Combining this identity with (V.19) we set

$$A^0 := (\sigma_i\rho_i)^{-1}\, d(\sigma_i\rho_i) + (\sigma_i\rho_i)^{-1}\, A^{i,\infty}\, (\sigma_i\rho_i) \quad \text{ in } B_{\rho/2}(x_i)\,.$$

Clearly A^0 extends to a $W^{1,2}$ $\mathcal{G}$-valued 1-form in B^4, moreover, following the arguments in the proof of Theorem V.2, the restriction of A^0 to ∂B^4 is gauge equivalent to η. This concludes the proof of Theorem V.5 in the absence of concentration points.

The general case with possible concentration: $\{p_1 \cdots p_N\} \neq \emptyset$. Following the arguments in the previous case, for any $\delta > 0$ we exhibit a subsequence $A^{k'}$, a covering by balls $B_{\rho_\delta}(x_i)$ of $B^4 \setminus \cup_{l=1}^N B^4_\delta(p_l)$ and a family of gauge changes g_i^k such that

$$\begin{cases} \forall i \in I \quad (A^k)^{g_i^k} \rightharpoonup A^{i,\infty} & \text{weakly in } W^{1,2}(B_\rho(x_i)) \\ \forall i \neq j \qquad\quad g_{ij}^k \rightharpoonup g_{ij}^\infty & \text{weakly in } W^{2,(2,1)}(B_{3\rho/4}(x_i) \cap B_{3\rho/4}(x_i)))\,. \end{cases}$$

The family g_{ij}^∞ defines again a $W^{2,(2,1)}$-co-chain that we can approximate in $C^0 \cap W^{2,2}$ by a smooth one $g_{ij}^\infty(t)$. Using the fact that the second homotopy group of the compact Lie group is trivial $\pi_2(G) = 0$ (see, for instance, [8], Chapter V, Proposition 7.5) we deduce that the co-chain $g_{ij}^{\infty}(t)$, defined on a covering of $B^4 \setminus \cup_{l=1}^N B^4_\delta(p_l)$. is trivial for the Čech cohomology for the co-chains on the presheaf of smooth G-valued functions. Following each step of the above argument we construct a $W^{1,2}$ $\mathcal{G}$-valued 1-form A^0 in $B^4 \setminus \cup_{l=1}^N B^4_\delta(p_l)$ which is gauge equivalent to $A^{i,\infty}$ in $B^4_{\rho/2}(x_i)$ for each $i \in I$ and whose restriction on $\partial B^4 \setminus \cup_{l=1}^N B^4_\delta(p_l)$ is also gauge equivalent to η. Moreover we have

$$\int_{B^4\setminus\cup_{l=1}^N B^4_\delta(p_l)} |F_{A^0}|^2 \, dx^4 \le \liminf_{k\to+\infty} \int_{B^4} |F_{A^k}|^2 \, dx^4 \, .$$

Using a diagonal argument with $\delta \to 0$ we can extend A^0 as a $\mathcal{G}$-valued 1-form in $W^{1,2}_{\mathrm{loc}}(B^4 \setminus \{p_1 \cdots p_N\})$ and still satisfying

$$\int_{B^4} |F_{A^0}|^2 \, dx^4 \le \liminf_{k\to+\infty} \int_{B^4} |F_{A^k}|^2 \, dx^4 \, . \tag{V.24}$$

We conclude the proof of Theorem V.5 by changing the gauge of A^0 in the neighborhood of each blow up point p_l making use of the following Theorem V.7, known as *point removability theorem*, which gives the existence of a change of gauge g in order to extend our connection 1-form $(A^0)^g$ as a $W^{1,2}$ $\mathcal{G}$-valued 1-form in the neighborhood of each p_l. We then paste together these $W^{1,2}$-gauges by using the same technique as the one we used in the case without blow up points in order to get a global $W^{1,2}$ representative of A^0 on B^4 gauge equivalent to η on ∂B^4 and satisfying (V.24). This concludes the proof of Theorem V.5. □

Theorem V.6 (Point removability). *Let* $A \in W^{1,2}_{\mathrm{loc}}(\wedge^1 \overline{B^4} \setminus \{0\}, \mathcal{G})$ *and* G *be a compact Lie group such that*

$$\int_{B^4} |dA + A\wedge A|^2 \, dx^4 < +\infty \, ,$$

then there exists a gauge change $g \in W^{2,2}_{\mathrm{loc}}(\overline{B^4} \setminus \{0\}, G)$ *such that*

$$A^g \in W^{1,2}(\wedge^1 B^4, \mathcal{G}) \, .$$ □

Remark V.1. Point removability results play an important rôle in the analysis of conformally invariant variational problems. This is a natural consequence of due to the existence of point concentration which is inherent to the conformal invariance. These results are often formulated for the critical points of conformally invariant Lagrangians and in the present case it has been first proved by K. Uhlenbeck for Yang–Mills fields (see [58]). Observe that here we are not assuming that A is satisfying a particular equation. □

Remark V.2. Beyond geometric analysis, point removability results play also an important rôle in complex geometry. One could for instance quote the work of Bando [3] about the possibility to extend a Hermitian holomorphic structure $F_A^{0,2} = 0$ with L^2 bounded curvature on a the punctured ball $B^4 \setminus \{0\}$ as a smooth holomorphic bundle throughout the origin. Beside the holomorphicity condition

$F_A^{0,2} = 0$ no further "equation" is assumed and in particular the Einstein equation $\omega \cdot F_A^{1,1} = c\, I$ is not assumed and the connection form is not necessarily a Yang–Mills field. $\square$

Remark V.3. A similar point removability result can be established at the boundary. It suffices to extend carefully the gauge about the "singular point" at the boundary in order to reduce to the interior case. Details are left to the reader. $\square$

Proof of Theorem V.6. Without loss of generality we can assume that

$$\int_{B^4} |F_A|^2 \, dx^4 < \delta\,,$$

where $\delta > 0$ will be fixed later on in the proof. Denote for $i \geq 2$

$$T_i := B^4_{2^{-i+2}}(0) \setminus B^4_{2^{-i-2}}(0)\,.$$

From Theorem IV.1 there exists $\delta > 0$ such that, on each annulus T_i there exists a change of gauge g_i such that there exists $g_i \in W^{2,2}(T_i, G)$ such that

$$\begin{cases} \displaystyle\int_{T_i} 2^{2\,i} |A^{g_i}|^2 + \sum_{k,l=1}^4 |\partial_{x_k} A_l^{g_i}|^2 \, dx^m \leq C_G \int_{T_i} |dA + A \wedge A|^2 \, dx^4 \\ d^* A^{g_i} = 0 \qquad \text{in } T_i \\ \iota^*_{\partial T_i}(*A^{g_i}) = 0\,. \end{cases} \tag{V.25}$$

On $T_i \cap T_{i+1} = B^4_{2^{-i+1}} \setminus B^4_{2^{-i-2}}$ the transition function $g_{i\,i+1} = g_i (g_{i+1})^{-1}$ satisfy

$$A^{g_{i+1}} = (g_{i\,i+1})^{-1} \, dg_{i\,i+1} + (g_{i\,i+1})^{-1} \, A^{g_i} \, g_{i\,i+1}\,. \tag{V.26}$$

Hence (V.25) imply

$$2^{2\,i} \int_{B^4_{2^{-i+1}} \setminus B^4_{2^{-i-2}}} |dg_{i\,i+1}|^2 \, dx^4 \leq C \int_{B_{2^{-i+2}} \setminus B_{2^{-i-3}}} |dA + A \wedge A|^2 \, dx^4\,. \tag{V.27}$$

Taking the adjoint of the covariant derivative of equation (V.26)

$$-\Delta g_{i\,i+1} = A^{g_i} \cdot dg_{i\,i+1} - dg_{i\,i+1} \cdot A^{g_{i+1}}\,,$$

and, arguing as in the first part of the proof of Theorem V.5, we deduce the existence of $\overline{g_{i\,i+1}} \in G$ such that

$$\begin{aligned} &\|g_{i\,i+1} - \overline{g_{i\,i+1}}\|_{L^\infty(T_i \cap T_{i+1})} \\ &\quad \leq C\; 2^i \|dg_{i\,i+1}\|_{L^2(T_i \cap T_{i+1})} + C \sum_{k,l=1}^4 \|\partial^2_{x_k x_l} g_{i\,i+1}\|_{L^{2,1}(T_i \cap T_{i+1})} \\ &\quad \leq C \left[\int_{T_i \cup T_{i+1}} |dA + A \wedge A|^2 \, dx^4 \right]^{1/2} \leq C\, \sqrt{\delta}\,. \end{aligned} \tag{V.28}$$

We now modify the gauge change g_i as follows. Precisely, for any $i \in \mathbb{N}$, we denote

$$\overline{\sigma_i} := \overline{g_{1\,2}}\; \overline{g_{2\,3}} \cdots \overline{g_{i-1\,i}} \in G\,.$$

Observe that

$$A^{g_i\,\overline{\sigma_i}^{-1}} = \overline{\sigma_i}\,A^{g_i}\,\overline{\sigma_i}^{-1}\,. \tag{V.29}$$

Hence $A^{g_i\,\overline{\sigma_i}^{-1}}$ is still a Coulomb gauge satisfying

$$\left\{\begin{array}{l}\displaystyle\int_{T_i} 2^{2\,i}|A^{g_i\,\overline{\sigma_i}^{-1}}|^2+\sum_{k,l=1}^4|\partial_{x_k}A_l^{g_i\,\overline{\sigma_i}^{-1}}|^2\,dx^m\le C_G\int_{T_i}|dA+A\wedge A|^2\,dx^4\\ d^*A^{g_i\,\overline{\sigma_i}^{-1}}=0 \qquad \text{in } T_i\\ \iota^*_{\partial B^m}(*A^{g_i\,\overline{\sigma_i}^{-1}})=0\,.\end{array}\right. \tag{V.30}$$

Denote $h_i := g_i\,\overline{\sigma_i}^{-1}$ the transition functions on $T_i\cap T_{i+1}$ for these new gauges are given by

$$h_{i\,i+1} := g_i\,\overline{\sigma_i}^{-1}\,\overline{\sigma_{i+1}}\,(g_{i+1})^{-1} = g_i\,\overline{g_{i\,i+1}}\,(g_{i+1})^{-1}\,.$$

Using (V.28) we have

$$\|h_{i\,i+1}-e\|_{L^\infty(T_i\cap T_{i+1})}\le C\,\left[\int_{T_i\cup T_{i+1}}|dA+A\wedge A|^2\,dx^4\right]^{1/2}\le C\,\sqrt{\delta}\,. \tag{V.31}$$

Exactly as for g_i, using the identity

$$A^{h_{i+1}} = (h_{i\,i+1})^{-1}\,dh_{i\,i+1}+(h_{i\,i+1})^{-1}\,A^{h_i}\,h_{i\,i+1}\,, \tag{V.32}$$

together with (V.29) and (V.30) we obtain

$$\begin{aligned}&2^i\|dh_{i\,i+1}\|_{L^2(T_i\cap T_{i+1})}+\sum_{k,l=1}^4\|\partial^2_{x_kx_l}h_{i\,i+1}\|_{L^{2,1}(T_i\cap T_{i+1})}\\ &\le C\,\left[\int_{T_i\cup T_{i+1}}|dA+A\wedge A|^2\,dx^4\right]^{1/2}\le C\,\sqrt{\delta}\,,\end{aligned} \tag{V.33}$$

where e is the content function on $T_i\cap T_{i+1}$ equal to the neutral element of G. Having chosen δ small enough we ensure that the transition functions of this new set of trivialization are contained in a neighborhood of the neutral element into which the exponential map defines a diffeomorphism and there exist $U_{i\,i+1}$ such that $h_{i\,i+1}=\exp(U_{i\,i+1})$ and

$$\begin{aligned}&\|U_{i\,i+1}\|_{L^\infty(T_i\cap T_{i+1})}+2^i\|dU_{i\,i+1}\|_{L^2(T_i\cap T_{i+1})}+\sum_{k,l=1}^4\|\partial^2_{x_kx_l}U_{i\,i+1}\|_{L^{2,1}(T_i\cap T_{i+1})}\\ &\le C\,\left[\int_{T_i\cup T_{i+1}}|dA+A\wedge A|^2\,dx^4\right]^{1/2}\le C\,\sqrt{\delta}\,.\end{aligned} \tag{V.34}$$

Let ρ be a smooth function on $\mathbb{R}_+$ identically equal to 1 between 0 and $\sqrt{2}$ and compactly supported in $[0,2]$. On B^4 we define

$$\rho_i(x) := \rho(|x|\,2^i)\quad V_i := B_{2^{-i+3/2}}\setminus B_{2^{-i}}\quad \text{and } \tau_i := \exp(\rho_i\,U_{i\,i+1})\,.$$

With these notations we have

$$\text{on}\quad V_{i+1}\cap V_i = B_{2^{-i+1/2}}\setminus B_{2^{-i}}\qquad \text{we have } \tau_i = h_{i\,i+1} \text{ and } \tau_{i+1} = e\,.$$

Hence on $V_{i+1}\cap V_i$ we have

$$\begin{aligned} A^{h_i\,\tau_i} &= \tau_i^{-1}\, d\tau_i + \tau_i^{-1}\, A^{h^i}\,\tau_i \\ &= (h_{i\,i+1})^{-1}\, dh_{i\,i+1} + (h_{i\,i+1})^{-1}\, A^{h_i}\, h_{i\,i+1} = A^{h_{i+1}} = A^{h_{i+1}\,\tau_{i+1}}\,, \end{aligned} \tag{V.35}$$

and the 1-form $\hat{A}$ equal to $A^{h_i\,\tau_i}$ on each annulus V_i defines a global $W^{1,2}_{\text{loc}}$ connection 1-form on $B^4\setminus\{0\}$ gauge equivalent to A. Clearly, for $k=1,2$, we have the pointwise estimate

$$|d^k\tau_i| \le C\,\sum_{l=0}^{k} 2^{i\,l}\,|d^{k-l}U_{i\,i+1}|\quad \text{ on } V_i\,.$$

Combining this fact together with (V.30), (V.34) and (V.35) we obtain

$$\int_{V_i} 2^{2\,i}|A^{h_i\,\tau_i}|^2 + \sum_{k,l=1}^{4}|\partial_{x_k}A_l^{h_i\,\tau_i}|^2\,dx^4 \le C_G\,\int_{T_i\cup T_{i+1}}|dA + A\wedge A|^2\,dx^4\,.$$

Summing over i gives

$$\int_{B^4}|x|^{-2}\,|\hat{A}|^2 + \sum_{k,l=1}^{4}|\partial_{x_k}\hat{A}_l|^2\,dx^4 \le C_G\,\int_{B^4}|dA + A\wedge A|^2\,dx^4\,,$$

$\hat{A}$ is then in $W^{1,2}(\wedge^1 B^4, \mathcal{G})$ and this concludes the proof of theorem V.6. □

VI. The Yang–Mills equation in sub-critical and critical dimensions

VI.1. Yang–Mills fields

Until now we have produced solutions to the Yang–Mills Plateau problem in dimensions less or equal to four but we have not addressed issues related to the special properties that should be satisfied by these solutions. Maybe one of the first question that should be looked at is whether these minima define smooth *equivariant horizontal plane distributions* or not.

In order to study the regularity of solutions to the Yang–Mills Plateau problem we have first to produce the *Euler Lagrange equation* attached to this variational problem. This is the so-called *Yang–Mills equation*.

Definition VI.1. Let G be a compact Lie group and A be an L^2 connection 1-form on B^m into the Lie Algebra $\mathcal{G}$ of G. Assume that

$$\int_{B^m}|dA + A\wedge A|^2\;dx^m < +\infty\,,$$

we say that A is a Yang–Mills field if

$$\forall\,\xi\in C_0^\infty(\wedge^1 B^m,\mathcal{G})$$
$$\frac{d}{dt}\int_{B^m}|d(A+t\xi)+(A+t\xi)\wedge(A+t\xi)|^2\ dx^m\,|_{t=0}\ =0\,. \qquad \square$$

Observe that this definition makes sense for any $A\in L^2$ such that $F_A\in L^2$, indeed we have for any ξ in $C_0^\infty(\wedge^1 B^m,\mathcal{G})$

$$F_{A+t\xi}=F_A+t\ \left(d\xi+A\wedge\xi+\xi\wedge A\right)+t^2\ \xi\wedge\wedge\xi\quad\in L^2(\wedge^2 B^m,\mathcal{G})\,.$$

For such a $A\in L^2$ and for any ξ in $C_0^\infty(\wedge^1 B^m,\mathcal{G})$ we denote by $d_A\xi$ the following 2-form

$$d_A\xi(X,Y):=d\xi(X,Y)+[A(X),\xi(Y)]+[\xi(X),A(Y)]\,.$$

So we have for instance

$$d_A\xi(\partial_{x_i},\partial_{x_j})=\partial_{x_i}\xi_j-\partial_{x_j}\xi_i+[A_i,\xi_j]+[\xi_i,A_j]\,.$$

We have then the following proposition.

Proposition VI.1 (Yang–Mills Equation). *Let $A\in L^2(\wedge^1 B^m,\mathcal{G})$ such that $F_A\in L^2(\wedge^2 B^m,\mathcal{G})$. The connection 1-form A is a Yang–Mills field if*

$$\forall\,\xi\in C_0^\infty(\wedge^1 B^m,\mathcal{G})$$
$$\int_{B^4} d_A\xi\cdot F_A=0\,, \tag{VI.36}$$

which is equivalent to

$$d_A^*F_A=0\qquad in\ \mathcal{D}'(B^m)\,. \tag{VI.37}$$

In coordinates this reads

$$\forall\ i=1\cdots m\qquad \sum_{j=1}^m \partial_{x_j}(F_A)_{ij}+[A_j,(F_A)_{ij}]=0\,. \tag{VI.38}$$

$\square$

The Yang–Mills equation (VI.36) is also written symbolically as follows

$$d^*F_A+[A,\llcorner F_A]=0\,,$$

where $\llcorner$ is referring to the contraction operation between tensors with respect to the flat metric on B^m. The proof of the last statement of the proposition goes as follows (VI.36) in coordinates is equivalent to

$$\sum_{i,j=1}^m\int_{B^m}\left\langle\partial_{x_i}\xi_j-\partial_{x_j}\xi_i+[A_i,\xi_j]+[\xi_i,A_j],(F_A)_{ij}\right\rangle\ dx^m=0\,,$$

using integration by parts and the fact that the Killing metric, invariant under adjoint action, satisfies $\langle U,[V,W]\rangle=\langle W,[U,V]\rangle$ we obtain

$$\sum_{i,j=1}^m\int_{B^m}\langle-\partial_{x_i}(F_A)_{ij}+[(F_A)_{ij},A_i],\xi_j\rangle+\left\langle\partial_{x_j}(F_A)_{ij}+[A_j,(F_A)_{ij}],\xi_i\right\rangle dx^m=0\,,$$

which implies (VI.38).

The gauge invariance of the integrant of Yang–Mills Lagrangian implies that (VI.38) is solved for A if and only if it is solved for any gauge transformation A^g of A. More generally we have the following

$$\begin{cases} \forall\, A \in W^{1,2}(\wedge^1 B^m, \mathcal{G}) \qquad \forall\, g \in W^{2,2}(B^m, G) \\ d^*_{(A^g)} F_{A^g} = g^{-1}\, d^*_A F_A\, g\,, \end{cases} \tag{VI.39}$$

where we recall that $A^g := g^{-1}dg + g^{-1}A\,g$.

The Yang–Mills equation (VI.37) has to be compared with the *Bianchi identity* to which it is a kind of "dual equation". This is a *structure equation* which holds for any connection 1-form.

Proposition VI.2. [Bianchi identity] *For any $A \in L^2(\wedge^1 B^m, \mathcal{G})$ such that $F_A \in L^2(\wedge^2 B^m, \mathcal{G})$ the following identity holds*

$$d_A F_A = 0\,,$$

where $d_A F_A$ is the 3-form given by

$$\begin{aligned} d_A F_A(X,Y,Z) := dF_A(X,Y,Z) + [A(X), F_A(Y,Z)] + [A(Y), F_A(Z,X)] \\ + [A(Z), F_A(X,Y)]\,. \end{aligned}$$

□

The proof of the *Bianchi identity* goes as follows. We have

$$\begin{aligned} dF_A(X,Y,Z) &= d(A\wedge A)(X,Y,Z) \\ &= [dA(X,Y), A(Z)] + [dA(Y,Z), A(X)] + [dA(Z,X), A(Y)] \\ &= [F_A(X,Y), A(Z)] + [F_A(Y,Z), A(X)] + [F_A(Z,X), A(Y)]\,, \end{aligned}$$

where we have used the *Jacobi identity*

$$[[A(X), A(Y)], A(Z)] + [[A(Y), A(Z)], A(X)] + [[A(Z), A(X)], A(Y)] = 0\,.$$

This concludes the proof of the *Bianchi identity*.

In the particular case where G is **Abelian**, the Yang–Mills equation together with the Bianchi identity is equivalent to the harmonic map form system

$$\begin{cases} d^* F_A = 0 & \textbf{Yang–Mills} \\ dF_A = 0 & \textbf{Bianchi} \end{cases},$$

whose solutions are known to be analytic in every dimension. We are now asking about the same regularity issue in the <u>non-Abelian</u> case.

VI.2. The regularity of $W^{1,2}$ Yang–Mills fields in sub-critical and critical dimensions

Due to the huge gauge invariance, the "intrinsic" Yang–Mills equation (VI.38) as such cannot generate any kind of improved regularity for an arbitrary solution[10]. Nevertheless the "breaking this invariance" by taking the Coulomb gauge will make

[10] Assuming one solution A would be smooth and taking any arbitrary other "non-smooth" gauge g the new expression of the connection A^g is not smooth though it still solves Yang–Mills.

Yang–Mills PDE elliptic and will generate regularity under the ad hoc assumptions. Precisely we have the following result.

Theorem VI.7. *Let G be a compact Lie group and $m\geq 3$. Let $A\in W^{1,m/2}(\wedge^1 B^m,\mathcal{G})$ be a solution of the Yang–Mills equation* (VI.37) *satisfying the Coulomb condition $d^*A=0$ then A is C^∞ in B^m.* □

Proof of Theorem VI.7. We assume that A is Coulomb and satisfy the Yang–Mills equation (VI.37). Hence we have

$$d^*dA+d^*(A\wedge A)+[A,\llcorner dA]+[A,\llcorner A\wedge A]=0\,.$$

Using the fact that $d^*A=0$ the Yang–Mills equation in this Coulomb gauge reads then

$$\Delta A=d^*(A\wedge A)+[A,\llcorner dA]+[A,\llcorner(A\wedge A)]\,, \tag{VI.40}$$

and Theorem VI.7 is now the direct consequence of the following result. □

Theorem VI.8. *Let $m>2$ and $N\in\mathbb{N}^*$. Let $f\in C^\infty(\mathbb{R}^N\times(\mathbb{R}^m\otimes\mathbb{R}^N),\mathbb{R}^N)$ and let $g\in C^\infty(\mathbb{R}^N,\mathbb{R}^N)$ such that there exists $C>0$ satisfying*

$$|f(\xi,\Xi)|\leq C\ |\xi|\ |\Xi| \qquad \text{and} \qquad |g(\xi)|\leq C\ |\xi|^3\,. \tag{VI.41}$$

Let $u\in W^{1,m/2}(B^m,\mathbb{R}^N)$ satisfying

$$\Delta u=f(u,\nabla u)+g(u)\,, \tag{VI.42}$$

then u is C^∞. □

Proof of Theorem VI.8. We start with the following embedding

$$W^{1,m/2}(B^m)\hookrightarrow L^m(B^m)\,.$$

We claim that there exists $\alpha>0$ such that

$$\sup_{x_0\in B^4_{1/2}(0)\ ;\ 0<\rho<1/4}\rho^{-4\,\alpha}\int_{B_\rho(x_0)}|u|^m(x)\ dx^m<+\infty\,. \tag{VI.43}$$

Let $\epsilon>0$ to be fixed later. There exists $\rho_0>0$ such that

$$\sup_{x_0\in B^m_{1/2}(0)\ ;\ 0<\rho<\rho_0}\int_{B^m_\rho(x_0)}[|u|^m(x)+|\nabla u|^{m/2}(x)]dx^m<\varepsilon\,. \tag{VI.44}$$

Let now $x_0\in B^m_{1/2}(0)$ and $\rho<\rho_0$ arbitrary. On $B_\rho(x_0)$ we consider φ to be the solution of

$$\begin{cases}\Delta\varphi=f(u,\nabla u)+g(u) & \text{in } B^m_\rho(x_0)\\ \quad\varphi=0 & \text{on}\quad \partial B^m_\rho(x_0)\,.\end{cases} \tag{VI.45}$$

Classical elliptic estimates (see [16]) give the existence of a constant independent of ρ such that

$$\begin{aligned}\|\varphi\|_{L^m(B^m_\rho(x_0))}&\leq C\ \|f(u,\nabla u)+g(u)\|_{L^{m/3}(B^m_\rho(x_0))}\\ &\leq C\ \|u\|_{L^m(B^m_\rho(x_0))}\ \|\nabla u\|_{L^{m/2}(B^m_\rho(x_0))}+C\ \|u\|^3_{L^m(B^m_\rho(x_0))}\,.\end{aligned} \tag{VI.46}$$

The difference $v := u - \varphi$ is harmonic on $B^m_\rho(x_0)$. Hence $|v|^m$ is subharmonic

$$\Delta |v|^m = m\,|v|^{m-2}|\nabla v|^2 + \frac{m}{2}\,\left(\frac{m}{2} - 1\right)\,|v|^{m-4}\,|\nabla |v|^2|^2 \geq 0\,.$$

This gives that

$$\forall\ r < \rho \qquad \int_{\partial B_r(x_0)} \frac{\partial |v|^m}{\partial r} \geq 0\,,$$

which implies that

$$\frac{d}{dr}\left[\frac{1}{r^m}\int_{B^4_r(x_0)} |v|^m(x)\ dx^m\right] \geq 0\,.$$

So we have in particular

$$\int_{B_{\rho/4}(x_0)} |v|^m(x)\ dx^m \leq 4^{-m}\int_{B_\rho(x_0)} |v|^m(x)\ dx^m\,. \tag{VI.47}$$

From this inequality we deduce

$$\begin{aligned}\int_{B_{\rho/4}(x_0)} |u|^m(x)\ dx^m &\leq 2^{m-1}\ \int_{B_{\rho/4}(x_0)} [|v|^m + |\varphi|^m]\ dx^m \qquad \text{(VI.48)}\\ &\leq 2^{-m-1}\int_{B_\rho(x_0)} |v|^m(x)\ dx^m + 2^{m-1}\ \int_{B_\rho(x_0)} |\varphi|^m\ dx^m\\ &\leq 2^{-2}\int_{B_\rho(x_0)} |u|^m(x)\ dx^m +\ 2^m\int_{B_\rho(x_0)} |\varphi|^m\ dx^m\,.\end{aligned}$$

Combining (VI.46) and (VI.48) we then have

$$\begin{aligned}&\int_{B_{\rho/4}(x_0)} |u|^m(x)\ dx^m\\ &\quad\leq \left[2^{-2} + C_0\ [\|\nabla u\|^m_{L^{m/2}(B_\rho)} + \|u\|^{2m}_{L^m(B_\rho)}]\right]\int_{B_\rho(x_0)} |u|^m(x)\ dx^m\,.\end{aligned} \tag{VI.49}$$

We choose $\varepsilon > 0$ such that $C_0\varepsilon^2 \leq 2^{-1}$ and we have then established that for any $\rho < \rho_0$

$$\int_{B_{\rho/4}(x_0)} |u|^m(x)\ dx^m \leq \frac{1}{2}\int_{B_\rho(x_0)} |u|^m(x)\ dx^m\,. \tag{VI.50}$$

Iterating this inequality gives (VI.43). Inserting the *Morrey bound* (VI.43) into the equation (VI.42) gives

$$\sup_{x_0\in B^m_{1/2}(0)\ ;\ 0<\rho<1/4} \rho^{-m\,\alpha/3}\int_{B_\rho(x_0)} |\Delta u|^{m/3}(x)\ dx^m < +\infty\,. \tag{VI.51}$$

The *Adams–Sobolev* embeddings (see [1]) give then the existence of $p > m/2$ such that $\nabla u \in L^p_{\text{loc}}(B^m_{1/2}(0))$. It is easy then to see that the PDE (VI.42) becomes sub-critical for $W^{1,p}$ (with $p > m/2$) in m dimensions and we can apply a similar bootstrap arguments to obtain the desired regularity for u. This concludes the proof of Theorem VI.8. $\Box$

Remark VI.4. The proof of the regularity of Yang–Mills fields in the critical four dimensions is "soft" in comparison with the proof of the regularity of the "cousin problem": the harmonic maps between a surface and a manifold. Both equations are critical respectively in four and two dimensions but the analysis of the harmonic map equation is made more delicate by the fact that the non-linearity in the harmonic map equation is in L^1 which is a space which does not behave "nicely" with respect to Calderòn–Zygmund operations. There is no such a difficulty for Yang–Mills. What is delicate however is to construct a "good gauge" in which Yang–Mills equation becomes elliptic. In a somewhat parallel way the difficulty posed by the harmonic maps equation was overcome by the author by solving a gauge problem (see [41]).

One consequence of the previous regularity result and the point removability result V.6 is the following *point removability* theorem for Yang–Mills fields in four dimension

Theorem VI.9 (Point removability for Yang–Mills in conformal dimension). *Let A be a weak solution in $W^{1,2}_{\mathrm{loc}}(\wedge^1 B^4, \mathcal{G})$ to Yang–Mills equation*

$$d_A^* F_A = d^* F_A + [A, \llcorner F_A] = 0 \qquad \text{in } \mathcal{D}'(B^4 \setminus \{0\}) .$$

Assume

$$\int_{B^4} |dA + A \wedge A|^2 \, dx^4 < +\infty ,$$

then there exists $g \in W^{2,2}_{\mathrm{loc}}(B^4 \setminus \{0\})$ such that

$$A^g \in C^\infty(B^4) ,$$

and A^g solves the Yang–Mills equation strongly in the whole ball B^4. □

Proof of Theorem VI.9. The point removability result V.4 gives the existence of a $W^{2,2}_{\mathrm{loc}}$-gauge such that $A^g \in W^{1,2}(\wedge^1 B^4, \mathcal{G})$. Using Uhlenbeck's Coulomb gauge IV.1 extraction theorem we can assume that A^g satisfies the Coulomb condition $d^* A^g = 0$ in $\mathcal{D}'(B^4)$. So A^g is a $W^{1,2}$-solution of a system of the form

$$\Delta u = f(u, \nabla u) + g(u) \quad \text{in } \mathcal{D}'(B^4 \setminus \{0\}) ,$$

where f and g are smooth maps satisfying (VI.41).

The distribution $\Delta u - f(u, \nabla u) - g(u)$ is supported in $\{0\}$. Hence by a classical result in distribution theory this distribution is a <u>finite</u> linear combination of derivatives of Dirac masses:

$$-\Delta u + f(u, \nabla u) + g(u) = \sum_{|\alpha \leq N} C_\alpha \, \partial_\alpha \delta_0 , \tag{VI.52}$$

where $N \in \mathbb{N}$, $\alpha = (\alpha_1 \cdots \alpha_4) \in \mathbb{N}^4$, $|\alpha| := |\alpha_1| + \cdots + |\alpha_4|$, $C_\alpha \in \mathbb{R}^4$ and ∂_α denotes the partial derivative

$$\partial_\alpha := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \, \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \, \frac{\partial^{\alpha_3}}{\partial x_3^{\alpha_3}} \, \frac{\partial^{\alpha_4}}{\partial x_4^{\alpha_4}} .$$

Let χ be an arbitrary smooth compactly supported function in $B^4_1(0)$. Denote $\chi_\varepsilon(x) := \chi(x/\varepsilon)$, multiply equation (VI.52) by this function and integrate over B^4 gives

$$\sum_{|\alpha|\le N} C_\alpha \frac{\partial^\alpha\chi(0)}{\varepsilon^{|\alpha|}} = \frac{1}{\varepsilon}\int_{B^4_\varepsilon} \nabla\chi(x/\varepsilon)\cdot\nabla u + \int_{B^4_\varepsilon} \chi_\varepsilon\ \left[f(u,\nabla u)+g(u)\right] .$$

Hence, using $\nabla u \in L^2$ and $u\in L^4$ we have

$$\left|\sum_{|\alpha|\le N} C_\alpha \frac{\partial^\alpha\chi(0)}{\varepsilon^{|\alpha|}}\right| = o(\varepsilon) .$$

Since $\partial^\alpha\chi(0)$ are arbitrary, this implies that $C_\alpha = 0$ for any α. So the equation $-\Delta u + f(u,\nabla u) + g(u) = 0$ holds on the whole ball and we can apply Theorem VI.8 to $u = A^g$ and obtain that it is C^∞ which concludes the proof of Theorem VI.9. □

It is clear that the solutions to the Yang–Mills Plateau problems satisfy the Yang–Mills equation and hence we have the following corollary.

Corollary VI.1. *Let G be a compact Lie group and $m\le 4$. For any 1-form $\eta \in H^{1/2}(\wedge^1\partial B^m,\mathcal{G})$ the following minimization problem*

$$\inf\left\{\int_{B^m} |F_A|^2\ dx^m\ ;\ \iota^*_{\partial B^m} A = \eta^g \quad \textit{for some } g\in H^{3/2}(\partial B^m, G)\right\} \tag{VI.53}$$

is achieved by a 1-form $A^0 \in W^{1,2}(\wedge^1 B^m,\mathcal{G})$ which is C^∞ in any local $W^{1,2}$-Coulomb gauge inside the ball B^m. □

VII. Concentration compactness and energy quantization for Yang–Mills fields in critical dimension

The goal of this section is to study establish the behavior of sequences of Yang–Mills fields of uniformly bounded energy in critical dimension four. There are three main problematics attached to this study

- Modulo extraction of subsequence, do we have strong converge to a limiting Yang–Mills?
- If the strong convergence does not hold where is located the lack of strong convergence in the base?
- How much Yang–Mills energy is lost at the limit?

We have already several tools and results at hand that we established in the previous sections in order to provide a relatively precise answer to these three questions. The proof or our main result in this section is based in particular on the following "quantitative reformulation" of the regularity theorem (Theorem VI.7) which belongs to the family of the so-called ϵ-regularity results for *conformally invariant problems.*

Theorem VII.1 (ϵ-regularity for Sobolev solutions to Yang–Mills in conformal dimension). *Let G be a compact Lie group, there exists $\varepsilon_{G,4} > 0$ such that for any $\mathcal{G}$-valued 1-forms A in $W^{1,2}(\wedge^1 B_1^4(0), \mathcal{G})$ satisfying the Yang–Mills equation*

$$d_A^* F_A = d^* F_A + [A, \llcorner F_A] = 0 \quad \text{in} \quad \mathcal{D}'(B_1^4(0))\,.$$

and the small energy condition

$$\int_{B_1^4(0)} |F_A|^2 \; dx^4 < \varepsilon_{G,4}\,,$$

then there exists a gauge g in which the following estimates holds: for any $l \in \mathbb{N}$ there exists $C_l > 0$ such that

$$\|\nabla^l (A)^g\|^2_{L^\infty(B_{1/2}(0))} \leq C_l \int_{B_1^m(0)} |F_A|^2 \; dx^4\,. \tag{VII.1}$$

□

Proof of Theorem VII.1. We choose $\varepsilon_{G,4} > 0$ that will be definitively fixed a bit later in the proof to be at least smaller than the $\epsilon_G > 0$ of the *Coulomb gauge* extraction result theorem (Theorem IV.1). We now work in this *Coulomb gauge* and we omit to mention the superscript g. So, from now on until the end of the proof, we are then assuming that we have

$$\int_{B_1^4} |A|^2 \; dx^4 + \sum_{i,j=1}^{4} |\partial_{x_i} A_j|^2 \, dx^4 \leq C_G \int_{B_1^4} |F_A|^2 \, dx^4 < \varepsilon_{G,4}\,, \tag{VII.2}$$

from which we deduce in particular

$$\int_{B_1^4} |A|^4 \; dx^4 \leq C_0 \left[\int_{B_1^4} |F_A|^2 \, dx^4\right]^2 , \tag{VII.3}$$

for some constant $C_0 > 0$. Recall that in the Coulomb gauge we are choosing, the connection form A satisfies the elliptic system (VI.40) to which we can apply the arguments of the proof of Theorem VI.8 that we are going to follow closely keeping track this time of each estimate. In particular, having chosen $\varepsilon_{G,4}$ small enough we have inequality (VI.50) which holds for $u = A$ and for any $B_\rho(x_0) \subset B_1(0)$ and we deduce

$$\begin{aligned}
&\forall\, x_0 \in B_{3/4}(0) \quad \forall\, \rho < 1/4 \\
&\int_{B_\rho(x_0)} |A|^4 \; dx^4 \leq 2\, \rho^\alpha \int_{B_1^4(0)} |A|^4 \; dx^4 \\
&\qquad\qquad \leq 2\, C_0\, \rho^\alpha \left[\int_{B_1^4(0)} |F_A|^2 \, dx^4\right]^2 ,
\end{aligned} \tag{VII.4}$$

where $\alpha = \log 2/\log 4$. Inserting this inequality in the Yang–Mills PDE in Coulomb gauge (VI.40) we obtain the existence of a constant $C_1 > 0$ such that

$$\forall\, x_0 \in B_{3/4}(0) \quad \forall\, \rho < 1/4$$
$$\int_{B_\rho(x_0)} |\Delta A|^{4/3}\, dx^4 \le C_1\, \rho^{\alpha/3} \left[\int_{B_1^4(0)} |F_A|^2\, dx^4\right]^{4/3} . \tag{VII.5}$$

Combining (VII.4) and (VII.5) we deduce from *Adams–Morrey inequalities* (see [1])

$$\|\nabla A\|_{L^p(B_{3/5}(0))} \le C \sup_{x_0\in B_{3/4}(0)\,;\,\rho<1/4} \left[\rho^{-\alpha/3}\int_{B_\rho(x_0)} |\Delta A|^{4/3}\, dx^4\right]^{3/4} + C\, \|A\|_{L^2(B_1(0))}\,, \tag{VII.6}$$

where

$$p = \frac{16 - 4\,\alpha/3}{8-\alpha} > 2\,.$$

Hence we have for this $p > 2$

$$\|\nabla A\|_{L^p(B_{3/5}(0))} \le C \left[\int_{B_1^4(0)} |F_A|^2\, dx^4\right]^{1/2} .$$

Since $p > 2$ the non-linear elliptic system (VI.40) becomes sub-critical in four dimensions and a standard bootstrap argument gives (VII.1). This concludes the proof of Theorem VII.1. □

The previous ϵ-regularity result is the main step for proving the following *concentration compactness* theorem for sequences of Yang–Mills fields.

Theorem VII.2 (Concentration compactness for Yang–Mills Fields in conformal dimension). *Let (M^4, h) be a closed four-dimensional Riemannian manifold and P a principal smooth G bundle over M^4. Let ∇^k be a sequence of Yang–Mills connections satisfying*

$$\limsup_{k\to+\infty} \int_{M^4} |F_{\nabla^k}|_h^2\, dvol_h < +\infty\,.$$

Then there exists a subsequence $\nabla^{k'}$, a smooth G-bundle P^∞ over (M^4, h), a smooth Yang–Mills connection ∇^∞ of P^∞ and finitely many points $\{p_1 \cdots p_N\}$ in M^4 such that for any contractible open set $D^4 \subset M^4 \setminus \{p_1 \cdots p_N\}$ with compact closure, there exists a sequence of trivialization of P over D^4 for which

$$A^{k'} \to A^\infty \qquad \textit{strongly in } C^l(D^4) \quad \forall\, l \in \mathbb{N}\,,$$

where $A^{k'}$ (resp. A^∞) is the connection 1-form associated to $\nabla^{k'}$ (resp. ∇^∞) in this sequence of local trivializations of P (resp. P^∞) over D^4. Moreover we have *the following weak convergence in Radon measure*

$$\mu^{k'} := |F_{\nabla^{k'}}|_h^2 \, \mathrm{dvol}_h \rightharpoonup \mu^\infty := |F_{\nabla^\infty}|_h^2 \, \mathrm{dvol}_h + \nu \,, \tag{VII.7}$$

where ν is a non-negative atomic measure supported by the points p_j

$$\nu := \sum_{j=1}^N f_j \ \delta_{p_j} \,. \tag{VII.8}$$

□

Proof of Theorem VII.2. We follow step by step the proof of theorem V.5 replacing for the choice of the covering the Uhlenbeck Coulomb gauge threshold $\varepsilon_G(M^4, h)$ by the smaller positive constant $\varepsilon_{G,4}(M^4, h)$ given by the ϵ-regularity result VII.1 on the manifold (M^4, h). Observe that, because of the epsilon regularity result, the Coulomb gauges $(A^k)^{g_i^k}$ are pre-compact for any C^l-topology on each ball $B_\rho(x_i)$. This gives also the pre-compactness of the transition functions g_{ij}^k in any of the C^l topologies. Hence, the co-cycles g_{ij}^k converge in any of these topology to the limiting (now smooth) co-cycle g_{ij}^∞ which defines a smooth G-bundle P^∞ over $M^4 \setminus \{p_1 \cdots p_N\}$. Moreover the limiting collection of 1-forms $A^{j,\infty}$ defines a connection ∇^∞ on P^∞ satisfying also the Yang–Mills equation which passes obviously to the limit under C^∞ convergence. The gauge invariant quantities such as $|F_{\nabla^k}|^2$ converge also to the corresponding limiting quantities and we have then, modulo extraction of a further subsequence, the existence of a limiting radon measure ν supported on the points p_j exclusively such that (VII.9) holds. Finally applying the point removability theorem and once again the ϵ-regularity we extend ∇^∞ globally on M^4 as a Yang–Mills smooth connection of the bundle P^∞ which also obviously extends throughout the p_j as a smooth bundle. This concludes the proof of Theorem VII.2. □

Finally, we identify the *concentration atomic measure* by proving that the weights f_j in front of the Dirac masses δ_{p_j} are the sums of Yang–Mills energies of Yang–Mills fields over S^4, the so-called "bubbles". Precisely we have the following *energy identity result* which was first established for instantons in [56] and for Yang–Mills fields in general in [42]. The proof we present below is using the *interpolation Lorentz spaces* following a technic introduced in [24] and [25].

Theorem VII.3 (Energy quantization for Yang–Mills Fields in conformal dimension.). *Let (M^4, h) be a closed four-dimensional Riemannian manifold and P a principal smooth G bundle over M^4. Let ∇^k be a sequence of Yang–Mills connections of uniformly bounded Yang–Mills energy converging strongly away from finitely many points $\{p_1 \cdots \partial_N\}$ to a limiting Yang–Mills connection ∇^∞ as described in Theorem* VII.2*. Let ν be the atomic concentration measure*

$$\nu := \sum_{j=1}^N f_j \ \delta_{p_j} \,,$$

satisfying

$$\mu^{k'} := |F_{\nabla^{k'}}|_h^2 \, \mathrm{dvol}_h \rightharpoonup \mu^\infty := |F_{\nabla^\infty}|_h^2 \, \mathrm{dvol}_h + \nu \, .$$

Then for each $j = 1 \cdots N$ *there exists finitely many* G*-Yang–Mills connections* $(D_j^i)_{i=1\cdots N_j}$ *over* S^4 *such that*

$$\forall\, j = 1 \cdots N \quad f_j = \sum_{i=1}^{N_j} \int_{S^4} |F_{D_j^i}|^2 \, \mathrm{dvol}_{S^4} \, . \tag{VII.9}$$

□

Remark VII.1. The energy quantization result can be established by direct geometric arguments in the special case of instantons in four dimensions (see [15]).

Proof of Theorem VII.3. Since the result is local, the metric in the domain does not play much role and we will present the proof for $M^4 = B_1^4(0)$ equipped with the flat metric and assuming moreover that there is exactly one limiting blow-up point, $N = 1$, which coincide with the origin, $p_1 = 0$. We also express the connection ∇^k in s.

Recall that we denote by $\varepsilon_{G,4}$ the positive constant given by the ϵ-regularity theorem (Theorem VII.1). We detect the "most concentrated bubble" precisely let

$$\rho^k := \inf \left\{ \rho \, ; \, \exists x \in B_1^4(0) \quad \text{s. t.} \int_{B_\rho(x)} |F_{A^k}|^2 \, dx^4 = \frac{\varepsilon_{G,4}}{2} \right\} .$$

Since we are assuming that blow-up is happening exactly at the origin we have that

$$\begin{cases} \rho^k \to 0 \quad \text{and} \\[2mm] \exists\, x^k \to 0 \quad \text{s.t.} \displaystyle\int_{B_{\rho^k}(x^k)} |F_{A^k}|^2 \, dx^4 = \frac{\varepsilon_{G,4}}{2} \, . \end{cases}$$

We choose a sequence x^k that we call *center of the first bubble* and ρ^k is called the *critical radius of the first bubble.* Let

$$\hat{A}^k(y) := \rho^k \sum_{i=1}^{4} A_i^k(\rho^k \, y + x^k) \, dy_i \, .$$

Due to the scaling invariance of the Yang–Mills Lagrangian in four dimensions, $\hat{A}^k$, which is the pull-back of A^k by the dilation map $D^k(y) := \rho^k \, y + x^k$, is a Yang–Mills fields moreover

$$\max_{y \in B_{1/(2\rho^k)}(0)} \int_{B_1^4(y)} |F_{\hat{A}^k}|^2 \, dy^4 = \int_{B_1^4(0)} |F_{\hat{A}^k}|^2 \, dy^4 = \frac{\varepsilon_{G,4}}{2} \, . \tag{VII.10}$$

Applying the ϵ-regularity theorem (Theorem VII.1) we deduce that on *Uhlenbeck's Coulomb gauges* g which exists on each unit ball $B_1(y)$ since the $\epsilon_{G,4}$ has been taken smaller than ϵ_G from Theorem IV.1

$$\forall\, l \in \mathbb{N} \qquad \sup_{y \in B_{1/(2\rho^k)}(0)} \|\nabla^l (\hat{A})^g\|_{L^\infty(B_{1/2})} \leq C_l \, ,$$

where C_l is independent of k. Hence, locally in Coulomb gauge, modulo extraction of a subsequence, the sequence $\hat{A}^k$ converges strongly in any C^l_{loc} topology on $\mathbb{R}^4$ to a limiting Yang–Mills connection $\hat{A}^\infty$ satisfying

$$\int_{B^4_1(0)} |F_{\hat{A}^k}|^2 \; dy^4 = \frac{\varepsilon_{G,4}}{2}$$

and which is therefore non-trivial. Let now π be the stereographic projection with respect to the north pole, due to the *conformal invariance* of Yang–Mills energy $\tilde{A} := \pi^* \hat{A}^\infty$ is a non trivial Yang–Mills Field on $S^4 \setminus \{\text{south pôle}\}$. Since $\tilde{A}$ is a smooth $\mathcal{G}$-valued 1-form with finite Yang–Mills energy and satisfying the Yang–Mills equation we can apply the *point removability* result for Yang–Mills fields, theorem VI.9 and conclude that $\tilde{A}$ extends to a global smooth Yang–Mills G-connection D^1_1 over the whole S^4 which is our *first bubble* and using again the conformal invariance of Yang–Mills energy we have

$$\frac{\varepsilon_{G,4}}{2} \le \int_{S^4} |F_{D^1_1}|^2 \; dvol_{S^4} = \lim_{R\to+\infty} \lim_{k\to+\infty} \int_{B_{R\,\rho^k}(x^k)} |F_{A^k}|^2 \; dx^4 \, . \tag{VII.11}$$

We have now to study the loss of Yang–Mills energy in the so-called *neck region* between the *first bubble* and the *macroscopic solution* A^∞ to which A^k converges away from zero. Precisely we are studying

$$\begin{aligned} &\lim_{k\to+\infty} \int_{B^4} |F_{A^k}|^2 dx^4 - \int_{B^4} |F_{A^\infty}|^2 dx^4 - \int_{S^4} |F_{D^1_1}|^2 \, \mathrm{dvol}_{S^4} \\ &\quad = \lim_{R\to+\infty} \lim_{k\to+\infty} \int_{B_{R^{-1}}(x^k)\setminus B_{R\,\rho^k}(x^k)} |F_{A^k}|^2 \; dx^4 \, . \end{aligned} \tag{VII.12}$$

For any G there is a minimal Yang–Mills energy among all non-trivial Yang–Mills Fields. This can be proved easily observing that if the energy is less than the $\varepsilon_{G,4}$ threshold, the connections can be represented by a global smooth Yang–Mills 1-form on S^4 to which the C^l estimates of Theorem VII.1 apply. Hence since A satisfies globally on S^4 the PDE (VI.40) and for small enough Yang–Mills energy this implies that A is a harmonic 1-form on S^4 which gives that it is a trivial Yang–Mills fields. Denote

$$\mathrm{YM}(G, S^4) = \min \left\{ \int_{S^4} |F_D|^2 \, \mathrm{dvol}_{S^4} \; ; \; D \text{ is a non zero } G \text{ Yang–Mills Field} \right\} .$$

To simplify the presentation we assume that

$$\lim_{k\to+\infty} \int_{B^4} |F_{A^k}|^2 dx^4 - \int_{B^4} |F_{A^\infty}|^2 dx^4 - \int_{S^4} |F_{D^1_1}|^2 \, \mathrm{dvol}_{S^4} < \; \mathrm{YM}(G, S^4) \, , \tag{VII.13}$$

or in other words

$$\lim_{R\to+\infty} \lim_{k\to+\infty} \int_{B_{R^{-1}}(x^k)\setminus B_{R\,\rho^k}(x^k)} |F_{A^k}|^2 \; dx^4 < \; \mathrm{YM}(G, S^4) \, . \tag{VII.14}$$

Alternatively we would have to go through some standard and fastidious induction procedure to remove all the bubbles one by one – each of them taking at least an amount of $\mathrm{YM}(G,S^4)$ Yang–Mills energy – and we would be anyway reduced at the end to study the loss of energy in annuli region where (VII.14) holds (such a procedure is described, for instance in, [5, Proposition III.1] in the framework of *Willmore surfaces*). Under the assumption (VII.14) the goal is ultimately to prove

$$\lim_{R\rightarrow+\infty}\lim_{k\rightarrow+\infty}\int_{B_{R^{-1}}(x^k)\setminus B_{R\,\rho^k}(x^k)}|F_{A^k}|^2\ dx^4=0\,,$$

that will finish the proof of the theorem. We are now going to prove the following claim.

Claim 1:

$$\forall\,\delta>0\quad\exists\,R_\delta>1\quad\text{s. t.}\quad\forall\ r\in[R_\delta\,\rho^k,R_\delta^{-1}]$$

$$\limsup_{k\rightarrow+\infty}\int_{B_{2\,r}(x^k)\setminus B_r(x^k)}|F_{A^k}|^2\ dx^4<\delta\,.$$

Proof of claim 1. We argue by contradiction. Assume there exists $\delta_0>0$ such that for all $R>1$ there exists $r^k\in[R\,\rho^k,R^{-1}]$

$$\limsup_{k\rightarrow+\infty}\int_{B_{2\,r^k}(x^k)\setminus B_{r^k}(x^k)}|F_{A^k}|^2\ dx^4>\delta_0\,.$$

Since we can find a sequence r_k for any $R>1$, using a diagonal argument and the extraction of a subsequence we can assume that

$$\frac{r^k}{\rho^k}\rightarrow+\infty\qquad\text{and}\qquad r^k\rightarrow 0\,.$$

Now we introduce

$$s^k:=\inf\left\{\begin{array}{c} s\ ;\ \exists\,x\in B_{2\,r^k}(x^k)\setminus B_{r^k}(x^k)\qquad\text{s.t.}\\ \displaystyle\int_{B_s(x)}|F_{A^k}|^2\ dx^4=\min\left\{\frac{\varepsilon_{G,4}}{2},\frac{\delta_0}{16}\right\}\,.\end{array}\right\}$$

Let $\tilde{x}^k$ be a point in the dyadic annulus $B_{2\,r^k}(x^k)\setminus B_{r^k}(x^k)$ where this infimum is achieved. We clearly have

$$B_{s^k}(\tilde{x}^k)\subset B_{R^{-1}}(x^k)\setminus B_{R\,\rho^k}(x^k)$$

for any $R>1$ and k large enough. Dilating the Yang–Mills connection A^k about $\tilde{x}^k$ at a rate $(s^k)^{-1}$ we again obtain a limiting non-trivial Yang–Mills field, a second *bubble*, $\tilde{A}^\infty$ either on $\mathbb{R}^4$ or on $\mathbb{R}^4\setminus\{0\}$ depending whether s^k/r^k tends to zero or not. In any case we have

$$\lim_{R\rightarrow+\infty}\lim_{k\rightarrow+\infty}\int_{B_{R^{-1}}(x^k)\setminus B_{R\,\rho^k}(x^k)}|F_{A^k}|^2\ dx^4\geq\int_{\mathbb{R}^4}|F_{\tilde{A}^\infty}|^2\ dx^4\geq\min\left\{\frac{\varepsilon_{G,4}}{2},\frac{\delta_0}{16}\right\}\,,$$

and the *point removability* theorem (Theorem VI.9) for Yang–Mills fields implies that $\tilde{A}^\infty$ extends to a non-trivial Yang–Mills Field on S^4

$$\int_{\mathbb{R}^4} |F_{\tilde{A}^\infty}|^2 \, dx^4 \geq \mathrm{YM}(G, S^4)\,,$$

which contradicts the fact that we are working under the assumption that there is only one *bubble* (i.e., assumption (VII.14)). So we have proved claim 1.

Combining claim 1 and the ϵ-regularity theorem (Theorem VII.1) we obtain

$$\begin{aligned} &\forall\, \delta > 0 \quad \exists\, R_\delta > 1 \quad \text{s.t.} \\ &\forall\, x \in B_{R_\delta^{-1}} \setminus B_{R_\delta\, \rho^k}(x^k) \qquad |x|^2\ |F_{A^k}|^2(x) < \delta\,. \end{aligned}$$

Consider an Uhlenbeck Coulomb gauge $(A^k)^{g^k}$ in the annulus

$$B_{2\, R_\delta\, \rho^k}(x^k) \setminus B_{R_\delta\, \rho^k}(x^k).$$

Introduce χ to be a cut-off function such that

$$\begin{cases} \chi(x) \equiv 1 & \text{in } R^4 \setminus B^4_2(0) \\ \chi(x) \equiv 0 & \text{in } B^4_1(0)\,, \end{cases}$$

and let $\chi^k(x) := \chi(R_\delta\, \rho^k\ (x - x^k))$. Extend $(A^k)^{g^k}$ in $B_{2\, R_\delta\, \rho^k}(x^k)$ by taking

$$\tilde{A}^k := \chi^k (A^k)^{g^k}\,.$$

Using again the C^l estimates (VII.1) of ϵ-regularity theorem (Theorem VII.1) for the Uhlenbeck Coulomb gauge $(A^k)^g$ in the annulus $B_{2\, R_\delta\, \rho^k}(x^k) \setminus B_{R_\delta\, \rho^k}(x^k)$ we obtain for any $x \in B_{2\, R_\delta\, \rho^k}(x^k) \setminus B_{R_\delta\, \rho^k}(x^k)$

$$\begin{aligned} &|x|\ |\tilde{A}^k|(x) + |x|^2\ |\nabla \tilde{A}^k|(x) \\ &\quad \leq C\ \left[\int_{B_{4\, R_\delta\, \rho^k}(x^k) \setminus B_{R_\delta\, \rho^k/2}(x^k)} |F_{A^k}|^2\ dx^4\right]^{1/2}\,. \end{aligned} \tag{VII.15}$$

We have then produced an extension $\tilde{A}^k$ of A^k inside the ball $B_{2\, R_\delta\, \rho^k}(x^k)$ equal to A^k in $B_{R_\delta^{-1}} \setminus B_{R_\delta\, \rho^k}(x^k)$ and satisfying

$$\| |x|^2\ |F_{\tilde{A}^k}| \|_{L^\infty(B_{R_\delta^{-1}}(x^k))} < \sqrt{\delta}\,.$$

This implies in particular

$$\|F_{\tilde{A}^k}\|_{L^{2,\infty}(B_{R_\delta^{-1}}(x^k))} \leq C\ \sqrt{\delta}\,, \tag{VII.16}$$

where $C > 0$ is a constant independent of δ and k. Taking δ small enough we can apply Theorem IV.4 and find a gauge that we denote simply $\overline{A}^k$ and which satisfies

$$\left\{\begin{array}{l}\displaystyle\int_{B_{R_\delta^{-1}}(x^k)}|\overline{A}^k|^2+\sum_{i,j=1}^4|\partial_{x_i}\overline{A}^k_j|^2\,dx^m\le C_G\int_{B_{R_\delta^{-1}}(x^k)}|F_{\overline{A}^k}|^2\,dx^4\\ d^*\overline{A}^k=0\qquad\text{in } B_{R_\delta^{-1}}(x^k)\\ \iota^*_{\partial B_{R_\delta^{-1}}(x^k)}(*\overline{A}^k)=0\,.\end{array}\right.\tag{VII.17}$$

Using the gauge invariance of Yang–Mills integrant (VI.39) together with the C^l estimates (VII.1) of ϵ-regularity theorem (Theorem VII.1) applied to the Uhlenbeck Coulomb gauge $(A^k)^g$ in the annulus $B_{2\,R_\delta\,\rho^k}(x^k)\setminus B_{R_\delta\,\rho^k}(x^k)$ we obtain on $B_{R_\delta^{-1}}(x^k)$

$$\begin{aligned}|d^*_{\overline{A}^k}F_{\overline{A}^k}|&\le C\;|\nabla\chi^k|\;|(A^k)^{g^k}|^2+C\;|\chi^k|\;|(A^k)^{g^k}|^3\\&\le C\;(R_\delta\,\rho^k)^{-3}\;\mathbf{1}_{k,\delta}\;\delta\end{aligned}\tag{VII.18}$$

where $\mathbf{1}_{k,\delta}$ is the characteristic function of $B_{2\,R_\delta\,\rho^k}(x^k)\setminus B_{R_\delta\,\rho^k}(x^k)$. This implies the following estimate

$$\|d^*_{\overline{A}^k}F_{\overline{A}^k}\|_{L^{(4/3,1)}(B_{R_\delta^{-1}}(x^k))}\le\;C\;\int_{B_{2\,R_\delta\,\rho^k}(x^k)\setminus B_{R_\delta\,\rho^k}(x^k)}|F_{\overline{A}^k}|^2\,dx^4\,.\tag{VII.19}$$

where we recall that $L^{(4/3,1)}$ is the Lorentz space whose dual is the *Marcinkiewicz weak* L^4 space: $L^{4,\infty}$. Using the embedding (IV.15) for $p=2$ and $m=4$

$$W^{1,2}(B^4)\hookrightarrow L^{4,2}(B^4)\,,$$

we obtain from (VII.17) the estimate

$$\|\overline{A}^k\|^2_{L^{4,2}(B_{R_\delta^{-1}}(x^k))}\le C_G\int_{B_{R_\delta^{-1}}(x^k)}|F_{\overline{A}^k}|^2\,dx^4\,.\tag{VII.20}$$

Using now one of the embeddings (IV.13):

$$L^2(B^4)\cdot L^{4,2}(B^4)\hookrightarrow L^{4/3,1}(B^4)\,,$$

we obtain

$$\|d^*d\overline{A}^k\|_{L^{4/3,1}(B_{R_\delta^{-1}}(x^k))}\le C_G\int_{B_{R_\delta^{-1}}(x^k)}|F_{\overline{A}^k}|^2\,dx^4\,.\tag{VII.21}$$

Combining this fact with the three lines of (VII.17) together with classical elliptic estimates in Lorentz spaces (see [53]) gives

$$\|F_{\overline{A}^k}\|^2_{L^{2,1}(B_{R_\delta^{-1}}(x^k))}\le C_G\int_{B_{R_\delta^{-1}}(x^k)}|F_{\overline{A}^k}|^2\,dx^4\,.$$

Combining this inequality with the estimate (VII.16) of the curvature in the dual space $L^{2,\infty}$ in the *neck region* we obtain

$$\forall\,\delta>0\quad \exists\, R_\delta>1 \qquad \text{s. t.}$$

$$\limsup_{k\to+\infty}\int_{B_{R_\delta^{-1}}(x^k)\setminus B_{2\,R_\delta\,\rho^k}(x^k)}|F_{A^k}|^2\,dx^4\le C\,\sqrt{\delta}$$

from which we deduce

$$\lim_{R\to+\infty}\lim_{k\to+\infty}\int_{B_{R^{-1}}(x^k)\setminus B_{R\,\rho^k}(x^k)}|F_{A^k}|^2\ dx^4=0\,.$$

This implies

$$|F_{A^k}|^2\ dx^4\rightharpoonup\mu^\infty:=|F_{A^\infty}|^2\ dx^4+\int_{S^4}|F_{D_1^1}|^2\ dvol_{S^4}\ \delta_0\,.$$

This completes the proof of Theorem VII.3. □

VIII. The resolution of the Yang–Mills Plateau problem in super-critical dimensions

VIII.1. The absence of $W^{1,2}$ local gauges

We can reformulate the sequential weak closure of $W^{1,2}$ connections we proved in the previous sections for the dimensions up to 4 in the following way. Let G be a compact Lie group and (M^m,h) a compact Riemannian manifold. Introduce the space of so-called *Sobolev connections* defined by

$$\mathfrak{A}_G(M^m):=\left\{\begin{array}{c}A\in L^2(\wedge^1M^m,\mathcal{G})\ ;\ \displaystyle\int_{M^m}|dA+A\wedge A|_h^2\ dvol_h<+\infty\\ \text{locally}\ \exists\ g\in W^{1,2}\qquad \text{s.t.}\qquad A^g\in W^{1,2}\,.\end{array}\right\}$$

It is clear that any Sobolev $W^{1,2}$-connection of a smooth G-bundle in the classical sense of [15] defines an element in $\mathfrak{A}_G(M^m)$ for $m\le 4$. This follows from the existence of a global representative in L^p $(p<4)$ given by Theorem IV.5.

We now prove the converse: that any element $A\in\mathfrak{A}_G(M^m)$ for $m\le 4$ define a smooth bundle and a $W^{1,2}$-Sobolev connection in this bundle in the classical sense given in [15]. Precisely we have the following result.

Proposition VIII.1. *Let G be a compact Lie group and (M^m,h) be a Riemannian manifold of dimension $m\le 4$. Let $A\in\mathfrak{A}_G(M^m)$ then there exists a smooth bundle E over M^m and there exists a $W^{1,2}$-connection ∇ of E such that A is locally gauge equivalent to ∇.* □

Proof of Proposition VIII.1. We cover M^m by a locally finite family of open balls $B_{\rho_i}(x_i)$ realizing a good covering (in the sense of [6]) such that on each of these sets A has a $W^{1,2}$ representative and $\int_{B_{\rho_i}(x_i)}|F_A|_h^2\,\mathrm{dvol}_h<\varepsilon$ where ε is positive number small enough in order to ensure the existence of a Coulomb $W^{1,2}$-representative.

We denote by g_i the gauge change such that $A^{g_i} = g_i^{-1}dg_i + g_i^{-1}\,A\,g_i$ is such a Coulomb $W^{1,2}$-representative. On $B_{\rho_i}(x_i)\cap B_{\rho_j}(x_i)$ we denote $g_{ij} := g_i^{-1}g_j$. Following word by word the arguments of the proof of Theorem V.5 we have that $g_{ij}\in C^0\cap W^{1,2}(B_{\rho_i}(x_i)\cap B_{\rho_j}(x_i),G)$ and the family satisfy moreover obviously the co-cycle condition $g_{ij} = g_{ik}\,g_{kj}$.

Having taken $\varepsilon>0$ small enough we can ensure that the C^0 norms of the g_{ij} – possibly on the intersection of slightly smaller balls – are small and there exists an equivalent smooth cocycle (h_{ij}) (i.e., $h_{ij}=(\sigma_i)^{-1}g_{ij}\sigma_j$ where $\sigma_i\in C^0\cap W^{1,2}(B_{\rho_i}(x_i),G)$ and $h_{ij}\in C^\infty$.) The cocycle (h_{ij}) associated to the good covering $B_{\rho_i}(x_i)$ defines a smooth bundle E over M^m. We have

$$A^{g_i\sigma_i}\in W^{1,2}(\wedge^1 B_{\rho_i}(x_i),\mathcal{G})\quad\text{ and }\quad A^{g_j\sigma_j}=h_{ij}^{-1}dh_{ij}+h_{ij}^{-1}A^{g_i\sigma_i}h_{ij}\,.$$

Hence by definition $(A^{g_i\sigma_i})$ defines a $W^{1,2}$-connection on E. It is obviously locally gauge equivalent to A. This concludes the proof of the proposition. □

Using the same ingredients[11] as the one used to prove theorem V.2 one establishes the following result which is the boundary-free counterpart of theorem V.2.

Theorem VIII.1. *For $m\le 4$ the space $\mathfrak{A}_G(M^m)$ is weakly sequentially closed below any given Yang–Mills energy level: precisely For any $A^k\in\mathfrak{A}_G(M^m)$ satisfying*

$$\limsup_{k\to+\infty} YM(A^k)=\int_{M^m}|dA^k+A^k\wedge A^k|_h^2\ dvol_h<+\infty\,,$$

there exists a subsequence $A^{k'}$ and a Sobolev connection $A^\infty\in\mathfrak{A}_G(M^m)$ such that

$$d(A^{k'},A^\infty):=\inf_{g\in W^{1,2}(M^m,G)}\int_{M^m}|A^{k'}-(A^\infty)^g|_h^2\ dvol_h\longrightarrow 0\,,$$

moreover

$$YM(A^\infty)\le\liminf_{k'\to 0} YM(A^{k'})\,.$$ □

Remark VIII.1. Observe that the space $\mathfrak{A}_G(M^m)$ contains for instance global L^2 1-forms taking values into the Lie algebra $\mathcal{G}$ that correspond to smooth connections of some principal G-bundle over M^m. If the Yang–Mills energy of a sequence of such smooth connections is uniformly bounded, we can extract a subsequence converging weakly to a Sobolev connection and corresponding possibly to <u>another</u> G-bundle. This possibility of "jumping" from one bundle to another, as predicted for instance in the *concentration compactness* result Theorem VII.2, is encoded in the definition of $\mathfrak{A}_G(M^m)$. □

Because of this weak closure property the space $\mathfrak{A}_G(M^m)$ is the ad hoc space for minimizing Yang–Mills energy in dimension less or equal than four. This is however not the case in higher dimension. We have the following proposition.

[11] Observe that the topology of the underlying bundles defined by each A^k and given by Proposition VIII.1 plays no role in the proof and that the arguments of Theorem V.2 carry over in a straightforward way. Again our approach is to work with analytical objects, i.e., global L^2-1-forms, ignoring the geometric structures that each of this 1-form is inducing.

Proposition VIII.2. *For $m > 4$ the space $\mathfrak{A}_{SU(2)}(M^m)$ is* <u>*not*</u> *weakly sequentially closed below a given Yang–Mills energy level: precisely there exists $A^k \in \mathfrak{A}_{SU(2)}(M^m)$ satisfying*

$$\limsup_{k\to+\infty} YM(A^k) = \int_{M^m} |dA^k + A^k \wedge A^k|^2 \, dx^m < +\infty\,,$$

and a $su(2)$-valued form $A^\infty \in L^2$ such that

$$d(A^{k'}, A^\infty) := \inf_{g\in W^{1,2}(M^m,SU(2))} \int_{M^m} |A^{k'} - (A^\infty)^g|_h^2 \, dvol_h \longrightarrow 0\,,$$

but in every neighborhood U of every point of M^m there is <u>*no*</u> *g such that $(A^\infty)^g \in W^{1,2}(U)$ (i.e., $A^\infty \notin \mathfrak{A}_{SU(2)}(M^m)$).* □

The proposition is not difficult to prove but we prefer to illustrate this fact by a small cartoon. We consider a sequence A^k of smooth 1-forms on B^5, the unit five-dimensional ball, into $su(2)$ such that $\limsup_{k\to+\infty} YM(A^k) < +\infty$. The drawing is representing the flow lines of the divergence free vector field associated to the closed *Chern* 4-*form*: $Tr(F_{A^k} \wedge F_{A^k})$. In this cartoon A^k is weakly converging in L^2 to some limit $su(2)$-valued 1-form A^∞ such that $F_{A^\infty} \in L^2$ but this 1-form satisfies

$$d\left(Tr(F_{A^\infty} \wedge F_{A^\infty})\right) = 8\pi^2 \, [\delta_P - \delta_N] \neq 0\,,$$

where P and N are two distinct points of B^5 – the red dots in the Figure 3 cartoon. Hence for almost every small radii $r > 0$ we have for instance

$$\int_{\partial B^5_r(P)} Tr(F_{A^\infty} \wedge F_{A^\infty}) = 8\,\pi^2\,.$$

Assume there would exist $g \in W^{1,2}$ such that $(A^\infty)^g \in W^{1,2}$ in the neighborhood of P. The gauge invariance of the Chern form gives that

$$\int_{\partial B^5_r(P)} Tr(F_{(A^\infty)^g} \wedge F_{(A^\infty)^g}) = 8\pi^2\,.$$

However we have seen in Section III that the fact that $\iota^*_{\partial B^5_r}(A^\infty)^g$ is in

$$W^{1,2}(\wedge^1 \partial B^5_r, su(2))$$

for almost every r – due to the Fubini theorem – imposes

$$\int_{\partial B^5_r(P)} Tr(F_{(A^\infty)^g} \wedge F_{(A^\infty)^g}) = 0\,,$$

which is a contradiction.

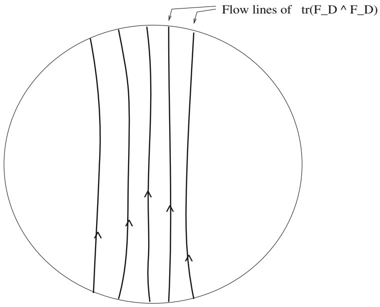


Fig. 1: Sequence of smooth connections time 1

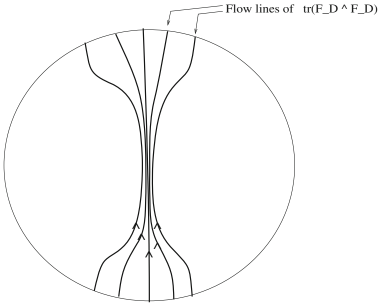


Fig. 2: Sequence of smooth connections time 2

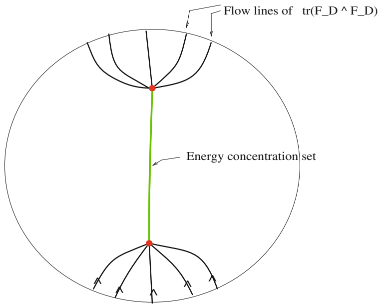


Fig. 3: Sequence of smooth connections – the limit.

In the above cartoon[12] the limiting 1-form A^∞ is a 1-form of a smooth connection but on a $SU(2)$-bundle which is only defined over $B^5 \setminus \{P\} \cup \{N\}$.

[12] A reader interested in having a rigorous implementation of this cartoon could take for instance a sequence of axially symmetric smooth maps u_k from B^5 into S^4 with uniformly bounded $W^{1,4}$-energy and weakly converging to a map u^∞, smooth away from P and N and such that

$$d(u^* \omega_{S^4}) = |S^4| \, [\delta_P - \delta_N]$$

where ω_{S^4} is the volume form on S^4. If ∇_0 is a smooth connection of a non trivial $SU(2)$-bundle E_0 over S^4, the cartoon is realized by A^k a 1-form representing the pull back of the connection by u^k over the trivial bundle $(u^k)^{-1} E_0$

The conclusions we can draw at this point are the following. In dimension $m \leq 3$, the Yang–Mills energy is controlling the bundle in the sense that from a sequence of connections of a given bundle with uniformly bounded L^2-curvature one can extract a weakly converging sequence to a limiting Sobolev connection of the same bundle. In dimension four, under the same assumptions one can extract a weakly converging sequence to a limiting Sobolev connection but possibly from *another bundle*. The last example is illustrating the fact that, starting in five dimension, the Yang–Mills energy is not coercive enough to control the bundle structure: The Yang–Mills energy control does not prevent the corresponding bundle to degenerate and to have local twists at the limit. In fact, instructed from what happens with the Dirichlet energy of maps from B^3 into S^2 one could even construct a sequence of smooth connections whose weak limit is given by a form satisfying

$$\mathrm{supp}\left[d\left(Tr(F_{A^\infty} \wedge F_{A^\infty})\right)\right] = B^5 \quad !$$

This means that the limiting bundle has become everywhere singular on B^5.

In order to find an ad hoc weakly sequentially closed space of connections below any Yang–Mills energy level, we have then to weaken the notion of *Sobolev connections*. Sobolev connections are singular but the underlying bundle was smooth. We see that the L^2 control of the curvature does mot preserve the smoothness of the bundle at the limit in dimension $m > 4$. Since we do not impose anything more that having bounded Yang–Mills energy, we have to relax the notion of bundle, the carrier of the connection, and allow the bundle to have singularities. Contrary to classical differential geometry where the bundle precede the connection, inspired by the "philosophy" of Proposition VIII.1, we shall work with one-forms and "generate" singular bundles from these 1-forms.

The new variational objects will be called *Weak connections*. The quest for finding the right space is similar to the one at the origin of *Geometric Measure Theory* when the class of *integer rectifiable currents* – i.e., submanifolds with singularities in a way – have been produced in order to solve the *Classical Plateau Problem* in super critical dimension $m > 2$. We are looking for a **geometric measure theoretic version of bundles and connections**.

VIII.2. Tian's results on the compactification of the space of smooth Yang–Mills fields in high dimensions

The need of developing a *Geometric measure theoretic* version of bundle and connections beyond the too small class of *Sobolev connections* on smooth bundles has been already encountered in the study of the compactification of the *moduli space* of smooth Yang–Mills fields by G. Tian in [57].

Theorem VIII.2. *Let G be a compact Lie group and A^k be a sequence of $\mathcal{G}$-valued 1-forms in B^m. Assume A^k are all smooth solutions to the Yang–Mills equation*

$$d^*_{A^k} F_{A^k} = 0\,.$$

Assume

$$\limsup_{k\to+\infty}\int_{B^m}|dA^k+A^k\wedge A^k|^2\ dx^m<+\infty\,.$$

Then there exists[13] *a subsequence* $A^{k'}$ *and a limiting* $\mathcal{G}$*-valued* 1*-forms* A^∞

$$d(A^{k'},A^\infty):=\inf_{g\in W^{1,2}(M^m,G)}\int_{M^m}|A^{k'}-(A^\infty)^g|_h^2\ dvol_h\longrightarrow 0\,. \tag{VIII.1}$$

Moreover there exists a $m-4$ *rectifiable closed subset of* B^m, K, *of finite* $m-4$ *Hausdorff measure,* $\mathcal{H}^{m-4}(K)<+\infty$ *such that in*

$$\forall\ B_r(x_0)\subset B^m\setminus K\qquad \exists\ g\in W^{1,2}(B_r(x_0),G)\qquad s.\ t.$$
$$(A^\infty)^g\ \textit{is a smooth solution of Yang–Mills equation in}\ B_r(x_0)\,.$$

and the following weak convergence as Radon measure holds

$$|F_{A^{k'}}|^2\ dx^m\rightharpoonup|F_{A^\infty}|^2\ dx^m+f\ \mathcal{H}^{m-4}\llcorner K\,, \tag{VIII.2}$$

where $\mathcal{H}^{m-4}\llcorner K$ *is the restriction to* K *of the* $m-4$ *Hausdorff measure and* f *is a* $\mathcal{H}^{m-4}\llcorner K$ *measurable function* $f(x)\in L^\infty(K)$. □

This result is very close to a similar result proved in [23] by F.H. Lin for harmonic maps in super-critical dimension $m\geq 3$.

Proof of Theorem VIII.2. The starting point of the proof of Theorem VIII.2 is the following monotonicity formula computed first by P. Price in [38].

Proposition VIII.3 (Monotonicity formula). *Let* $m\geq 4$ *and* A *be a* $\mathcal{G}$*-valued 1-forms in* $B_1^m(0)$ *assume that* A *is a smooth solution to the Yang–Mills equation*

$$d_A^*F_A=0\qquad \textit{in}\ B_1^m(0)\,.$$

then the following monotonicity formula holds for any point p *in* $B_1^m(0)$ *and any* $r<\operatorname{dist}(p,\partial B_1^m(0))$

$$\frac{d}{dr}\left[\frac{1}{r^{m-4}}\int_{B_r^m(p)}|F_A|^2\ dx^m\right]=\frac{4}{r^{m-4}}\int_{\partial B_r^m(p)}|F_A\llcorner\partial_r|^2\,\mathrm{dvol}_{\partial B_r^m}\,, \tag{VIII.3}$$

where $\partial_r=r^{-1}\sum_{i=1}^m x_i\,\partial_{x_i}$ *and*

$$F_A\llcorner\partial_r:=r^{-1}\sum_{i,j=1}^m(F_A)_{ij}\ x_i\ dx_j\,.$$ □

[13]The weak compactness of the sequence of Yang–Mills connection as it is presented in Tian's work [57] does not involve the gauge invariant L^2-distance (VIII.1). Nevertheless by combining the ϵ regularity result VIII.5 with the L^2-control of some gauge by the Yang–Mills energy given in [35] one can recast the weak convergence mentioned in [57] into (VIII.1).

The *monotonicity formula* is a direct consequence of the *stationarity condition* which is satisfied by any smooth critical point of the Yang–Mills Lagrangian. Precisely the *stationarity condition* says

$$\forall\, X \in C^\infty_0(B^m, \mathbb{R}^m) \qquad \frac{d}{dt}\int_{B^m} |\Psi_t^* F_A|^2 \Big|_{t=0} = 0\,, \tag{VIII.4}$$

where Ψ_t is the flow of X.

This condition is equivalent to the following *stationary equation*:

$$\int_{B^m} |F_A|^2 \,\mathrm{div}\, X - 4 \sum_{ijk=1}^{m} \langle (F_A)_{ij}, (F_A)_{kj}\rangle\ \partial_{x_i} X^k \ dx^m\,. \tag{VIII.5}$$

The monotonicity formula is obtained by applying (VIII.4) to the each vector field X of the following form: On the ball $B_r(p)$ the vector field X is equal to the radial one, $X = \partial_r$ for canonical coordinates centered at p which generates the dilations centred at p, and it realizes a smooth interpolation to 0 outside $B_{r+\delta}(p)$ for any $\delta > 0$ (see [38]). Once such a vector field is chosen one computes (VIII.4) and make δ tend to zero. This computation gives then (VIII.3).

The second ingredient of the proof is the extension of the *Coulomb gauge extraction* in dimension larger than 4 to the framework of the so-called Morrey spaces where the $m-4$ densities of Yang–Mills energy are assumed to be small everywhere and at any scale. The following result has been established first in [26] by Yves Meyer and the author of the present survey and independently by Terence Tao and Gang Tian in [54].

Theorem VIII.3 (Coulomb gauge extraction). *Let $m \geq 4$ and G be a compact Lie group, there exists $\varepsilon_{m,G} > 0$ such that for any smooth $\mathcal{G}$-valued 1-forms A in $B^m_1(0)$ satisfying the small Morrey energy condition*

$$\|F_A\|^2_{M^0_{2,4}(B^m_1(0))} := \sup_{p\in B^m_1(0),\ r>0} \frac{1}{r^{m-4}} \int_{B^m_r(p)\cap B^m_1(0)} |F_A|^2 \ dx^m < \varepsilon_{m,G}\,,$$

then there exists a gauge $g \in W^{2,2}(B^m_{1/2}(0), G)$ such that

$$\sup_{p\in B^m_1(0),\ r>0} \frac{1}{r^{m-4}} \int_{B^m_r(p)\cap B^m_1(0)} \sum_{i,j=1}^{m} |\partial_{x_i}(A^g)_j|^2 \leq C\ \|F_A\|^2_{M^0_{2,4}(B^m_1(0))}\,,$$

and

$$\begin{cases} d^*(A)^g = 0 \qquad in\ B^m_1(0) \\ \iota^*_{\partial B^m_1(0)} * (A)^g = 0\,, \end{cases}$$

where the constant C only depends on m and G. □

In this Coulomb gauge any Yang–Mills smooth connection 1-form satisfies

$$\Delta A^g = d^*(A^g \wedge A^g) + [A^g, \llcorner\, dA^g] + [A^g, \llcorner (A^g \wedge A^g)]\,. \tag{VIII.6}$$

We shall make now use of the following generalization of Theorem VI.8 to Morrey spaces.

Theorem VIII.4. *Let $m \geq 4$ and $N \in \mathbb{N}^*$. Let $f \in C^\infty(\mathbb{R}^N \times (\mathbb{R}^m \otimes \mathbb{R}^N), \mathbb{R}^N)$ and let $g \in C^\infty(\mathbb{R}^N, \mathbb{R}^N)$ such that there exists $C > 0$ satisfying*

$$|f(\xi, \Xi)| \leq C\ |\xi|\ |\Xi| \qquad \text{and} \qquad |g(\xi)| \leq C\ |\xi|^3\,. \tag{VIII.7}$$

There exists $\varepsilon > 0$ such that for any u in $L^4 \cap W^{1,2}(B^m, \mathbb{R}^N)$ satisfying

$$\sup_{p \in B_1^m(0),\ r>0} \frac{1}{r^{m-4}} \int_{B_r^m(p) \cap B_1^m(0)} |\nabla u|^2\ dx^m < \varepsilon_{m,G}\,, \tag{VIII.8}$$

and

$$\Delta u = f(u, \nabla u) + g(u)\,, \tag{VIII.9}$$

then we have for any $l \in \mathbb{N}$ the existence of $C_l > 0$ such that

$$\|\nabla^l u\|^2_{L^\infty(B^m_{1/2}(0))} \leq C_l \sup_{p \in B_1^m(0),\ r>0} \frac{1}{r^{m-4}} \int_{B_r^m(p) \cap B_1^m(0)} |\nabla u|^2\ dx^m\,. \tag{VIII.10}$$

□

The proof of this theorem is more or less identical to the one of Theorem VI.8 replacing the different spaces in four dimensions by their *Morrey counterparts* in higher dimension, bearing in mind that Calderòn–Zygmund theory extends with the natural exponents to these spaces (see [27]).

Combining the *monotonicity formula*, Theorem VIII.3 and Theorem VIII.4 applied to the PDE (VIII.6), adapting the arguments we followed for proving the corresponding result – Theorem VII.1 – from the 4-D counterparts of theorem VIII.4 in the conformal dimension four, we obtain the following ϵ-regularity result[14]:

Theorem VIII.5 (ϵ-regularity for smooth Yang–Mills). *Let $m \geq 4$ and G be a compact Lie group, there exists $\varepsilon_{m,G} > 0$ such that for any smooth $\mathcal{G}$-valued 1-forms A in $B_1^m(0)$ satisfying the Yang–Mills equation*

$$d_A^* F_A = 0 \quad \text{in } B_1^m(0)\,.$$

and the small energy condition

$$\int_{B_1^m(0)} |F_A|^2\ dx^m < \varepsilon_{m,G}\,,$$

then there exists a gauge g in which the following estimates holds: for any $l \in \mathbb{N}$ there exists $C_l > 0$ such that

$$\|\nabla^l (A)^g\|^2_{L^\infty(B_{1/2}(0))} \leq C_l \int_{B_1^m(0)} |F_A|^2\ dx^m\,.$$

□

[14] An ϵ-regularity theorem for smooth Yang–Mills fields has first been obtained by H. Nakajima (see [30]). It is however a "gauge invariant result" which gives only an L^∞ bound on the curvature under the small energy assumption but is not providing any control of the connection in some gauge. The proof of Nakajima ϵ-regularity for smooth Yang–Mills fields is following the arguments originally introduced by R. Schoen in [46] for proving the corresponding result for smooth harmonic map. It is using the *Bochner Formula* as a starting point together with the *maximum principle* and the *Moser iteration technique*.

Proof of Theorem VIII.2 *continued.* Let

$$E_r^k := \left\{ p \in B^m \ ; \ \frac{1}{r^{m-4}} \int_{B_r(p)} |F_{A^k}| \ dx^m \geq \varepsilon_{m,G} \right\} ,$$

where $\varepsilon_{m,G}$ is the epsilon in the ϵ-regularity theorem VIII.5. The *monotonicity formula* implies

$$\forall k \in \mathbb{N} \quad \forall\ r < \rho \qquad E_r^k \subset E_\rho^k \,.$$

Hence, by a standard diagonal argument we can extract a subsequence such that $E_{k',2^{-j}}$ converges to a limiting closed set $E_{\infty,2^{-j}}$ which of course satisfy

$$E_{2^{-j-1}}^\infty \subset E_{2^{-j}}^\infty \,.$$

Let

$$K := \bigcap_{j \in \mathbb{N}} E_{2^{-j}}^\infty$$

A classical Federer–Zimmer covering argument gives

$$\mathcal{H}^{m-4}(K) < +\infty \,.$$

With the ϵ-regularity theorem at hand, extracting possibly a further subsequence following a diagonal argument, we can ensure that $A^{k'}$ converges locally away from K in every C^l-norm in the Coulomb gauges constructed in Theorem VIII.3 and we have

$$\mu^{k'} := |F_{A^{k'}}|^2 \ dx^m \rightharpoonup \mu^\infty := |F_{A^\infty}|^2 \ dx^m + \nu \,,$$

where ν is a Radon measure supported in the closed set K. Because of the Radon measure convergence, the monotonicity formula (VIII.3) satisfied by A^k can be transferred to the measure μ^∞

$$\forall\, p \in B_1^m(0) \quad \forall\ B_r^m(p) \subset B_1^m(0) \quad \frac{d}{dr}\left[\frac{\mu^\infty(B_r^m(p))}{r^{m-4}} \ dx^m\right] \geq 0 \,,$$

from which we deduce

$$\theta^{m-4}(\mu^\infty, p) := \lim_{r \to 0} \frac{\mu^\infty(B_r^m(p))}{r^{m-4}} \geq 0 \qquad \text{exists for every } p \in B_1^m(0) \,.$$

Observe that

$$K = \left\{ p \in B_1^m(0) \ ; \ \theta^{m-4}(\mu^\infty, p) > 0 \right\} \,. \tag{VIII.11}$$

Using the monotonicity we have that for any $\rho \in (0,1)$ and any $p \in K \cap B_\rho(0)$

$$\nu(B_r^m(p)) \leq \frac{r^{m-4}}{(1-\rho)^{m-4}} \lim_{k' \to +\infty} \int_{B_1^m(0)} |F_{A^{k'}}|^2 \ dx^m \,.$$

We deduce from this inequality that ν is *absolutely continuous* with respect to $\mathcal{H}^{m-4} \llcorner K$, the restriction to K of the $m-4$-Hausdorff measure. Let $\delta > 0$, define

$$G_\delta := \left\{ p \in B_1^m(0) \ ; \ \delta < \limsup_{r \to 0} \frac{1}{r^{m-4}} \int_{B_r^m(p)} |F_{A^\infty}|^2 \ dx^m \right\} \,.$$

Considering for any $\delta>0$ for any $\eta>0$ and any $p\in G_\delta$ a radius $0<r_p^\eta<\eta$ such that
$$\int_{B^m_{r_p^\eta}(p)}|F_{A^\infty}|^2\ dx^m\ge\frac{\delta}{2}\ (r_p^\eta)^{m-4}.$$
For any $\eta>0$ we extract from the covering $(B^m_{r_p^\eta}(p))_{p\in G_\delta}$ a Besicovitch subcovering $(B^m_{r_i}(p))_{i\in I}$ of G_δ in such a way that there exists an integer $N_m>0$ depending only on m such that each point of B^m is covered by at most N_m balls of this sub-covering. We then have
$$\begin{aligned}\mathcal{H}^{m-4}(G_\delta)&\le\frac{2}{\delta}\limsup_{\eta\to 0}\sum_{i\in I}\int_{B^m_{r_i}(p)}|F_{A^\infty}|^2\ dx^m\\&\le\frac{2\,N_m}{\delta}\limsup_{\eta\to 0}\int_{\mathrm{dist}(x,K)<\eta}|F_{A^\infty}|^2\ dx^m\,.\end{aligned}$$
Since K is closed and has Lebesgue measure zero we deduce that
$$\mathcal{H}^{m-4}\left(\bigcap_{\delta>0}G_\delta\right)=0\,,$$
or in other words
$$\text{for }\mathcal{H}^{m-4}\text{ almost every }p\in B^m\quad\limsup_{r\to 0}\frac{1}{r^{m-4}}\int_{B^m_r(p)}|F_{A^\infty}|^2\ dx^m=0\,.$$
Using the characterization of K given by (VIII.11) and the fact that ν is *absolutely continuous* with respect to $\mathcal{H}^{m-4}\llcorner K$ we deduce from the previous fact that
$$\text{for }\nu\text{ almost every }p\in B^m\quad\lim_{r\to 0}\frac{\nu(B_r(p))}{r^{m-4}}\quad\text{exists and is positive}\,.$$
The following result, which is an important contribution to *Geometric measure theory* was proved by D. Preiss in [37] and answered positively to a conjecture posed by Besicovitch. It permits to conclude the proof of Theorem VIII.2. □

Theorem VIII.6. *Let ν be a Borel non-negative measure in B^m. Assume that ν almost everywhere the n-dimensional density of ν exists and is positive then ν is supported by an n-dimensional rectifiable subset in B^m.* □

In order to have a more complete description of the *compactification of the moduli space* of smooth Yang–Mills fields in supercritical dimension three main questions are left open in this theorem

i) What are the special geometric properties satisfied by the set K?

ii) What is the energy defect $f(x)$?

iii) What is the regularity of A^∞ – modulo gauge – throughout K?

In subcritical dimension $m<4$, due to the analysis we have exposed in the previous sections, we have that $K=\emptyset$, $f=0$ and modulo gauge transformation the limiting connection extends to a smooth Yang–Mills field over the whole ball.

In critical four dimensions these questions are answered by the point removability theorem (Theorem VI.9): the set K is made of isolated points, the function f is the sum of the Yang–Mills energies of Yang–Mills fields over S^4 (see [42]) and, modulo gauge transformation, the limiting connection extends to a smooth Yang–Mills field over the whole ball.

In super-critical dimension it is expected that K with the multiplicity f defines a so-called *stationary varifold* (see [50]). This belief comes from the fact that smooth Yang–Mills fields satisfy the *stationarity condition* (VIII.4) and it is expected that this condition should still be satisfied by the weak limit A^∞ itself, and hence, due to the Radon measure convergence (VIII.2), it would be "transferred" to the measure $f\ \mathcal{H}^{m-4}\llcorner K$. This last fact is equivalent to the stationarity of (K, f).

Regarding the regularity of A^∞ a result of T. Tao and G. Tian [54] asserts that A^∞ is a smooth Yang–Mills connection of a smooth bundle defined away of a closed subset $L \subset K$ satisfying $\mathcal{H}^{m-4}(L) = 0$. This partial regularity result is probably not optimal but this optimality or non-optimality is an open problem (a similar open question exists for *stationary harmonic maps* – see [43]).

The author has proved in [42] that a uniform L^1 control of the *Hessian* of the curvature leads to the energy quantisation along K. Precisely the following theorem holds.

Theorem VIII.7 ([42]). *Let A^k be a sequence of smooth Yang–Mills Fields on B^m satisfying*

$$\limsup_{k\to+\infty} \int_{B^m} |\nabla^2 F_{A^k}| + |F_{A^k}|^2\ dx^m < +\infty\,.$$

Then, modulo extraction of a subsequence, there exists a $m-4$ rectifiable $K \subset B^m$ and a $\mathcal{H}^{m-4}\llcorner K$ measurable function $f \in L^\infty(K)$. such that

$$|F_{A^k}|^2\ dx^m \rightharpoonup |F_{A^\infty}|^2\ dx^m + f\ d\mathcal{H}^{m-4}\llcorner K\,,$$

where F_{A^∞} is the curvature of a smooth Yang–Mills field of a G-bundle over B^m. Moreover

For $\mathcal{H}^{m-4}$- a.e. $x \in K \quad \exists\, N_x \in \mathbb{N} \quad \nabla_1 \cdots \nabla_{N_x}$ and

$$f(x) = \sum_{j=1}^{N_x} \int_{S^4} |F_{\nabla_j}|^2\, \mathrm{dvol}_{S^4}\,,$$

where ∇_j are Yang–Mills connections of some smooth G-bundles over S^4.

Recently, Aaron Naber and Daniele Valtorta in [29] have been able to prove the following uniform L^1 control of the *Hessian* which had been conjectured by the author in [42].

Theorem VIII.8 ([29]). *Let A be a smooth Yang–Mills Fields of a smooth G-bundle over a closed manifold (M^m, h) such that*

$$\int_{M^m} |F_A|^2\, \mathrm{dvol}_h \le \Lambda\,,$$

then

$$\int_{M^m} |\nabla^2 F_A| \,\mathrm{dvol}_h \leq C(\Lambda, \dim(G), (M^m, h))\,. \qquad \square$$

Combining Theorem VIII.7 and Theorem VIII.8 implies the energy quantisation along K (i.e., the conclusion of Theorem VIII.7) for sequences of smooth Yang–Mills connections under Yang–Mills energy control only.

VIII.3. The Ω-anti-self-dual instantons

In four dimension a special class of solutions to the Yang–Mills equation, the *anti-self-dual instantons* have been considered by S. Donaldson in the early 80s to produce invariants of differential four-dimensional manifolds. On a four-dimensional Riemannian manifold (M^4, h) for some given $SU(n)$-bundle over M^m we consider connections A solutions to the equation

$$*_h F_A = -F_A\,. \tag{VIII.12}$$

This equation is issued from an elliptic complex (see [13]) and is the natural generalization in four dimension of the *flat connection equation* $F_A = 0$ on Riemann surfaces considered to classify holomorphic complex structures over such a surface. It defines special solutions to Yang–Mills equation. Indeed taking the covariant exterior derivative with respect to A, using the Bianchi identity $d_A F_A = 0$ one obtains $d^{*_h} F_A = 0$. The *anti-self-dual equation* (VIII.12) is generalized in higher dimension

$$*_h F_A = -\ \Omega \wedge F_A\,, \tag{VIII.13}$$

where Ω is a <u>closed</u> $m-4$ form. Again, due to *Bianchi identity*, by taking the covariant exterior derivative with respect to the connection A, using *Leibnitz identity* on 1-derivations and the fact that Ω is closed, we obtain the Yang–Mills equation $d^{*_h} F_A = 0$. The Ω-anti-self-dual equation is not elliptic in general. There are however special situations of geometric interest when the base manifold has a restricted holonomy.

Hermitian Yang–Mills fields. Let (M^{2n}, h) be an even-dimensional Riemannian manifold. We assume that M^{2n} is equipped with an integrable complex structure J_M – i.e., the brackets operation leaves the space of $1-0$ vector fields of $TM \otimes \mathbb{C}$ invariant

$$\forall\ X, Y \text{ vector fields} \qquad J\,[X - i\,J\,X, Y - i\,J\,Y] = i\,[X - i\,J\,X, Y - i\,J\,Y]\,.$$

Finally we assume that (M^{2m}, h, J_M) is Kähler: $\omega(\cdot,\cdot) := g(\cdot, J_M\cdot)$ is a closed 2-form. It defines a non degenerate 2-form and $\omega^m/m! = \mathrm{dvol}_g$. Let

$$\Omega := \omega^{m-2}/(m-2)!$$

Consider a Hermitian vector bundle E associated to a principal $SU(n)$ bundle over M^m with projection map $\pi\ : E \to M^{2m}$. A connection ∇ is Ω anti self-dual if and only if it satisfies the Hermitian Yang–Mills equations:

$$*F_\nabla = -\,\Omega\wedge F_\nabla \quad\Longleftrightarrow\quad \left\{\begin{aligned} F_\nabla^{0,2} &= 0\\ \omega\cdot F_\nabla^{1,1} &= 0\end{aligned}\right.$$

where $F_\nabla^{0,2}$ (resp. $F_\nabla^{1,1}$) is the $0-2$ (resp. $1-1$) part of the curvature (the space $T^*M\otimes\mathbb{C}$ is decomposed according to the eigenspaces of J_M for the eigenvalues i and $-i$) so the Hermitian Yang–Mills equation

$$\left\{\begin{array}{lll} F_\nabla^{0,2}=0 & \Leftrightarrow & \forall\, X,Y \quad F_\nabla(X-iJX,Y-iJY)=0\\ \omega\cdot F_\nabla^{1,1}=0 & \Leftrightarrow & \displaystyle\sum_{l=1}^m (F_\nabla)_{\epsilon_l,J\,\epsilon_l}=0 \quad \text{where } \omega=\sum_{l=1}^m \epsilon_l\wedge J\epsilon_l\end{array}\right.$$

and ε_l denotes an orthonormal basis of (T^*M^{2m},h).

The *Hermitian Yang–Mills* equation can be interpreted as follows. The *equivariant horizontal distribution of plane* H associated to ∇ defines an almost complex structure J_∇ on E in the following way

$$\forall\,\xi\in E\ \forall X\in T_\xi E \quad J_\nabla(X):=J_E(X^V)+(J_M(\pi_*X))^H\,,$$

we recall that X^V is the projection onto the tangent vertical space (the kernel of the projection π_*) with respect to the horizontal plane H, and J_E is the complex structure on the tangent vertical space defined by the $SU(n)$ structure group of the bundle. The first part of the equation *Hermitian Yang–Mills* can be reformulated as follows

$$\begin{aligned}&\forall\; X,Y \text{ vector fields in } M^{2m} \quad F^{0,2}(X,Y)=0 \quad\Longleftrightarrow\\ &J_\nabla\,[(X-i\,J_M\,X)^H,(Y-i\,J_M\,Y)^H]=i\,[(X-i\,J_M\,X)^H,(Y-i\,J_M\,Y)^H]\,,\end{aligned}$$

which is equivalent to say that J_∇ is <u>integrable</u> and that the Hermitian bundle is holomorphic.

If A is a $su(n)$-valued 1-form representing ∇ in an orthonormal trivialization, this integrability condition implies, by switching to a local holomorphic trivialization, that there is a gauge change g (non-unitary anymore but taking values into $Gl(n,\mathbb{C})$) such that locally

$$g^{-1}\overline{\partial}g+g^{-1}A^{0,1}g=(A^g)^{0,1}=0$$

and the $0-1$ part of the connection ∇ in this holomorphic trivialization coincide with $\overline{\partial}$. Since A is taking values into $su(n)$, we have that $A^{1,0}=-\overline{A^{0,1}}^t$. Thus we obtain the fact that

$$(A^{1,0})^g=h^{-1}\,\partial h\,,$$

where $h:=\overline{g}^t\,g$ is taking values into invertible self-dual matrices. So the so-called "Einstein part" of the equation $\omega\cdot F_\nabla^{1,1}=0$ becomes equivalent to the non-linear elliptic equation[15]

$$\omega \cdot \overline{\partial}[h^{-1}\,\partial h] = 0\,.$$

***SU*(4)-instantons in Calaby–Yau 4-folds.** In high energy physics and later in geometry (see the PhD thesis of C. Lewis, [14, 57]) the following generalization of *instantons* has been introduced. Consider an $SU(n)$ principal bundle P over (M^8, g) a *Calabi–Yau manifold* of complex dimension four – i.e., (M^8, g) has holonomy $SU(4)$. Such a manifold possesses an integrable complex structure J_M for which (M^8, g, J_M) is Kähler and it possesses in addition a global holomorphic $4-0$ form θ of unit norm – unique modulo unit complex number multiplication – and satisfying

$$\theta \wedge \overline{\theta} = \frac{\omega^4}{4!}\,.$$

The 4-form Ω defined by

$$\Omega := 4\,\Re(\theta) + \frac{\omega^2}{2}$$

is closed, parallel and of unit *co-mass* (see [57, Lemma 4.4.1)]: $\forall\ x \in M$

$$1 = \|\Omega\|_*(x) = \sup\left\{ \frac{\langle \Omega, u_1 \wedge u_2 \wedge u_3 \wedge u_4\rangle}{\prod_{i=1}^4 |u_i|} \quad ; \quad u_i \in T_xM \setminus \{0\} \right\}.$$

The holomorphic $4-0$ form θ defines an isometry of the space of $0-2$ forms on M^8 such that

$$\forall\, \alpha\,,\ \beta \in \wedge^{0,2}M^8 \qquad \alpha \wedge *_\theta \beta = \alpha \cdot \beta\ \overline{\theta}\,.$$

Some basic computation gives the Ω-anti self dual equation in this case is equivalent to

$$*F_\nabla = -\ \Omega \wedge F_\nabla \quad \Longleftrightarrow \quad \left\{ \begin{array}{r} (1 + *_\theta)\, F_\nabla^{0,2} = 0 \\ \omega \cdot F_\nabla^{1,1} = 0\,, \end{array} \right.$$

which is also known as the $SU(4)$ *instanton equation.*

VIII.4. Tian's regularity conjecture on Ω-anti-self-dual instantons

Tian's result in the case of Ω anti-self dual instantons for a closed $m-4$-form Ω of co-mass less than 1 is the following.

Theorem VIII.9. *Let (M^m, h) be a compact Riemannian manifold. Let Ω be a smooth $m-4$ closed form in M^m. Assume Ω has co-mass less than* 1

$$\||\Omega|_*\|_{L^\infty(M^m)} \le 1\,,$$

where

$$|\Omega|_*(x) := \sup\left\{ \frac{\langle \Omega, u_1 \wedge \cdots \wedge u_{m-4}\rangle}{\prod_{i=1}^{m-4} |u_i|} \quad ; \quad u_i \in T_xM \setminus \{0\} \right\}.$$

[15] Recall that in a Kähler manifold $\omega \cdot \overline{\partial}\partial = \Delta$

Let E be a Hermitian vector bundle associated to an $SU(n)$-bundle over M^m and let ∇^k be a sequence of smooth $SU(n)$-connections satisfying the Ω anti self dual instantons equation

$$*_h F_{\nabla^k} = -\,\Omega\wedge F_{\nabla^k} \quad \text{in } M^m\,.$$

Then, modulo extraction of a subsequence, there exists a $m-4$ rectifiable closed subset of M^m, K, of finite $m-4$ Hausdorff measure, $\mathcal{H}^{m-4}(K)<+\infty$, a Hermitian bundle E_0 defined over $M^m\setminus K$ and a smooth $SU(n)$ connection ∇^∞ of E_0 such that

$$*_h F_{\nabla^\infty} = -\,\Omega\wedge F_{\nabla^\infty} \quad \text{in } M^m\setminus K\,, \tag{VIII.14}$$

moreover

$$\forall\ B_r(x_0)\subset M^m\setminus K \quad \exists \ \text{a sequence of trivializations s. t.}$$

$$A^{k'}\to A^\infty \quad \text{strongly in } C^l(B^m) \quad \forall l\in\mathbb{N}\,,$$

where $\nabla^{k'}\simeq d+A^{k'}$ in these trivializations and $\nabla^\infty\simeq d+A^\infty$ in a trivialization of E_0 over B^m. Moreover there exists an integer rectifiable current C such that for any smooth $m-4$ form φ on M^m

$$\int_{M^m} Tr(F_{\nabla^{k'}}\wedge F_{\nabla^{k'}})\wedge\varphi \to \int_{M^m} Tr(F_{\nabla^\infty}\wedge F_{\nabla^\infty})\wedge\varphi + 8\pi^2\, C(\varphi)\,,$$

and the current C is calibrated by Ω

$$C(\Omega)=\mathbf{M}(C)=\sup\{C(\varphi)\ ;\ \||\varphi|_*\|_{L^\infty(M^m)}\le 1\}\,,$$

where $\mathbf{M}$ is the mass of the current C. Finally the following convergence holds weakly as Radon measures

$$|F_{\nabla^{k'}}|^2_h\,\mathrm{dvol}_h \rightharpoonup |F_{\nabla^\infty}|^2_h\ dvol_h + 8\pi^2\ \Theta(C)\ \mathcal{H}^{m-4}\llcorner K\,,$$

where $\Theta(C)$ is the integer-valued L^1 function with respect to the restriction of the $m-4$ Hausdorff measure to K and which is giving the multiplicity of the current C at each point. □

Observe that no bound is a priori needed for the Yang–Mills energy of the sequence. We have indeed

$$\begin{aligned} YM(\nabla^k) &= -\int_{M^m} Tr(F_{\nabla^k}\wedge *_h F_{\nabla^k})\,\mathrm{dvol}_h\\ &= \int_{M^m} Tr(F_{\nabla^k}\wedge F_{\nabla^k})\wedge\Omega\,. \end{aligned}$$

Since Ω is closed this integral only depends on the cohomology class of the other closed form $Tr(F_{\nabla^k}\wedge F_{\nabla^k})$ which is the second Chern class of E and which is independent of ∇^k.

Remark VIII.2. Regarding the limiting bundle E_0, it is important to insist on the fact that there is no reason for E_0 to extend through K over the whole manifold M^m as a smooth bundle. Hence there is a priori no meaning to give to ∇^∞ over the whole manifold and therefore the Ω-anti-self dual equation (VIII.14) cannot even hold in a distributional sense throughout K if we do not relax the notion of bundle and connections. □

An important question directly related to the regularity of the limiting configuration (E_0, ∇^∞, C) is the following open problem.

Open problem. *Show that the limiting current C has no boundary:*

$$\partial C = 0\,.$$ □

The resolution of this conjecture would imply that the limiting connection (though defined on $M^m \setminus \mathrm{supp}(C)$ only) satisfies globally the stationary equation VIII.5 (i.e., more precisely the version of (VIII.5) on (M^m, h)). One proves that the open question about proving that $\partial C = 0$ is equivalent to

$$d\,(Tr(F_{\nabla^\infty} \wedge F_{\nabla^\infty})) = 0\,.$$

This last identity should be equivalent to the following

Strong approximation property: Does there exist a sequence D^k of smooth $SU(n)$-connection of smooth Hermitian bundles over M^m such that

$$\lim_{k\to+\infty}\ \inf_{g\in W^{1,2}} \int_{M^m} |\nabla^\infty - (D^k)^g|_h^2 + |F_{\nabla^\infty} - g^{-1} F_{D^k} g|_h^2 \,\mathrm{dvol}_h = 0\,?$$

If we would know that $\partial C = 0$ then C defines a *calibrated integral cycle.* The optimal regularity for *calibrated* and *semi-calibrated* integral cycles of dimension two has been proved respectively in [44] and [12]. Such cycles have at most isolated point singularities. More generally, the calibrated condition implies that such a cycle is homologically *mass minimizing* (see [18]). From this later fact, using Almgren regularity result [2] proved also by C. De Lellis and E. Spadaro ([9], [10] and [11]), we would obtain that C is the integration along a rectifiable set which is a smooth dimension $m-4$ sub-manifold away from a co-dimension two singular set with smooth integer multiplicity away from that set. This result is optimal: integration along holomorphic curves in $\mathbb{C}P^n$ is a calibrated integral cycle for the Fubini–Study–Kähler form and can have isolated singularities which are of co-dimension two within the curve.

In his paper Tian made the following conjecture.

Tian regularity conjecture. *Let (E_0, ∇^∞) be the weak limit of smooth Ω-anti-self dual instantons on a bundle E. Then the limiting bundle E_0 and the limiting connections ∇^∞ extend to smooth bundle, resp. smooth connection, away from a closed co-dimension six set L in M^m.* □

There is one case which has been completely settled and where the conjecture has been proved. This is the case of *Hermitian Yang–Mills* fields. In that case the currents defined by

$$\varphi \longrightarrow \int_{M^{2m}} Tr(F_{\nabla^k} \wedge F_{\nabla^k}) \wedge \varphi$$

is a $(m-2)-(m-2)$ positive cycle (i.e., calibrated by $\omega^{m-2}/(m-2)!$), This condition is of course preserved at the limit and then

$$\varphi \longrightarrow \int_{M^m} Tr(F_{\nabla^\infty}\wedge F_{\nabla^\infty})\wedge\varphi + 8\pi^2\, C(\varphi)$$

is also a $(m-2)-(m-2)$ positive cycle (i.e., calibrated by $\omega^{m-2}/(m-2)!$). The points of non zero density correspond to the support of C and using an important result by Y.T. Siu [51] we obtain that C is the integration along a holomorphic sub-variety of complex co-dimension two. Using now a result by S. Bando and Y.T. Siu [4] we obtain that E_0 extends to an analytic reflexive sheaf over the whole M^{2m} it is then locally free (i.e., it realizes a smooth bundle) away from a closed complex co-dimension three subset of M^{2m} which is included in K (see, for instance, [22]). They also prove a point removability asserting that ∇^∞ defines a smooth connection on the part where the sheaf is free which proves the *Tian regularity conjecture* in the special case of *Hermitian Yang–Mills* fields.

VIII.5. The space of weak connections

As we saw in the previous section, in dimension larger than four, "bundles with singularities" arise naturally as "carriers" of limits of smooth Yang–Mills fields with uniformly bounded energy over smooth bundles. If now we remove the assumption to be Yang–Mills and just follow sequences of connections with uniformly bounded Yang–Mills energy over smooth bundles we have seen in the beginning of this section that the limiting carrying bundle can have twists everywhere on the base! Similarly, taking a sequence of closed sub-manifolds with uniformly bounded volume the limit "escape" from the space of smooth sub-manifold and can be singular. The main achievement of the work of Federer and Fleming has been to introduce a class of objects, the *integral cycles* which complete the space of closed sub-manifold with uniformly bounded volume and which was suitable to solve the *Plateau problem* in a general framework. The purpose of the work [34] is to define a class of *weak bundles* and *weak connections* satisfying a closure property under uniformly bounded Yang–Mills energy and suitable to solve the *Yang–Mills Plateau problem.*

We introduce the following stratified definition.

Definition VIII.2. Let G be a compact Lie group and (M^m,h) a compact Riemannian manifold. For $m\le 4$ the space of *weak connections* $\mathcal{A}_G(M^m)$ is defined to coincide with the space of *Sobolev connections* defined by

$$\mathfrak{A}_G(M^m) := \left\{\begin{array}{l} A\in L^2(\wedge^1 M^m,\mathcal{G})\ ;\ \displaystyle\int_{M^m}|dA+A\wedge A|^2_h\, \mathrm{dvol}_h<+\infty \\ \text{locally } \exists\, g\in W^{1,2} \quad \text{s.t.} \quad A^g\in W^{1,2}\,. \end{array}\right\}$$

For $m > 4$ we define the space of *weak connections* $\mathcal{A}_G(M^m)$ to be

$$\mathcal{A}(M^m) := \left\{ \begin{array}{l} A \in L^2(\wedge^1 M^m, \mathcal{G}) \ ; \ \displaystyle\int_{M^m} |dA + A \wedge A|_h^2 \, \mathrm{dvol}_h < +\infty \\ \forall \ p \in M^m \quad \text{for a.e. } r > 0 \quad \iota^*_{\partial B_r(p)} A \in \mathcal{A}_G(\partial B^m_r(p)) \, , \end{array} \right\}$$

where $B^m_r(p)$ denotes the geodesic ball of center p and radius $r > 0$ and $\iota^*_{\partial B_r(p)} A$ is the restriction of the 1-form A on the boundary of $B^m_r(p)$. □

As an illustration of this space, tt is not difficult to check that the limiting connections ∇^∞ from Tian's closure theorem are in $\mathcal{A}_G$. We have the following result which justifies the previous definition.

Theorem VIII.10. *For any $m \in \mathbb{N}^*$ the space of weak connection $\mathcal{A}_G(B^m)$ is weakly sequentially closed bellow any Yang–Mills energy level. Precisely, let $A^k \in \mathcal{A}_G(B^m)$ such that*

$$\limsup_{k \to +\infty} \int_{B^m} |dA^k + A^k \wedge A^k|^2 \ dx^m < +\infty \, ,$$

then there exists a subsequence k' and $A^\infty \in \mathcal{A}_G(B^m)$ such that

$$\delta(A^{k'}, A^\infty) := \sup_{f \in \mathrm{Lip}_1(B^m, \mathbb{R}^{m-4})} \ \inf_{g \in W^{1,2}(B^m, G)} \int_{B^m} |A^{k'} - (A^\infty)^g \wedge f^* \omega|^2 \ \frac{dx^m}{|f^*\omega|} \longrightarrow 0 \, ,$$

where $\mathrm{Lip}_1(B^m)$ is the space of Lipschitz function on B^m with norm bounded by 1 and $\omega := dx_1 \wedge \cdots \wedge dx_{m-4}$. □

Remark VIII.3. In the statement of Theorem VIII.10 δ can be replaced by the more coercive gauge invariant L^2 distance (VIII.1) when $G = SU(2)$. It is expected this fact to be true at least for $G = SU(n)$ and n arbitrary. It is related to an interesting problem in multilinear algebra (see [35]). □

Remark VIII.4. Theorem VIII.10 can be seen as a **Rellich–Kondrachov** type result for weak connections with L^2-bounded curvatures. □

A proof of Theorem VIII.10 is given in [34] and [35]. It is using the following strong approximation theorem whose proof is rather technical.

Theorem VIII.11. *Let $A \in \mathcal{A}_G(B^m)$ then there exists A^k which are the connection $\mathcal{G}$ 1-forms on B^m associated to smooth connections of a sequence of smooth bundles over B^m minus polyhedral chains of codimension five such that*

$$\lim_{k \to +\infty} \ \inf_{g \in W^{1,2}(B^m, G)} \int_{B^m} |A - (A^k)^g|_h^2 + |F_A - g^{-1} F_{A^k} g|^2 \ dx^m = 0 \, .$$ □

These two last results take their roots in the following weak sequential closure, respectively strong approximation, results proved in [32].

Theorem VIII.12. *Let $p > 1$ and let F_k be a sequence of L^p 2-forms in B^3 such that*

$$\forall \ p \in B^3 \quad \text{for a.e. } B_r(p) \subset B^3 \quad \int_{\partial B_r(p)} F_k \in 2\pi \, \mathbb{Z} \, .$$

Assume $F_k \rightharpoonup F_\infty$ weakly in L^p then

$$\forall\, p \in B^3 \quad \textit{for a.e. } B_r(p) \subset B^3 \quad \int_{\partial B_r(p)} F_\infty \in 2\pi\,\mathbb{Z}\,. \qquad \square$$

The strong approximation result says the following.

Theorem VIII.13. *Let $p > 1$ and let F be an L^p 2-forms in B^3 such that*

$$\forall\, p \in B^3 \quad \textit{for a.e. } B_r(p) \subset B^3 \quad \int_{\partial B_r(p)} F \in 2\pi\,\mathbb{Z}\,.$$

then there exists a sequence A^k of smooth connections of a sequence of smooth $U(1)$ bundles over B^3 minus finitely many points such that

$$\lim_{k\to+\infty} \int_{B^3} |F_{A^k} - F|^p \; dx^3 = 0\,. \qquad \square$$

VIII.6. The resolution of the Yang–Mills Plateau problem in five dimensions

In the present section we are solving the “Schoen–Uhlenbeck type” question introduced for the first time in [48] for harmonic maps. We call it in our framework “Yang–Mills Plateau Problem”. We present the result in five dimension which has appeared in [34] while the higher-dimensional case is under preparation.

In order to solve the Yang–Mills Plateau problem in the five-dimensional ball B^5, we introduce a sub-space of $\mathcal{A}_{SU(n)}(B^5)$ of weak connections admitting a trace.

Let η be a weak connection 1-form of $\mathcal{A}_G(S^4)$ we denote by $\mathcal{A}^\eta_{SU(n)}(B^5)$ the subspace of $\mathcal{A}_{SU(n)}(B^5)$ made of weak connection 1-forms A such that

$$\lim_{r\to 1^-} \inf_{g\in W^{1,2}(S^4,G)} \int_{S^4} |D_r^* A - g^{-1}\, dg - g^{-1}\eta\, g|^2 \; dvol_{S^4}\,,$$

where $D_r^* A$ is the pull-back on $S^4 \simeq \partial B^5_1(0)$ of the restriction of A to $\partial B^5_r(0)$ by the dilation of ratio r. It is proved in [34] that for any $\eta \in \mathcal{A}_G$ the space $\mathcal{A}^\eta_G$ is weakly sequentially closed under any Yang–Mills energy level. Hence we have the following theorem.

Theorem VIII.14 (Existence for YM minimizers). *Let G be a compact Lie group and let η be an arbitrary Sobolev connection 1-form in $\mathcal{A}_G(S^4)$ then the following minimization problem is achieved*

$$\inf\left\{\int_{B^5} |dA + A\wedge A|^2 \; dx^5 \;;\; A \in \mathcal{A}^\eta_G(B^5)\right\}. \qquad \text{(VIII.15)}$$

$\square$

Solutions to this minimization problem will be called *solutions to the Plateau problem* for the boundary connection η.

Once the existence of the minimizer is established it is legitimate to ask about the regularity for these solutions to the *Yang–Mills Plateau problem.* Before to give the optimal regularity we give an intermediate result which holds in the general class of *stationary Yang–Mills* in the space of *weak connections* $\mathcal{A}_G(B^5)$.

Theorem VIII.15 (ϵ-regularity for weak stationary YM fields in 5D). *Let G be a compact Lie group, there exists $\varepsilon_G > 0$ such that for any weak connection 1-forms A in $\mathcal{A}_G(B^5)$ satisfying weakly the Yang–Mills equation*

$$d^*F_A + [A, \llcorner F_A] = 0 \quad \text{in } \mathcal{D}'(B^5_1(0)) \,.$$

assume also that it satisfies the stationarity condition

$$\forall\ X \in C^\infty_0(B^5, \mathbb{R}^5) \qquad \frac{d}{dt}\int_{B^5} |\Psi_t^* F_A|^2 \bigg|_{t=0} = 0\,,$$

where Ψ_t is the flow of X, and the small energy condition

$$\int_{B^5_1(0)} |F_A|^2\ dx^5 < \varepsilon_G\,,$$

then there exists a gauge $g \in W^{1,2}(B^5_{1/2}(0), G)$ in which the following estimates hold: for any $l \in \mathbb{N}$ there exists $C_l > 0$ such that

$$\|\nabla^l (A)^g\|^2_{L^\infty(B^5_{1/2}(0))} \le C_l \int_{B^5_1(0)} |F_A|^2\ dx^5 \,. \qquad \square$$

This theorem is a consequence the *monotonicity formula* deduced from the *stationarity condition* and a weak version of the *Coulomb gauge* extraction theorem VIII.3 which generalizes the work of T. Tao and G. Tian [54] for *admissible connections* or the work of Y. Meyer and the author [26] for *approximable connections* to the general framework of *weak connections* in $\mathcal{A}_G(B^5)$.

With the ϵ-*regularity* result at hand, using a *Luckhaus lemma* together with a *Federer dimension reduction method* following the proof of the regularity result of R. Schoen and K. Uhlenbeck for minimizing harmonic map [48] we obtain the following theorem.

Theorem VIII.16 (Regularity for minimizers of the Yang–Mills Plateau problem). *Let G be a compact Lie group, and let η be the connection 1-form associated to a smooth connection of some G-bundle over ∂B^5 then the minimizers of the Yang–Mills Plateau problem* (VIII.15) *are smooth connections ∇ of a smooth bundle defined on B^5 minus finitely many points. For $G = SU(2)$ we have*

$$d(tr(F_\nabla \wedge F_\nabla)) = \sum_i \varepsilon_i\ \delta_{a_i} \quad \text{with} \quad \varepsilon_i \in \{-1, +1\}\,,$$

where a_i are the point singularities. $\square$

It is proved in [34] that the theorem is optimal in the sense that there exists a smooth $su(2)$-valued 1-form η – defining then a connection of the trivial bundle over S^4 – such that any solution to the Plateau problem (VIII.15) has isolated singularities.

VIII.7. Weak holomorphic structures over complex manifolds

With the definition of weak connection at hands one can reformulate a more general version of **Tian's regularity conjecture** as follows.

Conjecture. *Let (M^m, h) be a compact Riemannian manifold and Ω a closed $m-4$ form of co-mass less than one. Let A be a* **weak connection** *in $\mathcal{A}(M^m)$ for the Lie group G and solving*

$$*_h F_A = -\,\Omega \wedge F_A \qquad \textit{in } \ \mathcal{D}'(M^m)\,.$$

Does A define a smooth connection of some smooth G-bundle over the whole manifold minus a closed subset of co-dimension six.

A case of special interest for instance is given by the space of *Weak Hermitian Yang–Mills $SU(n)$-connections* over a Kähler manifold (M^{2p}, J, ω): the elements in $\mathcal{A}_{SU(n)}(M^{2p})$ satisfying

$$\begin{cases} F_A^{0,2} = 0 \\ \omega \cdot F_A^{1,1} = 0 \end{cases} \qquad \text{in } \ \mathcal{D}'(M^{2p})\,. \tag{VIII.16}$$

In order to study this question it is of interest to first concentrate on density properties within the following space of *weak holomorphic structures*

$$\mathcal{A}^{1,1}_{SU(n)}(M^{2p}) := \left\{A \in \mathcal{A}_{SU(n)}(M^{2p}) \quad ; \quad F_A^{0,2} = 0 \quad \text{a. e.}\right\}\,.$$

The following result asserts that weak holomorphic structures in two complex dimension (the dimension for which the Yang–Mills energy is critical) are defining classical smooth holomorphic structures and are strongly approximable by smooth ones.

Theorem VIII.17 ([31]). *Let (M^4, J, ω) be a Kähler two-dimensional manifold and $A \in \mathcal{A}^{1,1}_{SU(n)}(M^{2p})$ and denote by E the smooth $SU(n)$-bundle defined by A and by ∇ the associated Sobolev connection* (*see Proposition* VIII.1). *Then there exists a smooth holomorphic structure $\mathcal{E}$ on E such that*

$$\nabla = \partial_0 + h^{-1}\partial_0 h + \overline{\partial}_{\mathcal{E}}\,,$$

where h and h^{-1} are $W^{2,p}$-sections of $GL(E)$ for any $p < 2$, ∂_0 and $\overline{\partial}_{\mathcal{E}}$ are respectively the $1-0$ and the $0-1$ parts of the Chern connection associated to $\mathcal{E}$ and the choice of some smooth Hermitian structure on the bundle. Moreover there exists a sequence of smooth connections ∇^k satisfying

$$F^{0,2}_{\nabla^k} = 0\,,$$

and

$$\lim_{k\to+\infty}\ \inf_{g\in W^{1,2}(M^4,G)} \int_{M^4} |A - (\nabla^k)^g|_h^2 + |F_A - g^{-1}F_{\nabla^k} g|_h^2 \ dvol_h = 0\,,$$

where $h(\cdot,\cdot) = \omega(\cdot, J\cdot)$. □

Remark VIII.5. It is natural to expect the above Theorem VIII.17 to be extended in higher dimension but allowing singularities and replacing in the result the smooth holomorphic bundle $\mathcal{E}$ by a the more general notion of coherent sheaf. $\Box$

Finally we are closing this section by mentioning two open problems relating weak $1-1$ curvatures and the strong approximability property.

Open Problem 1. *Let F be a $1-1$ real 2-form in $\mathbb{C}^2$ such that $F \in L^p_{\text{loc}}(\mathbb{C}^2)$ for $1 \leq p$ and satisfying for almost every two spheres S in $\mathbb{C}^2$*

$$\int_S F \in \mathbb{Z}\,. \tag{VIII.17}$$

Prove that F is a closed form[16]

The result should still hold in the almost complex framework. As a matter of illustration we mention the following open question. In [45], Tian and the author of these notes establish the optimal partial regularity result for $W^{1,2}$-*pseudo-holomorphic* maps from an *almost Kähler*[17] four-dimensional real manifold (M^4, J, ω) into a projective space $\mathbb{C}P^n$ assuming $u^*\omega_{\mathbb{C}P^n}$ is closed[18] where $\omega_{\mathbb{C}P^n}$ is the Fubini–Study–Kähler 2-form on $\mathbb{C}P^n$.

A priori only the condition (VIII.17) is satisfied for $F := u^*\omega$ but it is natural to ask whether $d(u^*\omega) = 0$ always hold when u is pseudo-holomorphic and if the closedness assumption made in [45] is redundant or not.

Open Problem 2. *Let (M^{2p}, J, ω) be a p-dimensional complex manifold equipped with a Kähler structure and let $A \in \mathcal{A}^{1,1}_{SU(n)}(M^{2p})$. Does the following identity*[19] hold true

$$d(tr(F_A \wedge F_A)) = 0\,. \tag{VIII.18}$$

$\Box$

[16]If $p \geq 4/3$, using Hölder inequality, one proves that under these assumptions $\int_S F$ is equal to zero for almost every 2-spheres and then F is closed.

[17]The form ω, compatible with J, is closed but J is not necessarily integrable.

[18]The condition

$$d(u^*\omega) = 0$$

is obviously always satisfied for any smooth maps. It is equivalent to the local strong approximability by smooth maps for a map in $W^{1,2}(M^4, \mathbb{C}P^n)$. Observe that it is not satisfied by every map in $W^{1,2}(M^4, \mathbb{C}P^n)$. For the particular case of $W^{1,2}$-*pseudo-holomorphic* maps the local strong approximability is proven in [45] to be equivalent to the stationarity condition. This is very special to the pseudo-holomorphic case. In general there exists minimizing harmonic maps from B^3 into $\mathbb{C}P^1$, and hence stationary, where $d(u^*\omega_{\mathbb{C}P^1}) \neq 0$.

[19]This identity is equivalent to the following strong approximation question: does there exists a sequence of smooth $SU(n)$-connections ∇^k satisfying

$$\lim_{k\to+\infty} \inf_{g\in W^{1,2}(M^4,G)} \int_{M^4} |A - (\nabla^k)^g|^2_h + |F_A - g^{-1} F_{\nabla^k} g|^2_h \; dvol_h = 0\,?$$

If (VIII.18) would always hold true under the assumption that $A \in \mathcal{A}^{1,1}_{SU(n)}(M^{2p})$, we could then deduce from such an approximation property the Tian's regularity conjecture in the particular case of *Weak Hermitian Yang–Mills* $SU(n)$ *connections* over a closed p-Kähler manifold in it's full generality[20].

References

[1] Adams, David R. "A note on Riesz potentials." Duke Math. J. 42 (1975), no. 4, 765–778.

[2] Almgren's Big Regularity Paper, Ed. Vladimir Scheffer, Jean E. Taylor World Scientific Monograph Series in Mathematics: Volume 1,World Scientific, 2000.

[3] Bando, Shigetoshi "Removable singularities for holomorphic vector bundles" Tohoku Math. J. (2) Volume 43, Number 1 (1991), 61–67.

[4] Bando, Shigetoshi; Siu, Yum-Tong "Stable sheaves and Einstein–Hermitian metrics", in "Geometry and Analysis on Complex Manifolds," T. Mabuchi, J. Noguchi, and T. Ochiai, eds., World Scientific, River Edge, NJ, 1994, 39–50.

[5] Bernard, Yann; Rivière, Tristan Energy quantization for Willmore surfaces and applications. Ann. of Math. (2) 180 (2014), no. 1, 87–136.

[6] Bott, Raoul; Tu, Loring *Differential forms in Algebraic Topology* GTM 82, Springer, (1982).

[7] Brezis, Haïm, "Functional Analysis, Sobolev Spaces and Partial Differential Equations" Universitext, Springer 2011.

[8] Bröcker, Theodor; tom Dieck, Tammo Representations of compact Lie groups. Graduate Texts in Mathematics, 98. Springer-Verlag, New York, 1995.

[9] De Lellis, Camillo; Spadaro, Emanuele "Regularity of area minimizing currents I, L^p gradient estimate" Geom. Funct. Anal. 24 (2014), no. 6, 1831–1884.

[10] De Lellis, Camillo; Spadaro, Emanuele "Regularity of area minimizing currents II, center manifold" Ann. of Math. (2) 183 (2016), no. 2, 499–575.

[11] De Lellis, Camillo; Spadaro, Emanuele "Regularity of area minimizing currents III, blow up" Ann. of Math. (2) 183 (2016), no. 2, 577–617.

[12] De Lellis, Camillo; Spadaro, Emanuele; Spolaor, Luca "Regularity theory for 2-dimensional almost minimal currents III: blowup" Ann. of Math. (2) 183 (2016), no. 2, 577–617.

[13] S.K. Donaldson and P.B. Kronheimer "The geometry of Four-Manifolds" Oxford, (1990).

[14] S.K. Donaldson and R.P. Thomas, "Gauge theory in higher dimensions". The geometric universe (Oxford, 1996), 31–47, Oxford Univ. Press, Oxford, 1998.

[15] Freed, Daniel S.; Uhlenbeck, Karen K. Instantons and four-manifolds. Second edition. Mathematical Sciences Research Institute Publications, 1. Springer-Verlag, New York, 1991.

[20]The case of "admissible" *Weak Hermitian Yang–Mills* $SU(n)$ *connections* which are weak limits of smooth ones is already solved by the mean of Siu's results in [51].

[16] Gilbarg, David; Trudinger, Neil S. Elliptic partial differential equations of second order. Reprint of the 1998 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2001.

[17] Grafakos, Loukas "Classical Fourier Analysis." Graduate Texts in Mathematics 249, Springer (2009).

[18] Harvey, Reese; Lawson Blaine. "Calibrated geometries." Acta. Math. **148** (1982), 47–157

[19] Isobe, Takeshi "Topological and analytical properties of Sobolev bundles, I: the critical case." Ann. Glob. Anal. Geom. (35) **3** (2009), 277–337.

[20] Isobe, Takeshi "Topological and analytical properties of Sobolev bundles. II. Higher dimensional cases." Rev. Mat. Iberoamericana, **26** (2010), no. 3, 729–798.

[21] Iwaniec Tadeusz; Martin Gaven John "Geometric Function Theory and Non-linear Analysis" Oxford 2001.

[He] Hélein, Frédéric, Harmonic maps, conservation laws and moving frames, Cambridge Tracts in Mathematics, 150, Cambridge University Press, 2002.

[22] Kobayashi, Shoshichi "Differential Geometry of Complex Vector Bundles," Princeton University Press, 1987.

[23] Lin, Fanghua"Gradient estimates and blow-up analysis for stationary harmonic maps" Ann. Math., **149**, (1999), 785–829.

[24] Lin, Fanghua; Rivière, Tristan "A Quantization property for static Ginzburg–Landau Vortices" Comm. Pure App. Math., **54**, (2001), 206–228.

[25] Lin, Fanghua and Rivière, Tristan "Energy Quantization for Harmonic Maps" Duke Math. Journal, **111** (2002) 177–193.

[26] Meyer, Yves; Rivière, Tristan "A partial regularity result for a class of stationary Yang–Mills Fields in high dimension." Rev. Mat. Iberoamericana, **19** (2003), no. 1, 195–219.

[27] C.B. Morrey "Multiple integrals in the calculus of variations" Springer, New York (1966).

[28] Müller, Christoph; Wockel, Christoph "Equivalences of smooth and continuous principal bundles with infinite-dimensional structure group". Adv. Geom. 9 (2009), no. 4, 605–626.

[29] Naber, Aaron; Valtorta, Daniele "Energy Identity for Stationary Yang–Mills" arXiv:1610.02898.

[30] Nakajima, Hiraku "Compactness of the moduli space of Yang–Mills connections in higher dimensions" J. Math. Soc. Japan 40 (1988), 383–392.

[31] Paunoiu, Alexandru; Rivière, Tristan "Weak Holomorphic Structures over Kähler Surfaces" in preparation.

[32] Petrache, Mircea; Rivière, Tristan "Weak closure of singular abelian L^p-bundles in 3 dimensions." Geom. Funct. Anal., 21 (2011), 1419–1442.

[33] Petrache, Mircea; Rivière, Tristan "Global gauges and global extensions in optimal spaces." Anal. PDE 7 (2014), no. 8, 1851–1899.

[34] Petrache, Mircea; Rivière, Tristan "The resolution of the Yang–Mills Plateau problem in super-critical dimensions" Adv. Math. 316 (2017), 469–540.

[35] Petrache, Mircea; Rivière, Tristan "The Space of Weak Connections in High Dimensions" arXiv:1812.04432.

[36] Petrache, Mircea; Van Schaftingen, Jean "Controlled singular extension of critical trace Sobolev maps from spheres to compact manifolds." Int. Math. Res. Not. IMRN 2017, no. 12, 3647–3683.

[37] Preiss, David "Geometry of measures in R^n: distribution, rectiifiability, and densities, Ann. of Math. 125 (1987), 537–643.

[38] Price, Peter "A Monotonicity Formula for Yang–Mills Fields" manuscripta math. 43, (1983) 131–166.

[39] Rivière, Tristan "Conformally Invariant Variational Problems." arXiv:1206.2116 (2012).

[40] Rivière, Tristan Weak immersions of surfaces with L^2-bounded second fundamental form. Geometric analysis, 303–384, IAS/Park City Math. Ser., 22, Amer. Math. Soc., Providence, RI, 2016.

[41] Rivière, Tristan "Conservation laws for conformally invariant variational problems." Invent. Math., 168 (2007), no. 1, 1–22.

[42] Rivière, Tristan "Interpolation spaces and energy quantization for Yang–Mills fields." Comm. Anal. Geom., 10 (2002), no. 4, 683–708

[43] Rivière, Tristan; Michael Struwe: Partial regularity for harmonic maps and related problems. Comm. Pure and Applied Math., 61 (2008), no. 4, 451–463.

[44] Rivière, Tristan; Tian, Gang "The singular set of 1-1 integral currents". Ann. of Math. (2) 169 (2009), no. 3, 741–794.

[45] Rivière, Tristan; Tian, Gang The singular set of J-holomorphic maps into projective algebraic varieties. J. Reine Angew. Math. 570 (2004), 47–87.

[46] Schoen, Richard "Analytic aspects of the harmonic maps problem", Math. Sci. Res. Inst. Publ. 2 Springer-Verlag, New York, 321–358, 1984.

[47] Schoen, Richard; Uhlenbeck, Karen Approximation theorems for Sobolev mappings, unpublished manuscript.

[48] Schoen, Richard; Uhlenbeck, Karen A regularity theory for harmonic maps. J. Differential Geom. 17 (1982), no. 2, 307–335.

[49] Sedlacek, Steven "A direct method for minimizing the Yang–Mills functional over 4-manifolds", Comm. Math. Phys. **86** no. 4 (1982), 515–527.

[50] Simon, Leon "Lectures on geometric measure theory." Proceedings of the Centre for Mathematical Analysis, Australian National University, 3. Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[51] Siu, Yum-Tong "Analyticity of sets associated to Lelong numbers and the ex- tension of closed positive currents" Invent. Math. 27 (1974), 53–156

[52] Stein, Elias "Singular integrals and differentiability properties of functions" Ann. Math. Studies 63, Princeton University Press (1970).

[53] Stein Elias; Weiss, Guido "Introduction to Fourier analysis on Euclidean spaces." Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971.

[54] Tao, Terence; Tian, Gang. "A singularity removal theorem for Yang–Mills fields in higher dimensions." Journ. Amer. Math. Soc., **17** (2004), no. 3, 557–593.

[55] Tartar, Luc "An introduction to Sobolev spaces and interpolation spaces." Lecture Notes of the Unione Matematica Italiana, 3. Springer, Berlin; UMI, Bologna, 2007.

[56] Taubes, Clifford Henry "Metrics, connections and gluing theorems." Vol. 89. Amer Math. Soc., 1996.

[57] Tian, Gang *Gauge theory and calibrated geometry.* I. Ann. of Math. (2) **151** (2000), no. 1, 193–268.

[58] Uhlenbeck, Karen "Removable singularities in Yang–Mills fields", Comm. Math. Phys. **83** (1982), no. 1, 11–29.

[59] Uhlenbeck, Karen "Connections with L^p bounds on curvature", Comm. Math. Phys. **83** (1982), no. 1, 31–42.

[60] Wang, Yu "Sharp estimate of global Coulomb gauge" https://arxiv.org/abs/1808.01685

[61] White, Brian "Homotopy classes in Sobolev spaces and the existence of energy minimizing maps", Acta Math. 160 (2008) no. 1, 1–17.

Tristan Rivière
Forschungsinstitut für Mathematik
ETH Zentrum
CH-8093 Zürich, Switzerland
e-mail: `tristan.riviere@math.ethz.ch`

Progress in Mathematics, Vol. 333, 381–443

Tian's Properness Conjectures: An Introduction to Kähler Geometry

Yanir A. Rubinstein

Dedicated to Gang Tian on the occasion of his 60*th birthday*

Abstract. This manuscript served as lecture notes for a minicourse in the 2016 Southern California Geometric Analysis Seminar Winter School. The goal is to give a quick introduction to Kähler geometry by describing the recent resolution of Tian's three influential properness conjectures in joint work with T. Darvas. These results – inspired by and analogous to work on the Yamabe problem in conformal geometry – give an analytic characterization for the existence of Kähler–Einstein metrics and other important canonical metrics in complex geometry, as well as strong borderline Sobolev type inequalities referred to as the (strong) Moser–Trudinger inequalities.

Mathematics Subject Classification (2010). 32Q20, 58E11 (primary); 53C25, 53C55, 14J50, 32W20, 32U05 (secondary).

Keywords. Kähler geometry; Kähler–Einstein metrics; Moser–Trudinger inequality; Tian's properness conjectures.

Prologue

Harmonic functions are special. They enjoy a high degree of regularity, and in some vague sense are considered to be more aesthetically pleasing than an arbitrary function. In geometry, one similarly seeks aesthetically pleasing structures on a given space. A typical example is that of an Einstein structure. Among all Riemannian structures Einstein structures are special in many ways; the interested reader is referred to the book by Besse [23].

Harmonic functions can be defined as solutions to the Laplace equation. A fundamental result in analysis is that harmonic functions are also characterized as minimizers of the Dirichlet energy. This result is fundamental in many ways. First, the Dirichlet energy makes sense for functions whose gradient is merely square integrable while Laplace's equation requires two derivatives to exist pointwise. Second, it gives an approach to actually constructing harmonic functions.

Kähler–Einstein metrics can be defined as solutions to a fully nonlinear analogue of the Laplace equation. A nonlinear analogue of the Dirichlet energy was introduced by Mabuchi 30 years ago. A basic aspect of this analogy is that the Euler–Lagrange equation for the Mabuchi functional is precisely the Kähler–Einstein equation. One possible way to view these lectures is as an attempt to explain some aspects of this analogy in more detail. In doing so, we strive to give a quick – and at least partly introductory – course in Kähler geometry.

A second prologue

The Kähler–Einstein problem is also strongly motivated by an analogy with the Yamabe problem, that is, of course, itself motivated by the classical Dirichlet problem described above. The following table serves as an overall guidance to the Kähler–Einstein problem, especially for those familiar with the resolution of the Yamabe problem. Our goal in these lectures is to describe the right column of this table, culminating in a complete analytic description of "Tian's properness conjectures" and "Tian's Moser–Trudinger conjecture" at the bottom right.

A few remarks are in place. First, this table is highly schematic, and its main purpose is to highlight some possible analogies between the two analytic problems. Second, the infimum in the definition of $\mu_{[\omega]}$ is, of course, solely for the analogy, since $\mu_{[\omega]} = \frac{1}{2}\int_M R_{g_{\omega_\varphi}}\omega_\varphi^n / \int_M \omega_\varphi^n$ is a cohomological invariant of the Kähler class. Third, an alternative sufficient and necessary condition that appeared very recently and after these lectures were delivered can now be described in terms of an invariant that is different but related to $\alpha_{[\omega]}$ coming from algebraic geometry and K-stability [35, 55, 99]. Yet, this last characterization is purely algebraic, and so it is less pertinent to the analogy with the Yamabe problem. Finally, one may also discuss the more general problem of constant scalar curvature (csc) Kähler metrics. While this is beyond the scope of the lectures, it is worth mentioning briefly the state-of-the-art on this problem at the time the lectures were given. Indeed, in [45] aside from solving the Kähler–Einstein case we also were able to reduce the general csc problem to the regularity of weak minimizers of the Mabuchi energy. Shortly afterwards, our techniques played a rôle in the resolution of this regularity problem, and hence of the analytic characterization of constant scalar curvature Kähler metrics [18, 34]. Some of these important developments are already described in the survey [42], while others just appeared and seem to involve important new ideas beyond the scope of these lectures. In the same breath let me guide the reader that Darvas' survey [42] complements the present one in the sense that it gives the necessary detail on some of the geometric pluripotential methods that are needed in the proof of the properness conjectures; this is also the reason that I do not go into many of those details here.

Finally, I should of course point out that the analogy between the Kähler–Einstein problem and the Yamabe problem is an old one that goes back to T. Aubin and pursued by many others, e.g., Futaki, Bourguignon, Tian. Indeed, in his thesis work [1], Aubin studied in the very same paper the prescription of curvature in a conformal class and in a Kähler class. In many senses this is a pioneering paper that is largely overlooked nowadays. In it Aubin even solved the Calabi conjecture under the assumption of nonnegative bisectional curvature. Later, already in the first edition of his book, Aubin described the Yamabe and Kähler–Einstein problem side by side [4]. Soon afterward Aubin introduced his functionals I and J which play a crucial role in these lecture notes. The discerning reader of his paper [5] will realize that his familiarity with the Yamabe problem could very well have been crucial in this ingenious construction. Subsequently, Futaki introduced his famous eponymous invariants inspired by those of Kazdan–Warner in the conformal world [56, 63] (and Bourguignon later explained that they both can be constructed in a unified manner [16]; this could actually have been another line in the table in the next page, but it was hard already to fit the table into one page!). Several years later, when Tian introduced his α-invariant he pointed out its relation to the Moser–Trudinger inequality in the conformal world [93, p. 225] and the inspiration the Yamabe problem has had on his resolution of the Kähler–Einstein problem in dimension two [94, p. 104].

	Yamabe problem	Kähler–Einstein problem
structure	Riemannian manifold (M^n, g)	Kähler manifold $(M^{2n}, \mathrm{J}, \omega)$
class	conformal class $(M^n, [g])$ $= \left\{ Fg \,:\, C^\infty(M) \ni F > 0 \right\}$	Kähler class $(M^{2n}, [\omega])$ $= \left\{ \omega + \sqrt{-1}\partial\bar\partial\varphi > 0 : \varphi \in C^\infty(M) \right\}$
problem	(non)-existence of constant scalar curvature metric in a conformal class $(M^n, [g])$	(non)-existence of a Kähler–Einstein metric in a Kähler class $(M^{2n}, [\omega])$
equation	$u^{N-2}g$ has constant curvature μ (here $N := \frac{2n}{n-2}$) $\Updownarrow$ $\mu = u^{N-1}\big(R_g u - (N+2)\Delta_g u\big)$ (here R_g = scalar curvature of g)	$\omega_\varphi := \omega + \sqrt{-1}\partial\bar\partial\varphi$ has constant scalar Ricci curvature μ $\Updownarrow$ $\mu = -\log \dfrac{\omega_\varphi^n}{\omega^n e^{f_\omega}} \Big/ \varphi$ (here $\Delta_\omega f_\omega = \frac{1}{2}(R_{g_\omega} - 2n\mu_{[\omega]})$ where $\mu_{[\omega]}$ is defined below and $g_\omega(\,\cdot\,,\,\cdot\,) := \omega(\cdot, \mathrm{J}\cdot))$
functional	Yamabe energy $u^{N-2}g \mapsto Y(u) :=$ $\dfrac{\int_M \Big((N+2)\lvert\nabla_g u\rvert^2 + R_g u^2\Big) dV_g}{\Big(\int_M u^N dV_g\Big)^{2/N}}$ (invariant under $u \mapsto cu, c > 0$)	Mabuchi energy $\omega_\varphi \mapsto E(\varphi) := \displaystyle\int_M \log \frac{\omega_\varphi^n}{\omega^n e^{f_\omega}} \omega_\varphi^n$ $\displaystyle -\mu \int_M \sqrt{-1}\partial\varphi \wedge \bar\partial\varphi$ $\displaystyle \wedge \sum_{k=0}^{n-1} \frac{\omega^{n-1-k} \wedge {\omega_\varphi}^k}{(n+1)/(k+1)}$ (invariant under $\varphi \mapsto \varphi + c, c \in \mathbb{R}$)
sign invariant	$\displaystyle \mu_{[g]} := \inf_{u>0} \int_M R_{ug} dV_{ug} \Big/ \Big(\int_M dV_{ug}\Big)^{\frac{2}{N}}$	$\displaystyle \mu_{[\omega]} := \inf_{\varphi>0} \frac{1}{2} \int_M R_{g_{\omega_\varphi}} \omega_\varphi^n \Big/ \int_M \omega_\varphi^n$
solution $(\mu \leq 0)$	always exists	always exists
sufficient criterion $(\mu > 0)$	exists if $\mu_{[g]} < \mu_{[g_{S^n}]}$ (Aubin) where $\mu_{[g_{S^n}]} = n(n-1)\mathrm{vol}(S^n(1))^{\frac{2}{n}}$	exists if $\alpha_{[\omega]} > \dfrac{n\mu_{[\omega]}}{n+1}$ (Tian) where $\alpha_{[\omega]} :=$ $\displaystyle \sup_{c>0} \Big\{ c \,:\, \sup_{\substack{\omega_\varphi>0 \\ \sup\varphi=-1}} \int_M e^{-c\varphi}\omega^n < \infty \Big\}$
necessary condition	Aubin's criterion always holds (if $[g] \neq [g_{S^n}]$)	Tian's properness conjectures?
strong borderline Sobolev	Aubin's strong Moser–Trudinger inequality on S^n	Tian's Moser–Trudinger conjecture?

1. Introduction

The main motivation for these lectures are three conjectures: Tian's properness conjectures and Tian's Moser–Trudinger conjecture. Consider the space

$$\mathcal{H} = \{\omega_\varphi := \omega + \sqrt{-1}\partial\bar{\partial}\varphi \,:\, \varphi \in C^\infty(M),\, \omega_\varphi > 0\} \tag{1}$$

of all Kähler metrics representing a fixed cohomology class on a compact Kähler manifold (M, J, ω).

Motivated by results in conformal geometry and the direct method in the calculus of variations, in the 90s Tian introduced the notion of "properness on $\mathcal{H}$" [95, Definition 5.1] in terms of the Aubin nonlinear energy functional J [5] and the Mabuchi K-energy E [69] as follows (both functionals are defined § 3 below, see (7) and (12)).

Definition 1.1. The functional $E : \mathcal{H} \to \mathbb{R}$ is said to be proper if

$$\forall\, \omega_j \in \mathcal{H}, \quad \lim_j J(\omega_j) \to \infty \quad \Longrightarrow \quad \lim_j E(\omega_j) \to \infty. \tag{2}$$

Tian made the following influential conjecture [95, Remark 5.2], [97, Conjecture 7.12]. Denote by $\mathrm{Aut}(M, \mathrm{J})_0$ the identity component of the group of automorphisms of (M, J), and denote by $\mathrm{aut}(M, \mathrm{J})$ its Lie algebra, consisting of holomorphic vector fields.

Conjecture 1.2. *Let (M, J, ω) be a Fano manifold. Let K be a maximally compact subgroup of* $\mathrm{Aut}(M, \mathrm{J})_0$. *Then $\mathcal{H}$ contains a Kähler–Einstein metric if and only if E is proper on the subset $\mathcal{H}^K \subset \mathcal{H}$ consisting of K-invariant metrics.*

Tian's conjecture is central in Kähler geometry since it predicts an analytic characterization of Kähler–Einstein manifolds. Appropriate analogues of this conjecture in conformal geometry are known and were crucial in the solution of the famous Yamabe problem concerning the existence of constant scalar curvature metrics in conformal classes. We briefly discuss this in Section 16.

The conjecture has attracted much attention over the past two and a half decades including motivating much work on equivalence between algebro-geometric notions of stability and existence of canonical metrics, as well as on the interface of pluripotential theory and Monge–Ampère equations. While the algebraic-geometric characterization of Kähler–Einstein manifolds has been finally obtained [35, 99], the analytic characterization of Conjecture 1.2 has remained open. We refer to the surveys [73, 75, 81, 92, 98].

Conjecture 1.2 (which we refer to as the Tian's first properness conjecture) gives a characterization of Kähler–Einstein manifolds in terms of the Mabuchi energy. Thus, it can be seen as the analogue of the properness of the Yamabe energy which led to the resolution of the Yamabe problem. Another central theorem in conformal geometry is Aubin's strong Moser–Trudinger inequality on spheres. Tian's second properness conjecture suggests a Kähler geometry analogue of this inequality on any Kähler–Einstein manifold, and is referred to as the Moser–Trudinger inequality for Kähler–Einstein Fano manifolds, which we now describe.

Denote by Λ_1 the real eigenspace of the smallest positive eigenvalue of $-\Delta_\omega$, and set

$$\mathcal{H}_\omega^\perp := \left\{ \varphi \in \mathcal{H} \,:\, \int \varphi\psi\omega^n = 0,\, \forall \psi \in \Lambda_1 \right\}.$$

When ω is Kähler–Einstein, it is well known that Λ_1 is in a one-to-one correspondence with holomorphic gradient vector fields [57]. Tian made the following conjecture in the 90's [96, Conjecture 5.5], [97, Conjecture 6.23],[98, Conjecture 2.15].

Conjecture 1.3. *Suppose* (M, J, ω) *is Fano Kähler–Einstein. Then for some* C, $D > 0$,

$$E(\varphi) \geq CJ(\varphi) - D, \qquad \varphi \in \mathcal{H}_\omega^\perp.$$

By the end of these lectures we will present results that resolve both Conjectures 1.2 and 1.3. For Conjecture 1.2, the special case when K is trivial has already been known for almost 20 years from the work of Tian and Tian–Zhu [96, 100]. Treating the general case has remained open since. Somewhat surprisingly, Conjecture 1.2 was actually disproved by Darvas and the author recently [45]. More precisely, Theorem 6.5 establishes precisely for which manifolds Conjecture 1.2 holds, giving a converse to a result of Phong et al. [74]. Next, and this is the second main result of [45], an alternative formulation to Conjecture 1.2 is established, giving the sought after analytic characterization for Kähler–Einstein metrics. This is stated in Theorem 12.4. Finally, the Moser–Trudinger inequality for Kähler–Einstein manifolds is established, confirming Conjecture 1.3 [45]. This is stated in Corollary 16.7

We leave out a few relevant topics due to space and time limitations. Notably, we mostly do not delve into the pluripotential theoretic and Bergman kernel aspects of the proof of Theorem 12.4, for which we refer to Darvas' survey that has appeared in the meantime [42]. On the other hand, our treatment is rather self-contained and reviews most of the basics, giving an opportunity for the interested reader for a rapid introduction to current research in Kähler geometry.

2. Kähler and Fano manifolds

In these lectures all manifolds will be assumed to be complex. Complex manifolds are just like topological or differentiable manifolds except that the transition functions between the different charts in the atlas are required to be holomorphic in both directions (i.e., biholomorphic). Thus, all our manifolds will be of even dimension. In dimension two all complex manifolds are also *Kähler*. In higher dimensions, however, the latter condition is rather subtle and reflects the existence of a Riemannian structure highly compatible with the given complex structure – we explain this next.

Let (M, J, g) denote a complex manifold together with a Riemannian metric g on M that is compatible with J in the sense that

$$g(x, y) = g(\mathrm{J}x, \mathrm{J}y), \quad \forall\, x, y \in \Gamma(M, TM),$$

where $\Gamma(M, TM)$ denotes smooth vector fields on M. Since $\mathrm{J}^2 = -I$, the formula $\omega := \omega = g(\mathrm{J}\,\cdot\,, \,\cdot\,)$ defines a 2-form on M. We call (M, J, g) a Kähler manifold if the form ω is a closed 2-form,

$$d\omega = 0.$$

Kähler manifolds have many other equivalent characterizations; we refer the reader to [80, § 2.1.4].

In these lectures we will be interested in the curvature of Kähler manifolds. In particular, we will be interested in trying to understand *when Kähler manifolds admit Einstein metrics.* Just as the Riemannian metric can be transformed using the complex structure into a skew-symmetric form, so can the Ricci curvature tensor Ric g. We denote the Ricci form by

$$\mathrm{Ric}\,\omega := \mathrm{Ric}\ g(\mathrm{J}\,\cdot\,, \,\cdot\,).$$

Thus, on a Kähler manifold the Einstein equation

$$\mathrm{Ric}\ g = cg,$$

transforms to

$$\mathrm{Ric}\,\omega = c\omega. \tag{3}$$

Let $z_1, \ldots, z_n$ be local holomorphic coordinates on a neighborhood in M. In those coordinates express the form ω,

$$\omega = \sqrt{-1} g_{i\bar{j}} dz^i \wedge \overline{dz^j}.$$

As discovered by Schouten [83], Schouten and van Dantzig [84, 85], and Kähler [66] (see [80, p. 35] for more references) the Ricci form then has the following expression

$$\mathrm{Ric}\,\omega = -\sqrt{-1}\partial\bar{\partial} \log \det[g_{i\bar{j}}]. \tag{4}$$

In fact, the proof is not hard. First, recall some useful formulæ:

Exercise 2.1. Let D be a constant coefficient first-order operator defined on some domain in $\mathbb{R}^m$ and let A be a matrix-valued function on the same domain. Then,

$$D \log \det A = A^{ij} DA_{ij}, \quad \text{and} \quad DA^{ij} = -A^{it} DA_{ts} A^{sj},$$

where A^{ij} is the coefficient in the ith row and jth column of the inverse matrix of A.

Therefore,

$$\partial\bar{\partial} \log \det[g_{i\bar{j}}] = -g^{i\bar{t}} g_{t\bar{s},k} g^{s\bar{j}} g_{i\bar{j},\bar{l}} + g^{i\bar{j}} g_{i\bar{j},k\bar{l}}.$$

Now, $d\omega = 0$ implies that both $\partial\omega = \bar{\partial}\omega = 0$. Thus, $g_{i\bar{j},k} = g_{k\bar{j},i}$ and $g_{i\bar{j},\bar{k}} = g_{i\bar{k},\bar{j}}$.

Exercise 2.2. Complete the proof of (4).

Thus, if η is any Kähler form such that locally $\eta = \sqrt{-1}h_{i\bar{j}}\overline{dz^i} \wedge dz^j$ then

$$\operatorname{Ric}\omega - \operatorname{Ric}\eta = \sqrt{-1}\partial\bar{\partial}\log\frac{\det[h_{i\bar{j}}]}{\det[g_{i\bar{j}}]} = \sqrt{-1}\partial\bar{\partial}\log\frac{\eta^n}{\omega^n} \tag{5}$$

is an exact two form on M since $\log\frac{\eta^n}{\omega^n}$ is a globally defined smooth function on M as $\frac{\eta^n}{\omega^n} > 0$. Therefore, the Ricci form of any Kähler metric is not only a closed two-form (as is evident from (4)), but also lies in a fixed cohomology class. Up to a constant factor, this class is called the first Chern class of M and is denoted by $2\pi c_1(M)$.

The point of this discussion is that Einstein metrics on a Kähler manifold can exist only if the equality of cohomology classes

$$\mu[\omega] = 2\pi c_1(M) \tag{6}$$

holds. Now, as a rule of thumb, Einstein metrics of negative Ricci curvature exist in abundance on Riemannian manifolds, while Einstein metrics of positive Ricci curvature are quite harder to come by [23]. Somewhat analogously, it is easier to prove existence of Kähler–Einstein metrics of negative Ricci curvature, i.e., when $\mu < 0$: a fundamental theorem of Aubin and Yau states that then (6) also implies that there exists a unique Kähler–Einstein metric whose cohomology class is $[\omega]$ [2, 3, 105]. Around the same time, Yau also showed that the same is true when $\mu = 0$ [105]. A couple years earlier Aubin proved this under a restrictive assumption on the bisectional curvature [1]. However, when $\mu > 0$ it was shown by Matsushima already in the 50s that (6) is not sufficient. Kähler manifolds for which (6) holds with $\mu > 0$ are called Fano manifolds. Thus, it is natural to ask:

Question 2.3. *When does a Fano manifold admit a Kähler–Einstein metric?*

As just explained, if such a Kähler–Einstein metric exists it must have positive Ricci curvature. (Conversely, if a Kähler manifold admits a Kähler metric of positive Ricci curvature it is Fano.) In these lectures we describe an answer to this question. The key player will be the Mabuchi energy, which we now turn to describe.

3. The Mabuchi energy

Before defining the Mabuchi energy we introduce several other basic functionals.

The two most basic functionals, introduced by Aubin [5], are defined by

$$\begin{aligned} J(\varphi) = J(\omega_\varphi) &:= V^{-1}\int_M \varphi\omega^n - \frac{V^{-1}}{n+1}\int_M \varphi\sum_{l=0}^{n}\omega^{n-l}\wedge\omega_\varphi^l, \\ I(\varphi) = I(\omega_\varphi) &:= V^{-1}\int_M \varphi(\omega^n - \omega_\varphi^n). \end{aligned} \tag{7}$$

Here,

$$V = \int \omega_\varphi^n,$$

is a constant independent of $\omega_\varphi \in \mathcal{H}$. Whenever we integrate without mentioning the domain we mean integrating over M.

Exercise 3.1. Show that V is $n!$ times the volume of M with respect to the Riemannian metric g. (See the end of the proof of Proposition 2.1 in [38] for a solution.)

The notation $J(\varphi) = J(\omega_\varphi)$ (and similarly for I) is justified by the fact that $J(\varphi) = J(\varphi + c)$ for any $c \in \mathbb{R}$. These two functionals, as well as their difference, are mostly equivalent, in the sense that,

$$\frac{1}{n^2}(I - J) \leq \frac{1}{n(n+1)} I \leq \frac{1}{n} J \leq I - J \leq \frac{n}{n+1} I \leq nJ. \tag{8}$$

Remark 3.2. We will be rather sloppy and often say "$\varphi \in \mathcal{H}$" when we really mean $\omega_\varphi \in \mathcal{H}$. However, see Remark 8.17 where we start being more precise.

A closely related functional is the Aubin–Mabuchi functional, introduced by Mabuchi [69, Theorem 2.3],

$$\mathrm{AM}(\varphi) := V^{-1} \int_M \varphi \omega^n - J(\varphi) = \frac{V^{-1}}{n+1} \sum_{j=0}^{n} \int_M \varphi \, \omega^j \wedge \omega_\varphi^{n-j}, \tag{9}$$

Exercise 3.3. Prove the integration by parts formula

$$\int g\sqrt{-1}\partial\bar{\partial} f \wedge \alpha^j \wedge \beta^{n-j-1} = \int f\sqrt{-1}\partial\bar{\partial} g \wedge \alpha^j \wedge \beta^{n-j-1},$$

whenever α, β are smooth closed (1, 1)-forms and $f, g \in C^2(M)$. Then, show that if $\delta \mapsto \varphi(\delta)$ denotes a C^1 curve in $\mathcal{H}$, (in the sense that $\delta \mapsto \varphi(\delta)(x)$ is C^1 map for each $x \in M$, and $\omega_{\varphi(\delta)} \in \mathcal{H}$ for each δ),

$$\frac{d}{d\delta} \mathrm{AM}(\varphi(\delta)) = V^{-1} \int \frac{d}{d\delta} \varphi(\delta) \omega^n_{\varphi(\delta)}. \tag{10}$$

Denote by

$$\mathrm{Ent}(\nu, \chi) = \frac{1}{V} \int_M \log \frac{\chi}{\nu} \chi, \tag{11}$$

the entropy of the measure χ with respect to the measure ν (where here $V = \int_M \chi = \int_M \nu$).

The Mabuchi energy (sometimes also called the K-energy as in Mabuchi's original article)

$$E : \mathcal{H} \to \mathbb{R},$$

is defined by [81, (5.27)],

$$E(\omega_\varphi) = E(\varphi) := \mathrm{Ent}(e^{f_\omega} \omega^n, \omega_\varphi^n) - \mu \mathrm{AM}(\varphi) + \mu V^{-1} \int_M \varphi {\omega_\varphi}^n. \tag{12}$$

Here, f_ω is a smooth function depending on ω that we define next. For historical reasons, let us point that the original definition of Mabuchi [69] is different than (12); the equivalence of the two definitions was first shown by Tian [95, 97] and later also by Chen [32].

Definition 3.4. The Ricci potential of ω, f_ω, satisfies

$$\sqrt{-1}\partial\bar{\partial} f_\omega = \mathrm{Ric}\ \omega - \mu\omega, \tag{13}$$

where it is convenient to require the normalization $\int e^{f_\omega}\omega^n = \int \omega^n$.

Exercise 3.5. Show that

$$\mathrm{AM}(\varphi) = (I - J)(\varphi) + V^{-1}\int \varphi {\omega_\varphi}^n, \tag{14}$$

and therefore the last two terms in (12) equal $-\mu(I-J)(\omega,\omega_\varphi)$, so

$$E(\omega_\varphi) = \mathrm{Ent}(e^{f_\omega}\omega^n, \omega_\varphi^n) - \mu(I-J)(\omega,\omega_\varphi). \tag{15}$$

From this formula, we see that understanding the K-energy essentially means understanding the interplay between the entropy and the Aubin functional $I - J$ (or, the equivalent functionals I or J, recall (8)). This is in some sense the holy grail, the difficult analytical question at the heart of Tian's conjecture. A first (and fundamental) result in this direction is Theorem 3.8 below, however only after much more work do we obtain a clearer picture of this relationship, culminating in Theorem 12.4.

Exercise 3.6. Show that indeed $E(\omega_\varphi) = E(\varphi)$, i.e., that $E(\varphi + C) = E(\varphi)$ for any $C \in \mathbb{R}$.

There is another way to write the K-energy:

$$E(\varphi) := \mathrm{Ent}(\omega^n, \omega_\varphi^n) + s_0\mathrm{AM}(\varphi) - \frac{1}{V}\sum_{j=0}^{n-1}\int_M \varphi\mathrm{Ric}\,\omega \wedge {\omega_\varphi}^j \wedge \omega^{n-1-j}, \tag{16}$$

where $s_0 = V^{-1}\int_M s_\omega\omega^n$ is the average scalar curvature. Of course,

$$V^{-1}\int_M s_\omega\omega^n = V^{-1}\int_M n\mathrm{Ric}\,\omega\wedge\omega^{n-1} = n\mu$$

so $s_0 = n\mu$.

Exercise 3.7. Show that (16) coincides with (12) when $\mu[\omega] = 2\pi c_1(M)$ $(= [\mathrm{Ric}\,\omega])$.

The point of (16) is that it actually makes sense on any Kähler manifold. We will however stick to the first formula in these lectures for simplicity.

3.1. The K-energy when $\mu < 0$

Using Exercise 3.5, Conjecture 1.2 is seen to hold in the case $\mu < 0$ as follows. First, convexity of the exponential function implies that

$$\int \log f d\nu \le \log \int f d\nu,$$

whenever $d\nu$ is a probability measure (so $\int d\nu = 1$), so

$$\mathrm{Ent}(\nu,\chi) = -\int \log\frac{\nu}{\chi}\frac{\chi}{V} \ge -\log\int\frac{\nu}{\chi}\frac{\chi}{V} = 0, \tag{17}$$

i.e., the entropy is always nonnegative. Therefore,

$$E(\varphi) = \mathrm{Ent}(e^{f_\omega}\omega^n, \omega_\varphi^n) - \mu(I-J)(\varphi) \geq \frac{|\mu|}{n} J(\varphi), \tag{18}$$

as desired.

3.2. The K-energy when $\mu \geq 0$

We now describe a technique, due to Tian [97, § 7], to treat some of the cases when $\mu \geq 0$ by showing the entropy itself is always proper. Since

$$\mathrm{Ent}(e^{f_\omega}\omega^n, \omega_\varphi^n) \geq \mathrm{Ent}(\omega^n, \omega_\varphi^n) - \sup f_\omega, \tag{19}$$

it suffices to estimate $\mathrm{Ent}(\omega^n, \omega_\varphi^n)$. Since the functionals I and J are interchangeable as far as properness goes (recall (8)), what we would like to show is:

Theorem 3.8. *There exists a positive β, C such that*

$$\mathrm{Ent}(\omega^n, \omega_\varphi^n) \geq \beta I(\varphi) - C = -\beta V^{-1} \int_M \Big(\varphi - V^{-1} \int_M \varphi\omega^n\Big)\omega_\varphi^n - C.$$

Rewriting the functional I in this way is useful for the following reason:

$$\begin{aligned}
\beta I(\varphi) - \mathrm{Ent}(\omega^n, \omega_\varphi^n) &= \int \log e^{\log \frac{\omega^n}{\omega_\varphi^n} - \beta(\varphi - V^{-1}\int_M \varphi\omega^n)} \omega_\varphi^n / V \\
&\leq \log \int e^{\log \frac{\omega^n}{\omega_\varphi^n}} e^{-\beta(\varphi - V^{-1}\int_M \varphi\omega^n)} \omega_\varphi^n / V \\
&= \log \int e^{-\beta(\varphi - V^{-1}\int_M \varphi\omega^n)} \omega^n / V,
\end{aligned} \tag{20}$$

and so the question reduces to whether there exists a positive β such that the functional

$$\varphi \mapsto \int e^{-\beta(\varphi - V^{-1}\int_M \varphi\omega^n)} \omega^n$$

is uniformly bounded on $\mathcal{H}$. Observe that we have managed to eliminate the dependence on the measure ω_φ^n. To be more precise, the question is now about integrability properties of functions in $\mathcal{H}$ with respect to a *fixed* measure. We treat this question in the next subsection. Before doing so, observe that an affirmative answer implies the K-energy E is proper whenever $\mu = 0$. When $\mu > 0$, using (8), an affirmative answer implies

$$E(\varphi) \geq (\beta - n\mu/(n+1)) I(\varphi). \tag{21}$$

Thus, if β can be taken larger than $n\mu/(n+1)$ then the K-energy is proper even when $\mu \geq 0$.

3.3. Tian's invariant

The preceding question is equivalent to the following:

Question 3.9. *Is*

$$\alpha(M,[\omega]) = \sup\Big\{\beta : \sup_{\varphi\in\mathcal{H}}\int_M e^{-\beta(\varphi - V^{-1}\int_M \varphi\omega^n)}\omega^n < C(\beta) \\ \text{for some constant } C(\beta)>0\Big\} \tag{22}$$

positive?

By definition, the number $\alpha(M,[\omega])$ is an invariant of the Kähler class $[\omega]$. It was introduced by Tian, who answered Question 3.9 affirmatively [93, Proposition 2.1].

Theorem 3.10. $\alpha(M,[\omega]) > 0$.

As explained above, Theorem 3.10 implies Theorem 3.8.

Before going into the detailed proof of this theorem we observe that the last statement of the previous subsection can be stated as follows.

Theorem 3.11. *Suppose* (6) *holds. The K-energy is proper whenever* $\alpha(M,[\omega]) > n\mu/(n+1)$.

Thanks to Theorem 3.10, Theorem 3.11 treats in a unified fashion the negative, zero, and some positive cases.

Remark 3.12. Using (16) instead of (12) one may generalize Theorem 3.11 to cohomology classes nearby $c_1(M)/\mu$, as shown recently by Dervan [47, Theorem 1.3].

We now turn to proving Theorem 3.10. The key is an elementary, but by no means trivial, result on subharmonic functions in the plane from Hörmander's book [64, Theorem 4.4.5]. Denote the ball of radius R about the origin in $\mathbb{C}$ by

$$B_R(0).$$

Theorem 3.13. *Let* $R>0$ *and let* ψ *be a smooth subharmonic function defined on* $B_R(0)\subset\mathbb{C}$, *satisfying*

$$\begin{aligned} \psi(0) &\geq -1, \\ \psi &\leq 0, \ \text{ on } B_R(0). \end{aligned} \tag{23}$$

Then for every $\rho\in[R/2, e^{-1/2}R)$ *there exists a constant* C *depending only on* R,ρ *such that*

$$\int_{B_\rho(0)} e^{-\psi}\sqrt{-1}dz\wedge\overline{dz} \leq C. \tag{24}$$

Proof. Let $\tilde\psi := \psi + 1$ (so that $\tilde\psi \leq 1$ and $\tilde\psi(0)\geq 0$). We prove (24) for $\tilde\psi$ which is the same thing as (24) for ψ up to a factor of e.

The Riesz (or Poisson) representation of a smooth function $f : B_R(0) \to \mathbb{R}$ takes the form

$$\begin{aligned} f(z) = \frac{1}{2\pi} \int_{B_R(0)} \log \left| \frac{Rz - R\zeta}{R^2 - z\bar{\zeta}} \right| \Delta f(\zeta) \frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} \\ + \int_0^{2\pi} \frac{R^2 - |z|^2}{|z - Re^{\sqrt{-1}\theta}|^2} f(Re^{\sqrt{-1}\theta}) \frac{d\theta}{2\pi}. \end{aligned} \tag{25}$$

Now we consider (25) for $f = \tilde{\psi}$ and try to obtain bounds for each of the terms.

Second term: First we show that the second term in (25) for $f = \tilde{\psi}$ is actually itself uniformly bounded when $z \in B_{R/2}(0)$. Indeed, putting $z = 0$ and $f = \tilde{\psi}$ in (25),

$$0 \leq \tilde{\psi}(0) = \frac{1}{2\pi} \int_{B_R(0)} \log \frac{|\zeta|}{R} \Delta\tilde{\psi}(\zeta) \frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} + \int_0^{2\pi} \tilde{\psi}(Re^{\sqrt{-1}\theta}) \frac{d\theta}{2\pi}.$$

Hence

$$2\pi \geq 2\pi - 2\pi\tilde{\psi}(0) = \int_{B_R(0)} \log \frac{R}{|\zeta|} \Delta\tilde{\psi}(\zeta) \frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} + \int_0^{2\pi} (1 - \tilde{\psi}(Re^{\sqrt{-1}\theta}))d\theta. \tag{26}$$

Since $\tilde{\psi} \leq 1$ the second integrand is nonnegative. So is the first, since $\Delta\tilde{\psi} \geq 0$. So each of the terms is nonnegative and hence bounded from above by 2π. Therefore,

$$\int_0^{2\pi} |\tilde{\psi}(Re^{\sqrt{-1}\theta})| d\theta \leq \int_0^{2\pi} (1 - \tilde{\psi}(Re^{\sqrt{-1}\theta}))d\theta + \int_0^{2\pi} 1 \cdot d\theta \leq 2\pi + 2\pi = 4\pi,$$

hence

$$\begin{aligned} \left| \int_0^{2\pi} \frac{R^2 - |z|^2}{|z - \zeta|^2} \tilde{\psi}(Re^{\sqrt{-1}\theta}) \frac{d\theta}{2\pi} \right| &\leq \sup_{z \in B_r(0),\, |\zeta| = R} \frac{R^2 - |z|^2}{|z - \zeta|^2} \int_0^{2\pi} |\tilde{\psi}(Re^{\sqrt{-1}\theta})| \frac{d\theta}{2\pi} \\ &\leq C(r, R) \cdot \frac{1}{2\pi} 4\pi = 6, \end{aligned} \tag{27}$$

where $C(r, R)$ is some constant depending only on r, R.

First term: The first term in (25) is not uniformly bounded, however we will show it is uniformly exponentially integrable in the sense of the statement of the theorem. We split this first term into two parts one of which will be actually uniformly bounded (all we really need is a uniform bound from below):

Claim 3.14. *For each z such that $|z| < r < \rho$ one has*

$$\left| \frac{1}{2\pi} \int_{B_R(0) \setminus B_\rho(0)} \log \left| \frac{Rz - R\zeta}{R^2 - z\bar{\zeta}} \right| \Delta\tilde{\psi} \frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} \right| \leq C,$$

for some constant $C > 0$ depending only on r, ρ, R.

Proof. Recall that in (26) each of the terms was bounded by 2π, so

$$\int_{B_R(0)} \log \frac{R}{|\zeta|} \Delta\tilde{\psi}(\zeta) \frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} \le 2\pi.$$

Thus, as $\log(1+b) \ge Cb$ for some constant $C = C(\epsilon) \in (0,1)$ when $b \in (0,\epsilon)$,

$$\int_{B_R(0)\setminus B_{R(1-\epsilon)}(0)} (C\frac{|R-|\zeta||}{|\zeta|}) \Delta\tilde{\psi}(\zeta) \frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} \le 2\pi.$$

We have,

$$C\int_{B_R(0)} \frac{|R-|\zeta||}{|\zeta|} \Delta\tilde{\psi}(\zeta) \frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} \le 2\pi,$$

in particular,

$$C\int_{B_R(0)} |1-|\zeta|/R| \Delta\tilde{\psi}(\zeta) \frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} \le 2\pi. \tag{28}$$

Now,

$$\Big|\frac{R^2 - z\bar{\zeta}}{Rz - R\zeta}\Big| \in \partial B_1(0), \quad \forall \zeta \in \partial B_R(0),$$

as can be checked from the fact that for each z such that $|z|<1$ the map $\zeta \mapsto \frac{1-z\bar{\zeta}}{z-\zeta}$ is a Möbius map, i.e., sends $\partial B_1(0)$ to itself and then scaling. Thus, if $z \in B_r(0)$ with $r<1$, there exists $C>0$ possibly depending on r, ρ, R such that

$$\Big|\log\Big|\frac{Rz-R\zeta}{R^2 - z\bar{\zeta}}\Big|\Big| \le \begin{cases} C\big|1-|\zeta|/R\big| & \text{for } \zeta \in (R(1-\epsilon), R), \\ C & \text{for } \zeta \in (\rho, R(1-\epsilon)). \end{cases}$$

Then,

$$\begin{aligned} \Big|\frac{1}{2\pi}\int_{B_R(0)\setminus B_\rho(0)} &\log\Big|\frac{Rz-R\zeta}{R^2 - z\bar{\zeta}}\Big| \Delta\tilde{\psi}\frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta}\Big| \\ &\le C\frac{1}{2\pi}\int_{B_R(0)\setminus B_{R(1-\epsilon)}(0)} \big|1-|\zeta|/R\big| \Delta\tilde{\psi}\frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta} \\ &\quad + C\frac{1}{2\pi}\int_{B_{R(1-\epsilon)}(0)\setminus B_\rho(0)} \Delta\tilde{\psi}\frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta}. \end{aligned}$$

The first term on the right-hand side is uniformly bounded by (28). The second term is uniformly bounded by Claim 3.15 below. □

Since $\rho < e^{-1/2}R$ from now and on we write

$$\rho = e^{-1/2-\epsilon}R,$$

with $\epsilon > 0$ small, say $\epsilon = 1/500$, where 500 could be replaced by a generous quantity (cf. [7, Proposition 8.1]). In order to estimate $e^{-\tilde{\psi}}$ we only need to estimate for each z such that $|z| < r$ the exponential term (note the minus sign)

$$\exp\Big(-\frac{1}{2\pi}\int_{B_{e^{-1/2-\epsilon}R}(0)} \log\Big|\frac{Rz-R\zeta}{R^2 - z\bar{\zeta}}\Big| \Delta\tilde{\psi}\frac{\sqrt{-1}}{2} d\zeta \wedge d\bar{\zeta}\Big). \tag{29}$$

This can be done using Jensen's inequality, or just the arithmetic mean-geometric mean inequality. For that need to normalize the measure so that it integrates to 1 (i.e., becomes a probability measure). That is need to divide by

$$a := \frac{1}{2\pi}\int_{B_{e^{-1/2-\epsilon}R}(0)} \Delta\tilde{\psi}\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta},$$

i.e., the mass of the harmonic measure on this ball. It is well known that the mass of the harmonic measure on a compact subset of a ball is uniformly bounded by a constant depending on the distance to the boundary of the ball whenever the function is uniformly bounded from above on the whole ball and its value is fixed at one point. Moreover,

$$a < 2\,!$$

Indeed, this is the reason to choose $\rho = e^{-1/2-\epsilon}R$:

$$\begin{aligned} a = \frac{1}{2\pi}\int_{B_\rho(0)} \Delta\tilde{\psi}\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta} &\le \frac{1}{2\pi}\int_{B_\rho(0)} \frac{\left(2\log\frac{R}{|\zeta|}\right)}{1+2\epsilon}\Delta\tilde{\psi}\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta} \\ &\le \frac{2}{1+2\epsilon}\cdot\frac{2\pi}{2\pi} = \frac{2}{1+2\epsilon}, \end{aligned}$$

since earlier we proved the first term of (26) is bounded by 2π (note all we did was insert a term between the large parenthesis which is bigger than 1).

For an earlier reference we state the following claim whose proof is identical to the computation of a.

Claim 3.15. *For every $\epsilon \in (0,1)$ there is a constant $C = C(\epsilon)$ such that*

$$\frac{1}{2\pi}\int_{B_{R(1-\epsilon)}(0)} \Delta\tilde{\psi}\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta} \le C.$$

Exercise 3.16. Compute the constant $C(\epsilon)$ in the previous claim.

So we come back to (29), and apply the arithmetic mean-geometric mean inequality:

$$\begin{aligned} (29) &= \exp\left(\int_{B_{e^{-1/2-\epsilon}R}(0)} -a\cdot\log\left|\frac{Rz-R\zeta}{R^2-z\bar{\zeta}}\right|\Delta\tilde{\psi}\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta}/(2\pi a)\right) \\ &\le \int_{B_{e^{-1/2-\epsilon}R}(0)} \left|\frac{R^2-z\bar{\zeta}}{Rz-R\zeta}\right|^a \Delta\tilde{\psi}\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta}/(2\pi a) \\ &\le C\int_{B_{e^{-1/2-\epsilon}R}(0)} \frac{1}{|z-\zeta|^a}\Delta\tilde{\psi}\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta}/(2\pi a). \end{aligned}$$

Now this itself may not be bounded, however, it is in L^1 in z – more precisely in $L^1(B_r(0))$ – and this is what we want to show. It is crucial here that $a < 2$ or in other words to chose $\rho = e^{-1/2-\epsilon}R$ earlier. To be precise, we integrate now in z to get

$$
\begin{aligned}
&\int_{B_{1/2R}(0)}\int_{B_{e^{-1/2-\epsilon}R}(0)} \frac{1}{|z-\zeta|^a}\Delta\tilde{\psi}(\zeta)\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta}/(2\pi a)\wedge\frac{\sqrt{-1}}{2}dz\wedge d\bar{z}/2\pi\\
&\quad\le \int_{B_{(1/2+e^{-1/2-\epsilon})R}(0)}\int_{B_{e^{-1/2-\epsilon}R}(0)}\\
&\qquad\times\frac{1}{|\xi|^a}\Delta\tilde{\psi}(\zeta)\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta}/(2\pi a)\wedge\frac{\sqrt{-1}}{2}d\xi\wedge d\bar{\xi}/2\pi\\
&\quad=\frac{2\pi}{2-a}[(1/2+e^{-1/2-\epsilon})R]^{2-a}\int_{B_{e^{-1/2-\epsilon}R}(0)}\Delta\tilde{\psi}(\zeta)\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta}/(2\pi a)\\
&\quad\le\frac{2\pi}{2-a}[(1/2+e^{-1/2-\epsilon})R]^{2-a}=C(\epsilon).
\end{aligned}
$$

Note Fubini's theorem applies thanks to the last estimate, so the change of order of integration is justified, and so the original integral is bounded, concluding the proof of Theorem 3.13. □

Exercise 3.17. Show that the fraction in (27) is bounded above by a constant depending only on r/R as claimed and blows up as r approaches 0. When $r = R/2$ show that this constant is equal to 3 (it is even simpler to see it must be ≤ 4).

Exercise 3.18. Compute the Green kernel of $B_R(0)$ and then derive the Riesz representation formula (25) starting with the identity (cf. [58, § 2.4], [101])

$$f(x) = -\int_{B_R(0)} G(x,y)\Delta f(y)dy + \int_{\partial B_R(0)} \partial_r G(x,y)f(y)dy.$$

Exercise 3.19. Show that Theorem 3.13 holds for any subharmonic function f by using the fact that the Riesz representation (25) holds with the same expression by interpreting $\Delta f(\zeta)\frac{\sqrt{-1}}{2}d\zeta\wedge d\bar{\zeta}$ as the harmonic measure associated to f (a positive measure with respect to which the Green kernel is integrable) [77, Theorem 4.5.1].

Theorem 3.13 can be extended to any dimension as follows [64, Theorem 4.4.5].

Corollary 3.20. *The same result holds in $\mathbb{C}^n$ with constants that might additionally depend on n. In other words, if ψ is a smooth plurisubharmonic function on $B_R(0)\subset\mathbb{C}^n$, satisfying*

$$
\begin{aligned}
\psi(0) &\ge -1,\\
\psi &\le 0, \ \ \text{on } B_R(0),
\end{aligned}
\tag{30}
$$

then for every $\rho\in[R/2, e^{-1/2}R)$ there exists a constant C depending only on R,ρ,n such that

$$\int_{B_\rho(0)} e^{-\psi}\sqrt{-1}dz_1\wedge\overline{dz_1}\wedge\cdots\wedge\sqrt{-1}dz_n\wedge\overline{dz_n}\le C. \tag{31}$$

Remark 3.21. In both Theorem 3.13 and Corollary 3.20 one may drop the smoothness assumption on ψ: indeed convolve ψ with a smooth mollifier and then observe the integrals on the left-hand side converge in the limit.

Proof. Write

$$\begin{aligned}&\int_{B_r(0)\subseteq\mathbb{C}^n} e^{-\psi}\\ &\quad=\int_{\partial B_1(0)\subseteq\mathbb{C}^n} dV_{S_r^{2n-1}}(\lambda)\int_{B_r(0)\subseteq\mathbb{C}}|w|^{2n-2}e^{-\psi(\lambda w)}\frac{\sqrt{-1}}{2}dw\wedge d\bar{w}/2\pi,\end{aligned}\tag{32}$$

and $|w|^{2n-2}\leq r^{2n-2}$ is uniformly bounded, so we can apply the previous result for $n=1$. To obtain (32), note that we are integrating over a $2n+1$-dimensional manifold on the right-hand side and on a $2n$-dimensional one on the left-hand side. We normalize by 2π, the area of S^1 since each point is counted "S^1 times", since if wish to write $z=\lambda w$ with $w\in\mathbb{C}$, $\lambda\in S^{2n-1}$, and $|z|=|w|$, $|\lambda|=1$ then w is only determined in $\mathbb{C}$ up to multiplication by a number in S^1! This is not quite precise since we should normalize by an S^1 of varying radius! So need to actually divide by $2\pi|w|$, however the change of variables introduces a factor $|w|^{2n-1}$; so we get (32). $\square$

Proof of Theorem 3.10. Let the injectivity radius of (M,ω) be $6r$ (so at each point exists a geodesic ball of that radius). Choose an r-net of M, that is a collection of points $\{x_j\}_{j=1}^N$ such that $M=\cup_j B_r(x_j)$. For each $\varphi\in\mathcal{H}$ one has $n+\Delta_\omega\varphi>0$. Hence Green's formula says that [6, Theorem 4.13 (a), p. 108]

$$\varphi(x)=V^{-1}\int_M\varphi\omega^n+V^{-1}\int -G(x,y)\Delta_\omega\varphi(y)\omega^n(y)\leq V^{-1}\int_M\varphi\omega^n+nA_\omega,\tag{33}$$

where $G(x,y)\geq -A_\omega$ and $\int_M G(x,y)\omega^n(y)=0$ for each $x\in M$. Note that A_ω is a constant depending only on (M,ω), in other words, the Green kernel is uniformly bounded from below [6, Theorem 4.13 (d), p. 108]. Since we are also assuming $\sup\varphi=0$, we obtain (the right-hand side of (33) is independent of x)

$$V^{-1}\int_M\varphi\omega^n+nA_\omega\geq 0.$$

Hence, since φ is non-positive,

$$\int_{B_r(x_j)}\varphi\omega^n\geq\int_M\varphi\omega^n\geq -nA_\omega V,$$

that is,

$$\sup_{B_r(x_j)}\varphi\geq\frac{-nA_\omega V}{\mathrm{vol}(B_r(x_j))}.\tag{34}$$

Choose a local Kähler potential ψ_j satisfying $\sqrt{-1}\partial\bar{\partial}\psi_j=\omega|_{B_{5r}(x_j)}$. Let $C_2:=\sup_j\sup_{B_{5r}(x_j)}\psi_j$. Therefore, since $\varphi\leq 0$,

$$\psi_j+\varphi\leq C_2,\ \text{on}\ \ B_{5r}(x_j).$$

Also, let $y_j \in B_r(x_j)$ be such that (using (34))

$$\varphi(y_j) \geq \frac{-nA_\omega V}{\mathrm{vol}_\omega(B_r(x_j))}.$$

We may also assume without loss of generality that $\psi_j(y_j) = 0$, otherwise we add a constant to ψ_j (and then C_2 possibly increases).

Now the function $\varphi+\psi_j$ is plurisubharmonic (psh for short) on $B_{5r}(y_j)$ (recall that φ is not psh only ω-psh so we add to it a local potential for ω in order to be able to apply Hörmander's result). Also, since $B_{4r}(y_j) \subseteq B_{5r}(x_j)$, we have that on the set $B_{4r}(y_j)$ it holds

$$\varphi + \psi_j - C_2 \leq 0$$

$$(\varphi + \psi_j - C_2)(y_j) \geq -C_2 - \frac{nA_\omega V}{\mathrm{vol}(B_r(x_j))}.$$

Therefore can apply Hörmander's result to $f_j := (\varphi + \psi_j - C_2)/(C_2 + \frac{nA_\omega V}{\mathrm{vol}(B_r(x_j))})$ on $B_{4r}(y_j)$ (note $C_2 \geq 0$, $A_\omega > 0$, so we are not dividing by zero), namely obtain that

$$\int_{B_{2r}(y_j)} e^{-f_j}\ \omega^n\big|_{B_{2r}(y_j)} < C_j$$

(instead of $2r$ could have taken any number in the range $[2r, e^{-1/2}4r)$). Patching these up, using the fact that $B_{2r}(y_j) \supseteq B_r(x_j)$ and M is covered by the latter, we obtain that regardless of φ, one has

$$\int_M e^{-a\varphi}\omega^n < C,$$

where $a := \min_j \frac{1}{\left(C_2 + \frac{nA_\omega V}{\mathrm{vol}(B_r(x_j))}\right)}$, and consequently $\alpha(M, [\omega]) \geq a > 0$. □

4. The Kähler–Einstein equation

The Kähler–Einstein equation (3) is a fourth-order equation in terms of the Kähler potential. The remarkable formula (4) for the Ricci form, however, allows to integrate it to a second-order equation. Indeed, subtracting $\mathrm{Ric}\,\omega$ from both sides of the equation and using (5) yields

$$\mathrm{Ric}\ \omega_\varphi - \mathrm{Ric}\,\omega = \sqrt{-1}\partial\bar\partial \log \frac{\omega^n}{\omega_\varphi^n} = \mu\omega_\varphi - \mathrm{Ric}\,\omega = \mu\sqrt{-1}\partial\bar\partial\varphi - \sqrt{-1}\partial\bar\partial f_\omega,$$

where f_ω is called the Ricci potential of ω, and satisfies $\sqrt{-1}\partial\bar\partial f_\omega = \mathrm{Ric}\ \omega - \mu\omega$, where it is convenient to require the normalization $\int e^{f_\omega}\omega^n = \int \omega^n$. We thus obtain the Kähler–Einstein equation,

$$\omega_\varphi^n = \omega^n e^{f_\omega - \mu\varphi}, \qquad \text{on } M \tag{35}$$

for a global smooth function φ (called the Kähler potential of ω_φ relative to ω). The function f_ω, in turn, is given in terms of the reference geometry and is thus known. Observe that, strictly speaking, the right-hand side of (35) should be

$$\omega^n e^{f_\omega - \mu\varphi + C}$$

For some constant C; whenever $\mu \neq 0$ we can incorporate the constant C by subtracting C/μ from φ since the left-hand side of (35) is invariant under this. When $\mu = 0$, the constant C must be zero by (13).

Note also that the solution φ is only determined up to a constant when $\mu = 0$, while it is uniquely determined when $\mu \neq 0$ by (13).

We close this section by noting a relationship between the K-energy and Kähler–Einstein metrics: the Euler–Lagrange equation of the K-energy is precisely the Kähler–Einstein equation. Indeed,

$$\begin{aligned}
\frac{d}{d\epsilon}\Big|_{\epsilon=0} E(\varphi(\epsilon)) &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} \mathrm{Ent}(e^{f_\omega}\omega^n, \omega^n_{\varphi(\epsilon)}) \\
&\quad - \mu V^{-1}\int_M \dot\varphi\omega^n_\varphi + \mu V^{-1}\int_M \dot\varphi\omega^n_\varphi + \mu V^{-1}\int_M \varphi\Delta\dot\varphi\omega^n_\varphi \\
&= V^{-1}\int_M \left(\Delta_{\omega_\varphi}\dot\varphi + \log\frac{\omega^n_\varphi}{e^{f_\omega}\omega^n}\Delta\dot\varphi\right)\omega^n_\varphi + \mu V^{-1}\int_M \varphi\Delta\dot\varphi\omega^n_\varphi \qquad (36)\\
&= V^{-1}\int_M \left(\log\frac{\omega^n_\varphi}{e^{f_\omega}\omega^n}\Delta\dot\varphi\right)\omega^n_\varphi + \mu V^{-1}\int_M \varphi\Delta\dot\varphi\omega^n_\varphi \\
&= -V^{-1}\int_M \dot\varphi\Delta_{\omega_\varphi} f_{\omega_\varphi}\omega^n_\varphi,
\end{aligned}$$

where we used Exercise 2.1 and the following exercise.

Exercise 4.1. Show that the Ricci potential satisfies the following equation

$$f_{\omega_\varphi} = \log\frac{e^{f_\omega}\omega^n}{\omega^n_\varphi} - \mu\varphi - \log V^{-1}\int_M e^{f_\omega - \mu\varphi}\omega^n. \qquad (37)$$

Thus, the Euler–Lagrange equation of the K-energy is precisely

$$\Delta_{\omega_\varphi} f_{\omega_\varphi} = 0, \quad \text{i.e., as } M \text{ is compact} \quad f_{\omega_\varphi} = \text{const},$$

that, recalling Definition 3.4, means that $\mathrm{Ric}\,\omega_\varphi = \mu\omega_\varphi$.

5. Properness implies existence

In this section we prove the easier part of Conjecture 1.2:

Theorem 5.1. *If the Mabuchi energy E is proper on $\mathcal{H}^K$ then there exists a K-invariant Kähler–Einstein metric in $\mathcal{H}$.*

This result is due to Tian [96], even though it is in some sense already implicit in Ding–Tian [50]. The proof we give follows the same ideas as in the original proof, with some simplifications in the presentation.

In particular, in combination with Theorem 3.11, we obtain as a corollary a theorem of Tian [93, Theorem 2.1]:

Corollary 5.2. *Let $\mu > 0$ and suppose that* (6) *holds. Whenever*

$$\alpha(M, 2\pi c_1(M)/\mu) > n\mu/(n+1)$$

there exists Kähler–Einstein metric cohomologous to $2\pi c_1(M)/\mu$.

We also obtain as a corollary the theorems of Aubin and Yau:

Corollary 5.3. *Let $\mu \leq 0$ and suppose that* (6) *holds. Then there exists a unique Kähler–Einstein metric cohomologous to ω.*

Remark 5.4. The proof of Corollary 5.3 we will give will not directly use the fact that the K-energy is proper whenever $\mu \leq 0$ (which holds according to Theorem 3.11). In fact, the proof of Corollary 5.3 will be more or less a step in the proof of Theorem 5.1.

5.1. A two-parameter continuity method

We will give a somewhat nonstandard proof of Theorem 5.1 using a two-parameter continuity method instead of the more standard proofs that use one-parameter continuity methods or the Ricci flow equation. Namely, we consider the two-parameter family of equations,

$$\omega_\varphi^n = e^{tf_\omega + c_t - s\varphi}\omega^n, \quad c_t := -\log \frac{1}{V}\int_M e^{tf_\omega}\omega^n, \quad (s,t) \in A, \tag{38}$$

where

$$A := (-\infty, 0] \times [0,1] \ \cup [0,\mu] \times \{1\}$$

is the parameter set – the union of a semi-infinite rectangle and an interval, and show that there exists a unique solution $\varphi(s,t)$ for each $(s,t) \in A$ once we require that the solution be continuous in the parameters s, t. Of course, we are really looking to show the existence of $\varphi(\mu, 1)$. Thus, the strategy is to first construct $\varphi(s,t)$ for other values of (s,t) for which existence is easier to show and then perturbing the equation and eventually arriving at the equation for the values $(\mu, 1)$. Hence, the name 'continuity method.'

To show existence for the sub-rectangle $(-\infty, 0] \times [0,1]$ is easier and is precisely what proves Corollary 5.3. Also, one could restrict to the sub-rectangle $(-S, \mu] \times [0,1]$ for any value $S > 0$ as far as proving the existence theorems is concerned. Working on A requires no more work and is somewhat more natural and canonical, since the value $s = -\infty$ corresponds, in a sense that can be made precise [104, Proposition 7.3],[65, § 9],[13], to the initial reference metric ω. Thus, one can view this continuity method as starting with the given reference metric and deforming it to the Kähler–Einstein metric. In fact, the one-parameter continuity method with $t = 1$ fixed and s varying between $-\infty$ and μ can be viewed as the continuity method analogue of the Kähler–Ricci flow, and is called the Ricci continuity method, introduced in [79] and further developed in [65]. One of the reasons we work with the two-parameter family in these lectures is that then the existence of solutions for some parameter values is automatic. Indeed,

$$\varphi(s,0)=0, \qquad s\in(-\infty,0]. \tag{39}$$

Working with the one-parameter Ricci continuity method is harder since it is non-trivial to show the existence of solutions for some parameter value $(s,1)$. The full strength of the Ricci continuity method however goes beyond that of the two-parameter family in that the former can be used to show existence of Kähler–Einstein edge metrics, which are a natural generalization of Kähler–Einstein metrics that allows for a singularity along a complex submanifold of codimension one. In that context the two-parameter continuity method does not always seem to work.

Exercise 5.5. Show that for each $(s,t)\in A$,

$$\operatorname{Ric}\omega_{\varphi(s,t)}=(1-t)\operatorname{Ric}\omega+s\omega_{\varphi(s,t)}+(\mu t-s)\omega. \tag{40}$$

Note that this implies that if indeed, as claimed, $\varphi(s,1)\to 0$ as $s\to-\infty$ then $f_\omega - s\varphi$ should be small, i.e., $\varphi\approx f_\omega/s$ in that regime.

5.2. Openness

Let

$$\mathrm{PSH}(M,\omega)=\{\varphi\in L^1(M,\omega^n)\,:\,\varphi \text{ is upper semicontinuous and } \omega_\varphi\geq 0\}$$

denote the set of ω-plurisubharmonic functions on M.

Define $M_{s,t}: C^{2,\gamma}\cap\mathrm{PSH}(M,\omega)\to C^{0,\gamma}$ by

$$M_{s,t}(\varphi):=\log\frac{\omega_\varphi^n}{\omega^n}-tf_\omega+s\varphi-c_t, \quad (s,t)\in A.$$

If $\varphi(s,t)\in C^{2,\gamma}\cap\mathrm{PSH}(M,\omega)$ is a solution of (38), we claim that its linearization

$$DM_{s,t}|_{\varphi(s,t)}=\Delta_{\varphi(s,t)}+s: C^{2,\gamma}\to C^{0,\gamma}, \quad (s,t)\in A, \tag{41}$$

is an isomorphism when $s\neq 0$ and $s<\mu$. If $s=0$, this map is an isomorphism if we restrict on each side to the codimension one subspace of functions with integral equal to 0 with respect to $\omega^n_{\varphi(0,t)}$. Furthermore, we also claim that $C^{2,\gamma}\cap\mathrm{PSH}(M,\omega)\times A\ni(\varphi,s,t)\mapsto M_{s,t}(\varphi)\in C^{0,\gamma}$ is a C^1 mapping. Given these claims, the Implicit Function Theorem then guarantees the existence of a solution $\varphi(\tilde{s},\tilde{t})\in C^{2,\gamma}$ for all $(\tilde{s},\tilde{t})\in A$ sufficiently close to (s,t). This solution must necessarily be contained in $\mathrm{PSH}(M,\omega)$ since $M_{s,t}(\varphi(s,t))=0$ means that (38) holds, so in particular $\omega^n_{\varphi(s,t)}>0$, and by the continuity of $\omega_{\varphi(s,t)}$ in the parameters it follows that all the eigenvalues of the metric stay positive along the deformation.

We concentrate on the first claim, since the second claim is easier.

Now, $DM_{s,t}$ is an elliptic operator and there is a classical and well-developed theory for those kind of operators acting on Hölder spaces [58]. In particular, such an operator has a generalized inverse, or Green kernel. Also, it is Fredholm of index 0. Using the existence of a Green kernel shows that $C^{2,\gamma}$ decomposes as a direct sum $\{Gf : f \in C^{0,\gamma}\} \oplus K_{s,t}$, where $K_{s,t}$ denotes the kernel of the operator $DM_{s,t}$. Thus, whenever $K_{s,t} = \{0\}$ the operator is an isomorphism.

The nullspace of $DM_{s,t}$ is clearly trivial when $s < 0$ since the spectrum of $\Delta_{\varphi(s,t)}$ is contained in $(-\infty, 0]$. When $s = 0$ the nullspace consists of the constants and so is an isomorphism when restricted to functions of zero average. Thus, the claim is verified whenever $s \le 0$. To deal with the case when s is positive the following lemma is needed.

Lemma 5.6. *Suppose that $M_{s,1}(\varphi(s,1)) = 0$ and that $s \in (0, \mu)$. Then the spectrum of $\Delta_{\varphi(s,1)}$ is contained in $(-\infty, -s)$.*

Proof. Let ψ be an eigenfunction of $\Delta_{\omega_{\varphi(s,t)}}$ with eigenvalue $-\lambda_1$. By standard theory, ψ is smooth. The Bochner–Weitzenböck formula states that

$$\frac{1}{2}\Delta_g|\nabla_g f|_g^2 = \mathrm{Ric}\,(\nabla_g f, \nabla_g f) + |\nabla^2 f|_g^2 + \nabla f \cdot \nabla(\Delta_g f).$$

Since $\Delta_g = 2\Delta_\omega$ and $|\nabla^2 f|_g^2 = 2|\nabla^{1,0}\nabla^{1,0} f|^2 + 2(\Delta_\omega f)^2$, this becomes

$$\Delta_\omega|\nabla^{1,0}\psi|_g^2 = 2\mathrm{Ric}\,(\nabla^{1,0}\psi, \nabla^{0,1}\psi) + 2|\nabla^{1,0}\nabla^{1,0}\psi|^2 + 2\lambda_1^2\psi^2 - 4\lambda_1|\nabla^{1,0}\psi|_\omega^2. \quad (42)$$

Integrating (42) and using that $\mathrm{Ric}\,\omega(s) > s\omega(s)$ when $s < \mu$ by (40),

$$\int \Big((2s - 4\lambda_1)|\nabla^{1,0}\psi|_\omega^2 + 2\lambda_1^2\psi^2\Big)\omega^n < 0.$$

Now,

$$\int 2\lambda_1^2\psi^2\omega^n = -\int 2\lambda_1\psi\Delta_\omega\psi\omega^n = \int 2\lambda_1|\nabla^{1,0}\psi|_\omega^2.$$

Thus,

$$\int (2s - 2\lambda_1)|\nabla^{1,0}\psi|_\omega^2\omega^n < 0,$$

so we see that $\lambda_1 > s$. □

Remark 5.7. Here we see why we cannot use the rectangle

$$(-\infty, \mu] \times [0, 1]$$

containing A: we run into trouble with openness. If we had chosen ω to have nonnegative Ricci curvature we could have also worked on the larger trapezoid

$$(-\infty, 0] \times [0, 1] \cup \{(s, t) \in [0, \mu] \times [0, 1] \,:\, \mu t \ge s\}.$$

Producing such an ω is possible by applying Corollary 5.3 with $\mu = 0$ (whose proof does not require these arguments).

5.3. An L^∞ bound in the sub-rectangle

The following two lemmas will be sufficient for our purposes.

Lemma 5.8. *Suppose that $\varphi(s,t)$ is a solution of* (38). *Then whenever $s<0$,*

$$\varphi(s,t) < C(1+1/|s|),$$

for some uniform constant C independent of s and t.

Proof. Let p be a point where the maximum of φ is achieved. Then, $\sqrt{-1}\partial\bar\partial\varphi(p) \le 0$. Thus,

$$\frac{\omega_\varphi^n}{\omega^n}(p) \le 1, \quad \text{i.e.,} \quad t f_\omega(p) + c_t - s\varphi(s,t)(p) \le 0,$$

so,

$$\max \varphi(s,t) = \varphi(s,t)(p) \le (-c_t - t\min f_\omega)/|s|.$$

Similarly, if q is a point where the minimum is achieved,

$$\min \varphi(s,t) = \varphi(s,t)(q) \ge (-c_t - t\max f_\omega)/|s|.$$

concluding the proof. □

Lemma 5.9. *Suppose that $\varphi(0,t)$ is a solution of* (38). *Then,*

$$\max_M |\varphi(0,t)| < C,$$

for some uniform constant C independent of t.

Proof. As remarked earlier, the solution in this case is a priori only unique up to a constant. However, we fixed the normalization by requiring that the solution be continuous in the parameters s,t. We will eventually show that there are solutions $\varphi(s,t)$ for all s less than 0 and all $t\in[0,1]$. Therefore, $\varphi(s,t)$ converges pointwise to the solution $\varphi(0,t)$ for each fixed t as s tends to 0, and so this solution is actually unique. In particular, since the latter change sign, so must the former. Thus, it is enough to estimate the oscillation of $\varphi(0,t)$ in order to estimate the L^∞ norm of $\varphi(0,t)$, i.e., it suffices to estimate the minimum of

$$\varphi(0,t) - \max\varphi(0,t) - 1.$$

This bound, due to Yau [105], then follows just as in [97, § 5]. □

5.4. An L^∞ bound in the interval

Lemma 5.10. *Let $t=1$. The K-energy is monotonically decreasing in s.*

Proof. When $t=1$, $c_t=0$. Then,

$$\begin{aligned}\frac{d}{ds}E(\varphi(s,1)) &= \frac{d}{ds}\mathrm{Ent}(e^{f_\omega}\omega^n, e^{f_\omega - s\varphi}\omega^n)\\ &\quad - \mu V^{-1}\int_M \dot\varphi{\omega_\varphi}^n + \mu V^{-1}\int_M \dot\varphi{\omega_\varphi}^n + \mu V^{-1}\int_M \varphi\Delta\dot\varphi{\omega_\varphi}^n\\ &= V^{-1}\int_M(-\varphi - s\dot\varphi - s\varphi\Delta\dot\varphi){\omega_\varphi}^n + \mu V^{-1}\int_M \varphi\Delta\dot\varphi{\omega_\varphi}^n.\end{aligned}$$

Differentiating (38) yields

$$\varphi(s,t) = -(\Delta + s)\dot{\varphi}(s,t). \tag{43}$$

Thus,

$$\begin{aligned}\frac{d}{ds}E(\varphi(s,1)) &= V^{-1}\int_M (\Delta\dot{\varphi} - (\mu - s)\Delta\dot{\varphi}(\Delta + s)\dot{\varphi})\omega_\varphi{}^n \\ &= -(\mu - s)V^{-1}\int_M \dot{\varphi}\Delta(\Delta + s)\dot{\varphi}\omega_\varphi{}^n < 0,\end{aligned}$$

since $\dot{\varphi}$ is not constant as can be seen from (43) and (38), while $\Delta + s$ is a negative operator for $s < \mu$ thanks to Lemma 5.6. □

By the previous lemma, the K-energy actually decreases along the interval. By properness this implies that the functional $I - J$ stays uniformly bounded along the interval once we know $\varphi(0,1)$ exists. This will indeed be the case as we will show (in the first step of the proof) existence in the sub-rectangle for all values $s \leq 0$. Now, the explicit formula (12) for the K-energy hence implies that the entropy is bounded from above along the interval,

$$\mathrm{Ent}(e^{f_\omega}\omega^n, \omega^n_{\varphi(s,1)}) < C, \tag{44}$$

thus,

$$\int \big((t-1)f_\omega - s\varphi(s,1) + c_t\big)\omega^n_\varphi < C, \tag{45}$$

or

$$\int -\varphi(s,1)\omega^n_{\varphi(s,1)} < C(1 + 1/s). \tag{46}$$

Observe that here we may assume that $s > \epsilon > 0$ since by openness about the value $(0,1)$ we have existence for small positive values of s. Going back to (7) now shows that

$$\int \varphi(s,1)\omega^n < C(1 + 1/s), \tag{47}$$

the mean value inequality (33) shows that

$$\max \varphi(s,1) < C(1 + 1/s). \tag{48}$$

It remains to estimate $\min \varphi(s,1)$. Now, as in the proof of Lemma 5.9, we set

$$\alpha(s) := \max \varphi(s,1) - \varphi(s,1) + 1.$$

A standard Moser iteration argument now yields an estimate [97, § 5]

$$||\alpha(s)||_{L^\infty(M,\omega^n_{\varphi(s,1)})} \leq C\big(||\alpha(s)||_{L^1(M,\omega^n_{\varphi(s,1)})}\big),$$

but

$$||\alpha(s)||_{L^1(M,\omega^n_{\varphi(s,1)})} = \max \varphi(s,1) + 1 + \int -\varphi(s,1)\omega^n_{\varphi(s,1)} < C(1 + 1/s).$$

Thus, we have proven the following.

Lemma 5.11. *Suppose that $\varphi(s,1)$ exists for $s \in (0,\epsilon]$ for some $\epsilon > 0$. Then for every $s > \epsilon$,*

$$\max_M |\varphi(s,1)| < C(1+1/s).$$

5.5. Second-order estimates

The reference for this subsection is [81, § 7.2–7.4,7.7].

We say that ω, ω_φ are uniformly equivalent if

$$C_1\omega \le \omega_\varphi \le C_2\omega, \tag{49}$$

for some constants $C_2 \ge C_1 > 0$. We start with a simple result which shows that a Laplacian estimate can be interpreted geometrically. Denote

$$\mathrm{tr}_\omega \eta := \mathrm{tr}\big([g_{i\bar{j}}]^{-1}[h_{k\bar{l}}]\big),$$

where $\omega = g_{i\bar{j}} dz^i \wedge \overline{dz^j}$, $\eta = h_{i\bar{j}} dz^i \wedge \overline{dz^j}$ in local coordinates. Similarly, denote

$$\det{}_\omega \eta := \det\big([g_{i\bar{j}}]^{-1}[h_{k\bar{l}}]\big).$$

Exercise 5.12. Show that (49) is implied by either

$$n + \Delta_\omega \varphi = \mathrm{tr}_\omega \omega_\varphi \le C_2, \quad \text{and} \quad \det{}_\omega \omega_\varphi \ge C_1 C_2^{n-1}/(n-1)^{n-1}, \tag{50}$$

or,

$$n - \Delta_{\omega_\varphi} \varphi = \mathrm{tr}_{\omega_\varphi} \omega \le 1/C_1, \quad \text{and} \quad \det{}_\omega \omega_\varphi \le C_1 C_2^{n-1}(n-1)^{n-1}. \tag{51}$$

Exercise 5.13. Conversely, show that (49) implies

$$\mathrm{tr}_\omega \omega_\varphi \le nC_2, \quad \text{and} \quad \det{}_\omega \omega_\varphi \ge C_1^n,$$

as well as

$$\mathrm{tr}_{\omega_\varphi} \omega \le n/C_1, \quad \text{and} \quad \det{}_\omega \omega_\varphi \le C_2^n.$$

(Indeed, $\sum(1+\lambda_j) \le A$, and $\Pi(1+\lambda_j) \ge B$ implies $1+\lambda_j \ge (n-1)^{n-1}B/A^{n-1}$; conversely, $\Pi(1+\lambda_j) \ge \big(\frac{1}{n}\sum\frac{1}{1+\lambda_j}\big)^{-n} \ge C_1^n$.)

The quantities $\mathrm{tr}_\omega \omega_\varphi$ and $\det_\omega \omega_\varphi$ have a nice geometric interpretation. To see that, we will study the geometry of the identity map $\iota : M \to M$! Consider $\partial\iota^{-1}$ either as a map from $T^{1,0}M$ to itself, or as a map from $\Lambda^n T^{1,0}M$ to itself. Alternatively, it is section of $T^{1,0\star}M \otimes T^{1,0}M$, or of $\Lambda^n T^{1,0\star}M \otimes \Lambda^n T^{1,0}M$, and we may endow these product bundles with the product metric induced by ω on the first factor, and by ω_φ on the second factor. Then, (50) means that the norm squared of $\partial\iota^{-1}$, in its two guises above, is bounded from above by C_2, respectively bounded from below by $C_1 C_2^{n-1}/(n-1)^{n-1}$. Similarly, the quantities $\mathrm{tr}_{\omega_\varphi}\omega$ and $\det_{\omega_\varphi}\omega$ and (51) can be interpreted in terms of $\partial\iota$.

Now, for us the quantities

$$\det{}_\omega \omega_{\varphi(s,t)} = e^{tf_\omega + c_t - s\varphi(s,t)} \quad \text{and} \quad \det{}_{\omega_{\varphi(s,t)}} \omega = e^{-tf_\omega - c_t + s\varphi(s,t)}$$

are already uniformly bounded thanks to the uniform estimate on $||\varphi(s,t)||_{L^\infty}$ obtained in § 5.3–5.4. Thus, according to Exercise 5.12, it remains to find an upper bound for either $|\partial\iota^{-1}|^2$ or $|\partial\iota|^2$ (from now on we just consider maps on $T^{1,0}M$).

The standard way to approach this is by using the maximum principle, and thus involves computing the Laplacian of either one of these two quantities. The classical approach, due to Aubin [1–3] and Yau [105], is to estimate the first, while a more recent approach is to estimate the second [65, 79], and this builds on using and finessing older work of Lu [67] and Bando–Kobayashi [8]. Both of these approaches are explained in a unified manner in [81, § 7]. The result we need is [81, Corollary 7.8 (i)].

Lemma 5.14. *Let $\varphi \in C^4(M) \cap \mathrm{PSH}(M, \omega)$. Suppose that*

$$\mathrm{Ric}\,\omega_\varphi \geq -C_1\omega - C_2\omega_\varphi, \tag{52}$$

and

$$\max_M \mathrm{Bisec}_\omega \leq C_3. \tag{53}$$

Then

$$-n < \Delta_\omega\varphi \leq (C_1 + n(C_2 + 2C_3 + 1))e^{(C_2+2C_3+1)\,\mathrm{osc}\,\varphi} - n. \tag{54}$$

To see that this result is applicable, observe first that (53) holds simply because M is compact and ω is smooth. Second, according to (40)

$$\begin{aligned}\mathrm{Ric}\,\omega_{\varphi(s,t)} &= (1-t)\mathrm{Ric}\,\omega + s\omega_{\varphi(s,t)} + (\mu t - s)\omega \\ &\geq s\omega_{\varphi(s,t)} + (\mu t + (1-t)C_4 - s)\omega,\end{aligned}$$

where C_4 is a lower bound for the Ricci curvature of ω, i.e., satisfying

$$\mathrm{Ric}\,\omega \geq C_4\omega.$$

Therefore, Lemma 5.14 holds with

$$C_1 = \max\{0, |\mu t + (1-t)C_4 - s|\}, \quad C_2 = \max\{0, |s|\}, \quad C_3 = C_3(\omega).$$

5.6. Higher-order compactness via Evans–Krylov's estimate

To show our solutions along the continuity method are smooth, it suffices to improve the Laplacian estimate to a $C^{2,\gamma}$ estimate for some $\gamma > 0$. Indeed, then it is standard to see that the solutions automatically have uniform $C^{k,\gamma}$ estimates for each k (Exercise 5.18). This is obtained via the standard Evans–Krylov estimate, adapted to the complex setting. The standard references for this are the lecture notes of Siu [87] and Błocki [20, § 5], as well as the treatment of the real Monge–Ampère equation by Gilbarg and Trudinger [58], with the modification by Wang–Jiang [102], Błocki [19].

Lemma 5.15. *Let $\psi \in C^4(M) \cap \mathrm{PSH}(M, \omega)$ be a solution to $\omega_\psi^n = \omega^n e^F$. Then*

$$||\psi||_{C^{2,\gamma}} \leq C, \tag{55}$$

where $\gamma > 0$ and C depend only on M, ω, $||\Delta_\omega\psi||_{C^0}, ||\psi||_{C^0}$, and $||F||_{C^2}$.

Proof. For concreteness, we carry out the proof for out particular F in (38). For each pair (s, t) define $h = h(s, t)$ by

$$\log h := tf_\omega - s\varphi + c_t + \log\det[\psi_{i\bar{j}}], \tag{56}$$

where ψ is a local Kähler potential for ω on some fixed neighborhood (we will obtain our estimate only on this neighborhood, but then cover M with finitely many such). Set $u := \psi + \varphi$. Each $C^{k,\gamma}$ norm of ψ is bounded by a constant $C = C(k, \gamma, \omega)$, so to get the desired bound on φ is tantamount to bounding u.

Let $\eta = (\eta_1, \ldots, \eta_n) \in \mathbb{C}^n$ be a unit vector, and consider u as a function of $(z_1, \ldots, z_n) \in \mathbb{C}^n$. Then,

$$(\log \det[u_{i\bar{j}}])_{\eta\bar{\eta}} = -u^{i\bar{l}} u^{k\bar{j}} u_{\eta i\bar{j}} u_{\bar{\eta} k\bar{l}} + u^{i\bar{j}} u_{\eta\bar{\eta} i\bar{j}}.$$

(repeated differentiation is justified since by assumption φ and hence u belong to $C^4(M)$). Since

$$\log \det[u_{i\bar{j}}] = \log \det[\psi_{i\bar{j}} + \varphi_{i\bar{j}}] = \log h,$$

and letting

$$w := u_{\eta\bar{\eta}}, \tag{57}$$

we thus have

$$u^{i\bar{j}} w_{i\bar{j}} \geq (\log h)_{\eta\bar{\eta}} = \frac{h_{\eta\bar{\eta}}}{h} - \frac{|h_\eta|^2}{h^2}, \tag{58}$$

which can be rewritten in divergence form,

$$(hu^{i\bar{j}} w_i)_{\bar{j}} \geq \eta^{\bar{l}} (\eta^k h_k)_{\bar{l}} - g, \qquad g := \frac{|h_\eta|^2}{h}. \tag{59}$$

Theorem 5.16 ([58, Theorem 8.18]). *Let $\Omega \subset \mathbb{R}^m$, and assume $B_{4\rho} = B_{4\rho}(y) \subset \Omega$. Let $L = D_i(a^{ij} D_j + b^i) + c^i D_i + d$ be strictly elliptic, $\lambda I < [a^{ij}]$, with $a^{ij}, b^i, c^i, d \in L^\infty(\Omega)$, satisfying*

$$\sum_{i,j} |a^{ij}|^2 < \Lambda^2, \quad \lambda^{-2} \sum (|b^i|^2 + |c^i|^2) + \lambda^{-1} |d| \leq \nu^2.$$

Then if $U \in W^{1,2}(\Omega)$ is nonnegative and satisfies $LU \leq g + D_i f^i$, with $f^i \in L^q, g \in L^{q/2}$ with $q > m$, then for any $p \in [1, \frac{m}{m-2})$,

$$\begin{aligned}\rho^{-m/p} ||U||_{L^p(B_{2\rho})} &< C(\inf_{B_\rho} U + \rho^{1-m/q} ||f||_{L^q(B_{2\rho})} + \rho^{2-2n/q} ||g||_{L^{q/2}(B_{2\rho})}) \\ &< C(\inf_{B_\rho} U + \rho ||f||_{L^\infty(B_{2\rho})} + \rho^2 ||g||_{L^\infty(B_{2\rho})}),\end{aligned}$$

with $C = C(n, \Lambda/\lambda, \nu\rho, q, p)$.

Lemma 5.17. *Let w be defined by* (57)*. Suppose that $s > S$. One has*

$$\sup_{B_{2\rho}} w - \frac{1}{|B_\rho|} \int_{B_\rho} w\omega^n \leq C\big(\sup_{B_{2\rho}} w - \sup_{B_\rho} w + \rho(\rho + 1)\big), \tag{60}$$

with $C = C(M, \omega, S, ||\varphi(s,t)||_{C^0(M)}, ||\Delta_\omega \varphi(s,t)||_{C^0})$.

Proof. By (59) $v := \sup_{B_{2\rho}} w - w$ satisfies

$$(hu^{i\bar{j}} v_i)_{\bar{j}} \leq g - \eta^{\bar{l}} (\eta^k h_k)_{\bar{l}}, \tag{61}$$

where $g := \frac{|h_\eta|^2}{h}$. In general, there are positive bounds on $[u_{i\bar{j}}]$ and $[u^{i\bar{j}}]$, depending only on $||\Delta_\omega \varphi(s,t)||_{C^0}$, and hence similar positive bounds on $[a^{i\bar{j}}] := h[u^{i\bar{j}}]$ and its inverse, in these coordinates, depending only on $S, M, \omega, ||\Delta_\omega\varphi(s,t)||_{C^0}$, and $||\varphi(s,t)||_{C^0}$ (since the latter two quantities control $||\varphi(s,1)||_{C^{0,1}}$ by interpolation). This, together with Theorem 5.16, gives the desired inequalities provided that $v \in W^{1,2}(M,\omega^n)$, which is automatic as v and ω^n are smooth and M is compact. The lemma follows. □

Now, let $\{V_j\}_{j=1}^n$ be smooth vector fields on M that span $T^{1,0}M$ over M and that on a local chart are given by $V_k := \frac{\partial}{\partial z_k}$, $k = 1,\ldots,n$, and denote

$$M(\rho) := \sup_{|\zeta|,|Z|\in(0,\rho)} \sum_{j=1}^n V_j\overline{V_j}u, \quad m(\rho) := \inf_{|\zeta|,|Z|\in(0,\rho)} \sum_{j=1}^n V_j\overline{V_j}u$$

Our goal is to show that $\nu(\rho) := M(\rho) - m(\rho)$ is Hölder continuous with respect to g_ω, i.e., $\nu(\rho) \le C\rho^{\gamma'}$, for some $\gamma' > 0$, or equivalently that $\nu(\rho) \le (1-\epsilon)\nu(2\rho) + \sigma(\rho)$, for some $\epsilon \in (0,1)$ and some non-decreasing function σ [58, Lemma 8.23]. Let

$$M_\eta(\rho) := \sup_{|\zeta|,|Z|\in(0,\rho)} u_{\eta\bar\eta}, \quad m_\eta(\rho) := \inf_{|\zeta|,|Z|\in(0,\rho)} u_{\eta\bar\eta}, \quad \nu_\eta(\rho) := M_\eta(\rho) - m_\eta(\rho).$$

Equation (60) implies

$$\sup_{B_{2\rho}} w - \frac{1}{|B_\rho|}\int_{B_\rho} w \le C\big(\nu_\eta(2\rho) - \nu_\eta(\rho) + \rho(\rho+1)\big), \tag{62}$$

and so it remains to obtain a similar inequality for $w - \inf_{B_{2\rho}} w$.

Note that $DF|_A.(A-B) \le F(A) - F(B)$, by concavity of $F(A) := \log\det A$ on the space of positive Hermitian matrices. Since $DF|_{\nabla^{1,1}u} = (\nabla^{1,1}u)^{-1}$, we have

$$u^{i\bar j}(y)(u_{i\bar j}(y) - u_{i\bar j}(x)) \le \log\det u_{i\bar j}(y) - \log\det u_{i\bar j}(x) \le |h|_{C^{0,1}}|y-x|. \tag{63}$$

We now decompose $(u^{i\bar j})$ as a sum of rank one matrices. This will result in the previous equation being the sum of pure second derivatives for which we can apply our estimate from the previous step. By uniform ellipticity this decomposition can be done uniformly in y [87, p. 103],[20]. Namely, we can fix a set $\{\gamma_k\}_{k=1}^N$ of unit vectors in $\mathbb{C}^n$ (which we can assume contains $\gamma_1 = \eta$ as well as a unitary frame of which η is an element) and write

$$(u^{i\bar j}(y)) = \sum_{k=1}^N \beta_k(y)\gamma_k^*\gamma_k,$$

with $\beta_k(y)$ uniformly positive depending only on n, λ and Λ. Thus (63) gives

$$\begin{aligned} w(y) - w(x) &\le C|y-x| - \sum_{k=2}^N \beta_k(y)(u_{\gamma_k\bar\gamma_k}(y) - u_{\gamma_k\bar\gamma_k}(x)) \\ &\le C|y-x| + \sum_{k=2}^N \beta_k(y)(\sup_{B_{2\rho}} u_{\gamma_k\bar\gamma_k} - u_{\gamma_k\bar\gamma_k}(y)). \end{aligned}$$

Now let $w(x) = \inf_{B_{2\rho}} w$, and average over B_ρ to get, using (62),

$$\frac{1}{|B_\rho|}\int_{B_\rho} w - \inf_{B_{2\rho}} w \le C\Big(\sum_{k=2}^{N} \nu_{\eta_k}(2\rho) - \nu_{\eta_k}(\rho) + \rho(\rho+1)\Big). \tag{64}$$

Combining this with (62), and summing over $k = 1, \ldots, N$ we thus obtain an estimate on $\nu(\rho)$ of the desired form. Hence $\Delta_\omega\varphi(s) \in C^{0,\gamma'}$ for some $\gamma' > 0$. In fact our proof actually showed that $\varphi_{\eta\bar\eta} \in C^{0,\gamma'}$ for any η. Hence, by polarization we deduce that also $\varphi_{i\bar j} \in C^{0,\gamma'}$, for any i, j. Hence, $|\Delta_\omega\varphi(s,t)|_{C^{0,\gamma'}} \le C = C(M, \omega, S, ||\Delta_\omega\varphi(s,t)||_{C^0(M)}, ||\varphi(s,t)||_{C^0(M)})$. This concludes the proof of Lemma 5.15. □

Exercise 5.18. Suppose that $\varphi \in C^\infty(M)$ satisfies $\omega_\varphi^n = e^F\omega^n$ and that

$$||\varphi|_{C^{2,\gamma}} \le C.$$

Show that there exists C' such that

$$||\varphi|_{C^{3,\gamma}} \le C' = C'(M, \omega, ||F||_{C^{1,\alpha}}).$$

(Hint: Let D be a first-order operator with constant coefficients in some holomorphic coordinate chart. Write the Monge–Ampère equation in those coordinates as

$$\log\det[u_{i\bar j}] = \log\det[\psi_{i\bar j}] + F =: \tilde F,$$

as in the proof of Lemma 5.15 and apply D to this equation. By Exercise 2.1 this then gives a Poisson type equation for Du,

$$u^{i\bar j}(Du)_{i\bar j} = D\tilde F.$$

This is not quite a Poisson equation since the Laplacian on the left-hand side depends on u itself! However, since we already have uniform $C^{0,\gamma}$ estimates on $[u^{i\bar j}]$ and $[u_{i\bar j}]$ the usual Schauder estimates [58] give

$$||Du||_{C^{2,\gamma}} \le C\big(||Du||_{C^{0,\gamma}} + ||\tilde F||_{C^{0,\gamma}}\big).$$

Since this holds for

$$D \in \left\{\frac{\partial}{\partial z^1}, \ldots, \frac{\partial}{\partial z^n}, \frac{\partial}{\overline{\partial z^1}}, \ldots, \frac{\partial}{\overline{\partial z^n}}\right\},$$

we are done.)

Applying the previous exercise repeatedly yields the following improvement of Lemma 5.15:

Corollary 5.19. *Let $\psi \in C^{k+1}(M) \cap \mathrm{PSH}(M, \omega)$ be a solution to $\omega_\psi^n = \omega^n e^F$. Then*

$$||\psi||_{C^{k,\gamma}} \le C, \tag{65}$$

where $\gamma > 0$ and C depend only on M, ω, $||\Delta_\omega\psi||_{C^0}, ||\psi||_{C^0}$, and $||F||_{C^{k-1}}$, and C depends additionally also on k.

5.7. Properness implies existence

We now complete the proof of one direction of Conjecture 1.2. Let K be a connected compact subgroup of the automorphism group. Recall that

$$\mathcal{H}^K \subset \mathcal{H}$$

consists of all K-invariant elements of $\mathcal{H}$. We denote by

$$C_K^{k,\gamma}$$

the subset of $C^{k,\gamma}$ consisting of K-invariant functions. Denote by

$$B \subset A$$

the subset of parameter values (s,t) for which there exists a K-invariant $C^{2,\gamma}$ solution $\varphi(s,t)$ to (41). Note that $(-\infty,0]\times\{0\}\subset B$ since $\varphi(s,0)=0$ and we can always assume that ω is K-invariant, for instance by taking an arbitrary Kähler metric and averaging it with respect to the Haar measure of G [62, p. 88].

Next, observe that the openness arguments of § 5.2 run through unchanged for K-invariant solutions. This is because $M_{s,t}: C^{2,\gamma}\cap \mathrm{PSH}(M,\omega)\to C^{0,\gamma}$ defined by

$$M_{s,t}(\varphi) := \log\frac{\omega_\varphi^n}{\omega^n} - tf_\omega + s\varphi - c_t, \quad (s,t)\in A,$$

actually maps $C_K^{2,\gamma}\cap \mathrm{PSH}(M,\omega)$ to $C_K^{0,\gamma}$, and therefore

$$DM_{s,t}|_{\varphi(s,t)} = \Delta_{\varphi(s,t)} + s, \quad (s,t)\in A,$$

maps $C_K^{2,\gamma}$ to $C_K^{0,\gamma}$. In conclusion then, B is a nonempty open subset of A. Moreover, if

$$A_S := (-S,-1/S]\times[0,1], \tag{66}$$

we have that $B\cap A_S$ is a nonempty open subset of A_S for any value $S>1$.

First, we show that $A_S\subset B$. Indeed, let $(s,t)\in\partial(B\cap A_S)$, and let $\{(s_j,t_j)\}_j\subset B\cap A_S$ be a subsequence converging to (s,t). According to Lemma 5.8,

$$\sup_j \max_M |\varphi(s_j,t_j)| < C(1+S).$$

Then, according to Lemma 5.14,

$$\sup_j \max_M |\Delta_\omega\varphi(s_j,t_j)| < C = C(M,\omega,S).$$

Thus, according to Lemma 5.15,

$$\sup_j \max_M ||\varphi(s_j,t_j)||_{C^{2,\gamma}} < C = C(M,\omega,S).$$

Therefore, for every $\alpha\in(0,\gamma)$, the functions $\varphi(s_j,t_j)$ converge to $\varphi(s,t)$ in the $C^{2,\alpha}$ topology, and moreover $\varphi(s,t)\in C^{2,\gamma}$. Thus, $(s,t)\in B$. This completes the proof that $A_S\subset B$, So we have shown that

$$\cup_{S>1}A_S = (-\infty,0)\times[0,1]\subset B.$$

Observe that this actually concludes the proof of Corollary 5.3 whenever $\mu<0$ thanks to the elliptic regularity results mentioned below.

Second, we show that actually $A_\infty \subset B$. Indeed, $(0,0) \in B$, and by openness also $\{(0,t) : 0 < t < \epsilon\} \subset B$, for some $\epsilon > 0$. Applying now Lemma 5.9 instead of Lemma 5.8, we get just as in the previous paragraph

$$\sup_t \max_M ||\varphi(0,t)||_{C^{2,\gamma}} < C = C(M,\omega).$$

Thus, as before it follows that $A_\infty \subset B$. Observe that the solutions we constructed are continuous in the parameters $s,t \in A_\infty$, in particular even up to $s = 0$, since we use the openness argument that relies on the implicit function theorem that necessarily produces solutions that depend continuously on the parameters. This actually concludes the proof of Corollary 5.3 (again, thanks to the elliptic regularity results).

Third, we treat the remaining piece in A. First, by openness $\{(s,1) : 0 \leq s < \epsilon\} \subset B$, for some $\epsilon = \epsilon(M,\omega) > 0$. Therefore, by Lemma 5.11 together with the higher-order estimates (as in the preceding paragraphs)

$$\sup_s \max_M ||\varphi(s,1)||_{C^{2,\gamma}} < C = C(M,\omega).$$

Once again, this is enough to conclude that $[0,\mu] \times \{1\} \subset B$. Thus,

$$B = A,$$

as desired.

Finally, by Corollary 5.19 (standard elliptic regularity results), the $C^{2,\gamma}$ solutions we constructed are actually smooth. Thus, $\varphi(\mu,1) \in \mathcal{H}^K$, and $\omega_{\varphi(\mu,1)}$ is K-invariant Kähler–Einstein metric. This concludes the proof of Theorem 5.1.

6. A counterexample to Tian's first conjecture and a revised conjecture

Theorem 5.1 shows that properness implies existence. This is one direction of Tian's first conjecture (Conjecture 1.2). The special case when there are *no* automorphisms (by which we mean ones homotopic to the identity, i.e., $\mathrm{Aut}(M,\mathrm{J})_0 = \{\mathrm{id}\}$) of the other, harder, direction of Conjecture 1.2 was established by Tian [96] under a technical assumption that was removed by Tian–Zhu [100]. This gave considerable plausibility to the conjecture. We now explain another reason why the general case of the conjecture seems plausible.

6.1. Why Tian's conjecture is plausible

First we explain why it is natural (in fact, necessary!) for this harder converse direction to only try to establish properness on $\mathcal{H}^K$ and not on all of $\mathcal{H}$. For this, observe first that E is invariant under the action of $\mathrm{Aut}(M,\mathrm{J})_0$ whenever a Kähler–Einstein metric exists:

Claim 6.1. *Suppose (M,J,ω) is Fano Kähler–Einstein with $\mu[\omega] = 2\pi c_1(M)$ and $\mu > 0$. Then $E(g^\star\omega_\varphi) = E(\omega_\varphi)$ for all $g \in \mathrm{Aut}(M,\mathrm{J})_0$ and $\varphi \in \mathcal{H}$.*

Proof. By (36),

$$\left.\frac{d}{dt}\right|_{t=0} E((\exp_I tX)^\star\omega_\varphi) = -V^{-1}\int_M \psi^X_{\omega_\varphi}\Delta_{\omega_\varphi} f_{\omega_\varphi}\omega^n_\varphi \\ = -V^{-1}\int_M \psi^X_{\omega_\varphi}(s_{\omega_\varphi}-n\mu) f_{\omega_\varphi}\omega^n_\varphi. \tag{67}$$

By a theorem of Futaki the functional

$$\eta \mapsto \int_M \psi^X_\eta (s_\eta - n\mu) f_\eta \eta^n$$

is constant on $\mathcal{H}$ [26, 29, 56]. Since it is zero at ω ($\mathrm{Ric}\,\omega = \mu\omega$ implies $s_\omega = n\mu$), it is identically zero. Thus,

$$\left.\frac{d}{dt}\right|_{t=0} E((\exp_I tX)^\star\omega_\varphi) = 0. \tag{68}$$

Now, actually

$$\left.\frac{d}{dt}\right|_{t=s} E((\exp_I tX)^\star\omega_\varphi) = 0 \tag{69}$$

for every s. Indeed,

$$\left.\frac{d}{dt}\right|_{t=s} E((\exp_I tX)^\star\omega_\varphi) = \left.\frac{d}{dt}\right|_{t=0} E((\exp_I (s+t)X)^\star\omega_\varphi) \\ = \left.\frac{d}{dt}\right|_{t=0} E((\exp_I tX)^\star((\exp_I sX)^\star\omega_\varphi)) \\ = 0, \tag{70}$$

by replacing ω_φ by $(\exp_I sX)^\star\omega_\varphi$ in (68). Since $\mathrm{Aut}(M,\mathrm{J})_0$ is a covered by its one-parameter subgroups, the statement follows. □

On the other hand, the Aubin functional is not invariant under the action of automorphisms. In fact, it might blow up along a one-parameter subgroup. The following lemma is due to Bando–Mabuchi [9, Lemma 6.2].

Lemma 6.2. *Let $\omega \in \mathcal{H}$ be arbitrary and suppose $\eta \in \mathcal{H}$ is Kähler–Einstein with $\mu > 0$. The function $F_\eta : \mathrm{Aut}(M,\mathrm{J})_0 \to \mathbb{R}_+$,*

$$F_\eta : g \mapsto (I-J)(g^\star\eta)$$

is proper (when we identify $\mathrm{Aut}(M,\mathrm{J})_0$ with its η-orbit in $\mathcal{H}$ and endow this subset of $\mathcal{H}$ with the $C^{2,\gamma}(M,\omega)$-topology).

Proof. Indeed, suppose that $I-J$ is bounded on a sequence $\omega_j = \omega_{\varphi_j}$ of Kähler–Einstein metrics in $\mathrm{Aut}(M,\mathrm{J})_0.\eta \subset \mathcal{H}$. By Claim 6.1 and (15),

$$C = E(\varphi_1) = E(\varphi_j) = \mathrm{Ent}(e^{f_\omega}\omega^n, \omega^n_{\varphi_j}) - \mu(I-J)(\varphi_j).$$

Now, normalize φ_j so that (recall (35))

$$\omega^n_{\varphi_j} = \omega^n e^{f_\omega - \mu\varphi_j}$$

(this fixes φ_j since $\mu > 0$ and the right-hand side must integrate to V). Plugging back into the formula for $E(\varphi_j)$ gives

$$C = E(\varphi_j) = -\mu \int \varphi_j \omega_{\varphi_j}^n - \mu(I - J)(\varphi_j),$$

or,

$$-\mu \int \varphi_j \omega_{\varphi_j}^n = C + \mu(I - J)(\varphi_j) \leq C',$$

by assumption that $(I - J)(\varphi_j)$ is uniformly bounded. But now

$$\begin{aligned} V^{-1} \int \varphi_j \omega^n &= -V^{-1} \int -\varphi_j \omega_{\varphi_j}^n + I(\varphi_j) \\ &\leq -V^{-1} \int -\varphi_j \omega_{\varphi_j}^n + \frac{n+1}{n}(I - J)(\varphi_j) \\ &\leq -(I - J)(\varphi_j) + \frac{n+1}{n}(I - J)(\varphi_j) \\ &= \frac{1}{n+1}(I - J)(\varphi_j) \leq \frac{C}{n+1}. \end{aligned} \tag{71}$$

Thus, by (33),

$$\max \varphi_j \leq V^{-1} \int \varphi_j \omega^n + C' < C''.$$

Now a Moser iteration argument just as in § 5.4 applies (the Sobolev and Poincaré constants of the Kähler–Einstein metrics of Ricci curvature equal to $\mu > 0$ are all uniform) to give

$$-\min \varphi_j \leq \frac{C}{V} \int -\varphi_j \omega_{\varphi_j}^n + C.$$

Combining the last two equations,

$$\operatorname{osc} \varphi_j = \max \varphi_j - \min \varphi_j \leq C + \frac{C}{V} \int -\varphi_j \omega_{\varphi_j}^n + C'' \leq C''',$$

using the display prior to (71). Since φ_j must change signs (from the normalization for φ_j inherent in $\omega_{\varphi_j}^n = \omega^n e^{f_\omega - \mu\varphi_j}$ and the one for f_ω in Definition 3.4), we have showed that

$$||\varphi_j||_{L^\infty} < C,$$

and consequently

$$||\varphi_j||_{C^{k,\gamma}} < C(k, \gamma),$$

for all k, α, which when $k = 2$ gives

$$C^{-1}\omega \leq \omega_j \leq C\omega.$$

Thus, endowing $\mathrm{Aut}(M, \mathrm{J})_0$ with, say, the $C^{2,\gamma}$-topology we see that the preimage under of F_η of compact sets in $\mathbb{R}_+$ are compact in the $C^{2,\gamma}$-topology, i.e., by definition F_η (the original F_η considered as a map on the group $\mathrm{Aut}(M, \mathrm{J})_0$) is proper. □

Corollary 6.3. *Suppose* (M, J, η) *is Fano Kähler–Einstein with* $\mu[\omega] = 2\pi c_1(M)$ *and* $\mu > 0$ *and that* $\mathrm{Aut}(M, \mathrm{J})_0$ *is nontrivial. Then*

$$(I - J) : \{g^\star\eta \, : \, g \in \mathrm{Aut}(M, \mathrm{J})_0\} \to \mathbb{R}_+$$

is unbounded from above.

Proof. Indeed, by Corollary 14.7 below F_η (defined in Lemma 6.2) descends to a function on $\mathrm{isom}(M, g)$, still denoted by F_η,

$$F_\eta(X) = (I - J)\big((\exp_I \mathrm{J}X)^\star\eta\big).$$

Since this function is still proper and $\mathrm{isom}(M, g)$ is a non-compact vector space, F_η must be unbounded. □

Remark 6.4. There is actually no particular need to look at the orbit of a Kähler–Einstein metric to show unboundedness; the same is true for the orbit of *any* metric as long as a Kähler–Einstein exists. Indeed, if $\alpha, \omega, \eta \in \mathcal{H}$, with η Kähler–Einstein,

$$E(g^\star\alpha) = E(\omega, g^\star\alpha) = E(\omega, g^\star\eta) + E(g^\star\eta, g^\star\alpha) = E(\omega, g^\star\eta) + E(\eta, \alpha).$$

Thus, $E(g^\star\alpha)$ is unbounded if and only if $E(g^\star\eta)$ is (as $E(\eta, \alpha)$ is some fixed constant).

6.2. A counterexample

However, surprisingly, Tian's first conjecture (which was stated as a theorem in [96, Theorem 4.4]) was recently disproved by Darvas and the author by establishing the following optimal version of Tian's conjecture.

Theorem 6.5. *Suppose* (M, J, ω) *is Fano with* $\mu[\omega] = 2\pi c_1(M)$ *and* $\mu > 0$*, and that* K *is a maximal compact subgroup of* $\mathrm{Aut}(M, \mathrm{J})_0$ *with* $\omega \in \mathcal{H}^K$*. The following are equivalent:*

(i) *There exists a Kähler–Einstein metric in* $\mathcal{H}^K$ *and* $\mathrm{Aut}(M, \mathrm{J})_0$ *has finite center.*
(ii) *There exists* $C, D > 0$ *such that* $E(\eta) \geq CJ(\eta) - D, \ \eta \in \mathcal{H}^K$.

Thus, restricting to the K-invariant potentials is necessary, but not sufficient, to guarantee properness.

Remark 6.6. The estimate in (ii) gives a concrete version of the properness condition (2). The direction (i) ⇒ (ii) is due to Phong et al. [74, Theorem 2], building on earlier work of Tian [96] and Tian–Zhu [100] in the case $\mathrm{Aut}(M, \mathrm{J})_0 = \{\mathrm{id}\}$, who obtained a weaker inequality in (ii) with J replaced by J^δ for some $\delta \in (0, 1)$ (for more details see also the survey [98, p. 131]).

Example 6.7. [45, Example 2.2] Let M denote the blow-up of $\mathbb{P}^2$ at three non colinear points. It is well known that it admits Kähler–Einstein metrics (see, e.g., [103]). In fact, one way to see this is by showing that Tian's invariant is equal to 1 for an appropriately chosen group of symmetries [10] and then apply Corollary 5.2 (with $\mu = 1$). According to [51, Theorem 8.4.2],

$$\mathrm{Aut}(M, \mathrm{J})_0 = (\mathbb{C}^\star)^2. \tag{72}$$

We will explain this fact in a moment. Given this, we see that $\mathrm{Aut}(M, \mathrm{J})_0$ is equal to its center which is clearly not finite. Thus, Conjecture 1.2 fails for M by Theorem 6.5. Following the appearance of [45], X.-H. Zhu informed the author that using toric methods one can give an alternative proof that Conjecture 1.2 fails in the special case of toric Fano n-manifolds that satisfy $\mathrm{Aut}(M, \mathrm{J})_0 = (\mathbb{C}^\star)^n$.

To see (72), observe that automorphisms homotopic to the identity map preserve the cohomology class of divisors. Thus, they preserve each of the three exceptional divisors. In particular, they descend to automorphisms of $\mathbb{P}^2$ which preserve the three blowup points. By that we mean that if $f \in \mathrm{Aut}(M, \mathrm{J})_0$ then $\pi \circ f \circ \pi^{-1} \in \mathrm{Aut}(\mathbb{P}^2)$. Now automorphisms of $\mathbb{P}^2$ are represented by invertible three-by-three matrices, up to a nonzero complex number. We may assume in this representation that the three points are then $[1 : 0 : 0], [0 : 1 : 0], [0 : 0 : 1]$ (since they are not collinear!). Thus, each such automorphism is represented by a diagonal matrix. Since the matrix is invertible, and determined up to a nonzero complex number, that matrix can be taken to be

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad a, b \in \mathbb{C}^\star.$$

Conversely, the blow-up of $\mathbb{P}^2$ at three non colinear points is a toric manifold so its automorphism group contains a copy of $(\mathbb{C}^\star)^2$. Thus, (72) is established.

These results motivate a reformulation of Tian's original conjecture. To present this reformulation we first make an excursion to infinite-dimensional metric geometry in the next sections. In Section 12 we return to state the reformulated conjecture, whose proof is described in Section 15.

7. Infinite-dimensional metrics on $\mathcal{H}$

Approaching problems in Kähler geometry through an infinite-dimensional perspective goes back to Calabi in 1953 [28] and later Mabuchi in 1986 [69]. These works proposed two different weak Riemannian metrics of L^2 type which have been studied extensively since.

The most widely studied such metric is the Mabuchi metric [69],

$$g_{\mathrm{M}}(\nu, \eta)|_\varphi := \int_M \nu\eta\, \omega_\varphi^n, \quad \nu, \eta \in T_\varphi \mathcal{H}_\omega \cong C^\infty(M), \tag{73}$$

discovered independently also by Semmes [86] and Donaldson [52] (see, e.g., [80, Chapter 2] for an exposition and further references).

Calabi's metric is given by

$$g_{\mathrm{C}}(\nu, \eta)|_\varphi := \int_M \Delta_\varphi \nu \Delta_\varphi \eta\, \frac{\omega_\varphi^n}{n!}. \tag{74}$$

This metric was introduced by Calabi in the 1950s in talks and in a research announcement [28]. It might seem a little less natural at first since it involves more derivatives than the Mabuchi metric. However, from a Riemannian geometric point of view it is actually more natural, since it is simply the L^2 metric on the level of Riemannian metrics, as the following simple result shows. To state this result we let

$$\mathcal{M}$$

denote the infinite-dimensional space of all smooth Riemannian metrics on M. The Ebin metric, also called the L^2 metric [53] (cf. [48]) is defined by

$$g_{\mathrm{E}}(h,k)|_g := \int_M \operatorname{tr}(g^{-1}hg^{-1}k)dV_g, \tag{75}$$

where $g \in \mathcal{M}$, $h,k \in T_g\mathcal{M}$ and $T_g\mathcal{M} \cong \Gamma(\mathrm{Sym}^2T^\star M)$, the space of smooth, symmetric $(0,2)$-tensor fields on M.

Proposition 7.1 ([38, Proposition 2.1]). *Consider the inclusion $\iota_{\mathcal{H}} : \mathcal{H} \hookrightarrow \mathcal{M}$. Then, $\iota_{\mathcal{H}}^\star g_{\mathrm{E}} = 2g_{\mathrm{C}}$.*

In other words, $(\mathcal{H}, 2g_{\mathrm{C}})$ is isometrically embedded in $(\mathcal{M}, g_{\mathrm{E}})$, or what is the same, the metric g_{C} is induced by the metric g_{E}.

On the other hand, the Mabuchi metric is more natural from a symplectic or complex geometry point of view. As shown by Semmes and Donaldson, the Mabuchi metric can be considered as an infinite-dimensional analogue of the symmetric space metric structure on spaces of the form $G^{\mathbb{C}}/G$ where G is a compact Lie group, but where the group is now infinite-dimensional, more specifically the group of Hamiltonian diffeomorphisms of (M,ω). We refer the reader to [52, 86], [97, Chapter 4], [92]. In another vein, the Mabuchi metric is also natural from the point of view of semi-classical complex geometry, also referred to as Kähler quantization sometimes. We refer the reader to [54, 75, 80, 82].

8. Metric completions of $\mathcal{H}$

Historically, Calabi claimed that the completion of his metric "consists of the positive semidefinite Kähler metrics defining the same principal class," i.e., of

$$\{\omega_\varphi := \omega + \sqrt{-1}\partial\bar\partial\varphi \,:\, \varphi \in C^\infty(M),\, \omega_\varphi \geq 0\}.$$

Except from this single line published in his short talk abstract in 1953 [28], there has been no study or even conjectures in the literature concerning metric completions of $\mathcal{H}$. The first article in this direction is due to Clarke–Rubinstein in 2011 [38], that we now turn to discuss.

8.1. The Calabi metric completion

Denote by $d_{\mathrm{C}} : \mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$ the distance function associated to metric g_{C}. It is defined as follows. A curve $[0,1] \ni t \mapsto \alpha_t \in \mathcal{H}$ is called smooth if $\alpha(t,z)$ is smooth in both t and z. Denote $\dot\alpha_t := \partial\alpha(t)/\partial t$. The length of a smooth curve $t \to \alpha_t$ is

$$\ell_{\mathrm{C}}(\alpha) := \int_0^1 \sqrt{g_{\mathrm{C}}(\dot\alpha_t, \dot\alpha_t)|_{\alpha_t}}\,dt. \tag{76}$$

Definition 8.1. The path length distance of $(\mathcal{H}, g_{\mathrm{C}})$ is defined by

$$d_{\mathrm{C}}(\omega,\eta) := \inf \big\{ \ell_{\mathrm{C}}(\alpha) \, : \, \alpha : [0,1] \to \mathcal{H} \text{ is a smooth curve with } \alpha(0)=\omega,\, \alpha(1)=\eta \big\}.$$

We refer to the pseudometric d_{C} as the *Calabi metric.*

Remark 8.2. As observed already by Calabi, the Calabi–Yau Theorem implies that $(\mathcal{H}, g_{\mathrm{C}})$ is isometric to a portion of a sphere in $L^2(M,\omega^n)$, and therefore the Calabi (pseudo)-metric is actually a metric, justifying the above name (see, e.g., [38, pp. 1488–1489] or [30]). Even though we refer to d_{C} and to g_{C} by the same name, we hope it will be clear below to which one we are referring to from the context.

The Calabi metric completion is given by the following theorem due to Clarke–Rubinstein [38, Theorem 5.6].

Theorem 8.3. *The metric completion of $(\mathcal{H}, d_{\mathrm{C}})$ is given by*

$$\overline{(\mathcal{H}, d_{\mathrm{C}})} \cong \{\varphi \in \mathcal{E}(M,\omega) \, : \, \omega_\varphi^n \text{ is absolutely continuous with respect to } \omega^n \text{ and } \omega_\varphi^n/\omega^n \in L^1(M,\omega^n)\},$$

and is a strict subset of

$$\mathcal{E}(M,\omega) := \left\{ \varphi \in \mathrm{PSH}(M,\omega) \, : \, \lim_{j\to\infty} \int_{\{\varphi \le -j\}} (\omega + \sqrt{-1}\partial\bar\partial \max\{\varphi, -j\})^n = 0 \right\}.$$

Furthermore, convergence with respect to d_{C} is characterized as follows. A sequence $\{\omega_{\varphi_k}\} \subset \mathcal{H}$ converges to $\omega_\varphi \in \mathcal{H}$ with respect to d_{C} if and only if $\omega_{\varphi_k}^n \to \omega_\varphi^n$ in the L^1 sense, i.e.,

$$\int_M \left| \frac{\omega_{\varphi_k}^n}{\omega^n} - \frac{\omega_\varphi^n}{\omega^n} \right| \omega^n \to 0.$$

Remark 8.4. Observe that the metric completion turns out to be considerably larger than what Calabi claimed. We also note that Theorem 8.3 was motivated by the computation of the metric completion of the ambient space $(\mathcal{M}, g_{\mathrm{E}})$ obtained in Clarke's thesis [37]. It is interesting to note that his result does not directly imply Theorem 8.3 as one might suspect from Proposition 7.1.

Remark 8.5. The space $\mathcal{E}(M,\omega)$ was introduced by Guedj–Zeriahi [60, Definition 1.1]. The statement of Theorem 8.3 of course assumes that the measure ω_φ^n can be defined for each $\varphi \in \mathcal{E}(M,\omega)$. This is indeed the case, but requires considerable background from pluripotential theory. One defines

$$\omega_\varphi^n := \lim_{j\to-\infty} \mathbf{1}_{\{\varphi > j\}} (\omega + \sqrt{-1}\partial\bar\partial \max\{\varphi, j\})^n.$$

By definition, $\mathbf{1}_{\{\varphi > j\}}(x)$ is equal to 1 if $\varphi(x) > j$ and zero otherwise, and the measure $(\omega + \sqrt{-1}\partial\bar\partial \max\{\varphi, j\})^n$ is defined by the work of Bedford–Taylor [11] since $\max\{\varphi, j\}$ is bounded. The limit is then well defined as a Borel measure; for more details we refer to [60, p. 445].

What is perhaps more interesting than computing the metric completion itself, is the fact that this computation yields nontrivial geometric information [38, Theorem 6.3].

Definition 8.6. We say that (M, J) is Calabi–Ricci unstable (or CR-unstable) if there exists a Ricci flow trajectory that diverges in $\overline{(\mathcal{H}, d_{\mathrm{C}})}$. Otherwise, we say (M, J) is CR-stable.

Theorem 8.7. *A Fano manifold (M, J) is CR-stable if and only if it admits a Kähler–Einstein metric. Moreover, if it is CR-unstable then any Ricci flow trajectory diverges in $\overline{(\mathcal{H}, d_{\mathrm{C}})}$.*

Theorem 8.7 might seem rather abstract, however it shows that convergence in the metric completion is fundamental geometrically. In addition, it can be stated entirely in terms of an a priori estimates without any reference to the metric completion [38, Corollary 6.9]:

Corollary 8.8. *The Ricci flow* (78) *converges smoothly if and only if*

$$||s - n||_{L^1(\mathbb{R}_+, L^2(M, \omega(t)))} < \infty, \tag{77}$$

where $s = s(t)$ denotes the scaler curvature of $(M, \omega(t))$.

This improves a result of Phong et al. [74], where (77) is replaced by

$$||s - n||_{L^1(\mathbb{R}_+, C^0(M))} < \infty,$$

which was proved by completely different methods. The novelty in Corollary 8.8 is that it uses supposedly "soft" infinite-dimensional geometry to prove actual "hard" a priori estimates for a PDE. Of course, the catch is that some analysis does go into computing the metric completion and, aside from that, some PDE techniques are still needed in the proof of Corollary 8.8. But, nevertheless, the idea that some PDE estimates can be explained using infinite-dimensional geometry seems attractive.

Exercise 8.9. Show that the length of the curve $t \mapsto \omega_{\varphi(t)}$ with respect to the Calabi metric is equal to

$$||s - n||_{L^1(\mathbb{R}_+, L^2(M, \omega(t)))}$$

if $\omega_{\varphi(t)}$ satisfies the Ricci flow equation

$$\frac{\partial \omega(t)}{\partial t} = -\mathrm{Ric}\, \omega(t) + \mu \omega(t), \quad \omega(0) = \omega \in \mathcal{H}. \tag{78}$$

Also, show that any solution of (78) that starts in $\mathcal{H}$ remains in $\mathcal{H}$ [64]. Thus, it makes sense to write $\omega(t) = \omega_{\varphi(t)}$.

Thus, Corollary 8.8 shows that convergence of the flow is equivalent to having finite distance in the Calabi metric.

Exercise 8.10. Rewrite (78) in the form of a complex Monge–Ampère equation

$$\omega_\varphi^n = \omega^n e^{f_\omega - \mu\varphi + \dot{\varphi}}, \quad \varphi(0) = \text{const}. \tag{79}$$

We remark that, depending on the context, the choice of the constant $\varphi(0)$ might involve some care (see [36, § 10.1],[72, § 2]).

Exercise 8.11. Assuming the theory of short-time existence for (78) (which replaces the openness arguments for the continuity method) show that for every $\omega \in \mathcal{H}$ the equation (79) admits a solution for all $t > 0$ whenever $\mu < 0$. To do this, use Exercise 8.10 as well as the results of § 5. Moreover, show that as t tends to infinity, the solutions $\omega(t)$ converge to the Kähler–Einstein metric.

Recently, Darvas generalized Calabi's metric to a two-parameter family of Finsler metrics, given by

$$||\eta||_{\varphi}^{\mathrm{C},p,q} := \left(\int_M |\Delta_{\omega_\varphi} \eta|^p \left(\frac{\omega_\varphi^n}{\omega^n}\right)^q \frac{\omega^n}{n!} \right)^{1/q}, \tag{80}$$

and computed the corresponding metric completions, directly generalizing Theorem 8.3. Denote by $d_{\mathrm{C,p,q}} : \mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$ the path-length distance function associated to (80).

Theorem 8.12 ([41, Theorem 1.1]). *Let $p, q \in (1, \infty)$ and $q \leq p$. The metric completion of $(\mathcal{H}, d_{\mathrm{C,p,q}})$ is given by*

$$\overline{(\mathcal{H}, d_{\mathrm{C,p,q}})} \cong \{\varphi \in \mathcal{E}(M, \omega) : \omega_\varphi^n \text{ is absolutely continuous with respect to } \omega^n \text{ and } \omega_\varphi^n/\omega^n \in L^q(M, \omega^n)\}.$$

Furthermore, convergence with respect to $d_{\mathrm{C,p,q}}$ is characterized as follows. A sequence $\{\omega_{\varphi_k}\} \subset \mathcal{H}$ converges to $\omega_\varphi \in \mathcal{H}$ with respect to $d_{\mathrm{C,p,q}}$ if and only if $\omega_{\varphi_k}^n \to \omega_\varphi^n$ in the L^q sense, i.e.,

$$\int_M \left| \frac{\omega_{\varphi_k}^n}{\omega^n} - \frac{\omega_\varphi^n}{\omega^n} \right|^q \omega^n \to 0.$$

In particular, the metric completion is independent of p! This immediately yields, by the same results of [38] that lead to Corollary 8.8, the following improvement to Corollary 8.8 [41, Theorem 1.1].

Corollary 8.13. *The Ricci flow* (78) *converges smoothly if and only if*

$$||s - n||_{L^1(\mathbb{R}_+, L^1(M, \omega(t)))} < \infty. \tag{81}$$

Exercise 8.14. Show that the length of the curve $t \mapsto \omega_{\varphi(t)}$ with respect to $d_{\mathrm{C},1,1}$ is equal to

$$||s - n||_{L^1(\mathbb{R}_+, L^1(M, \omega(t)))}$$

if $\omega_{\varphi(t)}$ satisfies the Ricci flow equation (78).

It would be interesting to obtain a proof of Corollary 8.13 using direct flow methods. At the same time, it is remarkable that such metric completion techniques can lead to new estimates on geometric flows. We believe that this circle of ideas should find more applications in other geometric and analytic settings.

8.2. The Mabuchi metric completion

As remarked earlier, the Calabi metric is more closely tied with the Riemannian geometry of M, and indeed convergence in the Calabi metric is related to convergence of the associated Riemannian volume forms. The Mabuchi metric, on the other hand, is more closely tied with the complex geometry of M, and so completely different methods would be needed to compute the Mabuchi metric completion. Using sophisticated techniques from pluripotential theory this was carried through by Darvas. A special case was also obtained around the same time by Guedj [59]. Define,

$$\mathcal{H}_\omega = \{\varphi \,:\, \varphi \in C^\infty(M),\, \omega_\varphi > 0\}, \tag{82}$$

and

$$\mathcal{E}_2 := \left\{\varphi \in \mathcal{E}(M,\omega) \,:\, \int \varphi^2 \omega_\varphi^n < \infty\right\}.$$

A curve $[0,1] \ni t \mapsto \varphi(t) \in \mathcal{H}_\omega$ is called smooth if $\varphi(t,z) = \varphi(t)(z) \in C^\infty([0,1]\times M)$. Denote $\dot\varphi(t) := \partial\varphi(t)/\partial t$. The length of a smooth curve $t \mapsto \varphi(t)$ is

$$\ell_{\mathrm{M}}(\alpha) := \int_0^1 \sqrt{g_{\mathrm{M}}(\dot\varphi(t),\dot\varphi(t))|_{\varphi(t)}}dt. \tag{83}$$

Definition 8.15. The path length distance of $(\mathcal{H}_\omega, d_{\mathrm{M}})$ is defined by

$$d_{\mathrm{M}}(\varphi_0,\varphi_1) := \inf\big\{\ell_{\mathrm{M}}(\varphi) \,:\, \varphi : [0,1] \to \mathcal{H}_\omega \text{ is a smooth curve with } \varphi(0) = \varphi_0,\, \varphi(1) = \varphi_1\big\}.$$

We call the pseudometric d_{M} the *Mabuchi metric.*

The metric completion of the Mabuchi metric is given by the following theorem of Darvas [39, Theorem 1] which also justifies the name given to d_{M} above.

Theorem 8.16. $(\mathcal{H}_\omega, d_{\mathrm{M}})$ *is a metric space. Moreover, the metric completion of* $(\mathcal{H}_\omega, d_{\mathrm{M}})$ *equals* $(\mathcal{E}_2, d_{\mathrm{M},2})$*, where*

$$d_{\mathrm{M},2}(\varphi_0,\varphi_1) := \lim_{k\to\infty} d_{\mathrm{M}}(\varphi_0(k),\varphi_1(k)), \tag{84}$$

for any smooth decreasing sequences $\{\varphi_i(k)\}_{k\in\mathbb{N}} \subset \mathcal{H}$ *converging pointwise to* $\varphi_i \in \mathcal{E}_2, i = 0,1$.

Of course, the statement should be understood as also including the claims that: (i) (84) is well defined independently of the choices of the approximating sequences, (ii) convergence in the metric completion is characterized as follows: $\{\varphi_j\} \subset \mathcal{E}_2$ converges to $\varphi \in \mathcal{E}_2$ if $\lim_j d_{\mathrm{M},2}(\varphi_j,\varphi) = 0$.

The first part Theorem 8.16, namely the statement that $(\mathcal{H}_\omega, d_{\mathrm{M}})$ is a metric space, was essentially proven by Chen in 2000 [33] following a conjecture by Donaldson in 1999 [52].

Remark 8.17. The space $\mathcal{H}$ (1) is the space of Kähler forms, while the space $\mathcal{H}_\omega$ (82) is the space of Kähler potentials. In many instances one can go back and forth between the two carelessly, however in some situations some care is needed. One may also identify the latter as a subspace of the former in several ways, but again some care is needed in doing so. For example,

$$\mathcal{H}_\omega \cap \{\mathrm{AM} = 0\} \tag{85}$$

is a d_{M}-totally geodesic submanifold (hypersurface) of $\mathcal{H}_\omega$ [69, Proposition 2.6.1], [52, § 3]. The submanifold (85) can be naturally identified with $\mathcal{H}$. Sometimes, though, we will use identifications different from (85).

In the vein of Remark 8.17, we distinguish between solutions of (78), which we continue to refer to as solutions to the Ricci flow, and solutions of (79), which we refer to as solutions to the Kähler–Ricci flow.

Exercise 8.18. Does the map $\omega(t) \mapsto \varphi(t)$ that sends solutions of (78) to solutions of (79), come from the identification of $\mathcal{H}$ with (85)?

Theorem 8.16 has already found several geometric applications. The first is the following analogue of Theorem 8.7 for the Mabuchi metric, due to Darvas [39, Theorem 6.1].

Definition 8.19. We say that (M, J) is Mabuchi–Ricci unstable (or MR-unstable) if there exists a Kähler–Ricci flow trajectory that diverges in $\overline{(\mathcal{H}, d_{\mathrm{M}})}$. Otherwise, we say (M, J) is MR-stable.

Theorem 8.20. *A Fano manifold (M, J) is MR-stable if and only if it admits a Kähler–Einstein metric. Moreover, if it is MR-unstable then any Ricci flow trajectory diverges in $\overline{(\mathcal{H}, d_{\mathrm{C}})}$.*

Exercise 8.21. Show that the length of the curve $t \mapsto \varphi(t)$ with respect to d_{M} is equal to

$$||f_{\omega_{\varphi(t)}}||_{L^1(\mathbb{R}_+, L^2(M, \omega_{\varphi(t)}))} \tag{86}$$

if $\varphi(t)$ satisfies (79) (which by Exercise 8.10 implies that $\omega_{\varphi(t)}$ satisfies the Ricci flow equation (78)). As observed by Darvas, Theorem 8.20 together with the arguments of [38] imply the following analogue of Corollary 8.13 first obtained by McFeron [71]: the flow (79) converges if and only if (86) is finite.

In fact, the following improvement of the last statement in Exercise 8.21 is due to Darvas. It follows from [39, Theorem 6.1] together with later work of Darvas surveyed in § 9:

Theorem 8.22. *The Kähler–Ricci flow* (79) *converges smoothly if and only if*

$$||f||_{L^1(\mathbb{R}_+, L^1(M, \omega_{\varphi(t)}))} < \infty, \tag{87}$$

where $f = f_{\omega_{\varphi(t)}}$ is the Ricci potential along the flow (*recall Definition* 3.4).

Other applications for Theorem 8.16 include the work of Streets [90], and more recently Berman–Darvas–Lu [18], who show that one gains new insight on the long time behavior of the Calabi flow by placing it in the context of the Mabuchi metric completion; the work of Darvas–He [43], where the asymptotic behavior of the Kähler–Ricci flow in the metric completion is related to destabilizing geodesic rays. We refer the reader to the survey [81] for more references.

9. The Darvas metric and its completion

Perhaps surprisingly, a key observation of Darvas is that not a Riemannian, but rather a *Finsler* metric, encodes the asymptotic behavior of the Aubin functional J. This is discussed in Section 10. In this section we introduce the Darvas metric and survey some of its basic properties. In later sections, through considerable more technical work, we survey later work of Darvas–Rubinstein that shows that the same is also true for essentially all energy functionals on $\mathcal{H}$ whose critical points are precisely various types of canonical metrics in Kähler geometry. In fact, as pointed out in [45, Remark 7.3], the same kind of statement is in general false for the much-studied Riemannian metrics of Calabi and Mabuchi. Thus, the Darvas metric turns out to be fundamental.

The Darvas metric is a weak Finsler metric on $\mathcal{H}_\omega$ given by [40],

$$\|\nu\|_\varphi^{\mathrm{D}} := V^{-1}\int_M |\nu|\omega_\varphi^n, \quad \nu \in T_\varphi\mathcal{H}_\omega = C^\infty(M). \tag{88}$$

As in § 8.2, define the length of a smooth curve $t \mapsto \varphi(t)$,

$$\ell_{\mathrm{D}}(\alpha) := \int_0^1 \int_M |\dot\varphi(t)|\omega^n_{\varphi(t)} \wedge dt. \tag{89}$$

Definition 9.1. The path length distance of $(\mathcal{H}_\omega, d_{\mathrm{D}})$ is defined by

$$d_{\mathrm{D}}(\varphi_0,\varphi_1) := \inf\big\{\ell_1(\alpha)\ :\ \alpha : [0,1] \to \mathcal{H}_\omega \\ \text{is a smooth curve with } \ \alpha(0)=\varphi_0,\, \alpha(1)=\varphi_1\big\}.$$

We call the pseudometric d_{D} the *Darvas metric.*

The following result of Darvas justifies this name. To state the result, consider $[0,1]\times\mathbb{R}\times M$ as a complex manifold of dimension $n+1$, and denote by $\pi_2 : [0,1]\times\mathbb{R}\times M \to M$ the natural projection.

Theorem 9.2 ([40, Theorem 3.5]). *$(\mathcal{H}_\omega, d_{\mathrm{D}})$ is a metric space. Moreover,*

$$d_{\mathrm{D}}(\varphi_0,\varphi_1) = \|\dot\varphi_0\|_{\varphi_0} \ge 0, \tag{90}$$

with equality iff $\varphi_0=\varphi_1$, where $\dot\varphi_0$ is the image of $(\varphi_0,\varphi_1)\in\mathcal{H}_\omega\times\mathcal{H}_\omega$ under the Dirichlet-to-Neumann map for the Monge–Ampère equation,

$$\varphi \in \mathrm{PSH}(\pi_2^\star\omega, [0,1]\times\mathbb{R}\times M), \quad (\pi_2^\star\omega + \sqrt{-1}\partial\bar\partial\varphi)^{n+1} = 0, \quad \varphi|_{\{i\}\times\mathbb{R}} = \varphi_i,\ i=0,1. \tag{91}$$

Remark 9.3. (i) The Dirichlet-to-Neumann operator simply maps (φ_0,φ_1) to the initial tangent vector of the curve

$$t \mapsto \varphi(t) \equiv \varphi_t$$

that solves (91).
(ii) One needs to make sense of the expression $\dot\varphi_0$ in (90) since there is no guarantee that φ_t will be smooth in t. Since φ (considered as a function on $[0,1]\times\mathbb{R}\times M$) is

$\pi_2^*\omega$-psh and independent of the imaginary part of the first variable, it is convex in t. Thus,

$$\dot{\varphi}_0(x) := \lim_{t\to 0^+} \frac{\varphi(t,x)-\varphi_0(x)}{t}, \tag{92}$$

with the limit well defined since the difference quotient is decreasing in t.

The metric completion of the Darvas metric is given by the next result [40, Theorem 2]. The proof is similar in spirit to that of Theorem 8.16, but involves considerable additional technicalities stemming, at least intuitively, from the fact that $x \mapsto x^2$ is a smooth function while $x \mapsto |x|$ is only Liphscitz; partly due to this dealing with an L^1 type metric is technically harder in this setting.

Theorem 9.4. *The metric completion of* $(\mathcal{H}_\omega, d_{\mathrm{D}})$ *equals* $(\mathcal{E}_1, d_{\mathrm{D}})$, *where*

$$d_{\mathrm{D}}(\varphi_0,\varphi_1) := \lim_{k\to\infty} d_{\mathrm{D}}(\varphi_0(k),\varphi_1(k)),$$

for any smooth decreasing sequences $\{\varphi_i(k)\}_{k\in\mathbb{N}} \subset \mathcal{H}_\omega$ *converging pointwise to* $\varphi_i \in \mathcal{E}_1, i=0,1$. *Moreover, for each* $t\in(0,1)$, *define*

$$\varphi_t := \lim_{k\to\infty} \varphi_t(k),\ \ t\in(0,1), \tag{93}$$

where $\varphi_t(k)$ *is the solution of* (91) *with endpoints* $\varphi_i(k), i=0,1$. *Then* $\varphi_t \in \mathcal{E}_1$, *and the curve* $t \to \varphi_t$ *is well defined independently of the choices of approximating sequences and is a* d_{D}*-geodesic.*

10. The Aubin functional and the Darvas distance function

Finally we come to the fact stated at the beginning of the previous section relating the Darvas metric to the Aubin functional.

The subspace

$$\mathcal{H}_0 := \mathrm{AM}^{-1}(0) \cap \mathcal{H}_\omega \tag{94}$$

is isomorphic to $\mathcal{H}$ (1), the space of Kähler metrics (recall Remark 8.17). We use this isomorphism to endow $\mathcal{H}$ with a metric structure, by pulling back the Darvas metric defined on $\mathcal{H}_\omega$.

Proposition 10.1 ([40, Remark 6.3]). *There exists* $C>1$ *such that for all* $\varphi\in\mathcal{H}_0$ (*recall* (94)),

$$\frac{1}{C}J(\varphi) - C \le d_{\mathrm{D}}(0,\varphi) \le CJ(\varphi)+C.$$

We refer the reader to [45, Proposition 5.5] for a proof.

Given the equivalence of J and d_{D} on $\mathcal{H}_0$ it is natural to expect that this should extend to the metric completion. This is indeed the case. This amounts to two things: (i) one can extend Aubin's functional J to the metric completion in a continuous way with respect to the d_{D}-topology, (ii) $\mathcal{H}_0$, considered as a submanifold of $\mathcal{H}$ endowed with the metric induced by d_{D}, is a totally geodesic metric space whose completion coincides with $\mathcal{E}_1 \cap \mathrm{AM}^{-1}(0)$, which in turn requires verifying that the Aubin–Mabuchi functional AM can be extended to $\mathcal{E}_1$ in a continuous way with respect to the d_{D}-topology. These facts are contained in the following Lemma [45, Lemma 5.2].

Lemma 10.2.

(i) $\mathrm{AM}, J : \mathcal{H}_\omega \to R$ *each admit a unique* d_{D}*-continuous extension to* $\mathcal{E}_1$ *and these extensions still satisfy* (9) *and* (7) (*in the sense of pluripotential theory*).

(ii) *The subspace* $(\mathcal{E}_1 \cap \mathrm{AM}^{-1}(0), d_{\mathrm{D}})$ *is a complete geodesic metric space, coinciding with the metric completion of* $(\mathcal{H}_0, d_{\mathrm{D}})$ (*recall* (94)).

Consequently, from now on we denote by AM, J the unique d_{D}-continuous extensions to $\mathcal{E}_1$ given by the previous lemma.

Corollary 10.3. *There exists* $C > 1$ *such that for all* $\varphi \in \mathcal{E}_1 \cap \mathrm{AM}^{-1}(0)$,

$$\frac{1}{C} J(\varphi) - C \leq d_{\mathrm{D}}(0, \varphi) \leq C J(\varphi) + C.$$

Next, we discuss a concrete formula for the d_{D} metric relating it to the Aubin–Mabuchi energy and also give a concrete growth estimate for d_{D}. First we need to introduce the following rooftop type envelope for $u, v \in \mathcal{E}_1$:

$$P(u,v)(z) := \sup \big\{ w(z) \,:\, w \in \mathrm{PSH}(M, \omega),\ w \leq \min\{u, v\} \big\}.$$

Note that $P(u,v) \in \mathcal{E}_1$ [39, Theorem 2]. Darvas shows the following beautiful "Pythagorean" formula for d_{D}, as well as a very useful growth estimate [40, Corollary 4.14, Theorem 3].

Proposition 10.4. *Let* $u, v \in \mathcal{E}_1$*. Then,*

$$d_{\mathrm{D}}(u,v) = \mathrm{AM}(u) + \mathrm{AM}(v) - 2\mathrm{AM}(P(u,v)). \tag{95}$$

Also, there exists $C > 1$ *such that for all* $u, v \in \mathcal{E}_1$,

$$C^{-1} d_{\mathrm{D}}(u,v) \leq \int_M |u - v| \omega_u^n + \int_M |u - v| \omega_v^n \leq C d_{\mathrm{D}}(u,v). \tag{96}$$

11. Quotienting the metric completion by a group action

We now incorporate automorphisms into the picture. Since automorphisms induce isometries of the various infinite-dimensional metrics we have studied so far it is natural to consider the associated quotient spaces from the metric geometry point of view. In addition, the various functionals we have studied also admit natural descents to the quotient spaces.

11.1. The action of the automorphism group on $\mathcal{H}$

Let $\mathrm{Aut}_0(M, \mathrm{J})$ denote the connected component of the complex Lie group of automorphisms (biholomorphisms, i.e., homeomorphisms that are holomorphic and admit a holomorphic inverse) of (M, J). Denote by $\mathrm{aut}(M, \mathrm{J})$ the Lie algebra of $\mathrm{Aut}_0(M, \mathrm{J})$, consisting of infinitesimal automorphisms, i.e., real vector fields X satisfying $\mathcal{L}_X \mathrm{J} = 0$, equivalently,

$$\mathrm{J}[X, Y] = [X, \mathrm{J}Y], \quad \forall\, X \in \mathrm{aut}(M, \mathrm{J}),\ \forall\, Y \in \mathrm{diff}(M), \tag{97}$$

where $\mathrm{diff}(M)$ denotes all smooth vector fields on M. Thus $\mathrm{aut}(M, \mathrm{J})$ is a complex Lie algebra with complex structure J.

The automorphism group $\mathrm{Aut}(M, \mathrm{J})_0$ acts on $\mathcal{H}$ by pullback:

$$f.\eta := f^\star \eta, \qquad f \in \mathrm{Aut}(M, \mathrm{J})_0, \quad \eta \in \mathcal{H}. \tag{98}$$

Given the one-to-one correspondence between $\mathcal{H}$ and $\mathcal{H}_0$, the group $\mathrm{Aut}(M, \mathrm{J})_0$ also acts on $\mathcal{H}_0$. The action is described in the next lemma.

Lemma 11.1. *For $\varphi \in \mathcal{H}_0$ and $f \in \mathrm{Aut}(M, \mathrm{J})_0$ let $f.\varphi \in \mathcal{H}_0$ be the unique element such that $f.\omega_\varphi = \omega_{f.\varphi}$. Then,*

$$f.\varphi = f.0 + \varphi \circ f, \qquad f \in \mathrm{Aut}(M, \mathrm{J})_0, \quad \varphi \in \mathcal{H}_0. \tag{99}$$

Proof. Note that (99) is a Kähler potential for $f^\star\omega_\varphi$. Indeed, $f \in \mathrm{Aut}(M, \mathrm{J})$ implies that $f^\star\sqrt{-1}\partial\bar\partial\varphi = \sqrt{-1}\partial\bar\partial\varphi \circ f$. That $\mathrm{AM}(f.0 + \varphi \circ f) = 0$ follows from Exercise 11.2 as we have,

$$\begin{aligned}\mathrm{AM}(f.0 + \varphi \circ f) &= \mathrm{AM}(f.0 + \varphi \circ f) - \mathrm{AM}(f.0)\\ &= \int_M \varphi \circ f \sum_{j=0}^n f^\star\omega^{n-j} \wedge f^\star\omega_\varphi^j = \mathrm{AM}(\varphi) = 0.\end{aligned}$$

(Of course, $\mathrm{AM}(f.0) = 0$ since by definition $f.0 \in \mathcal{H}_0$.) □

Exercise 11.2. Show that

$$\mathrm{AM}(v) - \mathrm{AM}(u) = \frac{V^{-1}}{n+1} \int_M (v - u) \sum_{k=0}^n \omega_u^{n-k} \wedge \omega_v^k. \tag{100}$$

Among other things, this formula shows that AM is monotone, i.e.,

$$u \le v \quad \Rightarrow \quad \mathrm{AM}(u) \le \mathrm{AM}(v). \tag{101}$$

Lemma 11.3. *The action of* $\mathrm{Aut}(M, \mathrm{J})_0$ *on $\mathcal{H}_0$ is a d_{D}-isometry.*

Proof. From (99),

$$\frac{d}{dt} f.\varphi_t = \dot\varphi_t \circ f,$$

for any smooth path $t \mapsto \varphi_t$ in $\mathcal{H}_0$. Thus, the d_{D}-length of $t \mapsto f.\varphi_t$ is

$$V^{-1} \int_{[0,1]\times M} |\dot\varphi_t \circ f| f^\star\omega_{\varphi_t}^n \wedge dt = V^{-1} \int_{[0,1]\times M} |\dot\varphi_t| \omega_{\varphi_t}^n \wedge dt,$$

equal to the d_{D}-length of φ_t. □

Suppose G is a subgroup of $\mathrm{Aut}(M, \mathrm{J})_0$. By the previous lemma G acts on $\mathcal{H}$ by d_{D}-isometries, hence induces a pseudometric on the orbit space $\mathcal{H}/G$,

$$d_{\mathrm{D},G}(Gu, Gv) := \inf_{f,g\in G} d_{\mathrm{D}}(f.u, g.v).$$

Here, we denote by Gu the orbit of u under the action of G. Naturally, Gu is an element of the orbit space $\mathcal{H}/G$. Thus, $d_{\mathrm{D},G}$ measures the distance between orbits.

It is natural to expect that the group action extends to the metric completion. This is indeed the case.

Lemma 11.4. *Let (X,ρ) and (Y,δ) be two complete metric spaces, W a dense subset of X and $f: W \to Y$ a C-Lipschitz function, i.e.,*

$$\delta(f(a), f(b)) \le C\rho(a,b), \quad \forall\, a, b \in W. \tag{102}$$

Then f has a unique C-Lipschitz continuous extension to a map $\bar{f}: X \to Y$.

Proof. Let $w_k \in W$ be a Cauchy sequence converging to some $w \in X$. Lipschitz continuity gives

$$\delta(f(w_k), f(w_l)) \le C\rho(w_k, w_l),$$

hence $\bar{f}(w) := \lim_k f(w_k) \in Y$ is well defined and independent of the choice of approximating sequence w_k. Choose now another Cauchy sequence $z_k \in W$ with limit $z \in X$, plugging in w_k, z_k in (102) and taking the limit gives that $\bar{f}: X \to Y$ is C-Lipschitz continuous. □

Lemma 11.5. *The action of $\mathrm{Aut}(M,\mathrm{J})_0$ on $\mathcal{H}_0$ has a unique d_{D}-isometric extension to the metric completion $\overline{(\mathcal{H}_0, d_{\mathrm{D}})} = (\mathcal{E}_1 \cap \mathrm{AM}^{-1}(0), d_{\mathrm{D}})$.*

Proof. Because $\mathrm{Aut}(M,\mathrm{J})_0$ acts by d_{D}-isometries, each $f \in \mathrm{Aut}(M,\mathrm{J})_0$ induces a 1-Lipschitz continuous self-map of $\mathcal{H}_0$. By Lemma 11.4, such maps have a unique 1-Lipschitz extension to the completion $\mathcal{E}_1 \cap \mathrm{AM}^{-1}(0)$ and the extension is additionally a d_{D}-isometry. By density, the laws governing a group action have to be preserved as well. □

For any Lie subgroup K of the isometry group of (M, g_ω) define the subspace

$$\mathcal{H}_\omega^K := \{\varphi \in \mathcal{H}_\omega \,:\, \varphi \text{ is invariant under } K\}, \tag{103}$$

and similarly define $\mathcal{H}_0^K = \mathcal{H}^K \cap \mathrm{AM}^{-1}(0)$. According to Theorem 9.4, the d_{D}-metric completion of $\mathcal{H}_\omega^K$ is

$$\mathcal{E}_1^K := \{u \in \mathcal{E}_1 \,:\, u \text{ is invariant under } K\}.$$

The next result follows using the arguments in the proofs of Lemmas 10.2 and 11.5.

Lemma 11.6. *The metric completion of $(\mathcal{H}_0^K, d_{\mathrm{D}})$ is $\mathcal{E}_1^K \cap \mathrm{AM}^{-1}(0)$.*

11.2. The Aubin functional on the quotient space

Let $G \subset \mathrm{Aut}(M,\mathrm{J})_0$ be a subgroup. Following Zhou–Zhu [107, Definition 2.1] and Tian [98, Definition 2.5], define the descent of J to $\mathcal{H}/G$,

$$J_G(Gu) := \inf_{g \in G} J(g.u).$$

By Lemma 10.2 this functional can be extended to a functional

$$J_G : \mathcal{E}_1 \cap \mathrm{AM}^{-1}(0)/G \to \mathbb{R},$$

still satisfying

$$J_G(Gu) = \inf_{g \in G} J(g.u). \tag{104}$$

We now see that the key inequality between the Aubin functional and the Darvas distance function (Proposition 10.1) descends to the metric completion of the quotient space.

Lemma 11.7. *For $u \in \mathcal{E}_1 \cap \mathrm{AM}^{-1}(0)$ we have*

$$\frac{1}{C} J_G(Gu) - C \leq d_{\mathrm{D},G}(G0, Gu) \leq C J_G(Gu) + C, \tag{105}$$

where $d_{\mathrm{D},G}$ is the pseudometric of the quotient $\mathcal{E}_1 \cap \mathrm{AM}^{-1}(0)/G$.

Proof. By Lemma 11.3,

$$d_{\mathrm{D},G}(G0, Gu) = \inf_{f \in G} d_{\mathrm{D}}(0, f.u).$$

The result now follows from Proposition 10.1. □

12. A modified conjecture

At last, we return to Conjecture 1.2 and pick up the discussion from where we left it at the end of Section 6. Lemma 11.7 motivates the following modification of Conjecture 1.2.

Definition 12.1. Let $F : \mathcal{H} \to \mathbb{R}$ be G-invariant.

- We say F is $d_{\mathrm{D},G}$*-proper* if for some $C, D > 0$,

$$F(u) \geq C d_{\mathrm{D},G}(G0, Gu) - D.$$

- We say F is J_G*-proper* if for some $C, D > 0$,

$$F(u) \geq C J_G(Gu) - D.$$

Conjecture 12.2. *Let (M, J, ω) be a Fano manifold. Set $G := \mathrm{Aut}(M, \mathrm{J})_0$. There exists a Kähler–Einstein metric in $\mathcal{H}$ if and only if the descent of the Mabuchi energy E to the quotient space $\mathcal{H}/G$ is $d_{\mathrm{D},G}$-proper (equivalently, J_G-proper).*

Note that according to Lemma 11.7 both notions of properness are indeed equivalent. Also, the G-invariance condition can be considered as a version of the Futaki obstruction [56].

Albeit being a purely analytic criterion, properness should be morally equivalent to properness in a metric geometry sense, namely, that the Mabuchi functional should grow at least linearly relative to some metric on $\mathcal{H}$, and this is precisely the content of Conjecture 12.2.

Remark 12.3. We now come back to the analogy with the Dirichlet energy alluded to in the Prologue. There we seek to minimize the Dirichlet energy, say on the unit ball in $\mathbb{R}^n$,

$$E(f) := \int_{B_1(0)} \sum_{i=1}^{n} (\partial_{x_i} f)^2 dx^1 \wedge \cdots dx^n.$$

The space of competitors $\mathcal{H}$ is now the space of smooth functions with prescribed boundary values $g \in C^\infty(\partial B_1(0))$,

$$\mathcal{H} := \{f \in C^\infty(B_1(0)) \, : \, f|_{\partial B_1(0)} = g\}.$$

In some sense, the prescribed boundary values can be morally thought of as the analogue for fixing a Kähler class. What is the analogue of the Aubin functional? In this case it is just E itself, i.e., we put $J = E$, so an analogue of Conjecture 1.2 is trivial here. However, the direct method in the calculus of variations motivates replacing J (which is the $W^{1,2}$ seminorm) with the $W^{1,2}$ norm. Namely, we consider the metric

$$(h, k) := \int \sum_{i=1}^{n} \partial_{x_i} h \partial_{x_i} k dx^1 \wedge \cdots dx^n + \int hk dx^1 \wedge \cdots dx^n.$$

The path-length distance is then just the one coming from the norm $W^{1,2}$, and the properness inequality is a consequence of the Poincaré inequality. This then implies that a minimizer exists in the $W^{1,2}$ completion of $\mathcal{H}$. The Euler–Lagrange equation is precisely the Laplace equation with prescribed boundary data. Elliptic regularity theory then shows the minimizer must be an element of $\mathcal{H}$ itself, hence a smooth harmonic function agreeing with g on the boundary.

In the remainder of these notes, we sketch the resolution of Conjecture 12.2 due to Darvas–Rubinstein [45].

Theorem 12.4. *Conjecture* 12.2 *holds.*

The proof of this result is completed in Section 15.

Remark 12.5. The easier implication "J_G-proper $\Rightarrow$ existence of Kähler–Einstein" is due to Tian [98, Theorem 2.6] and is a modification of the proof of Theorem 5.1. Our proof of Theorem 12.4 also furnishes a new proof of this fact. In the special case of toric Fano manifolds, a variant of the converse direction is due to Zhou–Zhu [107, Theorem 0.2].

13. A general existence/properness principle

Motivated by Remark 12.3, we approach Conjecture 12.2 using an abstract metric geometry framework. While seemingly abstract it turns out to be a powerful way of dealing with several different minimization problems in Kähler geometry.

Notation 13.1. The data $(\mathcal{R}, d, F, G)$ is defined as follows.

(A1) $(\mathcal{R}, d)$ is a metric space with a distinguished element $0 \in \mathcal{R}$, whose metric completion is denoted $(\overline{\mathcal{R}}, d)$.

(A2) $F : \mathcal{R} \to R$ is lower semicontinuous (lsc). Let $F : \overline{\mathcal{R}} \to R \cup \{+\infty\}$ be the largest lsc extension of $F : \mathcal{R} \to R$:

$$F(u) = \sup_{\varepsilon > 0} \left(\inf_{\substack{v \in \mathcal{R} \\ d(u,v) \le \varepsilon}} F(v) \right), \quad u \in \overline{\mathcal{R}}.$$

For each $u, v \in \mathcal{R}$ define also

$$F(u,v) := F(v) - F(u).$$

(A3) The set of minimizers of F on $\overline{\mathcal{R}}$ is denoted

$$\mathcal{M} := \Big\{ u \in \overline{\mathcal{R}} \ : \ F(u) = \inf_{v \in \overline{\mathcal{R}}} F(v) \Big\}.$$

(A4) Let G be a group acting on $\mathcal{R}$ by $G \times \mathcal{R} \ni (g,u) \to g.u \in \mathcal{R}$. Denote by $\mathcal{R}/G$ the orbit space, by $Gu \in \mathcal{R}/G$ the orbit of $u \in \mathcal{R}$, and define $d_G : \mathcal{R}/G \times \mathcal{R}/G \to \mathbb{R}_+$ by

$$d_G(Gu, Gv) := \inf_{f,g \in G} d(f.u, g.v).$$

Hypothesis 13.2. The data $(\mathcal{R}, d, F, G)$ satisfies the following properties.

(P1) For any $\varphi_0, \varphi_1 \in \mathcal{R}$ there exists a d-geodesic segment $[0,1] \ni t \mapsto \varphi_t \in \overline{\mathcal{R}}$ for which $t \mapsto F(\varphi_t)$ is continuous and convex on $[0,1]$.
(P2) If $\{\varphi_j\}_j \subset \overline{\mathcal{R}}$ satisfies $\lim_{j\to\infty} F(\varphi_j) = \inf_{\overline{\mathcal{R}}} F$, and for some $C > 0$, $d(0, \varphi_j) \leq C$ for all j, then there exists a $u \in \mathcal{M}$ and a subsequence $\{\varphi_{j_k}\}_k$ d-converging to u.
(P3) $\mathcal{M} \subset \mathcal{R}$.
(P4) G acts on $\mathcal{R}$ by d-isometries.
(P5) G acts on $\mathcal{M}$ transitively.
(P6) If $\mathcal{M} \neq \emptyset$, then for any $u, v \in \mathcal{R}$ there exists $g \in G$ such that $d_G(Gu, Gv) = d(u, g.v)$.
(P7) For all $u, v \in \mathcal{R}$ and $g \in G$, $F(u,v) = F(g.u, g.v)$.

The following result will provide the aforementioned framework for dealing with many minimization problems.

Theorem 13.3. *Let $(\mathcal{R}, d, F, G)$ be as in Notation* 13.1 *and satisfying Hypothesis* 13.2. *Then $\mathcal{M}$ is nonempty if and only if $F : \mathcal{R} \to \mathbb{R}$ is G-invariant, and for some $C, D > 0$,*

$$F(u) \geq C d_G(G0, Gu) - D, \quad \textit{for all } u \in \mathcal{R}. \tag{106}$$

One direction in this theorem is easy. Namely, if (106) holds, then F is bounded from below. By (A2),

$$\inf_{v \in \overline{\mathcal{R}}} F(v) = \inf_{v \in \mathcal{R}} F(v). \tag{107}$$

This, combined with (106), the G-invariance of F and the definition of d_G implies there exists $\varphi_j \in \mathcal{R}$ such that $\lim_j F(\varphi_j) = \inf_{\overline{\mathcal{R}}} F$ and $d(0, \varphi_j) \leq d_G(G0, G\varphi_j) + 1 < C$ for C independent of j. By (P2), $\mathcal{M}$ is non-empty. For the other direction we refer the reader to [45, Theorem 3.4].

We have set up things in such a way that the modified properness conjecture, Conjecture 12.2, would become a corollary of Theorem 13.3 applied to the following data

$$\mathcal{R} = \mathcal{H}_0, \quad d = d_1, \quad F = E, \quad G := \mathrm{Aut}_0(M, \mathrm{J}), \tag{108}$$

if this data satisfies the hypothesis of Theorem 13.3. In the next sections we verify that this is indeed the case. Property (P4) has already been verified in Lemma 11.5. In the next few sections we verify the remaining hypothesis of Theorem 13.3.

14. Applying the general existence/properness principle

In this section we briefly motivate – in the context of the Kähler–Einstein problem – some of the key assumptions in the general existence/properness principle. The point is to convince the reader that this principle fits naturally/seamlessly with classical/foundational results in Kähler geometry.

First, a seemingly harmless condition, tucked into the "notation" part of Theorem 13.3, is that the functional we are trying to minimize on the metric completion should be the greatest lower semicontinuous extension (with respect to the path-length metric) of the functional we are trying to study originally on the "regular" objects $\mathcal{R}$. This turns out to be quite a technical thing to verify. At first, this might cause confusion: indeed any functional admits such an extension by means of the abstract formula

$$F(u) = \sup_{\varepsilon>0}\Bigg(\inf_{\substack{v\in\mathcal{R}\\ d(u,v)\leq\varepsilon}} F(v)\Bigg), \quad u \in \overline{\mathcal{R}}. \tag{109}$$

However, the issue is to verify that this abstract formula, say in the case of the Mabuchi energy, coincides with the original defining formula (12) which initially only makes sense on the space of smooth potentials $\mathcal{R} = \mathcal{H}$. This is because only then can we actually verify that this extended functional satisfies the other hypothesis in Theorem 13.3 (without an explicit formula it is not clear how to proceed). Fortunately, condition (A2) for (108) does hold by the following result [45, Proposition 5.21].

Proposition 14.1. *Formula* (12) *coincides with formula* (109) *on $\mathcal{E}_1$. In other words, formula* (12) *gives the greatest d_1-lsc extension of $E : \mathcal{H} \to R$ to $\mathcal{E}_1$.*

Remark 14.2. The analogue of this result for the Mabuchi metric d_{M} can be found in [18].

Second, property (P1) holds for the Mabuchi energy due to a result of Berman–Berndtsson [14, Theorem 1.1]. In fact, we remark that it is well known that the geodesic between smooth endpoints has considerable regularity (as compared to just being in $\overline{\mathcal{R}}$) [24, 33]. In [14] it is shown that the Mabuchi energy is convex along such partially regular geodesics.

Third, property (P2) stipulates precompactness of sublevel sets of the Mabuchi energy with respect to the Darvas metric. Pre-compactness with respect to other functionals is a key result in the works [17, 25], and can be adapted to show the aforementioned pre-compactness [45, Proposition 5.28].

Fourth, property (P3) stipulates regularity of minimizers of the Mabuchi energy in the metric completion. This follows from the regularity result of Berman

[12, Theorem 1.1] combined with the characterization of the metric completion of Darvas (Theorem 9.4).

Fifth, property (P5), modulo property (P3), amounts to the classical Bando–Mabuchi theorem on uniqueness of Kähler–Einstein metrics up to automorphisms.

Sixth, property (P7) says that the Mabuchi functional is exact, or of "Bott–Chern" type, and this is precisely Mabuchi's original theorem on his functional [68, Theorem 2.4]. For an expository treatment we referred to [81, § 5].

Finally, property (P6) is a new ingredient, and so we go into more detail, sketching property (P6) for (108). It fits nicely into our framework since it shows precisely the role of another classical result in Kähler geometry, namely, Matsushima's classical theorem about the automorphism group of a Kähler–Einstein manifold. The key result in showing (P6) is the following [45, Proposition 6.8].

Proposition 14.3. *Let* $(M, \mathrm{J}, \omega, g)$ *be Kähler–Einstein. Define* $(\mathcal{R}, d, F, G)$ *by* (108)*, and suppose that* (A1)–(A4) *and* (P4) *hold. Finally, assume the following:*

(i) *For each* $X \in \mathrm{isom}(M, g)$*,* $t \mapsto \exp_I tJX.\omega$ *is a* d_{D}*-geodesic whose speed depends continuously on* X*.*

(ii) $\mathrm{Aut}(M, \mathrm{J})_0 \times \mathrm{Aut}(M, \mathrm{J})_0 \ni (f, g) \mapsto d(f.u, g.v)$ *is a continuous map for every* $u, v \in \mathcal{H}$*.*

Then property (P6) *holds.*

Condition (i) is essentially a corollary of (90), while (ii) follows from (96). Property (P6) stipulates that a certain infimum over the group G is attained. Thus, for the proof of Proposition 14.3 we decompose the group G into a compact part and a non-compact part in such a way that the compact part acts by d-isometries while on the non-compact part (but finite-dimensional!) we have d-properness. Then, together with conditions (i) and (ii), the existence of a minimizer is guaranteed.

The aforementioned decomposition of the group into a compact and a non-compact part is stated in Corollary 14.7 below. It should be well known and relies on classical results that we now recall. First, we recall Matsushima's classical theorem [70, Théorème 1]. We refer to Gauduchon [57] for more details. Let $g(\cdot\,,\cdot) = \omega(\cdot\,, \mathrm{J}\,\cdot)$ denote the Riemannian metric associated to (M, J, ω). Denote by $\mathrm{Isom}(M, g)_0$ the identity component of the isometry group of (M, g). Since M is compact so is $\mathrm{Isom}(M, g)_0$ [76, Proposition 29.4]. Denote by $\mathrm{isom}(M, g)$ the Lie algebra of $\mathrm{Isom}(M, g)_0$.

Theorem 14.4. *Let* $(M, \mathrm{J}, \omega, g)$ *be a Fano Kähler manifold. Suppose* g *is a Kähler–Einstein metric. Then,*

$$\mathrm{aut}(M, \mathrm{J}) = \mathrm{isom}(M, g) \oplus \mathrm{J}\,\mathrm{isom}(M, g). \tag{110}$$

The following result is classical, and we only state its Kähler–Einstein version, whose proof we sketch.

Theorem 14.5. *Let* $(M, \mathrm{J}, \omega, g)$ *be Kähler–Einstein. Then any maximally compact subgroup of* $\mathrm{Aut}(M, \mathrm{J})_0$ *is conjugate to* $\mathrm{Isom}(M, g)_0$*.*

Proof. By a Theorem of Iwasawa–Malcev [91, Theorem 32.5], if G is a connected Lie group then its maximal compact subgroup must be connected and any two maximal compact subgroups are conjugate. But then by Theorem 14.4 $\mathrm{Isom}(M,g)_0$ has to be a maximal compact subgroup of $\mathrm{Aut}(M,\mathrm{J})_0$. □

Next, we need a version of the classical Cartan decomposition [27, Proposition 32.1, Remark 31.1].

Theorem 14.6. *Let S be a compact connected semisimple Lie group. Denote by $(S^{\mathbb{C}},\mathrm{J})$ the complexification of S, namely the unique connected complex Lie group whose Lie algebra is the complexification of that of $\mathfrak{s}$, the Lie algebra of S. Then the map C from $S\times\mathfrak{s}$ to $S^{\mathbb{C}}$ given by*

$$(s,X)\mapsto C(s,X):=s\exp_I \mathrm{J}X \tag{111}$$

is a diffeomorphism.

Combining Theorems 14.4, 14.5 and 14.6 we obtain the decomposition of $\mathrm{Aut}(M,\mathrm{J})_0$ into a compact and a non-compact part that is needed for the proof of Proposition 14.3. For details on how the following result yields Proposition 14.3 we refer to [45, § 6], where a more general result is proven in the constant scalar curvature setting (when the Cartan type decomposition is not given by classical results and we construct instead a “partial Cartan decomposition” that may only be surjective).

Corollary 14.7. *Let (M,J,ω,g) be Kähler–Einstein. Then the map C from*

$$\mathrm{Isom}(M,g)_0\times \mathrm{isom}(M,g) \quad \textit{to} \quad \mathrm{Aut}(M,\mathrm{J})_0$$

given by

$$(s,X)\mapsto C(s,X):=s\exp_I \mathrm{J}X \tag{112}$$

is a diffeomorphism.

15. A proof of Tian’s modified first conjecture

As already explained at the end of Section 13, and as we started to elaborate in the previous section, we prove Theorem 12.4 by applying Theorem 13.3 to data (108). Thus, it only remains to verify that this data satisfies the hypothesis of Theorem 13.3.

First, we go over Notation 13.1. First, in (A1), $\overline{\mathcal{R}}=\mathcal{E}_1\cap\mathrm{AM}^{-1}(0)$ by Theorem 9.4 and Lemma 10.2. Observe that (A2) holds by Proposition 14.1. In (A3), the minimizers of F are denoted by $\mathcal{M}$. Finally, (A4) holds since $G\subset\mathrm{Aut}(M,\mathrm{J})_0$ implies that if $g\in G$ and $\eta\in\mathcal{H}$ then $g.\eta$ is both Kähler and cohomologous to η, i.e., $g.\eta\in\mathcal{H}$. Thus, it remains to verify Hypothesis 13.2.

Properties (P1)–(P7) were all verified in § 14 with the exception of property (P4), that itself follows from Lemma 11.3.

Finally, we need to justify why we did not state E must be $\mathrm{Aut}(M,\mathrm{J})_0$-invariant in Theorem 12.4, while it is needed to apply Theorem 13.3. This follows from Futaki's theorem [56, p. 437]. Indeed, as in the proof of Claim 6.1

$$\frac{d}{dt}E((\exp_I tX)^\star\omega_\varphi) = C_X,$$

for some $\mathbb{R} \ni C_X$ depending on X but not on $\omega_\varphi \in \mathcal{H}$. Also,

$$\frac{d}{dt}E((\exp_I -tX)^\star\omega_\varphi) = -C_X.$$

Thus, unless this derivative, i.e., C_X, is zero for every $X \in \mathrm{aut}(M,\mathrm{J})$ and $\omega_\varphi \in \mathcal{H}$, the functional E cannot be bounded from below. Now, properness of E with respect to any nonnegative functional implies E is bounded from below. Thus, J_G-properness of E implies it is $\mathrm{Aut}(M,\mathrm{J})_0$-invariant.

16. A proof of Tian's second conjecture: the Moser–Trudinger inequality

We now explain the proof of Tian's second properness conjecture. First, let us recall the statement.

Denote by Λ_1 the real eigenspace of the smallest positive eigenvalue of $-\Delta_\omega$, and set

$$\mathcal{H}_\omega^\perp := \{\varphi \in \mathcal{H} : \int \varphi\psi\omega^n = 0,\ \forall\psi \in \Lambda_1\}.$$

Conjecture 16.1. *Suppose (M,J,ω) is Fano Kähler–Einstein. Then for some C, $D > 0$,*

$$E(\varphi) \geq CJ(\varphi) - D, \qquad \varphi \in \mathcal{H}_\omega^\perp. \tag{113}$$

Observe that no invariance properties are assumed, and the functionals are not taken on the quotient space. Instead, an orthogonality assumption is made.

Conjecture 1.3 was originally motivated by results in conformal geometry related to the determination of the best constants in the borderline case of the Sobolev inequality. By restricting to functions orthogonal to the first eigenspace of the Laplacian, Aubin was able to improve the constant in the aforementioned inequality on spheres [6, p. 235]. This can be seen as the sort of coercivity of the Yamabe energy occuring in the Yamabe problem, and it clearly fails without the orthogonality assumption due to the presence of conformal maps. Conjecture 1.3 stands in clear analogy with the picture in conformal geometry, by stipulating that coercivity of the K-energy holds in 'directions perpendicular to holomorphic maps' (when ω is Kähler–Einstein, it is well known that Λ_1 is in a one-to-one correspondence with holomorphic gradient vector fields, in fact this is how Matsushima's Theorem 14.4 is proven [31, 57]). It can be thought of as a higher-dimensional fully nonlinear generalization of the classical Moser–Trudinger inequality.

It is a rather simple consequence of the work of Bando–Mabuchi [9] that when a Kähler–Einstein metric exists, J_G-properness implies J-properness on $\mathcal{H}_\omega^\perp$

[96, Corollary 5.4],[107, Lemma A.2],[98, Theorem 2.6]. We now explain how to carry this through. The key is to study the Aubin functional restricted to orbits of $\mathrm{Aut}(M,\mathrm{J})_0$ and identify the minimizers and relate them to the first eigenspace.

Fix $\eta\in\mathcal{H}$. Let $F_\eta:\mathrm{Aut}(M,\mathrm{J})_0\to\mathbb{R}_+$ be given by

$$F_\eta(g):=(I-J)(g^\star\eta)=V^{-1}\frac{1}{n+1}\int_M\sqrt{-1}\partial\varphi_g\wedge\bar\partial\varphi_g\wedge\sum_{l=0}^{n-1}(n-l)\omega^{n-l-1}\wedge(g^\star\eta)^l,$$

where $\varphi_g\in\mathcal{H}_\omega$ is such that $g^\star\eta=\omega+\sqrt{-1}\partial\bar\partial\varphi_g$ (i.e., where the $I-J$ energy of $g^\star\eta$ with respect to the reference form ω).

Lemma 16.2. *Suppose $(M,\mathrm{J},\eta=\omega_\psi)$ is Fano Kähler–Einstein. Then*

$$h\in\mathrm{Aut}(M,\mathrm{J})_0$$

is a critical point of F_η precisely if $-\varphi_h\in\mathcal{H}^\perp_{h^\star\eta}$.

Proof. Using (14) and (10),

$$\begin{aligned}\frac{d}{d\delta}(I-J)(\varphi(\delta))&=\frac{d}{d\delta}\mathrm{AM}(\varphi)-\frac{d}{d\delta}V^{-1}\int\varphi(\delta)\omega^n_{\varphi(\delta)}\\&=V^{-1}\int\frac{d}{d\delta}\varphi(\delta)\omega^n_{\varphi(\delta)}-V^{-1}\int\big(\frac{d}{d\delta}\varphi(\delta)+\varphi(\delta)\Delta_{\omega_{\varphi(\delta)}}\frac{d}{d\delta}\varphi(\delta)\big)\omega^n_{\varphi(\delta)}\\&=-V^{-1}\int\varphi(\delta)\Delta_{\omega_{\varphi(\delta)}}\frac{d}{d\delta}\varphi(\delta)\omega^n_{\varphi(\delta)}.\end{aligned}\tag{114}$$

Writing $g_t=h\exp_I tX$ with $X\in\mathrm{aut}(M,\mathrm{J})$, observe that

$$\begin{aligned}\sqrt{-1}\partial\bar\partial\dot\varphi_h&=\sqrt{-1}\partial\bar\partial\dot\varphi_{g_0}\\&=\frac{d}{dt}\Big|_0(\omega+\sqrt{-1}\partial\bar\partial\varphi_{g_t})\\&=\frac{d}{dt}\Big|_0 g_t^\star\eta=\frac{d}{dt}\Big|_0(\exp_I tX)^\star(h^\star\eta)=\sqrt{-1}\partial\bar\partial\psi^X_{h^\star\eta}.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{d}{dt}\Big|_0F_\eta(g_t)&=-V^{-1}\int\varphi_h\Delta_{h^\star\eta}\dot\varphi_h(h^\star\eta)^n=-V^{-1}\int\varphi_h\Delta_{h^\star\eta}\psi^X_{h^\star\eta}(h^\star\eta)^n\\&=V^{-1}\int\varphi_h\psi^X_{h^\star\eta}(h^\star\eta)^n.\end{aligned}\tag{115}$$

Since this holds for all $X\in\mathrm{aut}(M,\mathrm{J})$, it follows that $\varphi_h\in\mathcal{H}^\perp_{h^\star\eta}$. □

Lemma 16.3. *Suppose $(M,\mathrm{J},\eta=\omega_\psi)$ is Fano Kähler–Einstein. By Theorem* 14.5 *then $\mathrm{Aut}(M,\mathrm{J})_0=K^{\mathbb{C}}$ for a maximally compact subgroup K. Suppose $\omega\in\mathcal{H}^K$. Then F_η has a unique critical point which is a global minimum.*

Proof. We start with the following observation.

Exercise 16.4. If $g\in\mathrm{Aut}(M,\mathrm{J})_0$ preserves ω then

$$(I-J)(g^\star\eta)=(I-J)(\eta).$$

Thus, using the Cartan decomposition (Corollary 14.7), F_η descends to a function on $\mathrm{isom}(M, g)$, still denoted by F_η,

$$F_\eta(X) = (I - J)\big((\exp_I \mathrm{J}X)^\star\eta\big).$$

Now, we show that the function $(\exp_I t\mathrm{J}X)^\star\eta$ satisfies a useful equation.

The Hodge decomposition implies that every $X \in \mathrm{aut}(M, \mathrm{J})$ can be uniquely written as [57]

$$X = X_H + \nabla\psi^X_\omega - \mathrm{J}\nabla\psi^{\mathrm{J}X}_\omega, \tag{116}$$

where ∇ is the gradient with respect to the Riemannian metric associated to J and ω, and X_H is the g_ω-Riemannian dual of a g_ω-harmonic 1-form.

By (116) and the fact that $X \in \mathrm{isom}(M, g_\eta)$ (here g_η denotes the Riemannian metric associated to J and η) it follows that

$$\mathrm{J}X = \nabla\psi^{\mathrm{J}X}_\eta \tag{117}$$

is a gradient (with respect to g_η) vector field [69, Theorem 3.5]. We set

$$\omega_{\varphi(t)} := \omega(t) = \exp_I t\mathrm{J}X.\eta.$$

Thus,

$$\dot\omega(t) = \frac{d}{dt}\exp_I t\mathrm{J}X.\eta = \mathcal{L}_{\mathrm{J}X}\eta \circ \exp_I t\mathrm{J}X = \sqrt{-1}\partial\bar\partial\psi^{\mathrm{J}X}_\eta \circ \exp_I t\mathrm{J}X, \tag{118}$$

and

$$\begin{aligned}\ddot\omega(t) &= \sqrt{-1}\partial\bar\partial\big((\mathrm{J}X)(\psi^{\mathrm{J}X}_\eta)\big) \circ \exp_I t\mathrm{J}X = \sqrt{-1}\partial\bar\partial\big(d\psi^{\mathrm{J}X}_\eta(\mathrm{J}X)\big) \circ \exp_I t\mathrm{J}X\\ &= \sqrt{-1}\partial\bar\partial|\nabla\psi^{\mathrm{J}X}_\eta|^2 \circ \exp_I t\mathrm{J}X,\end{aligned}$$

since the η-Riemannian dual of $d\psi^{\mathrm{J}X}_\eta$ is $\nabla\psi^{\mathrm{J}X}_\eta$. Thus,

$$\ddot\varphi(t) - |\nabla\dot\varphi(t)|^2_{\omega_{\varphi(t)}} = 0.$$

Next, we can generalize this computation slightly to obtain an equation for the function $(\exp_I \mathrm{J}((1-t)Y + tZ))^\star\eta$. By (116) and the fact that $X \in \mathrm{isom}(M, g_\eta)$ (here g_η denotes the Riemannian metric associated to J and η) it follows that

$$\mathrm{J}(Z - Y) = \nabla\psi^{\mathrm{J}(Z-Y)}_\eta \tag{119}$$

is a gradient (with respect to g_η) vector field [69, Theorem 3.5]. We set

$$\omega_{\varphi(t)} := \omega(t) = (\exp_I \mathrm{J}((1-t)Y + tZ))^\star\eta.$$

Thus,

$$\begin{aligned}\dot\omega(t) &= \frac{d}{dt}(\exp_I \mathrm{J}((1-t)Y + tZ))^\star\eta\\ &= \mathcal{L}_{\mathrm{J}(Z-Y)}\eta \circ \exp_I \mathrm{J}((1-t)Y + tZ)\\ &= \sqrt{-1}\partial\bar\partial\psi^{\mathrm{J}(Z-Y)}_\eta \circ \exp_I \mathrm{J}((1-t)Y + tZ),\end{aligned} \tag{120}$$

and

$$\begin{aligned}\ddot\omega(t) &= \sqrt{-1}\partial\bar\partial\big((\mathrm{J}(Z-Y))(\psi^{\mathrm{J}(Z-Y)}_\eta)\big) \circ \exp_I t\mathrm{J}(Z-Y)\\ &= \sqrt{-1}\partial\bar\partial|\nabla\psi^{\mathrm{J}(Z-Y)}_\eta|^2 \circ \exp_I \mathrm{J}((1-t)Y + tZ),\end{aligned}$$

Thus, again,

$$\ddot{\varphi}(t) - |\nabla\dot{\varphi}(t)|^2_{\omega_{\varphi(t)}} = 0. \tag{121}$$

Observe that

$$F_\eta((1-t)Y + tZ) = (I - J)\big((\exp_I \mathrm{J}((1-t)Y + tZ))^\star \eta\big).$$

Therefore,

$$\begin{aligned} \frac{d}{dt}\Big|_0 F_\eta((1-t)Y + tZ) &= -V^{-1}\int \varphi(t)\Delta_{g_t^\star\eta}\dot{\varphi}(t)(g_t^\star\eta)^n \\ &= -V^{-1}\int \varphi(t)\Delta_{\omega(t)}\dot{\varphi}(t)\omega(t)^n \\ &= -V^{-1}\int \dot{\varphi}(t)\Delta_{\omega(t)}\varphi(t)\omega(t)^n \\ &= V^{-1}\int \dot{\varphi}(t)n(\omega - \omega(t))\wedge\omega(t)^{n-1}, \end{aligned} \tag{122}$$

where $g_t := \exp_I \mathrm{J}((1-t)Y + tZ)$, since $g_t^\star\eta = \omega(t)$. Also, using (121),

$$\begin{aligned} &\frac{d^2}{dt^2}\Big|_0 F_\eta((1-t)Y + tZ) \\ &\quad = V^{-1}\int \ddot{\varphi}(t)n(\omega - \omega(t))\wedge\omega(t)^{n-1} \\ &\qquad - V^{-1}\int \dot{\varphi}(t)n\sqrt{-1}\partial\bar{\partial}\dot{\varphi}(t)\wedge\omega(t)^{n-1} \\ &\qquad + V^{-1}\int \dot{\varphi}(t)n(n-1)(\omega - \omega(t))\wedge\sqrt{-1}\partial\bar{\partial}\dot{\varphi}(t)\wedge\omega(t)^{n-2} \\ &\quad = nV^{-1}\int |\nabla\dot{\varphi}|^2\omega\wedge\omega(t)^{n-1} - nV^{-1}\int |\nabla\dot{\varphi}|^2\omega(t)^n \\ &\qquad + nV^{-1}\int \sqrt{-1}\partial\dot{\varphi}(t)\wedge\sqrt{-1}\bar{\partial}\dot{\varphi}(t)\wedge\omega(t)^{n-1} \\ &\qquad + V^{-1}\int \dot{\varphi}(t)n(n-1)\omega\wedge\sqrt{-1}\partial\bar{\partial}\dot{\varphi}(t)\wedge\omega(t)^{n-2} \\ &\qquad - V^{-1}\int \dot{\varphi}(t)n(n-1)\sqrt{-1}\partial\bar{\partial}\dot{\varphi}(t)\wedge\omega(t)^{n-1} \\ &\quad = nV^{-1}\int |\nabla\dot{\varphi}|^2\omega\wedge\omega(t)^{n-1} - nV^{-1}\int |\nabla\dot{\varphi}|^2\omega(t)^n \\ &\qquad + V^{-1}\int |\nabla\dot{\varphi}|^2\omega(t)^n \\ &\qquad + V^{-1}\int \dot{\varphi}(t)n(n-1)\omega\wedge\sqrt{-1}\partial\bar{\partial}\dot{\varphi}(t)\wedge\omega(t)^{n-2} \\ &\qquad + (n-1)V^{-1}\int |\nabla\dot{\varphi}|^2\omega(t)^n \end{aligned}$$

$$\begin{aligned}
&= nV^{-1}\int |\nabla\dot\varphi|^2\omega\wedge\omega(t)^{n-1}\\
&\quad - n(n-1)V^{-1}\int \sqrt{-1}\partial\dot\varphi(t)\wedge\sqrt{-1}\bar\partial\dot\varphi(t)\wedge\omega\wedge\omega(t)^{n-2}\\
&= \frac{n}{V}\int \Big(|\nabla\dot\varphi|^2\omega(t)-(n-1)\sqrt{-1}\partial\dot\varphi(t)\wedge\sqrt{-1}\bar\partial\dot\varphi(t)\Big)\wedge\omega\wedge\omega(t)^{n-2}\\
&\geq \frac{n}{V}\int \sqrt{-1}\partial\dot\varphi(t)\wedge\sqrt{-1}\bar\partial\dot\varphi(t)\wedge\omega\wedge\omega(t)^{n-2}>0, \qquad (123)
\end{aligned}$$

since if α,β are two positive (1,1)-forms then $(\mathrm{tr}_\alpha\beta)\alpha-\beta\geq 0$, in general, so have $|\nabla\dot\varphi|^2\omega(t)-n\sqrt{-1}\partial\dot\varphi(t)\wedge\sqrt{-1}\bar\partial\dot\varphi(t)\geq 0$. Thus, F_η is strictly convex on the vector space $\mathrm{isom}(M,g)$. Now, observe that it is a proper function by Lemma 6.2. Since a proper strictly convex function attains a unique minimum, the proof is complete. □

Exercise 16.5. Prove the formula (see, e.g., [80, p. 140])

$$(I-J)(\omega,\eta)=J(\eta,\omega),$$

where $(I-J)(\omega,\eta)$ is just $(I-J)(\varphi)$ for any φ such that $\eta=\omega_\varphi$, while $J(\eta,\omega)$ is just J (recall (7)) "of" ω "with respect to" the reference η, in the sense that

$$J(\eta,\omega)=V^{-1}\int_M\varphi\eta^n-\frac{V^{-1}}{n+1}\int_M\psi\sum_{l=0}^n\eta^{n-l}\wedge\omega^l,$$

where ψ satisfies $\omega=\eta_\psi$.

Proposition 16.6. *Suppose* (M,J,η) *is Fano Kähler–Einstein. If* E *is* J_G*-proper then* (113) *holds.*

Proof. According to Lemma 16.2, the functional

$$g\mapsto(I-J)(\omega,g^\star\eta)$$

has a critical point at the identity $g=\mathrm{id}$ if $\eta=\omega-\sqrt{-1}\partial\bar\partial\varphi$ when $\varphi\in\mathcal{H}_\eta^\perp$. Now, by Exercise 16.5, this is tantamount to the functional

$$g\mapsto J(g^\star\eta,\omega) \qquad (124)$$

having a critical point at the identity $g=\mathrm{id}$ if $\eta=\omega-\sqrt{-1}\partial\bar\partial\varphi$ when $\varphi\in\mathcal{H}_\eta^\perp$.

Suppose now that indeed $\varphi\in\mathcal{H}_\eta^\perp$. Then the functional (124) has a critical point at $g=\mathrm{id}$. By Lemma 16.3, this is the unique minimum of this functional. Thus, using also $\mathrm{Aut}(M,\mathrm{J})_0$-invariance of J yields

$$J(\varphi)=:J(\eta,\eta_\varphi)=J(\eta,\omega)=\inf_{g\in G}J(g^\star\eta,\eta_\varphi)=\inf_{g\in G}J(\eta,g^\star\eta_\varphi).$$

The last expression is precisely $J_G(\varphi)$ (with respect to the reference metric η (not ω!)). By assumption E is J_G-proper, so, say, for concreteness,

$$E(\varphi)\geq CJ_G(\eta)-D=CJ(\varphi)-D,$$

as desired. (Observe that the proof also gives the converse, namely that if (113) holds then E is J_G-proper.) □

Therefore, Theorem 12.4 and Proposition 16.6 confirm Tian's conjecture.

Corollary 16.7. *Conjecture* 1.3 *holds.*

Acknowledgment

Thanks go to T. Darvas for many helpful discussions and a stimulating collaboration on many of the results presented here; to B. Clarke for a fruitful collaboration on the results presented in § 8.1 that stimulated some of the later developments; to B. Berndtsson for help with Theorem 3.13; to J. Streets for organizing the Winter School as well as to the numerous students who attended it and asked stimulating questions; and to the staff at MSRI for providing an excellent research environment (supported by NSF grant DMS-1440140), where these notes were completed in January 2016. Finally, it is a real privilege to dedicate this article to Gang Tian whose theorems, conjectures and vision have shaped so much of this field, and whose guidance and encouragement have been invaluable in my own pursuits over the years. This research was also supported by BSF grant 2012236, a Sloan Research Fellowship, and NSF grants DMS-1206284,1515703.

References

[1] T. Aubin, Métriques Riemanniennes et courbure, J. Differential Geom. 4 (1970), 383–424.

[2] _____, Équations du type Monge–Ampère sur les variétés kählériennes compactes, C. Rendus Acad. Sci. Paris 283 (1976), 119–121.

[3] _____, Équations du type Monge–Ampère sur les variétés kählériennes compactes, Bull. Sci. Math. 102 (1978), 63–95.

[4] T. Aubin, Nonlinear analysis on manifolds. Monge–Ampère equations, Springer, 1982.

[5] _____, Réduction du cas positif de l'équation de Monge–Ampère sur les variétés kählériennes compactes à la démonstration d'une inégalité, J. Funct. Anal. 57 (1984), 143–153.

[6] _____, Some nonlinear problems in Riemannian Geometry, Springer, 1998.

[7] R. Bamler, Cheeger–Colding–Naber theory, preliminary draft, 2016.

[8] S. Bando, R. Kobayashi, Ricci-flat Kähler metrics on affine algebraic manifolds, Lecture notes in Math. 1339 (1988), 20–31.

[9] S. Bando, T. Mabuchi, Uniqueness of Kähler–Einstein metrics modulo connected group actions, in: Algebraic Geometry, Sendai, 1985, Advanced Studies in Pure Mathematics 10, 1987, pp. 11–40.

[10] V.V. Batyrev, E.N. Selivanova, Einstein–Kähler metrics on symmetric toric Fano manifolds, Crelle's J. 512 (1999), 225–236.

[11] E. Bedford, B.A. Tayor, A new capacity for plurisubharmonic functions, Acta Math. 149 (1982), 1–40.

[12] R.J. Berman, A thermodynamical formalism for Monge–Ampère equations, Moser–Trudinger inequalities and Kähler–Einstein metrics. Adv. Math. 248 (2013), 1254–1297.

[13] _____, From Monge–Ampère equations to envelopes and geodesic rays in the zero temperature limit, preprint, arxiv:1307.3008.

[14] R.J. Berman, R. Berndtsson, Convexity of the K-energy on the space of Kähler metrics, preprint, arxiv:1405.0401.

[15] R.J. Berman, S. Boucksom, Growth of balls of holomorphic sections and energy at equilibrium, Invent. Math. 181 (2010), 337–394.

[16] J.-P. Bourguignon, Invariants intégraux fonctionnels pour des équations aux dérivées partielles d'origine géométrique, Lecture Notes in Math. 1209 (1986), 100–108.

[17] R.J. Berman, S. Boucksom, V. Guedj, A. Zeriahi, A variational approach to complex Monge–Ampère equations, Publ. Math. Inst. Hautes Études Sci. 117 (2013), 179–245.

[18] R.J. Berman, T. Darvas, C.H. Lu, Convexity of the extended K-energy and the large time behaviour of the weak Calabi flow, preprint, arxiv:1510.01260.

[19] Z. Błocki, Interior regularity of the complex Monge–Ampère equation in convex domains, Duke Math. J. 105 (2000), 167–181.

[20] Z. Błocki, The complex Monge–Ampère equation on compact Kähler manifolds, Lecture notes, February 2007.

[21] S. Boucksom, P. Eyssidieux, V. Guedj, A. Zeriahi, Monge–Ampère equations in big cohomology classes, Acta Math. 205 (2010), 199–262.

[22] B. Berndtsson, A Brunn–Minkowski type inequality for Fano manifolds and some uniqueness theorems in Kähler geometry, Invent. math. 200 (2015), 149–200.

[23] A.L. Besse, Einstein manifolds, Springer, 1987.

[24] Z. Błocki, On geodesics in the space of Kähler metrics, in: Advances in Geometric Analysis (S. Janeczko et al., Eds.), International Press, 2012, pp. 3–20.

[25] S. Boucksom, R. Berman, P. Eyssidieux, V. Guedj, A. Zeriahi, Kähler–Einstein metrics and the Kähler–Ricci flow on log Fano varieties, preprint, arxiv:1111.7158.

[26] J.-P. Bourguignon, Invariants intégraux fonctionnels pour des équations aux dérivées partielles d'origine géométrique, Lecture Notes in Mathematics 1209, Springer, 1986, pp. 100–108.

[27] D. Bump, Lie groups, Springer, 2004.

[28] E. Calabi, The variation of Kähler metrics. I. The structure of the space; II. A minimum problem, Bull. Amer. Math. Soc. 60 (1954), 167–168.

[29] _____, Extremal metrics. II, in: Differential geometry and complex analysis (I. Chavel et al., Eds.), Springer, 1985, pp. 95–114.

[30] S. Calamai, The Calabi's metric for the space of Kähler metrics, preprint, arxiv:1004.5482.

[31] I.A. Cheltsov, Y.A. Rubinstein, Asymptotically log Fano varieties, Adv. Math. 285 (2015), 1241–1300.

[32] X.X. Chen, On the lower bound of the Mabuchi energy and its application, Int. Math. Res. Notices (2000), 607–623.

[33] X.X. Chen, The space of Kähler metrics, J. Differential Geom. 56 (2000), 189–234.

[34] X.X. Chen, J. Cheng, On the constant scalar curvature Kähler metrics, II, existence results, preprint, arxiv:1801.00656.

[35] X.-X. Chen, S.K. Donaldson, S. Sun, Kähler–Einstein metrics on Fano manifolds, J. Amer. Math. Soc. 28 (2015), 183–278.

[36] X.-X. Chen, G. Tian, Ricci flow on Kähler–Einstein surfaces, Inv. Math. 147 (2002), 487–544.

[37] B. Clarke, The completion of the manifold of Riemannian metrics, J. Different. Geom. 93 (2013), 203–268.

[38] B. Clarke, Y.A. Rubinstein, Ricci flow and the metric completion of the space of Kähler metrics, Amer. J. Math. 135 (2013), 1477–1505.

[39] T. Darvas, The Mabuchi completion of the space of Kähler potentials, preprint, arxiv:1401.7318.

[40] ______, The Mabuchi geometry of finite energy classes, preprint, Adv. Math. 285 (2015), 182–219.

[41] ______, Comparison of the Calabi and Mabuchi geometries and applications to geometric flows, 2015.

[42] ______, Geometric pluripotential theory on Kähler manifolds, preprint on the author's webpage.

[43] T. Darvas, W. He, Geodesic rays and Kähler–Ricci trajectories on Fano manifolds, preprint, arxiv:1411.0774

[44] T. Darvas, Y.A. Rubinstein, Kiselman's principle, the Dirichlet problem for the Monge–Ampère equation, and rooftop obstacle problems, J. Math. Soc. Japan 68 (2016), 773–796.

[45] ______, Tian's properness conjectures and Finsler geometry of the space of Kähler metrics, J. Amer. Math. Soc. 30 (2017), 347–387.

[46] J.P. Demailly, Complex Analytic and Differential Geometry, https://www-fourier.ujf-grenoble.fr/∼demailly/manuscripts/agbook.pdf.

[47] R. Dervan, Alpha invariants and coercivity of the Mabuchi functional on Fano manifolds, preprint, arxiv:1412.1426.

[48] B.S. DeWitt, Quantum theory of gravity. I. The canonical theory, Phys. Rev. 160 (1967), 1113–1148.

[49] W.-Y. Ding, Remarks on the existence problem of positive Kähler–Einstein metrics, Math. Ann. 282 (1988), 463–471.

[50] W.Y. Ding, G. Tian, The generalized Moser–Trudinger inequality, in: Nonlinear Analysis and Microlocal Analysis (K.-C. Chang et al., Eds.), World Scientific, 1992, pp. 57–70.

[51] I.A. Dolgachev, Classical algebraic geometry: a modern view, Cambridge University Press, 2012.

[52] S.K. Donaldson, Symmetric spaces, Kähler geometry and Hamiltonian dynamics, in: Northern California Symplectic Geometry Seminar (Ya. Eliashberg et al., Eds.), Amer. Math. Soc., 1999, pp. 13–33.

[53] D.G. Ebin, The manifold of Riemannian metrics, in: Global analysis (S.S. Chern et al., Eds.), Proceedings of Symposia in Pure and Applied Mathematics 15 (1970), pp. 11–40.

[54] F. Ferrari, S. Klevtsov, S. Zelditch, Random Kähler Metrics, preprint, arxiv: 1107.4575.

[55] K. Fujita, Y. Odaka, On the K-stability of Fano varieties and anticanonical divisors, to appear in Tohoku Math. Journal.

[56] A. Futaki, An obstruction to the existence of Einstein Kähler metrics, Invent. Math. 73 (1983), 437–443.

[57] P. Gauduchon, Calabi's extremal metrics: an elementary introduction, manuscript, 2010.

[58] D. Gilbarg, N.S. Trudinger, Elliptic partial differential equations of second order, Reprint of the 1998 edition, Springer, 2001.

[59] V. Guedj, The metric completion of the Riemannian space of Kähler metrics, preprint, arxiv:1401.7857.

[60] V. Guedj, A. Zeriahi, The weighted Monge–Ampère energy of quasiplurisubharmonic functions, J. Funct. Anal. 250 (2007), 442–482.

[61] W. Hazod, K.H. Hofmann, H.P. Scheffler, M. Wustner, H. Zeuner, Normalizers of compact subgroups, the existence of commuting automorphisms, and applications to operator semistable measures, J. Lie Theory 8 (1998), 189–209.

[62] S. Helgason, Groups and geometric analysis, Amer. Math. Soc., 2002.

[63] J.L. Kazdan, F.W. Warner, Curvature functions for compact 2-manifolds, Ann. of Math. 99 (1974), 14–47.

[64] L. Hörmander, An introduction to complex analysis in several variables, Third edition, North-Holland, 1990.

[65] T. Jeffres, R. Mazzeo, Y.A. Rubinstein, Kähler–Einstein metrics with edge singularities, (with an appendix by C. Li and Y.A. Rubinstein), Annals of Math. 183 (2016), 95–176.

[66] E. Kähler, Über eine bemerkenswerte Hermitesche Metrik, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg 9 (1933), 173–186.

[67] Y.-C. Lu, Holomorphic mappings of complex manifolds, J. Diff. Geom. 2 (1968), 299–312.

[68] T. Mabuchi, K-energy maps integrating Futaki invariants, Tôhoku Math. J. 38 (1986), 575–593.

[69] _____, Some symplectic geometry on compact Kähler manifolds I, Osaka J. Math. 24, 1987, 227–252.

[70] Y. Matsushima, Sur la structure du groupe d'homéomorphismes analytiques d'une certaine variété kählérienne, Nagoya Math. J. 11 (1957), 145–150.

[71] D. McFeron, The Mabuchi metric and the Kähler–Ricci flow, Proc. Amer. Math. Soc. 142 (2014), 1005–1012.

[72] D.H. Phong, N. Sešum, J. Sturm, Multiplier ideal sheaves and the Kähler–Ricci flow, Comm. Anal. Geom. 15 (2007), 613–632.

[73] D.H. Phong, J. Song, J. Sturm, Complex Monge–Ampère equations, in: Survey in Differential Geometry XVII, Int. Press, 2012, pp. 327–410.

[74] D.H. Phong, J. Song, J. Sturm, B. Weinkove, The Moser–Trudinger inequality on Kähler–Einstein manifolds, Amer. J. Math. 130 (2008), 651–665.

[75] D.H. Phong, J. Sturm, Lectures on stability and constant scalar curvature, in: Current developments in mathematics 2007, Int. Press, 2009, pp. 101–176.

[76] M.M. Postnikov, Geometry VI. Riemannian geometry, Springer, 2001.

[77] T. Ransford, Potential theory in the complex plane, Cambridge University Press, 1995.

[78] Y.A. Rubinstein, On energy functionals, Kähler–Einstein metrics, and the Moser–Trudinger–Onofri neighborhood, J. Funct. Anal. 255, special issue dedicated to Paul Malliavin (2008), 2641–2660.

[79] _____, Some discretizations of geometric evolution equations and the Ricci iteration on the space of Kähler metrics, Adv. Math. 218 (2008), 1526–1565.

[80] _____, Geometric quantization and dynamical constructions on the space of Kähler metrics, Ph.D. Thesis, Massachusetts Institute of Technology, 2008.

[81] _____, Smooth and singular Kähler–Einstein metrics, in: Geometric and Spectral Analysis, (P. Albin et al., Eds.), Contemp. Math. 630, Amer. Math. Soc. and Centre de Recherches Mathématiques, 2014, pp. 45–138.

[82] Y.A. Rubinstein, S. Zelditch, The Cauchy problem for the homogeneous Monge–Ampère equation, I. Toeplitz quantization, J. Differential Geom. 90 (2012), 303–327.

[83] J.A. Schouten, Ueber unitäre Geometrie, Proceedings Koninklijke Nederlandse Akademie van Wetenschappen Amsterdam 32 (1929), 457–465.

[84] J.A. Schouten, D. van Dantzig, Ueber die Differentialgeometrie einer Hermiteschen Differentialform und ihre Beziehungen zu den Feldgleichungen der Physik, Proceedings Koninklijke Nederlandse Akademie van Wetenschappen Amsterdam 32 (1929), 60–64.

[85] _____, Über unitäre Geometrie, Mathematische Annalen 103 (1930), 319–346.

[86] S. Semmes, Complex Monge–Ampère and symplectic manifolds, Amer. J. Math. 114 (1992), 495–550.

[87] Y.-T. Siu, Lectures on Hermitian–Einstein metrics for stable bundles and Kähler–Einstein metrics, Birkhäuser, 1987.

[88] Y.-T. Siu, The existence of Kähler–Einstein metrics on manifolds with positive anticanonical line bundle and a suitable finite symmetry group, Ann. of Math. 127 (1988), 585–627.

[89] J. Song, The α-invariant on toric Fano manifolds, Amer. J. Math. 127 (2005), 1247–1259.

[90] J. Streets, Long time existence of minimizing movement solutions of Calabi flow, Adv. Math. 259 (2014) 688–729.

[91] M. Stroppel, Locally compact groups, EMS Textbooks in Mathematics, 2006.

[92] R.P. Thomas, Notes on GIT and symplectic reduction for bundles and varieties, in: Surveys in Differential Geometry: Essays in memory of S.-S. Chern (S.-T. Yau, Ed.), Int. Press, 2006, pp. 221–273.

[93] G. Tian, On Kähler–Einstein metrics on certain Kähler manifolds with $C_1(M) > 0$, Inv. Math. 89 (1987), 225–246.

[94] _____, On Calabi's conjecture for complex surfaces with positive first Chern class, Invent. Math. 101 (1990), 101–172.

[95] _____, The K-energy on hypersurfaces and stability, Comm. Anal. Geom. 2 (1994), 239–265.

[96] _____, Kähler–Einstein metrics with positive scalar curvature, Invent. Math. 130 (1997), 1–37.

[97] ______, Canonical Metrics in Kähler Geometry, Birkhäuser, 2000.

[98] ______, Existence of Einstein metrics on Fano manifolds, in: Metric and Differential Geometry (X.-Z. Dai et al., Eds.), Springer, 2012, pp. 119–159.

[99] ______, K-stability and Kähler–Einstein metrics, Comm. Pure. Appl. Math. 68 (2015), 1085–1156.

[100] G. Tian, X. Zhu, A nonlinear inequality of Moser–Trudinger type, Calc. Var. PDE 10 (2000), 349–354.

[101] G. Wanby, Subharmonic and strongly subharmonic functions in case of variable coefficients, Math. Scand. 22 (1968), 283–309.

[102] R. Wang, J. Jiang, Another approach to the Dirichlet problem for equations of Monge–Ampère type, Northeastern Math. J. 1 (1985), 27–40.

[103] X.-J. Wang, X.-H. Zhu, Kähler–Ricci solitons on toric manifolds with positive first Chern class, Adv. Math. 188 (2004), 87–103.

[104] D. Wu, Kähler–Einstein metrics of negative Ricci curvature on general quasi-projective manifolds, Comm. Anal. Geom. 16 (2008), 395–435.

[105] S.-T. Yau, On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampère equation, I, Comm. Pure Appl. Math. 31 (1978), 339–411.

[106] A. Zeriahi, Volume and capacity of sublevel sets of a Lelong class of plurisubharmonic functions, Indiana Univ. Math. J. 50 (2001), 671–703.

[107] B. Zhou, X. Zhu, Relative K-stability and modified K-energy on toric manifolds, Adv. Math. 219 (2008) 1327–1362.

Yanir A. Rubinstein
University of Maryland
e-mail: yanir@alum.mit.edu

Progress in Mathematics, Vol. 333, 445–463

Ancient Solutions in Geometric Flows

Natasa Sesum

Abstract. In this survey paper we discuss ancient solutions to different geometric flows, such as the Ricci flow, the mean curvature flow and the Yamabe flow. We survey the classification results of ancient solutions in the Ricci flow and the mean curvature flow. We also discuss methods for constructing new ancient solutions to the Yamabe flow, indicating that the classification results for this flow are impossible to expect.

Mathematics Subject Classification (2010). 53C44.

Keywords. Ancient solutions, Ricci flow, mean curvature flow.

1. Introduction

A solution to a geometric evolution equation such as the MCF, the Ricci flow, or the Yamabe flow is called *ancient* if it exists for all $t \in (-\infty, t_0]$, for some $t_0 \leq +\infty$. If $t_0 = +\infty$, the solution is called *eternal.* While solutions starting from arbitrary smooth initial data can be constructed on a short enough time interval for all these flows, the requirement that a solution should exist for *all* time $t \leq t_0$, combined with some sort of positive curvature condition, turns out to be very restrictive. In a number of cases there are results which state that the list of possible ancient solutions to some given geometric flow consists of self-similar solutions ("solitons") and a shorter list of non self similar solutions.

Definition 1.1. Assume that a geometric flow, say for example the Ricci flow, develops a singularity at time $T < \infty$. We say the singularity is Type I if $\limsup_{t\to T}(T-t)\sup_M |\mathrm{Rm}| < \infty$, otherwise we say the singularity is Type II.

We study singularities in geometric flows using the blow up analysis. In the case of a Type I singularity, the sequence of a rescaled solution around the singularity subconverges to an ancient solution. In the case of Type II singularity there exists a rescaled sequence around the singularity (see [27]) that subconverges to an eternal solution.

All geometric flows tend to develop finite time singularities and in order to use the flows to understand topological and other properties of an underlying manifold,we need to better understand singularities. As mentioned above, ancient and eternal solutions arise as singularity models in geometric flows. More precisely, ancient solutions play an important role in understanding the singularity formation in geometric flows, as such solutions are usually obtained after performing a blow up near points where the curvature is very large. In fact, Perelman's famous work on the Ricci flow [40] shows that the high curvature regions in the three-dimensional Ricci flow are modeled on ancient solutions which have nonnegative curvature and are κ-noncollapsed. Similar results for mean curvature flow were obtained in [31, 45, 46] assuming mean convexity and embeddedness. That is why understanding those solutions, and more specifically, their classification, is important for understanding singularities of geometric flows.

For instance, for two-dimensional Ricci flow, Daskalopoulos, Hamilton and Sesum [22] classified all compact ancient solutions. It turns out the complete list contains only the shrinking sphere solitons and the King–Rosenau solutions [35, 42]. The latter are not solitons and can be visualized as two steady solitons, called "cigars", coming from spatial infinities and glued together. In [18] the authors obtained the full classification of simple closed embedded convex ancient solutions to the curve shortening flow. They are either the family of contracting circles or Angenent ovals, known as well as "paper clips". The latter can be visualized as two grim repears (which are translators) glued together, moving apart from each other, as time approaches negative infinity.

The higher-dimensional analogue of the curve shortening flow is the mean curvature flow. In [4] we showed that all rotationally symmetric noncollapsed closed ancient solutions that are not self-similar have unique asymptotics as $t \to -\infty$. Using that, in a subsequent paper [5], we show that all uniformly two-convex noncollapsed ancient closed solutions to the mean curvature flow that are not self-similar are rotationally symmetric and unique up to isometries, scalings and translations in time.

The outline of this survey article is as follows. In Section 2 we survey classification results of ancient solutions in two dimensions and discuss what is known and expected in higher dimensions. In Section 3 we survey classification results of ancient solutions in the mean curvature flow. In Section 4 we survey our work on ancient solutions in the Yamabe flow.

2. Ancient solutions to the Ricci flow

We consider an ancient solution of the Ricci flow

$$\frac{\partial g_{ij}}{\partial t} = -2\, R_{ij} \tag{2.1}$$

on a compact two-dimensional surface that exists for time $t \in (-\infty, T)$ and becomes singular at $t = T$, for some $T < \infty$.

2.1. Ancient closed solutions

In two dimensions we have $R_{ij} = \frac{1}{2} R\, g_{ij}$, where R is the scalar curvature of the surface. Moreover, on an ancient non-flat solution we have $R > 0$. It is well known ([15], [26]) that the surface also becomes extinct at T and it becomes spherical, which means that after a normalization, the normalized flow converges to a spherical metric, to which we will refer as to the limiting sphere.

Since $R > 0$, by the Uniformization theorem and the fact that the Ricci flow in dimension two preserves the conformal class, we can parametrize the Ricci flow by the limiting sphere at time T, that is, we can write

$$g(\cdot, t) = u(\cdot, t)\, g_{S^2}.$$

The spherical metric can be written as

$$g_{S^2} = d\psi^2 + \cos^2 \psi\, d\theta^2 \tag{2.2}$$

where ψ, θ denote the global coordinates on the sphere. An easy computation shows that (2.1) is equivalent to the following evolution equation for the conformal factor $u(\cdot, t)$, namely

$$u_t = \Delta_{S^2} \log u - 2 \qquad \text{on } S^2 \times (-\infty, T) \tag{2.3}$$

where Δ_{S^2} denotes the Laplacian on S^2. Let us recall, for future reference, that the only nonzero Christoffel symbols for the spherical metric (2.2) are

$$\Gamma^2_{12} = \Gamma^2_{21} = -\tan \psi, \quad \Gamma^1_{22} = \frac{\sin 2\psi}{2}$$

where we use the indices $1, 2$ for the ψ, θ variables respectively. It follows that for any function f on the sphere we have

$$\Delta_{S^2} f = f_{\psi\psi} - \tan \psi\, f_\psi + \sec^2 \psi\, f_{\theta\theta}$$

which, in the case of a radially symmetric function $f = f(\psi)$, becomes

$$\Delta_{S^2} f = f_{\psi\psi} - \tan \psi\, f_\psi.$$

We will assume, in this article, that $g = u\, ds_p^2$ is an ancient solution to the Ricci flow (2.3) on the sphere which becomes extinct at time $T = 0$.

It is natural to consider the pressure function $v = u^{-1}$ which evolves by

$$v_t = v^2 \left(\Delta_{S^2} \log v + 2\right) \qquad \text{on } S^2 \times (-\infty, 0) \tag{2.4}$$

or, after expanding the Laplacian of $\log v$,

$$v_t = v\, \Delta_{S^2} v - |\nabla_{S^2} v|^2 + 2v^2 \qquad \text{on } S^2 \times (-\infty, 0). \tag{2.5}$$

Definition 2.1. We say that an ancient solution to the Ricci flow (2.1) on a compact surface M is type I, if it satisfies

$$\limsup_{t \to -\infty} \left(|t| \max_M R(\cdot, t)\right) < \infty.$$

A solution which is not of type I, is called type II.

Explicit examples of ancient solutions to the Ricci flow in two dimensions are:

1. **The contracting spheres**
   They are described on S^2 by a pressure v_S that is given by
   $$v_S(\psi, t) = \frac{1}{2(-t)} \tag{2.6}$$
   and they are examples of ancient type I shrinking Ricci solitons.
2. **The King–Rosenau solutions**
   They were discovered by J.R. King ([35]) and later, independently, by P. Rosenau ([42]). They are described on S^2 by a pressure v_K that has the form
   $$v_K(\psi, t) = a(t) - b(t)\,\sin^2\psi \tag{2.7}$$
   with $a(t) = -\mu\,\coth(2\mu t)$, $b(t) = -\mu\,\tanh(2\mu t)$, for some $\mu > 0$. These solutions are *not solitons.* We can visualize them as two cigars "glued" together to form a compact solution to the Ricci flow. They are type II ancient solutions.

In [18] we have proved the following classification result:

Theorem 2.2. *Let $g = u\,g_{S^2}$ be an ancient compact solution to the Ricci flow* (2.1)*. Then u is either one of the contracting spheres or one of the King–Rosenau solutions.*

Remark 2.3. The classification of two-dimensional, complete, non-compact ancient solutions of the Ricci flow was recently given in [19] (see also in [16, 28]). The result in Theorem 2.2 together with the results in [19] and [16] provide a complete classification of all ancient two-dimensional complete solutions to the Ricci flow, with the scalar curvature uniformly bounded at each time-slice.

In the course of proving Theorem 2.2 we first show a priori derivative estimates on any ancient solution v of (2.5), which hold uniformly in time, up to $t = -\infty$. These estimates turn out to play a crucial role throughout the proof of the Theorem. We also introduce a suitable Lyapunov functional and use it to show that the solution $v(\cdot, t)$ of (2.5) converges, as $t \to -\infty$, in the $C^{1,\alpha}$ norm, to a steady state v_∞. Next, we classify all backward limits v_∞. We show that there is a parametrization of the flow by a sphere, in which $v_\infty(\psi, \theta) = \mu\,\cos^2\psi$, for some $\mu \geq 0$ (ψ, θ are the global coordinates on S^2). When $\mu > 0$, then v_∞ represents the cylindrical metric. If $v_\infty(\psi, \theta) = \mu\,\cos^2\psi$, with $\mu > 0$, then we show v must be one of the King–Rosenau solutions. Finally, we show that if $v_\infty \equiv 0$, then the solution v must be one of the contracting spheres. Note that in proving Theorem 2.2 we relied on the fact that all our known solutions have been given in closed forms, which enabled us to construct various monotone quantities along the flow, that needed to vanish on our solutions and their backward limits.

A higher-dimensional analogue of the two-dimensional King–Rosenau solution is Perelman's ancient noncollapsed ancient solution. This is the rotationally symmetric ancient κ-noncollapsed solution on S^3 constructed in [41]. For the definition of κ-noncollapsed solutions to the Ricci flow see [40]. It can be showed that the backward asymptotic gradient shrinking Ricci soliton for this solution is a round cylinder $S^2 \times \mathbb{R}$. Note that Perelman's ancient noncollapsed solution is type II (backward in time), since $\sup_{M\times(-\infty,0]} |t||R(x,t)| = \infty$, whereas it forms a type I singularity (forward in time), as it shrinks to a round point. Perelman's ancient solution has backward in time limits which are the Bryant soliton and the round cylinder $S^2 \times \mathbb{R}$, depending on how the sequence of points and times about which one rescales are chosen. These are the only backward in time limits of Perelman's ancient solution.

Roughly speaking, Perelman's ancient solution is constructed as follows. Let $S^2(r)$ denote the round 2-sphere of radius r. For any $L \in (1,\infty)$ we construct a rotationally symmetric metric $g_L(0)$ on S^2 with weakly positive curvature operator which metrically looks like a long round cylinder $S^2(\sqrt{2}) \times [-L, L]$ with two caps B^2_+ and B^2_- smoothly attached to the boundary components $S^2(\sqrt{2}) \times \{-L, L\}$. Perelman's ancient solution is obtained by taking rescaled and time translated limit, as $L \to \infty$, of the solutions of the Ricci flow with initial metrics $g_L(0)$. There is some work involved in showing that this limit exists.

Conjecture 2.4. *The only closed three-dimensional κ-noncollapsed ancient solutions to the Ricci flow are either the family of contracting spheres or Perelman's solution.*

The same construction as above leads to the existence of Perelman's solutions in dimensions $n \geq 3$ as well. Perelman believes (see Section 1.3 of [41]) the analogous conjecture to Conjecture 2.4 holds in dimensions $n > 3$ as well, under the additional assumption of positivity of curvature operator.

2.2. Complete ancient solutions

Self-similar solutions play an important role in the study of the Ricci flow and have been extensively studied in connection with singularity formation ([25, 40, 41]). There are three types of self-similar solutions, which are referred to as shrinking solitons, steady solitons and expanding solitons. Shrinking solitons are special examples of ancient solutions. Spheres are a typical example of closed Ricci shrinkers and cylinders are examples of complete noncompact Ricci shrinkers. Steady solitons are special examples of eternal solutions. A steady Ricci soliton (M, g) is characterized by the fact that $2\mathrm{Ric} = \mathcal{L}_X(g)$ for some vector field X. If the vector field X is the gradient of a function, we say (M, g) is a steady gradient Ricci soliton.

The simplest example of a steady Ricci soliton is the cigar soliton in dimension two, which was found by Hamilton ([25]). Bryant ([14]) discovered a steady Ricci soliton in dimension three, which is the unique rotationally symmetric complete steady Ricci soliton in dimension three. Brendle in [8] showed that any three-dimensional complete steady gradient Ricci soliton which is non-flat and κ-noncollapsed must be rotationally symmetric and hence isometric to the Bryant soliton up to scaling. In [9] Brendle shows that if (M, g) is a steady gradient Ricci soliton of dimension $n \geq 4$ which has positive sectional curvature and is asymptotically cylindrical, then (M, g) is rotationally symmetric. In particular, (M, g) is isometric to the n-dimensional Bryant soliton up to scaling. In a recent paper [7], Brendle classifies all three-dimensional ancient complete noncompact noncollapsed Ricci flow solutions. He shows those solutions must be rotationally symmetric and then using rotational symmetry he shows they need to be steady solitons.

Conjecture 2.5. *Every complete three-dimensional κ-noncollapsed eternal solution to the Ricci flow has to be steady soliton.*

If Conjecture 2.4 and Conjecture 2.5 were completed, that combined with Brendle's result ([8]) would give us complete classification of three-dimensional κ-noncollapsed ancient solutions to the Ricci flow.

3. Ancient solutions to the Mean Curvature Flow

We consider now the Mean Curvature Flow. Recall that a family of immersed hypersurfaces $\boldsymbol{X} : M^n \times [0, T) \to \mathbb{R}^{n+1}$ evolves by Mean Curvature Flow (MCF) if it satisfies

$$\left(\frac{\partial \boldsymbol{X}}{\partial t}\right)^{\perp} = H\boldsymbol{\nu}, \tag{3.1}$$

where $\boldsymbol{\nu}$ is a unit normal vector of the surface $M_t = \boldsymbol{X}(M^n, t)$, H is the mean curvature in the direction of the normal $\boldsymbol{\nu}$, and $\big(\boldsymbol{X}_t(\xi, t)\big)^{\perp}$ is the component of the velocity $\boldsymbol{X}_t(\xi, t)$ that is perpendicular to M_t at $\boldsymbol{X}(\xi, t)$.

A smooth solution $\{M_t\}_{0 \leq t < T}$ to MCF exists on a sufficiently short time interval $0 \leq t < T$ for any prescribed smooth initial immersed hypersurface M_0. If the initial hypersurface M_0 is convex, then the solution M_t will also be convex. The simplest possible convex ancient solution is the shrinking sphere, i.e., if M_t is the sphere of radius $\sqrt{-2nt}$ centered at the origin, then $\{M_t\}_{t<0}$ is a self similar ancient solution. It is the only compact and convex self-similar solution to MCF. In the text below we will give the notion of a *non-collapsed* solution to MCF, which was introduced by B. Andrews in [1]. With this in mind we give the following definition.

Definition 3.1. An ancient oval is any ancient compact non-collapsed (in the sense of Definition 3.4) solution to MCF that is not self similar (i.e., that is not the sphere).

Note that the "non-collapsedness" condition from Definition 3.4 is necessary due to other "pancake" type examples which become collapsed as $t \to -\infty$ (see [6]) and which we will discuss more later. On the other hand, it has been shown in [31] that all non-collapsed ancient compact solutions to the mean curvature flow are convex, hence the ancient ovals are convex solutions.

3.1. Curve shortening flow

For Curve Shortening, i.e., MCF for curves in the plane, Angenent found such solutions (see [3] and also [39]). These solutions, which can be written in closed form, may be visualized as two "Grim Reapers" with the same asymptotes that approach each other from opposite ends of the plane. Daskalopoulos, Hamilton, and Sesum [18] classified all ancient convex solutions to Curve Shortening by showing that there are no other ancient ovals for Curve Shortening.

More precisely, consider an ancient embedded solution $\Gamma_t \subset \mathbb{R}^2$ of the curve shortening flow

$$\frac{\partial \mathbf{X}}{\partial t} = -\kappa \, \mathbf{N} \tag{3.2}$$

which moves each point $\mathbf{X}$ on the curve Γ_t in the direction of the inner normal vector $\mathbf{N}$ to the curve at P by a speed which is equal to the curvature κ of the curve.

Let $\Gamma_t \subset \mathbb{R}^2$ be an embedded ancient solution to the curve shortening flow (3.2). If s is the arclength along the curve and $\mathbf{X} = (x, y)$ we can express (3.2) as a system

$$\frac{\partial x}{\partial t} = \frac{\partial^2 x}{\partial s^2}, \qquad \frac{\partial y}{\partial t} = \frac{\partial^2 y}{\partial s^2}.$$

The evolution for the curvature κ of Γ_t is given by

$$\kappa_t = \kappa_{ss} + \kappa^3 \tag{3.3}$$

which is a strictly parabolic equation. Let θ be the angle between the tangent vector and the x axis. For convex curves we can use the angle θ as a parameter. It has been computed that

$$\kappa_t = \kappa^2 \, \kappa_{\theta\theta} + \kappa^3. \tag{3.4}$$

It turns out that the evolution of the family Γ_t is completely described by the evolution (3.4) of the curvature κ.

We will assume that Γ_t is an ancient solution of the curve shortening flow defined on $(-\infty, T)$. We will also assume that our extinction time $T = 0$.

It is natural to consider the *pressure* function

$$p := \kappa^2$$

which evolves by

$$p_t = p \, p_{\theta\theta} - \frac{1}{2} \, p_\theta^2 + 2 \, p^2. \tag{3.5}$$

In accordance with Definition 2.1 we say that an ancient solution to (3.4) is:

- of type I: if it satisfies $\sup_{t \in (-\infty, -1]} \sup_{\Gamma_t} |t||p(x,t)| < \infty$.
- of type II: if $\sup_{t \in (-\infty, -1]} \sup_{\Gamma_t} |t||p(x,t)| = \infty$.

The ancient solution to (3.4) defined by

$$p(\theta, t) = \frac{1}{2(-t)}$$

corresponds to a family of *contracting circles*. This solution is of type I and at the same time falls in a category of contracting self-similar solutions (these are solutions of the flow whose shapes change homothetically during the evolution). We will show in the next section the existence of compact ancient solutions to (3.4) that are not self-similar. Since they have been discovered by Angenent we will refer to them as to the *Angenent ovals*.

One very nice and important property of ancient solutions to the curve shortening flow is that $\kappa_t \geq 0$. This fact follows from Hamilton's Harnack estimate for convex curves ([27]). By the strong maximum principle, $\kappa(\cdot, t) > 0$ for all $t < 0$. If we start at any time $t_0 \leq 0$, Hamilton proved that

$$\kappa_t + \frac{\kappa}{2(t - t_0)} - \frac{\kappa_s^2}{k} \geq 0. \tag{3.6}$$

Letting $t_0 \to -\infty$ we get

$$\kappa_t \geq 0. \tag{3.7}$$

In [18] we provide the following classification of ancient convex solutions to the curve shottening flow.

Theorem 3.2 (Daskalopulos, Hamilton, Sesum). *Let $p(\theta, t) = \kappa^2(\theta, t)$ be an ancient solution to* (3.5)*, defining a family of embedded closed convex curves in $\mathbb{R}^2$ that evolve by the curve shortening flow. Then,*

(i) *either $p(\theta, t) = \frac{1}{(-2t)}$, which corresponds to contracting circles, or*

(ii) *$p(\theta, t) = \lambda(\frac{1}{1-e^{2\lambda t}} - \sin^2(\theta + \gamma))$, for two parameters $\lambda > 0$ and γ, which corresponds to the Angenent ovals.*

In the proof of Theorem 3.2 we found a monotone integral quantity along the flow whose limit as $t \to T$ and the limit as $t \to -\infty$ have been vanishing, forcing the quantity and its derivative to be identically zero along the flow. After that,a simple ODE argument concludes the proof of the theorem. The construction of our monotone integral quantity relied on the fact we knew the explicit formulas for all possible solutions.

3.2. Closed ancient solutions to the MCF

Natural questions to ask are whether there exists an analog of the Ancient Curve Shortening Ovals from [3, 39] in higher-dimensional Mean Curvature Flow and whether a classification of ancient ovals similar to the Daskalopoulos–Hamilton–Sesum [18] result is possible.

The existence question was already settled by White in [45] who gave a construction of ancient ovals for which

$$\frac{\text{in-radius } M_t}{\text{out-radius } M_t} \to 0 \quad \text{as } t \to -\infty.$$

Haslhofer and Hershkovits [32] provided recently more details on White's construction. If one represents $\mathbb{R}^{n+1}$ as $\mathbb{R}^{n+1} = \mathbb{R}^k \times \mathbb{R}^l$ with $k + l = n + 1$, then the White–Haslhofer–Hershkovits construction proves the existence of an ancient solution M_t with $O(k) \times O(l)$ symmetry. The construction of those solutions is similar to the construction of Perelman's closed ancient solutions for the Ricci flow. More precisely, they considered convex regions of increasing eccentricity and using limiting arguments, they proved the existence of ancient flows of compact, convex sets that are not self-similar. In contrast with the Ancient Curve Shortening Ovals, these solutions cannot be written in closed forms. Formal matched asymptotics, as $t \to -\infty$, were given by Angenent in [2].

The classification question is more complicated in higher dimensions.

Conjecture 3.3 (Uniqueness of ancient ovals). *For each (k, l) with $k + l = n + 1$ there is only one "ancient oval" solution with $O(k) \times O(l)$ symmetry, up to time translation and parabolic rescaling of space-time.*

Since the "ancient oval" solutions are not given in closed form and they are not solitons their classification as stated in the above conjecture poses a difficult question. In fact, up to now the only known classification results for ancient or eternal solutions involve either solitons or other special solutions that can be written in closed form.

In [4] we have made a partial progress towards the above conjecture by showing that any ancient, closed non-collapsed solution of MCF with $O(1) \times O(n)$ symmetry satisfies the *detailed asymptotic expansions* described in [2]. In particular, our results in that paper give precise estimates on the *extrinsic diameter* and *maximum curvature* of all such solutions near $t \to -\infty$.

Instead of an evolving family of convex hypersurfaces $\{M_t\}$ we can also think in terms of the evolving family $\{K_t\}$ of compact domains enclosed by M_t (thus $M_t = \partial K_t$). Andrews [1] introduced the following notion of "non-collapsedness" for any compact mean convex subset $K \subset \mathbb{R}^{n+1}$. Recall that a domain $K \subset \mathbb{R}^{n+1}$ with smooth boundary is mean convex if $H > 0$ on ∂K.

Definition 3.4. If $K \subset \mathbb{R}^{n+1}$ is a smooth, compact, mean convex domain and if $\alpha > 0$, then K is **α-noncollapsed** if for every $p \in \partial K$ there are closed balls $\bar{B}_{\text{int}} \subset K$ and $\bar{B}_{\text{ext}} \subset \mathbb{R}^{n+1} \backslash \text{Int}(K)$ of radius at least $\frac{\alpha}{H(p)}$ that are tangent to ∂K at p from the interior and exterior of K, respectively (in the limiting case $H(p) = 0$ this means that K is a half-space).

Every compact, smooth, strictly mean convex domain is α-noncollapsed for some $\alpha > 0$. Andrews showed that if the initial condition K_0 of a smooth compact mean curvature flow is α-noncollapsed, then so is the whole flow K_t for all later times t.

Definition 3.5. We say that a mean convex ancient solution $\{M_t\}_{t \in (-\infty, T]}$ to MCF is **noncollapsed** if there exists a constant $\alpha > 0$ so that the flow M_t is α-noncollapsed for all $t \in (-\infty, T]$, in the sense of Definition 3.4.

In order to say more about the classification of closed ancient noncollapsed solutions to the mean curvature flow, we need to understand first the geometry of those solutions and their more precise asymptotics. We first focus on hypersurfaces with $O(1) \times O(n)$ symmetry. More precisely, we first consider noncollapsed and therefore convex ancient solutions that are $O(1) \times O(n)$-invariant hypersurfaces in $\mathbb{R}^{n+1}$. Such hypersurfaces can be represented as

$$M_t = \big\{(x, x') \in \mathbb{R} \times \mathbb{R}^n : -d(t) < x < d(t), \|x'\| = U(x,t)\big\} \tag{3.8}$$

for some function $\|x'\| = U(x,t)$. The points $(\pm d(t), 0)$ are called *the tips* of the surface. The function $U(x,t)$, which we call the *profile* of the hypersurface M_t, is only defined for $x \in [-d(t), d(t)]$.

Any surface M_t defined by (3.8) is automatically invariant under $O(n)$ acting on $\mathbb{R} \times \mathbb{R}^n$. The surface will also be invariant under the $O(1)$ action on $\mathbb{R} \times \mathbb{R}^n$ if U is even, i.e., if $U(-x,t) = U(x,t)$.

Convexity of the surface M_t is equivalent to concavity of the profile U, i.e., M_t is convex if and only if $U_{xx} \le 0$.

For a family of surfaces defined by $\|x'\| = U(x,t)$, equation (3.1) for MCF holds if and only if the profile $U(x,t)$ satisfies the evolution equation

$$\frac{\partial U}{\partial t} = \frac{U_{xx}}{1 + U_x^2} - \frac{n-1}{U}. \tag{3.9}$$

We know by Huisken's result ([34]) that the surfaces M_t will contract to a point in finite time.

Self-similar solutions to MCF are of the form $M_t = \sqrt{T-t}\,\bar{M}$ for some fixed surface $\bar{M}$ and some "blow-up time" T. We rewrite a general ancient solution $\{M_t : t < t_0\}$ as

$$M_t = \sqrt{T-t}\,\bar{M}_{-\log(T-t)}. \tag{3.10}$$

The family of surfaces $\bar{M}_\tau$ with $\tau = -\log(T-t)$, is called a type-I or *parabolic blow-up* of the original solution M_t. These are again $O(1) \times O(n)$ symmetric with profile function u, which is related to U by

$$U(x,t) = \sqrt{T-t}\,u(y,\tau), \qquad y = \frac{x}{\sqrt{T-t}}, \quad \tau = -\log(T-t). \tag{3.11}$$

If the M_t satisfy MCF, then the hypersurfaces $\bar{M}_\tau$ evolve by the *rescaled MCF*

$$\boldsymbol{\nu} \cdot \frac{\partial \boldsymbol{X}}{\partial \tau} = H + \tfrac{1}{2}\boldsymbol{X} \cdot \boldsymbol{\nu}. \tag{3.12}$$

For the parabolic blow-up u this is equivalent with the equation

$$\frac{\partial u}{\partial \tau} = \frac{u_{yy}}{1 + u_y^2} - \frac{y}{2}\,u_y - \frac{n-1}{u} + \frac{u}{2}. \tag{3.13}$$

Regarding notation, we denote by $H(\cdot,t)$, $d(t)$, etc., the mean curvature and extrinsic diameter of the surface M_t, respectively, and by $\bar{H}(\cdot,\tau)$, $\bar{d}(\tau)$, etc., the mean curvature and extrinsic diameter of a corresponding parabolic blow-up $\bar{M}_\tau$, respectively. In general, we will use the bar to denote geometric quantities for $\bar{M}_\tau$.

The following theorem, which is shown in [4], describes certain geometric properties of the ancient solutions described above.

Theorem 3.6. *Let $\{M_t\}$ be any compact smooth noncollapsed ancient mean curvature flow with $O(1)\times O(n)$ symmetry. Then there exist uniform constants $c, C > 0$ so that the extrinsic diameter $\bar{d}(\tau)$, the area $\bar{\mathcal{A}}(\tau)$ and the maximum mean curvature $\bar{H}_{\max}(\tau)$ of the rescaled mean curvature flow $\bar{M}_\tau$ satisfy*

$$c\,\bar{d}(\tau) \le \bar{H}_{\max}(\tau) \le \bar{d}(\tau), \qquad c\,\bar{d}(\tau) \le \bar{\mathcal{A}}(\tau) \le C\,\bar{d}(\tau). \tag{3.14}$$

Corollary 6.3 in [44] implies that the dilations $\{\bar{X} \in \mathbb{R}^{n+1} \mid (-t)^{1/2}\bar{X} \in M_t\}$, of hypersurfaces M_t which evolve by (3.1) and which satisfy conditions of Theorem 3.6 that *sweep out the whole space*, converge as $t \to -\infty$

(a) to either a sphere of radius $\sqrt{2n}$ or
(b) a cylinder $S^{n-1} \times \mathbb{R}$, where S^{n-1} is a sphere of radius $\sqrt{2(n-1)}$.

In the present paper we show that any compact convex ancient solution to (3.1) as in Theorem 3.6 has unique asymptotics as $t \to -\infty$. The hope has been to use that to prove Conjecture 3.3. More precisely, we show that the following holds.

Theorem 3.7. *Let $\{M_t\}$ be any compact smooth noncollapsed ancient mean curvature flow as in Theorem 3.6. Then, either M_t is a family of contracting spheres or the solution $u(y,\tau)$ to (3.13), defined on $\mathbb{R} \times \mathbb{R}$, has the following asymptotics in the parabolic and the intermediate region:*

(i) **Parabolic region:** *For every $M > 0$,*

$$u(y,\tau) = \sqrt{2(n-1)}\Big(1 - \frac{y^2-2}{4|\tau|}\Big) + o(|\tau|^{-1}), \qquad |y| \le M$$

as $\tau \to -\infty$.

(ii) **Intermediate region:** *Define $z := y/\sqrt{|\tau|}$ and $\bar{u}(z,\tau) := u(z\sqrt{|\tau|},\tau)$. Then, $\bar{u}(z,\tau)$ converges, as $\tau \to -\infty$ and uniformly on compact subsets in z, to the function $\sqrt{2-z^2}$.*

(iii) **Tip region:** *Denote by p_t the tip of $M_t \subset \mathbb{R}^{n+1}$, and for any $t_* < 0$ we define the rescaled flow*

$$\tilde{M}_{t_*}(t) = \lambda(t_*)\big(M_{t_*+t\lambda(t_*)^{-2}} - p_{t_*}\big)$$

where $\lambda(t) := H(p_t,t) = H_{\max}(t)$. Then, as $t_* \to -\infty$, the family of solutions $M_{t_*}(\cdot)$ to MCF converges to the unique Bowl soliton, i.e., the unique rotationally symmetric translating soliton with velocity one.

While the α-noncollapsedness property for mean curvature flow is preserved forward in time, it is not necessarily preserved going back in time. Indeed, Xu-Jia Wang [44] exhibited examples of ancient compact convex mean curvature flow solutions $\{M_t \mid t < 0\}$, that are not uniformly α-noncollapsed for any $\alpha > 0$. Such solutions lie in slab regions. The methods in [44] rely on the level set flow. Recently, Bourni, Langford and Tinaglia [6] provided a detailed construction of the Xu-Jia Wang solutions by different methods, showing also that the solution they construct is unique within the class of rotationally symmetric mean curvature flows that lie in a slab of a fixed width. In the present paper we will not consider these ancient collapsed solutions and focus on the classification of ancient closed noncollapsed mean curvature flows. In [33] they classify graphical translators for the mean curvature flow in dimension three. They also construct new examples of those in higher dimensions.

Definition 3.8. We say that an ancient solution $\{M_t : -\infty < t < T\}$ is *uniformly 2-convex* if there exists a uniform constant $\beta > 0$ so that

$$\lambda_1 + \lambda_2 \geq \beta H, \qquad \text{for all } t \leq t_0. \tag{3.15}$$

Throughout the paper we will be using the following observation: *if an Ancient Oval M_t is uniformly* 2*-convex, then by results in* [44]*, the backward limit of its type-I parabolic blow-up must be a shrinking round cylinder $\mathbb{R} \times S^{n-1}$, with radius $\sqrt{2(n-1)}$.*

Using Theorem 3.7, in [5] we first prove the uniqueness of uniformly 2-convex Ancient ovals in the presence of symmetry.

Theorem 3.9 (Uniqueness of $\boldsymbol{O(n)}$-invariant Ancient Ovals). *Let $(M_1)_t$ and $(M_2)_t$, $-\infty < t < T$, be two $O(n)$-invariant Ancient Ovals with the same axis of symmetry* (*which is assumed to be the x_1-axis*) *whose profile functions $U_1(x,t)$ and $U_2(x,t)$ satisfy equation* (3.9) *and rescaled profile functions $u_1(y,\tau)$ and $u_2(y,\tau)$ satisfy equation* (3.13)*. Then, they are the same up to translations along the axis of symmetry* (*translations in x*)*, translations in time and parabolic rescaling.*

In [5] we also establish the following result.

Theorem 3.10 (Rotational symmetry of Ancient Ovals). *If $\{M_t : -\infty < t < 0\}$ is an Ancient Oval which is uniformly* 2*-convex, then it is rotationally symmetric.*

Combining Theorem 3.2 and Theorem 3.10 yields the following result that establishes the uniqueness of uniformly two convex Ancient ovals. More precisely, we have the following.

Theorem 3.11 (Uniqueness of Ancient Ovals). *Let $\{M_t, -\infty < t < T\}$ be a uniformly* 2*-convex Ancient Oval. Then it is unique up to rotation, scaling and translation in time and hence it must the solution constructed by White in* [45] *and later by Haslhofer and Hershkovits in* [32]*.*

3.3. Complete ancient solutions to the MCF

A special case of ancient solutions are solitons; these are solutions that move in a self-similar fashion under the evolution. In a recent paper [8], Brendle proved that every noncollapsed steady Ricci soliton in dimension three is rotationally symmetric, and hence is isometric to the Bryant soliton up to scaling. Using similar techniques, Haslhofer in [30] subsequently proved that every noncollapsed, convex translating soliton for the mean curvature flow in R^3 is rotationally symmetric, and hence coincides with the bowl soliton up to scaling and ambient isometries. A related uniqueness result for the bowl soliton was proved in an important paper by Wang in [44]; this relies on a completely different approach.

In [12] the authors classified all convex ancient solutions to mean curvature flow in R^3 under a noncollapsing assumption. They prove the following.

Theorem 3.12 (Brendle, Choi). *Let M_t, for $t \in (-\infty, 0)$, be a noncompact ancient solution of mean curvature flow in R^3 which is strictly convex and noncollapsed. Then M_t agrees with the bowl soliton, up to scaling and ambient isometries.*

By combining the previous theorem with earlier work of White [45, 46] (see also [31]), the following conclusion holds.

Corollary 3.13. *Consider an arbitrary closed, embedded, mean convex surface in R^3, and evolve it by mean curvature flow. At the first singular time, the only possible blow-up limits are shrinking round spheres; shrinking round cylinders; and the translating bowl soliton.*

In the course of proof of Theorem 3.12 the authors study the asymptotic behavior of the flow as $t \to -\infty$. To establish more precise asymptotics they study the linearization of the mean curvature equation around a round cylinder. Using those asymptotics they establish the Neck Improvement Theorem, which asserts that a neck becomes more symmetric under the evolution. This result does not require that the solution is ancient, it can be applied whenever we have a solution of mean curvature flow which is close to a cylinder on a sufficiently large parabolic neighborhood. They iterate the Neck Improvement Theorem to show M_t is rotationally symmetric. Then, by analyzing the rotationally symmetric solutions, Brendle and Choi show that such solutions agree with the Bowl soliton. They treat higher-dimensional cases in [13].

3.4. Sketch of the proof of Theorem 3.11

Our proof of Theorem 3.10 closely follows the arguments by Brendle and Choi in [12, 13] on the uniqueness of strictly convex, non-compact, uniformly 2-convex, and noncollapsed ancient mean curvature flow.

The proof of Theorem 3.2 is quite involved. Our method is based on a priori estimates for various distance functions between two given ancient solutions in appropriate coordinates and measured in weighted L^2 norms. We need to consider two different regions: the *cylindrical* region and the *tip* region. The tip region is divided in two sub-regions: the *collar* and the *soliton* region, because we use different weights in those regions, ensuring the weight is C^1 at the overlap of the regions. We give definitions of all the regions, review the equations in each region and define appropriate weighted L^2 norms with respect to which we prove coercive type estimates.

In the statement of Theorem 3.2 we claim the uniqueness of any two Ancient Ovals up to dilations and translations. In fact since equation (3.13) is invariant under translation in time, translation in space and also under cylindrical dilations in space-time, each solution $M_i(t)$ gives rise to a three parameter family of solutions

$$M_i^{\alpha\beta\gamma}(t) = e^{\gamma/2}\, \Phi_\alpha(M_i(e^{-\gamma}(t-\beta))), \tag{3.16}$$

where Φ_α is a rigid motion, that is just the translation of the hypersurface along x axis by value α. The theorem claims the following: *given two ancient oval solutions we can find* α, β, γ *and* $t_0 \in \mathbb{R}$ *such that*

$$M_1(t) = M_2^{\alpha\beta\gamma}(t), \qquad \text{for } t \le t_0.$$

The profile function $U_i^{\alpha\beta\gamma}$ corresponding to the modified solution $M_i^{\alpha\beta\gamma}(t)$ is given by

$$U_i^{\alpha\beta\gamma}(x,t) = e^{\gamma/2} U_i\Big(e^{-\gamma/2}(x-\alpha), e^{-\gamma}(t-\beta)\Big). \tag{3.17}$$

We rescale the solutions $M_i(t)$ by a factor $\sqrt{-t}$ and introduce a new time variable $\tau = -\log(-t)$, that is, we set

$$M_i(t) = \sqrt{-t}\, \bar{M}_i(\tau), \qquad \tau := -\log(-t). \tag{3.18}$$

These are again $O(n)$ symmetric with profile function u, which is related to U by

$$U(x,t) = \sqrt{-t}\, u(y,\tau), \qquad y = \frac{x}{\sqrt{-t}}, \quad \tau = -\log(-t). \tag{3.19}$$

If the U_i satisfy the MCF equation (3.9), then the rescaled profiles u_i satisfy (3.13), i.e.,

$$\frac{\partial u}{\partial \tau} = \frac{u_{yy}}{1+u_y^2} - \frac{y}{2} u_y - \frac{n-1}{u} + \frac{u}{2}.$$

Translating and dilating the original solution $M_i(t)$ to $M_i^{\alpha\beta\gamma}(t)$ has the following effect on $u_i(y,\tau)$:

$$u_i^{\alpha\beta\gamma}(y,\tau) = \sqrt{1+\beta e^\tau}\, u_i\left(\frac{y - \alpha e^{\tau/2}}{\sqrt{1+\beta e^\tau}}, \tau + \gamma - \log\big(1+\beta e^\tau\big)\right). \tag{3.20}$$

To prove the uniqueness theorem we look at the difference $U_1 - U_2^{\alpha\beta\gamma}$, or equivalently at $u_1 - u_2^{\alpha\beta\gamma}$. The parameters α, β, γ are chosen so that the projections of $u_1 - u_2^{\alpha\beta\gamma}$ onto positive eigenspace (that is spanned by two independent eigenvectors) and zero eigenspace of the linearized operator $\mathcal{L}$ at the cylinder are equal to zero at time τ_0, which is chosen sufficiently close to $-\infty$. Correspondingly, we denote the difference $U_1 - U_2^{\alpha\beta\gamma}$ by $U_1 - U_2$ and $u_1 - u_2^{\alpha\beta\gamma}$ by $u_1 - u_2$. What we actually observe is that the parameters α, β and γ can be chosen to lie in a certain range, which allows our main estimates to hold without having to keep track of these parameters during the proof.

4. Ancient solutions to Yamabe flow

Let (M, g_0) be a compact manifold without boundary of dimension $n \ge 3$. If $g = v^{\frac{4}{n-2}} g_0$ is a metric conformal to g_0, the scalar curvature R of g is given in terms of the scalar curvature R_0 of g_0 by

$$R = v^{-\frac{n+2}{n-2}} \big(-\bar{c}_n \Delta_{g_0} v + R_0\, v\big)$$

where Δ_{g_0} denotes the Laplace Beltrami operator with respect to g_0 and $\bar{c}_n = 4(n-1)/(n-2)$.

In 1989 R. Hamilton introduced the *Yamabe flow*

$$\frac{\partial g}{\partial t} = -R\,g \tag{4.1}$$

as an approach to solve the *Yamabe problem* on manifolds of positive conformal Yamabe invariant. It is the negative L^2-gradient flow of the total scalar curvature, restricted to a given conformal class. The flow may be interpreted as deforming a Riemannian metric to a conformal metric of constant scalar curvature, when this flow converges.

Hamilton [29] showed the existence of the normalized Yamabe flow (which is the re-parametrization of (4.1) to keep the volume fixed) for all time; moreover, in the case when the scalar curvature of the initial metric is negative, he showed the exponential convergence of the flow to a metric of constant scalar curvature.

Since then, there has been a number of works on the convergence of the Yamabe flow on a compact manifold to a metric of constant scalar curvature. Chow [15] showed the convergence of the flow, under the conditions that the initial metric is locally conformally flat and of positive Ricci curvature. The convergence of the flow for any locally conformally flat initially metric was shown by Ye [47].

More recently, Schwetlick and Struwe [43] obtained the convergence of the Yamabe flow on a general compact manifold under a suitable Kazdan–Warner type of condition that rules out the formation of bubbles and that is verified (via the positive mass Theorem) in dimensions $3 \leq n \leq 5$. The convergence result, in its full generality, was established by Brendle [10] and [11] (up to a technical assumption, in dimensions $n \geq 6$, on the rate of vanishing of Weyl tensor at the points at which it vanishes): starting with any smooth metric on a compact manifold, the normalized Yamabe flow converges to a metric of constant scalar curvature.

In the special case where the background manifold M_0 is the sphere S^n and g_0 is the standard spherical metric g_{S^n}, the Yamabe flow evolving a metric $g = v^{\frac{4}{n-2}}(\cdot,t)\,g_{S^n}$ takes (after rescaling in time by a constant) the form of the *fast diffusion equation*

$$(v^{\frac{n+2}{n-2}})_t = \Delta_{S^n} v - c_n v, \qquad c_n = \frac{n(n-2)}{4}. \tag{4.2}$$

Explicit examples of ancient solutions to the Yamabe flow on S^n are:

Contracting spheres: They are special solutions v of (4.2) which depend only on time t and satisfy the ODE

$$\frac{dv^{\frac{n+2}{n-2}}}{dt} = -c_n\,v.$$

They are given by

$$v_S(p,t) = \left(\frac{4}{n+2}\,c_n\,(T-t)\right)^{\frac{n-2}{4}}. \tag{4.3}$$

and represent a sequence of round spheres shrinking to a point at time $t = T$. They are shrinking solitons and type I ancient solutions.

King solutions: They were discovered by J.R. King [35]. They can be expressed on $\mathbb{R}^n$ in closed from, namely $g = \bar{v}_K(\cdot,t)^{\frac{4}{n-2}}\, g_{\mathbb{R}^n}$, where $\bar{v}_K$ is the radial function

$$\bar{v}_K(r,t) = \left(\frac{A(t)}{1 + 2B(t)\, r^2 + r^4}\right)^{\frac{n-2}{4}}. \tag{4.4}$$

and the coefficients $A(t)$ and $B(t)$ satisfy a certain system of ODEs. The King solutions are analogous to the King–Rosenau solution and they can be visualized as two shrinking solitons, called the Barenblatt solutions, coming from spatial infinities and glued together.

It has been showed by Daskalopoulos, Hamilton and Sesum [22] that the spheres and the King–Rosenau solutions are the only compact ancient solutions to the two-dimensional Ricci flow. The natural question to raise is whether the analogous statement holds true for the Yamabe flow, that is, whether the contracting spheres and the King solution are the only compact ancient solutions to the Yamabe flow. This occurs not to be the case as the following few results show.

In [20] we show the following result thus showing that the classification of closed ancient solutions to the Yamabe flow is very difficult, if not impossible. Unlike the above-mentioned closed ancient solutions, the Ricci curvature of the tower of bubbles solutions changes its sign (they still have nonnegative scalar curvature).

Theorem 4.1. *There exist infinitely many ancient radially symmetric solutions of the Yamabe flow* (4.2) *on S^n other than the contracting spheres* (4.3) *and the King solutions* (4.4). *Our new solutions, as $t \to -\infty$, may be visualized as two spheres joint by a short neck. Their curvature operator changes sign and they are type II ancient solutions. We refer to those solutions as to a tower of n moving bubbles.*

Since the towers of moving bubbles are shown to be type II ancient solutions, while the contracting spheres and the King solutions are of type I, one may still ask whether the latter two are the only ancient compact type I solutions of the Yamabe flow on S^n, equation (4.2). In [21] we observe that this is not the case, as we show the existence of other ancient compact type I solutions on S^n. We prove the following theorem.

Theorem 4.2. *There exist infinitely many ancient closed Type I solutions to the Yamabe flow that can be viewed, as time approaches $-\infty$, as two traveling waves, that is Yamabe shrinkers, glued together, with the cylinder in the middle.*

There exist infinitely many complete, noncompact rotationally symmetric Yamabe shrinkers that open up as cylinders at spatial infinity and in the previous theorem we can use any two of these Yamabe shrinkers to construct a closed ancient solution to the Yamabe flow. Theorem 4.1 and 4.2 imply the classification result of closed ancient solutions to the Yamabe flow is nearly impossible to obtain.

Acknowledgment

N. Sesum thanks the NSF for support in DMS-1056387.

References

[1] Andrews, B., *Noncollapsing in mean-convex mean curvature flow*; Geom. Topol., 16, (2012), 1413–1418.

[2] Angenent, S., *Formal asymptotic expansions for symmetric ancient ovals in mean curvature flow*, Networks and Heterogeneous Media, **8**, (2013) 1–8,

[3] Angenent, S., *Shrinking doughnuts*, in NONLINEAR DIFFUSION EQUATIONS AND THEIR EQUILIBRIUM STATES, 3 (GREGYNOG, 1989), 21–38, Birkhäuser Boston, Bost34'on, MA, 1992.

[4] Angenent, S., Daskalopoulos, P., Sesum, N., *Unique asymptotics of ancient convex mean curvature flow solutions*; to appear in J. Diff. Geom.

[5] Angenent, S., Daskalopoulos, P., Sesum, N., *Uniqueness of two-convex closed ancient solutions to the mean curvature flow*; arXiv:1804.07230.

[6] Bourni, T., Langford, M., Tinaglia, G., *On the existence of translating solutions of mean curvature flow in slab regions*; arxiv:1805.05173.

[7] Brendle, S., Ancient solutions to the Ricci flow in dimension 3; arXiv:1811.02559.

[8] Brendle, S., *Rotational symmetry of self-similar solutions to the Ricci flow*; Invent. Math. 194 (2013), no. 3, 731–764.

[9] Brendle, S., *Rotational symmetry of Ricci solitons in higher dimensions*; J. Differential Geom. 97 (2014), no. 2, 191–214.

[10] Brendle, S., *Convergence of the Yamabe flow for arbitrary initial energy*, J. Differential Geom. **69** (2005), 217–278.

[11] Brendle, S., *Convergence of the Yamabe flow in dimension* 6 *and higher*, Invent. Math. **170** (2007), 541–576.

[12] Brendle, S., Choi, K., *Uniqueness of convex ancient solutions to the mean curvature flow in* R^3; arXiv: 1711.00823v2.

[13] Brendle, S., Choi, K., *Uniqueness of convex ancient solutions to mean curvature flow in higher dimensions*; arXiv: 1804.00018.

[14] R.L. Bryant, *Ricci flow solitons in dimension three with* $SO(3)$*-symmetries*; available at www.math.duke.edu/ bryant/3DRotSymRicciSolitons.pdf

[15] Chow, B., *The Ricci flow on the* 2*-sphere*, J. Diff. Geom. **33** (1991), 325–334.

[16] Chu, S.C., *Type II ancient solutions to the Ricci flow on surfaces*, Comm. Anal. Geom. **15** (2007), 195–216.

[17] Courant, R., Hilbert, D., *Methods of mathematical physics.* Vol. I. Interscience Publishers, Inc., New York, N.Y., 1953.

[18] Daskalopoulos, P., Hamilton, R., Sesum, N., *Classification of compact ancient solutions to the curve shortening flow*; J. Differential Geom. 84 (2010), no. 3, 455–464.

[19] Daskalopoulos, P., Sesum, N., *Eternal solutions to the Ricci flow on* $\mathbb{R}^2$; Int. Math. Res. Not. **2006**, 1–20.

[20] Daskalopoulos, P., del Pino, M., Sesum, N., *Type II ancient compact solutions to the Yamabe flow*; to appear in J. reine ang. Math.; arXiv: arXiv:1209.5479.

[21] Daskalopulos, P., del Pino, M., Sesum, N., *New Type I ancient compact solutions of the Yamabe flow*; Math. Res. Lett. 18 (2011).

[22] Daskalopoulos, P., Hamilton, R., Sesum, N., *Classification of ancient compact solutions to the Ricci flow on surfaces*; J. Differential Geom. 91 (2012), no. 2, 171–214.

[23] Drugan, G., *Self-shrinking Solutions to Mean Curvature Flow.* Thesis (Ph.D.)-University of Washington. 2014. 123 pp. ISBN: 978-1321-10634-3

[24] Filippas, S., Kohn, R., *Refined asymptotics for the blowup of* $u_t - \Delta u = u^p$; Comm. Pure Appl. Math. 45 (1992), 821–869.

[25] Hamilton, R., *Formation of singularities in the Ricci flow*; Surveys in Differential Geometry, vol II, 7–136, International Press, Somerville MA (1995).

[26] Hamilton, R., *The Ricci flow on surfaces*; Mathematics and General Relativity **71** (1988), 237–261.

[27] Hamilton, R., *Harnack estimate for the mean curvature flow*; J. Diff. Geom. **41** (1995), 215–226.

[28] Hamilton, R., *Eternal solutions to the Ricci flow*, J. Diff. Geom. **38** (1993), 1–11.

[29] Hamilton, R.S., Lectures on geometric flows, 1989, unpublished.

[30] Haslhofer, R., *Uniqueness of the bowl soliton*, Geom. Topol. 19, 2393–2406 (2015).

[31] Haslhofer, R., Kleiner, B., *Mean curvature flow of mean convex hypersurfaces*; arXiv:1304.0926v2.

[32] Haslhofer, R., Hershkovits, O., *Ancient solutions of the mean curvature flow*; Comm. Anal. Geom. 24 (2016).

[33] Hoffman, D., Ilmanen, T., Martin, F., White, B., *Graphical translators for mean curvature flow*; arXiv: 1805.10860.

[34] Huisken, G., *Flow by mean curvature of convex surfaces into spheres*; J. Differential Geom. 20 (1984), no. 1, 237–266.

[35] King, J.R., *Exact polynomial solutions to some nonlinear diffusion equations*, Physica. D **64** (1993), 39–65.

[36] Kleene, S., Møller, N.M., *Self-shrinkers with a rotational symmetry.* Trans. Amer. Math. Soc. **366** (2014), no. 8, 3943–3963.

[37] Ladyzhenskaya, O.A., Solonnikov, V.A., Ural'ceva, N.N., *Linear and Quasi-linear Equations of Parabolic Type*; A.M.S. Transl. Math. Monogr. 23 (1968).

[38] Merle, F., Zaag, H., *Optimal estimates for blowup rate and behavior for nonlinear heat equations*; Comm. Pure Appl. Math., 51 (1998), 139–196.

[39] Nakayama, K., Iizuka, T. and Wadati, M., *Curve Lengthening Equation and its Solutions,* Journal of the Physical Society of Japan, 1994, **63**, 1311–1321.

[40] Perelman, G., *The entropy formula for the Ricci flow and its geometric applications*; arXiv:math/0211159.

[41] Perelman, G., *Ricci flow with surgery on three-manifolds*; arXiv:math/0303109.

[42] Rosenau, P., *Fast and superfast diffusion processes*, Phys. Rev. Lett. **74** (1995), 1056–1059.

[43] Schwetlick, H.; Struwe, M., *Convergence of the Yamabe flow for "large" energies*, J. Reine Angew. Math. **562** (2003), 59–100.

[44] Wang, X.-J., *Convex solutions to the mean curvature flow*; Ann. of Math. (2) 173 (2011), no. 3, 1185–1239.

[45] White, B., *The nature of singularities in mean curvature flow of mean convex sets*; J. Amer. Math. Soc., 16(1):123–138, 2003.

[46] White, B., *The size of the singular set in mean curvature flow of mean convex sets*; J. Amer. Math. Soc. 13 (2000).

[47] Ye, R., *Global existence and convergence of Yamabe flow*, J. Differential Geom. **39** (1994), 35–50.

Natasa Sesum
Department of Mathematics
Rutgers University
e-mail: `natasas@math.rutgers.edu`

Progress in Mathematics, Vol. 333, 465–469

The Kähler–Ricci Flow on $\mathbb{CP}^2$

Jian Song

Abstract. We give a direct proof of the convergence of the Kähler–Ricci flow on $\mathbb{CP}^2$ without assuming the existence of the Kähler–Einstein metric.

Mathematics Subject Classification (2010). Primary 53C55, 53C20.

Keywords. Kähler–Ricci flow, projective space.

1. Introduction

A Fano manifold is a closed Kähler manifold with positive First Chern class. The well-known Yau–Tian–Donaldson conjecture predicts the equivalence between the existence of the Kähler–Einstein metric and the algebraic K-stability. This conjecture was recently settled in [6, 21], extending Yau's solution to the Calabi conjecture [25]. The Ricci flow, introduced by Hamilton [9], is a canonical deformation to obtain Einstein or soliton metrics on Riemannian manifolds. In the Kähler case, the Kähler–Ricci flow provides an alternative proof for the Calabi conjecture [4]. In the Fano case, the Kähler–Ricci flow is proved to converge to a $\mathbb{Q}$-Fano Kähler–Einstein space with possible mild singularities [3, 8, 22]. The smooth convergence of the Fano Kähler–Ricci flow is obtained in [23, 24] assuming existence of Kähler–Einstein metric. In general, it is not clear how to directly prove general smooth convergence of the flow on $\mathbb{CP}^n$ without assuming the existence of Kähler–Einstein metric, algebraic K-stability or curvature assumption on the initial metric [7]. In the case of $\mathbb{CP}^2$, the smooth convergence is established in [15] assuming the Mabuch K-energy is bounded below, a condition equivalent to K-semistability. In this short note, we give a direct proof for the convergence of the Kähler–Ricci flow on $\mathbb{CP}^2$ without assuming the existence of Kähler–Einstein metrics by combining the blow-up argument and the classification of ALE Ricci flat Kähler surfaces.

Research supported in part by National Science Foundation grants DMS-1711439.

Let (X, ω_0) be a Fano manifold equipped with a Kähler metric $\omega_0 \in c_1(X)$. We consider the following normalized Kähler–Ricci flow on X

$$\begin{cases} \dfrac{\partial}{\partial t}\omega = -\operatorname{Ric}(\omega) + \omega, \\ \omega|_{t=0} = \omega_0 \,. \end{cases} \tag{1.1}$$

The following is the main result of the paper.

Theorem 1.1. *Let $\omega(t)$ be the global solution of the normalized Kähler–Ricci flow* (1.1) *on $\mathbb{CP}^2$ with initial Kähler metric $\omega_0 \in c_1(\mathbb{CP}^2)$. Then $\omega(t)$ converges smoothly to a Kähler–Einstein metric on $\mathbb{CP}^2$.*

The upshot of the proof for Theorem 1.1 is that we do not assume existence of Kähler–Einstein metric on $\mathbb{CP}^2$. The method can possibly be extended to $\mathbb{CP}^2$ blow-up at one point.

2. Proof

The main result in this section is to establish a uniform curvature bound for the Kähler–Ricci flow (1.1) on $\mathbb{CP}^2$.

Proposition 2.1. *For any initial Kähler metric $\omega_0 \in c_1(\mathbb{CP}^2)$, there exists $C > 0$ such that for*

$$\sup_{\mathbb{CP}^2 \times [0,\infty)} |Rm(g(t))|_{g(t)} < \infty.$$

We will prove the proposition by contradiction. Let $g(t)$ be the solution of (1.1) associated to the Kähler form $\omega(t)$. Suppose there exist a sequence $(z_j, t_j) \in \mathbb{CP}^2 \times [0, \infty)$ with $\lim_{j\to\infty} t_j = \infty$ such that

$$\lim_{j\to\infty} |Rm|_g(z_j, t_j) = \infty.$$

Without loss of generality, we can assume that

$$|Rm|_g(z_j, t_j) = \max_{\mathbb{CP}^2 \times [0,t_j]} |Rm|_g.$$

Let us recall the following result of Perelman [13, 16].

Lemma 2.2. *Let $g(t)$ be a global solution of the Kähler–Ricci flow* (1.1)*. Then there exists $C > 0$ such that for all $t \geq 0$,*

$$|R(t)| \leq C, \tag{2.1}$$

where $R(t)$ is the scalar curvature of $g(t)$.

We now apply the parabolic scaling to $g(t)$ by letting

$$g_j(t) = \lambda_j g(\lambda_j^{-1} t + t_j)).$$

Then $g_j(t)$ satisfies

$$\frac{\partial g_j}{\partial t} = -\operatorname{Ric}(g_j) + \lambda_j^{-1} g_j. \tag{2.2}$$

In particular,

$$\max_{\mathbb{CP}^2\times[-t_j,0]} |Rm_{g_j}|_g = |Rm(g_j)|_{g_j}(z_j,0) = 1.$$

Therefore we can apply Hamilton's compactness theorem. After possibly passing to a subsequence, $(\mathbb{CP}^2, g_j(t), z_j)$ converges to a smooth Kähler–Ricci flow $(X_\infty, g_\infty(t))$ for $t \in (-\infty, 0)$

$$\frac{\partial g_\infty}{\partial t} = -\operatorname{Ric}(g_\infty), \tag{2.3}$$

where X_∞ is a smooth complete Kähler surface.

Lemma 2.3. *For all $t \in (-\infty, 0)$, we have*

$$\operatorname{Ric}(g_\infty(t)) \equiv 0.$$

Proof. Let R and Ric be the scalar curvature and Ricci curvature of $g_\infty(t)$. Then

$$\frac{\partial R}{\partial t} = \Delta_{g_\infty} R + |\operatorname{Ric}|^2_{g_\infty}.$$

On the other hand, by Lemma 2.2,

$$R \equiv 0.$$

The lemma then immediately follows. □

In particular, $g_\infty(t)$ is a static solution and we use g_∞ for $g_\infty(t)$.

Lemma 2.4. *(X_∞, g_∞) is a Ricci flat ALE Kähler surface.*

Proof. Let z_∞ be the limiting point of z_j along the sequence $(\mathbb{CP}^2, g_j(t), z_j)$. By Perelman's κ-noncollapsing, there exists $\kappa > 0$ such that for any $r > 0$

$$\operatorname{Vol}_{g_\infty}(B_{g_\infty}(z_\infty, r)) \geq \kappa r^4.$$

Furthermore, the L^2-curvature of (X_∞, g_∞) is bounded from the smooth convergence. We can now blow down $(X_\infty, g_\infty, z_\infty)$ by letting

$$g_{\infty,k} = \epsilon_k g_\infty, \quad \lim_{k\to\infty} \epsilon_k = 0.$$

By the Cheeger–Colding–Tian theory [5], $(X_\infty, g_{\infty,k}, z_\infty)$ converges in pointed Gromov–Hausdorff sense to an orbifold Kähler surface Z with a cone metric g_Z. In fact, (Z, g_Z) is the tangent cone of (X_∞, g_∞) at infinity. In particular, there is at most one isolated orbifold point p as the limiting point of z_∞ and g_Z is a smooth Ricci-flat Kähler metric on $Z \setminus \{p\}$. Therefore the link of the cone (X_∞, g_∞) must be quotient of S^3 and (Z, g_Z) is the quotient space of Euclidean $\mathbb{C}^2$ by a finite subgroup of $U(2)$ by the standard theory of four manifolds [1, 10, 18]. The lemma immediately follows. □

The following lemma is obtained in [17] to generalizing the hyper-Kähler case [10].

Lemma 2.5. *Let (M, J, g) be a smooth ALE Ricci-flat Kähler surface whose tangent cone at the infinity is $\mathbb{C}^2/H$ for some finite subgroup of $U(2)$. Then (M, J) can be* *obtained as the minimal resolution of a one parameter $\mathbb{Q}$-Gorenstein deformation of $\mathbb{C}^2/H$.*

The following corollary follows immediately from Lemma 2.5 and the Hodge relation for complex surfaces by considering the exceptional curve in the resolution of Lemma 2.5.

Corollary 2.6. *There exists a holomorphic curve D of negative self-intersection in (X_∞, J') for some complex structure J'.*

Now we can complete the proof of Proposition 2.1. Since $(\mathbb{CP}^2, g_j(0), z_j)$ converges smoothly to (X_∞, g_∞). There exists $R > 0$ such that $D \subset B_{g_\infty}(z_\infty, R)$. This implies that there exists an element in $H_2(\mathbb{CP}^2)$ with negative self-intersection. Contradiction.

Finally, we are ready to prove Theorem 1.1. By Hamilton's compactness theorem, for any $t_j \to \infty$, after possibly taking a subsequence, $g(t + t_j)$ with $t \in [0, 1]$ converge smoothly to a Kähler–Ricci soliton metric $g_\infty(t)$ on a limiting Kähler surface (X_∞, J_∞) diffeomorphic to $\mathbb{CP}^2$. By Kodaira's classification of complex surfaces, (X_∞, J_∞) must be biholomorphic to $\mathbb{CP}^2$. Therefore the limiting metric g_∞ is a Kähler–Ricci soliton metric on $\mathbb{CP}^2$. On the other hand, the Lie algebra of the automorphism group for $\mathbb{CP}^n$ is $\eta = sl(n+1, \mathbb{C})$ and $\eta = [\eta, \eta]$. Direct computations show the Futaki invariant of $\mathbb{CP}^n$ vanishes (see [20]). Therefore there cannot be non-Einstein Kähler–Ricci soliton on $\mathbb{CP}^n$. This completes the proof of Theorem 1.1.

Acknowledgement

The author enjoys and learns a great deal from the very fruitful and exciting collaboration with Professor Gang Tian in the past fifteen years. This paper is dedicated to Professor Gang Tian on the occasion of his 60th birthday celebration.

References

[1] Anderson, M. *Ricci curvature bounds and Einstein metrics on compact manifolds*, J. Amer. Math. Soc. Vol. 2, No. 3(1989) 455–490

[2] Aubin, T. *Équations du type Monge–Ampère sur les variétés kählériennes compactes*, Bull. Sci. Math. (2) **102** (1978), no. 1, 63–95

[3] Bamler, R. *Convergence of Ricci flows with bounded scalar curvature*, arXiv: 1603.05235, to appear in Annals of Math.

[4] Cao, H.-D. *Deformation of Kähler metrics to Kähler–Einstein metrics on compact Kähler manifolds*, Invent. Math. 81 (1985), no. 2, 359–372

[5] Cheeger, J., Colding, T.H. and Tian, G. *On the singularities of spaces with bounded Ricci curvature*, Geom.Funct.Anal. Vol.12 (2002), 873–914

[6] Chen, X., Donaldson, S. and Sun, S. *Kähler–Einstein metrics on Fano manifolds. III: Limits as cone angle approaches 2π and completion of the main proof*, J. Amer. Math. Soc. 28 (2015), no. 1, 235–278

[7] Chen, X.X. and Tian, G. *Ricci flow on Kähler–Einstein surfaces* Invent. Math. 147 (2002), no. 3, 487–544

[8] Chen, X.X. and Wang, B. *Space of Ricci flows* (*II*), preprint

[9] Hamilton, R. *Three-manifolds with positive Ricci curvature*, J. Differential Geom. 17 (1982), no. 2, 255–306

[10] Kronheimer, P.B. *The construction of ALE spaces as hyper-Kähler quotients*, J. Differ. Geom. 29(3), 665–683 (1989)

[11] Nakajima, H. *Hausdorff convergence of Einstein* 4*-manifolds*, J. Fac. Sci. Univ. Tokyo 35 (1988) 411–424

[12] Perelman, G. *The entropy formula for the Ricci flow and its geometric applications*, arXiv:math.DG/0211159

[13] Perelman, G., unpublished work on the Kähler–Ricci flow

[14] Phong, D.H., Song, J., Sturm, J. and Weinkove, B. *The Kähler–Ricci flow and the* $\bar\partial$ *operator on vector fields*, J. Differential Geometry 81 (2009), no. 3, 631–647

[15] Ruan, W., Zhang, Y. and Zhang, Z. *Bounding sectional curvature along the Kähler–Ricci flow*, Commun. Contemp. Math. 11 (2009), no. 6, 1067–1077

[16] Sesum, N. and Tian, G. *Bounding scalar curvature and diameter along the Kähler Ricci flow* (*after Perelman*), J. Inst. Math. Jussieu 7 (2008), no. 3, 575–587

[17] Suvaina, I. *ALE Ricci-flat Kähler metrics and deformations of quotient surface singularities*, Ann. Global Anal. Geom. 41 (2012), no. 1, 109–123

[18] Tian, G. *On Calabi's conjecture for complex surfaces with positive first Chern class*, Invent. Math. 101, no. 1 (1990), 101–172

[19] Tian, G. *Kähler–Einstein metrics with positive scalar curvature*, Invent. Math. **130** (1997), no. 1, 1–37

[20] Tian, G. *Canonical metrics in Kähler geometry*, Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2000. vi+101 pp.

[21] Tian, G. *K-stability and Kähler–Einstein metrics*, Comm. Pure Appl. Math. 68 (2015), no. 7, 1085–1156

[22] Tian, G. and Zhang, Z. *Regularity of Kähler–Ricci flows on Fano manifolds*, Acta Math. 216 (2016), no. 1, 127–176

[23] Tian, G. and Zhu, X. *Convergence of Kähler–Ricci flow*, J. Amer. Math. Soc. 20 (2007), no. 3, 675–699

[24] Tian, G. and Zhu, X. *Convergence of the Kähler–Ricci flow on Fano manifolds*, J. Reine Angew. Math. 678 (2013), 223–245

[25] Yau, S.-T. *On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampère equation, I*, Comm. Pure Appl. Math. **31** (1978), 339–411

Jian Song
Department of Mathematics
Rutgers University
Piscataway, NJ 08854, USA
e-mail: `jiansong@math.rutgers.edu`

Progress in Mathematics, Vol. 333, 471–510

Pluriclosed Flow and the Geometrization of Complex Surfaces

Jeffrey Streets

Abstract. We recall fundamental aspects of the pluriclosed flow equation and survey various existence and convergence results, and the various analytic techniques used to establish them. Building on this, we formulate a precise conjectural description of the long time behavior of the flow on complex surfaces. This suggests an attendant geometrization conjecture which has implications for the topology of complex surfaces and the classification of generalized Kähler structures.

Mathematics Subject Classification (2010). 53C44, 53C55.

Keywords. Pluriclosed flow, complex surfaces.

1. Introduction

The Enriques–Kodaira classification of compact complex surfaces is a landmark achievement of 20th century mathematics, exploiting deep techniques from complex analysis, partial differential equations and algebraic geometry to give a descriptive classification of complex manifolds of complex dimension two. Despite being dubbed a classification, many questions remain unanswered, and the structure of complex surfaces remains an active area of research to this day. As the classical uniformization of Riemann surfaces profoundly intertwines complex structures and associated canonical Riemannian geometries, it is natural to try to associate canonical metrics to complex surfaces in order to provide further insight. The purpose of this article is to describe such a “geometrization conjecture” for compact complex surfaces. Specifically we aim to associate canonical families of Hermitian metrics on all complex surfaces via a universal geometric flow construction, and use properties of the resulting metrics to capture aspects of the underlying complex structure, and in particular yield the classification of Kodaira’s Class VII surfaces.

This strategy for understanding complex surfaces is inspired by Perelman's landmark resolution of Thurston's Geometrization Conjecture for 3-manifolds using Ricci flow [47, 58, 63–65]. Thurston [89] sought to decompose any compact 3-manifold into pieces, each of which admits canonical Riemannian metrics, the models of which are inspired by a thorough understanding of possible locally homogeneous spaces. In 1982, Hamilton introduced the Ricci flow equation and used it to classify compact 3-manifolds with positive Ricci curvature. Over the ensuing decades Hamilton further developed the theory of Ricci flow, eventually giving a conjectural description of the singularity formation and long time behavior of the flow, which would lead to a proof of Geometrization. Exploiting many deep insights into the structure of Ricci flow, Perelman achieved a precise description of this singularity formation, in particular yielding Geometrization. Amazingly, this one analytic tool provides both the topological decomposition of the underlying 3-manifold as well as an essentially canonical construction of the relevant geometric structures.

Inspired by this story, the author and Tian sought to bring the philosophy of geometric evolution to bear in understanding the topology and geometry of complex manifolds [80, 81]. Of course Ricci flow has already had a strong influence in complex, *Kähler*, geometry. The Ricci flow equation preserves Kähler geometry, and Cao initiated the study of Kähler–Ricci flow, [11] using it to reprove the Calabi–Yau [103] and Aubin–Yau Theorems [4, 103], canonically constructing Kähler–Einstein metrics on manifolds with $c_1 = 0$, $c_1 < 0$ respectively. More recently, Song–Tian initiated the analytic minimal model program [70], seeking to understand the algebraic minimal model program through the singularities of Kähler–Ricci flow. However, there are many examples of complex, *non-Kähler* manifolds, starting with the basic example of Hopf who constructed complex structures on $S^3 \times S^1$. The Kähler–Ricci flow cannot be employed on such manifolds, and one can show that for non-Kähler metrics the Ricci flow will not even preserve the condition that a metric is Hermitian, thus one must look elsewhere to apply the ideas of geometric evolutions. Thus the *pluriclosed flow* equation [80] was introduced as an extension of Kähler–Ricci flow which preserves natural conditions for Hermitian, non-Kähler metrics.

Following the general philosophy of geometric evolutions, one expects the limiting behavior of pluriclosed flow to meaningfully reflect aspects of the complex structure and topology of the underlying surface. Up to now results have been established which support the main conjecture on the maximal smooth existence time of the flow, and we will recount these below. Moreover we have phenomenological results in the locally homogeneous setting which suggest the general limiting behavior of the flow in many cases [7]. However, there is not yet a precise conjectural picture of how the flow behaves on most complex surfaces, most importantly Kodaira's Class VII$^+$ surfaces. These are minimal complex surfaces of negative Kodaira dimension and $b_2 > 0$. Conjecturally these are all diffeomorphic to $S^3 \times S^1 \# k\overline{\mathbb{CP}}^2$, with a complete description of the different complex structures.

Despite the relative simplicity of the underlying diffeotypes, these represent a large and varied class of surfaces, and their classification remains the main open problem in the Kodaira classification of surfaces. In this article we formulate precise conjectures on the limiting behavior of the pluriclosed flow on all complex surfaces, and announce a number of new global existence and convergence results in support of these conjectures.

Furthermore, it turns out that the pluriclosed flow is also related to Hitchin's "generalized geometry" program. The first hint of this was shown in [82], where the author and Tian showed that the pluriclosed flow equation actually preserves the delicate integrability conditions of Gualtieri/Hitchin's "generalized Kähler geometry," a subject with roots in mathematical physics. We will see in part III below that the conjectural framework for pluriclosed flow leads to a classification of generalized Kähler structures on complex surfaces. One example of particular interest is the case of $\mathbb{CP}^2$, where the global existence and convergence of the pluriclosed flow implies a uniqueness theorem for generalized Kähler structures, extending Yau's theorem on the uniqueness of the complex structure [102].

Part I: Pluriclosed flow

2. Existence and basic regularity properties

2.1. Definition and local existence

In this section we recount the rudimentary properties of the pluriclosed flow equation. To begin we define the pluriclosed condition.

Definition 2.1. Let (M^{2n}, g, J) be a complex manifold with Hermitian metric g, and associated Kähler form $\omega(X,Y) = g(X, JY)$. The metric g is *Kähler* if

$$d\omega = 0.$$

The metric g is *pluriclosed* if

$$\sqrt{-1}\partial\overline{\partial}\omega = dd^c\omega = 0,$$

where $d^c\omega = \sqrt{-1}(\overline{\partial} - \partial)\omega = -d\omega(J, J, J)$. As they are equivalent notions we will often refer to the Kähler form ω as being either Kähler or pluriclosed.

The Kähler condition is the simplest, and strongest, integrability condition for a Hermitian metric. The pluriclosed condition is essentially the only weakening of the Kähler condition which is linear in the Kähler form. As the aim is to understand all complex surfaces through the geometry of pluriclosed metrics, the following result of Gauduchon is of fundamental importance.

Theorem 2.2 ([30]). *Let (M^{2n}, g, J) be a compact complex manifold with Hermitian metric g. There exists a unique $u \in C^\infty(M)$ such that*

$$\sqrt{-1}\partial\overline{\partial}\left(e^{2u}\omega\right)^{n-1} = 0, \qquad \int_M u dV_g = 0.$$

In particular, applying this theorem in the case $n = 2$, we see that every compact complex surface admits pluriclosed metrics.

It is well known that for a Kähler manifold, the Levi-Civita connection preserves J, and is the unique Hermitian connection associated to the pair (g, J). When the metric is not Kähler, there is a natural one-parameter family of Hermitian connections associated to (g, J) ([32]). Two of these are particularly relevant to pluriclosed flow, namely the Bismut and Chern connections, defined by:

$$\begin{aligned} \left\langle \nabla^B_X Y, Z \right\rangle =&\ \left\langle \nabla^{LC}_X Y, Z \right\rangle + \tfrac{1}{2} d^c\omega(X, Y, Z),\\ \left\langle \nabla^C_X Y, Z \right\rangle =&\ \left\langle \nabla^{LC}_X Y, Z \right\rangle + \tfrac{1}{2} d\omega(JX, Y, Z). \end{aligned}$$

These connections induce curvature tensors Ω^B and Ω^C, and also connections on the canonical bundle, yielding in turn representatives of the first Chern class:

$$\rho^B_{\alpha\beta} = g^{\bar{j}i}\Omega^B_{\alpha\beta i\bar{j}}, \qquad \rho^C_{\alpha\beta} = g^{\bar{j}i}\Omega^C_{\alpha\beta i\bar{j}}.$$

By Chern–Weil theory we know that $d\rho^{B,C} = 0$, and $\rho^{B,C} \in c_1$. However, it is not the case in general that ρ^B is of type $(1, 1)$. Of course in the Kähler setting $\rho^B = \rho^C = \rho$, the Ricci form of the underlying metric.

The classical quest for canonical metrics on manifolds often centers around existence questions for Einstein metrics, i.e., solutions of

$$\mathrm{Rc} = \lambda g$$

for some constant λ. Hamilton [38] introduced the Ricci flow

$$\frac{\partial}{\partial t} g = -2\,\mathrm{Rc}$$

as a tool for constructing such metrics by the parabolic flow method. Cao [11] observed that the Ricci flow equation will preserve the Kähler condition, and can be expressed in terms of the Kähler form as

$$\frac{\partial}{\partial t}\omega = -\rho = \sqrt{-1}\partial\overline{\partial}\log\det g, \tag{2.1}$$

where the last expression holds in local complex coordinates.

Given the vast array of analytic tools and structure available for Ricci flow, one would want to use it to understand complex manifolds beyond the Kähler setting. However, for a general Hermitian metric the Ricci tensor is not of $(1, 1)$ type and thus the Ricci flow equation will not preserve the class of Hermitian metrics. For this reason we are forced to define a new equation if we are to preserve aspects of Hermitian, non-Kähler geometry. In [81] the author and Tian introduced a family of parabolic equations for Hermitian metrics on complex manifolds. Let $S_{i\bar{j}} = g^{\bar{l}k}\Omega^C_{k\bar{l}i\bar{j}}$, which is a kind of Ricci tensor defined using the Chern connection, and is always a $(1, 1)$ form. Furthermore, we let T denote the torsion of the Chern connection, and let $Q = T \star T$ denote an *arbitrary* quadratic in T which satisfies $Q \in \Lambda^{1,1}$. In [81] the author and Tian defined the class of flow equations:

$$\frac{\partial}{\partial t}\omega = -S + Q, \tag{2.2}$$

referring to these generally as *Hermitian curvature flow*. We showed that for an arbitrary choice of Q this equation is strictly parabolic and satisfies various natural analytic conditions such as smoothing estimates and stability near Kähler–Einstein metrics. In [81] a particular choice of Q was identified corresponding to the Euler equation of a certain Hilbert-type functional in Hermitian geometry. More recently Ustinovskiy [99] showed that for a different choice of Q the flow will preserve various curvature positivity conditions, leading to extensions of the classical Frankel conjecture into non-Kähler geomery. Given the rich diversity of Hermitian geometry, it is natural to expect that different flows, i.e., different choices of Q, would be needed to address different situations. In [80] the author and Tian identified a particular choice of Q which yields a flow which preserves the pluriclosed condition, specifically

$$Q^1_{i\bar{j}} = g^{\bar{l}k} g^{\bar{q}p} T_{ik\bar{p}} T_{\bar{j}\bar{l}q}.$$

As it turns out the resulting flow equation has several different useful manifestations, and we record the fundamental definition including some of these forms.

Definition 2.3. Let (M^{2n}, J) be a complex manifold. We say that a one-parameter family of pluriclosed metrics ω_t is a solution of *pluriclosed flow* if

$$\frac{\partial}{\partial t}\omega = - S + Q^1.$$

This is an example of Hermitian curvature flow, which is well posed for arbitrary Hermitian metrics. With pluriclosed initial data it can also be expressed using the Bismut curvature as

$$\frac{\partial}{\partial t}\omega = - \rho_B^{1,1}.$$

It is useful to furthermore express the flow in local complex coordinates, yielding

$$\frac{\partial}{\partial t}\omega = \partial\partial^*_\omega\omega + \overline{\partial\partial^*_\omega}\omega + \sqrt{-1}\partial\overline{\partial}\log\det g. \tag{2.3}$$

As an example of Hermitian curvature flow, there exist short time solutions on compact manifolds as remarked above. Arguing specifically in this setting, it turns out that the operator $\rho_B^{1,1}$, restricted to the class of pluriclosed metrics, is strictly elliptic. In fact, the symbol of the linearized operator is just the Laplacian with respect to the given metric. This renders pluriclosed flow a strictly parabolic equation, and so by appealing to general theory we can obtain short time existence on compact manifolds. Moreover, comparing (2.3) and (2.1), it is natural to expect that if we start the flow with Kähler initial data then it is a solution of Kähler–Ricci flow. This in fact holds, so to summarize:

Theorem 2.4 (cf. [80, Theorem 1.2]). *Let (M^{2n}, J) be a compact complex manifold. Suppose ω_0 is a pluriclosed metric on M. Then there exists $\epsilon > 0$ and a unique solution to pluriclosed flow ω_t with initial condition ω_0. If ω_0 is Kähler, then ω_t is the unique solution to Kähler–Ricci flow with initial data ω_0.*

Having obtained short time solutions, the main task is to describe the maximal smooth existence time, as well as the limiting behavior at this time, in general a formidable task. The remainder of this paper is devoted to giving precise conjectures on these questions in the specific case of complex surfaces.

Remark 2.5. We point out that a different geometric flow approach has been introduced to study complex surfaces, the Chern–Ricci flow:

$$\frac{\partial}{\partial t}\omega = -\rho^C.$$

This flow also preserves the pluriclosed condition, and reduces to a scalar PDE modeled on the parabolic complex Monge–Ampère equation. A crucial qualitative distinction between this flow and pluriclosed flow is that for Chern–Ricci flow the torsion is fixed as a tensor along the flow, i.e., $d\omega_t = d\omega_0$, whereas for pluriclosed flow the torsion tensor satisfies a parabolic PDE.

Despite this difference there are similarities between the results and expectations of pluriclosed flow and Chern–Ricci flow, especially in the cases of positive Kodaira dimension, where it is natural to rescale by "blowing down" the flow, which then yields $d\omega_t = e^{-t} d\omega_0$, allowing for convergence to Kähler metrics. Specifically, the analogue of Conjectures 6.3 and 7.1 below were shown to hold for Chern–Ricci flow ([97, Theorem 1.7], [98, Theorem 1.1]).

Differences start to appear in Kodaira dimension zero. For instance, as shown in [34], given an arbitrary Hermitian metric on the torus, the Chern–Ricci flow will exist globally and converge to a *Chern–Ricci flat*, but not necessarily *flat*, metric. Here the fact that $d\omega_t = d\omega_0$ prevents the flow converging to a Kähler metric when starting from a non-Kähler metric in this setting. This is related to the fact that there is an infinite-dimensional moduli space of pluriclosed Chern–Ricci flat metrics on the torus, using perturbations of the flat metric via $\overline{\partial}\alpha + \partial\overline{\alpha}$. Alternatively, Theorem 4.4 shows that the pluriclosed flow on the torus, with arbitrary initial data, exists globally and converges to a flat metric.

The differences become even more stark for Kodaira dimension $-\infty$, where for instance pluriclosed flow has fixed points on Hopf surfaces, whereas Chern–Ricci flow always encounters a finite time singularity. Also on $\mathbb{CP}^2$ the Fubini–Study metric is stable for normalized pluriclosed flow (also see Theorem 9.4 below for more general convergence results), whereas normalized Chern–Ricci flow will satisfy $d\omega_t = e^t d\omega_0$, and so cannot converge to Fubini–Study for non-Kähler initial data. In a strange quirk, the geometry of Inoue surfaces again requires a blowdown to obtain a geometric limit, so here again the expectations between the two flows agree, and the natural analogue of Conjecture 7.5 for Chern–Ricci flow was shown to hold for certain initial data [24].

2.2. Pluriclosed flow as a gradient flow

The original motivation for defining pluriclosed flow came from complex geometry, aiming to preserve natural properties of Hermitian metrics. As it turns out, this equation has a remarkable and useful relationship to the Ricci flow or, more precisely, the generalized Ricci flow which couples to the heat equation for a closed three-form.

Theorem 2.6 (cf. [83, Theorem 6.5]). *Let (M^{2n}, ω_t, J) be a solution to pluriclosed flow. Let $(g_t, H_t = d^c\omega_t)$ be the associated 1-parameter families of Riemannian metrics and Bismut torsion forms. Then*

$$\begin{aligned} \frac{\partial}{\partial t} g =& \ -2\operatorname{Rc} + \tfrac{1}{2} H^2 - L_{\theta^\sharp} g, \\ \frac{\partial}{\partial t} H =& \ \Delta_d H - L_{\theta^\sharp} H. \end{aligned} \tag{2.4}$$

Notice that we may apply a family of diffeomorphisms to remove the Lie derivative terms, yielding a solution of the system of equations

$$\begin{aligned} \frac{\partial}{\partial t} g =& \ -2\operatorname{Rc} + \tfrac{1}{2} H^2, \\ \frac{\partial}{\partial t} H =& \ \Delta_d H. \end{aligned} \tag{2.5}$$

This system of equations originally arose in mathematical physics in the context of renormalization group flow in the theory of σ-models. The author studied this system under the name "connection Ricci flow" [74] and "generalized Ricci flow" [76] due to the relationship to the curvature of the Bismut connection and generalized geometry. Notice that in the context of pluriclosed flow, if we apply the relevant gauge transformation to produce a solution (g_t, H_t) to (2.5), then g_t remains a pluriclosed metric, but with respect to the appropriately gauge-modified family of complex structures. This seemingly minor point is actually an essential and rich feature of the generalized Kähler–Ricci flow, explained in §9. But first, the primary analytic consequence of Theorem 2.6 is the realization of pluriclosed flow as a gradient flow. As shown in [62], (2.5) is the gradient flow of the first eigenvalue of a certain Schrödinger operator, extending Perelman's monotonicity formulas for Ricci flow. We briefly describe this construction, and the reader should consult ([62, 63]) for further detail.

To begin we define the functional

$$\mathcal{F}(g, H, f) = \int_M \left(R - \tfrac{1}{12} |H|^2 + |\nabla f|^2 \right) e^{-f} dV_g, \tag{2.6}$$

for $f \in C^\infty(M)$. This functional obeys a monotonicity formula when f obeys the appropriate conjugate heat equation, which is the adjoint of the heat equation with respect to the spacetime L^2 norm, in particular

$$\frac{\partial}{\partial t} f = \ -\Delta f - R + \tfrac{1}{4} |H|^2 . \tag{2.7}$$

Given (g_t, H_t) a solution of generalized Ricci flow, and f_t is an associated solution of (2.7), then one obtains the equation

$$\frac{d}{dt}\mathcal{F}(g_t, H_t, f_t) = \int_M \left[2\left|\mathrm{Rc} - \tfrac{1}{4}H^2 + \nabla^2 f\right|^2 + \left|d^* H + \nabla f \lrcorner H\right|^2\right] e^{-f} dV_g. \quad (2.8)$$

The right-hand side is of course nonnegative, so that $\mathcal{F}$ is monotone nondecreasing. Underlying this monotonicity is the fact that generalized Ricci flow is the gradient flow of the lowest eigenvalue λ of the operator $-4\Delta + R - \frac{1}{12}\left|H\right|^2$, characterized via

$$\lambda(g, H) := \inf_{\left\{f \mid \int_M e^{-f} dV_g = 1\right\}} \mathcal{F}(g, H, f). \quad (2.9)$$

Theorem 2.7 ([62, Proposition 3.4]). *Generalized Ricci flow is the gradient flow of* λ.

These monotonicity formulas have strong consequences for the singularity formation and conjectural framework for pluriclosed flow. In particular, this monotonicity formula motivates a key concept, that of a *generalized Ricci soliton*. These are triples (g, H, f) such that

$$\begin{aligned} \mathrm{Rc} - \tfrac{1}{4}H^2 + \nabla^2 f &= 0 \\ d^* H + \nabla f \lrcorner H &= 0, \end{aligned} \quad (2.10)$$

which by (2.8) correspond to critical points for $\mathcal{F}$ or λ. Given such a triple (g, H, f), the solution to generalized Ricci flow with this initial data evolves via pullback by the 1-parameter family of diffeomorphisms generated by $-\nabla f$, and so these solutions are self-similar, and thus a natural, more general notion of a fixed point for generalized Ricci flow.

3. Conjectural existence properties

In this section we describe the conjectural maximal smooth existence time for pluriclosed flow on compact complex manifolds, and give a more refined picture in the case of complex surfaces, following [83]. We first recall the fundamental theorem of Tian–Zhang [95] on the maximal smooth existence of Kähler–Ricci flow. We then extend the formal definition behind this result to the case of pluriclosed flow, and state the relevant conjectures. To finish we recall a result of [83] giving a characterization of the relevant positive cone in cohomology on complex surfaces which allows us to explicitly compute the formal existence time for the flow in part II.

3.1. Sharp local existence for Kähler–Ricci flow

To understand the formal picture of the long time existence and singularity formation of Kähler–Ricci flow, we first study the flow at the level of cohomology.

Definition 3.1. Let (M^{2n}, J) be a Kähler manifold. Let

$$H^{1,1}_{\mathbb{R}} := \frac{\left\{\operatorname{Ker} d : \Lambda^{1,1}_{\mathbb{R}} \to \Lambda^3\right\}}{\left\{\sqrt{-1}\partial\overline{\partial} f \mid f \in C^\infty\right\}}.$$

Furthermore, define the *Kähler cone* via

$$\mathcal{K} := \left\{[\psi] \in H^{1,1} \mid \exists\, \omega \in [\psi],\ \omega > 0\right\}.$$

The structure of the Kähler cone plays a fundamental role in understanding the singularity formation of the Kähler–Ricci flow. First note that an elementary consequence of the Kähler–Ricci flow equation is that

$$[\omega_t] = [\omega_0] - tc_1. \tag{3.1}$$

Thus the Kähler class moves along a ray in $\mathcal{K}$, and we thus obtain an upper bound for the possible smooth existence time:

Lemma 3.2. *Let (M^{2n}, g_0, J) be a Kähler manifold. Let*

$$\tau^*(\omega_0) := \sup\{t \geq 0 \mid [\omega_0] - tc_1 \in \mathcal{K}\}, \tag{3.2}$$

and let T denote the maximal smooth existence time for the Kähler–Ricci flow with initial condition g_0. Then $T \leq \tau^*(\omega_0)$.

Proof. Let ω_t denote the one parameter family of Kähler forms evolving by Kähler–Ricci flow with initial condition ω_0. If the flow existed smoothly for some time $t > \tau^*$, then in particular by (3.1) there exists a smooth positive definite metric in $[\omega_0] - tc_1$, contradicting the definition of τ^*. □

Informally this lemma states that the flow must go singular by the time the associated family of Kähler classes leaves the Kähler cone. In view of this, the natural question to ask is if singularities can possibly form without leaving the Kähler cone. The answer is no, due to Tian–Zhang, meaning that $\tau^*(\omega_0)$ is the maximal smooth existence time of the flow.

Theorem 3.3 ([95, Proposition 1.1]). *Let (M^{2n}, g_0, J) be a compact Kähler manifold. The maximal smooth solution of Kähler–Ricci flow with initial condition g_0 exists on $[0, \tau^*(\omega_0))$.*

The proof requires the development of a priori estimates for the metric along the flow. The fundamental role played by the formal considerations above is that, for times $t < \tau^*(\omega_0)$, it is possible to reduce the Kähler–Ricci flow to a scalar PDE modeled on the parabolic complex Monge–Ampère equation. This allows for various delicate applications of the maximum principle to obtain control over the metric as long as $t < \tau^*(\omega_0)$.

3.2. A positive cone and conjectural existence for pluriclosed flow

Now we follow the discussion of the previous subsection and investigate the formal existence time for pluriclosed flow. First we define the relevant positive cone, this time in an Aeppli cohomology space.

Definition 3.4. Let (M^{2n}, J) be a complex manifold. Define the *real* $(1,1)$ *Bott–Chern cohomology* via

$$H^{1,1}_{BC,\mathbb{R}} := \frac{\left\{ \operatorname{Ker} d : \Lambda^{1,1}_{\mathbb{R}} \to \Lambda^{3}_{\mathbb{R}} \right\}}{\left\{ \sqrt{-1}\partial\overline{\partial} f |\ f \in C^{\infty} \right\}}.$$

Furthermore, define the *real* $(1,1)$ *Aeppli cohomology* via

$$H^{1,1}_{\partial+\overline{\partial},\mathbb{R}} := \frac{\left\{ \operatorname{Ker} \sqrt{-1}\partial\overline{\partial} : \Lambda^{1,1}_{\mathbb{R}} \to \Lambda^{2,2}_{\mathbb{R}} \right\}}{\left\{ \partial\overline{\alpha} + \overline{\partial}\alpha \mid \alpha \in \Lambda^{1,0} \right\}}.$$

The restriction to real $(1,1)$ forms make these different spaces from what is usually referred to as Bott–Chern and Aeppli cohomology. Nonetheless in what follows we will drop the $\mathbb{R}$ from the notation and always mean these spaces defined above. Lastly, we define the $(1,1)$ *Aeppli positive cone* via

$$\mathcal{P} := \left\{ [\psi] \in H^{1,1}_{\partial+\overline{\partial}} \mid \exists\ \omega \in [\psi],\ \omega > 0 \right\}.$$

Note that $\mathcal{P}$ consists precisely of the $(1,1)$ Aeppli classes represented by pluriclosed metrics. Similarly to (3.1), we want to derive an ODE for the Aeppli classes associated to a solution of pluriclosed flow. First note that, on a general complex manifold, the first Chern class c_1 is usually considered as an element of $(1,1)$ Bott–Chern cohomology. There is a natural map $i : H^{1,1}_{BC} \to H^{1,1}_{\partial+\overline{\partial}}$, and using this we consider c_1 as an element of $(1,1)$ Aeppli cohomology, without further notation. Thus, given a solution ω_t of pluriclosed flow (2.3), we compute as an equation of $(1,1)$ Aeppli classes,

$$\frac{d}{dt}[\omega_t] = -[\rho_B^{1,1}] = [\partial\partial^*_{\omega_t}\omega_t + \overline{\partial\partial}^*_{\omega_t}\omega_t - \rho_C(\omega_t)] = -c_1.$$

Thus precisely as in the Kähler–Ricci flow case we obtain, as an equation of $(1,1)$ Aeppli classes,

$$[\omega_t] = [\omega_0] - tc_1. \tag{3.3}$$

This formal calculation allows us to derive an upper bound for the maximal smooth existence time of a solution to pluriclosed flow, whose proof is identical to that of Lemma 3.2.

Lemma 3.5. *Let (M^{2n}, g_0, J) be a complex manifold with pluriclosed metric. Let*

$$\tau^*(\omega_0) := \sup\{t \geq 0 \mid [\omega_0] - tc_1 \in \mathcal{P}\}, \tag{3.4}$$

and let T denote the maximal smooth existence time for the pluriclosed flow with initial condition g_0. Then $T \leq \tau^*(\omega_0)$.

What follows is the main conjecture guiding the study of pluriclosed flow. While Lemma 3.5 indicates the elementary fact that the maximal existence time for the flow can be no larger than $\tau^*(\omega_0)$, Conjecture 3.6 indicates that the flow is actually smooth up to time $\tau^*(\omega_0)$, i.e., that it equals the maximal existence time.

Conjecture 3.6 (Main Existence Conjecture). *Let (M^{2n}, g_0, J) be a compact complex manifold with pluriclosed metric. The maximal smooth solution of pluriclosed flow with initial condition g_0 exists on $[0, \tau^*(\omega_0))$.*

This conjecture first appeared in our joint work with Tian ([83, Conjecture 5.2]), inspired by Theorem 3.3.

3.3. Characterizations of positive cones

As a guide for the nature of singularity formation as one leaves the positive cone it is useful to have a characterization of the necessary and sufficient conditions for cohomology classes to lie in the appropriate positive cone. We give such a characterization in the case $n = 2$ in this subsection. First, for a given complex surface define

$$\Gamma = \frac{\{da \in \Lambda^{1,1}_{\mathbb{R}}\}}{\left\{\sqrt{-1}\partial\overline{\partial} f | f \in C^{\infty}\right\}}.$$

By ([86, Lemma 2.3]), This is identified with a subspace of $\mathbb{R}$, via the L^2 inner product with a pluriclosed metric ω. Note that if (M^4, J) is Kähler then the $\partial\overline{\partial}$-lemma holds and so $\Gamma = \{0\}$. Further arguments of ([86, Lemma 2.3]) in fact show that the vanishing of Γ implies the manifold is Kähler. Thus, on a non-Kähler surface we may choose a positive generator γ_0 for Γ, and since the space of pluriclosed metrics on M is connected, this orientation is well defined. This form γ_0 plays a key role in the characterization of the positive cone $\mathcal{P}$ in the next theorem.

Theorem 3.7 ([83, Theorem 5.6]). *Let (M^4, J) be a complex non-Kähler surface. Suppose $\phi \in \Lambda^{1,1}$ is pluriclosed. Then $[\phi] \in \mathcal{P}$ if and only if*

1. $\int_M \phi \wedge \gamma_0 > 0$
2. $\int_D \phi > 0$ *for every effective divisor with negative self-intersection.*

A natural question is whether there is a characterization of $\mathcal{P}$ in higher dimensions. In such cases it is not even clear how to define natural conditions which only depend on Aeppli cohomology classes. As the quantity $\int_M \omega \wedge \gamma_0$ will be fixed along a solution to pluriclosed flow, as a corollary of this theorem we obtain a clean characterization of the quantity τ^*:

Corollary 3.8. *Let (M^4, J) be a complex non-Kähler surface. Given ω_0 a pluriclosed metric, one has*

$$\tau^*(\omega_0) = \sup\left\{t \geq 0 \mid \int_D \omega_0 - tc_1 > 0 \text{ for } D^2 < 0\right\}.$$

4. (1,0)-form reduction

As described above, the proof of Theorem 3.3 rests on the key fact that the Kähler–Ricci flow, for times $t < \tau^*(\omega_0)$, the Kähler–Ricci flow can be reduced to a scalar equation. Due to the $\partial\overline{\partial}$-lemma, we know that locally Kähler metrics admit scalar potential functions, and thus one expects any variation of Kähler metrics to reduce to a variation of Kähler potentials, up to global topological considerations. Locally this PDE is modeled on the parabolic complex Monge–Ampère equation,

$$\frac{\partial f}{\partial t} = \log\det\sqrt{-1}\partial\overline{\partial}f, \tag{4.1}$$

where $\omega = \sqrt{-1}\partial\overline{\partial}f$. Using the scalar reduction, various maximum principle arguments are employed to obtain C^∞ estimates for f, and thus the Kähler–Ricci flow ([95]).

Turning to the pluriclosed flow, we first note that in non-Kähler geometry, Hermitian metrics, even pluriclosed, cannot be described by a single potential function. Thus it is not reasonable to expect to reduce the pluriclosed flow to a scalar PDE. Instead, pluriclosed metrics admit local potential $(1,0)$ forms. In particular, a short argument using the Dolbeault lemma shows that for any pluriclosed metric, locally there exists a $(1,0)$-form α such that $\omega = \overline{\partial}\alpha + \partial\overline{\alpha}$. Thus for a solution to pluriclosed flow, locally we can express $\omega_t = \omega_{\alpha_t} := \overline{\partial}\alpha_t + \partial\overline{\alpha}_t$, and using (2.3) one can show that α_t should locally satisfy the PDE

$$\frac{\partial\alpha}{\partial t} = \overline{\partial}^*_{g_\alpha}\omega_\alpha - \frac{\sqrt{-1}}{2}\partial\log\det g_\alpha. \tag{4.2}$$

Observe that this is a strict generalization of (4.1), where if the metric is Kähler and $\alpha_t = \frac{\sqrt{-1}}{2}\partial f$, then the PDE for α corresponds to that satisfied by the gradient of a function evolving by (4.1). As it turns out, equation (4.2) is degenerate parabolic, with the degeneracy arising from the redundancy wherein α and $\alpha + \partial f$ describe the same metric for $f \in C^\infty(M,\mathbb{R})$.

A full, positive resolution of Conjecture 3.6 will hinge on a complete understanding of equation (4.2). Using this reduced equation, we have achieved many global existence and convergence results which confirm Conjecture 3.6 in a variety of cases, and these are described below. We will not give a full account of the proofs of these results here, but rather describe one key estimate underlying all of these proofs. In establishing regularity of Kähler–Ricci flow, it is crucial to obtain the $C^{2,\alpha}$ estimate for the potential in the presence of $C^{1,1}$ estimates. This can be achieved using Calabi/Yau's C^3 estimate ([10, 103]) or the Evans–Krylov style of estimates ([23, 52]). On the other hand, thinking in terms of the metric tensor, this is a C^α estimate in the presence of uniform parabolicity bounds. As the pluriclosed flow is a parabolic system of equations for the Hermitian metric g, an estimate of this kind would be analogous to the DeGiorgi–Nash–Moser/Krylov–Safonov [14, 53, 54, 59, 61] estimate for uniformly parabolic equations. However, these results are false for general *systems* of equations [15]. Despite these challenges we are able to obtain this an estimate of this kind by uncovering a kind of convexity structure described below.

We give an informal statement of the result here, which combines ([75, Theorems 1.7, 1.8]), referring there for the precise claims.

Theorem 4.1. *Let (M^{2n}, J) be a compact complex manifold. Suppose g_t is a solution to the pluriclosed flow on $[0, \tau)$. Suppose there exist a constant $\lambda > 0$ such that*

$$\lambda g_0 \leq g_t.$$

Then there exist uniform C^∞ estimates for g_t on $[0, \tau)$.

The crucial point behind this theorem is a sharp differential inequality for a delicate combination of first derivatives of the potential α. In particular, consider the section of $\mathrm{Sym}^2(T \oplus T^*)$ defined by

$$W = \begin{pmatrix} g - \partial\alpha g^{-1}\overline{\partial\alpha} & \sqrt{-1}\partial\alpha g^{-1} \\ \sqrt{-1}g^{-1}\overline{\partial\alpha} & g^{-1} \end{pmatrix}.$$

The form of this matrix is inspired in part by Legendre transforms [84], and in part from its appearance as the "generalized metric," in generalized geometry [35], where $i\partial\alpha$ plays the role of the "B-field." A delicate computation shows that $LW \leq 0$, where L is the Laplacian of the time-dependent metric. Using this together with the fact that $\det W = 1$ it is possible to obtain a C^α estimate for the matrix W, which yields C^α estimates for g. Via a blowup argument one then obtains all higher-order estimates.

Thus using this theorem we see that the remaining obstacle to establish Conjecture 3.6 is to obtain a uniform lower bound for the metric along the flow. This is a significant hurdle, which is overcome in the Kähler setting through a delicate combination of maximum principles for the parabolic complex Monge–Ampère equation. There are many cases where Conjecture 3.6 can be established by exploiting some further underlying geometric structures. The simplest case to address is that of manifolds with nonpositive bisectional curvature:

Theorem 4.2 ([75, Theorem 1.1]). *Suppose (M^4, J, h) is a compact complex surface, with Hermitian metric h with nonpositive holomorphic bisectional curvature. Given g_0 a pluriclosed metric on M, the solution to normalized pluriclosed flow exists for all time.*

In general the limiting behavior at infinity in the above theorem can be delicate, involving contraction of divisors as well as collapsing. With a further curvature restriction we can obtain a convergence result as well:

Theorem 4.3 ([75, Theorem 1.1]). *Suppose (M^4, J, h) is a compact complex surface, with Hermitian metric h with constant negative bisectional curvature. Given g_0 a pluriclosed metric on M, the solution to normalized pluriclosed flow exists for all time and converges to h.*

Also, we can obtain convergence to a flat metric in the case of tori, which requires further a priori estimates and use of the $\mathcal{F}$-functional described in §2.2.

Theorem 4.4 ([75, Theorem 1.1]). *Let (M^4, J) be biholomorphic to a torus. Given g_0 a pluriclosed metric, the solution to pluriclosed flow with initial data g_0 exists on $[0, \infty)$ and converges to a flat Kähler metric.*

This theorem confirms the basic principle that pluriclosed flow cannot develop "local" singularities. For instance, on any manifold one can choose initial metrics whereby the Ricci flow develops topologically trivial neckpinches. The rigidity of pluriclosed metrics/pluriclosed flow apparently prevents the construction of such singular solutions.

5. Pluriclosed flow of locally homogeneous surfaces

Having described fundamental analytic aspects governing the regularity of pluriclosed flow, we now turn to describing the behavior of locally homogeneous metrics, which in principle give the prototypical behavior for different classes of complex surfaces. As we will see the pluriclosed flow naturally incorporates two classical points of view on canonical geometries: "geometric structures" in the sense of Thurston and (Kähler) Einstein metrics. In this section we recall fundamental aspects of Thurston geometries and how pluriclosed flow of locally homogeneous metrics relates to these structures.

5.1. Wall's classification

In 1985 Wall, directly inspired by the Thurston geometrization conjecture [89], sought to understand complex surfaces through the use of Thurston's "model geometries." By combining elements of Lie theory with results from the Kodaira classification he gave a complete classification of complex surfaces admitting model geometric structure. Following [100] we will briefly recall these results and later describe the relationship of pluriclosed flow to this classification. We recall the fundamental definition:

Definition 5.1. A *model geometry* is a triple (X, g, G) such that (X, g) is a complete, simply connected Riemannian manifold, G is a group of isometries acting transitively on (X, g), and G has a discrete subgroup Γ such that $\Gamma \backslash X$ is compact.

Let us very briefly describe the model geometries up to dimension four, only indicating the space X. In dimension one there is a unique geometry, the real line. In dimension two there are three, namely $S^2, \mathbb{R}^2$, and $\mathbb{H}^2$. In dimension three there are eight geometries, classified by Thurston [90]. First there are $S^3, \mathbb{R}^3$, and $\mathbb{H}^3$, as well as the product geometries $\mathbb{S}^2 \times \mathbb{R}$ and $\mathbb{H}^2 \times \mathbb{R}$. Another example is $\widetilde{SL}_2$, the universal cover of the unit tangent bundle of $\mathbb{H}^2$, with metric left-invariant under the natural Lie group structure. Also one has Nil^3, the unique simply connected nilpotent Lie group in three dimensions, as well as a solvable Lie group Sol^3, realized as an extension $\mathbb{R}^2 \ltimes_\alpha \mathbb{R}$, where $\alpha(t)(x,y) = (e^t x, e^{-t} y)$, again with left invariant metric.

In four dimensions these geometries were classified by Filipkiewicz [26]. First there are the irreducible symmetric spaces $S^4, \mathbb{H}^4, \mathbb{CP}^2, \mathbb{CH}^2$, as well as products of all lower-dimensional examples above. There is a four-dimensional nilpotent Lie group Nil^4 and a family of solvable Lie groups $\mathrm{Sol}^4_{m,n}$ which we will not describe as these do not admit compatible complex structures. Another class of solvable Lie groups arises, $\mathrm{Sol}^4_0 = \mathbb{R}^3 \ltimes_\delta \mathbb{R}$, where $\delta(t)(x,y,z) = (e^t x, e^t y, e^{-2t} z)$. Lastly one has Sol^4_1 which is the Lie group of matrices

$$\left\{ \begin{pmatrix} 1 & b & c \\ 0 & \alpha & a \\ 0 & 0 & 1 \end{pmatrix} \ : \ \alpha, a, b, c \in \mathbb{R}, \alpha > 0 \right\}.$$

Another geometry denoted F^4 is identified which we ignore as it admits no compact models.

In describing the existence of compatible complex structures on these geometries, some subtleties arise. We recall the main theorems of Wall [100] here.

Theorem 5.2. 1. *A model geometry X carries a complex structure compatible with the maximal connected group of isometries if and only if X is one of:*

$$\mathbb{CP}^2,\ \mathbb{CH}^2,\ S^2 \times S^2,\ S^2 \times \mathbb{R}^2,\ S^2 \times \mathbb{H}^2,\ \mathbb{R}^2 \times \mathbb{H}^2,$$
$$\mathbb{H}^2 \times \mathbb{H}^2,\ \widetilde{SL}_2 \times \mathbb{R}, \mathrm{Nil}^3 \times \mathbb{R},\ \mathrm{Sol}^4_0,\ \mathrm{Sol}^4_1.$$

2. $\mathbb{R}^4$ *admits a complex structure compatible with* $\mathbb{R}^4 \ltimes U_2$, *and* $S^3 \times \mathbb{R}$ *admits a complex structure compatible with* $U_2 \times \mathbb{R}$. *All other geometries admit no compatible complex structure.*
3. *In every case except* Sol^4_1 *the complex structure is unique up to isomorphism, and on* Sol^4_1 *there are two isomorphism classes. As complex manifolds these are denoted* Sol^4_1 *and* $(\mathrm{Sol}^4_1)'$
4. *There are Kähler metrics compatible with the complex and geometric structure precisely in the cases*

$$\mathbb{CP}^2,\ \mathbb{CH}^2,\ S^2 \times S^2,\ S^2 \times \mathbb{R}^2,\ S^2 \times \mathbb{H}^2,\ \mathbb{R}^2 \times \mathbb{H}^2, \mathbb{H}^2 \times \mathbb{H}^2.$$

5.2. Existence and convergence results

For locally homogeneous pluriclosed metrics, the pluriclosed flow will reduce to a system of ODE which greatly simplifies the analysis. We record theorem statements here assuming familiarity with the Kodaira classification of surfaces, described in more detail in part II. First there is a complete description of the long time existence behavior as well as the limiting behavior on the universal cover.

Theorem 5.3 ([7, Theorem 1.1]). *Let g_t be a locally homogeneous solution of pluriclosed flow on a compact complex surface which exists on a maximal time interval $[0,T)$. If $T < \infty$ then the complex surface is rational or ruled. If $T = \infty$ and the manifold is a Hopf surface, the evolving metric converges exponentially fast to a* *multiple of the Hopf metric (cf. 7.1). Otherwise, for the lift $\widetilde{g}$ of the flow to the universal cover there is a blowdown limit*

$$(\widetilde{g}_\infty)_t = \lim_{s \to \infty} s^{-1} \widetilde{g}_{st}$$

which is an expanding soliton in the sense that $\widetilde{g}_t = t\widetilde{g}_1$.

It is implicit in the proof that one can actually recover the underlying complex surface from the asymptotic behavior of the flow. In particular, if a locally homogeneous solution to pluriclosed flow encounters a finite time singularity, then the singularity is of type I and the underlying complex surface is rational or ruled. If it is type IIb, i.e., $|\mathrm{Rm}|\, t \to \infty$, then the underlying manifold is a diagonal Hopf surface, whereas if it is type III, i.e., $|\mathrm{Rm}|\, t < \infty$, then the underlying manifold is one of the remaining surfaces in the Wall classification, i.e., a torus, hyperlliptic, Kodaira, or Inoue surface. In fact, much sharper statements can be made concerning the Gromov–Hausdorff limits at infinity which recover precisely the underlying complex surface.

Theorem 5.4 ([7, Theorem 1.2]). *Let ω_t be a locally homogeneous solution of pluriclosed flow on a compact complex surface (M, J) which exists on $[0, \infty)$ and suppose (M, J) is not a Hopf surface. Let $\hat{\omega}_t = \frac{\omega_t}{t}$.*

1. *If the surface is a torus, hyperelliptic, or Kodaira surface, then the family $(M, \hat{\omega}_t)$ converges as $t \to \infty$ to a point in the Gromov–Hausdorff topology.*
2. *If the surface is an Inoue surface, then the family $(M, \hat{\omega}_t)$ converges as $t \to \infty$ to a circle in the Gromov–Hausdorff topology and moreover the length of this circle depends only on the complex structure of the surface.*
3. *If the surface is a properly elliptic surface where the genus of the base curve is at least 2, then the family $(M, \hat{\omega}_t)$ converges as $t \to \infty$ to the base curve with constant curvature metric in the Gromov–Hausdorff topology.*
4. *If the surface is of general type, then the family $(M, \hat{\omega}_t)$ converges as $t \to \infty$ to a product of Kähler–Einstein metrics on M.*

Part II: Geometrization of complex surfaces

First we recall the rudimentary aspects of the Kodaira classification of surfaces, referring the reader to the classic text [5] for a very detailed accounting. The first main tool for classifying complex surfaces is the Kodaira dimension. If (M^4, J) is a complex surface, let K denote the canonical bundle, and let $p_n = \dim H^0(K^{\otimes n})$ denote the plurigenera of M. If all p_n are zero, we say that the Kodaira dimension of M is $\mathrm{kod}(M) = -\infty$. Otherwise $\mathrm{kod}(M)$ is the smallest integer k such that $\frac{p_n}{n^k}$ is bounded, which is no greater than 2 in this case, giving the four possibilities, $-\infty, 0, 1$ and 2. Within these four classes one wants to understand which surfaces admit Kähler metrics. Applying the Kähler identities to Dolbeault cohomology, one obtains that the odd Betti numbers of a Kähler surface are even. Building on many results in the Kodaira classification, Siu [67] proved the converse: if $b_1(M)$ is even then the surface admits Kähler metrics. Later, direct proofs of this equivalence were given by Buchdahl [9], Lamari [55]. This leaves in principle eight classes of surfaces, although it follows from Grauert's ampleness criterion (cf. [5, IV.6]) that any surface with $\mathrm{kod}(M) = 2$ is projective, and hence Kähler, thus non-Kähler surfaces only occur for $\mathrm{kod}(M) \leq 1$, leaving seven classes. We will address each of these classes in turn, and state the conjectural behavior of pluriclosed flow for each case.

6. Conjectural limiting behavior on Kähler surfaces

It is natural to ask why one would bother studying pluriclosed flow on a Kähler surface, since one can detect via topological invariants whether a given complex surface admits Kähler metrics, and thus simply study Kähler–Ricci flow on these manifolds to produce canonical geometries. However, as we ultimately want to apply pluriclosed flow as an a priori geometrizing process relying on minimal hypotheses, understanding its properties with arbitrary initial data even on Kähler manifolds plays an important philosophical role. Moreover, Kähler surfaces are a rich class of complex manifolds on which to test the naturality and tameness of pluriclosed flow from a PDE point of view. Most importantly, as we will see in part III, there are concrete applications to understanding the classification of generalized Kähler manifolds that cannot be approached through the Kähler–Ricci flow.

Having said this, the guiding principle here is quite simple:

"*Pluriclosed flow behaves like Kähler–Ricci flow on Kähler manifolds.*"

To illustrate, note that Theorems 4.3 and 4.4 yield the same existence time, and converge to the same limits as, the Kähler–Ricci flow in those settings. The conjectures to follow below all follow this principle.

Remark 6.1. In the discussion below we make the assumption that the underlying complex surface is *minimal*, i.e., free of (-1)-curves. By a standard argument ([5, Theorem III.4.5]) one can perform a finite sequence of blowdowns to obtain a minimal complex surface. Thus from the point of view of complex geometry little is lost by considering only minimal surfaces. Of course from the point of view of analysis one still would like to know what happens to the flow in the general setting. An elementary calculation using the adjunction formula shows that pluriclosed flow homothetically shrinks the area of (-1) curves to zero in finite time. Conjecturally one expects that pluriclosed flow "performs the blowdown" in an appropriate sense. This has been confirmed in some special cases for Kähler–Ricci flow [71].

6.1. Surfaces of general type

By definition these are complex surfaces with Kodaira dimension two. These surfaces have $c_1^2(M) > 0$ and so it follows from Grauert's ampleness criterion (cf. [5, IV.6]) that all such are automatically projective, hence Kähler. In the case $c_1 < 0$

the existence of a Kähler–Einstein metric follows from the work of Aubin–Yau, and the construction of this metric via Kähler–Ricci flow follows from Cao [11].

Theorem 6.2 (Aubin–Yau [4, 103], Cao [11]). *Let (M^4, J) be a compact Kähler surface with $c_1 < 0$. There exists a unique metric $\omega_{\mathrm{KE}} \in -c_1$ satisfying $\rho_{\omega_{\mathrm{KE}}} = -\omega_{\mathrm{KE}}$. Moreover, given any Kähler metric ω_0, the solution to Kähler–Ricci flow with initial condition ω_0 exists on $[0, \infty)$, and*

$$\lim_{t\to\infty} \frac{\omega_t}{t} = \omega_{\mathrm{KE}}.$$

More generally, these surfaces can have $c_1 \leq 0$, admitting finitely many -2 curves, and their canonical models are orbifolds obtained by contraction of these curves. Tian–Zhang [95] proved in this setting that the Kähler–Ricci flow exists globally, converging after normalization to the unique oribfold Kähler–Einstein metric on the canonical model.

To describe the conjectured behavior of the pluriclosed flow, we first compute the formal existence time. By choosing a background metric with $\rho(\widetilde{\omega}) \leq 0$, given any pluriclosed metric ω_0 it follows that $\omega_0 - t\rho(\widetilde{\omega}) > 0$, and so $\tau^*(\omega_0) = \infty$. Thus we expect the pluriclosed flow to exist globally on such manifolds, and converge after normalization to the unique Kähler–Einstein metric on the canonical model. Theorem 4.3 confirms this conjecture for a large class of surfaces of general type.

Conjecture 6.3. *Let (M, J) be a complex surface of general type. Given ω_0 a pluriclosed metric, the solution to pluriclosed flow with initial condition ω_0 exists on $[0, \infty)$, and the solution to the normalized pluriclosed flow exists on $[0, \infty)$ and converges exponentially to ω_{KE}, the unique orbifold Kähler–Einstein metric on the associated canonical model.*

6.2. Properly Elliptic surfaces

By definition these are Kähler complex surfaces of Kodaira dimension one. For such surfaces there is a curve Σ together with a holomorphic map $\pi : M \to \Sigma$ such that the canonical bundle of M satisfies $K = \pi^* L$ for an ample line bundle L over Σ. Further, the generic fiber of π is a smooth elliptic curve, and we call such points regular. Near regular points we obtain a map to the moduli space of elliptic curves, and thus we can pull back the L^2 Weil–Petersson metric to obtain a semipositive $(1, 1)$ form ω_{WP} on the regular set. This form plays a key role in describing the limiting behavior of Kähler–Ricci flow on these surfaces, done by Song–Tian [69]:

Theorem 6.4 ([69, Theorem 1.1]). *Let $\pi : M \to \Sigma$ be a minimal elliptic surface of $\mathrm{kod}(M) = 1$ with singular fibers $M_{s_1} = m_1 F_1, \ldots, M_{s_k} = m_k F_k$ of multiplicity $m_i \in \mathbb{N}$, $i = 1, \ldots, k$. Then for any initial Kähler metric ω_0, the normalized Kähler–Ricci flow with this initial data exists on $[0, \infty)$ and satisfies:*

1. *ω_t converges to $\pi^* \omega_\infty \in -2\pi c_1(M)$ as currents for a positive current ω_∞ on Σ.*

2. *ω_∞ is smooth on Σ_{reg} and $\rho(\omega_\infty)$ is a well-defined current on Σ satisfying*

$$\rho(\omega_\infty) = -\omega_\infty + \omega_{\mathrm{WP}} + 2\pi \sum_{i=1}^{k} \frac{m_k - 1}{m_k}[s_i].$$

Turning to pluriclosed flow, we first note that since K_M is semiample, there exists a background metric with nonpositive Ricci curvature, and using this background metric it is clear that for any pluriclosed metric ω_0 one has $\tau^*(\omega_0) = \infty$. Furthermore, we conjecture the same limiting behavior as the Kähler–Ricci flow:

Conjecture 6.5. *Let $\pi : M \to \Sigma$ be a minimal elliptic surface of* $\mathrm{kod}(M) = 1$. *Given ω_0 a pluriclosed metric, the solution to normalized pluriclosed flow with initial condition ω_0 exists on* $[0, \infty)$, *and satisfies conclusions* (1) *and* (2) *of Theorem* 6.4.

6.3. Elliptic surfaces of Kodaira dimension zero

It follows from the Kodaira classification that a minimal Kähler surface of Kodaira dimension zero is finitely covered by either a torus or a $K3$ surface. In particular, $c_1 = 0$, and a Ricci flat (Calabi–Yau) metric is known to exist in each Kähler class by Yau [103]. The global existence and convergence of Kähler–Ricci flow to a Calabi–Yau metric follows from Cao [11]:

Theorem 6.6 (Yau [103], Cao [11]). *Let (M^4, J) be a compact Kähler surface with $c_1 = 0$. Given $[\omega]$ a Kähler class, there exists a unique metric $\omega_{\mathrm{CY}} \in [\omega]$ satisfying $\rho_{\omega_{\mathrm{CY}}} = 0$. Moreover, given any Kähler metric ω_0, the solution to Kähler–Ricci flow with initial condition ω_0 exists on $[0, \infty)$, satisfies $[\omega_t] \in [\omega]$, and moreover*

$$\lim_{t \to \infty} \omega_t = \omega_{\mathrm{CY}}.$$

For the pluriclosed flow, first note that as $c_1 = 0$, the Aeppli cohomology class is fixed along pluriclosed flow, and so $\tau^*(\omega_0) = \infty$ for any g_0, thus following the general principle we conjecture global existence, as well as convergence to a Calabi–Yau metric.

Conjecture 6.7. *Let (M^4, J) be a compact Kähler surface with* $\mathrm{kod}(M) = 0$. *Given ω_0 a pluriclosed metric on M, the pluriclosed flow with initial condition ω_0 exists on $[0, \infty)$, and converges to a Kähler Calabi–Yau metric ω_{CY}.*

Theorem 4.4 confirms this for tori, while ([77, Theorem 1.1]) establishes the long time existence and weak convergence for certain solutions to pluriclosed flow arising in generalized Kähler geometry (cf. §9 below).

6.4. Rational and ruled surfaces

By the Kodaira classification Kähler surfaces with $\mathrm{kod}(M) = -\infty$ are birational to $\mathbb{CP}^2$. An important case of such surfaces are Fano surfaces, i.e., surfaces with $c_1 > 0$, as these are candidates to admit Kähler–Einstein metrics. It follows from the Enriques–Kodaira classification that the smooth manifolds underlying Fano surfaces are $\mathbb{CP}^1 \times \mathbb{CP}^1$ and $\mathbb{CP}^2 \# k \overline{\mathbb{CP}}^2$, for $0 \leq k \leq 8$. The existence of a Kähler–Einstein metric in this setting is obstructed in general, in particular Matsushima showed the automorphism group must be reductive ([57]), with further obstructions due to Futaki [28]. This rules out $\mathbb{CP}^2 \# k\overline{\mathbb{CP}}^2$ for $k = 1, 2$. For the remaining cases, Tian settled the existence question:

Theorem 6.8 ([91]). *Let (M^4, J) be a compact Fano surface such that the Lie algebra of the automorphism group is reductive. There exists a metric $\omega_{KE} \in c_1$ satisfying $\rho_{\omega_{KE}} = \omega_{KE}$.*

Extending our point of view slightly, we can look for Kähler–Ricci solitons, i.e., solutions of

$$\rho_\omega = \omega + L_X \omega,$$

where X is a holomorphic vector field. Koiso [51] showed the existence of a soliton on $\mathbb{CP}^2 \# \overline{\mathbb{CP}}^2$ using a symmetry ansatz. Later Wang–Zhu [101] showed the existence of Kähler–Ricci solitons on all toric Kähler manifolds of positive first Chern class, in particular covering the case of $\mathbb{CP}^2 \# 2\overline{\mathbb{CP}}^2$.

It is natural to expect the Kähler–Ricci flow to converge to these canonical metrics, when they exist. There is a fairly complete picture of the Kähler–Ricci flow on Fano surfaces confirming this expectation:

Theorem 6.9 ([94, 96]). *Let (M^4, J) be a compact Fano surface admitting a Kähler–Ricci soliton (ω_{KRS}, X). Given $\omega_0 \in c_1$ a Kähler metric which is G_X-invariant, where G_X is the one-parameter subgroup generated by Im(X), the solution to normalized Kähler–Ricci flow exists on $[0, \infty)$ and converges to a Kähler–Ricci soliton.*

For the pluriclosed flow in the Fano setting, the normalized flow will fix the Aeppli cohomology class, and thus we expect global existence of the flow, and convergence to a soliton when one exists.

Conjecture 6.10. *Let (M^4, J) be a compact Fano surface admitting a Kähler–Ricci soliton (ω_{KRS}, X). Given $\omega_0 \in c_1$ a Kähler metric which is G_X-invariant, where G_X is the one-parameter subgroup generated by Im(X), the solution to normalized pluriclosed flow exists on $[0, \infty)$ and converges to a Kähler–Ricci soliton.*

As mentioned in the introduction, this conjecture would have consequences for the classification of generalized Kähler structures, and this will be detailed in part III below. As we describe in Theorem 9.4 below, the global existence of the normalized pluriclosed flow as well as a weak form of convergence can be established on $\mathbb{CP}^2$, which proves a kind of uniqueness for generalized Kähler structure on $\mathbb{CP}^2$.

Beyond Fano surfaces, the behavior of Kähler–Ricci flow more generally on rational and ruled surfaces can be quite delicate. Even on a fixed Hirzebruch surface the Kähler–Ricci flow can exhibit diverse behavior depending on the initial Kähler class. In particular, Song–Weinkove ([71, Theorem 1.1]) confirmed a conjecture of Feldman–Ilmanen–Knopf ([25]) and showed that the flow can collapse $\mathbb{CP}^1$ fibers, shrink to a point, or contract exceptional divisors, depending on the choice of initial Kähler class. More generally Song–Székelyhidi–Weinkove ([68]) established contraction of the fibers for more general ruled surfaces. Rather than delving into these possibilities we refer to the general principle that one expects the same kinds of singularity formation for the pluriclosed flow.

7. Conjectural limiting behavior on non-Kähler surfaces

We now turn to the case of non-Kähler surfaces. As it turns out, most complex surfaces outside of Class VII_+ admit locally homogeneous metrics, and thus admit special solutions to pluriclosed flow as described in Theorems 5.3, 5.4. It is natural to expect that the convergence behavior on such manifolds for arbitrary initial data is the same as that in the homogeneous case, and we formalize and expand on this below. Turning to the case of Class VII_+, we do not have homogeneous solutions available to guide us. Nonetheless based on some examples and carefully inspecting monotonicity formulas for pluriclosed flow, we motivate a conjectural description of all long time solutions, which is loosely related to the theory of complete Calabi–Yau metrics on the complement of canonical divisors in surfaces, pioneered by Tian–Yau [92, 93].

7.1. Properly elliptic surfaces

For elliptic fibrations with b_1 odd and $\mathrm{kod}(M) \geq 0$, it is known (cf. [100, Lemma 7.2], [8, Lemmas 1, 2]) that only multiple fibers can occur, and thus there is a finite covering of M which is an elliptic fiber bundle over a new curve which is a branched cover of the original base curve. Furthermore, the condition $\mathrm{kod}(M) = 1$ is equivalent to the Euler characteristic of the base curve being negative, and these are called properly elliptic surfaces. For these surfaces the canonical bundle is numerically effective, and so it follows from Corollary 3.8 that $\tau^*(\omega_0) = \infty$ for any initial metric, and thus we expect global existence of pluriclosed flow. Due to the lack of singular fibers, the expected limiting behavior in this case is actually simpler than that observed previously for the Kähler–Ricci flow on elliptic surfaces of Kodaira dimension one (cf. Theorem 6.4). Also, for these surfaces it follows from ([100, Theorem 7.4]) that M admits a geometric structure modeled on $\widetilde{SL}_2 \times \mathbb{R}$. Thus in particular Theorem 5.4 (3) gives a class of metrics exhibiting global existence and Gromov–Hausdorff convergence to the base curve. Moreover, by Theorem 5.3, there exists a blowdown limit of the solutions at infinity. These exhibit the interesting behavior that the blowdown limits on the universal covers are invariant metrics on $\mathbb{H} \times \mathbb{C}$, so that the geometric type "jumps" in the limit. This behavior is conjecturally universal.

Conjecture 7.1. *Let (M^4, J) be a properly elliptic surface with $\mathrm{kod}(M) = 1$ and b_1 odd. Given ω a pluriclosed metric on M, the solution to pluriclosed flow with this initial data exists on $[0, \infty)$, and $(M, \frac{\omega_t}{t})$ converges in the Gromov–Hausdorff topology to $(B, \omega_{\mathrm{KE}})$, the base curve with canonical orbifold Kähler–Einstein metric. On the universal cover of M there is a blowdown limit $\widetilde{\omega}_\infty(t) = \lim_{s\to\infty} s^{-1}\widetilde{\omega}(st)$ which is a locally homogeneous expanding soliton.*

Recently we were able to establish some cases of this conjecture, namely for initial data invariant under the torus action.

Theorem 7.2 ([72]). *Let (M^4, J) be a properly elliptic surface with* $\mathrm{kod}(M) = 1$ *and b_1 odd. Given ω a T^2-invariant pluriclosed metric on M, the solution to pluriclosed flow with this initial data exists on $[0, \infty)$, and $(M, \frac{\omega_t}{t})$ converges in the Gromov–Hausdorff topology to $(B, \omega_{\mathrm{KE}})$, the base curve with canonical orbifold Kähler–Einstein metric. On the universal cover of M there is a blowdown limit $\widetilde{\omega}_\infty(t) = \lim_{s \to \infty} s^{-1} \widetilde{\omega}(st)$ which is a locally homogeneous expanding soliton.*

7.2. Kodaira surfaces

Kodaira surfaces are the complex surfaces with $\mathrm{kod}(M) = 0$ and b_1 odd. These are elliptic fiber bundles over elliptic curves, called primary Kodaira surfaces, or finite quotients of such, called secondary Kodaira surfaces (cf. [5, V.5]). Wall showed ([100, Theorem 7.4]) that these surfaces admit a geometric structure modelled on $\mathrm{Nil}^3 \times \mathbb{R}$. This was refined by Klingler [48], who gave the precise description of the universal covers as well as the presentations of the lattice subgroups. To describe the pluriclosed flow on these surfaces we first compute the formal existence time. Topological considerations (cf. [5, p. 197]) show that the canonical bundle of primary Kodaira surfaces is trivial. Thus $K \cdot D = 0$ for all divisors D, and thus for any pluriclosed metric ω_0 it follows directly from Corollary 3.8 that $\tau^*(\omega_0) = \infty$. Thus we expect global existence of the flow in general. Theorem 5.4 confirms this, and gives the limiting behavior of the flow, in the locally homogeneous setting, which we conjecture to be the general behavior.

Conjecture 7.3. *Let (M^4, J) be a Kodaira surface. Given ω_0 a pluriclosed metric on M, the solution to pluriclosed flow with this initial data exists on $[0, \infty)$, and $(M, \frac{\omega_t}{t})$ converges in the Gromov–Hausdorff topology to a point. On the universal cover of M there is a blowdown limit $\widetilde{\omega}_\infty(t) = \lim_{s \to \infty} s^{-1} \widetilde{\omega}(st)$ which is a locally homogeneous expanding soliton.*

Similar to the case of properly elliptic surfaces, we can exploit the fibration structure to prove the conjecture for invariant initial data.

Theorem 7.4 ([72]). *Let (M^4, J) be a Kodaira surface. Given ω a T^2-invariant pluriclosed metric on M, the solution to pluriclosed flow with this initial data exists on $[0, \infty)$, and $(M, \frac{\omega_t}{t})$ converges in the Gromov–Hausdorff topology to a point. On the universal cover of M there is a blowdown limit $\widetilde{\omega}_\infty(t) = \lim_{s \to \infty} s^{-1} \widetilde{\omega}(st)$ which is a locally homogeneous expanding soliton.*

7.3. Class VII_0 surfaces

Surfaces of Class VII_0 are defined by the conditions $\mathrm{kod}(M) = -\infty$, $b_1 = 1$, and $b_2 = 0$. Hopf surfaces are of this type, and Inoue [43] gave classes of examples, and showed that surfaces in this class with no curves must be one of his examples. Later, using the methods of gauge theory, Li–Yau–Zheng [56] and Teleman [88] showed that all surfaces in this class are either Hopf surfaces or Inoue surfaces. We describe the conjectured behavior on these two classes of surfaces below.

7.3.1. Inoue surfaces. Inoue surfaces were introduced in [43], giving examples of complex surfaces with $b_1 = 1, b_2 = 0$, and containing no curves. These come in three classes, all of which are quotients of $\mathbb{H} \times \mathbb{C}$ by affine subgroups, and we will recall the construction of the simplest of these classes. First, fix $Z \in \mathrm{SL}(3, \mathbb{Z})$, with eigenvalues $\alpha, \beta, \overline{\beta}$ such that $\alpha > 1$ and $\beta \neq \overline{\beta}$. Choose a real eigenvector (a_1, a_2, a_3) for α and an eigenvector (b_1, b_2, b_3) for β. It follows that the three complex vectors (a_i, b_i) are linearly independent over $\mathbb{R}$. Using these we define a group of automorphisms G_Z of $\mathbb{H} \times \mathbb{C}$ generated by

$$\begin{aligned} g_0(w, z) &= (\alpha w, \beta z), \\ g_i(w, z) &= (w + a_i, z + b_i), \qquad i = 1, 2, 3. \end{aligned}$$

This action is free and properly discontinuous, and so defines a quotient surface $S_Z = \mathbb{H} \times \mathbb{C}/G_Z$. If we let G denote the subgroup generated by g_1, g_2, g_3, it is clear that this is isomorphic to $\mathbb{Z}^3$, and preserves the affine varieties

$$(w_0, z_0) + (\mathbb{R}(a_1, b_1) \oplus \mathbb{R}(a_2, b_2) \oplus \mathbb{R}(a_3, b_3)).$$

These spaces are parameterized by $\Im(w_0)$, and thus the quotient $\mathbb{H} \times \mathbb{C}/G$ is diffeomorphic to $T^2 \times \mathbb{R}_+$. Since g_0 preserves the fibers of this quotient, it follows that the quotient S_Z is a three-torus bundle over S^1, with $b_1 = 1$, $b_2 = 0$. Wall shows ([100, Proposition 9.1]) that these surfaces are precisely those with geometric structure Sol_0^4, Sol_1^4, or $(\mathrm{Sol}_1^4)'$. In particular for the example above one sees that each $g_i \in \mathrm{Sol}_0^4$.

To understand the pluriclosed flow on these surfaces, we first compute the formal existence time. As shown by Inoue ([43, §2–§4]), these surfaces contain no curves, thus it follows immediately from Corollary 3.8 that $\tau^*(\omega_0) = \infty$ for any ω_0. This long time existence was verified in the homogeneous setting by Boling (Theorem 5.3), and also in the commuting generalized Kähler setting ([79, Theorem 1.3], cf. Theorem 9.2 below). Boling also established interesting convergence behavior in the homogeneous setting (Theorem 5.4), showing that the pluriclosed flow collapses the three-torus fibers described above, yielding the base circle as the Gromov–Hausdorff limit. Moreover, he showed the existence of a blowdown limit on the universal cover which is an expanding soliton. We conjecture that this is the general behavior.

Conjecture 7.5. *Let (M^4, J) be an Inoue surface. Given ω_0 a pluriclosed metric on M, the solution ω_t to pluriclosed flow with this initial data exists on $[0, \infty)$. The family $(M, \frac{\hat{\omega}_t}{t})$ converges as $t \to \infty$ to a circle in the Gromov–Hausdorff topology and moreover the length of this circle depends only on the complex structure of the surface. On the universal cover of M there is a blowdown limit $\widetilde{\omega}_\infty(t) = \lim_{s \to \infty} s^{-1} \widetilde{\omega}(st)$ which is a canonical locally homogeneous expanding soliton.*

7.3.2. Hopf surfaces. Hopf surfaces by definition are all compact quotients of $\mathbb{C}^2 \backslash \{0\}$. A Hopf surface is called primary if its fundamental group is isomorphic to $\mathbb{Z}$. Since every Hopf surface has a finite cover which is primary, we restrict our discussion to the primary case. By ([50, Theorem 30]), the possible group actions on $\mathbb{C}^2 \backslash \{0\}$ take the form

$$\gamma(z_1, z_2) = (\alpha z_1, \beta z_2 + \lambda z_1^m), \quad \text{where } 0 < |\alpha| \leq |\beta| < 1, \qquad (\alpha - \beta^m)\lambda = 0.$$

We call the resulting quotient surface $(M, J_{\alpha,\beta,\lambda})$. It follows that M is always diffeomorphic to $S^3 \times S^1$, thus $b_1 = 1, b_2 = 0$, and the surfaces are not Kähler. We call the surface Class 1 if $\lambda = 0$, and Class 0 otherwise. Within Class 1 Hopf surfaces we call the surface diagonal if $|\alpha| = |\beta|$.

To describe the conjectured behavior of pluriclosed flow, we begin with a simple example. On $\mathbb{C}^2 \backslash \{0\}$, define the metric

$$\omega_{\mathrm{Hopf}} := \frac{\sqrt{-1}\partial\overline{\partial} \left|z\right|^2}{\left|z\right|^2}. \tag{7.1}$$

Written in this form, it is clear that this is a Hermitian metric which is invariant under $\gamma(z_1, z_2) = (\alpha z_1, \beta z_2)$, where $|\alpha| = |\beta|$, and thus descends to a metric on the diagonal Hopf surfaces. Elementary calculations show that this metric is pluriclosed. Of course $\mathbb{C}^2 \backslash \{0\} \cong S^3 \times \mathbb{R}$, and as it turns out the Riemannian metric corresponding to ω_{Hopf} is isometric to $g_{S^3} \oplus ds^2$, where g_{S^3} is the round metric on S^3 and s is a parameter on $\mathbb{R}$. Further calculations reveal that the torsion tensor H is a multiple of $dV_{g_{S^3}}$, the volume form on S^3. Since $\theta = \star H$ for complex surfaces, it follows that θ is a multiple of ds, and is in particular parallel. Putting these facts together and applying Theorem 2.6, it follows that ω_{Hopf} is a fixed point of pluriclosed flow, and in particular the pair (g, H) satisfies

$$\begin{aligned} \mathrm{Rc} - \tfrac{1}{4} H^2 &= 0, \\ d^* H &= 0. \end{aligned}$$

Thus we see explicitly a basic principle which appears to distinguish the non-Kähler setting from the Kähler setting: for this metric the torsion acts as a “balancing force” which cancels out the positive Ricci curvature of g_{S^3}. Thus this metric is Bismut–Ricci flat, while certainly not Ricci flat. In fact more is true: the Bismut connection is flat. As S^3 is a simple Lie group, there are flat connections $\nabla^{\pm}$ which are compatible with a bi-invariant metric and having torsion $T(X, Y) = \pm[X, Y]$. The connection ∇^+, after taking a direct sum with a flat connection on S^1, recovers the Bismut connection of ω_{Hopf}. It follows from work of Cartan–Schouten ([12, 13]), that triples (M^n, g, H) with flat Bismut connection are isometric to products of simple Lie groups and classically flat spaces, as in this example.

Surprisingly, the metrics described above, which exist on diagonal Hopf surfaces, are the *only* non-Kähler fixed points of pluriclosed flow, a result of Gauduchon–Ivanov ([33, Theorem 2]). The central point is to employ a Bochner argument to show that the Lee form is parallel. Thus it either vanishes identically, in which case the metric is Kähler and Calabi–Yau, or it is everywhere nonvanishing and yields a metric splitting of the universal cover. Looking back at the fixed point equation, one then observes that the geometry transverse to the Lee form has constant positive Ricci curvature, and so is a quotient of the round sphere. Thus the universal cover is isometric to $g_{S^3} \oplus ds^2$. This identifies the metric structure, and using a lemma of Gauduchon ([31, III Lemma 11]) one identifies the possible complex structures, yielding only the diagonal Hopf surfaces.

This rigidity result is surprising since it implies that a naive extension of the Calabi–Yau theorem/Cao's theorem is impossible in the context of pluriclosed flow. In particular, all Hopf surfaces have $b_2 = 0$ and thus $c_1 = 0$, and so $\tau^*(\omega_0) = \infty$ for all ω_0, and so one expects global existence of the pluriclosed flow. However, only the restricted class of diagonal Hopf surfaces admits fixed points. This observation, together with the Perelman-type monotonicity formula, inspired the author to look for soliton-type fixed points of pluriclosed flow, as described in §2. In [78] we were able to construct such soliton solutions of pluriclosed flow on all Class 1 Hopf surfaces, and show that Hopf surfaces are the only compact complex surfaces on which solitons can exist. Note that this is a key qualitative distinction from the case of Ricci flow, where all compact steady solitons are trivial, i.e., Einstein. While we were not able to rule out existence on Class 0 Hopf surfaces, we conjecture that they do not exist. It would be interesting to develop invariants akin to Futaki's invariants for Kähler–Einstein metrics to try to rule them out, or for that matter to rule out fixed points on non-diagonal Class 1 Hopf surfaces via such invariants.

With this background we can now state the conjectured convergence behavior. In particular, on diagonal Hopf surfaces the flow should converge to ω_{Hopf}. More generally, for Class 1 Hopf surfaces one expects convergence to a soliton, which is presumably unique. If it is true that solitons cannot exist on Class 0 Hopf surfaces, one cannot obtain convergence to a soliton in the usual sense. Due to the nature of Cheeger–Gromov convergence, it is however possible for the complex structure to "jump" in the limit. In particular, every Class 1 Hopf surface occurs as the central fiber in a family of Hopf surfaces, all other fibers of which are biholomorphic to the same Class 0 Hopf surface. The diffeomorphism actions necessary in taking Cheeger–Gromov limits on Class 0 surfaces should result in the limiting complex structure jumping to the central fiber, i.e., the associated Class 1 Hopf surface.

Conjecture 7.6. *Let (M^4, J) be a primary Hopf surface. Given ω_0 a pluriclosed metric on M, the solution ω_t to pluriclosed flow exists on $[0, \infty)$.*

1. *If (M, J) is a diagonal Hopf surface, ω_t converges in the C^∞ topology to ω_{Hopf}.*
2. *If $(M, J_{\alpha,\beta})$ is a Class 1 Hopf surface, $(M, J_{\alpha,\beta}, \omega_t)$ converges in the C^∞ Cheeger–Gromov topology to a unique steady soliton $(M, J_{\alpha,\beta}, \omega_S)$.*
3. *If $(M, J_{\alpha,\beta,\lambda})$ is a Class 0 Hopf surface, $(M, J_{\alpha,\beta,\lambda}, \omega_t)$ converges in the C^∞ Cheeger–Gromov topology to a unique steady soliton $(M^4, J_{\alpha,\beta}, \omega_S)$ on the Class 1 Hopf surface adjacent to $J_{\alpha,\beta,\lambda}$ as described above.*

Recently, in line with the results described above for properly elliptic and Kodaira surfaces, we are able to establish partial results confirming this behavior. In this case, for technical reasons, to establish the global existence we require an extra condition, namely that the torsion is nowhere vanishing.

Theorem 7.7 ([72]). *Let (M^4, J) be a diagonal Hopf surface. Given ω_0 a T^2-invariant pluriclosed metric with nowhere vanishing torsion, the solution to pluriclosed flow with this initial data exists on $[0, \infty)$ and converges to ω_{Hopf}.*

7.4. Class VII$_+$ surfaces

Surfaces of Class VII$_+$ are defined by the conditions $\mathrm{kod}(M) = -\infty$, $b_1 = 1$, and $b_2 > 0$. We briefly recall a general construction of surfaces of this type due to Kato [46], building on the initial construction of Inoue [44]. Let Π_0 denote blowup of the origin of the unit ball B in $\mathbb{C}^2$. Let Π_1 denote the blowup of a point $O_0 \in C_0 := \Pi_0^{-1}(0)$. Iteratively let Π_{i+1} denote blowup of a point $O_i \in C_i = \Pi_i^{-1}(O_{i-1})$. Let $\Pi : \hat{B} \to B$ denote the composition of these blowups. Choose a holomorphic embedding $\sigma : \overline{B} \to \hat{B}$ such that $\sigma(0) \in C_k$, the final exceptional divisor. Let $N = \hat{B} \backslash \sigma(\overline{B})$, which has two boundary components $\partial \hat{B}$ and $\sigma(\partial B)$. The map $\sigma \circ \Pi : \partial \hat{B} \to \sigma(\partial B)$ can be used to glue these two boundaries, producing a minimal compact complex surface $M = M_{\pi,\sigma}$, and such surfaces are referred to as *Kato surfaces.* These surfaces are diffeomorphic to $S^3 \times S^1 \# k\mathbb{CP}^2$, but the complex structures are minimal. It was conjectured by Nakamura ([60] Conjecture 5.5) that *all* complex surfaces of Class VII$_+$ are in fact Kato surfaces, and this remains the main open question in the Kodaira classification of surfaces.

One approach to resolving this conjecture focuses on a particular geometric feature shared by all Kato surfaces, that of a *global spherical shell* (GSS). This is a biholomorphism of an annulus in $\mathbb{C}^2 - \{0\}$ into M such that the image does not disconnect M. These are easily seen to exist in Hopf surfaces, and moreover in the construction of Kato surfaces as above, any annulus around the origin in B is a GSS. The relevance of GSS was exhibited by Kato [46], who showed that every surface admitting a global spherical shell is a degeneration of a blown up primary Hopf surface, and moreover is a Kato surface. Later it was shown by Dloussky, Oeljeklaus, and Toma [21] that if a complex surface of Class VII$_+$ admits b_2 complex curves, then it has a global spherical shell, and is hence a Kato surface. Given this, Teleman [85] proved the existence of a curve on all Class VII$_+$ surfaces with $b_2 = 1$, thus finishing their classification. Moreover, when $b_2 = 2$ Teleman [87] showed the presence of a cycle of rational curves, again yielding the classification for this case. This deep work remains the only proof of classification of these surfaces for $b_2 > 0$.

In our initial joint work with Tian [80], we conceived pluriclosed flow in part to address this classification problem. In the follow-up [83], we discussed an argument by contradiction whereby a Class VII$_+$ with no curves at all would violate natural existence conjectures for the pluriclosed flow. However, this line of argument via contradiction still leaves us with the question of what, even conjecturally, the flow may actually do in these settings. First we address the existence time. It follows from ([49, p. 755], [50, p. 683]) that one has the following topological characteristics for any Class VII$_+$ surface:

$$h^{0,1} = 1, \quad h^{1,0} = h^{2,0} = h^{0,2} = 0, \quad b_2^+ = 0, \quad c_1^2 = -b_2. \tag{7.2}$$

Since the intersection form is negative definite, it follows from the adjunction formula that $K \cdot D \geq 0$ for any divisor D, with equality if and only if D is a (-2)-curve. Corollary 3.8 then implies that $\tau^*(\omega_0) = \infty$ for any ω_0 (cf. [83, Proposition 5.7]).

To describe the conjectural limiting behavior, we recall basic facts on the configuration of curves in these surfaces. In the simplest case of parabolic Inoue surfaces, corresponding to a generic sequence of blowups in the construction described above, there is a cycle of (-2) curves C_i satisfying $K \cdot C_i = 0$, as well as an elliptic curve E satisfying $K \cdot E = 1$. In general, by blowing up the same point several times in the Kato construction, one can have rational curves of high negative self-intersection, which then satisfy $K \cdot C > 0$. We define

$$\Sigma_0 := \{\text{curves } C,\ K \cdot C = 0\},$$
$$\Sigma_{>0} := \{\text{curves } C,\ K \cdot C > 0\}.$$

It follows from the adjunction formula that Σ_0 consists of smooth rational (-2) curves, as well as rational curves with an ordinary double point and zero self-intersection.

Given these remarks on the structure of curves, a natural guess arises for the limiting behavior of pluriclosed flow which we now rule out. Along pluriclosed flow the area of curves in Σ_0 will remain fixed, while the area of curves in $\Sigma_{>0}$ will grow linearly. This is similar to the situation for surfaces of general type, where (-2) curves remain fixed and all other curves grow linearly. In that situation, one considers the normalized flow, for which the areas of (-2) curves decay exponentially and the areas of all other curves approach a fixed positive value. It was proved by Tian–Zhang ([95]) that in this setting the Kähler–Ricci flow converges to a Kähler–Einstein metric on the canonical model of the original surface, which is the orbifold given by contraction of all (-2) curves. Thus one might expect a similar picture for pluriclosed flow, with the normalized flow converging to a canonical metric on an orbifold obtained by contraction of the (-2) curves on the original surface. However, it follows from a short calculation (cf. [80, Proposition 3.8]) that along pluriclosed flow the integral of the Chern scalar curvature evolves by $-c_1^2 = b_2 > 0$. Thus for the normalized flow it approaches the value b_2. However, using the expanding entropy functional for generalized Ricci flow ([74]), if the normalized flow converged to some smooth metric on an orbifold, it would be a negative scalar curvature Kähler–Einstein metric, contradicting that the integral is $b_2 > 0$.

Instead, let us argue proceeding from the observation that the integral of Chern scalar curvature is asymptotically $b_2 t$. As this implies that the scalar curvature is becoming positive on average, this would force the volume to go to zero if not for the torsion acting as a "restoring force" as described in §7.3.2. Thus, for a point p where the scalar curvature is positive, bounded and bounded away from zero, one expects to be able to construct a nonflat limit of pointed solutions (M, ω_t, J, p). Arguing formally using the $\mathcal{F}$-functional monotonicity (2.8) one expects this limit to be a steady soliton. In almost every case, the set $\Sigma_{>0}$ is nonempty, and the areas of these curves grow linearly and so form the infinity of the resulting complete metric. In these cases one expects that limits bases at any point $p \notin \Sigma_{>0}$ will yield a nonflat steady soliton, while points in $\Sigma_{>0}$ will yield a flat limit. The exceptional cases are the Enoki surfaces described below which are certain exceptional compactifications of line bundles over elliptic curves. In the special case of parabolic Inoue surfaces the zero section of the line bundle is an elliptic curve, and forms $\Sigma_{>0}$. Outside of this special case the set $\Sigma_{>0}$ is empty, but nonetheless we expect some section of the line bundle to play the role of $\Sigma_{>0}$ as described above. We summarize:

Conjecture 7.8. *Let (M^4, J) be a compact surface of Class VII$_+$. Given ω_0 a pluriclosed metric on M the solution ω_t to pluriclosed flow with this initial data exists on $[0, \infty)$. For a generic point $p \in M$ as described above, the pointed spaces (M, J, ω_t, p) converge in the pointed C^∞ Cheeger–Gromov sense to a nonflat complete steady soliton $(M_\infty, J_\infty, \omega_\infty, p)$ for some smooth function f_∞. Furthermore,*

1. *The vector field $\theta^\sharp_\infty + \nabla f_\infty$ is J_∞-holomorphic.*
2. *M_∞ admits a compactification to a complex surface $(\overline{M}_\infty, \overline{J}_\infty)$.*
3. *The distribution orthogonal to $\theta^\sharp_\infty + \nabla f_\infty$ is integrable, and its generic leaf is an embedded submanifold of M_∞, whose closure in $\overline{M}_\infty$ is a global spherical shell.*

This is a natural extension of Conjecture 7.6. We note that for all steady solitons the vector field $\theta^\sharp + \nabla f$ is automatically holomorphic ([78, Proposition 3.4], which can be extended to the complete setting). Moreover, the twisted Lee form $e^{-f}(\theta + df)$ is automatically closed, and so this distribution is always integrable. Also, for the solitons on Hopf surfaces constructed in [78], one can verify that the leaves of the distribution $\ker e^{-f}(\theta + df)$ are global spherical shells. It would be interesting to determine if this was true for a general, non-Kähler complete steady soliton. Given the nature of the convergence process, assuming the generic leaf is a global spherical shell, this will yield the existence of one on the original complex surface (M, J), thus finishing the classification as described above.

Note that these conjectured limits bear a family resemblance to complete Calabi–Yau metrics, pioneered in work of Tian–Yau [92, 93]. A prototypical result of this kind exhibits a complete Calabi–Yau metric on $M \backslash D$, where M is smooth quasi-projective, and D is a smooth anticanonical divisor with $D^2 \geq 0$. Very loosely speaking the anticanonical divisor is a topological obstruction to the existence of a Ricci flat metric, and by removing it one can construct complete examples. Similarly, in this setting the curves in $\Sigma_{>0}$ form an obstruction to the existence of a "Calabi–Yau" type metric, which the flow naturally pushes to infinity. We note however that complete Calabi–Yau do not arise as limits of the Kähler–Ricci flow on compact complex surfaces. The difference in the expected qualitative behavior can be traced back to the simple fact that $c_1^2 = -b_2 \leq 0$ on these surfaces, which has a profound effect on the existence time and qualitative behavior on Class VII$_+$ surfaces versus other surfaces of Kodaira dimension $-\infty$.

For complete pluriclosed solitons as described, in line with a conjecture of Yau for complete Calabi–Yau metrics [104], it is natural to expect that these complete solitons admit compactifications $(\overline{M}_\infty, \overline{J}_\infty)$. These should be compact complex surfaces with $b_1 = 1$, thus Class VII_+ surfaces, although here again it is possible that the complex structure has "jumped" in the limit so that this compactification need not be biholomorphic to the original surface. Lastly, we note that not every Class VII surface admits holomorphic vector fields, and those that do are classified in [19, 20]. Despite this restriction, the conjectured convergence can still happen since we do not necessarily expect the vector field to extend smoothly to the compactification, and moreover the complex structure can change in the limit as mentioned.

Let us flesh this picture out for specific classes of Kato surfaces. First we consider the Enoki/Inoue surfaces, described as exceptional compactifications of line bundles over elliptic curves. We let $E = \mathbb{C}^* / \langle \alpha \rangle$, $0 < |\alpha| < 1$, be an elliptic curve. Fix some $n \geq 1$ and $t \in \mathbb{C}^n$, identified with a polynomial via $t(w) = \sum t_k w^k$. Using these one defines an automorphism of $\mathbb{C} \times \mathbb{C}^*$,

$$g_{n,\alpha,t}(z, w) = (w^n z + t(w), \alpha w).$$

Let $A_{n,\alpha,t}$ denote the quotient surface $\mathbb{C} \times \mathbb{C}^* / \langle g_{n,\alpha,t} \rangle$, which is an affine line bundle over E. Enoki showed [22], generalizing a previous construction of Inoue [44], that these can be compactified with a cycle of rational curves, yielding a compact surface $S_{n,\alpha,t}$ of Class VII_+. Generically this cycle of curves are the only curves present in the surface, but in the case $t = 0$ (also known as parabolic Inoue surfaces), the zero section of the line bundle is a smooth elliptic curve in $S_{n,\alpha,0}$, satisfying $K \cdot E = 1$. Thus according to the conjecture above for parabolic Inoue surfaces we expect the area of this elliptic curve to grow linearly and ultimately form the "infinity" of a limiting complete steady soliton. For more general Enoki surfaces $E_{n,\alpha,t}$, there is still a topological torus in the homotopy class of the cycle of rational curves, whose area goes to infinity, conjecturally resulting in the same steady soliton limit resulting from $E_{n,\alpha,0}$. It seems likely that the compactified limiting surfaces are all biholomorphic to $E_{n,\alpha,0}$.

Next consider the Inoue–Hirzebruch surfaces. These were initially constructed by Inoue [45], using ideas related to Hirzebruch's description of Hilbert modular surfaces [39]. These surfaces are constructed by resolving singularities of compactified quotients of $\mathbb{H} \times \mathbb{C}$. These surfaces always admit two cycles of rational curves (cf. [17] for a description of these surfaces and the structure of their curves), each of which contains curves of self-intersection ≤ -3. Thus according to the conjecture above we expect the limit soliton to have at least two ends, possibly more depending on the structure of the curves. In a given cycle of rational curves smooth -2 curves may intersect curves of high self-intersection. As their area stays fixed while the high self-intersection curve is pushed to infinity, this suggests that the -2 curve forms a finite area cusp end with the intersection point now at infinity.

The two examples of Enoki surfaces and Inoue–Hirzebruch surfaces form the extreme cases of Dloussky's index invariant [16], with the remaining cases called "intermediate surfaces". Despite the intricate structure of curves on these surfaces, the limiting picture is essentially the same as described in the examples above, with the curves in $\Sigma_{>0}$ being pushed to infinity, forming potentially several ends, and all -2 curves either contained in a compact region or intersecting curves at infinity with a cusp end.

Part III: Classification of generalized Kähler structures

8. Generalized Kähler geometry

Generalized Kähler geometry arose in work of Gates–Hull–Roček [29], in the course of their investigation of supersymmetric sigma models. Later, using the framework of Hitchin's generalized geometry [40], Gualtieri [36, 37] understood generalized Kähler geometry in terms of a pair of commuting complex structures on $TM \oplus T^*M$ satisfying further compatibility conditions. For our purposes here we will restrict ourselves to the "classical" formulation and not exploit the language of generalized geometry. Thus a generalized Kähler structure on a manifold M is a triple (g, I, J) of a Riemannian metric together with two integrable complex structures, such that

$$d^c_I \omega_I = -d^c_J \omega_J, \qquad dd^c_I \omega_I = 0,$$

where $\omega_I = g(I\cdot, \cdot)$, and $d^c_I = \sqrt{-1}(\overline{\partial}_I - \partial_I)$, with analogous definitions for J.

Associated to every generalized Kähler structure is a Poisson structure

$$\sigma = \tfrac{1}{2} g^{-1}[I, J]. \tag{8.1}$$

As shown by Pontecorvo [66] and Hitchin [41], σ is the real part of a holomorphic Poisson structure with respect to both I and J, in other words

$$\overline{\partial}_I \sigma^{2,0}_I = 0, \qquad \overline{\partial}_J \sigma^{2,0}_J = 0.$$

The vanishing locus of σ has profound implications for the structure of generalized Kähler manifolds, and it is natural to understand their classification in terms of its structure. The simplest case occurs when $\sigma \equiv 0$. In this case $[I, J] = 0$, and the endomorphism $Q = IJ$ satisfies $Q^2 = \mathrm{Id}$. Thus Q has eigenvalues ± 1, and the eigenspaces of Q yield a splitting $TM = T_+ \oplus T_-$. These summands are I-invariant, thus we obtain a further splitting

$$T_{\mathbb{C}} M = T^{1,0}_+ \oplus T^{0,1}_+ \oplus T^{1,0}_- \oplus T^{0,1}_-.$$

Complex surfaces admitting a holomorphic splitting of the tangent bundle were classified by Beauville [6]. Apostolov and Gualitieri [2] determined precisely which of these admits generalized Kähler structure of this type.

Theorem 8.1 ([2, Theorem 1]). *A compact complex surface (M, I) admits a generalized Kähler structure (g, I, J) with $[I, J] = 0$ if and only if (M, I) is biholomorphic to*

1. *A ruled surface which is the projectivization of a projectively flat holomorphic vector bundle over a compact Riemann surface,*
2. *A bi-elliptic surface,*
3. *A surface of Kodaira dimension one with b_1 even, which is an elliptic fibration over a compact Riemann surface, with singular fibers only multiple smooth elliptic curves,*
4. *A surface of general type, whose universal cover is biholomorphic to $\mathbb{H} \times \mathbb{H}$, with fundamental group acting diagonally on the factors.*
5. *A class* 1 *Hopf surface*
6. *An Inoue surface of type S_M.*

The next simplest case occurs when σ defines a nondegenerate bilinear form at all points. It follows that $\Omega = \sigma^{-1}$ is a symplectic form, which is the real part of a holomorphic symplectic form with respect to both I and J. The simplest example comes from hyperKähler geometry. If (M^{4n}, g, I, J, K) is hyperKähler, then the triple (M^{4n}, g, I, J) is generalized Kähler, and using the quaternion relations one computes that $\Omega = \omega_K$. As we will describe in §9.2, it is possible to deform this example to obtain a non-Kähler generalized Kähler structure with nondegenerate Poisson structure. The existence of a holomorphic symplectic form places strong restrictions on the underlying complex manifolds. For complex surfaces, it follows from ([1, Proposition 2]) that the only possible underlying complex surfaces are tori, $K3$ surfaces, or primary Kodaira surfaces. Considering the I and J-imaginary pieces of Ω, one overall obtains three independent self-dual forms, ruling out Kodaira surfaces which have $b_2^+ = 2$. Thus such structures exist only on tori and $K3$ surfaces.

Generally, the Poisson structure can experience "type change", that is, the rank can drop on some locus. As we are in four dimensions, and the rank jumps in multiples of 4, the Poisson structure will be nondegenerate outside of a locus

$$T = \{p \in M \mid I = \pm J\}.$$

It turns out that T is a complex curve in both (M, I) and (M, J). In the case the underlying surfaces are Kähler, T is the support of an anticanonical divisor, and so only Del Pezzo surfaces are possible backgrounds, and Hitchin [42] constructed generalized Kähler structures on these surfaces. If the underlying surfaces are non-Kähler T is the support of a numerically anticanonical divisor (cf. [18]). In fact T must be disconnected ([1, Proposition 4]), and by a result of Nakamura ([27, Lemma 3.3]), a surface with disconnected numerical anticanonical divisor must be either a class 1 Hopf surface, or a parabolic or hyperbolic Inoue surface. The existence of generalized Kähler structures in some of these cases was established by Fujiki–Pontecorvo [27].

9. Generalized Kähler–Ricci flow

As shown in ([82, Theorem 1.2]), the pluriclosed flow preserves generalized Kähler geometry. This arises as a consequence of Theorem 2.6. In particular, one notes that generalized Kähler structures consist of a metric which is pluriclosed with respect to two distinct complex structures, satisfying further integrability conditions. By solving pluriclosed flow on both complex manifolds, and applying gauge transformations one obtains two solutions to (2.5) with the same initial data, yielding a time-dependent triple (g_t, I_t, J_t) of generalized Kähler structures (cf. [82] for details). Unpacking the construction yields the following evolution equations, which have the notable feature that the complex structures must evolve along the flow.

Definition 9.1. A one-parameter family of generalized Kähler structures $(M^{2n}, g_t, I_t, J_t, H_t)$ is a solution of *generalized Kähler–Ricci flow* (GKRF) if

$$\begin{aligned} \frac{\partial}{\partial t} g &= -2\operatorname{Rc}^g + \frac{1}{2} H^2, & \frac{\partial}{\partial t} H &= \Delta_d H, \\ \frac{\partial}{\partial t} I &= L_{\theta_I^\sharp} I, & \frac{\partial}{\partial t} J &= L_{\theta_J^\sharp} J. \end{aligned} \tag{9.1}$$

As a special case of pluriclosed flow, in principle we have already described all of the conjectural long time existence and convergence behavior in part II. Nonetheless we will provide further discussion of long time existence results and refined descriptions of convergence behavior in this setting. Interestingly, the GKRF in the I-fixed gauge preserves the Poisson structure on complex surfaces, a natural fact since it is holomorphic and thus rigid. This can be shown using a case-by-case depending on the type (vanishing, nondegnerate, general) of the Poisson structure. Conjecturally this is true in all dimensions, and should follow from a direct computation. Depending on the type of the Poisson structure, the GKRF takes very different forms, which we will describe in turn below.

9.1. Commuting case

As discussed in §8, in the case $[I, J] = 0$ one obtains a splitting of the tangent bundle according to the eigenspaces of $Q = IJ$. This also induces a splitting of the cotangent bundle, which induces a splitting of the exterior derivative

$$d = \partial_+ + \overline{\partial}_+ + \partial_- + \overline{\partial}_-.$$

Arguing similarly to the $\partial\overline{\partial}$-Lemma in Kähler geometry, it is possible to obtain a local potential function describing generalized Kähler metrics in this setting. In particular if (g, I, J) is generalized Kähler then near any point there exists a smooth function f such that

$$\omega_I = \sqrt{-1}\left(\partial_+\overline{\partial}_+ - \partial_-\overline{\partial}_-\right) f.$$

This difference of sign indicates a fundamental distinction between Kähler geometry and generalized Kähler geometry in this setting: rather than being described locally by a plurisubharmonic function, the metric is described by a function which is plurisubharmonic in certain directions and plurisuperharmonic in others. Nonetheless, given this local description, one expects the pluriclosed flow to reduce to a scalar PDE in this setting. This was confirmed in ([79, Theorem 1.1]), and locally this PDE takes the form

$$\frac{\partial f}{\partial t} = \log \frac{\det \sqrt{-1}\partial_+ \overline{\partial}_+ f}{\det(-\sqrt{-1}\partial_- \overline{\partial}_- f)}. \tag{9.2}$$

We refer to this equation as "twisted Monge–Ampère," as it is a natural combination of Monge–Ampère operators for the different pieces of the metric. This PDE is still parabolic, but is mixed concave/convex, and thus many of the usual methods for analyzing fully nonlinear PDE do not directly apply. Nonetheless we were able to give a nearly complete picture of the long time existence of the flow on complex surfaces of this type.

Theorem 9.2 ([79, Theorem 1.3]). *Let (M^4, g_0, I, J) be a generalized Kähler surface satisfying $[I, J] = 0$ and $I \neq \pm J$. Suppose (M^4, I) is biholomorphic to one of:*

1. *A ruled surface over a curve of genus $g \geq 1$.*
2. *A bi-elliptic surface,*
3. *An elliptic fibration of Kodaira dimension one,*
4. *A compact complex surface of general type, whose universal cover is biholomorphic to $\mathbb{H} \times \mathbb{H}$,*
5. *An Inoue surface of type S_M.*

Then the solution to pluriclosed flow with initial condition g_0 exists on $[0, \tau^*(\omega_0))$.

Referring back to Theorem 8.1, the only cases not covered by this theorem are ruled surfaces over curves of genus 0 and Hopf surfaces. The reason for the restriction in the theorem is that we require one of the line subbundles $T_\pm$ to have nonpositive first Chern class to obtain some partial a priori control over the metric. Also we note that there is overlap between this theorem and Theorems 4.3, 4.4 above.

9.2. Nondegenerate case

On the other extreme of generalized Kähler geometry is the nondegenerate case described above. In §8 we explained that if (M^4, g, I, J, K) is hyperKähler, we can interpret (M^4, g, I, J) as a generalized Kähler structure. Of course the underlying pairs (g, I) and (g, J) are still Kähler, but Joyce showed (cf. [1]) that one can deform away from these examples to produce genuine examples of non-Kähler, generalized Kähler, structures with σ nondegenerate. Specifically, given f_t a one-parameter family of smooth functions, we define a family of vector fields X_t via

$$X_t = \sigma d f_t.$$

Let ϕ_t denote the one parameter family of diffeomorphisms generated by X_t, which one notes is Ω-Hamiltonian by construction (recall $\Omega = \sigma^{-1}$). The triple $(I, \phi_t^* J, \Omega)$ determines a generalized Kähler structure, with g_t determined algebraically by (8.1).

Surprisingly, the generalized Kähler–Ricci flow evolves by precisely this type of deformation. In particular, given a generalized Kähler structure (M^4, g, I, J) we let

$$p = \tfrac{1}{4} \operatorname{tr} IJ$$

denote the angle between I and J. One has $|p| \leq 1$, and the inequality is strict everywhere in the nondegenerate setting. In four dimensions the function p is constant if and only if g is hyperKähler ([66]). Thus we expect the generalized Kähler–Ricci flow in this setting to completely reduce to quantities involving the angle function, and indeed this is the case. Specifically, if we gauge-modify the generalized Kähler–Ricci flow to fix the complex structure I, then $J_t = \phi_t^* J$, where

$$\frac{d\phi}{dt} = \sigma d \log \frac{1+p}{1-p}.$$

This is roughly analogous to the usual scalar reduction for Kähler–Ricci flow, although here the Hamiltonian diffeomorphism ϕ_t depends on the entire history of the flow on $[0, t]$, and thus does not truly reduce to a single scalar.

We can give a complete description of the long time existence and a weak confirmation of the conjectured convergence behavior in this setting.

Theorem 9.3 ([77, Theorem 1.1]). *Let (M^4, g, I, J) be a nondegenerate generalized Kähler four-manifold. The solution to generalized Kähler–Ricci flow with initial data (g, I, J) exists for all time. Moreover, $(\omega_I)_t$ subconverges to a closed current.*

While one expects smooth convergence to a hyperKähler structure, the convergence behavior above at least shows that the flow contracts to the space of Kähler structures. In fact more convergence properties can be shown (cf. [77] for detail). This result was extended to arbitrary dimensions by the author and Apostolov [3]. The key observation is to show that a "generalized Calabi–Yau quantity," motivated by natural constructions in generalized geometry ([35]), governs the dynamics of the flow in the same manner that $\log \frac{1+p}{1-p}$ does in four dimensions.

9.3. General case

The general case involves Poisson structures with type change locus, which exist on Del Pezzo surfaces, Hopf surfaces, and parabolic and hyperbolic Inoue surfaces as described above. In all of these cases the conjectures on pluriclosed flow imply a connectedness result for the space of generalized Kähler structures. Note that on a given complex manifold, the space of Kähler metrics is convex by linear paths, whereas uniqueness and moduli questions for complex structures are of course much more subtle. In understanding the space of generalized Kähler metrics, these two problems are linked. Moreover, there is no linear structure to this space, with the natural class of deformations instead using Hamiltonian diffeomorphisms as described above. Thus the generalized Kähler–Ricci flow can potentially yield connectedness of the space of generalized Kähler structures, a nontrivial consequence due to the nonlinear structure of this space. Furthermore, the conjectural behavior for parabolic and hyperbolic Inoue surfaces described in part II suggests that the limiting complete steady soliton metrics associated to these surfaces should in fact be generalized Kähler, and thus generalized Kähler structures should play a key role in understanding the geometrization of complex surfaces.

As an example, consider the case of $\mathbb{CP}^2$, where uniqueness of the complex structure in known [102]. Hitchin [42] constructed generalized Kähler structures on $\mathbb{CP}^2$ by a modification of the Hamiltonian diffeomorphism method described above, deforming away from the standard Fubini Study structure. Given this, and even knowing the uniqueness of complex structure on $\mathbb{CP}^2$, it is still possible that the space of generalized Kähler metrics has multiple disconnected components, i.e., there may exist other generalized Kähler triples not arising by this deformation. This problem can be very naturally addressed using the generalized Kähler–Ricci flow, and in particular we can rule out the existence of such exotic generalized Kähler structures. The main input is a description of the long time existence and weak convergence behavior of generalized Kähler–Ricci flow in this setting.

Theorem 9.4 ([73]). *Let* $(\mathbb{CP}^2, g, I, J)$ *be a generalized Kähler structure. The solution to normalized generalized Kähler–Ricci flow with initial condition* (g, I, J) *exists on* $[0, \infty)$. *Moreover,* $(\omega_I)_t$ *subconverges to a closed current.*

As we detail in [73], there is a natural completion of the space of generalized Kähler metrics extending the usual completion of Kähler metrics in the space of closed currents. Our result yields connectivity of this space, and in fact that any point in this space is equivalent to the standard Fubini study structure by an extended Courant symmetry. This is a natural extension of the classical uniqueness Theorem of Yau [102] for complex structures on $\mathbb{CP}^2$ to generalized Kähler structures.

Acknowledgement

It is a pleasure to dedicate this article to Professor Gang Tian on the occasion of his 59th birthday. My collaboration with Tian obviously forms the underpinning of this article, and Tian's many works in Kähler geometry helped inspire the picture we paint here. We thank Georges Dloussky for many helpful conversations.

References

[1] Apostolov, V., Gauduchon, P., and Grantcharov, G. Bi-Hermitian structures on complex surfaces. *Proc. London Math. Soc.* (3) 79, 2 (1999), 414–428.

[2] Apostolov, V., and Gualtieri, M. Generalized Kähler manifolds, commuting complex structures, and split tangent bundles. *Comm. Math. Phys.* 271 , 2 (2007), 561–575.

[3] Apostolov, V., and Streets, J. The nondegenerate generalized Kähler Calabi–Yau problem. *ArXiv e-prints* (Mar. 2017).

[4] Aubin, T. équations du type Monge–Ampère sur les variétés kählériennes compactes. *Bull. Sci. Math.* (2) 102, 1 (1978), 63–95.

[5] Barth, W.P., Hulek, K., Peters, C.A.M., and Van de Ven, A. *Compact complex surfaces*, second ed., vol. 4 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics].* Springer-Verlag, Berlin, 2004.

[6] Beauville, A. Complex manifolds with split tangent bundle. In *Complex analysis and algebraic geometry.* de Gruyter, Berlin, 2000, pp. 61–70.

[7] Boling, J. Homogeneous solutions of pluriclosed flow on closed complex surfaces. *J. Geom. Anal.* 26, 3 (2016), 2130–2154.

[8] Brinzănescu, V. Néron-Severi group for nonalgebraic elliptic surfaces. II. Non-Kählerian case. *Manuscripta Math.* 84, 3-4 (1994), 415–420.

[9] Buchdahl, N. On compact Kähler surfaces. *Ann. Inst. Fourier (Grenoble)* 49, 1 (1999), vii, xi, 287–302.

[10] Calabi, E. Improper affine hyperspheres of convex type and a generalization of a theorem by K. Jörgens. *Michigan Math. J.* 5 (1958), 105–126.

[11] Cao, H.D. Deformation of Kähler metrics to Kähler–Einstein metrics on compact Kähler manifolds. *Invent. Math.* 81, 2 (1985), 359–372.

[12] Cartan, E., and Schouten, J.A. On Riemannian geometries admitting an absolute parallelism. *Proc. Amsterdam* 29 (1926), 933–946.

[13] Cartan, E., and Schouten, J.A. On the geometry of the group-manifold of simple and semi-simple groups. *Proc. Amsterdam* 29 (1926), 803–815.

[14] De Giorgi, E. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat.* (3) 3 (1957), 25–43.

[15] De Giorgi, E. Un esempio di estremali discontinue per un problema variazionale di tipo ellittico. *Boll. Un. Mat. Ital.* (4) 1 (1968), 135–137.

[16] Dloussky, G. Structure des surfaces de Kato. *Mém. Soc. Math. France (N.S.)*, 14 (1984), ii+120.

[17] Dloussky, G. Une construction élémentaire des surfaces d'Inoue–Hirzebruch. *Math. Ann.* 280, 4 (1988), 663–682.

[18] Dloussky, G. On surfaces of class VII_0^+ with numerically anticanonical divisor. *Amer. J. Math.* 128, 3 (2006), 639–670.

[19] Dloussky, G., Oeljeklaus, K., and Toma, M. Surfaces de la classe VII_0 admettant un champ de vecteurs. *Comment. Math. Helv.* 75, 2 (2000), 255–270.

[20] Dloussky, G., Oeljeklaus, K., and Toma, M. Surfaces de la classe VII_0 admettant un champ de vecteurs. II. *Comment. Math. Helv.* 76, 4 (2001), 640–664.

[21] Dloussky, G., Oeljeklaus, K., and Toma, M. Class VII_0 surfaces with b_2 curves. *Tohoku Math. J.* (2) 55, 2 (2003), 283–309.

[22] Enoki, I. Surfaces of class VII_0 with curves. *Tôhoku Math. J.* (2) 33, 4 (1981), 453–492.

[23] Evans, L.C. Classical solutions of fully nonlinear, convex, second-order elliptic equations. *Comm. Pure Appl. Math.* 35, 3 (1982), 333–363.

[24] Fang, S., Tosatti, V., Weinkove, B., and Zheng, T. Inoue surfaces and the Chern–Ricci flow. *J. Funct. Anal.* 271, 11 (2016), 3162–3185.

[25] Feldman, M., Ilmanen, T., and Knopf, D. Rotationally symmetric shrinking and expanding gradient Kähler–Ricci solitons. *J. Differential Geom.* 65, 2 (2003), 169–209.

[26] Filipkiewicz, R. *Four dimensional geometries.* PhD thesis, University of Warwick, 1984.

[27] Fujiki, A., and Pontecorvo, M. Anti-self-dual bihermitian structures on Inoue surfaces. *J. Differential Geom.* 85, 1 (2010), 15–71.

[28] Futaki, A. An obstruction to the existence of Einstein Kähler metrics. *Invent. Math.* 73, 3 (1983), 437–443.

[29] Gates, Jr., S.J., Hull, C.M., and Roček, M. Twisted multiplets and new supersymmetric nonlinear σ-models. *Nuclear Phys. B* 248, 1 (1984), 157–186.

[30] Gauduchon, P. La 1-forme de torsion d'une variété hermitienne compacte. *Math. Ann.* 267, 4 (1984), 495–518.

[31] Gauduchon, P. Structures de Weyl–Einstein, espaces de twisteurs et variétés de type $S^1 \times S^3$. *J. Reine Angew. Math.* 469 (1995), 1–50.

[32] Gauduchon, P. Hermitian connections and Dirac operators. *Boll. Un. Mat. Ital. B* (7) 11, 2, suppl. (1997), 257–288.

[33] Gauduchon, P., and Ivanov, S. Einstein–Hermitian surfaces and Hermitian Einstein-Weyl structures in dimension 4. *Math. Z.* 226, 2 (1997), 317–326.

[34] Gill, M. Convergence of the parabolic complex Monge–Ampère equation on compact Hermitian manifolds. *Comm. Anal. Geom.* 19, 2 (2011), 277–303.

[35] Gualtieri, M. *Generalized complex geometry.* PhD thesis, St. John's College, University of Oxford, 11 2003.

[36] Gualtieri, M. Generalized complex geometry. *Ann. of Math.* (2) 174, 1 (2011), 75–123.

[37] Gualtieri, M. Generalized Kähler geometry. *Comm. Math. Phys.* 331, 1 (2014), 297–331.

[38] Hamilton, R.S. Three-manifolds with positive Ricci curvature. *J. Differential Geom.* 17, 2 (1982), 255–306.

[39] Hirzebruch, F.E.P. Hilbert modular surfaces. *Enseignement Math.* (2) 19 (1973), 183–281.

[40] Hitchin, N. Generalized Calabi–Yau manifolds. *Q. J. Math.* 54, 3 (2003), 281–308.

[41] Hitchin, N. Instantons, Poisson structures and generalized Kähler geometry. *Comm. Math. Phys.* 265, 1 (2006), 131–164.

[42] Hitchin, N. Bihermitian metrics on del Pezzo surfaces. *J. Symplectic Geom.* 5, 1 (2007), 1–8.

[43] Inoue, M. On surfaces of Class VII_0. *Invent. Math.* 24 (1974), 269–310.

[44] Inoue, M. New surfaces with no meromorphic functions. 423–426.

[45] Inoue, M. New surfaces with no meromorphic functions. II. 91–106.

[46] Kato, M. Compact complex manifolds containing "global" spherical shells. I. In *Proceedings of the International Symposium on Algebraic Geometry* (*Kyoto Univ., Kyoto,* 1977) (1978), Kinokuniya Book Store, Tokyo, pp. 45–84.

[47] Kleiner, B., and Lott, J. Notes on Perelman's papers. *Geom. Topol.* 12, 5 (2008), 2587–2855.

[48] Klingler, B. Structures affines et projectives sur les surfaces complexes. *Ann. Inst. Fourier (Grenoble)* 48, 2 (1998), 441–477.

[49] Kodaira, K. On the structure of compact complex analytic surfaces. I. *Amer. J. Math.* 86 (1964), 751–798.

[50] Kodaira, K. On the structure of compact complex analytic surfaces. II. *Amer. J. Math.* 88 (1966), 682–721.

[51] Koiso, N. On rotationally symmetric Hamilton's equation for Kähler–Einstein metrics. In *Recent topics in differential and analytic geometry*, vol. 18 of *Adv. Stud. Pure Math.* Academic Press, Boston, MA, 1990, pp. 327–337.

[52] Krylov, N.V. Boundedly inhomogeneous elliptic and parabolic equations. *Izv. Akad. Nauk SSSR Ser. Mat.* 46, 3 (1982), 487–523, 670.

[53] Krylov, N.V., and Safonov, M.V. An estimate for the probability of a diffusion process hitting a set of positive measure. *Dokl. Akad. Nauk SSSR* 245, 1 (1979), 18–20.

[54] Krylov, N.V., and Safonov, M.V. A property of the solutions of parabolic equations with measurable coefficients. *Izv. Akad. Nauk SSSR Ser. Mat.* 44, 1 (1980), 161–175, 239.

[55] Lamari, A. Courants kählériens et surfaces compactes. *Ann. Inst. Fourier (Grenoble)* 49, 1 (1999), vii, x, 263–285.

[56] Li, J., Yau, S.-T., and Zheng, F. On projectively flat Hermitian manifolds. *Comm. Anal. Geom.* 2, 1 (1994), 103–109.

[57] Matsushima, Y. Sur la structure du groupe d'homéomorphismes analytiques d'une certaine variété kählérienne. *Nagoya Math. J.* 11 (1957), 145–150.

[58] Morgan, J., and Tian, G. *Ricci flow and the Poincaré conjecture*, vol. 3 of *Clay Mathematics Monographs.* American Mathematical Society, Providence, RI; Clay Mathematics Institute, Cambridge, MA, 2007.

[59] Moser, J. A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations. *Comm. Pure Appl. Math.* 13 (1960), 457–468.

[60] Nakamura, I. Classification of non-Kähler complex surfaces. *Sūgaku* 36, 2 (1984), 110–124. Translated in Sugaku Expositions **2** (1989), no. 2, 209–229.

[61] Nash, J. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.* 80 (1958), 931–954.

[62] Oliynyk, T., Suneeta, V., and Woolgar, E. A gradient flow for worldsheet nonlinear sigma models. *Nuclear Phys. B* 739, 3 (2006), 441–458.

[63] Perelman, G. The entropy formula for the Ricci flow and its geometric applications. *ArXiv Mathematics e-prints* (Nov. 2002).

[64] Perelman, G. Finite extinction time for the solutions to the Ricci flow on certain three-manifolds. *ArXiv Mathematics e-prints* (July 2003).

[65] Perelman, G. Ricci flow with surgery on three-manifolds. *ArXiv Mathematics e-prints* (Mar. 2003).

[66] Pontecorvo, M. Complex structures on Riemannian four-manifolds. *Math. Ann.* 309, 1 (1997), 159–177.

[67] Siu, Y.T. Every $K3$ surface is Kähler. *Invent. Math.* 73, 1 (1983), 139–150.

[68] Song, J., Székelyhidi, G., and Weinkove, B. The Kähler–Ricci flow on projective bundles. *Int. Math. Res. Not. IMRN*, 2 (2013), 243–257.

[69] Song, J., and Tian, G. The Kähler–Ricci flow on surfaces of positive Kodaira dimension. *Invent. Math.* 170, 3 (2007), 609–653.

[70] Song, J., and Tian, G. The Kähler–Ricci flow through singularities. *Invent. Math.* 207, 2 (2017), 519–595.

[71] Song, J., and Weinkove, B. The Kähler–Ricci flow on Hirzebruch surfaces. *J. Reine Angew. Math.* 659 (2011), 141–168.

[72] Streets, J. Pluriclosed flow on elliptic fibrations and Ricci–Yang–Mills flow, *to appear*.

[73] Streets, J. Uniqueness of generalized Kähler structure on $\mathbb{CP}^n$, *to appear*.

[74] Streets, J. Regularity and expanding entropy for connection Ricci flow. *J. Geom. Phys.* 58, 7 (2008), 900–912.

[75] Streets, J. Pluriclosed flow, Born–Infeld geometry, and rigidity results for generalized Kähler manifolds. *Comm. Partial Differential Equations* 41, 2 (2016), 318–374.

[76] Streets, J. Generalized geometry, T-duality, and renormalization group flow. *J. Geom. Phys.* 114 (2017), 506–522.

[77] Streets, J. Generalized Kähler–Ricci flow and the classification of nondegenerate generalized Kähler surfaces. *Adv. Math.* 316 (2017), 187–215.

[78] Streets, J. Classification of solitons for pluriclosed flow on complex surfaces. *ArXiv e-prints* (Feb. 2018).

[79] Streets, J. Pluriclosed flow on generalized Kähler manifolds with split tangent bundle. *J. Reine Angew. Math.* 739 (2018), 241–276.

[80] Streets, J., and Tian, G. A parabolic flow of pluriclosed metrics. *Int. Math. Res. Not. IMRN*, 16 (2010), 3101–3133.

[81] Streets, J., and Tian, G. Hermitian curvature flow. *J. Eur. Math. Soc.* (*JEMS*) 13, 3 (2011), 601–634.

[82] Streets, J., and Tian, G. Generalized Kähler geometry and the pluriclosed flow. *Nuclear Phys. B* 858, 2 (2012), 366–376.

[83] Streets, J., and Tian, G. Regularity results for pluriclosed flow. *Geom. Topol.* 17, 4 (2013), 2389–2429.

[84] Streets, J., and Warren, M. Evans–Krylov estimates for a nonconvex Monge–Ampère equation. *Math. Ann.* 365, 1-2 (2016), 805–834.

[85] Teleman, A. Donaldson theory on non-Kählerian surfaces and class VII surfaces with $b_2 = 1$. *Invent. Math.* 162, 3 (2005), 493–521.

[86] Teleman, A. The pseudo-effective cone of a non-Kählerian surface and applications. *Math. Ann.* 335, 4 (2006), 965–989.

[87] Teleman, A. Instantons and curves on class VII surfaces. *Ann. of Math.* (2) 172, 3 (2010), 1749–1804.

[88] Teleman, A.D. Projectively flat surfaces and Bogomolov's theorem on class VII_0 surfaces. *Internat. J. Math.* 5, 2 (1994), 253–264.

[89] Thurston, W.P. Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. *Bull. Amer. Math. Soc.* (*N.S.*) 6, 3 (1982), 357–381.

[90] Thurston, W.P. *Three-dimensional geometry and topology. Vol.* 1, vol. 35 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1997. Edited by Silvio Levy.

[91] Tian, G. On Calabi's conjecture for complex surfaces with positive first Chern class. *Invent. Math.* 101, 1 (1990), 101–172.

[92] Tian, G., and Yau, S.-T. Complete Kähler manifolds with zero Ricci curvature. I. *J. Amer. Math. Soc.* 3, 3 (1990), 579–609.

[93] Tian, G., and Yau, S.-T. Complete Kähler manifolds with zero Ricci curvature. II. *Invent. Math.* 106, 1 (1991), 27–60.

[94] Tian, G., Zhang, S., Zhang, Z., and Zhu, X. Perelman's entropy and Kähler–Ricci flow on a Fano manifold. *Trans. Amer. Math. Soc.* 365, 12 (2013), 6669–6695.

[95] Tian, G., and Zhang, Z. On the Kähler–Ricci flow on projective manifolds of general type. *Chinese Ann. Math. Ser. B* 27, 2 (2006), 179–192.

[96] Tian, G., and Zhu, X. Convergence of the Kähler–Ricci flow on Fano manifolds. *J. Reine Angew. Math.* 678 (2013), 223–245.

[97] Tosatti, V., and Weinkove, B. On the evolution of a Hermitian metric by its Chern–Ricci form. *J. Differential Geom.* 99, 1 (2015), 125–163.

[98] Tosatti, V., Weinkove, B., and Yang, X. Collapsing of the Chern–Ricci flow on elliptic surfaces. *Math. Ann.* 362, 3-4 (2015), 1223–1271.

[99] Ustinovskiy, Y. The Hermitian curvature flow on manifolds with non-negative Griffiths curvature. *ArXiv e-prints* (Apr. 2016).

[100] Wall, C.T.C. Geometric structures on compact complex analytic surfaces. *Topology* 25, 2 (1986), 119–153.

[101] Wang, X.-J., and Zhu, X. Kähler–Ricci solitons on toric manifolds with positive first Chern class. *Adv. Math.* 188, 1 (2004), 87–103.

[102] Yau, S.T. Calabi's conjecture and some new results in algebraic geometry. *Proc. Nat. Acad. Sci. U.S.A.* 74, 5 (1977), 1798–1799.

[103] Yau, S.T. On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampère equation. I. *Comm. Pure Appl. Math.* 31, 3 (1978), 339–411.

[104] Yau, S.T. The role of partial differential equations in differential geometry. In *Proceedings of the International Congress of Mathematicians* (*Helsinki,* 1978) (1980), Acad. Sci. Fennica, Helsinki, pp. 237–250.

Jeffrey Streets
Rowland Hall
University of California, Irvine
Irvine, CA 92617, USA
e-mail: jstreets@uci.edu

Progress in Mathematics, Vol. 333, 511–520

From Optimal Transportation to Conformal Geometry

Neil S. Trudinger

Dedicated to Gang Tian on the occasion of his 60th birthday

Abstract. In this paper we discuss the link between domain convexity in optimal transportation and the estimation of second derivatives in augmented Hessian equations, leading to the estimation of second derivatives in fully nonlinear Yamabe problems with boundary with boundary curvature conditions which may be also nonlinear.

Mathematics Subject Classification (2010). 35J66, 35J25, 53A30 .

Keywords. Nonlinear Yamabe problem with boundary, optimal transportation, augmented Hessian equations, local derivative estimates.

1. Introduction

In this paper, we explain the link between domain convexity in optimal transportation with the estimation of second derivatives in fully nonlinear Yamabe problems with boundary, arising from our study of oblique boundary value problems for augmented Hessian equations in [7–9]. Our presentation here is based on that in talks given most recently at the International Conference on Differential Geometry in celebration of Professor Gang Tian's 60th birthday at Sydney in January 2018 and previously at meetings in Seoul and Armidale in 2016 and Hangzhou in 2017.

We begin by formulating a general nonlinear Yamabe problem with boundary which extends the boundary mean curvature case studied in [2, 3, 12, 13, 15]. For an account of the history of the Yamabe problem and its nonlinear extensions the reader is referred for example to the most recent of these works [15]. Here we just recall that the original Yamabe problem concerned the conformal deformation of a metric on a compact Riemannian manifold, without boundary, to one with constant scalar curvature and was partially solved by the author in [22], following

Research supported by Australian Research Council (No. DP170100929).

the original work by Yamabe [31], and completely solved by Schoen in [20]. In the fully nonlinear version with boundary, which we formulate here, our nonlinearities will be described in terms of symmetric functions on cones in Euclidean space $\mathbb{R}^n$. Accordingly we let Γ denote an open, convex, symmetric cone $\subset \mathbb{R}^n$, with vertex at 0, containing the positive cone K^+. and f a positive, increasing function in $C^\infty(\Gamma) \cap C^0(\bar{\Gamma})$, vanishing on $\partial\Gamma$ and normalized to be positive one-homogeneous. Henceforth we refer to (f, Γ) as a symmetric pair in $\mathbb{R}^n$ and as a concave symmetric pair when f is also a concave function on Γ. Now we let (M, g) denote a compact, n-dimensional, $(n \geq 3)$, C^∞ Riemannian manifold with metric g and boundary ∂M. The Schouten tensor of (M, g) is given by

$$A_g = \frac{1}{n-2}\left(\mathrm{Ric}_g - \frac{1}{2(n-1)} Rg\right), \tag{1.1}$$

where Ric_g, R denote respectively the Ricci tensor and scalar curvature of (M, g). We denote by $\lambda(A_g)$ the eigenvalues of A_g and by $\kappa_g = (\kappa_1, \ldots, \kappa_{n-1})$ the principal curvatures of ∂M, both with respect to g. Letting (f, Γ), $(\tilde{f}, \tilde{\Gamma})$ be symmetric pairs on $\mathbb{R}^n$, $\mathbb{R}^{n-1}$ respectively, our general, (non-degenerate), Yamabe problem can be now expressed as follows:

To find a smooth conformal metric $\bar{g}$ so that $\lambda(A_{\bar{g}}) \in \Gamma$, $\kappa_{\bar{g}} \in \tilde{\Gamma}$ and $f(\lambda(A_{\bar{g}}))$, $\tilde{f}(\kappa_{\bar{g}})$ are given positive constants.

More generally we can seek to prescribe

$$f(\lambda(A_{\bar{g}})) = \psi, \quad \tilde{f}(\kappa_{\bar{g}}) = \tilde{\psi}, \tag{1.2}$$

for given positive $\psi \in C^\infty(M)$, $\tilde{\psi} \in C^\infty(\partial M)$.

Note that this problem is meaningful for bounded domains in $\mathbb{R}^n$, unlike the Yamabe problem for domains without boundary.

We may also consider various degenerate cases: (i) $\kappa_{\bar{g}} \in \bar{\tilde{\Gamma}}, \tilde{\psi} \geq 0$, (ii) $\lambda(A_{\bar{g}}) \in \bar{\Gamma}, \psi \geq 0$, (iii) f, (or $\tilde{f}$), degenerate with $D_i f \geq 0$, $\sum D_i f > 0$, $i = 1, \ldots, n$, while in the original cases of the Yamabe problem when f or $\tilde{f}$ is linear, that is ψ is the scalar curvature of M or $\tilde{\psi}$ is the mean curvature of ∂M, it is also meaningful to allow ψ or $\tilde{\psi}$ to be negative and $\Gamma = \mathbb{R}^n$ or $\tilde{\Gamma} = \mathbb{R}^{n-1}$.

The connection to nonlinear partial differential equations arises by considering conformal deformations of the form:

$$\bar{g} = e^{-2u} g, \quad u \in C^\infty(M), \tag{1.3}$$

leading to the nonlinear Neumann problem:

$$PDE: \quad F(U) = f(\lambda(U)) = e^{-2u}\psi, \quad \lambda(U) \in \Gamma \quad \text{in } M, \tag{1.4}$$

$$BC: \qquad \tilde{f}(D_\nu u I + \kappa_g) = e^{-u}\tilde{\psi}, \quad D_\nu u I + \kappa_g \in \tilde{\Gamma} \quad \text{on } \partial M, \tag{1.5}$$

where U denotes the augmented Hessian,

$$U = \nabla^2 u + du \otimes du - \frac{1}{2}|\nabla u|^2 g + A_g, \tag{1.6}$$

$I = (1, \ldots, 1)$ and ν denotes the unit inner normal to ∂M.

Our conditions on f and Γ imply that the PDE (1.4) is elliptic with respect to u. In the degenerate cases (ii) and (iii) above, the PDE (1.4) may become degenerate elliptic, in which case we would only expect solutions at most in $C^{1,1}(M)$. Using the properties of $\tilde{\Gamma}$ we may also write the boundary condition as a semilinear Neumann condition,

$$D_\nu u = \varphi(\cdot, u) \tag{1.7}$$

for some function $\varphi \in C^\infty(\partial M \times \mathbb{R})$.

We conclude this introduction with a list of basic examples of symmetric pairs.

(i) Elementary symmetric functions:

$\Gamma = \Gamma_k = \{\lambda \in \mathbb{R}^n \mid S_j(\lambda) > 0,\ j = 1, \ldots, k\},\ k = 1, \cdot, n,$

$f = f_k = (S_k)^{1/k},\ S_k(\lambda) = \sum_{1 \le i_1 < \cdots < i_k \le n} \lambda_{i_1} \cdots \lambda_{i_k},$

$k = 1:\ S_1(\lambda) = \sum \lambda_i,\ \Gamma_1 = \text{positive half-space},$

$k = n:\ \Gamma_n = \{\lambda_i > 0\} = \text{positive cone } K^+,\ S_n(\lambda) = \prod \lambda_i\,.$

(ii) Quotients:

$\Gamma = \Gamma_k,\ f = f_{k,l} = \left(\frac{S_k}{S_l}\right)^{\frac{1}{k-l}},\ ,0 < l < k \le n,$

$k = n, l = n-1:\ f_{n,n-1} = S_{n,n-1} = \left(\sum \frac{1}{\lambda_i}\right)^{-1},$ harmonic mean.

(iii) Negative means:

$\Gamma = K^+,\ f = f_\alpha = \left(\sum \lambda_i^\alpha\right)^{1/\alpha},\ -\infty < \alpha < 0,$

$\alpha = -1:\ f_{-1} = f_{n,n-1},$

$\alpha \to -\infty:\ f_{-\infty} = \min \lambda_i,\ (\text{degenerate}, D_i f \ge 0, \sum D_i f = 1, \text{a.e.}).$

(iv) General interpolants:

Let (f, Γ) be a symmetric pair on $\mathbb{R}^{\binom{n}{m}}$, $\binom{n}{m} = \frac{n!}{m!(n-m)!}$, $0 < m \le n$.

$\Gamma^{(m)} = \{\lambda \in \mathbb{R}^n \mid \{\lambda_{i_1} + \cdots + \lambda_{i_m} \mid 1 \le i_1 < \cdots < i_m \le n\} \in \Gamma\},$

$f^{(m)} = f(\{\lambda_{i_1} + \cdots + \lambda_{i_m} \mid 1 \le i_1 < \cdots < i_m \le n\})\,.$

(iv)$'$ Special cases:

$\mathcal{P}_m = \Gamma^{(m)}_{\binom{n}{m}} = \{\lambda \in \mathbb{R}^n \mid \sum_{s=1}^m \lambda_{i_s} > 0, 1 \le i_1 < \cdots < i_m \le n\}$

$\mathcal{P}_1 = \Gamma_n,\ \mathcal{P}_n = \Gamma_1,\ \Gamma_k \subset \mathcal{P}_{n-k+1}, 1 < k < n,$

$P_m(\lambda) = \prod_{i_1 < \cdots < i_m} \sum_{s=1}^m \lambda_{i_s},\quad f^{(m)} = f^{(m)}_{\binom{n}{m}} = (P_m)^{\frac{1}{\binom{n}{m}}}$

In general $(\Gamma^{(m)}, f^{(m)})$ interpolate between $(\Gamma, f) = (\Gamma^{(1)}, f^{(1)})$ and $(\Gamma_1, f_1) = (\Gamma^{(n)}, f^{(n)})$ where $f_1 = f(1)S_1$. Historically, examples (i) and (ii) arose in the pioneering works on Hessian equations [1, 5, 23] while examples (iii) and (iv)$'$ are from [4] and [21]. Note that our general construction here in example (iv) can be iterated further to give an abundance of further examples, although we are not aware of applications where these arise except for the special case of (iv)$'$.

Concavity: We remark that f is concave in examples (i)–(iii), (iv)$'$ and $f^{(m)}$ is concave in example (iv) whenever f is concave.

2. Optimal transportation

The theory of optimal transportation, (see for example [28, 29]), is based upon the notion of a cost function, which for our purposes here, we can assume is a smooth real-valued function c defined on the product of the closures of two bounded domains $\Omega, \Omega^* \subset \mathbb{R}^n$ such that $c_x(x,\cdot)$ is smoothly invertible on $\bar{\Omega}^*$, for all $x \in \bar{\Omega}$. We can then define a smooth mapping Y by

$$Y(x,p) = (c_x)^{-1}(x,p) \quad \text{for} \quad \mathrm{x} \in \Omega,\ \mathrm{p} \in \mathrm{c_x(x,\cdot)}(\Omega^*) \tag{2.1}$$

Finding an optimal transport map T involves solving, in an appropriate sense, the second boundary value problem for a Monge–Ampère type equation,

$$\det[D^2u - A(\cdot, Du)] = B(\cdot, Du), \quad D^2u > A(\cdot, Du), \quad \text{in} \quad \Omega \tag{2.2}$$

$$Tu(\Omega) := Y(\cdot, Du)(\Omega) = \Omega^*, \tag{2.3}$$

where

$$A(x,p) = c_{xx}(x,Y), \quad B(x,p) = |\det c_{x,y}(x,Y)|\rho(x)/\rho^*(Y) \tag{2.4}$$

and ρ, ρ^* are given positive densities on Ω, Ω^*, satisfying the mass balance condition,

$$\int \rho = \int \rho^*. \tag{2.5}$$

Corresponding to our notion of symmetric pair in Section 1, we have in (2.2), $\Gamma = K^+$ and $F = \det^{1/n}$ (for one-homogeneity).

For classical solvability of the boundary value problem (2.2), (2.3) there are two critical conditions introduced in our papers, [19, 24, 26]. The first is a notion of domain convexity. Namely, the domain Ω is c-convex, (uniformly c-convex), with respect to the target domain Ω^* if and only if $c_y(\cdot, y)(\Omega)$ is convex for all $y \in \Omega^*$, (uniformly convex for all $y \in \bar{\Omega}^*$).

Note that by defining $c^*(x,y) := c(y,x)$ we have a dual notion of c^*-convexity for Ω^* with respect to Ω.

The second is a notion of co-dimension one convexity for the matrix function A. Namely A is regular, (strictly regular), at (x,p) if and only if

$$D_{p_k p_l} A_{ij}(x,z,p)\xi_i\xi_j\eta_k\eta_l \geq 0, (>0), \quad \forall \xi, \eta \neq 0,\ \in \mathbb{R}^n, \quad \xi.\eta = 0. \tag{2.6}$$

As shown in [17, 19], the c and c^*-convexity of Ω and Ω^*, with respect to each other, and the regularity of A are in fact necessary for classical solvability of (2.2), (2.3).

The following classical existence theorem is proved in [6, 25, 26].

Theorem 2.1. *Assume:*

- *the domains Ω, Ω^* are smooth and uniformly c, c^*-convex with respect to each other;*
- *the mappings $c_x(x,\cdot)$, $c_y(\cdot,y)$ are smoothly invertible for all $x \in \Omega, y \in \Omega^*$;*
- *the matrix function A is regular for all $x \in \Omega, Y(x,p) \in \Omega^*$;*
- *the densities ρ and ρ^* are smooth and positive on $\bar{\Omega}$, $\bar{\Omega}^*$, satisfying the mass balance condition* (2.5).

Then there exists a unique, (*up to additive constants*)*, elliptic solution $u \in C^2(\bar{\Omega})$ of the second boundary value problem* (2.2), (2.3).

The key estimates in the proof of Theorem 2.1 are global second derivative estimates and an obliqueness estimate for the equivalent oblique boundary condition:

$$\mathcal{G}[u] := \Phi^* \circ Y(\cdot, Du) = 0 \quad \text{on} \quad \partial\Omega \tag{2.7}$$

where Φ^* is a smooth defining function for Ω^*.

3. Augmented Hessian equations

We describe here some relevant results from our study of oblique boundary value problems for augmented Hessian equations in [7–9].

The general set up is as follows. We let Ω denote a bounded smooth domain in $\mathbb{R}^n$, A a smooth mapping from $\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n$ to $\mathbb{S}^n$, the linear space of $n \times n$ real symmetric matrices, Γ an open, convex cone in $\mathbb{S}^n$, with vertex at 0, containing the positive cone K^+ and F a smooth increasing positive real function on Γ. Our augmented Hessian operator on $C^2(\Omega)$ is now defined by

$$\mathcal{F}[u] := F(D^2u - A(\cdot, u, Du)), \quad u \in C^2(\Omega), \tag{3.1}$$

with augmented Hessian

$$U := D^2u - A(\cdot, u, Du) \in \Gamma. \tag{3.2}$$

A function $u \in C^2(\Omega)$ satisfying (3.2) is called admissible and clearly this implies that the operator F is elliptic with respect to u.

Also we call $\mathcal{F}$ orthogonally invariant if Γ is invariant under orthogonal transformations and $F(r) = f(\lambda)$ for a symmetric function f where λ denotes the eigenvalues of $r \in \Gamma$. Unless there is confusion, we will also use Γ and $F = f$ to denote the corresponding cones in $\mathbb{R}^n$ and symmetric functions in the orthogonally invariant case. Our main examples are then our examples of symmetric pairs in Section 1.

Next letting G denote a smooth real function on $\partial\Omega \times \mathbb{R} \times \mathbb{R}^n$, our general boundary operator is now defined by

$$\mathcal{G}[u] = G(\cdot, u, Du), \quad u \in C^1(\bar{\Omega}). \tag{3.3}$$

The operator $\mathcal{G}$ is called oblique if $G_p.\nu > 0$ on $\partial\Omega \times \mathbb{R} \times \mathbb{R}^n$, where ν denotes the inner normal to $\partial\Omega$, and semilinear if

$$G(x, z, p) = \beta(x).p - \varphi(x, z) \tag{3.4}$$

for all $(x, z, p) \in \partial\Omega \times \mathbb{R} \times \mathbb{R}^n$, where β and φ are smooth functions on $\partial\Omega$ and $\partial\Omega \times \mathbb{R}$, respectively, normalized with $|\beta| = 1$.

Letting B denote a smooth positive function on $\Omega \times \mathbb{R} \times \mathbb{R}^n$, we consider oblique boundary value problems of the form:

$$PDE: \quad \mathcal{F}[u] = B(\cdot, u, Du), \; U \in \Gamma \text{ in } \Omega, \tag{3.5}$$

$$BC: \quad \mathcal{G}[u] = 0 \text{ on } \partial\Omega. \tag{3.6}$$

We explain now how the concept of c-convexity in optimal transportation leads to the natural notion of domain convexity for the boundary value problem (3.5), (3.6). First we express the c-convexity, (uniform c-convexity), of a smooth connected domain Ω as a boundary condition,

$$[-D_i\nu_j(x) + D_{p_k}A_{ij}(x, p)\nu_k(x)]\tau_i\tau_j \geq 0, (> 0), \tag{3.7}$$

for all $x \in \partial\Omega$, τ tangential at x and $G(x, p) := \Phi^* \circ Y(x, p) \geq 0$, where the smooth defining function Φ^* satisfies $\Phi^* > 0$ in Ω, $\Phi^* = 0$, $D\Phi^* \neq 0$ on $\partial\Omega$. This characterization was the starting point for global regularity in [26]. We observe now that it essentially depends only on the cone Γ, matrix function A and boundary operator $\mathcal{G}$. For general A in (3.2), we define the A-curvature of $\partial\Omega$ at $(x, z, p) \in \partial\Omega \times \mathbb{R} \times \mathbb{R}^n$ by

$$\mathcal{K}_A[\partial\Omega](x, z, p) = -\delta\nu(x) + P(x)[D_pA(x, z, p).\nu(x)]P(x), \tag{3.8}$$

where $\delta = D - \nu D_\nu$ is the tangential gradient and $P = I - \nu \circ \nu$ is the projection onto the tangent space. Then we define $\partial\Omega$ to be uniformly (Γ, A, G)- convex, with respect to $u \in C^0(\partial\Omega)$, at $x \in \partial\Omega$, if

$$\mathcal{K}_A(x, u(x), p) + \mu\nu \circ \nu \in \Gamma, \tag{3.9}$$

for some $\mu = \mu(x, p)$ and for all p such that $G(x, u(x), p) \geq 0$.

It follows that if A is regular at (x, z, p) then $\nu.D_p\mathcal{K}_A(x, z, p) \geq 0$ so that $\mathcal{K}_A[\partial\Omega]$ is nondecreasing in $p.\nu$ and thus we need only $G(x, u(x), p) = 0$ in (3.9) for the case of the semilinear Neumann problem, that is when $\beta = \nu$.

Moreover if A is strictly regular, then $\nu.D_p\mathcal{K}_A\tau.\tau > 0$, for all tangential τ, so that we obtain a natural association with A of oblique boundary operators, for example given by $G = F' \circ K_A$ for a smooth increasing function F' on an appropriate subset of $\mathbb{S}^n$, for example $F' = S_1$.

When Γ is orthogonally invariant, (so we can assume $\Gamma \subset \mathbb{R}^n$), defining

$$\Gamma' = \{\lambda' \in \mathbb{R}^{n-1} \mid (\lambda', \mu) \in \Gamma, \} \quad \text{for some } \mu > 0, \tag{3.10}$$

we see that (3.9) is equivalent to $\tilde{\kappa}' \in \Gamma'$, where $(\tilde{\kappa}', 0)$ denotes the eigenvalues of $\mathcal{K}_A$. As examples, we have for our fundamental cones in examples (i) and $(iv)'$,

$$\Gamma'_k = \Gamma_{k-1} \text{ for } k > 1, \; \Gamma'_1 = \mathbb{R}^n, \; \mathcal{P}'_m = \mathcal{P}_m, \; m < n. \tag{3.11}$$

The following interior and local boundary second derivative estimates are proved, in a more general form, in [7, 9].

Theorem 3.1. *Assume:*

- *u is a smooth admissible solution of the boundary value problem:*

$$\mathcal{F}[u] = \psi(\cdot, u), \ U \in \Gamma \text{ in } \Omega, \quad \inf \psi > 0, \tag{3.12}$$

$$\mathcal{G}[u] = \beta.Du - \varphi(\cdot, u) = 0 \text{ on } \partial\Omega, \quad \beta.\nu > 0; \tag{3.13}$$

- *F is positive, increasing, one-homogeneous and concave in $\Gamma \subset \mathcal{P}_{n-1}$, $F = 0$ on $\partial\Gamma$;*
- *A is strictly regular for $x \in B_R \cap \Omega$, $z = u(x)$, $p \in \mathbb{R}^n$, for some ball $B_R = B_R(x_0) \subset \mathbb{R}^n$;*
- *$\partial\Omega$ is uniformly (Γ, A, G)-convex with respect to u on $\partial\Omega \cap B_R$.*

Then we have the local second derivative estimate,

$$\sup_{\Omega \cap B_{R/2}} |D^2 u| \leq C, \tag{3.14}$$

where C is a constant depending on $\Omega, A, G, \Gamma, \psi$ and $|u|_{1;B_R \cap \Omega}$.

We remark that instead of assuming F is concave we may more generally assume that $\mu \circ F$ is concave for some smooth increasing function μ on $(0, \infty)$ and instead of assuming $F = 0$ on $\partial\Gamma$, we only need $F < \inf \psi$ on $\partial\Gamma$. From our previous remark about the Neumann case $\beta = \nu$, we only need to assume $\partial\Omega$ is uniformly (Γ, A, G)-convex with respect to u in the weaker sense that (3.9) need only hold for $p = Du(x)$, $x \in \partial\Omega \cap B_R$. Also such estimates are not true in general for the standard Hessian equation, $A = 0$, even in the global case when $\beta \neq \nu$, as is already known in the Monge–Ampère case, $F = \det^{1/n}$ [27, 30].

There are various alternative conditions for gradient estimates and these do not require any geometric conditions on the boundary $\partial\Omega$. Keeping in mind our application to the boundary value problem (1.4), (1.5), we have the following estimate, from [7], which extends that in [12] for (1.4), (1.7).

Theorem 3.2. *Assume*

- *u is a smooth admissible solution of the boundary value problem* (3.12), (3.13)*,*
- *F is positive, increasing and one-homogeneous in Γ, with $F = 0$ on $\partial\Gamma$,*
- *A is uniformly regular, in the sense that*

$$D_{p_k p_l} A_{ij}(x, z, p)\xi_i \xi_j \eta_k \eta_l \geq \lambda_0 |\xi|^2 |\eta|^2 - \tfrac{1}{\lambda_0}(\xi.\eta)^2, \tag{3.15}$$

for all $\xi, \eta \in \mathbb{R}^n$, $x \in B_R \cap \Omega$, $z = u(x)$, $p \in \mathbb{R}^n$, for some constant $\lambda_0 > 0$,

- *A has quadratic growth, in the sense that*

$$A_x, A_z, |p|A = O(|p|^2) \text{ as } p \to \infty, \quad A_z \leq 0. \tag{3.16}$$

Then we have the local gradient estimate

$$\sup_{\Omega \cap B_{R/2}} |Du| \leq C, \tag{3.17}$$

where C is a constant depending on $\Omega, A, G, \Gamma, \psi, \varphi$ and $|u|_{0;B_R \cap \Omega}$.

From Theorems 3.1 and 3.2, follows the existence of classical solutions under appropriate monotonicity conditions on φ and ψ or a priori L^∞ bounds.

Note that to dispense with the uniformly regular condition (3.15) in general, we also need F orthogonally invariant, $|\frac{\beta}{\beta.\nu} - \nu| < 1/\sqrt{n}$, $A = o(|p|^2)$ in (3.16) and either F concave and also $p.A_p \leq o(|p|^2)$ in (3.16) or $F_i / \sum F_i$ bounded away from zero, whenever the eigenvalue λ_i is negative, [7, 9]. In the special case when Γ is the positive cone K^+, we can simply replace (3.15) and (3.16) by a lower quadratic bound $A \geq O(|p|^2)$, [10].

4. Application to conformal geometry

Specializing to $\mathbb{R}^n$, we have

$$A = \frac{1}{2}|p|^2 I - p \otimes p, \tag{4.1}$$

so that

$$P[D_p A(\cdot, u, Du).\nu]P = D_\nu u(I - \nu \otimes \nu) = \varphi(\cdot, u)(I - \nu \otimes \nu) \tag{4.2}$$

on $\partial\Omega$. Consequently we have local second (and first) derivative bounds for solutions of (3.12) and (3.13), with $\beta = \nu$, if F is orthogonally invariant and $\tilde{\kappa} \in \Gamma'$, where $\tilde{\kappa}_i = \kappa_i + \varphi(\cdot, u)$. For the special case of the mean curvature, $\tilde{f} = S_1$, we have, from (1.5),

$$\tilde{\kappa}_i = \left(\kappa_i - \frac{1}{n-1}\sum \kappa_i\right) + \frac{1}{n-1}\tilde{\psi} e^{-u}, \tag{4.3}$$

which clearly includes the umbilic boundary case, $\kappa_i = \frac{1}{n-1}\sum \kappa_i$, [2, 12], by virtue of our assumed positivity of $\tilde{\psi}$. Note also that when $f = f_k$ and $\tilde{f} = f_l$ for $l \geq k-1$, no convexity condition on Ω is needed for second derivatives estimates. In the degenerate case $\tilde{\psi} \geq 0$, second derivative estimates are also proved in the umbilic case in [2, 12] and these would also follow as a limiting case of our methods. In fact, from our barrier constructions in [7], we would conjecture that the umbilic condition is also necessary in this case. In the “negative” case when the condition $\tilde{\kappa} \in \bar{\Gamma}'$ is violated, we would also expect not to have second derivative estimates at the boundary and a result in this direction has been already given in the umbilic case in [14]. We remark also that in the alternative degenerate cases, $\psi \geq 0$ or f degenerate, as in examples (ii) and (iii) in Section 1, our second derivative estimates in Theorem 3.1 continue to hold but we would not expect classical solutions $u \in C^2(\Omega)$.

Finally we remark that these results readily extend to Riemannian manifolds (M, g), which are conformally flat near their boundaries or when $f = f_n$, as already indicated in the umbilic case for $(\tilde{f}, \tilde{\Gamma}) = (S_1, \Gamma_1)$, in [12]. For further special cases, notably for examples (i) and (ii), this can also be achieved for general manifolds, as for the umbilic case in [11], utilizing the approach to second derivative estimates originating in [16] and a different approach to boundary normal derivative estimates from [18]. As in the Euclidean case, our boundary convexity condition is simply $\tilde{\kappa}_g := \kappa_g + \varphi(\cdot, u)(1, \ldots, n-1) \in \Gamma'$, where φ is given by (1.7), and holds automatically when ∂M is umbilic. To solve our nonlinear Yamabe problem, we would still need a priori solution bounds and so far this has only been achieved in the umbilic case, for locally conformally flat manifolds, in [15].

References

[1] L. Caffarelli, L. Nirenberg, J. Spruck, The Dirichlet problem for nonlinear second-order elliptic equations III: functions of eigenvalues of the Hessians, Acta Math., 155, 261–301, 1985.

[2] S.-Y.S. Chen, Boundary value problems for some fully nonlinear elliptic equations, Calc. Var. Partial Differ. Equ., 30(1), 1–15, 2007.

[3] S.-Y.S. Chen, Conformal deformation on manifolds with boundary, Geom. Funct. Anal., 19, 1029–1064, 2009.

[4] G. Gerhardt, Closed Weingarten hypersurfaces in Riemannian manifolds, J. Differ. Geom., 43, 612–641, 1996.

[5] N. Ivochkina, Solution of the Dirichlet problem for some equations of Monge–Ampère type, Mat. Sb., 128, 403–415, 1985; English translation in Math. USSR-Sb., 56, 403–415, 1987.

[6] F. Jiang, N.S. Trudinger, On Pogorelov estimates in optimal transportation and geometric optics, Bull. Math. Sci., 4, 407–431, 2014.

[7] F. Jiang, N.S. Trudinger, Oblique boundary value problems for augmented Hessian equations I, Bull. Math. Sci., 8, 353–411, 2018.

[8] F. Jiang, N.S. Trudinger, Oblique boundary value problems for augmented Hessian equations II, Nonlinear Anal., 154, 148–173, 2017.

[9] F. Jiang, N.S. Trudinger, Oblique boundary value problems for augmented Hessian equations III, Commun. Partial Differ. Equ.,44, 708–748, 2019.

[10] F. Jiang, N.S. Trudinger, N. Xiang, On the Neumann problem for Monge–Ampère type equations, Can. J. Math., 68, 1334–1361, 2016.

[11] Q. Jin, Local Hessian estimates for some conformally invariant fully nonlinear equations with boundary conditions, Differ. Int. Equ., 20, 121–132, 2007.

[12] Q. Jin, A. Li, Y.Y. Li, Estimates and existence results for a fully nonlinear Yamabe problem on manifolds with boundary, Calc. Var. Partial Differ. Equ., 28, 509–543, 2007.

[13] A. Li, Y.Y. Li, A fully nonlinear version of the Yamabe problem on manifolds with boundary, J. Eur. Math. Soc., 8(2), 295–316, 2006.

[14] Y.Y. Li, L. Nguyen, Counterexamples to C^2 boundary estimates for a fully nonlinear Yamabe problem on manifolds with boundary, Adv. Nonlinear Stud., 12, 783–797, 2012.

[15] Y.Y. Li, L. Nguyen, A fully nonlinear version of the Yamabe problem on locally conformally flat manifolds with umbilic boundary, Adv. Math., 251, 87–110, 2014.

[16] P.L. Lions, N.S. Trudinger, J. Urbas, The Neumann problem for equations of Monge–Ampère type, Commun. Pure Appl. Math., 39, 539–563, 1986.

[17] G. Loeper, On the regularity of solutions of optimal transportation problems, Acta Math., 202, 241–283, 2009.

[18] X.-N. Ma, G. Qiu, The Neumann problem for Hessian equations, Commun. Math. Phys. 366, 1–28, 2019.

[19] X.-N. Ma, N.S. Trudinger, X.-J. Wang, Regularity of potential functions of the optimal transportation problem, Arch. Rat. Mech. Anal., 177, 151–183, 2005.

[20] R. Schoen, Conformal deformation of a Riemannian metric to constant scalar curvature, J. Differ. Geom., 20, 479–495, 1984.

[21] J.-P. Sha, p-convex Riemannian manifolds, Invent. Math., 83, 437–447, 1986.

[22] N.S. Trudinger, Remarks concerning the conformal deformation of Riemannian structures on compact manifolds, Ann. Scuola Norm. Sup. Pisa, (3) 22, 265–274, 1968.

[23] N.S. Trudinger, On the Dirichlet problem for Hessian equations, Acta Math., 175, 151–164, 1995.

[24] N.S. Trudinger, Recent developments in elliptic partial differential equations of Monge–Ampère type, Proc. Int. Cong. Math., Madrid, 3, 291–302, 2006.

[25] N.S. Trudinger, A note on global regularity in optimal transportation, Bull. Math. Sci., 3(3), 551–557, 2013.

[26] N.S. Trudinger, X.-J. Wang, On the second boundary value problem for Monge–Ampère type equations and optimal transportation, Ann. Scuola Norm. Sup. Pisa Cl. Sci., (5) VIII, 143–174, 2009.

[27] J. Urbas, Nonlinear oblique boundary value problems for Hessian equations in two dimensions, Ann. Inst. Henri Poincaré – Analyse Non Linear, 12, 507–575, 1995.

[28] C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics, American Mathematical Society, Providence, Rhode Island, 2003.

[29] C. Villani, Optimal Transport – Old and new, Grundlehren, Springer, Berlin, 2008.

[30] X.-J. Wang, Oblique derivative problems for the equations of Monge–Ampère type, Chinese J. Contemp. Math., 13, 13–22, 1992.

[31] H. Yamabe, On a deformation of Riemannian structures on compact manifolds, Osaka Math. J.,12, 21–37, 1960.

Neil S. Trudinger
Mathematical Sciences Institute
The Australian National University
Canberra ACT 0200, Australia
e-mail: `Neil.Trudinger@anu.edu.au`

Progress in Mathematics, Vol. 333, 521–536

Special Lagrangian Equations

Yu Yuan

Dedicated to Professor Gang Tian on the occasion of his 60th birthday

Abstract. We survey special Lagrangian equation and its related fully nonlinear elliptic and parabolic equations: definition, geometric background, basic properties, and progress. These include the rigidity of entire solutions, a priori Hessian estimates, construction of singular solutions, existence, the counterparts in the parabolic-curvature flow-settings, and open problems.

Mathematics Subject Classification (2010). 35J96, 35K96, 53A10, 53C38, 53C44.

Keywords. Special Lagrangian equation, Monge–Ampère equation, fully nonlinear equation, minimal surface, calibrated geometry, curvature flows.

1. Introduction

1.1. Definition of the equation

We start with a scalar function u with its gradient Du and Hessian D^2u. The real symmetric matrix D^2u has n many real eigenvalues $\lambda_1, \dots, \lambda_n$. Adding them together, we have the Laplace equation

$$\Delta u = \lambda_1 + \cdots + \lambda_n = c;$$

multiplying them together, we have the Monge–Ampère equation

$$\ln \det D^2 u = \ln \lambda_1 + \cdots + \ln \lambda_n = c. \tag{1.1}$$

Switching from the logarithm function to the inverse tangent function, we then have the special Lagrangian equation

$$\arctan D^2 u = \arctan \lambda_1 + \cdots + \arctan \lambda_n = \Theta. \tag{1.2}$$

This work is partially supported by an NSF grant.

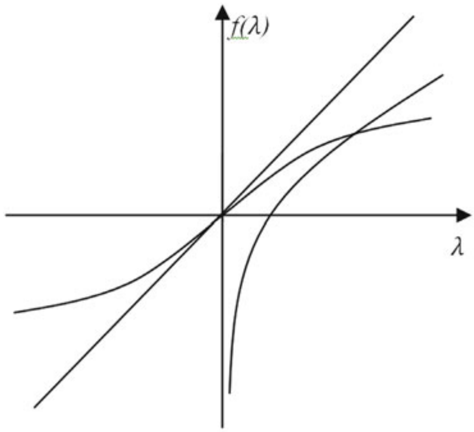


FIGURE 1. Elliptic equation corresponds to a monotonic function

The fundamental symmetric algebraic combination of those eigenvalues forms the general σ_k-equation

$$\sigma_k(\lambda) := \sum_{1\le i_1<\cdots<i_k\le n} \lambda_{i_1}\cdots\lambda_{i_k} = c.$$

General analytic combinations generate general second-order equations

$$F(D^2u) = f(\lambda) = 0. \tag{1.3}$$

If $f(\lambda)$ is monotonic in λ_i, then the equation is elliptic (Figure 1). In principle, when the defining function f is convex (or concave), the regularity of solutions is easier to study; otherwise, it is more complicated.

1.2. Special Lagrangian submanifold background of the equation

If a half-codimensional graph $(x, F(x)) \in \mathbb{R}^n \times \mathbb{R}^n$ has a potential u such that $F(x) = Du(x)$, then it is called a Lagrangian graph. Certainly, the vector field $F(x)$ having a potential is equivalent to it being irrotational. Meanwhile, if the tangent space T of the Lagrangian submanifold is perpendicular to JT at each point, with J being the complex structure of $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{C}^n$, then $F(x)$ has a potential. Special Lagrangian submanifold means its volume is minimizing compared to all submanifolds (Lagragian or not) with the same boundary.

Harvey–Lawson[14] showed that the "gradient" graph $(x, Du(x))$ is volume minimizing if and only if u satisfies special Lagrangian equation (1.2), by applying the fundamental theorem of calculus to a calibration, namely the real closed n form $\mathrm{Re}(e^{-\sqrt{-1}\Theta}dz_1\wedge\cdots\wedge dz_n)$. One obtains odd- as well as even-dimensional volume minimizing submanifolds from solving the special Lagrangian equation. Previously, the only known high-codimensional volume minimizing submanifolds were real even-dimensional complex submanifolds; the volume minimality was proved through applying the fundamental theorem of calculus to the real closed $2k$ form

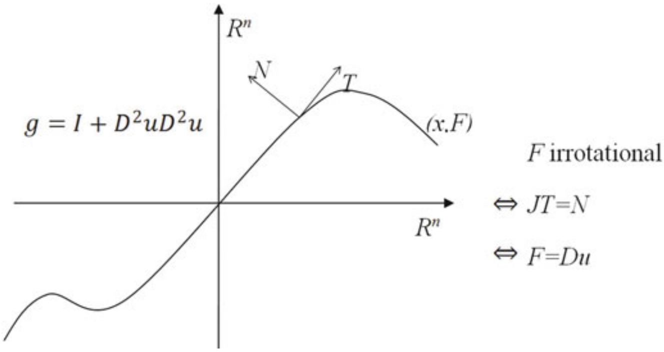


FIGURE 2. Lagrangian submanifold

$\frac{1}{k!}\omega^k$ by Wirtinger, where $\omega = \frac{1}{2\sqrt{-1}} \sum_{i=1}^n dz_i \wedge d\bar{z}_i$. Moreover, the volume minimality of codimensional one minimal graph $(x, f(x))$ over convex domains can also be proved by applying the fundamental theorem of calculus to variable coefficient n form

$$\frac{1}{\sqrt{1+|Df|^2}}\left[dx_1\wedge\cdots\wedge dx_n+\sum_{i=1}^n(-1)^{i-1}f_i dx_1\wedge\cdots\wedge\widehat{dx_i}\wedge\cdots\wedge dx_n\wedge dx_{n+1}\right].$$

This form is closed because f satisfies the minimal surface equation

$$\operatorname{div}\left(Df/\sqrt{1+|Df|^2}\right) = 0.$$

Interestingly, there is an analogous presentation for the Monge–Ampère equation. Indeed, consider spacelike Lagrangian submanifolds in $\mathbb{R}^n \times \mathbb{R}^n$ with pseudo-Euclidean ambient metric $dx^2 - dy^2$ or $dxdy$; we can show that a spacelike "gradient" graph of u is volume maximizing if and only if u satisfies Monge–Ampère equation (1.1). In passing, let us recall the potential $|x|^{-1}$ for the three-dimensional gravitational field $-(x_1,x_2,x_3)\,|x|^{-3}$ satisfies the Laplace equation $\Delta\,|x|^{-1} = 0$.

1.3. Algebraic form of the equation

From the eigenvalues $\lambda_1, \ldots, \lambda_n$ of D^2u we define a complex number

$$z := (1+\sqrt{-1}\lambda_1)\cdots(1+\sqrt{-1}\lambda_n) = (1-\sigma_2+\cdots)+\sqrt{-1}(\sigma_1-\sigma_3+\cdots).$$

Denoting the phase by $\Theta = \arctan D^2u$, z can also be written as

$$z = \sqrt{(1+\lambda_1^2)\cdots(1+\lambda_n^2)}(\cos\Theta+\sqrt{-1}\sin\Theta).$$

Obviously, z is perpendicular to complex number $-\sin\Theta+\sqrt{-1}\cos\Theta$ (Figure 3), such that u satisfies

$$\Sigma := \cos\Theta(\sigma_1-\sigma_3+\cdots)-\sin\Theta(1-\sigma_2+\cdots) = 0. \tag{1.4}$$

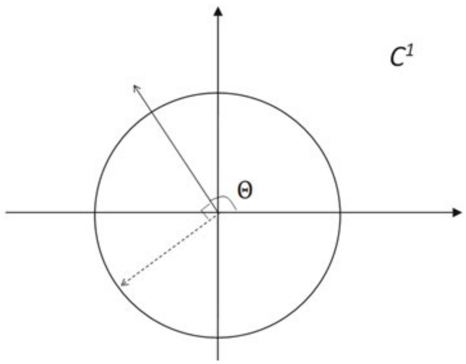


FIGURE 3. Phase $\Theta = \arctan D^2u$

Note that σ_k has a divergence structure; thus, when u satisfies (1.2), that is, Θ is a constant, (1.4) is also an equation in divergence form. In particular, equation (1.4) has the following special forms:

- $n = 2, \Theta = 0$: $\sigma_1 = 0$;
- $n = 2$ or 3, $\Theta = \pm\frac{\pi}{2}$: $\sigma_2 = 1$;
- $n = 3, \Theta = 0$ or $\pm\pi$: $\sigma_3 = \sigma_1$, that is $\det D^2u = \Delta u$.

It is worth noticing that the induced metric of the "gradient" graph of u is $g = I + (D^2u)D^2u$, such that its volume element becomes

$$\sqrt{\det g} = \sqrt{(1+\lambda_1^2)\cdots(1+\lambda_n^2)} = \cos\Theta(1 - \sigma_2 + \cdots) + \sin\Theta(\sigma_1 - \sigma_3 + \cdots).$$

When Θ is constant, the above volume element also has a divergence structure.

1.4. Level set of the equation

As mentioned above, the ellipticity of equation (1.3) means the defining function f is monotonic. Actually, we can also give a geometric description of the ellipticity. Consider the level set of f in λ-space; the ellipticity of the equation is equivalent to the fact that the normal of the level set $N := D_\lambda f$ falls into the positive cone Γ; namely all components of N are positive. Further, uniform ellipticity means N is uniformly inside the positive cone Γ, or all components of the unit normal $N/|N|$ have a fixed lower and upper bound. For example, Figure 4 illustrates the level sets of the three-dimensional special Lagrangian equations. In [32] we observed that the level set of the special Lagrangian equation is convex if and only if $|\Theta| \geq (n-2)\frac{\pi}{2}$. Naturally, $(n-2)\frac{\pi}{2}$ is called the critical phase. Solutions are better behaved when their equations are convex. Indeed, we have Bernstein type results for special Lagrangian equations with supercritical phase, and *a priori* estimates and regularity in the critical and supercritical cases. On the other hand, singular solutions do exist in the subcritical case.

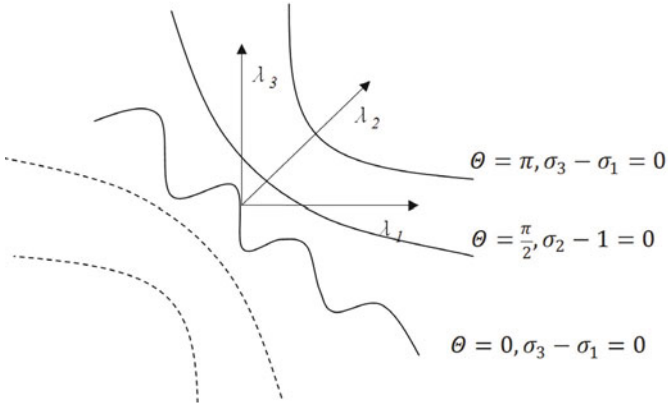


FIGURE 4. Level sets of Θ in λ-space ($n = 3$)

2. Results

2.1. Outline

Once equations are given, the first question to answer is the existence of solutions. Smooth ones cannot be reached at once, in general; worse, they may not even exist. The usual way to compromise is to first search for weak solutions, in the integral sense if the equation has divergence structure, or in the "pointwise integration by parts sense", namely, in the viscosity sense if the equation enjoys a comparison principle. After obtaining those weak solutions, one studies the regularity and other properties of the solutions, such as Liouville or Bernstein type results for entire solutions. All these depend on *a priori* estimates of derivatives of solutions:

$$\|D^2 u\|_{L^\infty(B_1)} \leq C(\|Du\|_{L^\infty(B_2)}) \leq C(\|u\|_{L^\infty(B_3)}).$$

Given the L^∞ bound of the Hessian, the ellipticity of the above fully nonlinear equations becomes uniform, we can apply Evans–Krylov–Safonov theory (for the ones with convexity/concavity, possibly without divergence structure) or Evans–Krylov–De Giorgi–Nash theory (for the ones with convexity/concavity and divergence structure) to obtain $C^{2,\alpha}$ estimates of solutions. For the special Lagrangian equation, this $C^{2,\alpha}$ estimate can also be achieved via geometric measure theory; for the Monge–Ampère equation, earlier in the 1950s, Calabi reached C^3 estimates by interpreting the cubic derivatives in terms of the curvature of the corresponding Hessian metric $g = D^2u$. In turn, iterating the classic Schauder estimates, one gains smoothness of the solutions, and even analyticity, if the smooth equations are also analytic.

2.2. Rigidity of entire solutions

The classic Liouville theorem asserts every entire harmonic function bounded from below or above is a constant. Thus every semiconvex harmonic function is a quadratic one, as its double derivatives are all harmonic with lower bounds, hence constants. Similarly, every entire (convex) solution to the Monge–Ampère equation $\det D^2u = 1$ is quadratic. This was first proved in two-dimensional case by Jörgens, later in low dimensions by Calabi, and in all dimensions by Pogorelov. Also, Cheng–Yau had a geometric proof. For the special Lagrangian equation $\arctan D^2u = \Theta$, Yuan [31] showed every entire convex solution is quadratic. Actually the convexity condition can be relaxed to a semiconvex one

$$D^2u \geq -\tan\frac{\pi}{6} - \varepsilon(n),$$

where $\varepsilon(n)$ is a small-dimensional constant. On the other hand, Yuan [32] replaced the convexity condition of solutions with the phase condition on the equation

$$|\Theta| > (n-2)\frac{\pi}{2}$$

for a rigidity result. This shows the phase $(n-2)\frac{\pi}{2}$ is indeed a critical one: all entire solutions to the special Lagrangian equation with supercritical phase must be quadratic. It is a Bernstein type result. Chang–Yuan [4] proved a similar Liouville type result for the σ_2-equation: If u is an entire solution to $\sigma_2(D^2u) = 1$ such that

$$D^2u \geq \left(\delta - \sqrt{\frac{2}{n(n-1)}}\right) I,$$

for any small fixed $\delta > 0$, then u is quadratic. In all the above rigidity results, certain convexity of the solutions u or lower bound of the Hessian D^2u is needed. Otherwise, there are counterexamples. For example, when $n = 2$, $u = \sin x_1 e^{x_2}$ is a nontrivial solution to $\arctan D^2u = 0$. Whereas for $n = 3$, Warren [27] found a precious explicit solution

$$u = (x_1^2 + x_2^2)e^{x_3} - e^{x_3} + \frac{1}{4}e^{-x_3}$$

to the equation $\arctan D^2u = \frac{\pi}{2}$ or $\sigma_2(D^2u) = 1$.

In the following, we present the idea of showing the rigidity of entire solutions to special Lagrangian equation in the two-dimensional case as an example. Given an entire solution u to $\arctan\lambda_1 + \arctan\lambda_2 = \Theta > 0$. First, notice that every dihedral angle $\arctan\lambda_1$ or $\arctan\lambda_2$ between the tangent plane of the "gradient" graph $(x, Du) \subset \mathbb{R}^2 \times \mathbb{R}^2$ and x plane has a lower bound $\Theta - \pi/2$. So after we rotate the x coordinate plane to another one $\bar{x} = x\cos\Theta/2 + y\sin\Theta/2$, the original tangent plane and the new coordinate $\bar{x}$ plane form the new dihedral angles $(\arctan\lambda_1 - \Theta/2, \arctan\lambda_2 - \Theta/2)$. Those two angles fall into the interval $(-\pi/2 + \Theta/2, \pi/2 - \Theta/2)$. This means the old "gradient" graph is still a graph in the new coordinate system $\bar{x}$ and $\bar{y} = -x\sin\Theta/2 + y\cos\Theta/2$. Further, it is another "gradient" graph $(\bar{x}, D\bar{u})$ corresponding to a new potential $\bar{u}$. It is easy to see the Hessian $D^2\bar{u}$ of the new potential $\bar{u}$ is bounded, and moreover, its eigenvalues satisfy equation $\arctan \bar{\lambda}_1 + \arctan \bar{\lambda}_2 = 0$. Thus, we have obtained an entire harmonic function $\bar{u}$ with bounded Hessian; in turn, $\bar{u}$ is quadratic. From this, we know the "gradient" graph is a plane. Therefore, the original entire solution u is quadratic.

For higher-dimensional special Lagrangian equation with supercritical phase, via a similar coordinate rotation, we get a new entire solution to special Lagrangian equation with critical phase. Applying Evans–Krylov's $C^{2,\alpha}$ estimates (really its scaled version in the entire space), we know the new Hessian is a constant matrix. Therefore, the original entire solution u is quadratic.

The above Liouville type result for the σ_2-equation can be proved in a similar way. As for the rigidity of entire semiconvex solutions to the special Lagrangian equation with subcritical phase, more effort is required, because the new equation loses convexity.

2.3. A priori estimates for Monge–Ampère equation

In the 1950s, Heinz [15] studied *a priori* estimates for the two-dimensional Monge–Ampère equation, a particular case is the following: If u is a solution to the equation $\det D^2 u = 1$ in the unit ball, then

$$|D^2 u(0)| \leq C(\|u\|_{L_\infty(B_1)}).$$

Later, this result was achieved in the higher-dimensional case by Pogorelov [19], but with a strict convexity restriction. Chou–Wang [10] proved similar estimates for "k-strictly" convex solutions to σ_k-equation by adapting Pogorelov's technique. Trudinger [22], Urbas [23], and Bao–Chen [1] obtained *a priori* Hessian bound in terms of the integral of the Hessian for solutions to σ_k-equation and its quotient forms. Bao–Chen–Guan–Ji [2] proved *a priori* Hessian estimates for strictly convex solutions to the quotient σ_n/σ_k type equations. If no strict convexity restriction is assumed, then Pogorelov [19] constructed his famous singular $C^{1,1-\frac{2}{n}}$ solution to the Monge–Ampère equation $\det D^2 u = 1$. Caffarelli provided merely Lipschitz solution to the Monge–Ampère equation with variable right-hand side. Furthermore, Caffarelli–Yuan obtained Lipschitz and $C^{1,\alpha}$, with α being any rational number in $(0, 1 - \frac{2}{n}]$, singular solutions to the Monge–Ampère equation $\det D^2 u = 1$.

2.4. A priori estimates for special Lagrangian equation with critical and supercritical phases

For special Lagrangian equation with critical and supercritical phases

$$\arctan D^2 u = \Theta, \quad |\Theta| \geq (n-2)\frac{\pi}{2}, \tag{2.1}$$

Wang–Yuan [24] proved the following *a priori* estimates for the Hessian (Figure 5): Suppose u is a smooth solution to special Lagrangian equation (2.1) in n-dimensional ($n \geq 3$) unit ball $B_1 \subset \mathbb{R}^n$. Then for $|\Theta| \geq (n-2)\frac{\pi}{2}$,

$$|D^2 u(0)| \leq C(n) \exp\left(C(n) \|Du\|^{2n-2}_{L^\infty(B_1)} \right); \tag{2.2}$$

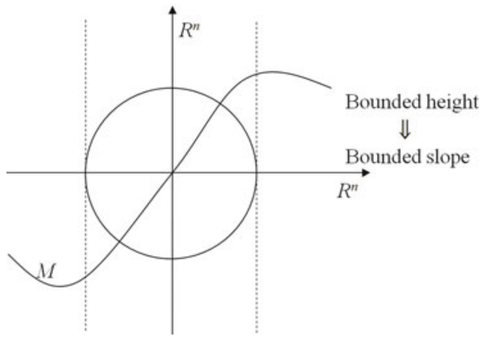


FIGURE 5. *A priori* estimate for Hessian D^2u

and for $|\Theta| = (n-2)\frac{\pi}{2}$,

$$|D^2u(0)| \leq C(n) \exp\left(C(n)\|Du\|_{L^\infty(B_1)}^{2n-4}\right). \tag{2.3}$$

Combined with the gradient estimates for equation (2.1) by Warren–Yuan [30]

$$\max_{B_R(0)} |Du| \leq C(n) \left(\operatorname{osc}_{B_{2R}(0)} \frac{u}{R} + 1\right),$$

we immediately obtain the estimate for D^2u in terms of solution u itself. Actually the gradient estimates for equation (2.1) can be improved slightly [33]

$$\max_{B_R(0)} |Du| \leq C(n) \operatorname{osc}_{B_{2R}(0)} \frac{u}{R}.$$

For $n = 3$, earlier on Warren–Yuan [29, 30] proved *a priori* Hessian estimates in the critical and supercritical cases. Chen–Warren–Yuan [9] showed similar estimates for convex solutions to the special Lagrangian equation. Warren–Yuan [28] derived Hessian estimates for solutions to two-dimensional special Lagrangian equation

$$|D^2u(0)| \leq C(2) \exp\left(\frac{C(2)}{|\sin\Theta|^{\frac{3}{2}}}\|Du\|_{L^\infty(B_1)}\right).$$

From the minimal surface example by Finn [13] via Heinz transformation [16], one sees that the above Hessian bound in terms of linear exponential of gradient is sharp. For $n \geq 3$, corresponding sharp Hessian estimates are not known. As applications of the above *a priori* estimates, we immediately know all C^0 viscosity solutions to (2.1) are smooth, and even analytic. For comparison, in the 1980s Caffarelli–Nirenberg–Spruck [3] obtained the interior regularity for solutions with C^4 smooth boundary data to the special Lagrangian equation (1.2) with $|\Theta| = [\frac{n-1}{2}]\pi$. Another direct consequence is that every entire solution with quadratic growth to critical phase special Lagrangian equation

$$\arctan D^2u = (n-2)\frac{\pi}{2}$$

is quadratic.

We briefly explain the possible reason and the idea in obtaining the Hessian estimates. Heuristically, the Hessian of any solution to (2.1) in certain norm is strongly subharmonic, such that its reciprocal is superharmonic. Thus, if this superharmonic quantity is zero somewhere, then it is zero everywhere. That is, if the Hessian is unbounded at one point, then it must be unbounded everywhere. Roughly, this contradicts the graphical picture of the corresponding "gradient" graph (x, Du). A key point in the argument is to show

$$\Delta_g \frac{1}{\sqrt{1+\lambda_{\max}^2}} \leq 0,$$

where $\lambda_{\max}$ is the maximal eigenvalue of D^2u, and Δ_g is the Laplace operator with respect to the induced metric of the Lagrangian submanifold. The above superharmonicity inequality is equivalent to the Jacobi inequality

$$\Delta_g \ln \sqrt{1+\lambda_{\max}^2} \geq |\nabla_g \ln \sqrt{1+\lambda_{\max}^2}|^2.$$

The outline of the argument is to start from the mean value inequality on the minimal Lagrangian graph, relying on the Sobolev inequality, Jacobi inequality, and the divergence structure of $\sigma_k(D^2u)$, then to control the integral average of the logarithm of the maximal eigenvalue $\ln \sqrt{1+\lambda_{\max}^2}$ in terms of the gradient of the solution. The process can be viewed as an arduous nonlinearization of the mean value equality proof for the *a priori* estimate of the Hessian in terms of the gradient of a harmonic function.

2.5. Singular solutions to special Lagrangian equation with subcritical phase

For the special Lagrangian equation with subcritical phase $|\Theta| < (n-2)\frac{\pi}{2}$, the above *a priori* Hessian estimates are not valid. Nadirashvili–Vladuct [17] first constructed $C^{1,\frac{1}{3}}$ singular solutions to three-dimensional special Lagrangian equation

$$\sum_{i=1}^{3} \arctan \lambda_i = 0.$$

For the three-dimensional special Lagrangian equation with arbitrary subcritical phase $|\Theta| \in (-\frac{\pi}{2}, \frac{\pi}{2})$, Wang–Yuan [25] constructed $C^{1,r}$ singular solutions, where $r = \frac{1}{2m-1} \in (0, \frac{1}{3}], m = 2, 3, \ldots$. To produce higher-dimensional singular solutions to subcritical special Lagrangian equation, we only need to add quadratics in terms of the extra variables to those three-dimensional singular solutions. The main new tool in [25] is a partial $U(n)$ coordinate rotation, the difficulty lies in proving that, after the rotation of preliminary solutions, the special Lagrangian submanifold is still a graph. The concrete construction goes as follows: first consider critical phase special Lagrangian equation $|\Theta| = \frac{\pi}{2}$; its algebraic equivalent form is the σ_2-equation

$$\sigma_2(D^2u) = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 = 1. \tag{2.4}$$

We construct a family of approximate polynomials P of order $2m$ such that the dihedral angles between the tangent plane of the corresponding "gradient" graph and the x coordinate plane are roughly $(0^-, \frac{\pi}{4}, \frac{\pi}{4})$ (Figure 6). Then taking this fam-

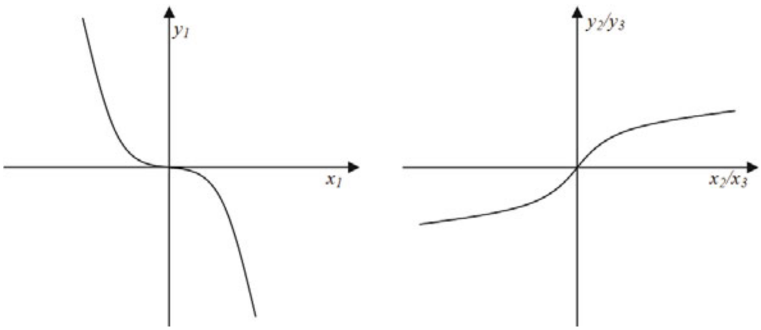


FIGURE 6. Construction of dihedral angles

ily of approximate solutions as initial data, we obtain a family of exact solutions u to equation (2.4) by Cauchy–Kowalevskaya. Next, we make a $U(3)$ coordinate rotation of $-\frac{\pi}{2}$, namely the Legendre transformation of u, to get singular $\tilde{u}$ with roughly the dihedral angles $(\frac{\pi}{2}^-, -\frac{\pi}{4}, -\frac{\pi}{4})$ satisfying the special Lagrangian equation with zero phase. Finally using a "horizontal" rotation which keeps the z_1 plane invariant, we can adjust the phase of $\tilde{u}$ to any subcritical one, to obtain the desired singular solutions.

3. Curvature flows with potential

3.1. Lagrangian mean curvature flow in Euclidean space

Under mean curvature flow, a submanifold is being deformed according to its mean curvature in the ambient space. The (effective) equation is

$$\partial_t X = \mathbf{H} = \Delta_g X,$$

where $X(\cdot, t)$ is a family of immersed submanifolds with time parameter, $\mathbf{H}$ is the mean curvature, and g is the induced metric from the ambient space. A known fact is that the Lagrangian structure of Lagrangian submanifolds is preserved under the mean curvature flow Smoczyk [20].

Meanwhile, we consider the following fully nonlinear parabolic equation satisfied by potential $u(x,t)$

$$\partial_t u = \arctan D^2 u. \tag{3.1}$$

Differentiating both sides of the equation with respect to space variables, we have

$$\partial_t (x, Du) = \sum_{i,j=1}^{n} g^{ij} \partial_{ij} (x, Du), \tag{3.2}$$

where parabolic coefficients g^{ij} are the inverse of the induced metric $g = I + D^2uD^2u$ of the "gradient" graph (x, Du) in Euclidean space $(\mathbb{R}^n \times \mathbb{R}^n, dx^2 + dy^2)$. The normal projection of the right-hand side of this equation (3.2) is the mean curvature, thus the effective part of the deformation of the "gradient" graph is indeed equal to its mean curvature. In dimension one, (3.1) and (3.2) respectively simplify to

$$\partial_t u = \arctan u_{xx} \quad \text{and } \partial_t u_x = \frac{u_{xxx}}{1+u_{xx}^2}.$$

For the initial value problem for the potential equation (3.1) of the Lagrangian mean curvature flow, in the periodic case, namely the gradient Du_0 of initial data u_0 is a lift to $\mathbb{R}^n$ of a map from $\mathbb{T}^n$ into itself, applying Krylov's theory for fully nonlinear uniformly parabolic equation with concavity, Smoczyk–Wang [21] showed, under the "uniform" convexity assumption $0 \leq D^2u_0 \leq C$ or equivalently

$$-(1-\delta)I_n \leq D^2u_0 \leq (1-\delta)I_n, \ \ \delta > 0,$$

on the initial data, the long time existence of solutions to equation (3.1). Chau–Chen–He [5] removed the periodicity assumption on Du_0; their *a priori* estimates deteriorate as $\delta \to 0$. For weak solutions to equation (3.1) with continuous initial data on $\mathbb{R}^n$, Chen–Pang [8] proved the long time existence and uniqueness of continuous viscosity solutions. For the standard heat equation $u_t = \Delta u$, it is worth noting here that there are the Tikhonov nonuniqueness example and the finite time blow-up solution $u\,(x,t) = \frac{1}{\sqrt{1-t}} \exp\left(\frac{x^2}{4(1-t)}\right)$. The contrasting phenomena can be explained by the heat conduction coefficient being uniform for the standard heat equation, but degenerate for fully nonlinear parabolic equation (3.1) when the spatial Hessian becomes unbounded. Moreover, saddle solutions to (3.1) could blow up in finite time at the second spatial derivative level.

Here, we explain a result on long time existence of smooth solutions with almost convexity by Chau–Chen–Yuan [6]. If initial potential u_0 satisfies

$$-(1+\eta)I \leq D^2u_0 \leq (1+\eta)I, \tag{3.3}$$

where $\eta = \eta(n)$ is a small-dimensional positive constant, then the potential equation (3.1) of the Lagrangian mean curvature flow has a unique long time solution $u(x,t) : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}^1$ such that u is smooth for $t > 0$; and moreover

1) $-\sqrt{3}I \leq D^2u(x,t) \leq \sqrt{3}I$ for any $t > 0$;
2) $\|D^l u\|_{L^\infty(\mathbb{R}^n)} \leq C_l t^{2-l}$, for any $t > 0, l \geq 3$;
3) $Du(x,t)$ is $C^{\frac{1}{2}}$ with respect to t at $t = 0$.

Relying on this result, via the $U(n)$ coordinate rotation technique described in the above, we immediately obtain long time existence of smooth solutions and related estimates to equation (3.1) with locally $C^{1,1}$ convex initial data or initial data u_0 with a large phase $\arctan D^2u_0 \geq (n-1)\frac{\pi}{2}$.

We point out that one cannot apply Krylov's theory for fully nonlinear uniformly parabolic equation with convexity here under the almost convexity (3.3), as the convexity condition fails. To overcome the difficulty, Chau–Chen–Yuan used approximation and the compactness of the solution space. The key tools are the uniqueness of solutions by Chen–Pang and the parabolic Schauder estimate for the potential equation (3.1) of the Lagrangian mean curvature flow with certain convexity but not the "full" convexity condition. And not surprisingly, the *a priori* estimates by Nguyen–Yuan [18] are based on the Bernstein–Liouville type results for the corresponding elliptic special Lagrangian equation.

3.2. Lagrangian mean curvature flow in pseudo-Euclidean space and Kähler–Ricci flow on Kähler manifold

We have introduced the parabolic version of the special Lagrangian equation

$$\partial_t v = \arctan D^2 v. \tag{3.4}$$

For the Monge–Ampère equation, we can consider its parabolic version too

$$\partial_t v = \ln\det D^2 v. \tag{3.5}$$

Again differentiating the equation with respect to spatial variables, we have

$$\partial_t (x, Dv) = \sum_{i,j=1}^{n} g^{ij} \partial_{ij} (x, Dv),$$

where parabolic coefficients g^{ij} are the inverse of the induced metric $g = D^2 v$ of the spacelike "gradient" graph in pseudo-Euclidean space $(\mathbb{R}^n \times \mathbb{R}^n, dxdy)$. Similarly, the normal projection of the right-hand side of this equation is the mean curvature; thus, the effective part of the deformation of the "gradient" graph is indeed equal to its mean curvature. We can also consider the parabolic complex Monge–Ampère equation, which is satisfied by a real-valued scalar function v on complex space $\mathbb{C}^m$

$$\partial_t v = \ln\det \partial\bar{\partial} v. \tag{3.6}$$

Differentiating the equation with respect to spatial variables twice, we have

$$\partial_t g_{i\bar{k}} = -R_{i\bar{k}}$$

where $g_{i\bar{k}} = v_{i\bar{k}}$ is the Kähler metric and $R_{i\bar{k}} = -\partial_i \bar{\partial}_k \ln\det \partial\bar{\partial} v$ is the Kähler–Ricci curvature. Thus the second-order parabolic potential equation (3.6) actually corresponds to the Kähler–Ricci flow in geometric analysis.

We investigate a class of self-similar solutions to the above three parabolic equations, that is, shrinking solitons in the form

$$v(x,t) = -tu\left(\frac{x}{\sqrt{-t}}\right).$$

If the above-defined v satisfies the three parabolic equations (3.4), (3.5), and (3.6) respectively, then the profile u respectively satisfies the following three elliptic equations:

$$\arctan D^2 u = \frac{1}{2} x \cdot Du(x) - u(x), \tag{3.7}$$

$$\ln\det D^2 u = \frac{1}{2} x \cdot Du(x) - u(x), \tag{3.8}$$

$$\ln\det \partial\bar{\partial} u = \frac{1}{2} x \cdot Du(x) - u(x). \tag{3.9}$$

For shrinking solitons, Chau–Chen–Yuan [7] proved the following rigidity result:

1) If u is an entire smooth solution to equation (3.7) on $\mathbb{R}^n$, then $u(x) = u(0) + \frac{1}{2}\left\langle D^2u(0)x, x\right\rangle$.
2) If u is an entire convex smooth solution to equation (3.8) on $\mathbb{R}^n$, and satisfies $D^2u(x) \geq \frac{2(n-1)}{|x|^2}$ near ∞, then u is quadratic.
3) If u is an entire complex convex (pluri-subharmonic, $\partial\bar{\partial}u \geq 0$) smooth solution to equation (3.9) on $\mathbb{C}^m$, and satisfies $\partial\bar{\partial}u(x) \geq \frac{2m-1}{2|x|^2}$ near ∞, then u is quadratic.

In fact, after differentiating the parabolic equations with respect to the time variable, Chau–Chen–Yuan observed that the phase function corresponding to shrinking solitons satisfies a second-order elliptic equation with an "amplifying" force term on the whole space. In dimension one, this elliptic equation can be interpreted in terms of acceleration being proportional to velocity. Hence, the changing rate of the phase function cannot be non-zero; in turn, the phase is constant. Further, notice that the right-hand side of the self-similar equation is the "excess of the potential from being quadratic" so we see that the smooth potential must be quadratic.

Let us explain more the argument for the above result by Chau–Chen–Yuan, using the first case as an example. Let $\Theta = \arctan D^2u$. Simple calculation shows that given solution u to equation (3.7), the phase function Θ satisfies

$$\sum\nolimits_{i,j=1}^{n} g^{ij}\partial_{ij}\Theta(x) = \frac{1}{2}x \cdot D\Theta(x), \tag{3.10}$$

Where, g^{ij} being the inverse of the induced metric $g = I + D^2uD^2u$, has an upper bound. The above second-order elliptic equation with the "amplifying " force term allows us to construct a suitable barrier, so that we can prove that Θ attains its minimum at a finite point. Then the strong minimum principle forces Θ to be a constant. Finally, Euler's theorem on homogeneous functions, applied to equation (3.7), leads to the desired quadratic conclusion of u.

As a matter of fact, in the above case of Monge–Ampère, the inverse square lower bound on the induced metrics is a concrete condition for the metric being complete. Now if we assume the metric is complete (abstractly), then the above rigidity result for the shrinking solitons in the Monge–Ampère case (complex as well as real) is also true. This is contained in Drugan–Lu–Yuan [12]. The further observation is that the radial derivative of the phase is the negative of the scalar curvature of the corresponding Kähler metric (3.10). On the other hand, the scalar curvature for self-shrinking solitons is nonnegative. In turn, the phase function attains its maximum at the origin. Similarly we arrive at the rigidity conclusion by applying the strong maximum principle. Heuristically, the non-negativity of scalar curvature R can be seen from its equation

$$\Delta_g R \leq \frac{1}{2}rR_r + R - \frac{1}{m}R^2.$$

If R attains its minimum somewhere, then $0 \leq R_{\min} - R_{\min}^2/m$. It follows that $R \geq 0$. The proof can actually be realized when the metric is complete.

Ding–Xin [11] and Wang [26] respectively proved a Bernstein type result for self-similar real Monge–Ampère equation (3.8) and one-dimensional complex Monge–Ampère equation (3.9); namely every entire solution is quadratic.

4. Problems

Problem 1. Can one find a pointwise argument for the *a priori* Hessian estimates to the special Lagrangian equation? Our proof is in integral form. If possible, it would represent a push-forward for a long time open problem on Hessian estimates for the quadratic symmetric Hessian equation $\sigma_2(D^2u) = 1$. The desire for such a pointwise way is because so far, we have not seen any structure in high dimensions ($n \geq 4$), as in the low-dimensional case ($n \leq 3$) for this equation, resulting in an effective mean value inequality to be employed. Recall for codimension one minimal surface equation $\operatorname{div}\left(Df/\sqrt{1+|Df|^2}\right) = 0$, one has the classic gradient estimates for solutions

$$|Df(0)| \leq C(n) \exp\left[C(n)\|f\|_{L^\infty(B_1)}\right].$$

The proof by Bombieri–De Giorgi–Miranda in the 1960s and its simplification by Trudinger in the 1970s are both in integral form. In the 1980s, Korevaar found a strikingly simple pointwise argument. They are all based on the Jacobi inequality

$$\Delta_g \ln\sqrt{1+|Df|^2} \geq |\nabla_g \ln\sqrt{1+|Df|^2}|^2.$$

Problem 2. Construction of nontrivial entire solutions to the special Lagrangian equation with critical phase $\arctan D^2u = (n-2)\pi/2$ in high dimensions ($n \geq 3$). The construction in dimension three by Warren is through separating variables with adjustment. The key for a systematic method is to search for nontrivial super and sub solutions. This is because we already have the follow-up tool to finish, namely the Hessian estimates in term of the solutions. A more urgent problem is the existence or nonexistence of nontrivial homogeneous-order two solutions to the special Lagrangian equation with subcritical phase in high dimension ($n \geq 5$). The rigidity and regularity for general special Lagrangian equation hinge on it.

Problem 3. Is every entire smooth solution to self-similar complex Monge–Ampère equation $\ln\det\partial\bar{\partial}u = \frac{1}{2}x\cdot Du(x) - u(x)$ quadratic? As mentioned above, it is indeed so in complex dimension one. Now there is known quite a lot of nontrivial entire solutions with corresponding Kähler metric being complete and non-flat to the complex Monge–Ampère equation $\ln\det\partial\bar{\partial}u = 0$, but the self-similar term on the right-hand side of the self-similar equation should still have a strong effect to force entire solutions to be trivial. Just as in the cases of self-similar codimension one minimal surface equation and self-similar special Lagrangian equation, rigidity is available, because of the self-similarity. Once self-similarity is removed, nontrivial entire solutions do exist in both cases.

References

[1] J.-G. Bao and J.-Y. Chen, Optimal regularity for convex strong solutions of special Lagrangian equations in dimension 3. Indiana Univ. Math. J. **52** (2003), no. 5, 1231–1249.

[2] J.-G. Bao, J.-Y. Chen, B. Guan, and M. Ji, Liouville property and regularity of a Hessian quotient equation. Amer. J. Math. **125** (2003), no. 2, 301–316.

[3] L. Caffarelli, L. Nirenberg, and J. Spruck, The Dirichlet problem for nonlinear second-order elliptic equations. III. Functions of the eigenvalues of the Hessian. Acta Math. **155** (1985), no. 3-4, 261–301.

[4] A.S.-Y. Chang, and Y. Yuan, A Liouville problem for the sigma-2 equation. Discrete Contin. Dyn. Syst. **28** (2010), no. 2, 659–664.

[5] A. Chau, J.-Y. Chen, and W.-Y. He, Lagrangian mean curvature flow for entire Lipschitz graphs. Calc. Var. Partial Differ. Eq. **44** (2012), 199–220.

[6] A. Chau, J.-Y. Chen and Y. Yuan, Lagrangian mean curvature flow for entire Lipschitz graphs II. Math. Ann., **357** (2013) 165–183.

[7] A. Chau, J.-Y. Chen and Y. Yuan, Rigidity of entire self-shrinking solutions to curvature flows. J. Reine Angew. Math., **664** (2012), 229–239.

[8] J.-Y. Chen and C. Pang, Uniqueness of viscosity solutions of a geometric fully nonlinear parabolic equation. C.R. Math. Acad. Soc. Paris Ser. **347** (2009), 1031–1034.

[9] J.-Y. Chen, M. Warren, and Y. Yuan, A priori estimate for convex solutions to special Lagrangian equations and its application. Comm. Pure Appl. Math. **62** (2009), no. 4, 583–595.

[10] K.-S. Chou and X.-J. Wang, A variational theory of the Hessian equation. Comm. Pure Appl. Math. **54** (2001), no. 9, 1029–1064.

[11] Q. Ding and Y.-L. Xin, The rigidity theorems for Lagrangian self-shrinkers. J. Reine Angew. Math. **692** (2014), 109–123.

[12] G. Drugan, P. Lu and Y. Yuan, Rigidity of complete entire self-shrinking solutions to Kähler–Ricci flow. Int. Math. Res. Not. IMRN **2015**, no. 12, 3908–3916.

[13] R. Finn, New estimates for equations of minimal surface type. Arch. Rational Mech. Anal. **14** (1963), 337–375.

[14] R. Harvey and H.B. Lawson, Calibrated geometries. Acta Mathematica **148** (1982), 47–157.

[15] E. Heinz, On elliptic Monge–Ampère equations and Weyl's embedding problem. J. Analyse Math. **7** (1959) 1–52.

[16] K. Jörgens, Über die Lösungen der Differentialgleichung $rt - s^2 = 1$. Math. Ann. **127** (1954), 130–134.

[17] N. Nadirashvili and Vlăduţ, Serge, Singular solution to Special Lagrangian Equations. Ann. Inst. H. Poincaré Anal. Non Linéaire **27** (2010), no. 5, 1179–1188

[18] T. Nguyen and Y. Yuan, *A priori* estimates for Lagrangian mean curvature flows. Int. Math. Res. Not. IMRN **2011**, no. 19, 4376–4383.

[19] A.V. Pogorelov, The Minkowski Multidimensional Problem. Translated from the Russian by Vladimir Oliker, Introduction by Louis Nirenberg, Scripta Series in Mathematics, V.H. Winston & Sons, Washington, DC, 1978.

[20] K. Smoczyk, Angle theorems for the Lagrangian mean curvature flow. Math. Z. **240** (2002), 849–883.

[21] K. Smoczyk and M.-T. Wang, Mean curvature flows of Lagrangian submanifolds with convex potentials. J. Differ. Geom. **62** (2002), 243–257.

[22] N.S. Trudinger, Regularity of solutions of fully nonlinear elliptic equations. Boll. Un. Mat. Ital. A (6) **3** (1984), no. 3, 421–430.

[23] J. Urbas, Some interior regularity results for solutions of Hessian equations, Calc. Var. Partial Differential Equations. **11** (2000), no. 1, 1–31.

[24] D.-K. Wang and Y. Yuan, Singular solutions to special Lagrangian equations with subcritical phases and minimal surface systems. Amer. J. Math. **135** (2013), no. 5, 1157–1177.

[25] D.-K. Wang and Y. Yuan, Hessian estimates for special Lagrangian equations with critical and supercritical phases in general dimensions. Amer. J. Math. **136** (2014), no. 2, 481–499.

[26] W.-L. Wang, Rigidity of entire self-shrinking solutions to Kähler–Ricci flow on the complex plane. Proc. Amer. Math. Soc. **145** (2017), no. 7, 3105–3108.

[27] M. Warren, Nonpolynomial entire solutions to σ_k equations. Comm. Partial Differential Equations. **41** (2016), no. 5, 848–853.

[28] M. Warren and Y. Yuan, Explicit gradient estimates for minimal Lagrangian surfaces of dimension two. Math. Z. **262** (2009), no. 4, 867–879.

[29] M. Warren and Y. Yuan, Hessian estimates for the sigma-2 equation in dimension 3. Comm. Pure Appl. Math. **62** (2009), no. 3, 305–321.

[30] M. Warren and Y. Yuan, Hessian and gradient estimates for three dimensional special Lagrangian equations with large phase, Amer. J. Math. **132** (2010), no. 3, 751–770.

[31] Y. Yuan, A Bernstein problem for special Lagrangian equations. Invent. Math. **150** (2002), 117–125.

[32] Y. Yuan, Global solutions to special Lagrangian equations. Proc. AMS **134** (2006), 1355–1358.

[33] Y. Yuan, Lecture Notes on Special Lagrangian Equations, 2015, unpublished.

Yu Yuan
University of Washington
Department of Mathematics
Box 354350
Seattle, WA 98195, USA
e-mail: `yuan@math.washington.edu`

Progress in Mathematics, Vol. 333, 537–544

Positive Scalar Curvature on Foliations: The Enlargeability

Weiping Zhang

Abstract. We generalize the famous result of Gromov and Lawson on the nonexistence of metric of positive scalar curvature on enlargeable manifolds to the case of foliations, without using index theorems on noncompact manifolds.

Mathematics Subject Classification (2010). 57R30, 58J20.

Keywords. Foliation, positive scalar curvature, enlargeability, relative index theorem.

0. Introduction

It has been an important subject in differential geometry to study when a smooth manifold carries a Riemannian metric of positive scalar curvature (cf. [8, Chap. IV]). A famous result of Gromov and Lawson [6], [7] states that an enlargeable manifold (in the sense of [7, Definition 5.5]) does not carry a metric of positive scalar curvature. In particular, there is no metric of positive scalar curvature on any torus, which is a classical result of Schoen–Yau [10] and Gromov–Lawson [6]. A generalization to foliations of the Schoen–Yau and Gromov–Lawson result on torus has been given in [11, Corollary 0.5]. In this paper, we further extend the above result of Gromov–Lawson on enlargeable manifolds to the case of foliations.

Let F be an integrable subbundle of the tangent vector bundle TM of a closed smooth manifold M. Let g^F be a Euclidean metric on F, and $k^F \in C^\infty(M)$ be the associated leafwise scalar curvature (cf. [11, (0.1)]). For any covering manifold $\pi : \widetilde{M} \to M$, one has a lifted integrable subbundle with metric $(\widetilde{F}, g^{\widetilde{F}}) = (\pi^* F, \pi^* g^F)$.

Definition 0.1. One calls (M, F) an enlargeable foliation if for any $\varepsilon > 0$, there is a covering manifold $\widetilde{M} \to M$ and a smooth map $f : \widetilde{M} \to S^{\dim M}(1)$ (the standard unit sphere), which is constant near infinity and has non-zero degree, such that for any $X \in \Gamma(\widetilde{F})$, $|f_*(X)| \leq \varepsilon |X|$.

When $F = TM$ and M is spin, this is the original definition of the enlargeability of M due to Gromov and Lawson [6], [7].

The main result of this paper can be stated as follows.

Theorem 0.2. *Let (M, F) be an enlargeable foliation. Then*
(i) *if TM is spin, then there is no g^F such that $k^F > 0$ over M;*
(ii) *if F is spin, then there is no g^F such that $k^F > 0$ over M.*

When $F = TM$, one recovers the classical theorem of Gromov–Lawson [6], [7] mentioned at the begining. In a recent paper [2], Benameur and Heitsch proved Theorem 0.2(ii) under the condition that (M, F) has a Hausdorff homotopy groupoid.

As a direct consequence of Theorem 0.2(i), one obtains an alternate proof, without using the families index theorem, of [11, Corollary 0.5] mentioned above (for the special case where the integrable subbundle on torus is spin, this result is due to Connes, as was stated in [5, p. 192]).

If M is enlargeable and (M, F) carries a transverse Riemannian structure, then Theorem 0.2(i) is trivial, as in this case, if there is g^F with $k^F > 0$ over M, then one can construct g^{TM} with $k^{TM} > 0$ over M, which contradicts with the Gromov–Lawson theorem. Thus, the main difficulty for Theorem 0.2 is that there might be no transverse Riemannian structure on (M, F). This is similar to what happens in [4] and [11], where one adapts the Connes fibration constructed in [4] to overcome this kind of difficulty.

Recall that we have proved geometrically in [11] that if M is oriented and there exists g^F with $k^F > 0$ over M, then under the condition that either TM or F is spin, one has $\widehat{A}(M) = 0$. The case where F is spin is a famous result of Connes [4, Theorem 0.2].

Our proof of Theorem 0.2 combines the methods in [6], [7] and [11]. It is based on deforming (twisted) sub-Dirac operators on the Connes fibration over $\widetilde{M}$. A notable difference with respect to [7], where the relative index theorem on noncompact manifolds plays an essential role, is that we will work with compact manifolds even for the noncompactly enlargeable situation. It will be carried out in Section 1.

1. Proof of Theorem 0.2

In this section, we first prove in Section 1.1 the easier case where (M, F) is a compactly enlargeable foliation, i.e., the covering manifold $\widetilde{M}$ in Definition 0.1 is compact. Then in Section 1.2 we show how to extend the arguments in Section 1.1 to the case where $\widetilde{M}$ is noncompact.

1.1. The case of compactly enlargeable foliations

Let F be an integrable subbundle of the tangent bundle TM of an oriented closed manifold M.

Let g^F be a metric on F and k^F be the scalar curvature of g^F. Let (E, g^E) be a Hermitian vector bundle on M carrying a Hermitian connection ∇^E.

Let $R^E = (\nabla^E)^2$ be the curvature of ∇^E.

For any $\varepsilon > 0$, we say (E, g^E, ∇^E) verifies the leafwise ε-condition if for any X, $Y \in \Gamma(F)$, the following pointwise formula holds on M,

$$\left|R^E(X,Y)\right| \leq \varepsilon\,|X|\,|Y|. \tag{1.1}$$

The following result extends slightly [11, Theorem 0.1] and [4, Theorem 0.2].[1]

Theorem 1.1. *If $k^F > 0$ over M and either TM or F is spin, then there exists $\varepsilon_0 > 0$ such that if (E, g^E, ∇^E) verifies the leafwise ε_0-condition, then*

$$\langle \widehat{A}(TM)\mathrm{ch}(E), [M]\rangle = 0.$$

Proof. The proof of this theorem is an easy modification of the proof given in [11] for the case of $E = \mathbf{C}|_M$. We only give a brief description, by following the notations given in [11]. Let $\delta > 0$ be such that $k^F \geq \delta$ over M. Without loss of generality, we may well assume that $\dim M$, $\mathrm{rk}(F)$ are divisible by 4, and that TM, F and TM/F are oriented with compatible orientations.

We assume first that TM is spin.

Following [4, §5] (cf. [11, §2.1]), let $\pi : \mathcal{M} \to M$ be the Connes fibration over M such that for any $x \in M$, $\mathcal{M}_x = \pi^{-1}(x)$ is the space of Euclidean metrics on the linear space T_xM/F_x. Let $T^V\mathcal{M}$ denote the vertical tangent bundle of the fibration $\pi : \mathcal{M} \to M$. Then it carries a natural metric $g^{T^V\mathcal{M}}$.

By using the Bott connection on TM/F, which is leafwise flat, one lifts F to an integrable subbundle $\mathcal{F}$ of $T\mathcal{M}$. Then g^F lifts to a Euclidean metric $g^{\mathcal{F}} = \pi^* g^F$ on $\mathcal{F}$.

Let $\mathcal{F}_1^\perp \subseteq T\mathcal{M}$ be a subbundle, which is transversal to $\mathcal{F} \oplus T^V\mathcal{M}$, such that we have a splitting $T\mathcal{M} = (\mathcal{F} \oplus T^V\mathcal{M}) \oplus \mathcal{F}_1^\perp$. Then $\mathcal{F}_1^\perp$ can be identified with $T\mathcal{M}/(\mathcal{F} \oplus T^V\mathcal{M})$ and carries a canonically induced metric $g^{\mathcal{F}_1^\perp}$. We denote $\mathcal{F}_2^\perp = T^V\mathcal{M}$.

Let $\mathcal{E} = \pi^* E$ be the lift of E which carries the lifted Hermitian metric $g^{\mathcal{E}} = \pi^* g^E$ and the lifted Hermitian connection $\nabla^{\mathcal{E}} = \pi^*\nabla^E$. Let $R^{\mathcal{E}} = (\nabla^{\mathcal{E}})^2$ be the curvature of $\nabla^{\mathcal{E}}$.

For any $\beta,\ \varepsilon > 0$, following [11, (2.15)], let $g^{T\mathcal{M}}_{\beta,\varepsilon}$ be the metric on $T\mathcal{M}$ defined by the orthogonal splitting,

$$T\mathcal{M} = \mathcal{F} \oplus \mathcal{F}_1^\perp \oplus \mathcal{F}_2^\perp, \qquad g^{T\mathcal{M}}_{\beta,\varepsilon} = \beta^2 g^{\mathcal{F}} \oplus \frac{g^{\mathcal{F}_1^\perp}}{\varepsilon^2} \oplus g^{\mathcal{F}_2^\perp}. \tag{1.2}$$

Now we replace the sub-Dirac operator constructed in [11, (2.16)] by the obvious twisted (by $\mathcal{E}$) analogue

$$\begin{aligned} D^{\mathcal{E}}_{\mathcal{F}\oplus\mathcal{F}_1^\perp,\beta,\varepsilon} : \Gamma\left(S_{\beta,\varepsilon}(\mathcal{F}\oplus\mathcal{F}_1^\perp)\widehat{\otimes}\Lambda^*\left(\mathcal{F}_2^\perp\right)\otimes\mathcal{E}\right) \\ \longrightarrow \Gamma\left(S_{\beta,\varepsilon}(\mathcal{F}\oplus\mathcal{F}_1^\perp)\widehat{\otimes}\Lambda^*\left(\mathcal{F}_2^\perp\right)\otimes\mathcal{E}\right), \end{aligned} \tag{1.3}$$

where $S_{\beta,\varepsilon}(\cdot)$ is the notation for the spinor bundle determined by $g^{T\mathcal{M}}_{\beta,\varepsilon}$.

[1] The case where F is spin is due to Connes, cf. [5, p. 192].

The analogue of [11, (2.28)] now takes the form

$$\left(D^{\mathcal{E}}_{\mathcal{F}\oplus\mathcal{F}_1^\perp,\beta,\varepsilon}\right)^2 = -\Delta^{\mathcal{E},\beta,\varepsilon} + \frac{k^{\mathcal{F}}}{4\beta^2} + \frac{1}{2\beta^2}\sum_{i,\,j=1}^{\mathrm{rk}(F)} R^{\mathcal{E}}(f_i,f_j)c_{\beta,\varepsilon}\left(\beta^{-1}f_i\right)c_{\beta,\varepsilon}\left(\beta^{-1}f_j\right) + O_R\left(\frac{1}{\beta}+\frac{\varepsilon^2}{\beta^2}\right), \tag{1.4}$$

where $-\Delta^{\mathcal{E},\beta,\varepsilon} \geq 0$ is the corresponding Bochner Laplacian, $k^{\mathcal{F}} = \pi^* k^F \geq \delta$ and $f_1, \ldots, f_{\mathrm{rk}(F)}$ is an orthonormal basis of $(\mathcal{F}, g^{\mathcal{F}})$. Moreover, the analogue of [11, (2.34)] now takes the form

$$\mathrm{ind}\left(P^{\mathcal{E}}_{R,\beta,\varepsilon,+}\right) = \left\langle \widehat{A}(TM)\mathrm{ch}(E), [M]\right\rangle. \tag{1.5}$$

From (1.1), (1.4), (1.5) and proceed as in [11, §2.2 and §2.3], one gets Theorem 1.1 for the case where TM is spin easily. As in [11, §2.5], the same proof applies to give a geometric proof for the case where F is spin, with an obvious modification of the (twisted) sub-Dirac operators (cf. [11, (2.58)]). □

Now for the proof of Theorem 0.2, one follows [6], [8] and chooses a complex vector bundle E_0 over $S^{\dim M}(1)$ such that

$$\left\langle \mathrm{ch}\left(E_0\right), \left[S^{\dim M}(1)\right]\right\rangle \neq 0. \tag{1.6}$$

From Definition 0.1 and [8, (5.8) of Chap. IV], one sees that for any $\varepsilon > 0$, one can find a compact covering $\widetilde{M} \to M$ and a map $f : \widetilde{M} \to S^{\dim M}(1)$ of non-zero degree such that $E = f^*(E_0)$ verifies the leafwise ε-condition. Thus, if there is g^F with $k^F > 0$ over M, then by Theorem 1.1 and in view of either [11, Theorem 0.1] (in the case where M is spin) or [4, Theorem 0.2] (in the case where F is spin), one has

$$\begin{aligned} 0 &= \left\langle \widehat{A}\left(T\widetilde{M}\right)\mathrm{ch}\left(E\right), \left[\widetilde{M}\right]\right\rangle \\ &= (\mathrm{rk}(E_0))\widehat{A}\left(\widetilde{M}\right) + \left\langle \widehat{A}\left(T\widetilde{M}\right) f^*\left(\mathrm{ch}\left(E_0\right) - \mathrm{rk}\left(E_0\right)\right), \left[\widetilde{M}\right]\right\rangle \\ &= \deg(f)\left\langle \mathrm{ch}\left(E_0\right), S^{\dim M}(1)\right\rangle, \end{aligned} \tag{1.7}$$

where the last equality comes from the definition of $\deg(f)$, as $\mathrm{ch}(E_0) - \mathrm{rk}(E_0)$ is a top class on $S^{\dim M}(1)$. This contradicts with (1.6) and completes the proof of Theorem 0.2 for compact $\widetilde{M}$.

Remark 1.2. Since any torus T^n is compactly enlargeable (cf. [8, p. 303]), the proof above already applies to give an alternate proof of [11, Corollary 0.5] on the nonexistence of any foliation with metric of positive leafwise scalar curvature on T^n.

1.2. The case where $\widetilde{M}$ is noncompact

We will deal with the case where $F = TM$ in detail. We will work with compact manifolds, thus giving a new proof of the Gromov–Lawson theorem [7, Theorem 5.8] in the case where $\widetilde{M}$ is noncompact. With this "compact" approach it is easy to prove the foliation extension as in Section 1.1.

We assume that $\widetilde{M}$ is noncompact. To simplify the notation, from now on we simply denote $\widetilde{M}$ by M, or rather M_ε to emphasize the dependence on ε. The key point is that the geometric data on M now comes from isometric liftings of geometric data on a compact manifold.

Thus for any $\varepsilon > 0$, let $f_\varepsilon : M_\varepsilon \to S^{\dim M}(1)$ be as in Definition 0.1. Let $K_\varepsilon \subset M_\varepsilon$ be a compact subset of M_ε such that $f(M_\varepsilon \setminus K_\varepsilon) = x_0$, where x_0 is a (fixed) point on $S^{\dim M}(1)$.[2] Following [7], we take a compact hypersurface H_ε in $M_\varepsilon \setminus K_\varepsilon$. We denote by M_{H_ε} the compact manifold with boundary H_ε containing K_ε.

Let M'_{H_ε} be another copy of M_{H_ε}. We glue M_{H_ε} and M'_{H_ε} along H_ε to get the double, which we denote by $\widehat{M}_{H_\varepsilon}$. Let $g^{T\widehat{M}_{H_\varepsilon}}$ be a metric on $T\widehat{M}_{H_\varepsilon}$ such that $g^{T\widehat{M}_{H_\varepsilon}}|_{M_{H_\varepsilon}} = g^{TM}|_{M_{H_\varepsilon}}$. The existence of $g^{T\widehat{M}_{H_\varepsilon}}$ is clear.[3] Let $S(T\widehat{M}_{H_\varepsilon})$ denote the corresponding spinor bundle.

We extend $f_\varepsilon : M_{H_\varepsilon} \to S^{\dim M}(1)$ to $f_\varepsilon : \widehat{M}_{H_\varepsilon} \to S^{\dim M}(1)$ by setting $f_\varepsilon(M'_{H_\varepsilon}) = x_0$.

Let (E_0, g^{E_0}) be a Hermitian vector bundle on $S^{\dim M}(1)$ verifying (1.6) and carrying a Hermitian connection ∇^{E_0}. Let $(E_1 = \mathbf{C}^k|_{S^{\dim M}(1)}, g^{E_1}, \nabla^{E_1})$, with $k = \mathrm{rk}(E_0)$, be the canonical Hermitian trivial vector bundle on $S^{\dim M}(1)$. Let $v : \Gamma(E_0) \to \Gamma(E_1)$ be an endomorphism such that $v|_{x_0}$ is an isomorphism. Let $v^* : \Gamma(E_1) \to \Gamma(E_0)$ be the adjoint of v with respect to g^{E_0} and g^{E_1}. Set

$$V = v + v^*. \tag{1.8}$$

Then the self-adjoint endomorphism $V : \Gamma(E_0 \oplus E_1) \to \Gamma(E_0 \oplus E_1)$ is invertible near x_0.

Let $(\xi, g^\xi, \nabla^\xi) = (\xi_0 \oplus \xi_1, g^{\xi_0} \oplus g^{\xi_1}, \nabla^{\xi_0} \oplus \nabla^{\xi_1}) = (f_\varepsilon^* E_0 \oplus f_\varepsilon^* E_1, f_\varepsilon^* g^{E_0} \oplus f_\varepsilon^* g^{E_1}, f_\varepsilon^* \nabla^{E_0} \oplus f_\varepsilon^* \nabla^{E_1})$ be the $\mathbf{Z}_2$-graded Hermitian vector bundle with Hermitian connection over $\widehat{M}_{H_\varepsilon}$ (here for simplicity, we do not make explicit the subscript ε in ξ, ξ_0 and ξ_1). Let $R^\xi = (\nabla^\xi)^2$ be the curvature of ∇^ξ. Set $V_{f_\varepsilon} = f_\varepsilon^* V$. Then

$$\left[\nabla^\xi, V_{f_\varepsilon}\right] = 0 \tag{1.9}$$

on M'_{H_ε}.

Let $D^\xi : \Gamma(S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi) \to \Gamma(S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi)$ be the canonically defined (twisted) Dirac operator (cf. [8]). Let

$$D^\xi_\pm : \Gamma((S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi)_\pm) \to \Gamma((S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi)_\mp)$$

---

[2] Up to an isometry of $S^{\dim M}(1)$, one can always assume that x_0 is fixed and does not depend on ε.

[3] Here we need not assume that $g^{T\widehat{M}_{H_\varepsilon}}$ is of product structure near M_{H_ε}.

be the obvious restrictions, where

$$(S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi)_+ = S_+(T\widehat{M}_{H_\varepsilon}) \otimes \xi_0 \oplus S_-(T\widehat{M}_{H_\varepsilon}) \otimes \xi_1,$$

while

$$(S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi)_- = S_-(T\widehat{M}_{H_\varepsilon}) \otimes \xi_0 \oplus S_+(T\widehat{M}_{H_\varepsilon}) \otimes \xi_1.$$

By the Atiyah–Singer index theorem [1] (cf. [8]) and [7], one has

$$\begin{aligned} \operatorname{ind}\left(D_+^\xi\right) &= \left\langle \widehat{A}\left(T\widehat{M}_{H_\varepsilon}\right)(\operatorname{ch}(\xi_0) - \operatorname{ch}(\xi_1)), \left[\widehat{M}_{H_\varepsilon}\right]\right\rangle \\ &= (\deg(f_\varepsilon)) \left\langle \operatorname{ch}(E_0), \left[S^{\dim M}(1)\right]\right\rangle, \end{aligned} \tag{1.10}$$

where the last equality comes from the definition of $\deg(f_\varepsilon)$ (cf. [7]).

Let k^{TM} denote the scalar curvature of g^{TM}. We assume that there is $\delta > 0$ such that $k^{TM} \geq \delta$ over M.

For any $\varepsilon > 0$, let $D_\varepsilon^\xi : \Gamma(S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi) \to \Gamma(S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi)$ be the deformed operator defined by

$$D_\varepsilon^\xi = D^\xi + V_{f_\varepsilon}. \tag{1.11}$$

Proposition 1.3. *There is* $\varepsilon_0 > 0$ *such that for any* $0 < \varepsilon \leq \varepsilon_0$*, one has* $\ker(D_\varepsilon^\xi) = \{0\}$.

Proof. Recall that $x_0 \in S^{\dim M}(1)$ is fixed and $V|_{x_0}$ is invertible. Let $U_{x_0} \subset S^{\dim M}(1)$ be a (fixed) sufficiently small open neighborhood of x_0 such that the following inequality holds on U_{x_0},

$$V^2 \geq \delta_1. \tag{1.12}$$

Let $\psi : S^{\dim 1}(1) \to [0,1]$ be a smooth function such that $\psi = 1$ near x_0 and $\operatorname{Supp}(\psi) \subset U_{x_0}$. Then $\varphi_\varepsilon = 1 - f_\varepsilon^*\psi$ is a smooth function on M_ε (and thus on M_{H_ε}), which extends to a smooth function on $\widehat{M}_{H_\varepsilon}$ such that $\varphi_\varepsilon = 0$ on M'_{H_ε}.

Following [3, p. 115], let $\varphi_{\varepsilon,1}$, $\varphi_{\varepsilon,2} : \widehat{M}_{H_\varepsilon} \to [0,1]$ be defined by

$$\varphi_{\varepsilon,1} = \frac{\varphi_\varepsilon}{(\varphi_\varepsilon^2 + (1-\varphi_\varepsilon)^2)^{\frac{1}{2}}}, \quad \varphi_{\varepsilon,2} = \frac{1-\varphi_\varepsilon}{(\varphi_\varepsilon^2 + (1-\varphi_\varepsilon)^2)^{\frac{1}{2}}}. \tag{1.13}$$

Then $\varphi_{\varepsilon,1}^2 + \varphi_{\varepsilon,2}^2 = 1$. Thus, for any $s \in \Gamma(S(T\widehat{M}_{H_\varepsilon})\widehat{\otimes}\xi)$, one has

$$\left\|D_\varepsilon^\xi s\right\|^2 = \left\|\varphi_{\varepsilon,1} D_\varepsilon^\xi s\right\|^2 + \left\|\varphi_{\varepsilon,2} D_\varepsilon^\xi s\right\|^2, \tag{1.14}$$

from which one gets

$$\begin{aligned} \sqrt{2}\left\|D_\varepsilon^\xi s\right\| &\geq \left\|\varphi_{\varepsilon,1} D_\varepsilon^\xi s\right\| + \left\|\varphi_{\varepsilon,2} D_\varepsilon^\xi s\right\| \\ &\geq \left\|D_\varepsilon^\xi(\varphi_{\varepsilon,1} s)\right\| + \left\|D_\varepsilon^\xi(\varphi_{\varepsilon,2} s)\right\| - \|c(d\varphi_{\varepsilon,1}) s\| - \|c(d\varphi_{\varepsilon,2}) s\|, \end{aligned} \tag{1.15}$$

where we identify $d\varphi_{\varepsilon,i}$, $i = 1, 2$, with the gradient of $\varphi_{\varepsilon,i}$.

Let $e_1, \ldots, e_{\dim M}$ be an orthonormal basis of $g^{T\widehat{M}_{H_\varepsilon}}$. Then by (1.11), one has

$$\left(D_\varepsilon^\xi\right)^2 = \left(D^\xi\right)^2 + \sum_{i=1}^{\dim M} c\left(e_i\right)\left[\nabla_{e_i}^\xi, V_{f_\varepsilon}\right] + V_{f_\varepsilon}^2. \tag{1.16}$$

From (1.16), one has for $j = 1, 2$ that

$$\begin{aligned}\left\|D_\varepsilon^\xi\left(\varphi_{\varepsilon,j}s\right)\right\|^2 = \left\|D^\xi\left(\varphi_{\varepsilon,j}s\right)\right\|^2 + \sum_{i=1}^{\dim M}\left\langle c\left(e_i\right)\left[\nabla_{e_i}^\xi, V_{f_\varepsilon}\right]\varphi_{\varepsilon,j}s,\, \varphi_{\varepsilon,j}s\right\rangle \\ + \left\|\varphi_{\varepsilon,j}V_{f_\varepsilon}s\right\|^2.\end{aligned} \tag{1.17}$$

By the Lichnerowicz formula [9] (cf. [8]), one has on M_{H_ε} that

$$\left(D^\xi\right)^2 = -\Delta^\xi + \frac{k^{TM}}{4} + \frac{1}{2}\sum_{i,\, j=1}^{\dim M} c\left(e_i\right)c\left(e_j\right)R^\xi\left(e_i, e_j\right), \tag{1.18}$$

where Δ^ξ is the corresponding Bochner Laplacian and $k^{TM} \geq \delta$ by assumption.

By Definition 0.1 and proceeding as in [8, (5.8) of Chap. IV], one finds on M_{H_ε} that

$$\begin{aligned}&\frac{1}{2}\sum_{i,\, j=1}^{\dim M} c\left(e_i\right)c\left(e_j\right)R^\xi\left(e_i, e_j\right) + \sum_{i=1}^{\dim M} c\left(e_i\right)\left[\nabla_{e_i}^\xi, V_{f_\varepsilon}\right] \\ &= \frac{1}{2}\sum_{i,\, j=1}^{\dim M} c\left(e_i\right)c\left(e_j\right)f_\varepsilon^*\left(R^{E_0}\left(f_{\varepsilon *}e_i, f_{\varepsilon *}e_j\right)\right) \\ &\quad + \sum_{i=1}^{\dim M} c\left(e_i\right)f_\varepsilon^*\left(\left[\nabla_{f_{\varepsilon *}e_i}^{E_0\oplus E_1}, V\right]\right) = O\left(\varepsilon\right).\end{aligned} \tag{1.19}$$

On the other hand, for any $1 \leq i \leq \dim M$, one verifies that

$$e_i\left(\varphi_\varepsilon\right) = -e_i\left(f_\varepsilon^*\psi\right) = -f_\varepsilon^*\left(\left(f_{\varepsilon *}e_i\right)\left(\psi\right)\right) = O(\varepsilon). \tag{1.20}$$

From (1.13) and (1.20), one finds that for $i = 1, 2$,

$$\left|c\left(d\varphi_{\varepsilon,i}\right)\right| = O(\varepsilon). \tag{1.21}$$

From (1.9), (1.12), (1.13), (1.15), (1.17)–(1.19) and (1.21), one deduces that there exists $\delta_2 > 0$ such that when $\varepsilon > 0$ is sufficiently small, one has (compare with [11, p. 1062])

$$\left\|D_\varepsilon^\xi s\right\| \geq \delta_2\|s\|, \tag{1.22}$$

which completes the proof of Proposition 1.3. □

From Proposition 1.3, one finds $\operatorname{ind}(D_+^\xi) = 0$, which contradicts with (1.10) where the right-hand side is non-zero. Thus, there should be no g^{TM} with $k^{TM} > 0$ over M. This completes the proof of Theorem 0.2 for the case of $F = TM$ (which is the original Gromov–Lawson theorem [7, Theorem 5.8]), without using the relative index theorem on noncompact manifolds in [7].

Now to prove Theorem 0.2(i), one simply combines the method in Section 1.1 with the doubling and gluing tricks above. The details are easy to fill. Theorem 0.2(ii) follows by modifying the sub-Dirac operator as in [11, §2.5].

The proof of Theorem 0.2 is completed.

Acknowledgment

This work was partially supported by NNSFC. The author would like to thank the referee for helpful suggestions.

References

[1] M.F. Atiyah and I.M. Singer, *The index of elliptic operators I.* Ann. of Math. 87 (1968), 484–530.

[2] M.-T. Benameur and J.L. Heitsch, *Enlargeability, foliations, and positive scalar curvature.* Invent. Math. 215 (2019), 367–382.

[3] J.-M. Bismut and G. Lebeau, *Complex immersions and Quillen metrics.* Inst. Hautes Études Sci. Publ. Math. (1991), no. 74, ii+298 pp.

[4] A. Connes, *Cyclic cohomology and the transverse fundamental class of a foliation.* in *Geometric Methods in Operator Algebras.* H. Araki ed., pp. 52–144, Pitman Res. Notes in Math. Series, vol. 123, 1986.

[5] M. Gromov, *Positive curvature, macroscopic dimension, spectral gaps and higher signatures.* in *Functional Analysis on the Eve of the 21st Century*, Gindikin, Simon (ed.) et al. Progress in Math. 132 (1996), 1–213, Birkhäuser, Basel.

[6] M. Gromov and H.B. Lawson, *Spin and scalar curvature in the presence of a fundamental group I.* Ann. of Math. 111 (1980), 209–230.

[7] M. Gromov and H.B. Lawson, *Positive scalar curvature and the Dirac operator on complete Riemannian manifolds.* Publ. Math. I.H.E.S. 58 (1983), 295–408.

[8] H.B. Lawson and M.-L. Michelsohn, *Spin Geometry.* Princeton Univ. Press, Princeton, NJ, 1989.

[9] A. Lichnerowicz, *Spineurs harmoniques.* C. R. Acad. Sci. Paris, Série A, 257 (1963), 7–9.

[10] R. Schoen and S.T. Yau, *On the structure of manifolds with positive scalar curvature.* Manuscripta Math. 28 (1979), 159–183.

[11] W. Zhang, *Positive scalar curvature on foliations.* Ann. of Math. 185 (2017), 1035–1068.

Weiping Zhang
Chern Institute of Mathematics & LPMC
Nankai University
Tianjin 300071, P.R. China
e-mail: `weiping@nankai.edu.cn`

Progress in Mathematics, Vol. 333, 545–585

Kähler–Einstein Metrics on Toric Manifolds and G-manifolds

Xiaohua Zhu

Dedicated to Professor Gang Tian on the occasion of his 60th birthday

Abstract. This is an expository paper. In the first part, we discuss variant approaches in the study of Kähler–Einstein metrics on toric Fano manifolds. In the second part, we discuss the existence problem of Kähler–Einstein metrics on G-manifolds via Mabuchi's K-energy. Our method can be regarded as an extension in toric Fano manifolds. Some remaining questions/problems are also discussed.

Mathematics Subject Classification (2010). Primary: 53C25; Secondary: 53C55, 58J05, 35J60.

Keywords. Toric manifolds, G-manifolds, Kähler–Einstein metrics, Kähler–Ricci solitons.

0. Introduction

Since the celebrated work of Yau [81] on the existence of Kähler–Einstein metrics on Kähler manifolds with negative or vanishing first Chern class, and that of Aubin [1] on compact complex manifolds with negative first Chern class, significant progress has been made in the study of Kähler–Einstein metrics on Fano manifolds, namely Kähler manifolds with positive first Chern class. The famous Yau–Tian–Donaldson's conjecture has been recently proved separately by Tian [70], and Chen, Donaldson and Sun [19]. This conjecture asserts that the existence of Kähler–Einstein metrics on Fano manifolds is equivalent to the K-stability. The notion of K-stability was first introduced by Tian by using special degenerations [68] and then reformulated by Donaldson in algebraic geometry via test-configurations [26]. For both special degenerations and test-configurations, one

The author was partially supported by NSFC Grants 11331001 and 11771019.

has to study infinitely possible degenerations of the manifold. A natural question is how to verify the K-stability in finite steps. In this survey, we hope to give a picture for this question through examples of toric Fano manifolds as well as G-manifolds from various views in differential geometry, algebraic geometry, and differential equation, etc.

The history in the study of Kähler–Einstein metrics on toric manifolds can go back to more than 30 years ago. In 1987, Mabuchi began to classify toric Fano 3-folds with vanishing Futaki invariant by using the classification of toric Fano 3-folds in [9, 49]. However, beside the Fubini–Study metric on $\mathbb{C}P^n$, the first example of Kähler–Einstein metrics on toric manifolds was constructed on the blow-up of $\mathbb{C}P^2$ at three points, $\mathbb{C}P^2\#3\overline{\mathbb{C}P^2}$, by using PDE from the work of Tian and Yau [71], and Siu [58]. Latterly, by using Tian's α-invariant criterion [66], Real extended the example $\mathbb{C}P^2\#3\overline{\mathbb{C}P^2}$ to $\mathbb{C}P^n\#(n+1)\overline{\mathbb{C}P^n}$ for any $n>2$ [56], where $\mathbb{C}P^n\#(n+1)\overline{\mathbb{C}P^n}$ denotes the manifold obtained from the complex projective space $\mathbb{C}P^n$ by blowing-up at $(n+1)$ generic points.

After the Mabuchi's work, Nakagawa [51, 52] classified toric Fano Kähler–Einstein manifolds of dimension four by using results of Tian–Yau and Real based on the classification of toric Fano 4-folds of Batyrev [10]. In 1999, Batyrev and Selivanova proved that any symmetric toric Fano manifolds admits a Kähler–Einstein metric [12]. A toric manifold M is called symmetric if its associated polytope is symmetric with respect to the barycenter of polytope. In particular, the Futaki invariant of M vanishes. Batyrev and Selivanova's method is also through the computation of Tian's α-invariant.

To construct a Kähler–Einstein metric on a Fano manifold, people usually use the continuity method to solve the following family of complex Monge–Ampère equations (cf. [58, 66, 69, 81]),

$$\det(g_{i\bar{j}}+\varphi_{i\bar{j}})=\det(g_{i\bar{j}})\exp\{h-t\varphi\},\ (g_{i\bar{j}}+\varphi_{i\bar{j}})>0, \tag{0.1}$$

where h is a Ricci potential of the background Kähler metric g in $2\pi c_1(M)$, and t is a parameter from 0 to 1. One can check that (0.1) is equivalent to a Kähler–Einstein metric equation when $t=1$. It is known that the existence problem of Kähler–Einstein metrics is reduced to a prior C^0-estimate for solutions φ_t of (0.1). Tian used this approach to solve the problem completely on Fano surfaces through computing his α-invariant and establishing the partial C^0-estimate on Kähler–Einstein Fano surfaces in 1990 [67].

In the case of toric Fano manifolds, (0.1) can be further reduced to the following real Monge–Ampère equations in a global Euclidean space,

$$\det(u_{ij})=\exp\{-tu-(1-t)\psi_0\}\quad \text{in } \mathbb{R}^n, \tag{0.2}$$

where $u = u_t = \psi_0 + \varphi_t$ and ψ_0 is a Kähler potential of torus-invariant background metric g_0. By estimating the minimum of the convex functions $w_t = tu_t + (1-t)\psi_0$ and the distance of the minimal points of w_t away from the original o, Wang and the author got a C^0-estimate for solutions φ_t in [79] (also see Section 2 below). As a result, this gives a complete resolution for the existence of Kähler–Einstein metrics. That is

Theorem 0.1. *There exists a unique Kähler–Ricci soliton up to holomorphic automorphisms on any toric Fano manifold. Moreover, the Kähler–Ricci soliton is Einstein if and only if the Futaki-invariant vanishes.*

Recall that a Kähler–Ricci soliton on a complex manifold M is a pair (X, ω), where X is a holomorphic vector field on M and g is a Kähler metric on M, such that

$$\mathrm{Ric}(\omega) - \omega = L_X(\omega), \tag{0.3}$$

where L_X is the Lie derivative along X. If $X = 0$, the Kähler–Ricci soliton becomes a Kähler–Einstein metric. The uniqueness theorem in [72, 73] states that a Kähler–Ricci soliton on a compact complex manifold, if it exists, must be unique modulo $\mathrm{Aut}(M)$.[1] Furthermore, X lies in the center of Lie algebra of the reductive part of $\mathrm{Aut}^0(M)$, which is the connected component of holomorphisms group $\mathrm{Aut}(M)$ containing the identity.

Theorem 0.1 has been generalized in various directions via the method in [79] in recent ten years. For examples, Podesta and Spiro proved the existence of Kähler–Ricci solitons on the torus bundle over a homogeneous space [55], Shi and the author proved the existence of Kähler–Ricci solitons on toric Fano orbifolds [62], Futaki, Ono and Wang proved the existence of transverse Sasaki–Ricci solitons on toric Sasaki manifolds [36], and very recently Deltroix proved the existence of Kähler–Einstein metrics on G-manifolds [33] (also see Section 5 below), etc. There are other applications of C^0-estimate in [79] found, such as in Ricci flow [88], in the study of singularities arising in (0.1) on toric Fano manifolds [42, 63]. We will discuss them in Subsection 3.1, 3.2, respectively.

In [84], Zhou and the author found another method to do C^0-estimate for solutions φ_t by proving the properness of Mabuchi's K-energy. The idea is based on a Tian's result of analytic criterion for the existence of Kähler–Einstein metrics via K-energy $K(\cdot)$ [69]. In case of toric Fano manifolds, Donaldson observed that $K(\cdot)$ is equivalent to a reduced K-energy $\mu(\cdot)$ via Legendre dual functions if one restricts the Kähler potentials in the space of torus invariant functions [26]. In [84], we actually found a way to verify the properness of $\mu(\cdot)$. More precisely, we give a criterion for the properness of $\mu(\cdot)$ in terms of polytope associated to the Kähler

[1] In the case of Kähler–Einstein metrics, this uniqueness theorem is due to Bando–Mabuchi [8].

class. The advantage is that the C^0-estimate problem returns to study the structure of polytope. Thus, this method works for any Kähler class on a toric manifold. As an application, we can prove the existence of weak minimizers of K-energy on toric manifolds [86].[2] Recently, we extended this method to G-manifolds and proved the existence of weak minimizers of K-energy [45]. Analogous criterion was also established for the modified K-energy associated to Kähler–Ricci solitons on toric manifolds [80], Mabuchi's extremal metrics on G-manifolds [47], and transverse Sasaki–Ricci solitons on G-Sasaki manifolds [46], respectively.

This paper is organized as follows. In Section 1, we recall some basic notations for toric manifolds. In Section 2, we discuss the C^0-estimate in [79]. In Section 3, we discuss another proof of Theorem 0.1 via Ricci flow in [88] and a Li's result for singular solutions of (0.1) [42]. These results can be regarded as applications of C^0-estimate in [79]. In Section 4, we discuss the method in [84] to prove Theorem 0.1 via K-energy. In Section 5, we discuss a recent result on the existence of Kähler–Einstein metrics on G-manifolds in [45]. In the last section as an appendix, we give some examples of Fano G-manifolds with a maximal torus action of rank 2 and discuss whether there are Kähler–Einstein metrics or Kähler–Ricci solitons on them.

Note. We are not able to discuss other interesting works for the construction of canonical metrics on toric manifolds, such as singular Kähler–Ricci solitons on a toric Q-Fano variety [13, 64], conical Kähler–Einstein metrics [25, 80], and Calabi's extremal metrics [21, 22, 29, 30] on a toric Fano manifold, etc. We refer the reader to those papers.

1. Preliminary on toric manifolds

An n-dimensional toric manifold M is a compactification of n-dimensional torus $T^{\mathbb{C}} = (\mathbb{C}^*)^n = (S^1)^n \times \mathbb{R}^n$ (cf. [53]). Then $T^{\mathbb{C}}$ acts naturally on M. Denote $K = (S^1)^n$. Thus a K-invariant Kähler metric g corresponds to a convex function ψ_0 in $\mathbb{R}^n$ such that its Kähler form ω_g is an extension of $\sqrt{-1}\partial\bar{\partial}\psi_0 = \sqrt{-1}\partial_z\partial_{\bar{z}}\psi_0$ on $(\mathbb{C}^*)^n$. Here $z = (z_1, \ldots, z_n)$ ($z_i = \log w_i = x_i + \sqrt{-1}\theta_i$) are the affine logarithmic coordinates on $T^{\mathbb{C}}$ and $(w_1, \ldots, w_n) \in (\mathbb{C}^*)^n$. Let $\mathcal{H}_K(\omega_g)$ be the set of such K-invariant Kähler potentials in $[\omega_g]$. Then it is easy to see that

$$\mathcal{H}_K(\omega_g) = \{\phi \in C^\infty(\mathbb{R}^n) |\ |\phi| < \infty \text{ and } \psi_0 + \phi \text{ is uniformly convex}\}.$$

[2] A recent work of Chen–Cheng shows that these minimizers are associated to the existence of canonical Kähler metrics with constant scalar curvature [17].

Let P be the image of the gradient of ψ_0 for $x \in \mathbb{R}^n$, namely,

$$\nabla\psi = \left(\frac{\partial\psi}{\partial x_1}, \ldots\ldots, \frac{\partial\psi}{\partial x_n}\right) = (y_1, \ldots\ldots, y_n).$$

Then P is a ploytope in $\mathbb{R}^n$, which is an intersection of d hyperplanes in $\mathbb{R}^n$ as follows,

$$P = \{x \in \mathbb{R}^n |\ l_\alpha(x) > \lambda_\alpha,\ \alpha = 1, \ldots, d\}, \tag{1.1}$$

where $l_\alpha(x) = \sum l_\alpha^i x_i$ with $l_\alpha = (l_\alpha^1, \ldots, l_\alpha^n) \in \mathbb{Z}^n$. Moreover, P satisfies the Delzant condition (cf. [2, 32, 37]). In particular, $\omega_g \in 2\pi c_1(M)$ if and only if $\lambda_\alpha = 2$ for all α after a translation to P. P is also called the moment polytope. It is easy to see that P is independent of the choice of $\psi = \psi_0 + \phi$ with $\phi \in \mathcal{H}_K(\omega_g)$, i.e., the choice of ω_ϕ as a moment polytope.

As an obstruction to the existence of Kähler–Einstein metrics, Futaki introduced the following holomorphic invariant (Futaki invariant) in 1983 [35],

$$F(X) = \int_M X(h)\omega_g^n,\ \forall\ X \in \eta(M), \tag{1.2}$$

where h is a Ricci potential of ω_g, and $\eta(M)$ is the Lie algebra of holomorphisms transformation group $\mathrm{Aut}(M)$, which consists of holomorphic vector fields on a compact Kähler manifold (M, g). In case of toric manifolds, we have (cf. [49] and [79]),

Lemma 1.1. *Let M be an n-dimensional toric Fano manifold and P the associated moment polytope as in* (1.1) *with all $\lambda_\alpha = 2$. Then the Futaki invariant vanishes if and only if the barycenter of P is the original. Namely,*

$$\int_P y_i dy = 0,\ i = 1, \ldots, n. \tag{1.3}$$

2. A priori C^0-estimate

In this section, we discuss two main technical lemmas in the proof of Theorem 0.1 by solving (0.2). For simplicity, we assume that the Futaki invariant vanishes. In general case, we shall modify (0.2) to an equation of Kähler–Ricci soliton type [72, 79].

By a Harnack inequality in [66] and higher-order estimates in [66, 81], it suffices to prove that for any solution φ_t of (0.1), it holds

$$\sup_M \varphi_t \leq C. \tag{2.1}$$

The proof includes two steps. First we show

Lemma 2.1. *We have*

$$m_t =: \inf_{\mathbb{R}^n} w_t(x) \leq C \tag{2.2}$$

for some $C > 0$ independent of $t \in [\epsilon_0, 1]$.

We will use the following well-known result for convex domains [50].

Lemma 2.2. *Let Ω be a bounded convex domain in $\mathbb{R}^n$. Then there is a unique ellipsoid E, called the minimum ellipsoid of Ω, which attains minimal volume among all ellipsoids containing Ω, such that*

$$\tfrac{1}{n}E \subset \Omega \subset E,$$

where αE denotes the α-dilation of E with concentrated center.

Let T be a linear transformation with $|T| = 1$, which leaves the center x_0 of E invariant, namely $T(x) = A(x - x_0) + x_0$ for some matrix A, such that $T(E)$ is a ball B_R. Then we have $B_{R/n} \subset T(\Omega) \subset B_R$ for two balls with concentrated center.

Proof of Lemma 2.1. For any nonnegative integer k, we denote

$$A_k = \{x \in \mathbb{R}^n : \ m_t + k \leq w(x) \leq m_t + k + 1\}.$$

Then for any $k \geq 0$, $\bigcup_{i=0}^k A_i = \{w < m_t + k + 1\}$ is convex. Observe that $\nabla w(\mathbb{R}^n) = P$ and by (1.1), the origin is contained in P. Hence for any $k \geq 0$, A_k is a bounded set and the minimum m_t is attained at some point in A_0.

By equation (0.2), we have

$$\det(w_{ij}) \geq t^n \det(u_{ij}) \geq t^n e^{-w},$$

where $d = \sup\{c_l y_l : \ y \in P\}$. Recall that $t \geq \epsilon_0$, we obtain

$$\det(w_{ij}) \geq C_0 e^{-m_t} \quad \text{in} \ \ A_0,$$

where $C_0 = \epsilon_0^n$. By Lemma 2.2, there exists a linear transformation $y = T(x)$ with $|T| = 1$, which leaves the center of the minimum ellipsoid of A_0 invariant, such that $B_{R/n} \subset T(A_0) \subset B_R$. The above equation is unchanged under the linear transformation. We claim

$$R \leq \sqrt{2}\, n C_0^{-1/2n}\, e^{m_t/2n}. \tag{2.3}$$

Indeed, let

$$v(y) = \frac{1}{2} C_0^{1/n} e^{-m_t/n} \left[|y - y_t|^2 - \left(\frac{R}{n}\right)^2 \right] + m_t + 1,$$

where y_t is the center of the minimum ellipsoid of A_0. Then

$$\det(v_{ij}) = C_0 e^{-m_t} \quad \text{in } T(A_0),$$

and $v \geq w$ on $\partial T(A_0)$. Hence by the comparison principle we have $v \geq w$ in $T(A_0)$. In particular we have

$$\begin{aligned} m_t &\leq w(y_t) \leq v(y_t) \\ &= -\frac{1}{2} C_0^{1/n} e^{-m_t/n} \left(\frac{R}{n}\right)^2 + m_t + 1. \end{aligned}$$

Hence (2.3) follows.

By the convexity of w, we have

$$T(A_k) \subset B_{2(k+1)R}.$$

We obtain

$$\begin{aligned} \int_{\mathbb{R}^n} e^{-w} &= \sum_k \int_{T(A_k)} e^{-w} \\ &\leq \sum_k e^{-m_t-k} |T(A_k)| \\ &\leq \omega_n \sum e^{-m_t-k} |2(k+1)R|^n \\ &= \omega_n \frac{(2R)^n}{e^{m_t}} \sum \frac{(k+1)^n}{e^k} \\ &\leq Ce^{-m_t/2}, \end{aligned}$$

where ω_n is the area of the sphere $\mathbb{S}^{n-1}$. We note that the above integration is invariant under any linear transformation T with $|T| = 1$. Returning to the coordinates x, by equation (0.2), we have

$$\begin{aligned} e^{-m_t/2} &\geq \frac{1}{C} \int_{\mathbb{R}^n} e^{-w} dx \\ &= \frac{1}{C} \int_M \omega_g^n = C_1, \end{aligned}$$

where we have used the transformation $y = \nabla u(x)$. Hence $m_t \leq C$. $\square$

Next by using the vanishing Futaki invariant we prove

Lemma 2.3. *Let $x^t = (x_1^t, \ldots, x_n^t) \in \mathbb{R}^n$ be the minimal point of $w = w_t$. Then*

$$|x^t| \leq C$$

for some uniform constant C.

Proof. By equation (0.2),

$$\int_{\mathbb{R}^n} e^{-w} dx = \int_P dy = \beta$$

for some constant β. Recall that $|\nabla w| \le d_0 := \sup\{|x| :\ x \in P\}$. Hence by (2.3) there exists $R > 0$ such that $\inf_{\partial B_R(x^t)} w \ge m_t + 1$. By convexity we have

$$|\nabla w(x)| \ge 1/R \quad \text{in} \ \ \mathbb{R}^n \setminus B_R(x^t).$$

Hence for any $\epsilon > 0$ small, there exists R_ϵ sufficiently large such that

$$\int_{\mathbb{R}^n \setminus B_{R_\epsilon}(x^t)} e^{-w} dx \le C \int_{\mathbb{R}^n \setminus B_{R_\epsilon}(x^t)} e^{-|x-x^t|/R} \le \epsilon,$$

where both R and R_ϵ are independent of t.

Observe that ψ_0 is a convex function defined on $\mathbb{R}^n$ satisfying $\nabla\psi_0(\mathbb{R}^n) = P$, and by the fact of the origin $0 \in P$. Hence for any small $\epsilon > 0$, there exists a large constant $C > 0$ such that if $|x^t| > C$,

$$\frac{\partial u^0}{\partial \xi} > \frac{1}{2} a_0 \quad \text{in} \ \ B_{R_\epsilon}(x^t),$$

where $\xi = \frac{x^t}{|x^t|}$ and $a_0 = \inf\{|x| :\ x \in \partial P\}$. To see the above inequality it suffices to consider the restriction of u^0 on the ray $\overrightarrow{ox^t}$. Hence

$$\int_{B_{R_\epsilon}(x^t)} \frac{\partial u^0}{\partial \xi} e^{-w} dx \ge \frac{1}{4} a_0 \beta,$$

and

$$\left| \int_{\mathbb{R}^n \setminus B_{R_\epsilon}(x^t)} \frac{\partial u^0}{\partial \xi} e^{-w} dx \right| \le d_0 \left| \int_{\mathbb{R}^n \setminus B_{R_\epsilon}(x^t)} e^{-w} dx \right| \le \epsilon d_0.$$

If $\epsilon > 0$ is sufficiently small, we obtain

$$\int_{\mathbb{R}^n} \frac{\partial u^0}{\partial \xi} e^{-w} dx > 0\,.$$

On the other hand, by the vanishing of the Futaki invariant (1.3) in Lemma 1.1 and equation (0.2),

$$\begin{aligned} 0 &= \int_P y_i dy \\ &= \int_{\mathbb{R}^n} \frac{\partial u}{\partial x_i} e^{-w} dx \\ &= -\frac{1-t}{t} \int_{\mathbb{R}^n} \frac{\partial u^0}{\partial x_i} e^{-w} dx. \end{aligned}$$

We obtain

$$\int_{\mathbb{R}^n} \frac{\partial u^0}{\partial \xi} e^{-w} dx = 0$$

for any unit vector $\xi \in \mathbb{R}^n$. We reach a contradiction if $|x^t| >> 1$. This completes the proof. $\square$

Proof of Theorem 0.1. We need to get the estimate (2.1). In fact, by $t\varphi_t(x_t) = m_t - \psi_0(x_t)$, we see from Lemma 2.1 and Lemma 2.3 that

$$\varphi_t(x_t) \le A_0, \tag{2.4}$$

for some uniform constant A_0. Note that $|\nabla\varphi|(x) \le 2|\nabla\psi(x)| \le 2\mathrm{diam}(P)$. Then

$$\varphi_t(x) \le A_0 + 2\mathrm{diam}(P),\ \forall\, x \in B_1(x_t) \subset \mathbb{R}^n. \tag{2.5}$$

On the other hand, by the mean value inequality,

$$\sup_M \varphi_t \le \frac{1}{V}\int_M \varphi_t\omega_g^n + C,$$

where C is a uniform constant. One can easily show that there exists a point $x \in B_1(x_t)$ such that

$$\sup_M \varphi_t \le \varphi_t(x) + C'.$$

Thus combining it with (2.5), we get (2.1). □

3. Generalization of Lemma 2.1 and its applications

In this section, we consider a more general equation of (0.2),

$$\det(u_{ij}) = \exp\{-tu - (1-t)\psi_0 + f(x,t)\} \quad \text{in } \mathbb{R}^n, \tag{3.1}$$

where $f(x,t)$ is a uniformly bounded smooth function on $\mathbb{R}^n$. Following the argument in the proof of Lemma 2.1, we actually prove

Lemma 3.1. *Let ϵ_0 be a small positive number and $u = u_t$ a convex solution of* (3.1) ($t \in [\epsilon_0, 1]$). *Suppose that*

$$\int_{\mathbb{R}^n} \det(u_{ij})dx \ge C_0.$$

Then there is uniform constant independent of t such that

$$m_t =: \inf_{\mathbb{R}^n} w_t(x) \le C. \tag{3.2}$$

We will give two applications of Lemma 3.1 in the following.

3.1. Deformation of Ricci flow

Ricci flow was introduced by Hamilton in his study of three-dimensional sphere geometry in 1982 [39]. On a Fano manifold (M,g) with $\omega_g \in 2\pi c_1(M)$, we usually study the following normalized Kähler–Ricci flow,

$$\frac{\partial\omega(\cdot,s)}{\partial s} = -\mathrm{Ric}(\omega(\cdot,s)) + \omega(\cdot,s),\ \omega(\cdot,0) = \omega_g. \tag{3.3}$$

(3.3) gives an approach to construct Kähler–Einstein metrics. It can be rewritten as a parabolic complex Monge–Ampère equation for solutions in a space $\mathcal{H}(\omega_g)$ of Kähler potentials,

$$\begin{aligned} &\frac{\partial \varphi}{\partial s} = \log \frac{\omega_\varphi^n}{\omega_g^n} + \varphi + h, \\ &\omega_\varphi = \omega_g + \sqrt{-1}\partial\bar{\partial}\varphi > 0, \varphi|_{s=0} = 0. \end{aligned} \tag{3.4}$$

In case of toric manifolds, the author proved the following smooth convergence of (3.3) [88].

Theorem 3.2. *Suppose that ω_g is K-invariant on a toric Fano manifold M. Then there is a family of holomorphisms $\sigma_t \in \mathrm{Aut}^0(M)$ such that $\sigma_t^*\omega_{\varphi_t}$ converge to a Kähler–Ricci soliton exponentially.*

Theorem 3.2 gives a parabolic proof of Theorem 0.1. To the best of authors' knowledge, Theorem 3.2 is the first result about the global convergence of Kähler–Ricci flow on a Fano manifold without any assumptions on the curvature. Under the assumption of the existence of Kähler–Einstein metrics or a Kähler–Ricci soliton, a general convergence result of Kähler–Ricci flow was proved by Tian and the author in [74–76].[3]

As the same as in the equation (0.1) arising in the continuity method, we need to get a C^0-estimate for $\varphi(x,t)$ in (3.3) [14]. By a Harnack inequality in [74, 88], it suffices to prove

$$\sup_{M\times[0,\infty)} \varphi(x,s) \le C. \tag{3.5}$$

But for the parabolic equation, we can modify (3.4) by a family of holomorphisms $\sigma = \sigma_s \in \mathrm{Aut}^0(M)$ to

$$\begin{aligned} &\sigma^*\frac{\partial \varphi}{\partial s} = \log \frac{\omega_{\tilde\varphi}^n}{\omega_g^n} + \tilde\varphi - h, \\ &\omega_{\tilde\varphi} = \omega_g + \sqrt{-1}\partial\bar{\partial}\tilde\varphi > 0,\ \tilde\varphi|_{s=0} = 0, \end{aligned} \tag{3.6}$$

where $\tilde\varphi(s,\cdot)$ is an induced Kähler potential of $\omega_{\tilde\varphi} = \sigma^*\omega_\varphi$ by σ_s. In toric case, we can take $\sigma_s \in T^{\mathbb{C}}$. Then the corresponding Kähler potential $u(x,s)$ of ω_φ becomes $\tilde u(x,s) = u(x + x_s, s)$ in the affine coordinates $x + \sqrt{-1}\theta = (x_1 + \sqrt{-1}\theta_1, \ldots, x_n + \sqrt{-1}\theta_n) \in \mathbb{C}^n$, where we choose points x_s such that $u(x_s,s) = \inf_M u(x,s)$. Thus $\tilde u(x,s) = \psi_0 + \tilde\varphi(x,s)$ satisfies a parabolic real Monge–Ampère equation,

$$\begin{aligned} &\sigma^*\frac{\partial u}{\partial s} = \log\det(\tilde u_{ij}) + \tilde u, \quad \text{in } \mathbb{R}^n, \\ &\tilde u(0,\cdot) = u_0. \end{aligned} \tag{3.7}$$

[3] Perelman announced this result in case of Kähler–Einstein manifolds in his first paper to solve the Poincaré conjecture [54].

We need the following Perelman result (cf. [60]).

Lemma 3.3. *There exists constant c_s for each $s > 0$ in* (3.4) *such that*

$$|\frac{\partial \varphi}{\partial s} + c_s| \le C,$$

where C is a uniform constant.

Remark 3.4. It was shown that c_s is uniformly bounded by adding a constant to h in (3.4) if the K-energy or modified K-energy is bounded below along the flow [23, 73]. Namely, $|\frac{\partial \varphi}{\partial s}| \le C$.

Proof of Theorem 3.2. Applying Lemma 3.3 to (3.7), $\bar{u}(x,s) = \tilde{u}(x,s) + c_s$ satisfies

$$\det(\bar{u}_{ij}) = e^{-\bar{u}+f}, \quad \text{in } \mathbb{R}^n,$$

where $f(x,s) = \sigma^* \frac{\partial u}{\partial s} + c_s$ is a uniformly bounded function by Lemma 3.3. Then by Lemma 3.1,

$$\bar{\varphi}(o,s) \le A$$

for some uniform constant, where $\bar{\varphi}(x,s) = \bar{u}(x,s) - u_0$. Thus as in the proof of Theorem 0.1, we see that

$$\sup_M \bar{\varphi}(x,s) \le C. \tag{3.8}$$

By the monotonicity of modified K-energy, we get (cf. [89]),

$$I(\bar{\varphi}(x,s)) \le C.$$

Therefore, by the Harnack inequality [74, 88], we obtain

$$|\bar{\varphi}(x,s)| \le C. \tag{3.9}$$

By (3.8), we can show that the modified K-energy defined in [73] is bounded below along the flow (3.4). Then by (3.9) and Remark 3.4, it follows that

$$|\tilde{\varphi}(x,s)| \le C. \tag{3.10}$$

Moreover, by equation (3.4), $\tilde{\varphi}$ satisfies

$$\frac{\partial \tilde{\varphi}}{\partial s} = \log \frac{\omega_{\tilde{\varphi}}^n}{\omega_g^n} + \tilde{\varphi} + h + X_s(\tilde{\varphi}), \tag{3.11}$$

where $X_s = \frac{d\sigma_s}{ds}$ is a family of holomorphic vector fields on M. On the other hand, by using a trick in [23] with the help of (3.10) (also see [88]), one can modify the holomorphisms σ_s such that $|X_s|_g$ is uniformly bounded. Thus by a result in [87] (also see [84]), we get

$$|X_s(\tilde{\varphi})| \le C. \tag{3.12}$$

By the higher-order estimate together with (3.10) and (3.12) as done in [72], we have

$$\|\tilde{\varphi}\|_{C^k} \leq C_k.$$

By Perelman's entropy [54], there is a subsequence of $\omega_{\tilde{\phi}}$ which converge to a Kähler–Ricci soliton (M, ω_{KS}) with respect to some holomorphic vector field X on M. Let $\sigma_s(X)^*\omega_{\varphi_s}$ be the family of induced metrics by $\sigma_s(X)$ generated by X. Then, by the uniqueness of Kähler–Ricci solitons [72], $\sigma_s(X)^*\omega_{\varphi_s}$ converges to ω_{KS} exponentially [89]. □

3.2. Singular solutions arising in the continuity method

If the Futaki invariant does not vanish, there is no solution of (0.1) when $t = 1$. Then there is a $\bar{T} \leq 1$ such that

$$\bar{T} = \{\bar{t}| \text{ there is a solution } \varphi \text{ of (0.1) for any } t \leq \bar{t}\}. \tag{3.13}$$

Thus the metric ω_{φ_t} will blow-up as $t \to \bar{T}$. It is interesting to analyze the behavior of ω_{φ_t} such as the partial regularity of Gromov–Hausdorff limit, the current limit in terms of weak pluri-subharmonic functions, the singular set structure of these limits, etc. More recently, Székeyihidi proved the partial regularity of Gromov–Hausdorff limit by showing that ω_{φ_t} has locally bounded Ricci curvature [59]. For a simple example of $CP^n\#\overline{CP^n}$, Shi and the author showed that the rotational metrics ω_{φ_t} arising in (0.1) must locally smoothly converge to a conical Kähler metric on $CP^n\#\overline{CP^n}$ when $t \to \bar{T}$ [62]. But, the limit is not a soliton metric, which is different to the situation of Kähler–Ricci flow as in Section 3.1 above.

We call a Kähler metric ω on $B_1(o) \setminus \{z_1 = 0\} \subset \mathbb{C}^n$ with conical singularities along the hyperplane $\{z_1 = 0\} \subset \mathbb{C}^n$ if there is a smooth function $u(z, \bar{z})$ on $B_1(o) \setminus \{z_1 = 0\} \subset \mathbb{C}^n$ such that $\omega = \omega_0 + \sqrt{-1}\partial\bar{\partial}u$ and $\tilde{u}(w_1, z', \bar{w}, \bar{z}') = u(w^\beta, z', \overline{w^\beta}, \bar{z}')$ can be locally extended to a $C^{2,\alpha}$ smooth function on the variables $w_1, \bar{w}_1, z', \bar{z}'$ near $w_1 = 0$ (cf. [28, 62]). Here $z' = (z_2, \ldots, z_n)$, $\omega_0 = \sqrt{-1}\partial\bar{\partial}|z|^2$ is the standard flat metric on $\mathbb{C}^n$, and $2\pi\beta \in (0, 2\pi]$ is called the conical angle.

In this subsection, we discuss a Li result of local convergence of ω_{φ_t} above on a toric Fano manifold [42]. Let g be the Fubini–Study metric of M induced by the Kodaira embedding of Fano toric manifold. Namely,

$$\omega_g = \sqrt{-1}\partial\bar{\partial}\log\Big(\sum_{i=1}^N |s_i|^2\Big),$$

where s_i is a basis of holomorphic sections of K_M^{-1}, each of which can be given by a defining section of combination of some infinity divisors D_i of M. Then Li proved the following theorem,

Theorem 3.5. *Let φ_t be a solution of* (0.1) *with a $(S^1)^n$-invariant background metric g. Let T be the number defined in* (3.13). *Then there exist a sequence of*

$t_i \to \bar{T}$ and $\sigma_{t_i} \in T^{\mathbb{C}}$ such that the induced Kähler potential $\tilde{\varphi}_{t_i}$ of $\omega_{\varphi_{t_i}}$ by σ_{t_i} converges locally C^∞ to a current solution φ_∞, which satisfies

$$\frac{\omega_\varphi^n}{\omega_g^n} = e^{-\bar{T}\varphi - h + F_\infty + c_0}, \tag{3.14}$$

where c_0 is a constant, and

$$F_\infty = (1-\bar{T}) \log \frac{\sum |s_i|^2}{\sum_{i'} |b_{i'} s_{i'}|^2}$$

with some positive numbers $b_{i'} < 1$, $i' \in \{1, \ldots, N\}$.

Proof. As in the proof of Theorem 0.1, we let

$$\tilde{u}_t = u_t(x + x_t) - u_t(x_t), \ \tilde{\varphi}_t(x) = \tilde{u} - \psi_0,$$

where x_t is the minimal point of $w_t = tu + (1-t)\psi_0$ as in Lemma 2.3, but $|x_t| \to \infty$. Then $\tilde{\varphi} = \tilde{\varphi}_t(x)$ satisfies an equation,

$$\frac{\omega_{\tilde{\varphi}}^n}{\omega_g^n} = e^{-t\tilde{\varphi} - h + F + m_t}, \tag{3.15}$$

where m_t are uniformly bounded constants by Lemma 2.1 and

$$F = F_t = (1-t) \log \frac{\sum |s_i|^2}{\sum_i |s_i(\sigma_t)|^2 - \sum_i |s_i|(x_t)^2}.$$

Moreover, $\tilde{\varphi}(o) = -\psi_0(o)$, and consequently, by the mean value inequality,

$$\sup_M \tilde{\varphi} \leq C.$$

On the other hand, by the argument in [66], Li proved the Harnack inequality [42],

$$-\inf_M \tilde{\varphi} \leq (n+1) \sup_M \tilde{\varphi} + C.$$

Thus we get the C^0-estimate,

$$|\tilde{\varphi}| \leq C.$$

Since (3.15) is equivalent to a Ricci equation,

$$\mathrm{Ric}(\omega_{\tilde{\varphi}}) = t\omega_{\tilde{\varphi}} + (1-t)\sigma^* \omega_g,$$

In particular, $\mathrm{Ric}(\omega_{\tilde{\varphi}}) \geq 0$. By the Schwartz Lemma [16, 43], we have

$$\Delta_{g'}(\mathrm{tr}_{g'}(\omega_g) \geq -a(\omega_g)\, \mathrm{tr}_{g'}(\omega_g),$$

where the constant $-a(\omega_g)$ depends on the upper bound of curvature of g and g' is the Kähler metric associated to $\omega_{\tilde{\varphi}}$. Let $f = \log \mathrm{tr}_{g'}(\omega_g) - A\tilde{\varphi}$, where A is a large constant. Then

$$\Delta_{g'} f \geq \frac{A}{2} e^f - C.$$

By the maximum principle, we get

$$\mathrm{tr}_{g'}(\omega_g) \leq C'.$$

As a consequence, by (3.15), Li proved

$$\frac{1}{C}\omega_g \leq \omega' \leq C(1 + e^{F_t})\omega_g. \tag{3.16}$$

Next, we claim that F_t converges to F_∞ locally smoothly. In fact, we can choose a basis of s_i generated by the vertices p_i of P, since M is a toric Fano manifold. Then $|s_i|^2(x) = e^{\langle p_i, x\rangle}$. It follows

$$\sum_i |s_i|(\sigma_t)^2 - \sum_i |s_i|(x_t)^2 = \sum \frac{e^{\langle p_i, x_t\rangle}}{\sum e^{\langle p_l, x_t\rangle}}|s_i|^2.$$

Clearly, there is a limit for each sequence of $\frac{e^{\langle p_i, x_t\rangle}}{\sum e^{\langle p_l, x_t\rangle}}$ as $|x_t| \to \infty$. We write it by b_i' if it is not zero. Hence, F_t converges to F_∞. Moreover, $e^{F_\infty} \in L^p(M)$ with some $p > 1$.

Define an analytic set by

$$D = \left\{x \in M | \ \sum\nolimits_{i'} b_i'|s_i|^2 = 0\right\}. \tag{3.17}$$

Then by (3.16), (3.15) is uniformly elliptic away from D. By the regularity of uniformly elliptic equation (cf. [38, 78]), $\tilde{\varphi}_t$ are uniformly C^k-bounded away from D. It follows that there exists a sequence of $t_i \to \bar{T}$ such that $\omega_{\tilde{\varphi}_{t_i}}$ converges locally to a solution φ_∞ of (3.14) smoothly on $M \setminus D$. Note that $\varphi_\infty \in L^\infty(M)$ and $e^{F_\infty} \in L^p(M)$ with some $p > 1$. Hence, φ_∞ is also a current solution of (3.14) on M (cf. [41]). □

Example 3.6. In [42], Li computed $\bar{T}$ and the limit set D (3.17) on the blow-up of $\mathbb{C}P^2$ at two points as follows, $T = \frac{4}{21}$,

$$D = \{x \in M | \ |z_1|^2|z_2|^2(|z_1|^2 + |z_2|^2) = 0\},$$

and z_1 and z_2 are two defining sections of exceptional divisors D_1 and D_2, respectively. Thus the limit solution φ_∞ satisfies the following type equation near divisors D_1 and D_2,

$$(\omega_g + \sqrt{-1}\partial\bar{\partial}\varphi)^n = \frac{h}{|z_1|^2|z_2|^2(|z_1|^2 + |z_2|^2)},$$

where h is a non-zero smooth function. It is easy to see that ϕ_∞ cannot be extended as a conical metric on the point $D_1 \cap D_2$.

Remark 3.7. $\omega_{\tilde{\varphi}_{t_i}}$ in Theorem 3.5 has a Hausdorff–Gromov limit $(M_\infty, \omega_\infty)$ since $\mathrm{Ric}(\omega_{\tilde{\varphi}_{t_i}}) \geq t_i\omega_{\tilde{\varphi}_{t_i}}$. It is easy to see that $\omega_\infty = \omega_{\varphi_\infty}$ away from D. Furthermore, one can show that the completion of $\omega_{\varphi_\infty}|_{M\setminus D}$ coincides with ω_∞ (cf. [57, 61, 77], etc.).

We may ask the following question.

Question 3.8. *Is it true that* ω_{φ_∞} *in Theorem* (3.5) *is a conical metric along the smooth part of* D*?*

If the answer is true to Question 3.8, the codimension of set of non-conical singularities of Hausdorff–Gromov limit $(M_\infty, \omega_\infty)$ is at least 4.

4. Reduced K-energy $\mu(u)$

In this section, we describe the third proof of Theorem 0.1 by studying the properness of K-energy in [84]. For simplicity, we also assume that the Futaki invariant vanishes. In general case, we can use the modified K-energy introduced in [24] instead of K-energy to prove the existence of Kähler–Ricci soliton [80].

The K-energy $K(\phi)$ was introduced by Mabuchi in 1987 [48]. It plays an important role in the study of Kähler–Einstein metrics (cf. [7, 69]). According to the definition,

$$K(\phi) = -\frac{1}{V}\int_0^1 \int_M \dot{\phi}_t (S(\phi_t) - n)\frac{\omega_{\phi_t}^n}{n!} \wedge dt, \ \forall \phi \in \mathcal{H}(\omega_g), \tag{4.1}$$

where ϕ_t is a path in $\mathcal{H}(\omega_g)$ which connects 0 and ϕ, and $S(\phi_t)$ is the scalar curvature of ω_{ϕ_t}. Recall two Aubin functionals [1], i.e., I-functional and J-functional, by

$$I_{\omega_g}(\phi) = \frac{1}{V}\int_M \phi \left(\frac{\omega_g^n}{n!} - \frac{\omega_\phi^n}{n!}\right),$$

$$J_{\omega_g}(\phi) = \frac{1}{V}\int_0^1 \int_M \dot{\phi}_s \left(\frac{\omega_g^n}{n!} - \frac{\omega_{\phi_s}^n}{n!}\right) \wedge ds.$$

It is known that [1, 65],

$$\frac{n}{n+1} I_{\omega_g}(\phi) \geq J_{\omega_g}(\phi) \geq \frac{1}{n+1} I_{\omega_g}(\phi), \ \forall \ \phi \in \mathcal{H}(\omega_g). \tag{4.2}$$

Then $K(\cdot)$ can also be rewritten as (cf. [68]),

$$K(\phi) = \frac{1}{V}\int_M \log\left(\frac{\omega_\phi^n}{\omega_g^n}\right)\frac{\omega_\phi^n}{n!} - (I_{\omega_g}(\phi) - J_{\omega_g}(\phi)) + \frac{1}{V}\int_M h_g \left(\frac{\omega_g^n}{n!} - \frac{\omega_\phi^n}{n!}\right). \tag{4.3}$$

Definition 4.1. Let K be a maximal compact subgroup of $\mathrm{Aut}^0(M)$ and $\mathcal{H}_K(\omega_g)$ the subset of K-invariant Kähler in $\mathcal{H}(\omega_g)$. Let G_0 be another subgroup of $\mathrm{Aut}^0(M)$. $K(\cdot)$ is called *proper on* $\mathcal{H}_K(\omega_g)$ *modulo* G_0 if there is a function $f(t)$ ($t \in [-A, \infty)$) with property $f(t) \to \infty$ as $t \to \infty$ such that

$$K(\phi) \geq \inf_{\sigma \in G_0} f(I(\phi_\sigma)), \ \forall \phi \in \mathcal{H}_K(\omega),$$

where ϕ_σ is an induced Kähler potential defined by

$$\omega_{\phi_\sigma} = \sigma^*(\omega_g + \sqrt{-1}\partial\bar{\partial}\phi) = \omega_g + \sqrt{-1}\partial\bar{\partial}\phi_\sigma.$$

The following theorem was proved in [69, 74].

Theorem 4.2. *Let (M, g) be a Fano manifold with $\omega \in 2\pi c_1(M)$. Let K be a maximal compact subgroup of $\mathrm{Aut}^0(M)$ and G_0 another subgroup of $\mathrm{Aut}^0(M)$. Then M admits a Kähler–Einstein metric if $K(\cdot)$ is proper on $\mathcal{H}_K(\omega_g)$ modulo G_0.*

4.1. The reduction of *K*-energy

Denote $\mathcal{H}_K(\omega_g) \subset \mathcal{H}(\omega_g)$ to be the set of K-invariant Kähler potentials on a toric Fano manifold (M, g) with $\omega_g \in 2\pi c_1(M)$, where $K = (S^1)^n$. Then $\mathcal{H}_K(\omega_g)$ is equal to the set

$$\{\phi \in C^\infty(\mathbb{R}^n)|\ |\phi| < \infty \text{ and } \psi_0 + \phi \text{ is uniformly convex}\}.$$

By using the Legendre transformation $\xi = (\nabla\psi_0)^{-1}(x)$, one sees that the function (Legendre dual function) defined by

$$u_0(x) = \langle \xi, \nabla\psi_0(\xi)\rangle - \psi_0(\xi) = \langle \xi(x), x\rangle - \psi_0(\xi(x)),\ \forall\ x \in P \tag{4.4}$$

is uniformly convex. Set the space of symplectic potentials by

$$\mathcal{C} = \{u = u_0 + f \mid u \text{ is a uniformly convex function in } \ P,\ f \in C^\infty(\overline{P})\}.$$

It was shown in [2] that there is a bijection between $\mathcal{C}$ and $\mathcal{H}_K(\omega_g)$.

Let

$$\mathcal{L}(u) = \int_{\partial P} u\, d\sigma - n\int_P u\, dx. \tag{4.5}$$

and

$$\mu(u) = -\int_P \log\det(u_{ij})\, dx + \mathcal{L}(u). \tag{4.6}$$

Then we have

Proposition 4.3. *Let u_ϕ be the Legendre dual function of $\varphi = \psi_0 + \phi$ for any $\phi \in \mathcal{H}_K(\omega_g)$. Then*

$$K(\phi) = \frac{(2\pi)^n}{V}\mu(u) + \text{const.} \tag{4.7}$$

Proof. By (4.3), a direct computation shows

$$\begin{aligned}
\mu(\phi) &= \frac{1}{V}\int_M \log\left(\frac{\omega_\phi^n}{\omega_g^n}\right)\frac{\omega_\phi^n}{n!} - \left[\frac{1}{V}\int_0^1\int_M \dot{\phi}_t\frac{\omega_{\phi_t}^n}{n!}\wedge dt - \frac{1}{V}\int_M \phi\frac{\omega_\phi^n}{n!}\right]\\
&\quad - \frac{1}{V}\int_M h_g\left(\frac{\omega_\phi^n}{n!} - \frac{\omega_g^n}{n!}\right)\\
&= \frac{1}{V}\int_M \log\left(\frac{\omega_\phi^n}{\omega_g^n}e^{\phi - h_g}\right)\frac{\omega_\phi^n}{n!} - \frac{1}{V}\int_0^1\int_M \dot{\phi}_t\frac{\omega_{\phi_t}^n}{n!}\wedge dt + +\frac{1}{V}\int_M h_g\frac{\omega_g^n}{n!}\\
&= \frac{1}{V}\int_M \log\left(\frac{\omega_\phi^n}{\omega_g^n}e^{\phi - h_g}\right)\frac{\omega_\phi^n}{n!} - \frac{1}{V}\int_0^1\int_M \dot{\phi}_t\frac{\omega_{\phi_t}^n}{n!}\wedge dt + \text{const.}
\end{aligned} \tag{4.8}$$

On the other hand,

$$h_g = -\psi_0 - \log\det(\psi_{0ij}) + C.$$

Then

$$\frac{\omega_\phi^n}{\omega_g^n} e^{\phi - h_g} = C\det(\varphi_{ij}) e^{\varphi}.$$

It follows that

$$\int_M \log\left(\frac{\omega_\phi^n}{\omega_g^n} e^{\phi-h_g}\right)\frac{\omega_\phi^n}{n!} = (2\pi)^n\left[\int_{\mathbb{R}^n} \log\det(\varphi_{ij})\det(\varphi_{ij})\, d\xi + \int_{\mathbb{R}^n} \varphi\det(\varphi_{ij})\, d\xi\right]. \tag{4.9}$$

By using the relations

$$\varphi = \sum_{i=1}^n x_i u_i - u,\ \det(\varphi_{ij})\, d\xi = dx, \quad \dot\phi_t = -\dot u_t,$$

where ϕ_t is a path in $\mathcal{H}_K(\omega_g)$ and u_t is the symplectic potential of $\varphi_t = \psi_0 + \phi_t$, we also get

$$\begin{aligned} &\int_0^1 \int_M \dot\phi_t \frac{\omega_{\phi_t}^n}{n!} \wedge dt \\ &\quad = (2\pi)^n \int_0^1 \int_{\mathbb{R}^n} \dot\phi_t \det(\varphi_{t\,ij})\, d\xi \wedge dt = -(2\pi)^n \int_P u dx + \text{const.}, \end{aligned} \tag{4.10}$$

$$\begin{aligned} &\int_{\mathbb{R}^n} \log\det(\varphi_{ij})\det(\varphi_{ij})\, d\xi + \int_{\mathbb{R}^n} \varphi\det(\varphi_{ij})\, d\xi \\ &\quad = -\int_P \log\det(u_{ij})\, dx + \int_P \left(\sum\nolimits_{i=1}^n x_i u_i - u\right) dx. \end{aligned} \tag{4.11}$$

Hence inserting (4.9)–(4.11) into (4.8), we obtain

$$K(\phi) = \frac{(2\pi)^n}{V}\left[-\int_P \log\det(u_{ij})\, dx + \int_P \sum_{i=1}^n x_i u_i\, dx\right] + C.$$

Integrating by parts, we deduce (4.7) immediately. □

Remark 4.4. $\mu(u)$ is usually called reduced K-energy on a toric manifold. (4.7) was first obtained by Donaldson for K-invariant Kähler potentials in general Kähler class $[\omega_g]$ on a toric manifold while the linear functional $\mathcal{L}(\cdot)$ replaced by

$$\mathcal{L}(u) = \int_{\partial P} u\, d\sigma - \bar R \int_P u\, dx,$$

where $\bar R$ is the average of scalar curvature of g. Here we give a proof by using the formula (4.3) for $2\pi c_1(M)$. This argument can be generalized to prove an analogy of (4.7) for the modified K-energy (cf. [80]).

4.2. Properness of $\mathcal{K}(\phi)$

In this subsection, we verify the properness of $K(\phi)$ via $\mu(u)$. First, we need the following lemma due to Donaldson [26].

Lemma 4.5. *Let $\mathcal{C}_\infty$ be a set of C^∞-convex functions on $\bar{P}$. Then there exists a constant $C > 0$ such that for any $u \in \mathcal{C}_\infty$, it holds*

$$\int_P \log\det(u_{ij})dx \leq \mathcal{L}_B(u) + C, \tag{4.12}$$

where $B = (u_0)^{ij}_{ij}$ is a bounded function on $\bar{P}$, and

$$\mathcal{L}_B(u) = \int_{\partial P} u d\sigma + \int_P Bu\, dx. \tag{4.13}$$

Proof. Let $f = u - u_0$. By the convexity of $-\log\det$, we have

$$\log\det(u_{ij}) \leq \log\det((u_0)_{ij}) + (u_0)^{ij} f_{ij}.$$

For any $\delta > 0$, let P_δ be the interior polygon with faces parallel to those of P separated by distance δ, then f is smooth over the closure of P_δ.

Integrating by parts,

$$\int_{P_\delta} (u_0)^{ij} f_{ij}\, dx = \int_{\partial P_\delta} (u_0)^{ij} f_i n_j\, d\sigma_0 - \int_{P_\delta} (u_0)^{ij}_j f_i\, dx.$$

Integrating by parts for the last two terms again, we have

$$\int_{P_\delta} (u_0)^{ij} f_{ij}\, dx = \int_{\partial P_\delta} (u_0)^{ij} f_i n_j\, d\sigma_0 - \int_{\partial P_\delta} (u_0)^{ij}_j n_i f d\sigma_0 + \int_{P_\delta} (u_0)^{ij}_{ij} f\, dx.$$

Note that

$$(u_0)^{ij} n_j\, d\sigma_0 \to 0, \quad -(u_0)^{ij}_j n_i d\sigma_0 \to\ d\sigma$$

as $\delta \to 0$ [26, 27]. Then

$$\int_{\partial P_\delta} (u_0)^{ij} f_i n_j\, d\sigma_0 \quad \text{and} \quad \int_{\partial P_\delta} (u_0)^{ij}_j n_i f\, d\sigma_0 \longrightarrow \int_{\partial P} f\, d\sigma$$

as $\delta \to 0$. In conclusion,

$$\int_P (u_0)^{ij} f_{ij}\, dx = \int_{\partial P} f\, d\sigma + \int_P Bf\, dx.$$

Hence,

$$\begin{aligned}
&\int_P \log\det(u_{ij})\, dx \\
&\quad \leq \int_{\partial P} u\, d\sigma + \int_P Bu\, dx + \int_{\partial P} u_0\, d\sigma - \int_P Bu_0\, dx + \int_P \log\det((u_0)_{ij}\, dx \\
&\quad = \int_{\partial P} u\, d\sigma + \int_P Bu\, dx + \text{const.}
\end{aligned}$$

$\square$

The vanishing of Futaki invariant implies that $\mathcal{L}(u)$ is invariant when adding u by a linear function. We call a convex function normalized at $0 \in P$ if $\inf_P u = u(0)$. Let $\tilde{\mathcal{C}}_\infty$ be the set of such normalized functions in $\mathcal{C}_\infty$. Then we prove

Lemma 4.6. *There exists a $\lambda > 0$ such that*

$$\mathcal{L}(u) \geq \lambda \int_{\partial P} u\, d\sigma, \quad u \in \tilde{\mathcal{C}}_\infty. \tag{4.14}$$

Proof. Since $d\sigma = \langle \vec{n}, x\rangle\, d\sigma_0$, we have

$$\int_{\partial P u} d\sigma = \int_P \operatorname{div}(xu)\, dx = \int_P \left(nu + \sum_{i=1}^n x_i u_i \right) dx.$$

It follows that

$$\mathcal{L}(u) = \int_P \sum_{i=1}^n x_i u_i\, dx = \int_P \left[\left(\sum_{i=1}^n x_i u_i - u \right) + u \right] dx \geq \int_P u\, dx. \tag{4.15}$$

We claim that (4.15) implies (4.14).

By the contradiction, we suppose that (4.14) is not true. Then there is a sequence of functions $\{u_k\}$ in $\tilde{\mathcal{C}}_\infty$ such that

$$\int_{\partial P} u_k\, d\sigma = 1 \tag{4.16}$$

and

$$\mathcal{L}(u_k) \longrightarrow 0, \text{ as } \ k \longrightarrow \infty. \tag{4.17}$$

By (4.16), there exists a subsequence (still denoted by $\{u_k\}$) of $\{u_k\}$, which converges locally uniformly to a convex function $u_\infty \geq 0$ on P. By (4.15), we have

$$\int_P u_k\, dx \leq \mathcal{L}(u_k) \longrightarrow 0.$$

Thus

$$\int_P u_\infty\, dx = 0.$$

Hence, we obtain $u_\infty \equiv 0$ in P. On the other hand,

$$\begin{aligned} \mathcal{L}(u_k) &= \int_{\partial P} u^k d\sigma - n \int_P u_k\, dx \\ &\longrightarrow \quad 1 - n \int_P u_\infty dx = 1 > 0. \end{aligned}$$

This contradicts with (4.17). Thus (4.14) is true and the lemma is proved. □

Proposition 4.7. *Suppose that the Futaki invariant vanishes on M. Then there exists C_δ such that*

$$\mu(u) \geq \delta \int_P u dy - C_\delta, \quad \forall u \in \tilde{\mathcal{C}}_\infty. \tag{4.18}$$

Proof. We compute the difference of $\mathcal{L}(u)$ and $\mathcal{L}_B(u)$ as follows,

$$\begin{aligned}|\mathcal{L}(u)-\mathcal{L}_B(u)| &= \left|\int_P (n+B)u\,dx\right| \\ &\le C'\int_P u\,dx \\ &\le (1+\delta)C_0C'\int_{\partial P} u\,d\sigma - \delta C'\int_P u\,dx\end{aligned}$$

where $C' = \|n+B\|_{L^\infty}$. Note

$$\int_P u\,dx \le C_0\int_{\partial P} u d\sigma,\ \forall\ u\in\tilde{\mathcal{C}}_\infty.$$

Then by (4.14), it follows

$$|\mathcal{L}(u)-\mathcal{L}_B(u)| \le \frac{(1+\delta)C_0C'}{\lambda}\mathcal{L}(u) - \delta C'\int_P u dx.$$

Thus

$$\left(1+\frac{(1+\delta)C_0C'}{\lambda}\right)\mathcal{L}(u) \ge \mathcal{L}_B(u) + \delta C'\int_P u\,dx.$$

Now let $r = \left(1+\frac{(1+\delta)C_0C'}{\lambda}\right)^{-1}$, we get

$$\mathcal{L}(u) \ge \mathcal{L}_B(ru) + r\delta C'\int_P u\,dx. \tag{4.19}$$

On the other hand, by applying the inequality (4.12) to ru, we have

$$-\int_P \log\det(u_{ij})\,dx \ge -\mathcal{L}_B(ru) - C + n\log r.$$

Hence, combining it with (4.19), we obtain

$$\mu(u) \ge r\delta C'\int_P u\,dx - C + n\log r. \qquad \square$$

By Proposition 4.7 and Proposition 4.3, we prove

Theorem 4.8. *There exist numbers $\delta > 0$ and C such that*

$$K(\phi) \ge \delta \inf_{\tau\in T^{\mathbb{C}}} I_{\omega_g}(\phi_\tau) - C,\ \forall\ \phi\in\mathcal{H}_K(\omega_g). \tag{4.20}$$

In particular, $K(\phi)$ is proper for any $\phi\in\mathcal{H}_K(\omega_g)$ modulo $T^{\mathbb{C}}$.

Theorem 4.8 and Theorem 4.2 imply Theorem 0.1.

With respect to Theorem 4.8, we propose the following conjecture.

Conjecture 4.9. *Let (M, g) be an n-dimensional Kähler–Einstein manifold with $\omega_g = 2\pi c_1(M)$. Then there are $\delta, C_\delta > 0$ such that for any K-invariant Kähler potential ϕ of ω_g it holds*

$$\mathcal{K}(\phi) \geq \delta \inf_{\tau \in Z(\mathrm{Aut}(M))} I(\phi_\tau) - C_\delta, \tag{4.21}$$

where K is a maximal compact subgroup of $\mathrm{Aut}^0(M)$ and $Z(\mathrm{Aut}(M))$ is the center of $\mathrm{Aut}^0(M)$).

Conjecture 4.9 can be regarded as a version of Tian's conjecture for K-invariant Kähler potentials proposed in [69]. We note that Conjecture 4.9 is true by a result of Tian [69], if $\mathrm{Aut}(M)$ is finite. Recently, Darvas and Rubinstein proved Tian's conjecture when $Z(\mathrm{Aut}(M))$ is replaced by $\mathrm{Aut}(M)$ [31].

5. Kähler–Einstein metrics on G-manifolds

In this section, we discuss a recent result of Delcroix for the existence of Kähler–Einstein metrics on a G-manifold M in [33], where G is a reductive complex Lie group. Here we will give another proof of Delcroix' Theorem (cf. Theorem 5.1 below) by verifying the properness of K-energy as we did in Section 4 [45]. M is called a *G-manifold* (*bi-equivariant compactification of G*) if it admits a holomorphic $G \times G$ action on M with an open and dense orbit isomorphic to G as a $G \times G$-homogeneous space. (M, L) is called a *polarized G-manifold* if L is a $G \times G$-linearized ample line bundle on M. In general, there are many different compactifications of G with an extended $G \times G$ action, and the compactification space may not be a smooth manifold, perhaps just an algebraic variety [3–5]. Clearly, toric manifolds are a class of simplest G-manifolds. We will discuss more examples in the Appendix at the end of this paper.

Let us introduce some notations for G-manifolds. Assume that $T^{\mathbb{C}}$ is a r-dimensional maximal complex torus of G and $\mathfrak{M}$ is its group of characters. Denote the roots system of $(G, T^{\mathbb{C}})$ in $\mathfrak{M}$ by Φ and choose a set of positive roots $\Phi_+ = \{\alpha_{(i)}\}_{i=1,\ldots,\frac{n-r}{2}}$. Let P be the polytope associated to (M, L), and P_+ the positive part of P defined by Φ_+ such that $P_+ = \{y \in P | \ \langle \alpha, y\rangle > 0, \ \forall\, \alpha \in \Phi_+\}$, Here $\langle \cdot, \cdot \rangle$ denotes the Cartan–Killing inner product on the dual space $\mathfrak{a}^*$ of the real part $\mathfrak{a}$ of Lie algebra $\mathfrak{t}^{\mathbb{C}}$ of $T^{\mathbb{C}}$. We call $W_\alpha = \{y \in \mathfrak{a}^* | \ \langle \alpha, y \rangle = 0\}$ the Weyl wall associated to $\alpha \in \Phi_+$. Set a function on $\mathfrak{a}^*$ by

$$\pi(y) = \prod_{\alpha \in \Phi_+} \langle \alpha, y \rangle^2, \ \ y \in \mathfrak{a}^*.$$

Clearly, $\pi(y)$ vanishes on $\partial P_+ \cap W_\alpha$ for each $\alpha \in \Phi_+$. Denote by $2P_+$ a dilation of P_+ at rate 2. We define a barycentre of $2P_+$ with respect to the weighted measure

$\pi(y)dy$ by

$$bar(2P_+) = \frac{\int_{2P_+} y\pi(y)\,dy}{\int_{2P_+} \pi(y)\,dy}.$$

More recently, Delcroix extended Theorem 0.1 to a Fano G-manifold M and proved the following theorem [33].

Theorem 5.1. *Let M be a Fano G-manifold. Then M admits a Kähler–Einstein metric if and only if*

$$bar(2P_+) \in 4\rho + \Xi, \tag{5.1}$$

where $\rho = \frac{1}{2}\sum_{\alpha\in\Phi_+} \alpha$ is a vector in $\mathfrak{a}^*$ and Ξ is the relative interior of the cone generated by Φ_+.

It was pointed by Delcroix that (5.1) implies that the Futaki invariant vanishes for holomorphic vector fields induced by $G \times G$, but the converse is not true in general. In fact, (5.1) is related to the K-stability and is determined by a generalized Futaki invariant for some test-configurations in terms of [26, 68] (cf. [33, 45]). In particular, M is K-unstable if $bar(2P_+) \notin \overline{4\rho + \Xi}$. In this section, we will use an argument in [84, 85] to derive an analytic obstruction to the existence of Kähler metrics with constant scalar curvature on a G-manifold in terms of convex piecewise linear functions. Then constructing a piecewise Weyl-invariant linear function as in [45], the analytic obstruction implies (5.1).

The sufficient part of Theorem 5.1 was proved by Delcroix using the method in [79] to establish an analogy of Lemma 2.1 and Lemma 2.3, respectively. However the reduced real Monge–Ampère equation from the Kähler–Einstein metric equation as in (0.2) is degenerate, which is defined in a cone $\mathfrak{a}_+ = \{x \in \mathfrak{a} |\ \alpha(x) > 0,\ \forall\, \alpha \in \Phi_+\}$ of $\mathfrak{a}$. Thus there need more delicate estimates to do. In [45], we gave another proof of Theorem 5.1 by verifying the properness of K-energy as done on a toric manifold in [84]. The method can work for the K-energy on a general polarized compactification (M, L) of G.

Let us introduce more notations before we state our main results in [45]. We divide $\partial(2P_+) \cap \partial(2P)$ into several pieces $\{F_A\}_{A=1}^{d_0}$ such that for any A, F_A lies on an $(r-1)$-dimensional hyperplane defined by $\langle y, u_A\rangle = \lambda_A$ for some primitive $u_A \in \mathfrak{N}$, where $\mathfrak{N}$ is the $\mathbb{Z}$-dual of $\mathfrak{M}$. Define a cone by $E_A = \{ty |\ t \in [0,1],\ y \in F_A\}$ for any A. It is clear that $2P_+ = \bigcup_{A=1}^{d_0} E_A$. Let

$$\Lambda_A = \frac{2}{\lambda_A}\left(1 + \langle 2\rho, u_A\rangle\right).$$

Then the average of scalar curvature $\bar{S}$ of $\omega_0 \in 2\pi c_1(L)$ is given by

$$\bar{S} = \frac{n\sum_A \Lambda_A \int_{E_A} \pi\, dy}{\int_{2P_+} \pi\, dy}.$$

Define a weighted barycentre $\widetilde{bar}$ of $2P_+$ by

$$\widetilde{bar} = \frac{\sum_A \Lambda_A \int_{E_A} y\pi \, dy}{\sum_A \Lambda_A \int_{E_A} \pi \, dy}.$$

Note that both bar and $\widetilde{bar}$ lie in $\mathfrak{a}^*$. Denote by bar_{ss} and $\widetilde{bar}_{ss}$ the projections of bar and $\widetilde{bar}$ on the semisimple part $\mathfrak{a}^*_{ss}$ of $\mathfrak{a}^*$, respectively. Then we have

Theorem 5.2. *Let (M, L) be a polarized compactification of G with vanishing Futaki invariant, and $\omega_0 \in 2\pi c_1(L)$ a $K \times K$-invariant Kähler metric, where K is a maximal compact subgroup of G, which complexifies G. Suppose that the polytope $2P_+$ satisfies the following conditions,*

$$\left(\min_A \Lambda_A \cdot \widetilde{bar}_{ss} - 4\rho\right) \in \Xi, \tag{5.2}$$

$$\left(\widetilde{bar}_{ss} - bar_{ss}\right) \in \bar{\Xi}, \tag{5.3}$$

$$(n+1) \cdot \min_A \Lambda_A - \bar{S} > 0. \tag{5.4}$$

Then the K-energy $K(\cdot)$ is proper on $\mathcal{H}_{K\times K}(\omega_0)$ modulo $Z(G)$, where $\mathcal{H}_{K\times K}(\omega_0)$ is the space of $K \times K$-invariant Kähler potentials in $2\pi c_1(L)$ and $Z(G)$ is the centre of G.

In case that M is Fano and $L = K_M^{-1}$, then $\bar{S} = n$ and $\Lambda_A = 1$ for all A. We have $\widetilde{bar} = bar$, thus (5.3), (5.4) are automatically satisfied. Moreover, (5.1) is equivalent to the vanishing of Futaki invariant and (5.2). Consequently, $K(\cdot)$ is proper modulo the action of $Z(G)$. Hence we get an alternative proof for the sufficient part of Theorem 5.1.

As mentioned above, we prove Theorem 5.2 by using the reduced K-energy $\mu(\cdot)$. One of the advantages of $\mu(\cdot)$ is that it can be defined on a complete space $\tilde{\mathcal{C}}_*$ of convex functions on $2P_+$. Following the argument in [86], we can discuss the semi-continuity property of $\mu(\cdot)$ and prove the following

Theorem 5.3. *$\mu(u)$ is lower semi-continuous on $\tilde{\mathcal{C}}_*$. Furthermore, if $K(\cdot)$ is proper on $\mathcal{H}_{K\times K}(\omega_0)$ modulo $Z(G)$, then there exists a minimizer of $\mu(\cdot)$ on $\tilde{\mathcal{C}}_*$.*

It is interesting to study the regularity of minimizers in Theorem 5.3. We guess that they are smooth in $2P_+$ if the dimension of the torus $T^{\mathbb{C}}$ is less than two. In case of toric surfaces, it is verified by Zhou in [82, 83].

In the following subsections, we outline a proof of Theorem 5.2 and prove the necessary part of Theorem 5.1. First we give a formula of scalar curvature under the Legendre transformation.

5.1. Reduced scalar curvature equation on $\mathfrak{a}_+$

Let Z be the closure of $T^{\mathbb{C}}$ in M. It is known that $(Z, L|_Z)$ is a polarized toric manifold with a Weyl action, and $L|_Z$ is a W-linearized ample toric line bundle on Z [3–5, 34]. Let $\omega_0 \in 2\pi c_1(L)$ be a $K\times K$-invariant Kähler form induced from (M, L) and P be the polytope associated to $(Z, L|_Z)$, which is defined by the moment map associated to ω_0. Then P is a W-invariant Delzant polytope in $\mathfrak{a}^*$. By the $K\times K$-invariance, for any $\phi \in \mathcal{H}_{K\times K}(\omega_0)$, the restriction of ω_ϕ on Z is a toric Kähler metric. It induces a smooth strictly convex function ψ on $\mathfrak{a}$, which is W-invariant [6].

By the KAK-decomposition ([40], Theorem 7.39), for any $g \in G$, there are k_1, $k_2 \in K$ and $x \in \mathfrak{a}$ such that $g = k_1 \exp(x) k_2$. Here x is uniquely determined up to a W-action. This means that x is unique in $\bar{\mathfrak{a}}_+$. Then we define a smooth $K \times K$-invariant function Ψ on G by

$$\Psi(\exp(\cdot)) = \psi(\cdot) : \ \mathfrak{a} \to \mathbb{R}.$$

Clearly Ψ is well defined since ψ is W-invariant. We usually call ψ *the function associated to* Ψ. It can be verified that Ψ is a Kähler potential on G such that $\omega = \sqrt{-1}\partial\bar\partial\Psi$ on G. Actually, we have the following lemma, which is due to Delcroix [34, Theorem 1.2].

Lemma 5.4. *Let Ψ be a $K\times K$ invariant function on G, and ψ the associated function on $\mathfrak{a}$. Let $\Phi_+ = \{\alpha_{(1)}, \ldots, \alpha_{(\frac{n-r}{2})}\}$. Then there are local holomorphic coordinates on G such that for $x \in \mathfrak{a}_+$, the complex Hessian matrix of Ψ is diagonal by blocks as follows,*

$$\mathrm{Hess}_{\mathbb{C}}(\Psi)(\exp(x)) = \begin{pmatrix} \frac{1}{4}\mathrm{Hess}_{\mathbb{R}}(\psi)(x) & 0 & & 0 \\ 0 & M_{\alpha_{(1)}}(x) & & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & & M_{\alpha_{(\frac{n-r}{2})}}(x) \end{pmatrix}, \tag{5.5}$$

where

$$M_{\alpha_{(i)}}(x) = \frac{1}{2}\langle \alpha_{(i)}, \nabla\psi(x)\rangle \begin{pmatrix} \coth \alpha_{(i)}(x) & \sqrt{-1} \\ -\sqrt{-1} & \coth \alpha_{(i)}(x) \end{pmatrix}.$$

By (5.5) in Lemma 5.4, we see that ψ is convex on $\mathfrak{a}$. The complex Monge–Ampère measure is given by $\omega_\phi^n = (\sqrt{-1}\partial\bar\partial\Psi)^n = MA_{\mathbb{C}}(\Psi)\, dV_G$, where

$$\mathrm{MA}_{\mathbb{C}}(\Psi)(\exp(x)) = \frac{1}{4^{r+p}} \mathrm{MA}_{\mathbb{R}}(\psi)(x) \frac{1}{\mathbf{J}(x)} \prod_{\alpha\in\Phi_+} \langle \alpha, \nabla\psi(x)\rangle^2 \tag{5.6}$$

and

$$J(x) = \prod_{\alpha \in \Phi_+} \sinh^2 \alpha(x).$$

Let

$$\chi(x) = -\log \mathbf{J}(x) = -2 \sum_{\alpha \in \Phi_+} \log \sinh \alpha(x).$$

Then applying Lemma 5.4 to the $K \times K$-invariant function

$$\tilde{\psi} = \log \det(\nabla^2 \psi) + 2 \sum_{\alpha \in \Phi_+} \log \alpha(\nabla \psi) + \chi(x),$$

$\mathrm{Ric}(\omega_\phi)$ can be expressed as

$$\begin{aligned} &- \mathrm{Hess}_{\mathbb{C}}(\log \det(\partial \bar{\partial} \Psi))(\exp(x)) \\ &= - \begin{pmatrix} \frac{1}{4}\mathrm{Hess}_{\mathbb{R}}(\tilde{\psi})(x) & 0 & & & 0 \\ 0 & \tilde{M}_{\alpha_{(1)}}(x) & & & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & & & \tilde{M}_{\alpha_{(\frac{n-r}{2})}}(x) \end{pmatrix}, \end{aligned} \tag{5.7}$$

where $x \in \mathfrak{a}_+$. Thus we get the following formula of scalar curvature of ω_ϕ,

$$S(\omega_\phi)|_{\exp(x)} = -\mathrm{tr}\left((\nabla^2 \psi)^{-1} \nabla^2 \tilde{\psi}\right) - \sum_{\alpha \in \Phi_+} \frac{\langle \alpha, \nabla \tilde{\psi} \rangle}{\langle \alpha, \nabla \psi \rangle}. \tag{5.8}$$

Let u be a Legendre function of ψ defined as in (4.4). Then by (5.8), one can show that

$$\begin{aligned} S(\omega_\phi) = &- u^{ij}_{,ij} - 2u^{ij}_{,j} \frac{\pi_{,i}}{\pi} - u^{ij} \frac{\pi_{,ij}}{\pi} \\ &- u_{,ik} \left. \frac{\partial^2 \chi}{\partial x^i \partial x^k} \right|_{x = \nabla u} - \left. \frac{\partial \chi}{\partial x^i} \right|_{x = \nabla u} \frac{\pi_{,i}}{\pi}. \end{aligned} \tag{5.9}$$

(5.9) can be rewritten as

$$S(\omega_\phi) = -\frac{1}{\pi} \left((u^{ij} \pi)_{,ij} + \frac{\partial}{\partial v_i} \left(\pi \left. \frac{\partial \chi}{\partial x^i} \right|_{x = \nabla u} \right) \right). \tag{5.10}$$

(5.9) can be regarded as Abreu's equation of scalar curvature on a G-manifold, which is defined on P_+. We note that u can be extended smoothly up to the Weyl walls W_β of P_+ although the functions π is degenerate and χ is singular on W_β, respectively.

5.2. A sketch of proof of Theorem 5.2

In this subsection, we outline a proof of Theorem 5.2 by using the argument in Section 4. First we have an analogy of Proposition 4.3 as follows.

Proposition 5.5. *Let $\phi \in \mathcal{H}_{K\times K}(\omega_0)$ and u be the Legendre function of $\psi = \psi_0 + \phi$. Let $V = \int_{2P_+} \pi\, dy$. Then*

$$K(\phi) = \frac{1}{V}\mu(u) + \text{const.},$$

where $\mu(u)$ is a reduced K-energy of $K(\phi)$ defined by

$$\mu(u) = \sum_A \int_{F_A} \Lambda_A \langle y, \nu_A\rangle u\pi\, d\sigma_0 - \int_{2P_+} \bar{S} u\pi\, dy - \int_{2P_+} \log\det(u_{ij})\pi\, dy + \int_{2P_+} \chi(\nabla u)\pi\, dy.$$

Proposition 5.5 can be proved by using the formula (5.9) of scalar curvature as in the proof of Proposition 4.3 [45].

For convenience, we write $\mu(u)$ as $\mu(u) = \mathcal{L}(u) + \mathcal{N}(u)$, where

$$\mathcal{L}(u) = \sum_A \int_{F_A} \Lambda_A \langle y, \nu_A\rangle u\pi\, d\sigma_0 - \int_{2P_+} \bar{S} u\pi\, dy - \int_{2P_+} 4\langle \rho, \nabla u\rangle \pi\, dy, \tag{5.11}$$

$$\mathcal{N}(u) = -\int_{2P_+} \log\det\left(u_{,ij}\right)\pi\, dy + \int_{2P_+} [\chi\left(\nabla u\right) + 4\langle \rho, \nabla u\rangle]\pi\, dy. \tag{5.12}$$

By integration by parts, we can rewrite $\mathcal{L}(u)$ as

$$\mathcal{L}(u) = \sum_A \int_{E_A} \left[\langle \Lambda_A y - 4\rho, \nabla u\rangle + (\Lambda_A n - \bar{S})u\right]\pi\, dy, \tag{5.13}$$

or

$$\mathcal{L}(u) = \sum_A \frac{2}{\lambda_A} \int_{F_A} \langle y, \nu_A\rangle u\pi\, d\sigma_0 - \int_{2P_+} \bar{S} u\pi\, dy + \int_{2P_+} 4\langle \rho, \nabla \pi\rangle u\, dy. \tag{5.14}$$

Next we estimate the terms of $\mathcal{L}(u)$ and $\mathcal{N}(u)$. For convenience, we denote the set of smooth convex W-invariant functions on $\bar{P}$ by $\mathcal{C}_{\infty,W}$. Clearly, $\mathcal{L}(u)$ is well defined on $\mathcal{C}_{\infty,W}$. Moreover, $\mathcal{N}(u)$ is also well defined as we will see below. We want to normalize u as follows. Let O be the origin of $\mathfrak{a}^*$. Note that the dual $\mathfrak{a}_t^*$ of center of Lie algebra $\mathfrak{g}$ of G is the fixed point set of the W-action. Then $\nabla u(O) \in \mathfrak{a}_t^*$ for any $u \in \mathcal{C}_{\infty,W}$. Thus we can normalize $u \in \mathcal{C}_{\infty,W}$ by

$$\tilde{u}(y) = u(y) - \langle \nabla u(O), y\rangle - u(O). \tag{5.15}$$

Clearly, $\tilde{u} \in \mathcal{C}_{\infty,W}$ and

$$\min_{2P} \tilde{u} = \tilde{u}(O) = 0. \tag{5.16}$$

The subset of normalized functions in $\mathcal{C}_{\infty,W}$ will be denoted by $\hat{\mathcal{C}}_{\infty,W}$.

By the argument in [84] together with the conditions (5.2)–(5.4) in Theorem 5.2, the following lemma was proved in [45].

Lemma 5.6. *Under the assumption of Theorem* 5.2*, there exists a positive constant* λ *such that*

$$\mathcal{L}(u) \geq \lambda \int_{\partial(2P_+)} \langle y, \nu \rangle u\pi \, d\sigma_0, \ \ \forall u \in \hat{\mathcal{C}}_{\infty,W}.$$

The following lemma gives a comparison between $\mathcal{L}(u)$ and $\mathcal{N}(u)$ (cf. [45, Proposition 4.4]).

Lemma 5.7. *There exist uniform constants* C_Λ, C_L, $C_0 > 0$ *such that for any* $u \in \hat{\mathcal{C}}_{\infty,+}$,

$$\mathcal{N}(u) \geq -C_L\mathcal{L}(u) - C_\Lambda \int_{\partial(2P_+)} u\langle y, \nu \rangle \pi \, d\sigma_0 - C_0 + \int_{2P_+} Qu\pi \, dy, \tag{5.17}$$

where

$$Q = -\left.\frac{\partial \chi}{\partial x^i}\right|_{x=\nabla u_0} \frac{\pi_{,i}}{\pi} - \left.\frac{\partial^2 \chi}{\partial x^i \partial x^k}\right|_{x=\nabla u_0} u_{0,ik} - u_0^{ij} \frac{\pi_{,ij}}{\pi}. \tag{5.18}$$

Since Q is singular and π vanishes along each W_α, we shall give an explicit estimate for the singular order of Q. Actually, we prove (cf. [45, Proposition 4.5]),

Lemma 5.8. *There are constants* $C_I, C_{II} > 0$ *independent of* u *such that*

$$\left| \int_{2P_+} Qu\pi \, dy \right| \leq C_I \int_{2P_+} \langle \rho, \nabla \pi \rangle u \, dy + C_{II} \int_{2P_+} u\pi \, dy, \ \ \forall u \in \hat{\mathcal{C}}_{\infty,W}.$$

By Lemma 5.7 and Lemma 5.8, we see that $\mathcal{N}(u)$ is well defined on $\mathcal{C}_{\infty,W}$. Combining Lemmas 5.6–5.8, as in the proof of Proposition 4.7, we finally prove

Proposition 5.9. *Under the assumption of Theorem* 5.2*, there exists* $\delta > 0$ *and* C_δ *such that*

$$\mathcal{K}(u) \geq \delta \int_{2P_+} u\pi \, dy - C_\delta, \ \forall \ u \in \hat{\mathcal{C}}_{\infty,W}. \tag{5.19}$$

Theorem 5.2 follows from Propositions 5.5 and 5.9 immediately.

5.3. Proof of the necessary part of Theorem 5.1

The following proposition gives an analytic obstruction to the existence of Kähler metrics with constant scalar curvature on a G-manifold in terms of convex piecewise linear functions.

Proposition 5.10. *Suppose that a* G*-manifold* M *admits a Kähler metric* ω_ϕ *with constant scalar curvature. Then for any convex* W*-invariant piecewise linear function* f *on* $2P$*, we have*

$$\mathcal{L}(f) \geq 0.$$

Moreover, the equality holds if and only if

$$f(v) = a^i v_i$$

for some $a = (a^i) \in \mathfrak{a}_z$, *where* $\mathfrak{a}_z = z(g) \cap \mathfrak{a}$ *and* $z(g)$ *is the center of Lie algebra of* G.

Proof. Note that a convex W-invariant piecewise linear function f on $2P$ can be written as

$$f = \max_{1 \le N \le N_0} \{f_N\},$$

where f_N is W-invariant such that

$$f_N|_{P_+}(v) = a_N^i v_i + c_N$$

for some constant vector $a_N = (a_N^i)$. Moreover, it can be showed that $a_N \in \overline{\mathfrak{a}_+}$ (cf. [45, Proposition 3.4]). Then we can divide $2P_+$ into τ_0 sub-polytopes $P_1, \ldots, P_{\tau_0}$ such that for each $\tau = 1, \ldots, \tau_0$, there is an $N(\tau) \in \{1, \ldots, N_0\}$ with

$$f|_{P_\tau} = f_{N(\tau)}.$$

For simplicity, we write f_τ as $f_{N(\tau)}$.

By a Calabi result [15], ω_ϕ is a $K \times K$-invariant metric. Then by (5.10), we have on each P_τ,

$$-\bar{S} \int_{P_\tau} f\pi \, dv = \int_{P_\tau} \left((u^{ij}\pi)_{,ij} + \frac{\partial}{\partial v_i} \left(\pi \left. \frac{\partial \chi}{\partial x^i} \right|_{x = \nabla u} \right) \right) f \, dv, \tag{5.20}$$

where $\bar{S} = S(\omega_\phi)$ is the average of $S(\omega_\phi)$ and u is a Legendre function of ψ. Note that $f_{,ij} = 0$ on each P_τ. Taking integration by parts, we get

$$\int_{P_\tau} (u^{ij}\pi)_{,ij} f \pi dv = \int_{\partial P_\tau} (u_{0,j}^{ij} \nu_i \pi + u^{ij} \pi_{,i} \nu_j) f \, d\sigma_0 - \int_{\partial P_\tau} u^{ij} \nu_i f_{,j} \pi \, d\sigma_0$$

and

$$\int_{P_\tau} \frac{\partial}{\partial v_i} \left(\pi \left. \frac{\partial \chi}{\partial x^i} \right|_{x = \nabla u} \right) f \, d\sigma_0 = \int_{\partial P_\tau} \nu_i \left. \frac{\partial \chi}{\partial x^i} \right|_{x = \nabla u} f\pi \, d\sigma_0 - \int_{p_\tau} \left. \frac{\partial \chi}{\partial x^i} \right|_{x = \nabla u} f_{,i} \pi \, dv.$$

Plugging the above relations into (5.20), it follows

$$\begin{aligned} -\bar{S} \int_{P_\tau} f\pi \, dv = &\int_{\partial P_\tau} \left(u_{0,j}^{ij} \nu_i \pi + u^{ij} \pi_{,i} \nu_j + \nu_i \pi \left. \frac{\partial \chi}{\partial x^i} \right|_{x = \nabla u} \right) f \, d\sigma_0 \\ &- \int_{\partial P_\tau} u^{ij} \nu_i f_{,j} \pi \, d\sigma_0 - \int_{P_\tau} \left. \frac{\partial \chi}{\partial x^i} \right|_{x = \nabla u} f_{,i} \pi \, dv. \end{aligned}$$

Thus summing over τ, by the argument of [84, Proposition 2.2], we obtain

$$-\bar{S}\int_{2P_+} f\pi\, dv = \sum_{\tau_1<\tau_2}\int_{\partial P_{\tau_1}\cap\partial P_{\tau_2}} \frac{u^{ij}(a^i_{\tau_1}-a^i_{\tau_2})(a^j_{\tau_1}-a^j_{\tau_2})}{|a_{\tau_1}-a_{\tau_2}|}\pi\, d\sigma_0 \qquad (5.21)$$
$$-\sum_A \Lambda_A \int_{\mathfrak{F}'_A\cap\partial P_+} f\langle v,\nu_A\rangle\pi\, d\sigma_0 - \sum_\tau\int_{P_\tau}\frac{\partial\chi}{\partial x^i}\Big|_{x=\nabla u} a^i_\tau \pi\, dv.$$

Recall (5.13). We see that

$$V_P\cdot\mathcal{L}(f) = \sum_A \Lambda_A\int_{\mathfrak{F}'_A\cap\partial 2P_+} f\langle v,\nu_A\rangle\pi\, d\sigma_0 - \bar{S}\int_{2P_+} f\pi\, dv - 4\sum_\tau\int_{P_\tau}\sigma(a_\tau)\pi\, dv. \qquad (5.22)$$

Note that for any $a_\tau = (a^i_\tau)\in\overline{\mathfrak{a}_+}$,

$$-a^i_\tau\frac{\partial\chi}{\partial x^i} - 4\sigma_i a^i_\tau = 2\sum_{\alpha\in\Phi_+}(\coth\alpha(x)-1)\alpha(a_\tau)\geq 0,\ \forall\ x\in\mathfrak{a}_+.$$

Hence, plugging (5.21) into (5.22), we derive

$$V_P\cdot\mathcal{L}(f) = \sum_{\tau_1<\tau_2}\int_{\partial P_{\tau_1}\cap\partial P_{\tau_2}}\frac{u_0^{ij}(a^i_{\tau_1}-a^i_{\tau_2})(a^j_{\tau_1}-a^j_{\tau_2})}{|a_{\tau_1}-a_{\tau_2}|}\pi\, d\sigma_0$$
$$+2\sum_\tau\sum_{\alpha\in\Phi_+}\int_{P_\tau}(\coth\alpha(x)-1)\alpha(a_\tau)\pi\, dv\geq 0. \qquad (5.23)$$

It is easy to see that the equality in (5.23) holds if and only there is an $a=(a^i)\in\overline{\mathfrak{a}_+}$ such that

$$a_\tau = a,\ \forall\ \tau \quad\text{and}\quad \alpha(a)=0,\ \forall\ \alpha\in\Phi_+.$$

The second relation means that $a\in\mathfrak{a}_z$. The proposition is proved. □

Proof of necessary part of Theorem 5.1. On the contrary, we assume that

$$bar(2P_+)-4\rho\notin\Xi.$$

Since the Futaki invariant vanishes, we also have

$$bar(2P_+)-4\rho\in\mathfrak{a}^*_{ss}.$$

Let $\{\alpha_{(1)},\ldots,\alpha_{(r)}\}$ be the simple roots in Φ_+. Without loss of generality, we can write

$$bar(2P_+)-4\rho = \lambda_1\alpha_{(1)}+\cdots+\lambda_r\alpha_{(r)},$$

where $\lambda_1\leq 0$. Let $\{\varpi_i\}$ be the fundamental weights for $\{\alpha_{(1)},\ldots,\alpha_{(r)}\}$ such that $\frac{2\langle\varpi_i,\alpha_{(j)}\rangle}{|\alpha_{(j)}|^2}=\delta_{ij}$. Define a W-invariant rational piecewise linear function f on $2P$ by

$$f(v) = \max_{w\in W}\{\langle w\cdot\varpi_1, v\rangle\}.$$

Then

$$f|_{2P_+} = \langle \varpi_1, v \rangle.$$

Note that $\varpi_1 \in \mathfrak{a}^*_{ss}$. However,

$$\mathcal{L}(f) = \frac{1}{2}|\alpha_{(1)}|^2 \lambda_1 \leq 0.$$

This contradicts to Proposition 5.10. Hence (5.1) is true. □

6. Appendix: Examples of Fano G-manifolds

In this appendix, we compute some examples of Fano G-manifolds with a maximal torus subgroup of rank 2. The most examples are from Delcroix' papers [33, 34]. In case of G with a torus subgroup of rank 1, there are only two examples, one is $SL_2(\mathbb{C})$, the other is $PSL_2(\mathbb{C})$. A wonderful Fano compactification of $SL_2(\mathbb{C})$ was described by Delcroix in his thesis (cf. [33]). The corresponding P_+ of $SL_2(\mathbb{C})$ is $[0, 3]$. The Fano compactification of $PSL_2(\mathbb{C})$ is just $\mathbb{CP}^3$ with $P_+ = [0, 2]$. Since they are both homogenous manifolds, there admit Kähler–Einstein metrics on them.

Example 6.1. The wonderful compactification of $\mathrm{PGL}_3(\mathbb{C})$ [34].

In this example, the corresponding roots system is A_2. We denote by α_1 and α_2 the simple roots. The third positive root is then $\alpha_1 + \alpha_2$, and $2\rho = 2(\alpha_1 + \alpha_2)$. For $p = x\alpha_1 + y\alpha_2$,

$$\prod_{\alpha \in \Phi_+} \langle \alpha, p \rangle^2 = (x - y/2)^2(-x/2 + y)^2(x/2 + y/2)^2.$$

The barycenter $bar(P_+)$ is given by

$$bar(P_+) = \frac{24641}{9888}(\alpha_1 + \alpha_2).$$

As a consequence, X_1 admits a Kähler–Einstein metric. Figure 1 gives a representation of P_+, where the cross is the barycenter, and the convex cone delimited by the dashed lines is $2\rho + \Xi$.

Example 6.2. $\mathrm{GL}_2(\mathbb{C})$. There are eight possible Fano compactifications $\mathrm{GL}_2(\mathbb{C})$.

A) There exist Kähler–Einstein metrics on first four compactifications, whose corresponding P_+ are given in [34] as follows.

Now we give the data. Let E_1, E_2 be the generator of $\mathfrak{M}$. We choose a coordinate on $\mathfrak{a}^*$ such that (x, y) is associated to the point $xE_1 + yE_2$. Then the only positive root is

$$\alpha_+ = 2\rho = (1, -1), \quad \text{and} \quad \mathfrak{a}^*_+ = \{x - y > 0\}.$$

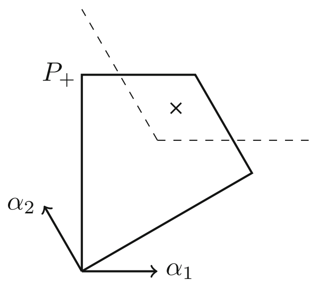


FIGURE 1

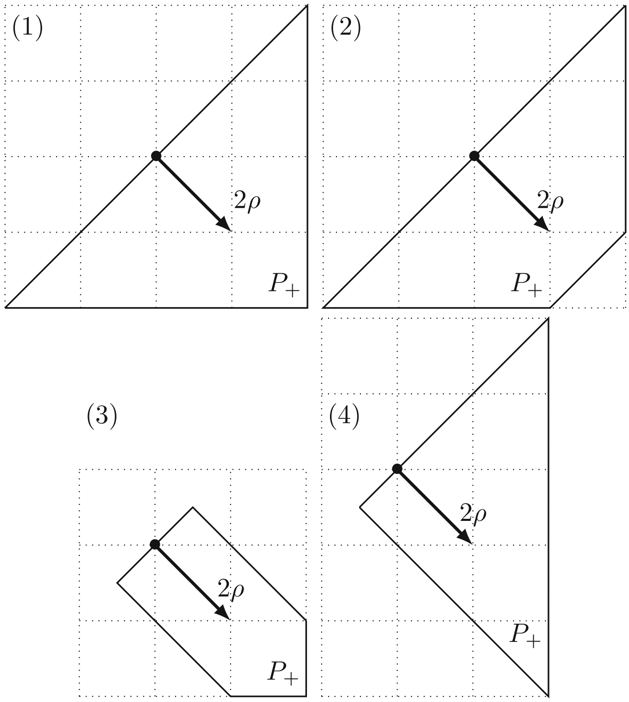


FIGURE 2

Thus

$$2\rho + \Xi = \{(t, -t)|t > 1\}, \quad \text{and} \quad \pi(x, y) = (x - y)^2.$$

Case (1). The polytope

$$P_+ = \{x - y > 0, 2 - x > 0, 2 + y > 0\},$$

and the barycenter

$$bar(P_+) = \left(\frac{6}{5}, -\frac{6}{5}\right).$$

Case (2). The polytope

$$P_+ = \{x - y > 0, 2 - x > 0, 2 + y > 0, 3 - x + y > 0\},$$

and the barycenter

$$bar(P_+) = \left(\frac{36}{35}, -\frac{36}{35}\right).$$

Case (3). The polytope

$$P_+ = \{x - y > 0, 1 + x + y > 0, 2 - x > 0, 2 + y \geq 0, 1 - x - y > 0\},$$

and the barycenter

$$bar(P_+) = \left(\frac{2343}{1750}, -\frac{2343}{1750}\right).$$

Case (4). The polytope

$$P_+ = \{x - y > 0, 2 - x > 0, 1 + x + y > 0\},$$

and the barycenter

$$bar(P_+) = \left(\frac{3}{2}, -\frac{3}{2}\right).$$

B) There exist Kähler–Ricci solitons on last four compactifications, whose corresponding P_+ are given as follows[4].

It is easy to check that each of P_+ above in Figure 3 does not satisfy the condition (5.1). Thus there is no Kähler–Einstein metric on the compactification associated to P_+. On the other hand, a version of Theorem 5.1 for the existence of Kähler–Ricci solitons was also established under a modification of (5.1) in [45] (also see [34]). A numerical computation shows that each of compactifications satisfies this modification condition. Hence, there exists a Kähler–Ricci soliton.

Now we describe the data. Since the soliton vector field lies in $\mathfrak{a}_t = \mathfrak{a} \cap \mathfrak{z}(\mathfrak{g})$, we may assume

$$X = a(1, 1).$$

In the following, we will approximate a by using software Maple.

Case (5). The polytope is

$$P_+ = \{x - y > 0, 2 - x > 0, 2 + y > 0, 1 + x + y > 0\},$$

and by a numerical computation, we have

$$a \in (-0.54596, -0.54595).$$

[4]The last two of them are constructed by Yan Li. I would like to thank him for telling me the result.

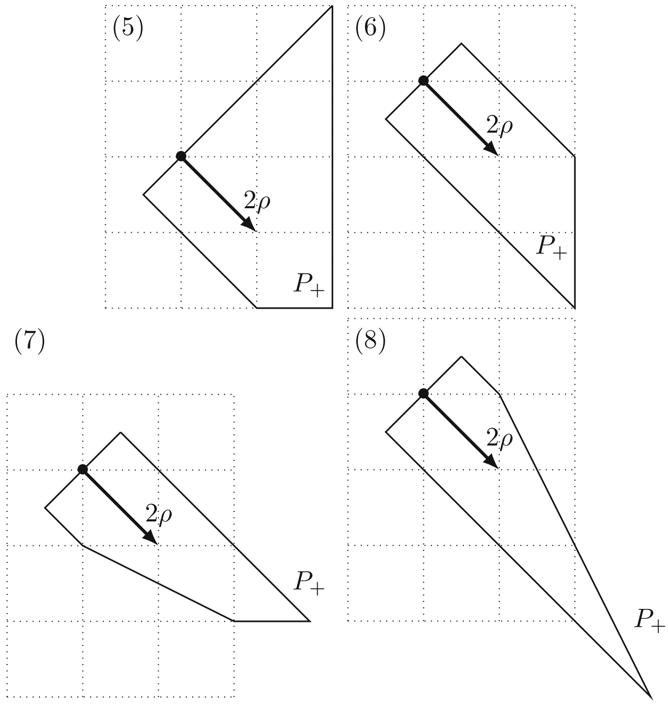


FIGURE 3

Thus the weighted barycenter can be approximated by

$$bar_X(P_+) \cong (1.30041, -1.30041).$$

Case (6). The polytope is

$$P_+ = \{x - y > 0, 2 - x > 0, 1 - x - y > 0, 1 + x + y > 0\},$$

and by a numerical computation, we have

$$a \in (0.1896710, 0.1896712).$$

Thus the weighted barycenter can be approximated by

$$bar_X(P_+) \cong (1.3354, -1.33354).$$

Case (7). The polytope is

$$P_+ = \{x - y > 0, 1 - x - y > 0, 1 + 2y + x > 0, 2 + y > 0, 1 + x + y > 0\},$$

and by a numerical computation, we have

$$a \in (-1.952245, -1.952235).$$

Thus the weighted barycenter can be approximated by

$$bar_X(P_+) \cong (1.35664, -1.35664).$$

Case (8). The polytope is

$$P_+ = \{x - y > 0, 1 + x + y > 0, 2 - 2x - y > 0, 1 - x = y > 0\},$$

and by a numerical computation, we have

$$a \in (2.35616, 2.35618).$$

Thus the weighted barycenter can be approximated by

$$bar_X(P_+) \cong (1.50000, -1.50000).$$

Example 6.3. $SO_4(\mathbb{C})$. There are three possible Fano compactifications of $SO_4(\mathbb{C})$ [34].

A) There is one smooth Fano compactification of $SO_4(\mathbb{C})$, of dimension six, which admits a Kähler–Einstein metric. P_+ is given as in Figure 4.

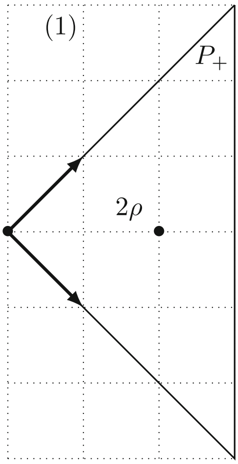


FIGURE 4

Now we describe the data. Choose a coordinate on $\mathfrak{a}^*$ such that the basis are the generator of $\mathfrak{M}$. Then the positive roots are

$$\alpha_1 = (1, -1),\ \alpha_2 = (1, 1), \quad \text{and} \quad 2\rho = (2, 0).$$

Thus

$$\mathfrak{a}_+^* = \{x > y > -x\},$$
$$2\rho + \Xi = \{-2 + x > y > 2 - x\},$$

and

$$\pi(x, y) = (x - y)^2 (x + y)^2.$$

Case (1). The polytope is

$$P_+ = \{y > -x, x > y, 2 - x > 0, 2 + y > 0\}.$$

Thus the barycenter is

$$bar(P_+) = \left(\frac{18}{7}, 0\right).$$

B) There are two smooth Fano compactifications of $\mathrm{SO}_4(\mathbb{C})$ with vanishing Futaki invariant but no Kähler–Einstein metrics. P_+ are given as Figure 5.

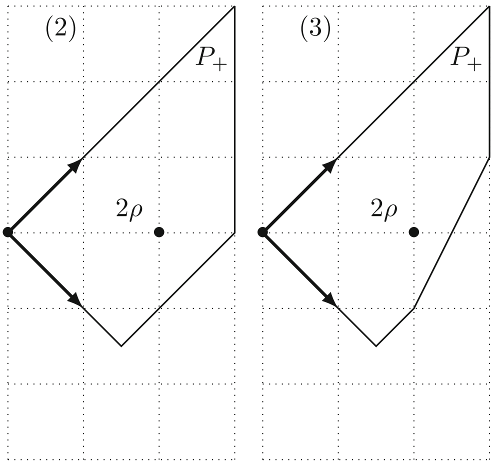


FIGURE 5

Both of P_+ above in Figure 5 do not satisfy (5.1). Moreover, The Futaki invariant vanishes since the center of automorphisms group are finite. Hence there are no Kähler–Ricci solitons on the compactifications associated to P_+ above.

Case (2). The polytope is

$$P_+ = \{y > -x, x > y, 2 - x > 0, 2 + y > 0, 3 - x + y > 0\}.$$

Thus the barycenter is

$$bar(P_+) = \left(\frac{489}{196}, \frac{15}{28}\right).$$

Case (3). The polytope is

$$P_+ = \{y > -x, x > y, 2 - x > 0, 2 + y > 0, 3 - x + y > 0, 5 - 2x + y > 0\}.$$

Thus the barycenter is

$$bar(P_+) = \left(\frac{102741}{43004}, \frac{16575}{23156}\right).$$

It is interesting to study the deformation of Kähler–Ricci flow (3.3) on two examples of Fano G-manifolds in Example 6.3-B as in Subsection 3.1. According to the Hamilton–Tian conjecture, the flow will converge to a Kähler–Ricci soliton with possible singularities of at least complex codimension 2. However, the limit of Kähler–Ricci soliton could not be smooth on such manifolds, otherwise, it is Einstein and still a compactification of $\mathrm{SO}_4(\mathbb{C})$, which keeps the complex structure [44]. But this is impossible by the classification in Examples 6.3! As a consequence, the Ricci flow will produce a singular point of type II on these two G-manifolds. To the best of author's knowledge, these are the first examples of Ricci flow with singularities of type II on compact Kähler manifolds in the literature. There are recent significant progresses by Tian and Zhang, Chen and Wang, and Bamler on the Hamilton–Tian conjecture, respectively. We refer the reader to their papers [11, 18, 77].

References

[1] Aubin, T., *Equations du type Monge–Ampère sur les variétés kählériennes compactes*, Bull. Sci. Math., **102** (1978), 63–95.

[2] Abreu, M., *Kähler geometry of toric varieties and extremal metrics*, Inter. J. Math., **9** (1998), 641–651.

[3] Alexeev, V. and Brion, M., *Stable reductive varieties I: Affine varieties*, Invent. Math., **157** (2004), 227–274.

[4] Alexeev, V. and Brion, M., *Stable reductive varieties II: Projective case*, Adv. Math., **184** (2004), 382–408.

[5] Alexeev, V. and Katzarkov, L., *On K-stability of reductive varieties*, Geom. Funct. Anal., **15** (2005), 297–310.

[6] Azad, H. and Loeb, J., *Plurisubharmonic functions and Kählerian metrics on complexification of symmetric spaces*, Indag. Math. (N.S.), **3(4)**(1992), 365–375.

[7] Berman, R., Boucksom, S., Eyssidieux, P., Guedj, V. and Zeriahi, A., *Kähler–Einstein metrics and the Kähler–Ricci flow on log Fano varieties*, arXiv:1111.7158v3, to appear in J. Reine Angew. Math.,

[8] Bando, S. and Mabuchi, T., *Uniqueness of Einstein Kähler metrics modulo connected group actions*, Algebraic geometry (Sendai, 1985), Adv. Stud. Pure Math., **10** (1987), 11–40.

[9] Batyrev, V.V., *Toric Fano threefolds*, Izv. Akad. Nauk SSSR, (Russian), Ser. Mat., **45** (1981), 704–717.

[10] Batyrev, V.V., *On the classification of toric Fano 4-folds*, J. Math. Sci., **94** (1999), 1021–1050.

[11] Bamler, R., *Convergence of Ricci flows with bounded scalar curvature*, arXiv: 1603.05235.

[12] Batyrev, V.V. and Selivanova, E.N., *Einstein–Kähler metrics on symmetric toric Fano manifolds*, J. Reine Angew. Math., **512** (1999), 225–236.

[13] Berman, R. and Berndtsson, B., *Real Monge–Ampère equations and Kähler–Ricci solitons on toric log Fano varieties*, Ann. Fac. Sci. Toulouse Math., **22** (2013), no. 4, 649–711.

[14] Cao, H.D., *Deformation of Kähler metrics to Kähler–Einstein metrics on compact Kähler manifolds*, Invent. Math., **81** (1985), 359–372.

[15] Calabi, E., *Extremal metrics II*, J. Differ. Geom. and complex analysis (I. Chavel, H.K. Farkas, eds), Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1985, 95–114.

[16] Chern S.S., *On holomorphic mappings of Hermitian manifolds of same dimension*, Proc. Symp. Pure Math., **11** (1968), Amer. Math. Soci., 157–170.

[17] Chen, X and Cheng, J., *On the constant scalar Kähler metrics, general automorphism group*, arXiv:1801.05907.

[18] Chen, X. and Wang, B., *Space of Ricci flows (II)*, arXiv:1405.6797.

[19] Chen, X., Donaldson S., and Sun, S., *Kähler–Einstein metrics on Fano manifolds, I, II, III* , J. Amer. Math. Soc., **28** (2015), 183–197, 199–234, 235–278.

[20] Chen, B., Han, Q., Li, A-M and Sheng L., *Interior estimates for the n-dimensional Abreu's equation*, Adv. Math., **251** (2014), 35–46.

[21] Chen, B., Han, Q., Li, A-M and Sheng L., *Prescribed scalar curvature for homogeneous toric bundles*, Arxiv 1603.01975.

[22] Chen, B., Li, A-M and Sheng L., *Extremal metrics on toric surfaces*, Arxiv 1008.2607.

[23] Chen, X. and Tian, G., *Ricci flow on Kähler–Einstein manifolds*, Duke Math. J., **131** (2006), 17–73.

[24] Cao, H.D., Tian, G., and Zhu, X.H., *Kähler–Ricci solitons on compact Kähler manifolds with $c_1(M) > 0$*, Geom. Anal. and Funct., **15** (2005), 697–719.

[25] Datar, V., Guo, B., Song, J. and Wang, X., *Connecting toric manifolds by conical Kähler–Einstein metrics*, Adv. Math., **323** (2018), 38–83.

[26] Donaldson, S., *Scalar curvature and stability of toric varieties*, Jour. Diff. Geom., **62** (2002), 289–348.

[27] Donaldson, S., *Interior estimates for solutions of Abreu's equation*, Collect. Math., **56** (2005), 103–142.

[28] Donaldson, S., *Kähler metrics with cone singularities along a divisor*, Essays in mathematics and its applications, 49–79, Springer, Heidelberg, 2012.

[29] Donaldson, S., *Extremal metrics on toric surfaces: a continuity method*, J. Differential Geom., **79** (2008), no. 3, 389–432.

[30] Donaldson, S., *Constant scalar curvature metrics on toric surfaces*, Geom. Funct. Anal., **19** (2009), no. 1, 83–136.

[31] Darvas, T. and Rubinstein, Y., *Tian's properness conjectures and Finsler geometry of the space of Kähler metrics*, J. Amer. Math. Soc., **30** (2017), 347–387.

[32] Delzant, T., *Hamiltoniens périodique et image convex de l'application moment*, Bull. Soc. Math. France, **116** (1988), 315–339.

[33] Delcroix, T., *Kähler–Einstein metrics on group compactifications*, Geom. Funct. Anal., **27** (2017), 78–129.

[34] Delcroix, T., *K-Stability of Fano spherical varieties*, arXiv:1608.01852.

[35] Futaki, A., *An obstruction to the existence of Kähler–Einstein metrics*, Invent. Math., **73** (1983), 437–443

[36] Futaki, A., Ono, H. and Wang, G., *Transverse Kähler geometry of Sasaki manifolds and toric Sasaki–Einstein manifolds*, Jour. Diff. Geom. **83** (2009), 585–635.

[37] Guillemin V., *Kähler structures on toric varieties*, Jour. Diff. Geom., **40** (1994), 285–309.

[38] Gilbarg, D. and Trudinger, N.S., *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001, Reprint of the 1998 edition.

[39] Hamilton, R.S., *Three-manifolds with positive Ricci curvature*, J Diff. Geom., **17** (1982), 255–306.

[40] Knapp, A., *Lie Groups beyond an introduction*, Birkhäuser Boston, Inc., Boston, 2002.

[41] [Ko] Kolodziej, S, The complex Monge–Ampère equation and pluripotential theory, Memoirs of the Amer. Math. Society, Number 840, v. 178 (2005), Amer. Math. Society., Providence, Rhode Island.

[42] Li, C., *On the limit behavior of metrics in the continuity method for the Kähler–Einstein problem on a toric Fano manifolds*, Compos. Math., **148** (2012), 1985–2003.

[43] Lu, Y.C., *Holomorphic mappings of complex manifolds*, J. Diff. Geom., **2** (1968), 299–312.

[44] Li, Y., Tian, G. and Zhu, X.H., *Singular limits of Kähler–Ricci flow on Fano G-manifolds*, J. of Func. Analysis, **275** (2018), 1023–1072.

[45] Li, Y., Zhou, B. and Zhu, X.H., *K-energy on polarized compactifications of Lie groups*, arxiv:1701.00306.

[46] Li, Y. and Zhu, X.H., *G-Sasaki manifolds and K-energy*, arXiv:1712.07934.

[47] Li, Y. and Zhou, B., *Mabuchi metrics and properness of the modified Ding functional*, arXiv:1709.03029.

[48] Mabuchi, T., *K-energy maps integrating Futaki invariants*, Tôhoku Math. J., **38** (1986), 575–593.

[49] Mabuchi, T., *Einstein–Kähler forms, Futaki invariants and convex geometry on toric Fano varieties*, Osaka J. Math., **24** (1987), 705–737.

[50] Miguel, De G., *Differentiation of integrals in* $\mathbb{R}^n$, Lectures in Math. **481** (1977), Springer-Verlag.

[51] Nakagawa, Y., *Einstein–Kähler toric Fano manifolds*, Tôhoku Math. J., **45** (1993), 297–310.

[52] Nakagawa, Y., *Classification of Einstein–Kähler toric Fano manifolds*, Tôhoku Math. J., **46** (1994), 125–133.

[53] Oda, T., *Convex bodies and algebraic geometry – an introduction to the theory of toric varieties*, Springer-Verlag, 1988.

[54] Perelman, G., *The entropy formula for the Ricci flow and its geometric applications*, arXiv:0211159.

[55] Podesta, F., Spiro, A., *Kähler–Ricci solitons on homogeneous toric bundles*, J. Reine Angew. Math., **642** (2010), 109–127.

[56] Real, C., *Métriques d'Einstein–Kähler sur des variétés à première classe de Chern positive*, J. Funct. Anal. **106** (1992), 145–188.

[57] Rong, X. and Zhang, Y., *Continuity of extremal transitions and flops for Calabi–Yau manifolds, Appendix B by Mark Gross*, J. Differential Geom. **89** (2011), no. 2, 233–269.

[58] Siu, Y.T., *The existence of Kähler–Einstein metrics on manifolds with positive anti-canonical line bundle and a suitable finite symmetry group*, Ann. Math., **127** (1988), 585–627.

[59] Szekelyhidi, G., *The partial C^0 estimate along the smooth continuity method*, J. Amer. Math. Soc., **29** (2016) 537–560.

[60] Sesum, N. and Tian, G., *Perelman's arguments for uniform bounds on scalar curvature and diameter along the Kähler Ricci flow* (*after Perelman*), J. Inst. Math., Jussieu, **7** (2008), 575–587.

[61] Song, J. and Wang, X.W., *The greatest Ricci lower bound, conical Einstein metrics and the Chern number inequality*, Geom. Topol., **20** (2016), no. 1, 49 C102.

[62] Shi, Y.L. and Zhu, X.H., *An example of singular metric arising from the blow-up limit in the continuity approach to Kähler–Einstein metrics*, J. Pacific Math., **250** (2011), 191–203.

[63] Shi, Y.L. and Zhu, X.H., *Kähler–Ricci solitons on toric Fano orbifolds*, Math. Zeit., **271** (2012), 1241–1251.

[64] Śuss, H., *Kähler–Einstein metrics on symmetric Fano T-varieties*, Adv. Math., **246** (2013), 100–113.

[65] Tian, G., *On Calabi's conjecture for complex surfaces*, Invent. Math., **101** (1990), 101–172.

[66] Tian, G., *On Kähler–Einstein metrics on certain Kähler manifolds with $C_1(M) > 0$*, Invent. Math., **89** (1987), no. 2, 225–246.

[67] Tian, G., *On Calabi's conjecture for complex surfaces*, Invent. Math., **101** (1990), 101–172.

[68] Tian, G., *Kähler–Einstein metrics with positive scalar curvature*, Invent. Math., **130** (1997), 1–37.

[69] Tian, G., *Existence of Einstein Metrics on Fano Manifolds*, in "Jeff Cheeger Anniversary Volume: Metric and Differential Geometry", Progress in Math., **297** (2012), 119–162.

[70] Tian, G., *K-stability and Kähler–Einstein metrics*, Comm. Pure Appl. Math., **68** (2015), 1085–1156.

[71] Tian, G. and Yau, S.T., *Kähler–Einstein metrics on complex surfaces with $C_1 > 0$*, Comm. Math. Phys., **112** (1987), 175–203.

[72] Tian, G. and Zhu, X.H., *Uniqueness of Kähler–Ricci solitons*, Acta Math., **184** (2000), 271–305.

[73] Tian, G. and Zhu, X.H., *A new holomorphic invariant and uniqueness of Kähler–Ricci solitons*, Comm. Math. Helv., **77** (2002), 297–325.

[74] Tian, G. and Zhu, X.H., *Convergence of Kähler–Ricci flow*, Jour. of Amer. Math. Soci., **20** (2007), 675–699.

[75] Tian, G. and Zhu, X.H., *Convergence of Kähler–Ricci flow on Fano manifolds*, J. Reine Angew. Math., **512** (2013), 585–627.

[76] Tian, G., Zhang, S., Zhang, Z., and Zhu, X.H., *Supremum of Perelman's entropy and Kähler–Ricci flow on a Fano manifolds*, Tran. Amer. Math. Soc., **365** (2013), 6669–6695.

[77] Tian, G. and Zhang, Z.L., *Regularity of Kähler–Ricci flows on Fano manifolds*, Acta Math., **216** (2016), 127–176.

[78] Wang Y., *A remark on $C^{2,\alpha}$-regularity of the complex Monge–Ampère equation*, arXiv:1111.0902.

[79] Wang X. and Zhu, X.H., *Kähler–Ricci solitons on toric manifolds with positive first Chern class*, Adv. Math., **188** (2004), 87–103.

[80] Wang, F., Zhou B. and Zhu, X.H., *Modified Futaki invariant and equivariant Riemann–Roch formula*, Adv. Math., **289**(2016), 1205–1235.

[81] Yau, S.T., *On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampère equation, I*, Comm. Pure Appl. Math., **31** (1978), 339–411.

[82] Zhou, B., *The first boundary value problem for Abreu's equation*, Int. Math. Res. Not., **7** (2012), 1439–1484.

[83] Zhou, B., *Variational solutions to extremal metrics on toric surfaces*, Math. Z., **283** (2016), 1011–1031.

[84] Zhou, B. and Zhu, X.H., *Relative K-stability and modified K-energy on toric manifolds*, Adv. Math., **219** (2008), 1327–1362.

[85] Zhou, B. and Zhu, X.H., *K-stability and extremal metrics*, Proc. AMS, **219** (2008), 1327–1362.

[86] Zhou, B. and Zhu, X.H., *Minimizing weak solutions for Calabi's extremal metrics on toric manifolds*, Calc. Var., **32** (2008), 191–217.

[87] Zhu, X.H., *Kähler–Ricci soliton type equations on compact complex manifolds with* $C_1(M) > 0$, J. Geom. Anal., **10** (2000), 759–774.

[88] Zhu, X.H., *Kähler–Ricci flow on a toric manifold with positive first Chern class*, Differential geometry, 323–336, Adv. Lect. Math. (ALM), **22**, Int. Press, Somerville, MA, 2012. arXiv:math/0703486.

[89] Zhu, X.H., *A global mean value inequality for plurisubharmonic functions on a compact Kähler manifold*, Recent developments in geometry and analysis, 495–507, Adv. Lect. Math. (ALM), **23**, Int. Press, Somerville, MA, 2012.

Xiaohua Zhu
School of Mathematical Sciences
Peking University
Beijing 100871, China
e-mail: xhzhu@math.pku.edu.cn

Progress in Mathematics, Vol. 333, 587–616

Some Questions in the Theory of Pseudoholomorphic Curves

Aleksey Zinger

On the occasion of Professor Gang Tian's 60th birthday

Abstract. This survey article, in honor of G. Tian's 60th birthday, is inspired by R. Pandharipande's 2002 note highlighting research directions central to Gromov–Witten theory in algebraic geometry and by G. Tian's complex-geometric perspective on pseudoholomorphic curves that lies behind many important developments in symplectic topology since the early 1990s.

Mathematics Subject Classification (2010). Primary 53D45; Secondary 14N35.

Keywords. Gromov–Witten invariants, pseudoholomorphic curves.

Symplectic topology is an area of geometry originating in and closely associated with classical mechanics. While long established, it has been flourishing especially since the introduction of pseudoholomorphic curves techniques in [41]. These techniques have led to an immense wealth of remarkable applications, mutually enriching interplay with algebraic geometry, and striking connections with string theory. They have in particular given rise to counts of such curves in symplectic manifolds, now known as the *Gromov–Witten invariants.* While many long-standing problems have been spectacularly resolved, new profound questions that could have been hardly imagined in the past have arisen in their place. This article, greatly influenced by G. Tian's perspective on the field, highlights a number of questions concerning pseudoholomorphic curves and their applications in symplectic topology, algebraic geometry, and string theory.

R. Pandharipande's ICM note [75] assembled three conjectures concerning structures in Gromov–Witten theory:

(P1) a Poincaré Duality for the tautological cohomology ring of the Deligne–Mumford moduli space $\overline{\mathcal{M}}_{g,n}$ of stable nodal n-marked genus g curves, known as the *Gorenstein property* of $R^*(\overline{\mathcal{M}}_{g,n})$;

Partially supported by NSF grant DMS 1500875.

(P2) integral counts of holomorphic curves in smooth complex projective threefolds, known as the *BPS states*;

(P3) algebraic restrictions on Gromov–Witten invariants, known as the *Virasoro constraints*.

Each of these conjectures presented a deep quandary requiring fundamentally new ideas to address.

The Gorenstein property is a triviality for $g=0$, since $\overline{\mathcal{M}}_{0,n}$ is a smooth projective variety and $R^*(\overline{\mathcal{M}}_{0,n})=H^*(\overline{\mathcal{M}}_{0,n})$. It is established for $g=1$ in [82] and shown to fail for $g=2$ whenever $n\geq 20$ in [83, 84]. The Virasoro constraints had been established for the Gromov–Witten invariants of manifolds with only even-dimensional cohomology in genus 0, of a point, of a curve, and of the complex projective space $\mathbb{P}^n$ before [75] in [37, 62, 71, 72], respectively, with the last case extended to arbitrary symplectic manifolds with semi-simple quantum cohomology in [92]. However, no geometric rationale behind this conjecture that might confirm it in general has emerged so far, and its testing outside of fairly standard cases in algebraic geometry has been limited by the available computational techniques. Just as (P1), the Virasoro Conjecture of (P3) may yet turn out to fail, at least for non-projective symplectic manifolds.

Unlike (P1) and perhaps (P3), (P2) is most naturally viewed from the symplectic topology perspective in which it splits into three parts. The extensive work on (P2) in algebraic geometry since [75] has not succeeded in confirming this conjecture even in special cases. On the other hand, fundamentally new approaches to the three different parts of (P2) have emerged in symplectic topology which should fully resolve its original formulation in a stronger formulation; see Section 2.

The questions collected in this article fall under four distinct, but related, topics:

(1) the topology of moduli spaces of pseudoholomorphic maps and applications to the mirror symmetry predictions of string theory and to the enumerative geometry of algebraic curves;

(2) integral counts of pseudoholomorphic curves in arbitrary compact symplectic manifolds;

(3) decomposition formulas for counts of pseudoholomorphic curves under "flat" degenerations of symplectic manifolds;

(4) applications of pseudoholomorphic curves techniques in symplectic topology and algebraic geometry.

Each of these topics involves fundamental issues concerning pseudoholomorphic curves and a deep contribution from G. Tian.

1. Topology of moduli spaces

A *symplectic form* on a $2n$-dimensional manifold X is a closed 2-form on X such that ω^n is a volume form on X. A *tame almost complex structure* on a symplectic manifold (X,ω) is a bundle endomorphism

$$J\colon TX \longrightarrow TX \quad \text{s.t.} \quad J^2=-\mathrm{Id}, \quad \omega(v,Jv)>0\ \forall\, v\in T_xX,\ x\in X,\ v\neq 0.$$

If Σ is a (possibly nodal) Riemann surface with complex structure $\mathfrak{j}$, a smooth map $u\colon\Sigma\longrightarrow X$ is called *J-holomorphic* if it solves the *Cauchy–Riemann equation* corresponding to $(J,\mathfrak{j})$:

$$\bar\partial_J u \equiv \frac{1}{2}\big(\mathrm{d}u + J\circ \mathrm{d}u\circ\mathfrak{j}\big)=0. \tag{1.1}$$

The image of such a map in X is called a *J-holomorphic curve.* GW-invariants are rational counts of such curves that depend only on (X,ω). They are generally obtained by counting smooth maps $u\colon\Sigma\longrightarrow X$ that solve locally deformed versions of (1.1) and pass through specified cycles in X as in (2.1) and/or satisfy other cohomological restrictions.

The most fundamental object in GW-theory is the moduli space $\overline{\mathfrak{M}}_{g,k}(A;J)$ of stable k-marked (*geometric*) genus g J-holomorphic maps in the homology class $A\in H_2(X)$. This compact space is generally highly singular. However, as shown in [59], $\overline{\mathfrak{M}}_{g,k}(A;J)$ still determines a rational homology class, called *virtual fundamental class* (*VFC*) and denoted by $[\overline{\mathfrak{M}}_{g,k}(A;J)]^{\mathrm{vir}}$. This class lives in an arbitrarily small neighborhood of $\overline{\mathfrak{M}}_{g,k}(A;J)$ in the naturally stratified configuration space $\mathfrak{X}_{g,k}(A)$ of smooth stable maps introduced in [59] and is independent of J. Integration of cohomology classes against $[\overline{\mathfrak{M}}_{g,k}(A;J)]^{\mathrm{vir}}$ gives rise to GW-invariants; see (2.1). The construction of [59] adapts the deformation-obstruction analysis from the algebro-geometric setting of [58] to symplectic topology via local versions of the inhomogeneous deformations the $\bar\partial_J$-equation introduced in [88, 89] and presents $[\overline{\mathfrak{M}}_{g,k}(A;J)]^{\mathrm{vir}}$ as the homology class of a space stratified by even-codimensional orbifolds. This approach is ideally suited for a range of concrete applications, some of which are indicated below, and can be readily extended via [104] beyond the so-called perfect deformation-obstruction settings. Alternative implementations of the key principles behind [58, 59] later appeared in [10, 11, 26, 77] and other works.

While $\overline{\mathfrak{M}}_{g,k}(A;J)$ is often called a "compactification" of its subspace

$$\mathfrak{M}_{g,k}(A;J)\subset\overline{\mathfrak{M}}_{g,k}(A;J)$$

of maps from smooth domains, $\mathfrak{M}_{g,k}(A;J)$ usually is not dense in $\overline{\mathfrak{M}}_{g,k}(A;J)$. For example,

$$\overline{\mathfrak{M}}_1(\mathbb{P}^n,d)\equiv\overline{\mathfrak{M}}_{1,0}\big(dL;J_{\mathbb{P}^n}\big),$$

where $L\in H_2(\mathbb{P}^n)$ is the standard generator and $J_{\mathbb{P}^n}$ is the standard complex structure on $\mathbb{P}^n$, is a quasi-projective variety over $\mathbb{C}$ containing $\mathfrak{M}_1(\mathbb{P}^n,d)$ as a Zariski open subspace; see [27]. For $m\in\mathbb{Z}^+$ with $m\le n$, the dimension of the Zariski open subspace $\mathfrak{M}_1^m(\mathbb{P}^n,d)$ of $\overline{\mathfrak{M}}_1(\mathbb{P}^n,d)$ consisting of maps u from a smooth

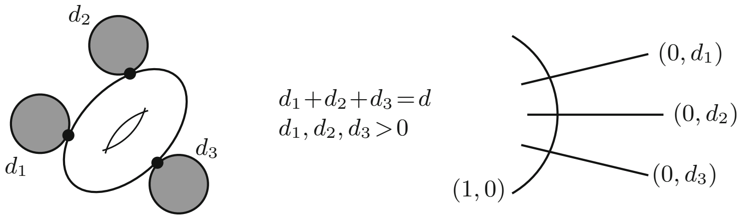


FIGURE 1. The domain of an element of $\mathfrak{M}_1^3(\mathbb{P},d)$ from the points of view of symplectic topology and algebraic geometry, with the first number in each pair in the second diagram denoting the genus of the associated smooth irreducible component and the second number denoting the degree of the restriction of the map to this component.

genus 1 curve Σ_P with m copies of $\mathbb{P}^1$ attached directly to Σ_P so that $u(\Sigma_P)\subset\mathbb{P}^n$ is a point is

$$\dim_{\mathbb{C}}\mathfrak{M}_1^m(\mathbb{P}^n,d)=(n+1)d+n-m\geq(n+1)d=\dim_{\mathbb{C}}\mathfrak{M}_1(\mathbb{P}^n,d);$$

see Figure 1. For example,

$$\mathfrak{M}_1^1(\mathbb{P}^n,d)\approx\mathcal{M}_{1,1}\times\mathfrak{M}_{0,1}(\mathbb{P}^n,d)\,.$$

Thus, $\mathfrak{M}_1(\mathbb{P}^n,d)$ is not dense in $\overline{\mathfrak{M}}_1(\mathbb{P}^n,d)$. This motivates the following deep question concerning the convergence of J-holomorphic maps in the sense of [41].

Question 1 ([88, p. 276]). *Is there a natural Hausdorff space $\overline{\mathfrak{M}}^0_{g,k}(A;J)$ of k marked J-holomorphic maps to X with images of arithmetic genus at least g containing $\mathfrak{M}_{g,k}(A;J)$ as an open subspace so that $\overline{\mathfrak{M}}^0_{g,k}(A;J)$ is compact whenever X is?*

The "natural" requirement in particular includes that

$$\bigsqcup_{\substack{B\in H_2(Y)\\ \iota_*B=A}}\overline{\mathfrak{M}}^0_{g,k}\big(B;J|_Y\big)=\big\{u\in\overline{\mathfrak{M}}^0_{g,k}(A;J)\colon \mathrm{Im}\;u\subset Y\big\}$$

for every inclusion $\iota\colon Y\longrightarrow X$ of an almost complex submanifold and relatedly that $\overline{\mathfrak{M}}^0_{g,k}(A;J)$ determines a fundamental class $[\overline{\mathfrak{M}}^0_{g,k}(A;J)]^{\mathrm{vir}}$. For $g=0$, the usual moduli spaces already have the desired properties and so

$$\overline{\mathfrak{M}}^0_{0,k}(A;J)=\overline{\mathfrak{M}}_{0,k}(A;J)\,.$$

We also note that $\overline{\mathfrak{M}}_{0,k}(\mathbb{P}^n,d)$ is a smooth irreducible quasi-projective variety containing $\mathfrak{M}_{0,k}(\mathbb{P}^n,d)$ as a Zariski dense open subspace and that

$$\overline{\mathfrak{M}}_{0,k}(\mathbb{P}^n,d)-\mathfrak{M}_{0,k}(\mathbb{P}^n,d)\subset\overline{\mathfrak{M}}_{0,k}(\mathbb{P}^n,d)$$

is a normal crossings divisor.

For $g=1$, Question 1 is answered affirmatively in [106, 107] by defining

$$\overline{\mathfrak{M}}^0_{1,k}(A;J)\subset\overline{\mathfrak{M}}_{1,k}(A;J)$$

and showing that $\overline{\mathfrak{M}}^0_{1,k}(A;J)$ determines a fundamental class. In particular, this subspace contains an element u of $\mathfrak{M}^m_{1,k}(A;J)$ if and only if the differentials of the restrictions of u to the m copies of $\mathbb{P}^1$ at the nodes attached to Σ_P span a subspace of $T_{u(\Sigma_P)}X$ of complex dimension less than m. This imposes no condition if $2m > \dim_{\mathbb{R}} X$. If $m \leq n$, this imposes a condition of complex codimension $n+1-m$ on $\mathfrak{M}^m_1(\mathbb{P}^n, d)$ and ensures that

$$\dim_{\mathbb{C}} \left(\overline{\mathfrak{M}}^0_1(\mathbb{P}^n, d) \cap \mathfrak{M}^m_1(\mathbb{P}^n, d)\right) = \dim_{\mathbb{C}} \mathfrak{M}_1(\mathbb{P}^n, d) - 1\,.$$

We also note that $\overline{\mathfrak{M}}^0_{1,k}(\mathbb{P}^n, d)$ is a singular irreducible quasi-projective variety containing $\mathfrak{M}_{1,k}(\mathbb{P}^n, d)$ as a Zariski dense open subspace and that

$$\overline{\mathfrak{M}}^0_{1,k}(\mathbb{P}^n, d) - \mathfrak{M}_{1,k}(\mathbb{P}^n, d) \subset \overline{\mathfrak{M}}^0_{1,k}(\mathbb{P}^n, d)$$

is a divisor. An explicit desingularization $\widetilde{\mathfrak{M}}^0_{1,k}(\mathbb{P}^n, d)$ of this space is constructed in [97] so that

$$\widetilde{\mathfrak{M}}^0_{1,k}(\mathbb{P}^n, d) - \mathfrak{M}_{1,k}(\mathbb{P}^n, d) \subset \widetilde{\mathfrak{M}}^0_{1,k}(\mathbb{P}^n, d)$$

is a normal crossings divisor. The numerical curve-counting invariants obtained by integrating cohomology classes against $[\overline{\mathfrak{M}}^0_{1,k}(A;J)]^{\mathrm{vir}}$ as in (2.1) are called *reduced genus 1 GW-invariants* in [107]. An algebro-geometric approach to these invariants is suggested in [96].

For sufficiently positive symplectic manifolds (X,ω), the *standard* genus 0 and *reduced* genus 1 GW-invariants with insertions pulled back from X only are *integer* counts of J-holomorphic counts in X for a generic ω-compatible almost complex structure J. The standard complex structure $J_{\mathbb{P}^n}$ on $\mathbb{P}^n$ works for these purposes. As demonstrated in [73, 88], the good properties of $\overline{\mathfrak{M}}_{0,k}(\mathbb{P}^n, d)$ indicated above are key to the enumeration of genus 0 curves in $\mathbb{P}^n$ and in particular establish Kontsevich's recursion for counts of such curves. The explicit constructions of $\overline{\mathfrak{M}}^0_{1,k}(A;J)$ and $\widetilde{\mathfrak{M}}^0_{1,k}(\mathbb{P}^n, d)$ in [97, 106] have opened the door for similar applications to the enumerative geometry of genus 1 curves.

For example, the Eguchi–Hori–Xiong recursion for counts of genus 1 curves in $\mathbb{P}^2$ is established in [74] by lifting Getzler's relation [35] from $\overline{\mathcal{M}}_{1,4}$ to $\overline{\mathfrak{M}}_{1,k}(\mathbb{P}^2, d)$ and obtaining a recursion for the genus 1 GW-invariants of $\mathbb{P}^2$; the latter are the same as the corresponding enumerative invariants in this particular case. Getzler's relation can also be lifted to $\overline{\mathfrak{M}}_{1,k}(A;J)$, $\overline{\mathfrak{M}}^0_{1,k}(A;J)$, and $\widetilde{\mathfrak{M}}^0_{1,k}(\mathbb{P}^n, d)$ to yield relations between the genus 0 GW and standard (resp. reduced) genus 1 GW-invariants from the first (resp. second/third) lift. The reduced genus 1 GW-invariants of $\mathbb{P}^n$ are the same as the corresponding enumerative invariants. As shown in [105], the difference between the standard and reduced genus 1 GW-invariants is a combination of the genus 0 GW-invariants; this combination takes a very simple form in complex dimension three. This leads to the following, very concrete question.

Question 2. *Can any of the above three lifts be used to obtain a recursion for the genus* 1 *standard or reduced GW-invariants of* $\mathbb{P}^n$ *for* $n \geq 3$ *and thus a* $\mathbb{P}^n$ *analogue of the Eguchi–Hori–Xiong recursion enumerating genus* 1 *curves?*

For $g=2$, [69] provides the affirmative answer to the main part of Question 1 by defining

$$\overline{\mathfrak{M}}^0_{2,k}(A;J) \subset \overline{\mathfrak{M}}_{2,k}(A;J)$$

and leaves no fundamental difficulty in constructing a fundamental class for this space. The description of this subspace is significantly more complicated than of its $g=1$ analogue. In addition to the simple "level 1" condition appearing in the $g=1$ case, this description involves a more elaborate "level 2" condition which depends on precisely how the "level 1" condition is satisfied relative to the involution and the Weierstrass points on the principal component Σ_P of the domain. While $\overline{\mathfrak{M}}^0_{2,k}(\mathbb{P}^n,d)$ is still a quasi-projective variety, it is no longer irreducible and $\mathfrak{M}_{2,k}(\mathbb{P}^n,d)$ is not dense in $\overline{\mathfrak{M}}^0_{2,k}(\mathbb{P}^n,d)$. However, this is not material for some applications.

While Question 1 concerns a foundational issue in GW-theory (and thus is of interest in itself), a satisfactory answer to this problem is key to relating GW-invariants of a compact symplectic submanifold Y of a compact symplectic manifold (X,ω) given as the zero set of a transverse bundle section to the GW-invariants of the ambient symplectic manifold X. If $\pi_{\mathcal{L}}: \mathcal{L} \longrightarrow X$ is a holomorphic vector bundle and $\iota_{\mathcal{L}}: X \longrightarrow \mathcal{L}$ is the inclusion as the zero section, there is a natural projection map

$$\widetilde{\pi}_{\mathcal{L}}: \mathcal{V}^A_{g,k}(\mathcal{L}) \equiv \overline{\mathfrak{M}}_{g,k}(\iota_{\mathcal{L}*}A;J) \longrightarrow \overline{\mathfrak{M}}_{g,k}(A;J),$$
$$\big[\widetilde{u}\colon \Sigma \longrightarrow \mathcal{L}\big] \longrightarrow \big[\pi_{\mathcal{L}}\circ\widetilde{u}\colon \Sigma\longrightarrow X\big].$$

The fiber of $\widetilde{\pi}_{\mathcal{L}}$ over an element $[u\colon \Sigma\longrightarrow X]$ is $H^0(\Sigma;u^*\mathcal{L})$, the space of holomorphic sections of the holomorphic bundle $u^*\mathcal{L}\longrightarrow\Sigma$. If X and $\mathcal{L}$ are sufficiently positive (such as $\mathbb{P}^n$ and sum of positive line bundles) and $g=0$, $\widetilde{\pi}_{\mathcal{L}}$ is in fact a vector orbi-bundle and

$$\sum_{\substack{B\in H_2(Y)\\ \iota_*B=A}} \iota_*\big[\overline{\mathfrak{M}}_{0,k}(B;J)\big]^{\mathrm{vir}} = e\big(\mathcal{V}^A_{0,k}(\mathcal{L})\big)\cap\big[\overline{\mathfrak{M}}_{0,k}(A;J)\big]^{\mathrm{vir}}. \tag{1.2}$$

This observation in [51], now known as the *Quantum Lefschetz Hyperplane Theorem* for genus 0 GW-invariants, was the starting point for the proofs of the genus 0 mirror symmetry prediction of [7] for the quintic threefold $X_5\subset\mathbb{P}^4$ in [36, 61].

Question 3. *Is there an analogue of the $g=0$ Quantum Lefschetz Hyperplane Theorem* (1.2) *for $g\geq 1$?*

While $\widetilde{\pi}_{\mathcal{L}}$ is not even a vector bundle for $g\geq1$ (even for sufficiently positive X and $\mathcal{L}$), it is shown in [60, 103] that the restriction

$$\widetilde{\pi}_{\mathcal{L}}: \mathcal{V}^A_{1,k}(\mathcal{L})\big|_{\overline{\mathfrak{M}}^0_{1,k}(A;J)} \longrightarrow \overline{\mathfrak{M}}^0_{1,k}(A;J) \tag{1.3}$$

carries a well-defined Euler class, which in turn relates the reduced genus 1 GW-invariants of the submanifold Y and the ambient manifold X:

$$\sum_{\substack{B\in H_2(Y)\\ \iota_*B=A}} \iota_*[\overline{\mathfrak{M}}^0_{1,k}(B;J)]^{\mathrm{vir}} = \mathrm{PD}_{[\overline{\mathfrak{M}}^0_{1,k}(A;J)]} e\big(\mathcal{V}^A_{1,k}(\mathcal{L})\big). \tag{1.4}$$

This is a Quantum Lefschetz Hyperplane Theorem for the reduced genus 1 GW-invariants introduced in [107]. Along with the comparison of the standard and reduced genus 1 GW-invariants established in [105], (1.4) provides a Quantum Lefschetz Hyperplane Theorem for the standard genus 1 GW-invariants. The latter is combined in [108] with the desingularization of the relevant special cases of (1.3) constructed in [97] to confirm the genus 1 mirror symmetry prediction of [2] for X_5 and to obtain similar mirror symmetry formulas for Calabi–Yau hypersurfaces in all projective spaces.

The concrete topological construction of virtual fundamental class in [59] is particularly convenient for the purposes of [103, 106, 107]. It readily handles the moduli spaces $\overline{\mathfrak{M}}^0_{1,k}(A;J)$, which are not virtually smooth, but are virtually stratified by smooth orbifolds of even codimensions. The representation of VFC by a geometric object in [59] also fits well with the comparisons carried out in [60, 103]. However, later variations on [59] of topological flavor, such as [26, 77], should also fit with [60, 103, 105–107].

A satisfactory affirmative answer to Question 1 for each $g\geq 2$, combined with the geometric virtual fundamental class perspective of [59], should readily lead to a Quantum Lefschetz Hyperplane Theorem and to computations of GW-invariants of projective complete intersections in the same genus g. In light of [69], there are no fundamental difficulties left to confirm the genus 2 mirror symmetry predictions of [2] for X_5 and other projective complete intersections by paralleling the genus 1 approach initiated in [106] and completed in [108]. The same approach should also yield confirmations of the mirror symmetry predictions of [99] for the real GW-invariants constructed in [32], after the additional topological subtleties typically arising in the real setting are addressed.

The methods of [69, 106] provide "level 1" and "level 2" obstructions to smoothing J-holomorphic maps from nodal domains and can be used to define natural closed subspaces

$$\overline{\mathfrak{M}}^0_{g,k}(A;J) \subset \overline{\mathfrak{M}}_{g,k}(A;J)$$

for $g\geq 3$, which refine Gromov's Compactness Theorem and determine fundamental classes giving rise to curve-counting invariants of compact symplectic manifolds. However, these sharper compactifications would still not be sufficiently small to exclude all J-holomorphic maps to X with images of arithmetic genus below g, but above 1. The associated reduced GW-invariants would then include lower-genus contributions, even for very positive almost complex structures J. Furthermore, there are indications in [69] that the answer to Question 1 may in fact be negative for an arbitrary almost complex structure J on X if $g>2$ (or perhaps slightly larger) and the dimension of X is sufficiently large.

On the other hand, an affirmative answer to Question 1 in full generality is not needed for specific applications, including to the enumerative geometry of positive-genus curves in the spirit of [73] and to the mirror symmetry predictions in the spirit of [105, 108]. While the complexity of a complete description of $\overline{\mathfrak{M}}^0_{g,k}(A;J)$, whenever it can be defined, would increase rapidly with the genus g, it is likely not to be needed for specific applications either. In particular, it appears feasible to set up a scheme paralleling the genus 1 approach initiated in [106] and completed in [108] that would compute all GW-invariants of X_5 modulo finitely many inputs in each genus g. This could potentially show that the generating functions F_g for these invariants satisfy the holomorphic anomaly equations as predicted in [2], without determining each specific F_g explicitly.

2. BPS states for arbitrary symplectic manifolds

GW-invariants of a symplectic manifold (X,ω) are in general rational numbers arising from families of J-holomorphic curves in X of possibly lower genus and/or "lower" degree (relative to the symplectic deformation equivalence class of ω). The *primary* genus 0 GW-invariants of positive symplectic manifolds (such as smooth Fano varieties) and of symplectic fourfolds arise only from J-holomorphic curves of the same genus and degree, for a generic ω-compatible almost complex structure J on X, and are integer counts of such curves. One might hope that the GW-invariants of (X,ω) in general are expressible in terms of some integer invariants of (X,ω) arising from J-holomorphic curves on X, for J generic at least in some non-empty open subset of such J's. The explicit prediction of [40] relating GW-invariants of Calabi–Yau (or CY) sixfolds (X,ω) to certain conjecturally integer counts could be interpreted in such a way; this prediction has since been extended to a number of other special cases in [49, 75, 76, 99].

For a compact symplectic manifold (X,ω), we denote by $\mathcal{J}_\omega$ the space of ω-compatible almost complex structures on X. For $g,k\in\mathbb{Z}^{\ge 0}$, $A\in H_2(X)$, $J\in\mathcal{J}_\omega$, and $i=1,\ldots,k$, let

$$\mathrm{ev}_i\colon \overline{\mathfrak{M}}_{g,k}(A;J)\longrightarrow X, \quad \mathrm{ev}_i\big([u,z_1,\ldots,z_k]\big)=u(z_i),$$

be the *evaluation map* at the ith marked point. We denote by

$$\begin{aligned} \mathrm{GW}^X_{g,A}\colon \mathcal{H}^*(X)\equiv\bigsqcup_{k=1}^{\infty}H^*(X)^{\oplus k}\longrightarrow\mathbb{Q},\\ \mathrm{GW}^X_{g,A}\big(\mu_1,\ldots,\mu_k\big)=\Big\langle\prod_{i=1}^{k}\mathrm{ev}_i^*\mu_i,\,\big[\overline{\mathfrak{M}}_{g,k}(A;J)\big]^{\mathrm{vir}}\Big\rangle, \end{aligned} \tag{2.1}$$

the *primary genus g degree A GW-invariants* of (X,ω); these multilinear functionals are graded symmetric. The number above vanishes unless

$$\sum_{i=1}^{k}\dim_{\mathbb{R}}\mu_i=\dim\big[\overline{\mathfrak{M}}_{g,k}(A;J)\big]^{\mathrm{vir}}=2\big(\langle c_1(X,\omega),A\rangle+k\big)+\dim_{\mathbb{R}}X-6. \tag{2.2}$$

In general, this number arises from the families of genus g' degree A' J-holomorphic curves in X that pass through generic pseudocycle representatives for the Poincaré duals of $\mu_1,\ldots,\mu_k$ in the sense of [104].

We denote the symplectic deformation equivalence class of a symplectic form ω on a manifold X by $[\omega]$ and let

$$\mathcal{A}\big([\omega]\big)=\big\{(g,A)\in\mathbb{Z}^{\ge0}\times(H_2(X)-\{0\}):\overline{\mathfrak{M}}_g(A;J)\neq\emptyset\ \forall\,J\in\mathcal{J}_{\omega'},\ \omega'\in[\omega]\big\}.$$

The genus g degree A GW-invariants of a compact symplectic manifold (X,ω) depend only on $[\omega]$ and vanish unless $(g,A)\in\mathcal{A}([\omega])$ or $A=0$. In general, they arise from families of connected J-holomorphic curves in X described by decorated graphs, i.e., tuples of the form

$$\Gamma=\big(\mathrm{V},\mathrm{Edg},\mathfrak{g}\colon\mathrm{V}\longrightarrow\mathbb{Z}^{\ge0},\mathfrak{d}\colon\mathrm{V}\longrightarrow H_2(X)-\{0\}\big). \tag{2.3}$$

In such a tuple, $\mathrm{V}(\Gamma)\equiv\mathrm{V}$ and $\mathrm{Edg}(\Gamma)\equiv\mathrm{Edg}$ are finite collections of *vertices* and *edges*, respectively; the latter are pairs of vertices, but of not necessarily distinct ones, and some pairs may appear multiple times in the collection Edg. The vertices and the edges index the irreducible components $\mathcal{C}_v$ of the curves and the nodes between them, respectively. The values of the maps $\mathfrak{g}$ and $\mathfrak{d}$ at $v\in\mathrm{V}$ specify the geometric genus of $\mathcal{C}_v$ and its degree, respectively. For a tuple as in (2.3), we define

$$g(\Gamma)=1+|\mathrm{Edg}|-|\mathrm{V}|+\sum_{v\in\mathrm{V}}\mathfrak{g}(v),\quad \mathfrak{g}_v(\Gamma)=\mathfrak{g}(v),\ \mathfrak{d}_v(\Gamma)=\mathfrak{d}(v)\ \forall\,v\in\mathrm{V}\,.$$

We denote by $\mathcal{P}([\omega])$ the collection of connected decorated graphs Γ as in (2.3) such that $(\mathfrak{g}(v),\mathfrak{d}(v))$ is an element of $\mathcal{A}([\omega])$ for every $v\in\mathrm{V}$.

For $(g,A)\in\mathcal{A}([\omega])$, let $\Gamma_0(g,A)$ be the unique connected edgeless graph with

$$\mathfrak{g}_v\big(\Gamma_0(g,A)\big)=g\qquad\text{and}\qquad\mathfrak{d}_v\big(\Gamma_0(g,A)\big)=A$$

for the unique vertex v. Define

$$\widetilde{\mathcal{P}}_{g,A}\big([\omega]\big)=\Big\{(\Gamma,\mathfrak{m})\colon\Gamma\in\mathcal{P}([\omega]),\ g(\Gamma)\le g,\ |\mathrm{Edg}(\Gamma)|\le(n-3)(g-g(\Gamma)),\\ \mathfrak{m}\in(\mathbb{Z}^+)^{\mathrm{V}(\Gamma)},\ \sum_{v\in\mathrm{V}}\mathfrak{m}_v\mathfrak{d}_v(\Gamma)=A\Big\},$$

where $2n\equiv\dim_{\mathbb{R}}X$. By Gromov's Compactness Theorem [41], this collection is finite for every $(g,A)\in\mathcal{A}(\omega)$. Let

$$\widetilde{\mathcal{P}}^{\star}_{g,A}\big([\omega]\big)\subset\widetilde{\mathcal{P}}_{g,A}\big([\omega]\big)$$

be the complement of the pair $(\Gamma_0(g,A),1)$.

For a graded symmetric multilinear functional

$$\mathrm{E}\colon\mathcal{H}^*(X)\longrightarrow\mathbb{Q}$$

and $\mu\in H^*(X)^{\oplus k_0}$, we denote by $\mathrm{E}(\mu,\cdot)$ the graded symmetric multilinear functional obtained by inserting additional k inputs after the k_0 inputs μ. For $m\in\mathbb{Z}^+$, define

$$\langle\mathrm{E}\rangle_m\colon\mathcal{H}^*(X)\longrightarrow\mathbb{Q},\qquad \langle\mathrm{E}\rangle_m(\mu)=m^k\mathrm{E}(\mu)\ \forall\,\mu\in H^k(X),\ k\in\mathbb{Z}^{\ge 0}\,.$$

For graded symmetric multilinear functionals $\mathrm{E}_1,\ldots,\mathrm{E}_r$ as above, let

$$\prod(\mathrm{E}_1,\ldots,\mathrm{E}_r)\colon\mathcal{H}^*(X)\longrightarrow\mathbb{Q}$$

be the graded symmetric multilinear functional obtained by distributing the k inputs between the r functionals $\mathrm{E}_1,\ldots,\mathrm{E}_r$, multiplying their outputs, and summing over all possible distributions with the appropriate signs depending on the degrees of the inputs.

For a symplectic form ω on X and $A,A^*\in H_2(X)$, we define $A\le_\omega A^*$ if $\omega'(A)\le\omega'(A^*)$ for some $\omega'\in[\omega]$. For the purposes of the question below, we identify the vertices V of each graph as in (2.3) with the set $\{1,\ldots,|\mathrm{V}|\}$.

Question 4. *Let (X,ω) be a compact symplectic manifold. Are there a collection*

$$C^{(g)}_{\mathfrak{g},\mathfrak{m}}\in\mathbb{Q}\qquad g\in\mathbb{Z}^{\ge0},\ (\mathfrak{g},\mathfrak{m})\in(\mathbb{Z}^{\ge0})^r\times(\mathbb{Z}^+)^r,\ r\in\mathbb{Z}^+,$$

of rational numbers and collections

$$\begin{aligned}\mathrm{E}^X_{\Gamma,\mathfrak{m}}\colon\mathcal{H}^*(X)&\longrightarrow\mathbb{Z}, &\qquad \Gamma\in\mathcal{P}\big([\omega]\big),\ \mathfrak{m}\in(\mathbb{Z}^+)^{\mathrm{V}(\Gamma)},\\ \mathrm{E}^X_{g,A}\colon\mathcal{H}^*(X)&\longrightarrow\mathbb{Z}, &\qquad (g,A)\in\mathcal{A}([\omega]),\end{aligned}$$

of graded symmetric multilinear functionals that depend only on $[\omega]$ and satisfy the following properties?

(E1) *for every $(g,A)\in\mathcal{A}([\omega])$,*

$$\mathrm{GW}^X_{g,A}=\mathrm{E}^X_{g,A}+\sum_{(\Gamma,\mathfrak{m})\in\widetilde{\mathcal{P}}^\star_{g,A}([\omega])}\mathrm{E}^X_{\Gamma,\mathfrak{m}}\,;\tag{2.4}$$

(E2) *for every $\Gamma\in\mathcal{P}([\omega])$ as in (2.3), there exist $N(\Gamma)\in\mathbb{Z}^{\ge0}$ and $\mu_{r;v}\in\mathcal{H}^*(X)$ with $r=1,\ldots,N(\Gamma)$ and $v\in\mathrm{V}$ such that*

$$\mathrm{E}^X_{\Gamma,\mathfrak{m}}=C^{(g)}_{\mathfrak{g}(\Gamma),\mathfrak{m}}\sum_{r=1}^{N(\Gamma)}\prod\Big(\big(\langle\mathrm{E}^X_{\mathfrak{g}(v),\mathfrak{d}(v)}\rangle_{\mathfrak{m}_v}(\mu_{r;v},\cdot)\big)_{v\in\mathrm{V}}\Big)\quad\forall\,\mathfrak{m}\in(\mathbb{Z}^+)^{\mathrm{V}(\Gamma)};\tag{2.5}$$

(E3) *for every $A\in H_2(X)$ with $\omega'(A)>0$ for all $\omega'\in[\omega]$,*

$$\sup\big\{g\in\mathbb{Z}^{\ge0}\colon\mathrm{E}^X_{g,A}\neq0\big\}<\infty;\tag{2.6}$$

(E4) *for all $g^*\in\mathbb{Z}^{\ge0}$ and $A^*\in H_2(X)$ there exists a subset $\mathcal{J}^{\mathrm{reg}}_\omega\subset\mathcal{J}_\omega$ of second category in a nonempty open subset of $\mathcal{J}_\omega$ so that for all $(g,A)\in\mathcal{A}([\omega])$ with $g\le g^*$ and $A\le_\omega A^*$, $J\in\mathcal{J}^{\mathrm{reg}}_\omega$, and $\mu_1,\ldots,\mu_k\in H^*(X)$ satisfying (2.2), there exist pseudocycle representatives f_i for the Poincaré duals of μ_i such that*
- *the set of genus g degree A J-holomorphic curves meeting the pseudocycles $f_1,\ldots,f_k$ is cut out transversely and thus is finite,*
- *the signed cardinality of this set is $\mathrm{E}^X_{g,A}(\mu_1,\ldots,\mu_k)$.*

For all $n\in\mathbb{Z}^{\ge0}$ and $A\in H_2(\mathbb{P}^n)$, there exists $g_A\in\mathbb{Z}^+$ so that every degree A $J_{\mathbb{P}^n}$-holomorphic map $u\colon\Sigma\longrightarrow\mathbb{P}^n$ from a smooth closed connected genus $g\ge g_A$ Riemann surface is a branched cover of a line $\mathbb{P}^1\subset\mathbb{P}^n$; this is a special of the classical Castelnuovo bound [38, p. 252]. In light of (E4), (E3) is an analogue of this bound for J-holomorphic curves in arbitrary symplectic manifolds.

For symplectic fourfolds, i.e., $n=2$ in the definition of the collection $\widetilde{\mathcal{P}}_{g,A}([\omega])$, (2.4) and (2.5) reduce to $\mathrm{GW}^X_{g,A}=\mathrm{E}^X_{g,A}$; (E4) is well known to hold in this case. For symplectic sixfolds, i.e., $n=3$, $(g,A)\notin\mathcal{A}([\omega])$ unless

$$\big\langle c_1(X,\omega),A\big\rangle=0 \qquad\text{or}\qquad \big\langle c_1(X,\omega),A\big\rangle>0\,. \tag{2.7}$$

In both cases, only edgeless connected graphs appear in (2.4). Precise predictions for the structure of (2.4) and (2.5) for symplectic sixfolds involve the coefficients $C_{h,A}(g)\in\mathbb{Q}$ specified by

$$\sum_{g=0}^{\infty}C_{h,A}(g)t^{2g}=\bigg(\frac{\sin(t/2)}{t/2}\bigg)^{2h-2+\langle c_1(X,\omega),A\rangle}. \tag{2.8}$$

In the second, *Fano*, case of (2.7), (2.4) and (2.5) were predicted in [75] to reduce to

$$\mathrm{GW}^X_{g,A}(\mu)=\sum_{h=0}^{g}C_{h,A}(g-h)\mathrm{E}^X_{h,A}(\mu)\qquad\forall\,\mu\in\mathcal{H}^*(X). \tag{2.9}$$

In the $g=0,1$ cases, this becomes

$$\mathrm{GW}^X_{0,A}(\mu)=\mathrm{E}^X_{0,A}(\mu),\;\;\mathrm{GW}^X_{1,A}(\mu)=\mathrm{E}^X_{1,A}(\mu)+\frac{2-\langle c_1(X,\omega),A\rangle}{24}\mathrm{E}^X_{0,A}(\mu), \tag{2.10}$$

respectively.

The first equation in (2.10) with $\mathrm{E}^X_{0,A}(\mu)$ described by (E4) is the original *definition* of $\mathrm{GW}^X_{0,A}(\mu)$ for Fano classes A in the basic case of the semi-positive symplectic manifolds (which include all symplectic sixfolds). The second equation in (2.10) holds with $\mathrm{E}^X_{1,A}(\mu)$ replaced by the reduced genus 1 GW-invariants $\mathrm{GW}^{X;0}_{1,A}(\mu)$ constructed in [107], which satisfy the first bullet in (E4) whenever (X,ω) is semi-positive; see [107, Theorem 1.1] and [106, Section 1.3], respectively. The existence of a subspace $\mathcal{J}^{\mathrm{reg}}_\omega\subset\mathcal{J}_\omega$ of second category satisfying (E4) for the Fano classes A on symplectic sixfolds is established in [109]. Since the system of equations (2.9) with all such classes A is invertible and the GW-invariants depend only on $[\omega]$, this implies that the resulting counts $\mathrm{E}^X_{h,A}(\mu)$ depend only on $[\omega]$ and thus affirmatively answers Question 4 with the exception of (E3) in the Fano case of (2.7).

The first, *CY*, case of (2.7) is much harder because degree $m\ge2$ covers of genus h degree A/m J-holomorphic curves $\mathcal{C}\subset X$ contribute to the genus g degree A GW-invariants of (X,ω). For $d\in\mathbb{Z}^+$, we denote by $\mathcal{P}(d)$ the set of partitions of d into positive integers $d_1\ge\cdots\ge d_k$. Each such partition ρ corresponds to a *Ferrers diagram*, i.e., a collection of boxes indexed by the set

$$S(\rho)=\big\{(i,j)\colon i\in1,\ldots,k,\;j\in1,\ldots,d_i\big\},$$

and to a *dual partition* $\rho'\equiv(d'_1\geq\cdots\geq d'_{k'})$ of d specified by

$$k'=d_1,\qquad d'_j=\max\big\{i=1,\ldots,k\colon d_i\geq j\big\}.$$

The *hooklength* of a box $(i,j)\in S(\rho)$ is defined to be

$$\ell_{ij}(\rho)=d_i+d_j-i-j+1\in\mathbb{Z}^+\,.$$

The degree d contribution $n^{(h)}_{h',d}\in\mathbb{Z}^+$ of a genus h curve to the genus h' curve count was predicted in [6] to be given by

$$\begin{split}\exp\Bigg(\sum_{d=1}^{\infty}\sum_{h'=h}^{\infty}n^{(h)}_{h',d}\bigg(\sum_{m=1}^{\infty}\frac{q^{md}}{m}\big(2\sin(mt/2)\big)^{2h'-2}\bigg)\Bigg)\\ =1+\sum_{d=1}^{\infty}q^d\bigg(\sum_{\rho\in\mathcal{P}(d)}\prod_{(i,j)\in S(\rho)}\Big(2\sin\big(\ell_{ij}(\rho)t/2\big)\Big)^{2h-2}\bigg).\end{split}\tag{2.11}$$

We note that

$$\exp\bigg(\sum_{m=1}^{\infty}\frac{q^m}{m}\big(2\sin(mt/2)\big)^{-2}\bigg)=1+\sum_{d=1}^{\infty}q^d\bigg(\sum_{\rho\in\mathcal{P}(d)}\prod_{(i,j)\in S(\rho)}\Big(2\sin\big(\ell_{ij}(\rho)t/2\big)\Big)^{-2}\bigg),$$

$$\sum_{d=1}^{\infty}\bigg(\sum_{m=1}^{\infty}\frac{q^{md}}{m}\bigg)=-\sum_{d=1}^{\infty}\ln\big(1-q^d\big)=\ln\bigg(\prod_{d=1}^{\infty}\big(1-q^d\big)^{-1}\bigg)=\ln\bigg(1+\sum_{d=1}^{\infty}q^d\big|\mathcal{P}(d)\big|\bigg);$$

the first identity above is the $t_1=t$, $t_2=t^{-1}$ case of [68, (4.5)]. Combining these two identities with the $h=0,1$ cases of (2.11), we obtain

$$n^{(0)}_{h',d}=\begin{cases}1, & \text{if } (h',d)=(0,1);\\ 0, & \text{otherwise;}\end{cases}\qquad n^{(1)}_{h',d}=\begin{cases}1, & \text{if } h'=1;\\ 0, & \text{otherwise.}\end{cases}\tag{2.12}$$

However, $n^{(h)}_{h',d}$ is generally nonzero for $h\geq 2$, $d\in\mathbb{Z}^+$, and some $h'>h$.

The primary GW-invariants (2.1) in the CY classes A are encoded by the rational numbers $N^X_{g,A}\equiv\mathrm{GW}_{g,A}()$, i.e., the GW-invariants with no insertions. In this case, (2.4) and (2.5) were predicted in [40] to reduce to

$$N^X_{g,A}=\sum_{\substack{m\in\mathbb{Z}^+\\ A/m\in\mathcal{A}([\omega])}}m^{2g-3}\sum_{h=0}^{g}\bigg(\sum_{\substack{d\in\mathbb{Z}^+\\ m/d\in\mathbb{Z}}}d^{3-2g}\sum_{h'=h}^{g}C_{h',0}(g-h')n^{(h)}_{h',d}\bigg)n^X_{h,A/m}\,,\tag{2.13}$$

where $n^X_{g,A}\equiv\mathrm{E}_{g,A}()$. For $m\in\mathbb{Z}^+$, we denote by $\langle m\rangle$ the sum of the positive divisors of m. By (2.12), the $g=0,1$ cases of (2.13) become

$$N^X_{0,A}=\sum_{\substack{m\in\mathbb{Z}^+\\ A/m\in\mathcal{A}([\omega])}}m^{-3}n^X_{0,A/m},\quad N^X_{1,A}=\sum_{\substack{m\in\mathbb{Z}^+\\ A/m\in\mathcal{A}([\omega])}}m^{-1}\bigg(\langle m\rangle n^X_{1,A/m}+\frac{1}{12}n^X_{0,A/m}\bigg),\tag{2.14}$$

respectively.

The system of equations (2.13) with all CY classes A on a symplectic six-fold (X,ω) is also invertible. Thus, it determines the numbers $n^X_{g,A}\in\mathbb{Q}$ from the number $N_{g,A}^X$. The original version of Question 4, known as the *Gopakumar–Vafa Conjecture* for projective CY threefolds, in fact predicted *only* the integrality of the numbers $n_{g,A}^X$ obtained in this way and the existence of a Castelnuovo-type bound for them. However, (E4) has been generally believed to be the underlying reason for the validity of this conjecture since its appearance in the late 1990s. Until [48], (E4) had also been central to every claim, including by the authors of [48] in the early 2000s, to establish the integrality part of this conjecture; all of these claims had quickly turned out to be erroneous.

A fundamentally new perspective on the integrality part of the Gopakumar–Vafa Conjecture for symplectic sixfolds is introduced in [48]. It completely bypasses the analytic step (E4) and appears to succeed in establishing the integrality of the numbers $n_{g,A}^X$ arising from (2.13) via local arguments that are generally topological in spirit. The existence of a subset $\mathcal{J}_\omega^{\mathrm{reg}} \subset \mathcal{J}_\omega$ of second category satisfying the first bullet in (E4) for symplectic CY sixfolds is treated in [101] following the general approach to this transversality issue in [13], but with additional technical input. However, it still remains to establish that the resulting counts of J-holomorphic curves satisfy the second bullet in (E4). Taking a geometric analysis perspective previously unexplored in GW-theory, [12] uses [86], which established an analogue of Gromov's Convergence Theorem for J-holomorphic maps without an a priori genus bound, to reduce the Castelnuovo-type bound (E3) for symplectic CY sixfolds to the existence of $J \in \mathcal{J}_\omega$ satisfying (E4).

Precise predictions for the structure of (2.4) and (2.5) have also been made in some cases for symplectic manifolds of real dimensions $2n \geq 8$. The genus 0 prediction for symplectic CY manifolds is a direct generalization of the first equation in (2.14) and is given by

$$\mathrm{GW}_{0,A}^X(\mu_1,\ldots,\mu_k) = \sum_{\substack{m\in\mathbb{Z}^+\\ A/m\in\mathcal{A}([\omega])}} m^{k-3}\mathrm{E}_{0,A/m}^X(\mu_1,\ldots,\mu_k) \tag{2.15}$$

for all $\mu_1,\ldots,\mu_k \in H^*(X)$; see [49, (2)]. The genus 1 predictions for symplectic CY manifolds of real dimensions 8 and 10 appear in [49] and [76], respectively. In contrast to the arbitrary genus GW-invariants of symplectic sixfolds in (2.9) and (2.13) and to the genus 0 GW-invariants of symplectic CY manifolds in (2.15), the genus 1 GW-invariants of symplectic CY manifolds (X,ω) of real dimensions $2n \geq 8$ include contributions from families of J-holomorphic curves in (X,ω) of positive dimensions ($2(n{-}3)$-dimensional families of genus 0 curves). This makes the analogues of (2.9), (2.13), and (2.15) in the last case significantly more complicated. All curves appearing in the relevant families of J-holomorphic curves are reduced in the sense of algebraic geometric geometry and have simple nodes if $n = 4, 5$. As noted in the last paragraphs of [76, Sections 1.2,2.2], non-reduced curves and curves with non-simple nodes appear in such families if $n \geq 6$. In order to obtain a precise prediction for the structure of (2.4) and (2.5) for the genus 1 GW-invariants of symplectic CY manifolds of real dimensions $2n \geq 12$, contributions from such curves to the genus 0 and genus 1 GW-invariants still need to be determined.

Question 4 readily extends to the real GW-invariants $\mathrm{GW}^{\phi}_{g,A}$ of compact real symplectic manifolds (X,ω,ϕ), whenever these invariants are defined. For example, the genus 0 real GW-invariants of real symplectic fourfolds constructed in [100] are just signed counts of J-holomorphic curves. So, the real analogues of (2.4) and (2.5) in this case also reduce to $\mathrm{GW}^{\phi}_{0,A}=\mathrm{E}^{\phi}_{0,A}$. Arbitrary genus real GW-invariants are constructed in [32] for many real symplectic manifolds, including the odd-dimensional projective spaces $\mathbb{P}^{2n-1}$ and quintic threefolds $X_5\subset\mathbb{P}^4$ cut out by real equations. It is established in [70] that the analogue of (2.9) for the Fano classes A on a real symplectic sixfold (X,ω,ϕ) is

$$\mathrm{GW}^{\phi}_{g,A}(\mu)=\sum_{\substack{0\le h\le g\\ g-h\in2\mathbb{Z}}}\widetilde{C}_{h,A}\big(\tfrac{g-h}{2}\big)\mathrm{E}^{\phi}_{h,A}(\mu)\qquad\forall\ \mu\in\mathcal{H}^*(X),\tag{2.16}$$

with the coefficients $\widetilde{C}_{h,A}(g)\in\mathbb{Q}$ defined by

$$\sum_{g=0}^{\infty}\widetilde{C}_{h,A}(g)t^{2g}=\bigg(\frac{\sinh(t/2)}{t/2}\bigg)^{h-1+\langle c_1(X,\omega),A\rangle/2}.\tag{2.17}$$

The invariants $\mathrm{E}^{\phi}_{h,A}(\mu)$ appearing in (2.16) are signed counts of real genus g degree A J-holomorphic curves $\mathcal{C}\subset X$.

The real Fano threefold case treated in [70] and [99, (5.41)] suggest that the real analogue of (2.13) should be

$$N^{\phi}_{g,A}=\sum_{\substack{m\in\mathbb{Z}^+-2\mathbb{Z}\\ A/m\in\mathcal{A}([\omega])}}m^{g-2}\sum_{\substack{0\le h\le g\\ g-h\in2\mathbb{Z}}}\bigg(\sum_{\substack{d\in\mathbb{Z}^+\\ m/d\in\mathbb{Z}}}d^{2-g}\sum_{\substack{h\le h'\le g\\ g-h'\in2\mathbb{Z}}}\widetilde{C}_{h',0}\big(\tfrac{g-h'}{2}\big)\widetilde{n}^{(h)}_{h',d}\bigg)n^{\phi}_{h,A/m},\tag{2.18}$$

for some $\widetilde{n}^{(h)}_{h',d}\in\mathbb{Z}$ (only the d odd cases matter). The right-hand sides of [99, (5.10),(5.28)] suggest that

$$\widetilde{n}^{(0)}_{h',d}=\begin{cases}1, & \text{if } (h',d)=(0,1);\\ 0, & \text{otherwise;}\end{cases}\qquad \widetilde{n}^{(1)}_{h',d}=\begin{cases}1, & \text{if } h'=1;\\ 0, & \text{otherwise.}\end{cases}$$

This would reduce the $g=0,1$ cases of (2.18) to

$$N^{\phi}_{0,A}=\sum_{\substack{m\in\mathbb{Z}^+-2\mathbb{Z}\\ A/m\in\mathcal{A}([\omega])}}m^{-2}n^{\phi}_{0,A/m},\qquad N^{\phi}_{1,A}=\sum_{\substack{m\in\mathbb{Z}^+-2\mathbb{Z}\\ A/m\in\mathcal{A}([\omega])}}m^{-1}\langle m\rangle n^{\phi}_{1,A/m}.$$

The numbers $\widetilde{n}^{(h)}_{h',d}$ should arise from a real analogue of (2.11), with the exponent on the left-hand side combining the real curve counts $\widetilde{n}^{(h)}_{h',d}$ and the complex curve counts $n^{(h)}_{h',d}$ to account for the real doublets of [33, Theorem 1.3]. The three theorems of [33, Section 1] should provide the necessary geometric input to adapt the approach of [6] for (2.13) to the real setting; related equivariant localization data is provided by [34, Section 4.2]. An analogue of (2.11) for the real setting has been obtained in [29].

The approach of [48] to the integrality of the numbers $n_{g,A}^X$ determined by (2.13) should be adaptable to other situations when the GW-invariants in question are expected to arise entirely from isolated J-holomorphic curves. These situations include the real genus 0 GW-invariants of many real symplectic manifolds and the real arbitrary genus GW-invariants of real symplectic CY sixfolds constructed in [28] and [32], respectively. In fact, the integrality of the numbers $\mathrm{E}_{0,A}^X(\mu)$ determined by (2.15) is already a (secondary) subject of [48]. On the other hand, the approach of [48] does not appear readily adaptable to situations when positive-dimensional families of J-holomorphic curves in X are expected to contribute to the GW-invariants in question. These situations include the genus 1 GW-invariants of symplectic CY manifolds of real dimensions 8 and 10 studied in [49] and [76], respectively. The approaches of [101] and [12] to the existence of a subset $\mathcal{J}_\omega^{\mathrm{reg}} \subset \mathcal{J}_\omega$ satisfying (E4) and to the Castelnuovo-type bound for the associated counts of J-holomorphic curves, respectively, appear more flexible in this regard.

Enumerative geometry of curves in projective varieties is a classical subject originating in the middle of the nineteenth century. However, the developments in this field had been limited to very low degrees until the emergence of GW-theory and its applications to enumerative geometry in the early 1990s. As the moduli spaces $\overline{\mathfrak{M}}_{g,k}(A;J)$ have fairly nice deformation-obstruction theory, the GW-invariants arising from these spaces are often amendable to computations. Whenever these invariants can be related to enumerative curve counts as in Question 4, computations of GW-invariants translate into direct applications to enumerative geometry. The most famous such application is perhaps Kontsevich's recursion for counts of genus 0 curves in $\mathbb{CP}^2$, stated in [52] and proved in [88]. Analogues of this recursion for counts of real genus 0 curves in $\mathbb{P}^2$ defined in [100] and in $\mathbb{P}^{2n-1}$ defined in [28] appear in [91] and [30, 31], respectively. The counts of genus g degree d curves arising from the proofs of the mirror symmetry predictions for the projective CY complete intersections in genus 0 in [36, 61] and in genus 1 in [85, 108] via (2.14) have been shown to match the classical enumerative counts for $g=0$, $d\le 3$ and for $g=1$, $d\le 4$; see [14]. The genus 0 real GW-invariants of real symplectic fourfolds defined in [100] and of many higher-dimensional real symplectic manifolds defined in [28] directly provide lower bounds for counts of genus 0 real curves; the arbitrary genus real GW-invariants defined in [32] provide such bounds in arbitrary genera via the relation (2.16) proved in [70]. For local CY manifolds, Question 4 points to intriguing number-theoretic properties of GW-invariants; see G. Martin's conjecture in [76, Section 3.2].

The coefficients $\widetilde{C}_{0,0}(g)$ in (2.17) are the coefficients of the renown A-series central to the Index Theorem [54, Theorem 3.13]; they in particular determine the index of the Dirac operator on a Spin bundle. The coefficients in (2.8) are closely related to the A-series as well. It is tempting to wonder if there is some connection between the multiply covered contributions encoded by (2.8) and by (2.17) and Dirac operators.

3. Symplectic degenerations and Gromov–Witten invariants

It is natural and essential to study the behavior of GW-invariants under reasonable degenerations and decompositions of symplectic manifolds, as pointed out in [93]. The standard example of such a decomposition is provided by the *symplectic sum construction* of [39]; it joins two symplectic manifolds X_1 and X_2 along a common *smooth symplectic divisor* X_{12} (i.e., a closed symplectic submanifold of real codimension 2) with dual normal bundles in the two manifolds into a symplectic manifold $X_1\#_{X_{12}}X_2$. In fact, the symplectic sum construction of [39] provides a symplectic fibration $\pi\colon \mathcal{Z}\longrightarrow\Delta$ over the unit disk $\Delta\subset\mathbb{C}$, whose central fiber $\mathcal{Z}_0$ is $X_1\cup_{X_{12}}X_2$ and the remaining fibers are smooth symplectic manifolds which are symplectically deformation equivalent to each other; see the first diagram in Figure 2. While the behavior of GW-invariants under the basic degenerations and decompositions associated with the construction of [39] was understood long ago and has since been followed up by numerous applications throughout GW-theory, the progress beyond these cases has been slow. The interest in finding usable decomposition formulas for GW-invariants in more general situations has grown considerably since the advent of the *Gross–Siebert program* [43] for a (fairly) direct approach to the mirror symmetry predictions of string theory.

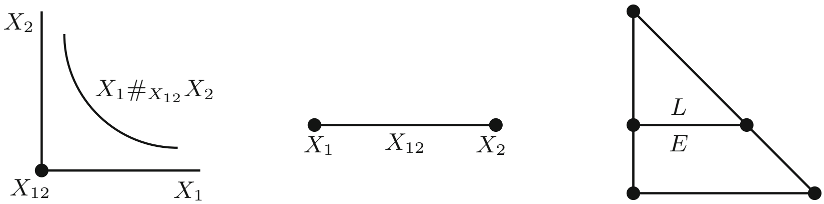


FIGURE 2. A 2-fold simple normal crossings variety $\mathcal{Z}_0=X_1\cup_{X_{12}}X_2$ with its smoothing $\mathcal{Z}_\lambda=X_1\#_{X_{12}}X_2$, its dual intersection complex, and a toric 2-fold decomposition of $\mathbb{P}^2$ into $\mathbb{P}^2$ and its one-point blowup $\widehat{\mathbb{P}}^2$ along a line $L\subset\mathbb{P}^2$ and the exceptional divisor $E\subset\widehat{\mathbb{P}}^2$.

A sequence of J-holomorphic curves in the smooth fibers $\mathcal{Z}_\lambda = X_1\#_{X_{12}}X_2$ of a symplectic fibration $\pi\colon \mathcal{Z}\longrightarrow\Delta$ associated with the construction of [39] with $\lambda\longrightarrow 0$ converges to curves in the singular fiber $\mathcal{Z}_0 = X_1\cup_{X_{12}}X_2$. Each of the irreducible components of a limiting curve either lies entirely in X_{12} or meets X_{12} in finitely many points (possibly none) and lies entirely in either X_1 or X_2. A key prediction in [93] concerning the behavior of the GW-invariants of $\mathcal{Z}_\lambda$ as $\lambda\longrightarrow 0$ is that they should arise only from J-holomorphic curves in $\mathcal{Z}_0$ with no irreducible components contained in X_{12} and with the irreducible components mapped into X_1 and X_2 having the same contacts with X_{12}; see Figure 3. In particular, there should be *no* direct contribution from the GW-invariants of X_{12}. The multiplicity with which such a limiting curve should contribute to the GW-invariants of $\mathcal{Z}_\lambda$ is determined in [9] based on a straightforward algebraic reason.

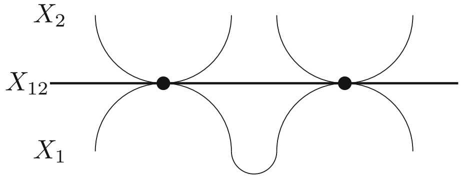


FIGURE 3. A connected curve in $\mathcal{Z}_0$ possibly contributing to the GW-invariants of $\mathcal{Z}_\lambda$.

Notions of stable J-holomorphic maps to *simple normal crossings* (or *SC*) projective varieties of the form $X_1 \cup_{X_{12}} X_2$ and of stable maps to X_i relative to a smooth projective divisor X_{12} are introduced in [56]. A *degeneration formula* relating the virtual cycles of the moduli spaces $\overline{\mathfrak{M}}_{g,k}(A_\lambda; J)$ with $A_\lambda \in H_2(\mathcal{Z}_\lambda)$ to the virtual cycles of the moduli spaces $\overline{\mathfrak{M}}_{g,k}(A_0; J)$ with $A_0 \in H_2(\mathcal{Z}_0)$ appears in [57]. A *splitting formula* decomposing the latter into the virtual cycles of the moduli spaces $\overline{\mathfrak{M}}_{g_i,k_i;\mathbf{s}_i}(X_{12}, A_i; J)$ of stable relative maps to (X_i, X_{12}) via a Kunneth decomposition of the diagonal

$$\Delta_{X_{12}} = \big\{(x,x) \colon x \in X_{12}\big\}$$

in X_{12}^2 is also established in [57]. The relative GW-invariants of (X_i, X_{12}) are in turn shown to reduce to the (absolute) GW-invariants of X_i and X_{12} in [63]. Thus, [56, 57, 63] fully establish the prediction of [93] in the projective category in the case of basic degenerations of the target as in Figure 2. An expository account of the symplectic topology perspective on the numerical reduction of the decomposition formula of [57] appears in [23].

The standard symplectic sum construction of [39] readily extends to the setting where the disjoint union $X_1 \sqcup X_2$ is replaced by a single symplectic manifold $(\widetilde{X}, \widetilde{\omega})$ and the two copies of the divisor X_{12} are replaced by a single smooth symplectic divisor $\widetilde{X}_{12} \subset \widetilde{X}$ with a symplectic involution ψ. The *NC symplectic variety* $\mathcal{Z}_{\psi;0} \equiv X_\psi$ is then obtained from $\widetilde{X}$ by identifying the points on $\widetilde{X}_{12}$ via ψ. This setting is discussed in Example 6.10 in the first two versions of [19]; a construction smoothing $\mathcal{Z}_{\psi;0}$ into symplectic manifolds $\mathcal{Z}_{\psi;\lambda}$ is a special case of the construction outlined in Section 7 of the first version of [20] and detailed in [22]. The reasoning behind the decomposition formulas for GW-invariants in the basic setting of the previous paragraph readily extends to provide a relation between the GW-invariants of a smoothing $\mathcal{Z}_{\psi;\lambda}$ of the NC symplectic variety X_ψ and the relative GW-invariants of $(\widetilde{X}, \widetilde{X}_{12})$. The only difference in the resulting formula is that a Kunneth decomposition of the diagonal $\Delta_{X_{12}} \subset X_{12}^2$ is replaced by a Kunneth decomposition of the ψ-diagonal

$$\widetilde{\Delta}_\psi = \big\{\big(\widetilde{x}, \psi(\widetilde{x})\big) : \widetilde{x} \in \widetilde{X}_{12}\big\};$$

the resulting sum of pairwise products of the GW-invariants of $(\widetilde{X},\widetilde{X}_{12})$ should then be divided by 2.

The decomposition formulas of [57] do not completely determine the GW-invariants of a smooth fiber $\mathcal{Z}_\lambda = X_1\#_{X_{12}}X_2$ in terms of the GW-invariants of (X_i,X_{12}) in many cases because of the so-called *vanishing cycles*: second homology classes in $\mathcal{Z}_\lambda$ which vanish under the projection to $\mathcal{Z}_0=X_1\cup_{X_{12}}X_2$. A refinement to the usual relative GW-invariants of (X,V) of [56] is suggested in [46] with the aim of resolving this unfortunate deficiency of the decompositions formulas of [57] in [47]. This refinement is constructed in [24] via a lifting

$$\widetilde{\mathrm{ev}}_X^V:\overline{\mathfrak{M}}_{g,k;\mathbf{s}}(V,A;J)\longrightarrow \widehat{V}_{X;\mathbf{s}}$$

of the relative evaluation map to a covering of $V_{\mathbf{s}}\equiv V^\ell$, where $\ell\in\mathbb{Z}^{\ge 0}$ is the length of the relative contact vector $\mathbf{s}$. This refinement sharpens the decomposition formulas of [57] by pulling back closed submanifolds

$$\widehat{V}^A_{X_1,X_2;\mathbf{s}}\subset\big(\widehat{V}_{X_1;\mathbf{s}}\times\widehat{V}_{X_2;\mathbf{s}}\big)\big|_{\Delta_V^\ell}, \tag{3.1}$$

with $V=X_{12}$ and $A\in H_2(X_1\#_{X_{12}}X_2)$, by $\widetilde{\mathrm{ev}}_{X_1}^V\times\widetilde{\mathrm{ev}}_{X_2}^V$; see [25, Section 1.2]. However, this does not necessarily lead to a decomposition of the GW-invariants of $X_1\cup_{X_{12}} X_2$ into the GW-invariants of (X_i,X_{12}) that completely describes the former in terms of the latter. The same approach provides a sharper version of the relation between the GW-invariants of a smoothing $\mathcal{Z}_{\psi;\lambda}$ of X_ψ and the GW-invariants of $(\widetilde{X},\widetilde{X}_{12})$ indicated in the previous paragraph. The submanifolds (3.1) in this case are replaced by certain submanifolds

$$\widehat{V}^A_{\widetilde{X};\mathbf{ss}}\subset\widehat{V}_{\widetilde{X};\mathbf{ss}}\big|_{\Delta_\psi^\ell},$$

with $V=\widetilde{X}_{12}$; the resulting relative invariants of $(\widetilde{X},\widetilde{X}_{12})$ should then be divided by 2.

Qualitative applications of the above refinements to relative GW-invariants and to the decomposition formula of [57] are described in [24, 25]. These refinements in principle distinguish between the GW-invariants of $\mathcal{Z}_\lambda$ in degrees A_λ differing by torsion. Torsion classes can also arise from the one-parameter families of smoothings $\mathcal{Z}_{\psi;\lambda}$ of X_ψ as above. *Quantitative* computation of GW-invariants in degrees differing by torsion has been a long-standing problem.

Question 5. *Is it possible to compute GW-invariants in degrees differing by torsion in some cases via the sharper version of the decomposition formula described in* [25] *and/or its analogue for the degenerations of the form $\mathcal{Z}_{\psi;\lambda}$ above?*

The Enriques surface X forms an elliptic fibration over $\mathbb{P}^1$ with 12 nodal fibers and 2 double fibers; see [64, Section 1.3]. The difference F_1-F_2 between the two double fibers is a 2-torsion class. A smooth genus 1 curve E has a fixed-point-free holomorphic involution ψ. The quotient

$$X_2\equiv\big(\mathbb{P}^1\times E\big)/\!\sim,\qquad (z,p)\longrightarrow\big(-z,\psi(p)\big),$$

forms an elliptic fibration over $\mathbb{P}^1$ with 2 double fibers. The blowup $\widetilde{X}$ of $\mathbb{P}^2$ at the 9-point base locus of a generic pencil of cubics is an elliptic fibration over $\mathbb{P}^1$ with 12 nodal fibers. The NC variety $\mathcal{Z}_0 \equiv X_2 \cup_E \widetilde{X}$ can be smoothed out to an Enriques surface $\mathcal{Z}_\lambda \equiv X$. The genus 1 GW-invariants of X are determined in [64] by applying the decomposition formula of [57] in this setting and using the Virasoro constraints. However, the computation in [64] does not distinguish between the map degrees differing by the torsion $F_1 - F_2$; this torsion arises from the vanishing cycles and thus is not detected by the decomposition formula of [57]. On the other hand, it may be possible to fully compute the genus 1 GW-invariants of X by refining the computation in [64] via the sharper version of this formula described in [25].

Another potential approach to a complete computation of the GW-invariants of the Enriques surface X is provided by the extension of the standard symplectic sum construction of [39] indicated above Question 5. Let ψ be a fixed-point-free holomorphic involution on a smooth fiber $F \approx E$ of $\widetilde{X} \longrightarrow \mathbb{P}^1$. The NC variety

$$\mathcal{Z}_{\psi;0} \equiv X_\psi \equiv \widetilde{X}/\sim, \qquad p \sim \psi(p) \ \forall\, p \in \widetilde{X}_{12} \equiv F,$$

has a $\mathbb{Z}^2$-collection of one-parameter families of smoothings $\mathcal{Z}_{\psi;\lambda}$. The total spaces of these families are $\mathbb{Z}_2$-quotients of the total families of the smoothings of $\widetilde{X} \cup_F \widetilde{X}$. The fibers $\mathcal{Z}_\lambda$ of one of the latter families are $K3$ surfaces. Thus, the fibers $\mathcal{Z}_{\psi;\lambda}$ in one of the families of smoothings of X_ψ should be Enriques surfaces (at least up to symplectic deformation equivalence). The extension of the standard degeneration formula of [57] indicated above applies to these families of smoothings and again distinguishes between the GW-invariants in degrees differing by the torsion $F_1 - F_2$.

The Gross–Siebert program [43] for a direct proof of mirror symmetry requires degeneration and splitting formulas for GW-invariants under degenerations $\pi\colon \mathcal{Z} \longrightarrow \Delta$ of algebraic varieties that are locally of the form

$$\begin{gathered} \pi\colon \big\{(\lambda, z_1, \ldots, z_k, p) \in \mathbb{C}^{k+2} \times \mathbb{C}^{n-k} \colon z_1 \cdots z_k = \lambda\big\} \longrightarrow \mathbb{C}, \\ \pi\big(\lambda, z_1, \ldots, z_k, p\big) = \lambda, \end{gathered} \tag{3.2}$$

around the central fiber $\mathcal{Z}_0 \equiv \pi^{-1}(0)$. The degenerations discussed above, i.e., the standard one associated with the symplectic sum construction of [39] and its extension indicated in [19, 20], correspond to $k = 2$ in (3.2). The central fiber of π for $k \geq 3$ in the algebro-geometric category is a more general NC variety; see Figure 4. Degeneration and splitting formulas for GW-invariants in this more general setting require notions of GW-invariants for (smoothable) NC varieties and for smooth varieties relative to NC divisors. A degeneration formula in the projective category extending that of [57] has finally appeared in the setting of the *logarithmic GW-theory* of [44] in [1]; the latter includes GW-invariants of smoothable NC varieties and of smooth varieties relative to NC divisors. However, a splitting formula for the GW-invariants of NC varieties in the projective category remains to be established.

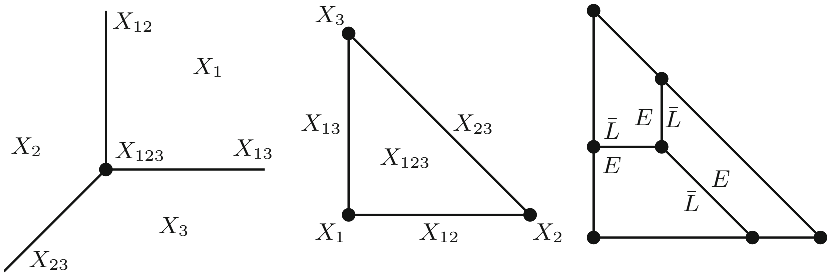


FIGURE 4. A 3-fold simple normal crossings variety $\mathcal{Z}_0$, its dual intersection complex, and a toric 3-fold decomposition of $\mathbb{P}^2$ into three copies of its one-point blowup $\widehat{\mathbb{P}}^2$ along the exceptional divisor $E\subset\widehat{\mathbb{P}}^2$ and the proper transform $\overline{L}\subset\widehat{\mathbb{P}}^2$ of a line $L\subset\mathbb{P}^2$.

The logarithmic GW-invariants of [44] are special cases of the GW-invariants of exploded manifolds introduced in [78]. Degeneration *and* splitting formulas for these invariants are studied in [79]. Based on the $k=2$ case established in [57], one might expect that all curves in $\mathcal{Z}_0$ contributing to the GW-invariants of $\mathcal{Z}_\lambda$ either

- have no irreducible components lying in the singular locus $\mathcal{Z}_0'$ of $\mathcal{Z}_0$ and meet at the smooth points of $\mathcal{Z}_0$ or at least
- have no irreducible components in $\mathcal{Z}_0'$.

As demonstrated in [79], even the weaker alternative does not hold in general. This makes any general splitting formula necessarily complicated; its $k=3$ case is described in [80]. A more geometric perspective on the GW-invariants of [78] appears in [45], without analogues of the crucial degeneration and splitting formulas.

The GW-invariants of exploded manifolds of [78] and their interpretation in some cases in [45] are essentially invariants of deformation equivalence classes of almost Kähler structures on manifolds. While these classes are much larger than the deformation equivalence classes of the algebro-geometric structures in [1, 44], GW-invariants are fundamentally invariants of the still larger deformation equivalence classes of symplectic structures. Purely topological notions of *NC symplectic divisors* and *varieties* are introduced in [19, 21], addressing a fundamental quandary of [42, p. 343] in the case of NC singularities. Crucial to the introduction of these long desired notions is the new perspective proposed in [19]:

A symplectic variety/subvariety should be viewed as a deformation equivalence class of objects with the same topology, not as a single object.

It is then shown in [19, 21] that the spaces of NC symplectic divisors and varieties are weakly homotopy equivalent to the spaces of almost Kähler structures, as needed for geometric applications.

The equivalence between the topological and geometric notions of NC symplectic variety established in [19, 21] immediately implies that any invariants arising from [45, 79] in fact depend only on the deformation equivalence classes of symplectic structures. These equivalences are also used in [20, 22] to establish a smoothability criterion for NC symplectic varieties. Direct approaches to constructing GW-invariants of symplectic manifolds relative to NC symplectic divisors in the perspective of [19, 21] and to obtaining degeneration and splitting formulas for the degenerations appearing in [20, 22] are discussed in [16] and [17], respectively.

The decomposition and splitting formulas for GW-invariants in [79] involve *exploded de Rham cohomology* of [81], which makes these formulas very hard to apply. The purpose of this elaborate modification of the ordinary de Rham cohomology is to correct the standard Kunneth decompositions of the diagonals of the strata of the singular locus $\mathcal{Z}_0'$ of the central fiber $\mathcal{Z}_0$ for the presence of lower-dimensional strata. This removes certain *degenerate contributions* to the Kunneth decompositions of the diagonals of the strata of $\mathcal{Z}_0'$. A local, completely topological approach to computing degenerate contributions in terms of the ordinary cohomology of the strata is presented in [102].

Question 6. *Is there a reasonably usable formula for general NC degenerations $\pi\colon \mathcal{Z} \longrightarrow \Delta$ of symplectic manifolds which splits the GW-invariants of a smooth fiber $\mathcal{Z}_\lambda$ into the GW-invariants of the strata of the central fiber $\mathcal{Z}_0$ that involves only the ordinary cohomology of the strata?*

The introduction of symplectic topology notions of NC divisors, varieties, and degenerations in [19–22] has made it feasible to study Question 6 entirely in the symplectic topology category, which is far more flexible than the algebraic geometry category of [1, 44] and the almost Kähler category of [45, 78]. A symplectic approach to this question should fit well with the topological approach of [102] to degenerate contributions. A splitting formula for GW-invariants of $\mathcal{Z}_\lambda$ resulting from such an approach should involve sums over finite trees with the edges labeled by integer weights and the vertices labeled by paths in the dual intersection complex of $\mathcal{Z}_0$ with additional de Rham cohomology data; these paths would correspond to the *tropical curves* of [79]. While such a formula would still be more complicated than in the standard case of [57], it should be more readily applicable than the presently available splitting formula of [79] that involves exploded de Rham cohomology.

Degeneration and splitting formulas for real GW-invariants under real degenerations of real symplectic manifolds have been obtained only in a small number of special cases. A fundamental difficulty for obtaining such formulas is that the standard notions of relative invariants of the complex GW-theory do not have direct analogues in the real GW-theory in most settings. Real GW-invariants of a real symplectic manifold (X,ω,ϕ) with simple contacts with a real symplectic divisor $V\subset X$ can be readily defined whenever the real GW-invariants of (X,ω,ϕ) are defined and V is disjoint from the fixed locus X^ϕ of ϕ. This observation lies behind the splitting formula and related vanishing result for some genus 0 real GW-invariants under special real degenerations of real symplectic manifolds obtained in [15].

The reduction of the complex relative GW-invariants of (X,V) to the complex GW-invariants of X and V in [63] suggests the possibility of expressing the real GW-invariants of a real symplectic sum $X_1\#_{X_{12}}X_2$ in terms of the real GW-invariants of X_1, X_2, X_{12}, whenever these are defined. If X_1 and X_2 are of real dimension four, then X is a real surface and the real GW-invariants of $X_1\#_{X_{12}}X_2$ should reduce to the real GW-invariants of X_1 and X_2. By [4, Theorem 7] and [3, Theorem 1.1], this is indeed the case for the genus 0 real GW-invariants if $X_{12} \approx \mathbb{P}^1$ is a real symplectic submanifold of self-intersection 2 in $X_2 = \mathbb{P}^1 \times \mathbb{P}^1$ and in some other settings with $X_{12} \approx \mathbb{P}^1$. Genus 0 real GW-invariants have been defined for many real symplectic sixfolds and for all real symplectic fourfolds. This leads to the following question.

Question 7. *Is it possible to express the genus 0 real GW-invariants of a real symplectic sum $X_1\#_{X_{12}}X_2$ of real symplectic sixfolds (X_i, ω_i, ϕ_i) along a common real symplectic divisor X_{12} in terms of the genus 0 real GW-invariants X_1, X_2, X_{12}, whenever the genus 0 real GW-invariants of the sixfolds are defined?*

4. Geometric applications

Pseudoholomorphic curves were originally introduced in [41] with the aim of applications in symplectic topology. These applications have included the Symplectic Non-Squeezing Theorem [41], classification of symplectic 4-manifolds [53, 65], distinguishing diffeomorphic symplectic manifolds [87], symplectic isotopy problem [90, 94], and applications in birational algebraic geometry [50, 95]. However, many deep related problems remain open.

Rational curves, i.e., images of J-holomorphic maps from chains of spheres, play a particularly important role in algebraic geometry. A smooth algebraic manifold X is called *uniruled* (resp. *rationally connected* or *RC*) if there is a rational curve through every point (resp. every pair of points) in X. According to [50], a uniruled algebraic variety admits a nonzero genus 0 GW-invariant with a point insertion (i.e., a count of stable maps in a fixed homology class which pass through a point and some other constraints). This implies that the uniruled property is invariant under symplectic deformations. The RC property is known to be invariant under integrable deformations of the complex structure [50]. It is a long-standing conjecture of J. Kollár that the RC property is invariant under symplectic deformations as well. It is unknown if every RC algebraic manifold admits a nonzero genus 0 GW-invariant with two point insertions; this would immediately imply Kollár's conjecture. The dimension three case of this conjecture is established in [95] by combining the special cases treated in [98] with the minimal model program.

As GW-invariants are symplectic invariants, it is natural to consider the parallel situation in symplectic topology. Given the flexibility of the symplectic category, this may also provide a different approach to Kollár's conjecture. A symplectic manifold (X,ω) is called *uniruled* (resp. *RC*) if for some ω-compatible almost complex structure J there is a genus 0 connected rational J-holomorphic curve through every point (resp. every pair of points) in X. This leads to the following two pairs of questions.

Question 8. *Let J be any almost complex structure on a uniruled (resp. RC) compact symplectic manifold (X,ω). Is there a connected rational J-holomorphic curve through every point (resp. every pair of points) in X?*

Question 9. *Does every uniruled (resp. RC) compact symplectic manifold (X,ω) admit a nonzero genus 0 GW-invariant with a point insertion (resp. two point insertions)?*

The affirmative answer to each case of Question 9 would immediately imply the affirmative answer to the corresponding case of Question 8. The uniruled case of Question 9 is known only under the rigidity assumptions that X is either Kähler [50] or admits a Hamiltonian S^1-action [66]. It is not difficult to construct J-holomorphic curves in a symplectic manifold that disappear as the almost complex structure J deforms. On the other hand, regular J-holomorphic curves do not disappear under small deformations of J, while J-holomorphic curves contributing to nonzero GW-invariants survive all deformations of J. Thus, the above four questions concern the fundamental issue of the extent of flexibility in the symplectic category with implications to birational algebraic geometry.

If $u\colon \mathbb{P}^1 \longrightarrow X$ is a J-holomorphic map into a Kähler manifold and for *some* $z\in\mathbb{P}^1$ the evaluation map

$$H^0(\mathbb{P}^1;u^*TX) \longrightarrow T_{u(z)}X, \qquad \xi \longrightarrow \xi(z), \tag{4.1}$$

is onto, then $H^1(\mathbb{P}^1;u^*TX)=0$, i.e., u is regular. This statement is key to the arguments of [50] in the algebraic setting. It in particular implies that if the rational J-holomorphic curves cover a nonempty open subset of a connected Kähler manifold, then they cover all of X. As shown in [67], the last implication can fail in the almost Kähler category. The first implication need not hold either, even if the evaluation homomorphism (4.1) is surjective for *every* $z\in\mathbb{P}^1$. However, the main results of [50] may still extend to the almost Kähler category. In particular, for the interplay between openness and closedness of various properties of complex structures exhibited in the proof of deformation invariance of the RC property for integrable complex structures in [50] to extend to a non-integrable complex structure, the vanishing of the obstruction space needs to hold only generically in a family of J-holomorphic maps covering X. This leads to a potentially even more fundamental problem in this spirit.

Question 10. *Let $\{u_\alpha\colon \mathbb{P}^1 \longrightarrow X\}$ be a family of J-holomorphic curves on a compact symplectic manifold (X,ω) that covers X. Is a generic member of this family a regular map?*

There are still many open questions concerning the geography and topology of symplectic manifolds The multifold smoothing constructions of [20, 22] may shed light on some of these questions. Just as the (2-fold) symplectic sum construction of [39], the multifold constructions could be used to build vast classes of non-Kähler symplectic manifolds with various topological properties. They might also be useful for studying properties of symplectic manifolds of algebro-geometric flavor, in the spirit of the perspective on symplectic topology initiated in [41].

Question 11 ([18, Question 14]). *Is every compact almost Kähler manifold with a rational J-holomorphic curve of a fixed homology class through every pair of points simply connected?*

By [8, Theorem 3.5], a compact RC Kähler manifold is simply connected. As noted by J. Starr, the fundamental group of a compact almost Kähler manifold (X,ω,J) as in Question 11 is finite. The multifold sum/smoothing constructions of [20, 22] can be used to obtain symplectic manifolds that are not simply connected from simply connected ones and thus may be useful in answering Question 11 negatively. The constructions of [20, 22] may also be useful in studying this question under the stronger assumption of the existence of a nonzero GW-invariant of (X,ω) with two point insertions.

As in the complex case, it is natural to expect that a real symplectic manifold (X,ω,ϕ) which has well-defined genus 0 real GW-invariants and is covered by real rational curves admits a nonzero genus 0 real GW-invariant with a real point insertion. However, the reasoning neither in [50], which relies on the positivity of intersections in complex geometry, nor in [66], which makes use of quantum cohomology, is readily adaptable to the real setting. Thus, there is not apparent approach to this problem at the present.

Another important question in real algebraic geometry is the existence of real rational curves on real even-degree complete intersections $X\subset\mathbb{P}^n$; this would be implied by the existence of a well-defined nonzero genus 0 real GW-invariant of X. However, the real analogue of the Quantum Lefschetz Hyperplane Principle (1.2) suggests that all such invariants should vanish. On the other hand, one may hope for some real analogue of the reduced/family GW-invariants of [5, 55], which effectively remove a trivial line bundle from the obstruction cone for deformations of J-holomorphic maps to X. The resulting reduced/family real invariants could well be nonzero.

Acknowledgment

G. Tian's perspectives on Gromov–Witten theory had a tremendous influence on the content of the present article in particular and the work of the author in general, and he is very grateful to G. Tian for generously sharing his insights on Gromov–Witten theory over the past two decades. The author would also like to thank J. Li, R. Pandharipande, and R. Vakil for introducing him to the richness of the algebro-geometric side of Gromov–Witten theory indicated by many of the questions in this article and P. Georgieva for acquainting him with the many related mysteries of the real sector of Gromov–Witten theory, as well as E. Brugallé, A. Doan, and C. Wendl for comments on parts of this article.

References

[1] D. Abramovich, Q. Chen, M. Gross, and B. Siebert, *Decomposition of degenerate Gromov–Witten invariants*, math/1709.09864

[2] M. Bershadsky, S. Cecotti, H. Ooguri, and C. Vafa, *Holomorphic anomalies in topological field theories*, Nucl. Phys. B405 (1993), 279–304

[3] E. Brugallé, *Surgery of real symplectic fourfolds and Welschinger invariants*, J. Singul. 17 (2018), 267–294

[4] E. Brugallé and N. Puignau, *On Welschinger invariants of symplectic* 4*-manifolds*, Comment. Math. Helv. 90 (2015), no. 4, 905–938

[5] J. Bryan and N. Leung, *The enumerative geometry of $K3$ surfaces and modular forms*, J. AMS 13 (2000), no. 2, 371–410

[6] J. Bryan and R. Pandharipande, *The local Gromov–Witten theory of curves*, J. AMS 21 (2008), no. 1, 101–136

[7] P. Candelas, X. de la Ossa, P. Green, and L. Parkes, *A Pair of Calabi–Yau manifolds as an exactly soluble superconformal theory*, Nuclear Phys. B359 (1991), 21–74

[8] F. Campana, *On twistor spaces of the class $\mathcal{C}$*, J. Diff. Geom. 33 (1991), no. 2, 541–549

[9] L. Caporaso and J. Harris, *Counting plane curves in any genus*, Invent. Math. 131 (1998), no. 2, 345–392

[10] B. Chen, A.-M. Li, and B.-L. Wang, *Virtual neighborhood technique for pseudo-holomorphic spheres*, math/1306.3276

[11] B. Chen and G. Tian, *Virtual manifolds and localization*, Acta Math. Sin. 26 (2010), no. 1, 1–24

[12] A. Doan and T. Walpulski, *Castelnuovo's bound and rigidity in almost complex geometry*, math/1809.04731

[13] E. Eftekhary, *On finiteness and rigidity of J-holomorphic curves in symplectic threefolds.* Adv. Math. 289 (2016), 1082–1105

[14] G. Ellingsrud and S. Strömme, *Bott's formula and enumerative geometry*, J. AMS 9 (1996), no. 1, 175–193

[15] M. Farajzadeh Tehrani, *Open Gromov–Witten theory on symplectic manifolds and symplectic cutting*, Adv. Math. 232 (2013), no. 1, 238–270

[16] M. Farajzadeh Tehrani, *Pseudoholomorphic maps relative to normal crossings symplectic divisors: compactification*, math/1710.00224

[17] M. Farajzadeh Tehrani, *Towards a degeneration formula for the Gromov–Witten invariants of symplectic manifolds*, math/1710.00599

[18] M. Farajzadeh Tehrani, M. McLean, and A. Zinger, *Singularities and semistable degenerations for symplectic topology*, C. R. Math. Acad. Sci. Paris 356 (2018), no. 4, 420–432

[19] M. Farajzadeh Tehrani, M. McLean, and A. Zinger, *Normal crossings singularities for symplectic topology*, Adv. Math. 339 (2018), 672–748

[20] M. Farajzadeh Tehrani, M. McLean, and A. Zinger, *The smoothability of normal crossings symplectic varieties*, math/1707.01464

[21] M. Farajzadeh-Tehrani, M. McLean, and A. Zinger, *Normal crossings singularities for symplectic topology, II*, in preparation

[22] M. Farajzadeh-Tehrani, M. McLean, and A. Zinger, *The smoothability of normal crossings symplectic varieties, II*, in preparation

[23] M. Farajzadeh Tehrani and A. Zinger, *On symplectic sum formulas in Gromov–Witten theory*, math/1404.1898

[24] M. Farajzadeh Tehrani and A. Zinger, *On the rim tori refinement of relative Gromov–Witten invariants*, math/1412.8204

[25] M. Farajzadeh Tehrani and A. Zinger, *On the refined symplectic sum formula for Gromov–Witten invariants*, math/1412.8205

[26] K. Fukaya and K. Ono, *Arnold Conjecture and Gromov–Witten invariant*, Topology 38 (1999), no. 5, 933–1048

[27] W. Fulton and R. Pandharipande, *Notes on stable maps and quantum cohomology*, Proc. Sympos. Pure Math. 62, Part 2, 45–96, AMS, 1997

[28] P. Georgieva, *Open Gromov–Witten invariants in the presence of an anti-symplectic involution*, Adv. Math. 301 (2016), 116–160

[29] P. Georgieva and E. Ionel, *A Klein TQFT: the local Real Gromov–Witten theory of curves*, math/1812.02505

[30] P. Georgieva and A. Zinger, *Enumeration of real curves in* $\mathbb{CP}^{2n-1}$ *and a WDVV relation for real Gromov–Witten invariants*, Duke Math. J. 166 (2017), no. 17, 3291–3347

[31] P. Georgieva and A. Zinger, *A recursion for counts of real curves in* $\mathbb{CP}^{2n-1}$*: another proof*, Internat. J. Math. 29 (2018), no. 4, 1850027

[32] P. Georgieva and A. Zinger, *Real Gromov–Witten theory in all genera and real enumerative geometry: construction*, Ann. of Math. 188 (2018), no. 3, 685–752

[33] P. Georgieva and A. Zinger, *Real Gromov–Witten theory in all genera and real enumerative geometry: properties*, math/1507.06633v4, to appear in J. Symplectic Geom. 17 (2019), no. 4

[34] P. Georgieva and A. Zinger, *Real Gromov–Witten theory in all genera and real enumerative geometry: computation*, math/1510.07568, to appear in J. Diff. Geom.

[35] E. Getzler, *Intersection theory on* $\overline{\mathcal{M}}_{1,4}$ *and elliptic Gromov–Witten invariants*, J. AMS 10 (1997), no. 4, 973–998

[36] A. Givental, *The mirror formula for quintic threefolds*, *California Symplectic Geometry Seminar*, AMS Transl. Ser. 196, 49–62, AMS, 1999

[37] A. Givental, *Gromov–Witten invariants and quantization of quadratic Hamiltonians*, Mosc. Math. J. 1 (2001), no. 4, 551–568

[38] P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, Wiley Classics Library, 1994

[39] R. Gompf, *A new construction of symplectic manifolds*, Ann. of Math. 142 (1995), no. 3, 527–595

[40] R. Gopakumar and C. Vafa, *M-theory and topological strings I,II*, hep-th/9809187, 9812127

[41] M. Gromov, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. 82 (1985), no. 2, 307–347

[42] M. Gromov, *Partial Differential Relations*, Springer-Verlag, 1986

[43] M. Gross and B. Siebert, *Affine manifolds, log structures, and mirror symmetry*, Turkish J. Math. 27 (2003), no. 1, 33–60

[44] M. Gross and B. Siebert, *Logarithmic Gromov–Witten invariants*, J. AMS 26 (2013), no. 2, 451–510

[45] E. Ionel, *GW-invariants relative normal crossings divisors*, Adv. Math. 281 (2015), 40–141

[46] E. Ionel and T. Parker, *Relative Gromov–Witten invariants*, Ann. of Math. 157 (2003), no. 1, 45–96

[47] E. Ionel and T. Parker, *The symplectic sum formula for Gromov–Witten invariants*, Ann. of Math. 159 (2004), no. 3, 935–1025

[48] E. Ionel and T. Parker, *The Gopakumar–Vafa formula for symplectic manifolds*, Ann. of Math. 187 (2018), no. 1, 1–64

[49] A. Klemm and R. Pandharipande, *Enumerative geometry of Calabi–Yau 4-folds*, Comm. Math. Phys. 261 (2006), no. 2, 451–516

[50] J. Kollár, *Rational Curves on Algebraic Varieties*, Modern Surveys in Mathematics, Springer-Verlag, 1996

[51] M. Kontsevich, *Enumeration of rational curves via torus actions*, The Moduli Space of Curves, Progr. Math. 129 (1995), 335–368

[52] M. Kontsevich and Y. Manin, *Gromov–Witten classes, quantum cohomology, and enumerative geometry*, Comm. Math. Phys. 164 (1994), no. 3, 525–562

[53] F. Lalonde and D. McDuff, *J-curves and the classification of rational and ruled symplectic 4-manifolds*, Contact and Symplectic Geometry, 3–42, Publ. Newton Inst. 8, 1996

[54] B. Lawson and M.-L. Michelsohn, *Spin Geometry*, Princeton Mathematical Series 38, 1989

[55] J. Lee, *Family Gromov–Witten invariants for Kähler surfaces*, Duke Math. J. 123 (2004), no. 1, 209–233

[56] J. Li, *Stable morphisms to singular schemes and relative stable morphisms*, J. Diff. Geom. 57 (2001), no. 3, 509–578

[57] J. Li, *A degeneration formula of GW-invariants*, J. Differential Geom. 60 (2002), no. 3, 199–293

[58] J. Li and G. Tian, *Virtual moduli cycles and Gromov–Witten invariants of algebraic varieties*, J. AMS 11 (1998), no. 1, 119–174

[59] J. Li and G. Tian, *Virtual moduli cycles and Gromov–Witten invariants of general symplectic manifolds*, Topics in Symplectic 4-Manifolds, 47–83, First Int. Press Lect. Ser., I, Internat. Press, 1998

[60] J. Li and A. Zinger, *On the genus-one Gromov–Witten invariants of complete intersections*, J. Diff. Geom. 82 (2009), no. 3, 641–690

[61] B. Lian, K. Liu, and S.-T. Yau, *Mirror Principle I*, Asian J. of Math. 1, No. 4 (1997), 729–763

[62] X. Liu and G. Tian, *Virasoro constraints for quantum cohomology*, J. Diff. Geom. 50 (1998), no. 3, 537–590

[63] D. Maulik and R. Pandharipande, *A topological view of Gromov–Witten theory*, Topology 45 (2006), no. 5, 887–918

[64] D. Maulik and R. Pandharipande, *New calculations in Gromov–Witten theory*, Pure Appl. Math. Q. 4 (2008), no. 2, 469–500

[65] D. McDuff, *The structure of rational and ruled symplectic four-manifolds*, J. AMS 3 (1990), 679–712

[66] D. McDuff, *Hamiltonian S^1-manifolds are uniruled*, Duke Math. J. 146 (2009), no. 3, 449–507

[67] M. McLean, *Local uniruledness does Not imply global uniruledness in symplectic topology*, available from the author's website

[68] H. Nakajima and K. Yoshioka, *Instanton counting on blowup, I: 4-dimensional pure gauge theory*, Invent. Math. 162 (2005), no. 2, 313–355

[69] J. Niu, *A sharp compactness theorem for genus-two pseudo-holomorphic maps*, PhD thesis, Stony Brook, 2016

[70] J. Niu and A. Zinger, *Lower bounds for the enumerative geometry of positive-genus real curves*, Adv. Math. 339 (2018), 191–247

[71] A. Okounkov and R. Pandharipande, *Virasoro constraints for target curves*, Invent. Math. 163 (2006), no. 1, 47–108, AMS, 2009

[72] A. Okounkov and R. Pandharipande, *Gromov–Witten theory, Hurwitz numbers, and matrix models*, Proc. Sympos. Pure Math. 80, Part 1, 325–414

[73] R. Pandharipande, *Intersections of $\mathbb{Q}$-divisors on Kontsevich's moduli space $\overline{M}_{0,n}(\mathbb{P}^r, d)$ and enumerative geometry*, Trans. AMS 351 (1999), no. 4, 1481–1505

[74] R. Pandharipande, *A geometric construction of Getzler's elliptic relation*, Math. Ann. 313 (1999), no. 4, 715–729

[75] R. Pandharipande, *Three questions in Gromov–Witten theory*, Proceedings of ICM, Vol. II, 503–512, 2002

[76] R. Pandharipande and A. Zinger, *Enumerative Geometry of Calabi–Yau 5-folds*, New Developments in Algebraic Geometry, Integrable Systems, and Mirror Symmetry, 239–288, Adv. Stud. Pure Math. 59, 2010

[77] J. Pardon, *An algebraic approach to virtual fundamental cycles on moduli spaces of pseudo-holomorphic curves*, Geom. Topol. 20 (2016), no. 2, 779–1034

[78] B. Parker, *Exploded fibrations*, Proceedings of Gokova Geometry-Topology Conference 2006 (2007), 52–90

[79] B. Parker, *Gromov–Witten invariants of exploded manifolds*, math/1102.0158

[80] B. Parker, *Gluing formula for Gromov–Witten invariants in a triple product*, math/1511.00779

[81] B. Parker, *de Rham theory of exploded manifolds*, Geom. Topol. 22 (2018), no. 1, 1–54

[82] D. Petersen, *The structure of the tautological ring in genus one*, Duke Math. J. 163 (2014), no. 4, 777–793

[83] D. Petersen, *Tautological rings of spaces of pointed genus two curves of compact type*, Compos. Math. 152 (2016), no. 7, 1398–1420

[84] D. Petersen and O. Tommasi, *The Gorenstein conjecture fails for the tautological ring of $\overline{\mathcal{M}}_{2,n}$*, Invent. Math. 196 (2014), no. 1, 139–161

[85] A. Popa, *The genus one Gromov–Witten invariants of Calabi–Yau complete intersections*, Trans. AMS 365 (2013), no. 3, 1149–1181

[86] T. Riviére and G. Tian, *The singular set of* 1-1 *integral currents*, Ann. of Math. 169 (2009), no. 3, 741–794

[87] Y. Ruan, *Symplectic topology on algebraic* 3*-folds*, J. Diff. Geom. 39 (1994), no. 1, 215–227

[88] Y. Ruan and G. Tian, *A mathematical theory of quantum cohomology*, J. Diff. Geom. 42 (1995), no. 2, 259–367

[89] Y. Ruan and G. Tian, *Higher genus symplectic invariants and sigma models coupled with gravity*, Invent. Math. 130 (1997), no. 3, 455–516

[90] B. Siebert and G. Tian, *On the holomorphicity of genus two Lefschetz fibrations*, Ann. of Math. 161 (2005), no. 2, 959–1020

[91] J. Solomon, *A differential equation for the open Gromov–Witten potential*, pre-print 2007

[92] C. Teleman, *The structure of* 2*D semi-simple field theories*, Invent. Math. 188 (2012), no. 3, 525–588, 2012

[93] G. Tian, *The quantum cohomology and its associativity*, Current Developments in Mathematics 1995, 361–401, Inter. Press.

[94] G. Tian, *Symplectic isotopy in four dimension*, First International Congress of Chinese Mathematicians, 143–147, AMS/IP Stud. Adv. Math. 20, 2001

[95] Z. Tian, *Symplectic geometry of rationally connected threefolds*, Duke Math. J. 161 (2012), no. 5, 803–843

[96] R. Vakil and A. Zinger, *A natural smooth compactification of the space of elliptic curves in projective space*, ERA AMS 13 (2007), 53–59

[97] R. Vakil and A. Zinger, *A desingularization of the main component of the moduli space of genus-one stable maps into* $\mathbb{P}^n$, Geom. Topol. 12 (2008), no. 1, 1–95

[98] C. Voisin, *Rationally connected* 3*-folds and symplectic geometry*, Astérisque 322 (2008), 1–21

[99] J. Walcher, *Evidence for tadpole cancellation in the topological string*, Comm. Number Theory Phys. 3 (2009), no. 1, 111–172

[100] J.-Y. Welschinger, *Invariants of real symplectic* 4*-manifolds and lower bounds in real enumerative geometry*, Invent. Math. 162 (2005), no. 1, 195–234

[101] C. Wendl, *Transversality and super-rigidity for multiply covered holomorphic curves*, math/1609.09867

[102] A. Zinger, *Counting rational curves of arbitrary shape in projective spaces*, Geom. Topol. 9 (2005), 571–697

[103] A. Zinger, *On the structure of certain natural cones over moduli spaces of genus-one holomorphic maps*, Adv. Math. 214 (2007), no. 2, 878–933

[104] A. Zinger, *Pseudocycles and integral homology*, Trans. AMS 360 (2008), no. 5, 2741–2765

[105] A. Zinger, *Standard vs. reduced genus-one Gromov–Witten invariants*, Geom. Topol. 12 (2008), no. 2, 1203–1241

[106] A. Zinger, *A sharp compactness theorem for genus-one pseudo-holomorphic maps*, Geom. Topol. 13 (2009), no. 5, 2427–2522

[107] A. Zinger, *Reduced genus-one Gromov–Witten invariants*, J. Diff. Geom. 83 (2009), no. 2, 407–460

[108] A. Zinger, *The reduced genus* 1 *Gromov–Witten invariants of Calabi–Yau hypersurfaces*, J. AMS 22 (2009), no. 3, 691–737

[109] A. Zinger, *A comparison theorem for Gromov–Witten invariants in the symplectic category*, Adv. Math. 228 (2011), no. 1, 535–574

Aleksey Zinger
Department of Mathematics
Stony Brook University
Stony Brook, NY 11794, USA
e-mail: azinger@math.stonybrook.edu