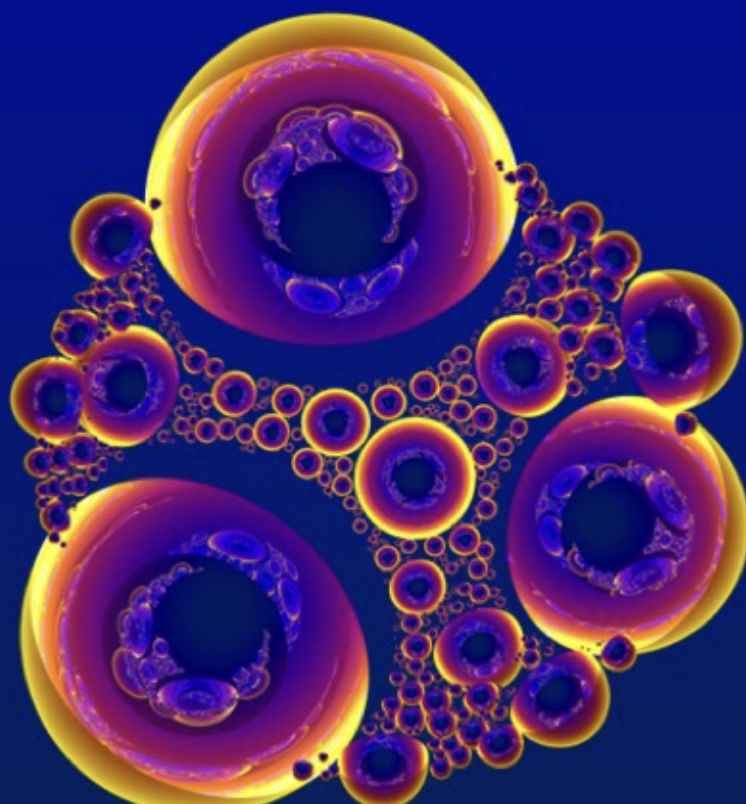
ARNOLD MATHEMATICAL JOURNAL



ASSOCIATION FOR MATHEMATICAL RESEARCH

IMS Stony Brook

2025

Volume 11, Issue 2

Arnold Mathematical Journal

Journal of **Institute of Mathematical Sciences**, Stony Brook University, NY
Published by **Association for Mathematical Research**, Davis, CA; Jenkintown PA.

Contents

Arnold Mathematical Journal
Volume 11, Issue 2, 2025

Non-fillability of Overtwisted Contact Manifolds via Polyfolds

Wolfgang Schmaltz Stefan Suhr Kai Zehmisch

Received 10 Mar 2024; Accepted 10 Nov 2024

Abstract: We prove that any weakly symplectically fillable contact manifold is tight. Furthermore we verify the strong Weinstein conjecture for contact manifolds that appear as the concave boundary of a directed symplectic cobordism whose positive boundary satisfies the weak-filling condition and is overtwisted. Similar results are obtained in the presence of bordered Legendrian open books whose binding–complement has vanishing second Stiefel–Whitney class. The results are obtained via polyfolds.

AMS Classification: 53D42; 53D40, 57R17, 53D45, 37J55, 34C25, 37C27

1 Introduction

In [22] Eliashberg introduced a dichotomy of closed contact 3-manifolds, the tight and overtwisted contact structures. He established in [22] an *h*-principle in the sense of

DOI: 10.56994/ARMJ.011.002.001

Gromov [33] for overtwisted contact structures. The higher dimensional analogue was defined by Borman–Eliashberg–Murphy [11]. One way to detect tight contact structures on a 3-manifolds is to find a weak symplectic filling. In view of the filling–by–holomorphic–discs technique such fillable contact manifolds cannot be overtwisted, see [23, 32] and cf. [29, Corollary 3.8]. In higher dimensions obstructions to overtwistedness in terms of semi-positive weak symplectic fillings were obtained by Niederkrüger [55] and Massot–Niederkrüger–Wendl [50]. The aim of this work is to remove the assumption of being semi-positive.

We consider not necessarily connected $(2n-1)$-dimensional contact manifolds (M,ξ) and assume that there is a contact form α on M defining ξ, i.e. ξ is the kernel of α. The restriction of $\mathrm{d}\alpha$ to ξ is a symplectic form providing ξ with the symplectic orientation via $(\mathrm{d}\alpha)^{n-1}$. The contact manifold (M,ξ) is oriented by $\alpha\wedge(\mathrm{d}\alpha)^{n-1}$. These notions are independent of the choice of contact form as long as the contact form equals $f\alpha$ for a positive smooth function f on M.

A compact $2n$-dimensional symplectic manifold (W,Ω) provided with the symplectic orientation Ω^n is called a **weak symplectic filling** of a given $(2n-1)$-dimensional contact manifold (M,ξ), if $\partial W = M$ as oriented manifolds, where ∂W carries the boundary orientation, such that the following condition is satisfied: For all choices of positive contact forms α for ξ the differential forms

$$\alpha\wedge\omega^{n-1} \quad\text{and}\quad \alpha\wedge\big(\mathrm{d}\alpha+\omega\big)^{n-1}, \quad\text{where}\quad \omega := \Omega|_{TM}\,,$$

are positive volume forms on M, see [20, 50]. Fixing a contact form α for ξ the latter is equivalent to

$$\alpha\wedge\big(f\mathrm{d}\alpha+\omega\big)^{n-1} > 0$$

for all non-negative smooth functions f on M. A contact manifold (M,ξ) is **weakly symplectically fillable** if it admits a weak symplectic filling.

If (M,ξ) contains an overtwisted disc, then (M,ξ) is called **overtwisted**; otherwise (M,ξ) is called **tight**, see [11] and Section 2.2.

Theorem 1. *Any weakly symplectically fillable contact manifold is tight.*

Potentially, Theorem 1 can be obtained with Pardon's [59] rigorously defined contact homology. An argument is indicated in Remark 1 below. We will prove Theorem 1 along the classical line of reasoning due to Gromov [32] and Eliashberg [23], cf. [69, 55, 50]. In fact, Theorem 1, will follow as a special case of Theorem 5.1.2 (ii). For that we remark, that a contact manifold, which contains an overtwisted disc, also contains a parallelisable small plastikstufe whose core is a torus, see the discussion in Section 2.2 and Theorem 2.2.1. A plastikstufe is an example of a bordered Legendrian open book such that the book fibration is trivial, the page is a product of an interval with the binding and the binding is the core, see Section 2.1. Whenever a bordered Legendrian open book in an ambient closed contact manifold is **small**, i.e. has a contractible neighbourhood, it was shown in [50, Theorem 4.4] that no *semi-positive* weak symplectic filling can exists. The way in which the theorem is formulated suggests the conjecture that the statement should be true even without the assumption of semi-positivity. Here we prove:

Theorem 2. *A contact manifold is not weakly symplectically fillable provided that it contains a small bordered Legendrian open book such that the complement of the binding has vanishing second Stiefel–Whitney class.*

Theorem 2 implies Theorem 1 and directly follows from Theorem 5.1.2 (ii). The examples of small bordered Legendrian open books given in [50, Proposition 5.9] all have vanishing second Stiefel–Whitney class though they are sometimes not orientable and, hence, are not spin, see Example 2.1.3. This leaves the question, whether there are contact manifolds that are (i) weakly symplectically fillable (and therefore tight with Theorem 1) and (ii) that admit a bordered Legendrian open book, whose binding–complement could be orientable, but is not spin. Note that non of the potential weak symplectic fillings can be semi-positive.

Restricting to weak symplectic fillings that are semi-positive for a moment Theorem 2 holds true also if the second Stiefel–Whitney class does not vanish. The reason is, that a compact 1-dimensional manifold has an even number of boundary components which is used in a typical Gromov–Witten–invariant type argument performed in a potential weak symplectic filling. Taking holomorphic discs with boundary on the bordered Legendrian open book that intersect a given path connecting the binding with the boundary inside a page yields a 1-dimensional moduli space. At the end of the path on the binding there is a foliation by boundary circles of Bishop discs; at the other end, which corresponds to the boundary of the page, no holomorphic disc homotopic to a Bishop disc does exist. After perturbing the almost complex structure no bubbling off takes place for the relevant moduli space in a semi-positive setting. In other words, the 1-dimensional moduli space is compact with an odd number of boundary components. This is not possible thus contradicts the presence of a weak symplectic filling.

Note that this is in contrast to the 1-dimensional branched manifolds that appear in the non semi-positive setting. Namely, in general, the solution space of a perturbed Cauchy–Riemann operator branches, because of the need of multisections near nodal holomorphic discs with multiply covered sphere bubbles of negative first Chern number. The vanishing assumption for the second Stiefel–Whitney class in the non semi-positive setting allows to orient the solution space (see Remark 7.4.1) resulting in an oriented compact 1-dimensional weighted branched manifold, which has an even number of boundary components, i.e. yields the desired contradiction to weak symplectic fillability. We remark, that moduli spaces of holomorphic discs in general are not orientable in contrast to the case of spheres, see [26].

As a consequence of Theorem 2 we can verify the conjecture stated in [50, Theorem 5.13 (a)] involving circular contactisations of Liouville domains also called **Giroux domaines**:

Theorem 3. *A contact manifold is not weakly symplectically fillable provided that it contains a domain that is obtained from a Giroux domain with disconnected boundary, where one boundary component is blown down via a contact cut.*

The construction in Example 2.1.3 yields a small bordered Legendrian open book whose binding–complement has trivial second Stiefel–Whitney class. Hence, Theorem 3 follows from Theorem 2.

Theorem 5.1.2 also verifies instances of the Weinstein conjecture, which asks for the existence of periodic Reeb orbits for all closed contact manifolds, see [65]. For a short historical review see [21, Section 1]. Our approach, besides the usage of polyfolds, is based on the work of Hofer [35], Albers–Hofer [6] and Niederkrüger–Rechtman [57], and yields so-called Reeb links: A **Reeb link** is a finite collection of parametrised periodic Reeb orbits each of which is oriented by the corresponding Reeb vector field and possibly multiply covered, see [2]. A Reeb link is called **null-homologous** if the link components counted multiplicity add up to zero in homology. The **strong Weinstein conjecture** as formulated by Abbas–Cieliebak–Hofer in [2] asserts the existence of a null-homologous Reeb link for all contact forms on all closed contact manifolds.

Theorem 4. *The strong Weinstein conjecture holds true for all contact manifolds that appear as the concave boundary of a directed symplectic cobordism whose positive end satisfies the weak-filling condition and that are at least one of the following:*

(i) an overtwisted contact manifold,

(ii) a contact manifold that contains a small bordered Legendrian open book such that the complement of the binding has vanishing second Stiefel–Whitney class,

(iii) a contact manifold that contains a domain that is obtained from a Giroux domain with disconnected boundary, where one boundary component is blown down via a contact cut.

The relevant notions related to directed symplectic cobordisms can be found in Section 5.1. Theorem 4 follows from Theorem 5.1.2 together with the remarks made for Theorem 1 and 3 above.

For the proof of Theorem 5.1.2, which implies Theorem 2 and 4, we will use the following alternative characterisation of weak symplectic fillability from [50]: A compact symplectic manifold (W, Ω) is a weak symplectic filling of a contact manifold $(M, \xi = \ker \alpha)$ if $\partial W = M$ as oriented manifolds and if there exists an Ω-tamed almost complex structure

J on W such that ξ is J-invariant and the restriction of $\mathrm{d}\alpha$ to ξ tames J. In this situation (W, Ω, J) is called a **tamed pseudo-convex** manifold, see [23] and cf. Section 4. This point of view allows the use of holomorphic disc fillings in the sense of Bishop, see Section 3.

It turns out that fillability questions can be perfectly described in the language of symplectic cobordisms, see Section 5. Assuming non-existence of Reeb links of the Reeb flows that appear on the negative ends of the symplectic cobordism the Gromov–Witten–invariant type polyfolds can be defined in the sense of Hofer–Wysocki–Zehnder [36, 37, 38, 39, 40, 41, 42, 43]. This was observed in [64] in the context of holomorphic spheres. Necessary modifications for the usage of holomorphic discs instead are worked out in Section 6 and 7. Special attention we pay to orientability questions. Similar to the polyfold version of the Deligne–Mumford space we introduce a Riemann moduli space of boundary un-noded stable discs with 3 ordered boundary marked points in Section 6. In Section 7 we define the relevant polyfold of stable boundary un-noded disc maps motivated by the absence of boundary disc bubbling in the Gromov compactification of the appearing moduli space of holomorphic discs.

Remark 1. Contact homology, as a formal concept, was introduced by Eliashberg–Givental–Hofer in [24] as contact-manifold-invariant having functorial properties. Symplectic cobordisms, which are directed from the negative to the positive end, induce structure preserving maps (e.g. unital) from the contact homology at the positive end to the contact homology at the negative end. To incorporate non-exact cobordisms and weak-filling boundary conditions a change of coefficients to a Novikov completion of the group ring of the second homology (or an adapted quotient thereof) of the symplectic cobordism resp. contact manifold is necessary by compactness reasons, see Bourgeois–van Koert [14, Section 1.1], Latschev–Wendl [49, Section 2], and Niederkrüger–Wendl [58, Section 2.5]. Applying this, one gets that a contact manifold with vanishing contact homology (i.e. with 1 being a boundary) cannot be weakly symplectically fillable (i.e. symplectically null-cobordant with empty negative end), as the contact homology of the empty contact manifold equals the coefficient ring and, therefore, never vanishes meaning $1 \neq 0$.

Combining Casals–Murphy–Presas' [17, Theorem 1.1] and Bourgeois–van Koert's [14, Theorem 1.3] shows that on every overtwisted contact manifold there exists a non-degenerate, defining contact form that admits a periodic Reeb orbit that bounds precisely one finite energy plane, which additionally is Fredholm regular, implying the vanishing of the contact homology. As Pardon rigorously defined contact homology in [59], this implies, as stated on [59, p. 835/6], the vanishing of the contact homology of overtwisted contact manifolds. Furthermore, after reworking [59] with group ring coefficients, this yields symplectic non-fillability even in the weak sense, i.e. Theorem 1.

Similarly and removing the word 'strong', part (i) of Theorem 4 could follow along the same line of reasoning because the vanishing of the contact homology implies the Weinstein conjecture for the underlying contact manifold. To obtain the strong Weinstein conjecture as verified in part (i) of Theorem 4 one could argue as above in the case of non-degenerate contact forms. In order to handle degenerate contact forms one could use an approximation argument as in [64, Section 6]. A filtered version of contact homology might yield the required energy, resp., action bounds.

On [59, p. 836] Pardon addresses the vanishing of contact homology in the presence of a small bordered Legendrian open book based on an idea of Bourgeois–Niederkrüger [13, p. 69]. It would be of interest, whether Pardon's approach to broken homolorphic discs with boundary via orientation local systems could remove the assumption on the second Stiefel–Whitney class, which we made in Theorem 2 and in part (ii) of Theorem 4 in order to make our approach via boundary un-noded stable holomorphic discs feasible. The Deligne–Mumford space elaborated in Section 6 might represent a first step in this direction.

2 Singular Legendrian foliations

2.1 Legendrian open books

Following [50, 56] we define:

Definition 2.1.1. *A* **relative open book decomposition** (B, ϑ) *of a connected manifold* N *with boundary* ∂N *consists of*

- *a non-empty codimension* 2 *submanifold* B *of* $\operatorname{Int} N$*, called the* **binding***,*
- *and a smooth, locally trivial fibration* $\vartheta\colon\ N \setminus B \to S^1$*, whose fibres* $\vartheta^{-1}(\theta)$*,* $\theta \in S^1$*, are called the* **pages***,*

such that the following conditions are satisfied:

(i) All pages intersect ∂N *transversally.*

(ii) The binding B *has a trivial tubular neighbourhood* $B \times D^2$ *in* N *in which* ϑ *is given by the angular coordinate in the* D^2*-factor.*

The pages $\vartheta^{-1}(\theta)$ in $N \setminus B$ are co-oriented by the orientation of S^1, i.e. the linearisation $T\vartheta$ maps positive normal vectors to positive tangent vectors of S^1.

As in [50, Section 4] and [56, Section I.4] we define:

Definition 2.1.2. *A connected compact* n*-dimensional submanifold* N *with boundary* ∂N *of a* $(2n-1)$*-dimensional contact manifold* (M, ξ) *carries a* **bordered Legendrian open book** (B, ϑ) *if* (B, ϑ) *is a relative open book decomposition of* N *such that*

- *the pages of* (B, ϑ) *are Legendrian submanifolds of* (M, ξ) *and*
- *the* **singular set** *of* $N \subset (M, \xi)$*, i.e. the set of all points* $p \in N$ *such that* $T_pN \subset \xi_p$*, is equal to* $B \cup \partial N$*.*

In particular, the binding of a bordered Legendrian open book is an isotropic submanifold of (M, ξ); the boundary ∂N is Legendrian. The complement

$$N^* := N \setminus \big(B \cup \partial N\big)$$

of $B \cup \partial N$ in N is the set of **regular points** of $N \subset (M, \xi)$, i.e. the set of all points $p \in N$ such that T_pN and ξ_p intersect transversally. The **characteristic distribution** $TN^* \cap \xi$ integrates by the Frobenius theorem to the so-called **characteristic foliation** on N^*. The **characteristic leaves**, which by definition are the leaves of the characteristic foliation, coincide with the pages of the open book (B, ϑ).

If, in addition, ξ is co-oriented, then $\xi|_{N^*}$ puts a co-orientation to the pages $\vartheta^{-1}(\theta)$ in $N \setminus B$. We will assume that this co-orientation coincides with the co-orientation induced by ϑ by possibly composing ϑ with a reflection on $S^1 = \partial D$.

It follows from [56, Theorem I.1.3] or [44, Theorem 1.4] that the germ of a contact structure (M, ξ) is unique near a submanifold $N \subset (M, \xi)$ (with boundary ∂N) that carries a bordered Legendrian open book (B, ϑ). The germ is uniquely determined by the **singular characteristic distribution** $\xi \cap TN$ given by the open book decomposition on N determined by (B, ϑ).

A bordered Legendrian open book (B, ϑ) is called **small** if the supporting submanifold $N \subset (M, \xi)$ is contained in a ball inside M.

Example 2.1.3. (A non-spin bordered Legendrian open book) In [50, Proposition 5.9] examples of contact manifolds (M, ξ) are constructed that contain a submanifold N, which carries a small bordered Legendrian open book. Some of the in [50, Proposition 5.9] constructed examples are indeed non-spin. In order to see this, we repeat the essential construction steps here.

The construction starts with a cylindrical Lagrangian submanifold L of an ideal Liouville domain V with disconnected boundary $\partial V = \partial_+ V \cup \partial_- V$ (see [50, Theorem C]) such that L has disconnected boundary $\partial L = \partial_+ L \cup \partial_- L$ with $\partial_\pm L \subset \partial_\pm V$. A perturbation of $L \times S^1$ inside the interior of the circular contactisation $V \times S^1$ of the ideal Liouville domain V – a so-called Giroux domain – followed by a contact cut along $\partial_- V \times S^1$ (see [50, Section 5.1]), say, yields a bordered Legendrian open book N. Gluing a Giroux domain along $\partial_+ V \times S^1$ with [50, Lemma 5.1] and eventually cutting remaining boundary components yields a contact embedding of N into a closed contact manifold (M, ξ).

In the process, L is the result of Polterovich surgery (see [60]) along, say, two transverse intersection points of a Hamiltonian deformation of two boundary parallel Lagrangian discs. If the dimension of L is even, then the Polterovich surgery result L necessarily is homotopy equivalent to a n-dimensional Klein bottle with two points removed, see [60, Paragraph 7]. Therefore, N^* is not orientable with $w_2(N^*) = 0$. If the dimension of L is odd, one can choose orientations such that L is homotopy equivalent to a n-dimensional Klein bottle or to $S^1 \times S^{n-1}$ each time with two points removed.

2.2 Overtwistedness

A $(2n-1)$-dimensional contact manifold (M, ξ) is called **overtwisted**, if (M, ξ) contains an overtwisted disc, see [11]. For example $\mathbb{R}^3$ equipped with the contact structure $\ker \alpha_{\mathrm{ot}}$,

$$\alpha_{\mathrm{ot}} := \cos r \, \mathrm{d}z + r \sin r \, \mathrm{d}\theta \, ,$$

is overtwisted, as $D^2_{\mathrm{ot}} := \{z = 0, r \leq \pi\}$ is an overtwisted disc.

For a $(n-2)$-dimensional closed smooth manifold Q we consider the contact manifold $\mathbb{R}^3 \times T^*Q$ equipped with contact structure $\xi_Q := \ker(\alpha_{\mathrm{ot}} + \lambda_{T^*Q})$, where we identify Q with the zero section in T^*Q and denote the Liouville 1-form of T^*Q by λ_{T^*Q}. Following [11, Section 10] we define the **model plastikstufe with core** Q to be the subset $P_Q := D^2_{\mathrm{ot}} \times Q$ of $\left(\mathbb{R}^3 \times T^*Q, \xi_Q\right)$.

We will say that (P_Q, ξ_Q) admits a contact embedding into a $(2n-1)$-dimensional contact manifold (M, ξ) if a neighbourhood of P_Q in $\left(\mathbb{R}^3 \times T^*Q, \xi_Q\right)$ does. In this case the image N of the model P_Q is called a **plastikstufe with core** Q and carries the structure of a bordered Legendrian open book with binding Q and pages corresponding to $I_\theta \times Q$, where I_θ is the straight line segment in $D^2_{\mathrm{ot}} \subset \mathbb{R}^2$ connecting 0 and $\pi e^{i\theta}$. For the original definition of a plastikstufe and the relation to bordered Legendrian open books we refer to [55, 56].

Theorem 2.2.1 (Borman–Eliashberg–Murphy [11])**.** *Let Q be a $(n-2)$-dimensional closed smooth manifold, whose complexified tangent bundle is trivial. Then any $(2n-1)$-dimensional overtwisted contact manifold admits a small plastikstufe with core Q.*

The assumption on Q is satisfied for any stably parallelisable manifold Q, cf. [9, Section 1.1]. For the converse of Theorem 2.2.1 we note:

Theorem 2.2.2 (Huang [45])**.** *If a contact manifold* (M, ξ) *contains a plastikstufe, then* (M, ξ) *is overtwisted.*

A forerunner version of this result, which is [45, Theorem 1.2], was given in [17]. Further, it is shown in [45, Theorem 1.3] that if (M, ξ) admits a bordered Legendrian open book and $\dim M = 5$, then (M, ξ) is overtwisted. In fact, a contact manifold (M, ξ) is overtwisted precisely if (M, ξ) contains a bordered Legendrian open book with pages diffeomorphic to $P \times \Sigma$, where P is a closed manifold and Σ a compact surface with boundary, see [45, Corollary 1.4].

2.3 Local model near the binding

Let (M, ξ) be a contact manifold. We consider a bordered Legendrian open book decomposition (B, ϑ) of a submanifold $N \subset (M, \xi)$. By Definition 2.1.1 the binding $B \subset N$ admits a tubular neighbourhood $B \times D^2$ on which the fibre projection ϑ is given by $(b, z = re^{i\theta}) \longmapsto \theta$.

By [55, Proposition 4] a neighbourhood of $B \times D^2$ in (M, ξ) is contactomorphic to a neighbourhood of $\{0\} \times D^2 \times B$ in $\big(\mathbb{R} \times \mathbb{C} \times T^*B, \ker \alpha_o\big)$, where

$$\alpha_o := \mathrm{d}t + \tfrac{1}{2}\big(x\mathrm{d}y - y\mathrm{d}x\big) + \lambda_{T^*B} \,,$$

denoting by $t, z = x + iy$ the coordinates on $\mathbb{R} \times \mathbb{C}$ and by λ_{T^*B} the Liouville 1-form on T^*B. The contactomorphism restricts to $(b, z) \longmapsto (0, z, b)$ on $B \times D^2$. Again we identify B with the zero section in T^*B.

2.4 Local model near the boundary

Consider a contact manifold (M, ξ). Let $N \subset (M, \xi)$ be a submanifold that supports a bordered Legendrian open book (B, ϑ). By Definition 2.1.1 the restriction of ϑ to ∂N induces a locally trivial fibration over S^1 with fibre F. Denoting the monodromy diffeomorphism by $\varphi\colon\ F \to F$ this fibration is equivalent to the mapping torus

$$M(\varphi) = \frac{[0,2\pi]\times F}{(2\pi, f) \sim \big(0, \varphi(f)\big)}$$

of φ. The induced diffeomorphism

$$\varphi^* := (T\varphi^{-1})^*\colon\ T^*F \to T^*F$$

naturally preserves the Liouville 1-form λ_{T^*F} so that the mapping torus

$$M(\varphi^*) = \frac{T^*[0,2\pi]\times T^*F}{\big((r,2\pi),u\big) \sim \big((r,0),(T\varphi^{-1})^*(u)\big)}$$

carries the Liouville form $r\mathrm{d}\theta + \lambda_{T^*F}$ and can be identified with $T^*\big(M(\varphi)\big)$. The corresponding Liouville vector field is of the form $r\partial_r + Y_{T^*F}$ and $M(\varphi^*)$ fibres naturally over $(T^*S^1, r\mathrm{d}\theta)$ with fibre projection map $[(r,\theta),u] \longmapsto [(r,\theta)]$.

We equip $\mathbb{R}\times M(\varphi^*)$ with the contact form α_φ induced by

$$\alpha_\varphi \equiv \mathrm{d}t + r\mathrm{d}\theta + \lambda_{T^*F}\ .$$

By [50, Lemma 4.6] a neighbourhood of $\partial N \subset (M,\xi)$ is contactomorphic to a neighbourhood of

$$\{0\}\times M(\varphi) \subset \big(\mathbb{R}\times M(\varphi^*), \ker\alpha_\varphi\big)\ ,$$

so that

$$\big(\{0\}\times\{r=0, u=0\}\big) \equiv \{0\}\times M(\varphi)$$

corresponds to ∂N and a neighbourhood of ∂N in $N\subset M$ corresponds to the quotient of the set $\{0\}\times\{r\le 0, u=0\}$ in $\{0\}\times M(\varphi^*)$. We orient $T^*S^1 \equiv \mathbb{R}\times S^1$ by $\mathrm{d}r\wedge\mathrm{d}\theta$. This matches the co-orientation conventions for the singular distribution determined by ξ and the pages of (B,ϑ) in Section 2.1.

3 Holomorphic discs

3.1 A germ of Bishop disc filling

Motivated by Section 2.3 we define a natural almost complex structure J on $\mathbb{R}\times(\mathbb{R}\times\mathbb{C}\times T^*B)$ that allows a lifting of obvious holomorphic discs similar to cf. [30, Section 2].

For that choose a Riemannian metric g_B on B. Denote by J_{T^*B} the almost complex structure on T^*B that is induced by the Levi-Civita connection of g_B, see [55, Appendix B] or [48, Section 5]. Observe that J_{T^*B} is compatible with the symplectic form $\mathrm{d}\lambda_{T^*B}$. Furthermore denoting by $g_B^\flat$ the dual metric of g_B the kinetic energy function on T^*B is defined by $k(u)=\frac{1}{2}g^\flat(u,u)$, $u\in T^*B$, and satisfies $\lambda_{T^*B}=-\mathrm{d}k\circ J_{T^*B}$. In other words, k is a strictly plurisubharmonic potential in the sense of [28, Section 3.1].

On the Liouville manifold

$$(V,\lambda_V):=\Big(\mathbb{C}\times T^*B,\frac{1}{2}\big(x\mathrm{d}y-y\mathrm{d}x\big)+\lambda_{T^*B}\Big)$$

we consider the almost complex structure

$$J_V=\mathrm{i}\oplus J_{T^*B}\ ,$$

which is compatible with the symplectic form $\mathrm{d}\lambda_V$. The function

$$\psi(z,u):=\frac{1}{4}|z|^2+k(u)$$

is a strictly plurisubharmonic potential ψ on (V,λ_V,J_V) satisfying $\lambda_V=-\mathrm{d}\psi\circ J_V$.

The contactisation $(\mathbb{R}\times V,\mathrm{d}t+\lambda_V)$ of (V,λ_V) is given by $(\mathbb{R}\times\mathbb{C}\times T^*B,\alpha_o)$ and the contact structure $\ker\alpha_o$ is spanned by vectors of the form $v-\lambda_V(v)\partial_t$, $v\in TV$. Using coordinates s on the first $\mathbb{R}$-factor of $\mathbb{R}\times(\mathbb{R}\times V)$ we define J by requiring $J(\partial_s)=\partial_t$ and

$$J\big(v-\lambda_V(v)\partial_t\big)=J_Vv-\lambda_V(J_Vv)\partial_t$$

for all $v\in TV$. In other words, J is a s-translation invariant almost complex structure on $\mathbb{R}\times(\mathbb{R}\times V)$ that preserves the contact distributions $\ker\alpha_o$ on all slices $\{s\}\times\mathbb{R}\times V$.

Remark 3.1.1. The form $\mathrm{d}(s\alpha_o) = \mathrm{d}s \wedge \alpha_o + s\mathrm{d}\alpha_o$ is symplectic on $\{s > 0\}$ and compatible with J. Therefore, the function $(s,t,z,u) \longmapsto \frac{1}{2}s^2$ is strictly plurisubharmonic on $\{s > 0\}$ because $\alpha_o = -\mathrm{d}s \circ J$.

By [55, Proposition 5] the **Niederkrüger map**

$$\Phi(s,t,z,u) = \big(s - \psi(z,u) + \mathrm{i}t, z, u\big)$$

is a biholomorphic map

$$\Phi\colon\ (\mathbb{R}\times\mathbb{R}\times\mathbb{C}\times T^*B, J) \longrightarrow (\mathbb{C}^2\times T^*B, \mathrm{i}\oplus J_{T^*B})\ ,$$

which maps the hypersurface $\{s = 0\}$ onto $\{s\circ\Phi^{-1} = 0\} = \{x_1 = -\psi(z,u)\}$. As in [55, Proposition 3.2] we consider a $(n-1)$-dimensional family of holomorphic discs

$$\{-\varepsilon^2\}\times\mathbb{D}_{2\varepsilon}\times\{b\}$$

in $(\mathbb{C}^2\times T^*B, \mathrm{i}\oplus J_{T^*B})$ with parameters $\varepsilon\in\mathbb{R}^+$, $b\in B$. Here, we denote by $\mathbb{D}_r\subset\mathbb{C}$ the closed disc with radius r and centre 0. Writing $\mathbb{D}$ for the closed unit disc $\mathbb{D}_1$ the disc family can be parametrised by

$$v_{\varepsilon,b}(z) = \big(-\varepsilon^2, 2\varepsilon\cdot z, b\big)$$

for $z\in\mathbb{D}$. The lifts $u_{\varepsilon,b} = \Phi^{-1}\circ v_{\varepsilon,b}$ via the Niederkrüger map are holomorphic maps

$$(\mathbb{D},\partial\mathbb{D}) \longrightarrow \Big(\mathbb{R}\times\mathbb{R}\times\mathbb{C}\times T^*B, \{0\}\times\{0\}\times\mathbb{C}^*\times B\Big)$$

given by

$$u_{\varepsilon,b}(z) = \Big(\varepsilon^2\big(|z|^2-1\big), 0, 2\varepsilon\cdot z, b\Big)\ .$$

We will refer to the $u_{\varepsilon,b}$ as **local Bishop discs**.

From [55, Proposition 6] we get **local uniqueness**:

Lemma 3.1.2. *For all simple J-holomorphic disc maps $u\colon\ \mathbb{D}\to\mathbb{R}\times\mathbb{R}\times\mathbb{C}\times T^*B$ such that $u(\partial\mathbb{D})\subset\{0\}\times\{0\}\times\mathbb{C}^*\times B$ there exist $\varepsilon\in\mathbb{R}^+$, $b\in B$ and a Möbius transformation $\varphi\colon\ (\mathbb{D},\partial\mathbb{D})\to(\mathbb{D},\partial\mathbb{D})$ such that $u = u_{\varepsilon,b}\circ\varphi$.*

Proof. Consider the Niederkrüger transform $v = \Phi(u)$ of u. The projection to the T^*B-factor is a J_{T^*B}-holomorphic disc with boundary on the zero section B and is therefore constant as by Stokes theorem the symplectic energy vanishes. So we are left with a holomorphic disc $v = (f + \mathrm{i}g, v_2)$ in $\mathbb{C}^2$ such that the restriction to $\partial\mathbb{D}$ satisfies $f + \frac{1}{4}|v_2|^2 = 0$ and $g = 0$, because

$$\Phi\big(\{0\} \times \{0\} \times \mathbb{C}^* \times B\big) = \big\{x_1 = -\tfrac{1}{4}|z|^2, y_1 = 0, u = 0\big\}\,.$$

The maximum and minimum principle implies that the harmonic function g vanishes identically so that f must be constant according to the classical Cauchy–Riemann equations. Write $f = -\varepsilon^2$ for some $\varepsilon > 0$. Observe, that ε indeed cannot vanish as a constant holomorphic disc is never simple. Hence, $\frac{1}{2\varepsilon}v_2$ is a holomorphic self-map of $(\mathbb{D}, \partial\mathbb{D})$. Again by the simplicity assumption the degree of the restriction of $\frac{1}{2\varepsilon}v_2$ to the boundary must be 1. The argument principle implies that $\frac{1}{2\varepsilon}v_2$ is an automorphism which is given by a Möbius transformation. □

We remark that the boundary circles

$$u_{\varepsilon,b}(\partial\mathbb{D}) = \{0\} \times \{0\} \times \partial\mathbb{D}_{2\varepsilon} \times \{b\}$$

of the local Bishop discs foliate $\{0\}\times\{0\}\times\mathbb{C}^*\times B$. Recall that a neighbourhood of $\{0\}\times\{0\}\times\{0\}\times B$ corresponds to a neighbourhood of the binding B in N^*.

3.2 Pseudo-convexity

We consider a $2n$-dimensional almost complex manifold (W, J) with non-empty boundary. Denote by M a boundary component $M \subset \partial W$ and by $\xi = TM \cap JTM$ the J-invariant hyperplane distribution along M. Assume that there exists a smooth 1-form α on M such that $\xi = \ker\alpha$ and that $\mathrm{d}\alpha$ is J-**positive** on complex lines in ξ in the sense that $\mathrm{d}\alpha(v, Jv) > 0$ for all $v \in \xi$, $v \neq 0$. In other words, (M, ξ) is a J-**convex hypersurface** of (W, J) as defined in [23].

In this situation, ξ is a contact structure with contact form α, cf. Remark 4.1.1. Denoting by R the Reeb vector field of α we additionally assume that $-JR$ is outward pointing. In other words, (M,ξ) is a J-**convex boundary component** of (W,J). Observe, that (W,J) is naturally oriented by the n-th power of any J-positive (and hence non-degenerate) 2-form on W. Therefore, the contact orientation $\alpha\wedge(\mathrm{d}\alpha)^{n-1}$ on M and the boundary orientation on $M\subset\partial W$ coincide.

Example 3.2.1. The almost complex manifold $\big((0,1]\times\mathbb{R}\times\mathbb{C}\times T^*B,J\big)$ as constructed in Section 3.1 has J-convex boundary with strictly plurisubharmonic function $\frac{1}{2}s^2$, see Remark 3.1.1.

In fact, if M is compact, then $M\subset\partial W$ is the regular zero-level set of a strictly plurisubharmonic function ϱ defined near M that is negative on the complement of M, see [19, Lemma 2.7]. The precomposition $\varrho\circ u$ with a holomorphic map $u\colon\; G\to(W,J)$, $G\subset\mathbb{C}$ open domain, is subharmonic where defined, and therefore satisfies the strong **maximum principle**. So, for example, non-constant holomorphic spheres in (W,J) are uniformly bounded away from the boundary component $M\subset\partial W$, which we assumed to be compact. Furthermore, in the case of a non-constant J-holomorphic disc $u\colon\;(\mathbb{D},\partial\mathbb{D})\to(W,M)$, we get $u(\operatorname{Int}\mathbb{D})\subset W\setminus M$ and the radial derivative

$$0<\mathrm{d}\big(\varrho\circ u\big)(\partial_r)=-\big(\mathrm{d}\varrho\circ J\big)\Big(Tu(\partial_\theta)\Big)$$

is positive along $\partial\mathbb{D}$ by the **boundary lemma of E. Hopf**. Because

$$\xi=\ker(\mathrm{d}\varrho)\cap\ker\big(-\mathrm{d}\varrho\circ J\big)$$

is co-oriented by the Reeb vector field of any contact form defining ξ as a co-oriented hyperplane distribution this means that the curve $u(\partial\mathbb{D})\subset M$ is an immersion **positively transverse** to ξ. In particular, any such holomorphic disc u such that $u|_{\partial\mathbb{D}}$ is an embedding must be simple, see [27, Lemma 4.5].

3.3 Uniqueness of the germ near the binding

Consider a $2n$-dimensional almost complex manifold (W,J) that has a compact J-convex boundary component (M,ξ) as described in Section 3.2. Assume that (M,ξ) contains a submanifold N supporting a bordered Legendrian open book (B,ϑ). By Section 2.3 a neighbourhood of the binding $B \subset W$ is diffeomorphic to a neighbourhood of $\{0\}\times\{0\}\times\{0\}\times B$ in $(-\infty,0]\times\mathbb{R}\times\mathbb{C}\times T^*B$ such that the restriction to the boundaries M and $\{0\}\times\mathbb{R}\times\mathbb{C}\times T^*B$ induces a contactomorphism. In addition, assume that the almost complex structure J of W corresponds to the one constructed in Section 3.1 under the diffeomorphism.

From [55, Proposition 7] we get **semi-global uniqueness**:

Lemma 3.3.1. *There exists a neighbourhood $U_B \subset W$ of B such that for all simple J-holomorphic disc maps $u\colon (\mathbb{D},\partial\mathbb{D}) \to (W,N^*)$ with $u(\mathbb{D})\cap U_B \neq \emptyset$ we have that $u(\mathbb{D})$ is contained in U_B and there exist $\varepsilon\in\mathbb{R}^+$, $b\in B$ and a Möbius transformation $\varphi\colon (\mathbb{D},\partial\mathbb{D}) \to (\mathbb{D},\partial\mathbb{D})$ such that $u = u_{\varepsilon,b}\circ\varphi$.*

Proof. Using the above diffeomorphism we describe such a neighbourhood U_B as subset of $(-\infty,0]\times\mathbb{R}\times\mathbb{C}\times T^*B$: For $x_1<0$, $y_1\in\mathbb{R}$ consider the complex hypersurfaces $\{x_1+\mathrm{i}y_1\}\times\mathbb{C}\times T^*B$ in $\big(\mathbb{C}^2\times T^*B, \mathrm{i}\oplus J_{T^*B}\big)$. The intersection with the real hypersurface $\{x_1 = -\psi(z,u)\}$ is the sphere bundle in $\underline{\mathbb{C}}\oplus T^*B$ given by $|x_1| = \frac{1}{4}|z|^2 + \frac{1}{2}g^\flat(u,u)$; the intersection with $\{x_1 \le -\psi(z,u)\}$, therefore, is the corresponding disc bundle in $\underline{\mathbb{C}}\oplus T^*B$. Here $\underline{\mathbb{C}}$ denotes the trivial complex line bundle over B. Hence, the complex hypersurfaces

$$H_{x_1,y_1} := \Phi^{-1}\Big(\{x_1+\mathrm{i}y_1\}\times\mathbb{C}\times T^*B\Big)\cap\{s\le 0\}$$

foliate the complement of $\{0\}\times\mathbb{R}\times\{0\}\times B$ in $\big((-\infty,0]\times\mathbb{R}\times\mathbb{C}\times T^*B, J\big)$ including a foliation by their real boundaries. Then, by definition, U_B corresponds to

$$U_B := \Big(\{0\}\times(-\delta,\delta)\times\{0\}\times B\Big)\cup\bigcup_{|x_1|,|y_1|<\delta} H_{x_1,y_1}$$

for $\delta>0$ sufficiently small, under the above mentioned identifying diffeomorphism.

Consider a simple J-holomorphic disc $u\colon (\mathbb{D},\partial\mathbb{D}) \to (W,N^*)$ and suppose that $G = u^{-1}(U_B)$ is not empty. Observe that $G\subset\mathbb{D}$ is open. Restricting to G we write $\Phi(u) =$

$(f+\mathrm{i}g, v_2, v_3)$. If g is constant in a neighbourhood of a point in G, then so is the holomorphic function $f+\mathrm{i}g$. With the identity theorem this implies that $f+\mathrm{i}g$ is constant on $G\setminus\partial\mathbb{D}$ and, hence, on G. Denoting the constant by $f_o+\mathrm{i}g_o$ this translates into $G=u^{-1}(H_{f_o,g_o})$ so that $G\subset\mathbb{D}$ is closed also. Hence, $G=\mathbb{D}$ by an open–closed argument, i.e. $u(\mathbb{D})\subset U_B$ and the claim follows with Lemma 3.1.2.

It remains to show that the complementary case, namely that $\{\mathrm{d}g=0\}$ has no interior points, cannot occur. Indeed, otherwise the holomorphic disc $u(\mathbb{D})$ and the complex hyperplane $H_{x_1^*,y_1^*}$ intersect along finitely many points for $x_1^*<0$, $y_1^*\neq 0$. For that observe with Section 3.2 and Example 3.2.1 (suitably shifted in the s-direction) that the intersection is along interior points of $u(\mathbb{D})$. This follows because $H_{x_1^*,y_1^*}$ and the boundary condition N^* for the holomorphic disc u are disjoint as

$$\Phi\Big(\{0\}\times\{0\}\times\mathbb{C}^*\times B\Big)\subset\mathbb{R}_-\times\{y_1=0\}\times\mathbb{C}^*\times B\,.$$

By positivity of local intersection numbers the total intersection number $u\bullet H_{x_1^*,y_1^*}$ is positive. On the other hand this total intersection number is equal to the homological intersection of $[u]\in H_2(W,M)$ and $[\partial c]\in H_{2n-2}W$, where the $(2n-1)$-chain c is given by

$$c=\bigcup_{x\in[x_1^*,0]}H_{x,y_1^*}\,.$$

Indeed, the maximum principle implies that $u(\operatorname{Int}\mathbb{D})$ does not intersect M so that $u(\mathbb{D})$ and $\partial c\cap M$ are disjoint. Hence,

$$u\bullet H_{x_1^*,y_1^*}=[u]\cdot[\partial c]=0\,.$$

This is a contradiction. In other words, $\{\mathrm{d}g=0\}$ has to have an interior point. □

3.4 Holomorphic model near the boundary

As in Section 3.1 we define an almost complex structure on $M(\varphi^*)$: Choose a Riemannian metric g_θ on F that smoothly depends on $\theta\in\mathbb{R}$ such that $g_{2\pi}=g_\theta=\varphi^*g_{\theta-2\pi}$ for all $\theta\in(2\pi-\varepsilon,2\pi+\varepsilon)$ and $\varepsilon>0$ small. For each $\theta\in\mathbb{R}$ we define the kinetic energy on T^*F by

$$k_\theta(u)=\frac{1}{2}g_\theta^\flat(u,u)\,,\quad u\in T^*F\,,$$

denoting by $g_\theta^\flat$ the dual metric of g_θ. For each $\theta \in [0, 2\pi]$ we construct an almost complex structure J_θ on T^*F as on [48, Section 5.3] that is compatible with $\mathrm{d}\lambda_{T^*F}$ and turns k_θ into a strictly plurisubharmonic potential on T^*F in the sense of [28, Section 3.1] meaning that

$$\lambda_{T^*F} = -\mathrm{d}k_\theta \circ J_\theta \ .$$

The almost complex structure J_θ is uniquely determined by g_θ and $\mathrm{d}\lambda_{T^*F}$; therefore, J_θ depends smoothly on θ and satisfies $J_{2\pi} = J_\theta = \varphi^* J_{\theta-2\pi}$ for $\theta \in (2\pi - \varepsilon, 2\pi + \varepsilon)$.

The metric on $[0, 2\pi] \times F$ obtained by taking the sum of the Euclidean metric and g_θ for each $\theta \in [0, 2\pi]$ descents to a metric g on the mapping torus $M(\varphi)$. The induced kinetic energy function on $M(\varphi^*)$ is denoted by

$$\psi(r, \theta, u) = \frac{1}{2} r^2 + k_\theta(u) \ .$$

The almost complex structure $\mathrm{i} \oplus J_\theta$ on $T^*[0, 2\pi] \times T^*F$ descents to a compatible almost complex structure J_g on $M(\varphi^*)$ and

$$r\mathrm{d}\theta + \lambda_{T^*F} \equiv -\mathrm{d}\psi \circ J_g$$

modulo second order terms in $|u|_\theta$, $u \in T^*F$, in the sense of [54, Appendix E.5]. Indeed, the derivative of $k_\theta(u)$ in θ-direction contributes a term that locally is a quadratic form in the coordinates of $u \in T^*F$. Consequently, the restriction of $\mathrm{d}r \wedge \mathrm{d}\theta + \mathrm{d}\lambda_{T^*F}$ to $\{u = 0\}$ in $M(\varphi^*)$ is equal to $-\mathrm{d}\big(\mathrm{d}\psi \circ J_g\big)$. As $\mathrm{d}r \wedge \mathrm{d}\theta + \mathrm{d}\lambda_{T^*F}$ is positive on $(\mathrm{i} \oplus J_\theta)$-complex lines and equals $-\mathrm{d}\big(\mathrm{d}\psi \circ J_g\big)$ modulo first order terms in $|u|_\theta$, $u \in T^*F$, we conclude that ψ is strictly plurisubharmonic in a neighbourhood of $\{u = 0\}$ in $M(\varphi^*)$.

We consider the almost complex manifold $(\mathbb{C} \times M(\varphi^*), \mathrm{i} \oplus J_g)$ provided with the Liouville form $s\mathrm{d}t + r\mathrm{d}\theta + \lambda_{T^*F}$, where we denote the coordinates on $\mathbb{C}$ by $s + \mathrm{i}t$. Observe that the almost complex structure

$$J = \mathrm{i} \oplus J_g$$

is compatible with $\mathrm{d}s \wedge \mathrm{d}t + \mathrm{d}r \wedge \mathrm{d}\theta + \mathrm{d}\lambda_{T^*F}$ and that the function

$$\Psi(s + \mathrm{i}t, r, \theta, u) = \frac{1}{2} s^2 + \psi(r, \theta, u)$$

is strictly plurisubharmonic in a neighbourhood of $\{u=0\}$ in $\mathbb{C}\times M(\varphi^*)$. This holds because

$$s\mathrm{d}t + r\mathrm{d}\theta + \lambda_{T^*F} \equiv -\mathrm{d}\Psi\circ J$$

modulo second order terms in $|u|_\theta$, $u \in T^*F$. Therefore, as above, the restriction of $\mathrm{d}s\wedge\mathrm{d}t + \mathrm{d}r\wedge\mathrm{d}\theta + \mathrm{d}\lambda_{T^*F}$ to $\{u=0\}$ in $\mathbb{C}\times M(\varphi^*)$ is equal to $-\mathrm{d}\big(\mathrm{d}\Psi\circ J\big)$.

By rescaling the metric g_θ on F by a constant if necessary we can assume that Ψ is strictly plurisubharmonic in a neighbourhood of $\{\Psi=\frac{1}{2}\}$. The hypersurfaces $\{\Psi=\frac{1}{2}\}$ and $\{s=1\}$ are transverse to $s\partial_s + r\partial_r + Y_{T^*F}$, which is a Liouville vector field w.r.t. the symplectic form $\mathrm{d}s\wedge\mathrm{d}t + \mathrm{d}r\wedge\mathrm{d}\theta + \mathrm{d}\lambda_{T^*F}$. Therefore, the contraction into the symplectic form induces contact forms on both hypersurfaces. The induced contact form on $\{s=1\}=\{1\}\times\mathrm{i}\mathbb{R}\times M(\varphi^*)$ is α_φ; the one on $\{\Psi=\frac{1}{2}\}$ is given by $-\mathrm{d}\Psi\circ J$ along $\{u=0\}$.

A reparametrisation of the flow of the Liouville vector field $s\partial_s + r\partial_r + Y_{T^*F}$ yields a contact embedding of $\{\Psi=\frac{1}{2}, s>0\}$ onto $\{s=1\}$ w.r.t. the induced contact structures, see [10, Appendix A.1]. The hypersurfaces $\{\Psi=\frac{1}{2}\}$ and $\{s=1\}$ intersect along $\{s=1, r=0, u=0\}$, on which the flow is stationary. Moreover, as we flow along the Liouville vector field $s\partial_s + r\partial_r + Y_{T^*F}$ we observe that the multi level set $\{\Psi=\frac{1}{2}, s>0, t=0, r\leq 0, u=0\}$ corresponds to $\{1\}\times\{0\}\times\{r\leq 0, u=0\}$ in $\{1\}\times\mathrm{i}\mathbb{R}\times M(\varphi^*)$ under the contact embedding. The latter set was used in Section 2.4 to describe the germ of contact structure near the boundary of a bordered Legendrian open book; the boundary being $\{\Psi=\frac{1}{2}, s>0, t=0, r=0, u=0\}$, which corresponds to $\{1\}\times\{0\}\times\{r=0, u=0\}$.

The hypersurface $\{\Psi=\frac{1}{2}\}$ carries a second contact structure given by

$$\ker\big(\mathrm{d}\Psi\big)\cap\ker\big(-\,\mathrm{d}\Psi\circ J\big)\,,$$

which turns $\{\Psi=\frac{1}{2}\}$ into a J-convex boundary of $\{\Psi\leq\frac{1}{2}\}$. The induced singular characteristic foliation on $\{\Psi=\frac{1}{2}, s>0, t=0, r\leq 0, u=0\}$ coincides with the one described in the preceding paragraph, where the contact distribution this time is taken w.r.t. $s\mathrm{d}t + r\mathrm{d}\theta + \lambda_{T^*F} \equiv -\mathrm{d}\Psi\circ J$ along $\{u=0\}$. By uniqueness of the germ of a contact structure formulated in [50, Lemma 4.6], a neighbourhood of $\{\Psi=\frac{1}{2}, s>0, t=0, r\leq 0, u=0\}$ is contactomorphic to the contact structure we considered first. In other words, given a contact manifold (M,ξ) containing a submanifold N that supports a bordered Legendrian open book, we obtain an alternative description of the germ of contact structure near the boundary ∂N presented in Section 2.4.

3.5 Holomorphically blocking boundary

Let (W,J) be a $2n$-dimensional almost complex manifold so that a given contact manifold (M,ξ) is a compact J-convex boundary component of W, see Section 3.2. Let N be a submanifold of (M,ξ) that supports a bordered Legendrian open book (B,ϑ). In view of Section 3.4 we assume that a neighbourhood of the boundary $\partial N \subset W$ is diffeomorphic to a neighbourhood of $\{1\}\times\{0\}\times\{r=0,u=0\}$ in $\{\Psi\leq\frac{1}{2},s>0\}$ such that $N\subset W$ and $\{\Psi=\frac{1}{2},s>0,t=0,r\leq 0,u=0\}$ correspond to each other diffeomorphically. Furthermore we assume that under the diffeomorphism the almost complex structure J of W corresponds to $\mathrm{i}\oplus J_g$ on $\{\Psi\leq\frac{1}{2}\}$ in $\mathbb{C}\times M(\varphi^*)$ inducing contactomorphisms on the boundary.

Similarly to [50, Lemma 4.7] we obtain the **blocking lemma**:

Lemma 3.5.1. *There exists a neighbourhood $U_{\partial N}\subset W$ of ∂N such that for all J-holomorphic disc maps $u\colon\ (\mathbb{D},\partial\mathbb{D})\to(W,N^*)$ with $u(\mathbb{D})\cap U_{\partial N}\neq\emptyset$ are constant.*

Proof. For $|s|<1$, $t\in\mathbb{R}$ consider the complex hypersurfaces

$$H_{s,t}\;:=\Big(\{s+\mathrm{i}t\}\times M(\varphi^*)\Big)\cap\{\Psi\leq\tfrac{1}{2}\}$$

of $\{\Psi\leq\frac{1}{2}\}$ in $(\mathbb{C}\times M(\varphi^*),\mathrm{i}\oplus J_g)$. Modulo the above identifying diffeomorphism $U_{\partial N}$ corresponds to

$$U_{\partial N}\;:=\Big(\{1\}\times(-\delta,\delta)\times M(\varphi)\Big)\cup\bigcup_{1-s,|t|<\delta}H_{s,t}$$

for $\delta>0$ sufficiently small.

Let $u\colon\ (\mathbb{D},\partial\mathbb{D})\to(W,N^*)$ be a J-holomorphic disc such that the open set $G=u^{-1}(U_{\partial N})$ is not empty. Write the restriction of u to G as $u=(u_1=f+\mathrm{i}g,u_2)$ w.r.t. $(\mathbb{C}\times M(\varphi^*),\mathrm{i}\oplus J_g)$. An argument similar to the last paragraph of the proof of Lemma 3.3.1 (that utilises J-convexity and positivity of intersections with the complex hyperplanes $H_{s,t}$) shows that

$\{\mathrm{d}g = 0\}$ has an interior point. As in the second paragraph of the proof of Lemma 3.3.1 this shows that $u_1 = f_o + \mathrm{i}g_o$ is constant and that $G = \mathbb{D}$, i.e. $u(\mathbb{D}) \subset U_{\partial N}$.

The projection map $M(\varphi^*) \to T^*S^1$ sends u_2 to a smooth map $v\colon \mathbb{D} \to T^*S^1$ such that $v(\partial\mathbb{D}) \subset \{r \leq 0\} \simeq S^1$. In particular, the degree of $v|_{\partial\mathbb{D}}$ must be zero so that $v(\partial\mathbb{D})$ is tangent to a fibre of T^*S^1. Therefore, $u(\partial\mathbb{D})$ admits a point of tangency with a page of the Legendrian open book on N. In view of the maximum principle by E. Hopf, which implies the positive transversality property formulated in Section 3.2, this implies that u must be constant. □

4 Tamed pseudo-convexity

Let (W, J) be a $2n$-dimensional almost complex manifold that admits a compact J-convex boundary component (M, ξ), see Section 3.2. We assume that there exists a symplectic form Ω on W that is J-positive on complex lines in TW, i.e. J is tamed by Ω, cf. [23]. Define an odd-symplectic form on M by setting $\omega := \Omega|_{TM}$.

4.1 Magnetic symplectisation

Let α be a defining contact form for ξ on M such that $\mathrm{d}\alpha$ is positive on complex lines in (ξ, J).

Remark 4.1.1. As ω is positive on complex lines in (ξ, J), it follows that for all non-zero $v \in \xi$ the contraction $\iota_v\omega$ does not vanish. In other words, ω restricts to a symplectic form on ξ. Choosing a symplectic basis for (ξ, ω) of the form $v_1, Jv_1, \ldots, v_{n-1}, Jv_{n-1}$ yields that ω^{n-1} is a positive volume form on (ξ, J). Furthermore the contact form α does not vanish on the characteristic foliation given by $\ker\omega$, i.e. ξ and $\ker\omega$ intersect transversally. In fact, according to our orientation conventions, $\alpha \wedge \omega^{n-1}$ is a positive volume form on M.

Observe, that the same reasoning applies to all 2-forms that are positive on complex lines in (ξ, J). This shows positivity of $\alpha \wedge (\mathrm{d}\alpha)^{n-1}$ which was used implicitly in Section 3.2.

The symplectic neighbourhood theorem for hypersurfaces implies that there exist

$\varepsilon > 0$ and a symplectic embedding

$$\Big((-\varepsilon, 0] \times M, \mathrm{d}(s\alpha) + \omega\Big) \longrightarrow (W, \Omega)$$

whose restriction to $\{0\} \times M$ is the inclusion of $M = \partial W$ into W, cf. [52, Exercise 3.36] and [50, Remark 2.7]. Observe that

$$\Big(\mathrm{d}(s\alpha) + \omega\Big)^n = \Big(\mathrm{d}s \wedge \alpha + s\mathrm{d}\alpha + \omega\Big)^n = n\mathrm{d}s \wedge \alpha \wedge \big(s\mathrm{d}\alpha + \omega\big)^{n-1}.$$

Therefore, we find $\varepsilon > 0$ and a large positive constant s_o such that, considered on $(-\varepsilon, \infty) \times M$, $\mathrm{d}(s\alpha) + \omega$ is symplectic on $(-\varepsilon, \varepsilon) \times M$ and $(s_o, \infty) \times M$. In fact, using Remark 4.1.1, $\alpha \wedge \big(s\mathrm{d}\alpha + \omega\big)^{n-1}$ is a positive volume form on M for all positive s because $\mathrm{d}\alpha$ and ω are positive on complex lines in (ξ, J). Therefore, $\mathrm{d}(s\alpha) + \omega$ is a symplectic form on $(-\varepsilon, \infty) \times M$. Via gluing along $\{0\} \times M \equiv M \subset \partial W$ using the above symplectic embedding we build a symplectic manifold, the so-called **magnetic completion**,

$$(\widetilde{W}, \widetilde{\Omega}) := (W, \Omega) \cup \Big([0, \infty) \times M, \mathrm{d}(s\alpha) + \omega\Big).$$

4.2 Truncating the magnetic completion

We continue the considerations in Section 4.1. In addition, let $N \subset (M, \xi)$ be a submanifold that carries the structure of a bordered Legendrian open book decomposition (B, ϑ). Choose contact embeddings of the model neighbourhood of the binding B (see Section 2.3) and of the alternative model neighbourhood of the boundary ∂N (see Section 3.4) into $(M, \xi = \ker \alpha)$. The push forward of the respective contact forms α_o and the restriction of $-\mathrm{d}\Psi \circ J$ to the tangent spaces of $\{\Psi = \frac{1}{2}\}$ are equal to $\mathrm{e}^h \alpha$ for a smooth function h on the images of the model neighbourhoods. Alter the contact form α by cutting off h to 0 near the boundary of these images, so that the considered contact embeddings – after shrinking the model neighbourhoods a bit – are in fact strict. Observe that this conformal change of the contact form α does not effect the property of $\mathrm{d}\alpha$ to be positive on complex lines in (ξ, J).

Let s_o be a positive real number and consider the truncation

$$\big(\hat{W}, \hat{\Omega}\big) := (W, \Omega) \cup \Big([0, 2s_o] \times M, \mathrm{d}(s\alpha) + \omega\Big) .$$

The resulting contact embeddings of the model neighbourhoods into $\{2s_o\} \times M$, which are strict up to conformal factor $2s_o$ w.r.t. the contact form $2s_o\alpha$, extend along collar directions to embeddings of the neighbourhood U_B constructed in Lemma 3.3.1 and of the neighbourhood $U_{\partial N}$ constructed in Lemma 3.5.1 into $(-\infty, 2s_o] \times M$ in the following way: For the embedding of U_B simply cross with the identity in the s-direction; for the embedding of $U_{\partial N}$ denote by Y the vector field on $\{\Psi \leq \frac{1}{2}\}$ obtained by multiplying the Reeb vector field of the contact form $-\mathrm{d}\Psi{\circ}J$ on the level sets of Ψ with $-J$ and follow the flow lines of Y in backward time (taking time logarithmically). Observe that Y coincides with the Liouville vector field $s\partial_s + r\partial_r + Y_{T^*F}$ along $\{u = 0\}$, see Section 3.4.

The images of the neighbourhoods are again denoted by U_B and $U_{\partial N}$. Taking s_o sufficiently large we achieve that $U_B \cup U_{\partial N}$ fit into $[s_o, 2s_o] \times M$. Push forward yields an almost complex structure J_o on $U_B \cup U_{\partial N}$ that allows the conclusions of Lemmata 3.3.1 and 3.5.1. We remark that J_o is invariant under translations in s-direction that shift off U_B; along $[s_o, 2s_o] \times \partial N$ the almost complex structure J_o is independent of s.

Moreover, the symplectic form $\mathrm{d}(s\alpha)$ is compatible with J_o on U_B and on $U_{\partial N}$, see Remark 3.1.1 and Section 3.4, resp. Furthermore J_o sends ∂_s to the Reeb vector field and preserves the contact distribution. Applying [52, Proposition 2.63(i)] to the symplectic bundle

$$\big(\xi, \mathrm{d}(s\alpha)\big) \longrightarrow [s_o, 2s_o] \times M$$

we extend J_o to an almost complex structure on $[s_o, 2s_o] \times M$ keeping the properties just listed. In particular, $\{2s_o\} \times M$ is a J_o-convex boundary independently of the conformal factor $2s_o$, see Section 3.2.

Placing the whole scenario to $[s_o, 2s_o] \times M$ for s_o sufficiently large we can additionally assume that J_o is also tamed by the symplectic form $\mathrm{d}(s\alpha) + \omega$ on $[s_o, 2s_o] \times M$. This is possible because for s_o sufficiently large $\mathrm{d}(s\alpha) = \mathrm{d}s \wedge \alpha + s\mathrm{d}\alpha$ dominates $\mathrm{d}(s\alpha) + \omega = \mathrm{d}s \wedge \alpha + (s\mathrm{d}\alpha + \omega)$

on J_o-complex lines. With [52, Proposition 2.51] we extend J_o to a tamed almost complex structure $\hat{J}$ on $(\hat{W},\hat{\Omega})$ that restricts to J on W and to J_o on $[s_o,2s_o]\times M$. Moreover, $\{2s_o\}\times M\subset\partial\hat{W}$ is a $\hat{J}$-convex boundary component. We will refer to the construction of $(\hat{W},\hat{\Omega},\hat{J})$ as **magnetic collar extension**.

4.3 Deforming the Truncation

Assuming exactness of $\omega|_{TN}$ in the situation of Sections 4.1 and 4.2 the symplectic form in the magnetic collar extension $(\hat{W},\hat{\Omega},\hat{J})$ can be assumed to satisfy

$$\hat{\Omega}=2s_o\mathrm{d}\alpha\qquad\text{on}\quad T\big(\{2s_o\}\times N\big)$$

after deformation:

Write $\omega=\mathrm{d}\beta$ in a neighbourhood of N taking a neighbourhood that strongly deformation retracts to N. Extend β to a 1-form on M that vanishes outside a larger neighbourhood. Define a 2-form $\eta:=\omega-\mathrm{d}(\varrho\beta)$ on $[s_o,2s_o]\times M$, where ϱ is a smooth function that vanishes on $\{s\leq s_o\}$, equals 1 on $\{s\geq 2s_o\}$, and satisfies $0\leq\varrho'\leq 2/s_o$. We claim that for s_o sufficiently large

$$\Big([s_o,2s_o]\times M,\mathrm{d}(s\alpha)+\eta\Big)$$

is symplectic. Indeed, spelling out the n-th power of $\mathrm{d}(s\alpha)+\eta$ we find

$$n\mathrm{d}s\wedge\left(\alpha-\varrho'\beta\right)\wedge\Big(s\mathrm{d}\alpha+\omega-\varrho\mathrm{d}\beta\Big)^{n-1}.$$

For s_o sufficiently large $\alpha-\varrho'\beta$ will be a contact form on all slices $\{s\}\times M$, $s\in[s_o,2s_o]$, so that, restricted to the related contact structures, $(s\mathrm{d}\alpha)^{n-1}$ is a positive volume form. Making s_o even larger $(s\mathrm{d}\alpha)^{n-1}$ dominates lower order terms in the $(n-1)$-st power of $s\mathrm{d}\alpha+\omega-\varrho\mathrm{d}\beta$ restricted to the mentioned contact structures.

Starting with the deformed symplectic structure corresponding to $\mathrm{d}(s\alpha)+\eta$ the tamed complex structure $\hat{J}$ can be constructed as in Section 4.2. In order to achieve that $\mathrm{d}(s\alpha)$ dominates

$$\mathrm{d}(s\alpha)+\eta=\mathrm{d}s\wedge\left(\alpha-\varrho'\beta\right)+\Big(s\mathrm{d}\alpha+\omega-\varrho\mathrm{d}\beta\Big)$$

on J_o-complex lines on $[s_o, 2s_o] \times M$ simply choose s_o sufficiently large.

We will refer to the construction of $(\hat{W}, \hat{\Omega}, 2s_o\alpha, \hat{J})$ as **deformed magnetic collar extension**.

4.4 Gromov compactness

We consider a deformed magnetic collar extension of the Ω-tamed almost complex structure (W, J) as in Section 4.3. The resulting tamed almost complex manifold together with the choice of contact form on the J-convex boundary component is denoted by (W, Ω, α, J). In particular, the restriction of the 2-forms Ω and $\mathrm{d}\alpha$ to the tangent spaces of the bordered Legendrian open book $(N, B, \vartheta) \subset (M, \xi = \ker\alpha)$ are equal as 2-form on N.

Notice, that the regular set N^* of N (see Section 2.1) is a **totally real** submanifold of (W, J), i.e. $TN^* \cap JTN^*$ is the zero section. Indeed, denoting by E the real linear span of v, Jv for a given tangent vector $v \in T_pN^*$ the only possibility for the complex line E to be tangent to $N^* \subset M$ is to be tangent to the page of (B, ϑ) through p, which is Legendrian w.r.t. $\xi = TM \cap JTM$. Positivity of $\mathrm{d}\alpha$ on complex lines in (ξ, J) implies the vanishing of v because the pages of (B, ϑ) are Legendrian submanifolds, and, hence, have Lagrangian tangent spaces inside $(\xi, \mathrm{d}\alpha)$.

Via the neighbourhood $U_B \subset W$ of the binding $B \subset N$ constructed in Lemma 3.3.1 we obtain an embedding relative boundary, a so-called **local Bishop filling**,

$$F\colon \Big((0,\delta) \times \mathbb{D} \times B, (0,\delta) \times \partial\mathbb{D} \times B\Big) \longrightarrow (W, N^*)$$

for some $\delta > 0$ such that for all $(\varepsilon, b) \in (0,\delta) \times B$ the maps $u_{\varepsilon,b} = F(\varepsilon, \,.\,, b)$ are J-holomorphic discs $(\mathbb{D}, \partial\mathbb{D}) \to (W, N^*)$. Furthermore F extends smoothly to a map defined on $[0,\delta) \times \mathbb{D} \times B$ that maps $\{0\} \times \mathbb{D} \times B$ to B via $(\varepsilon, z, b) \longmapsto b$.

We consider the **moduli space** $\mathcal{M}$ of J-holomorphic discs $u\colon (\mathbb{D}, \partial\mathbb{D}) \to (W, N^*)$ that are homologous in W relative N^* to one of the Bishop discs $u_{\varepsilon,b}$. For all $u \in \mathcal{M}$ we have the following:

1. The degree of the map $\vartheta \circ u\colon \partial\mathbb{D} \to S^1$, the so-called **winding number** of u, is equal

to 1. The boundary lemma of E. Hopf implies, that $u(\partial\mathbb{D})$ is an embedded curve in N^* positively transverse to ξ. Hence, u is simple, see Section 3.2.

2. We find a 2-chain C in N with boundary $-u(\partial\mathbb{D})$ such $u(\mathbb{D})+C$ is null-homologous in W. Therefore, by applying Stokes theorem twice, we get that the **symplectic energy** of u satisfies
$$0<\int_{\mathbb{D}} u^*\Omega = -\int_C \Omega = \int_{\partial\mathbb{D}} u^*\alpha\,,$$
as $\Omega = \mathrm{d}\alpha$ on TN. Denote by $\mathrm{d}\vartheta$ the pullback of $\mathrm{d}\theta$ along $\vartheta\colon\; N\setminus B \to S^1$. According to our co-orientation convention in Section 2.1 and the local models in Sections 2.3 and 2.4 we observe that $\alpha|_{TN} = f\mathrm{d}\vartheta$ for a non-negative function f on N, that vanishes precisely along the singular set $B\cup\partial N$. Consequently, we get **uniform energy bounds**
$$0<\int_{\mathbb{D}} u^*\Omega = \int_{\partial\mathbb{D}} u^*f\cdot(\vartheta\circ u)^*\mathrm{d}\theta \leq 2\pi\max f\,.$$

3. The restriction of u to $\partial\mathbb{D}$ is **uniformly bounded away from the boundary** ∂N as the intersection of $u(\mathbb{D})$ with $U_{\partial N}$ is empty by Lemma 3.5.1. Recall that if $u(\mathbb{D})$ intersects the neighbourhood U_B of the binding then u is a reparametrisation of a local Bishop disc $u_{\varepsilon,b}$, see Lemma 3.3.1. We **truncate** the moduli space $\mathcal{M}$ (keeping the notation) by removing all holomorphic discs that are reparametrisations of $u_{\varepsilon,b}$ for $(\varepsilon,b)\in(0,\delta/2)\times B$.

Remark 4.4.1. Under the assumption of **uniform** C^0**-bounds** for $\mathcal{M}$, i.e. that there exists a compact subset of W that contains all holomorphic discs $u(\mathbb{D})$, $u\in\mathcal{M}$, we obtain: Any sequence of holomorphic discs in $\mathcal{M}$ admits a Gromov converging subsequence, see [25]. Observe that by J-convexity (see Section 3.2) no sequence of holomorphic discs $u(\mathbb{D})$, $u\in\mathcal{M}$, can escape the boundary component M of W.

The total winding number of the Gromov limit $\mathbf{u}$ must be 1 according to the properties of Gromov convergence, see [25]. Moreover, all individual winding numbers of non-constant disc bubbles of $\mathbf{u}$ are positive by positive transversality, see the boundary lemma to E. Hopf in Section 3.2. Therefore, $\mathbf{u}$ contains precisely one disc component u_0 that is necessarily simple and of winding number 1. In particular, $u_0|_{\partial\mathbb{D}}$ is an embedding positively transverse to ξ.

If the image of $\mathbf{u}$ intersects U_B, then $\mathbf{u}$ does it along the disc component u_0 by the maximum principle, see Section 3.2. This implies that u_0 is one of the Bishop discs $u_{\varepsilon,b}$ by Lemma 3.3.1, so that there are in fact no bubbles in this situation. Indeed, potential sphere bubbles must be null-homologous or subject to an argument using the maximum principle. Therefore, any sequence of holomorphic discs in $\mathcal{M}$ that Gromov converges to $\mathbf{u}$ with $u_0(\mathbb{D})$ intersecting U_B converges in C^∞ to u_0 up to reparametrisation.

Consequently, $\mathcal{M}$ can be compactified to $\overline{\mathcal{M}}$ by adding all Gromov limits to $\mathcal{M}$. The resulting moduli space $\overline{\mathcal{M}}$ is compact in the sense that any sequence in $\overline{\mathcal{M}}$ admits a Gromov converging subsequence. The resulting limit objects $\mathbf{u}$ share the properties mentioned in the proceeding remark. In particular, we obtain **uniform gradient bounds near the boundary** w.r.t. a given background metric: There exists constants $\rho \in (0,1)$ and $C > 0$ such that for all $\mathbf{u}$ in $\overline{\mathcal{M}}$ we have that $|\nabla u_0| < C$ restricted to $\mathbb{D} \cap \{|z| \geq \rho\}$, where u_0 is the disc component of $\mathbf{u}$ parametrised such that $\vartheta \circ u_0(\mathrm{i}^k) = \mathrm{i}^k$ for $k = 0, 1, 2$.

5 Symplectic cobordisms

A **symplectic cobordism** is a compact $2n$-dimensional symplectic manifold (W, Ω) with boundary $M := \partial W$. We assume that (W, Ω) is *connected* and oriented via Ω^n. The **odd-symplectic** form $\omega := \Omega|_{TM}$ is closed and maximally non-degenerate with 1-dimensional kernel $\ker\omega$. The line bundle $\ker\omega$ on ∂W is trivialised by the restriction of the Hamiltonian vector field of a collar neighbourhood parameter to $M = \partial W \subset W$. Therefore, one finds a 1-form α on M that does not vanish on $\ker\omega$. The sign of α defines an orientation on each component of ∂W via the volume form $\alpha \wedge \omega^{n-1}$. If the boundary orientation (induced by the outward pointing normal) coincides with the one given by α we call the components of $M = \partial W$ **positive**, otherwise **negative**. We will write

$$\partial(W, \Omega) = -(M_-, \omega_-, \alpha_-) + (M_+, \omega_+, \alpha_+)$$

accordingly. Due to the symplectic neighbourhood theorem there exist collar neighbourhoods $[0,\varepsilon) \times M_-$, resp., $(-\varepsilon, 0] \times M_+$, such that the symplectic form Ω can be written as $\mathrm{d}(s\alpha_\mp) + \omega_\mp$ with $s \in [0,\varepsilon)$, resp., $s \in (-\varepsilon, 0]$. This allows gluing of symplectic cobordisms along boundary components of opposite sign that are orientation reversing odd-symplectomorphic. The most prominent examples of symplectic cobordisms have **contact type** boundary, i.e. the 1-form α on M can be chosen such that $\omega = \mathrm{d}\alpha$, see [51]. In the contact type context positive boundary components are called **convex**; negative ones **concave**.

5.1 A directed symplectic cobordism

We consider a symplectic cobordism (W, Ω). We assume that the boundary $\partial(W, \Omega)$ decomposes into concave boundary components, whose union we denote by $(M_-, \omega_-, \alpha_-)$, and positive boundary components, whose union we denote by $(M_+, \omega_+, \alpha_+)$. In addition, we require the **weak-filling condition**: α_+ can be chosen to be a contact form, which according to our conventions implies that $\alpha_+ \wedge (\mathrm{d}\alpha_+)^{n-1}$ and $\alpha_+ \wedge \omega_+^{n-1}$ are positive, such that

$$\alpha_+ \wedge \left(f_+ \mathrm{d}\alpha_+ + \omega_+\right)^{n-1} > 0$$

for all non-negative smooth functions f_+ on M_+. We call (W, Ω) a **directed symplectic cobordism**, cf. [24].

Remark 5.1.1. Observe that if M_- is empty, (W, Ω) will be a weak symplectic filling of the contact manifold (M_+, ξ_+) with contact structure $\xi_+ := \ker \alpha_+$. If additionally (M_+, ξ_+) is the convex boundary of the symplectic cobordism (W, Ω) with $M_- = \emptyset$ then (W, Ω) is a **strong symplectic filling** of (M_+, ξ_+).

As in [28, 64] such symplectic cobordisms (W, Ω) can be used to verify the strong Weinstein conjecture for the contact manifold that appears as the concave boundary component of (W, Ω). The 3-dimensional variant is related to the ball theorem, see [29, Theorem 2.2 and Corollary 3.8]:

Theorem 5.1.2. *Let (W, Ω) be a directed symplectic cobordism. Assume that the contact manifold (M_+, ξ_+) contains a submanifold N with non-empty boundary such that N supports a bordered Legendrian open book (B, ϑ) and such that $\omega_+|_{TN}$ is exact. Furthermore assume that one of the following conditions is satisfied:*

(i) The symplectic cobordism (W, Ω) is semi-positive.

(ii) The second Stiefel–Whitney class $w_2(TN^*)$ of $N^* = N \setminus (B \cup \partial N)$ vanishes; or N^* is orientable and there exists a class in $H^2(W; \mathbb{Z}_2)$ that restricts to $w_2(TN^*)$.

Then M_- necessarily is non-empty and for any ξ_--defining contact form there exists a null-homologous Reeb link.

Theorem 5.1.2 part (i) is contained in [50]. We include a variant of the argument in Section 5.3 as guideline for the polyfold proof of part (ii). Following [54] we call a $2n$-dimensional symplectic manifold (W, Ω) **semi-positive** if the first Chern class c_1 of (W, Ω) satisfies the following condition: $c_1(A) \geq 0$ for all spherical homology 2-classes A of W with $c_1(A) \geq 3 - n$ and $\Omega(A) > 0$. If $2n \leq 6$ the condition is automatic. The proof of part (ii) of Theorem 5.1.2 is postponed to Section 7.

We remark that the inclusion $N^* \subset W$ induces in cohomology the restriction map $H^2(W; \mathbb{Z}_2) \to H^2(N^*; \mathbb{Z}_2)$. The long exact cohomology sequence with $\mathbb{Z}_2$-coefficients yields, that the restriction map is surjective if and only if the co-boundary operator $H^2(N^*; \mathbb{Z}_2) \to H^3(W, N^*; \mathbb{Z}_2)$ vanishes. Furthermore the choice of an orientation on N^* induces an orientation on each page $\vartheta^{-1}(\theta)$, $\theta \in S^1$, via the co-orientation induced by ϑ. The binding is oriented via the boundary orientation of a page. Therefore, N is orientable if and only if N^* is.

Observe that the inclusion map $f\colon\; N^* \subset N$ induces $TN^* = f^*TN$, so that naturality of the Stiefel–Whitney classes implies $w_2(TN^*) = f^*w_2(TN)$. In particular, if $w_2(TN)$ vanishes so does $w_2(TN^*)$. Further, if the fibration ϑ on N^* is trivial (as it is the case for the plastikstufe), e.g. a product of the page $\vartheta^{-1}(1)$ with S^1, the second Stiefel–Whitney class $w_2(TN^*)$ is equal to $w_2\big(T\vartheta^{-1}(1)\big)$ according to the Whitney cross product formula for the total Stiefel–Whitney class, cf. [16, Chapter 17]. In general, taking a connection on the fibration ϑ we obtain a splitting of TN^* into the vertical $\ker T\vartheta$ and horizontal subbundle. As the horizontal subbundle is isomorphic to the trivial bundle $\vartheta^* TS^1$ Whitney sum formula for the total Stiefel–Whitney class implies $w_2(TN^*) = w_2(\ker T\vartheta)$. The square product of the 2-dimensional Klein bottle shows that the latter not always equals the second Stiefel–Whitney class of a fibre.

Remark 5.1.3. Consider a symplectic cobordism (W, Ω) that satisfies all the requirements of Theorem 5.1.2. Notice that if $M_- = \emptyset$, then (W, Ω) would be a weak symplectic filling of (M_+, ξ_+). Hence, by Theorem 5.1.2 no such weak symplectic filling can exist.

Remark 5.1.4. The exactness requirement for $\omega_+|_{TN}$ in Theorem 5.1.2 is fulfilled e.g. if (M_+, ξ_+) is the convex boundary of (W, Ω) or if the Legendrian open book N is small.

5.2 Completing the cobordism

We consider the directed symplectic cobordism (W, Ω) from Theorem 5.1.2. According to the contact type boundary condition along the negative boundary components of (W, Ω) the symplectic neighbourhood theorem allows a description of a collar neighbourhood of $M_- \subset W$ as $[0, \varepsilon) \times M_-$ such that the symplectic form equals $\Omega = \mathrm{d}(\mathrm{e}^s \alpha_-)$ with $s \in [0, \varepsilon)$. Therefore, a partial completion over the concave boundary of (W, Ω) can be given by

$$\Big((-\infty, 0] \times M_-, \mathrm{d}(\mathrm{e}^s \alpha_-)\Big) \cup (W, \Omega)$$

via gluing along $\{0\} \times M_- \equiv M_-$. Any other contact form defining $\xi_- = \ker \alpha_-$ on M_- can be realised as the restriction to the tangent spaces of a graph over M_- inside $(-\infty, \varepsilon) \times M_-$ up to a positive constant factor, cf. [28, Section 3.3]. A change of (W, Ω) by adding the super-level set of the graph to the directed symplectic cobordism will allow the verification of the strong Weinstein conjecture for a particular choice of contact form as announced in Theorem 5.1.2. In fact, we will assume that α_- is a generic perturbation of a given ξ_--defining contact form which allows an application of the Gromov–Hofer compactness theorem as formulated in [15]. This is justifiable with the Arzelà–Ascoli theorem, cf. [28, Section 6.4] and [64, Section 6].

On the **negative end** $\big((-\infty,0]\times M_-, \mathrm{d}(\mathrm{e}^s\alpha_-)\big)$ we choose a shift invariant almost complex structure J_- that sends the Liouville vector field ∂_s to the Reeb vector field of α_- and leaves ξ_- invariant such that J_- is compatible with the symplectic structure induced by $\mathrm{d}\alpha_-$ on ξ_-. Extend J_- to a tamed almost complex structure J on (W,Ω) such that $\xi_+ = TM_+ \cap JTM_+$ and the positive boundary (M_+,ξ_+) of (W,Ω) is J-convex, see [50, Theorem D]. Further, we glue the deformed magnetic collar extension constructed in Sections 4.2 and 4.3 along $M_+ \equiv \{0\}\times M_+$ to build (keeping the notation) a symplectic manifold

$$(\hat{W},\hat{\Omega}) := \Big((-\infty,0]\times M_-, \mathrm{d}(\mathrm{e}^s\alpha_-)\Big) \cup (W,\Omega) \cup \Big([0,2s_o]\times M_+, \mathrm{d}(s\alpha_+)+\eta\Big)\,.$$

The resulting almost complex structure is denoted by $\hat{J}$. Notice that $\hat{J}$ equals J_o on the neighbourhoods U_B and $U_{\partial N}$ of B and ∂N, resp. Here we think of B and ∂N as subsets of $M_+ \equiv \{2s_o\}\times M_+$. In particular, the results from Section 4.4 are available.

5.3 The semi-positivity case

We prove Theorem 5.1.2 under assumption (i). Recall the moduli space $\mathcal{M}$ introduced in Section 4.4. To cut out the Möbius reparametrisation group geometrically we define the moduli space $\mathcal{M}_{1,\mathrm{i},-1}$ to be the set of all holomorphic discs $u\in\mathcal{M}$ such that $\vartheta\circ u(\mathrm{i}^k)=\mathrm{i}^k$ for $k=0,1,2$. This allows to fix the disc reparametrisations for sequences in $\mathcal{M}_{1,\mathrm{i},-1}$ in the compactness formulation in Remark 4.4.1.

We choose a base point b_o of B and an embedded curve γ inside the page $\vartheta^{-1}(1)$ that connects B and ∂N such that γ is given by $\{b_o\}\times\big(D^2\cap\mathbb{R}_{\geq 0}\big)$ in the model description in Section 2.3. We define the moduli space $\mathcal{M}_\gamma$ to be the set of all holomorphic discs $u\in\mathcal{M}_{1,\mathrm{i},-1}$ with $u(1)\in\gamma$. In other words $\mathcal{M}_\gamma$ consists of all $u\in\mathcal{M}$ such that

$$u(1)\in\gamma\,,\quad \vartheta\circ u(\mathrm{i})=\mathrm{i}\,,\quad \vartheta\circ u(-1)=-1\,.$$

The Maslov index of the Fredholm problem defined by $\mathcal{M}$ equals 2 (see [55, Proposition 8] or [27]). With [68, Theorem 4.4] there is a generic choice of $\hat{J}$ such that with the first dimension formula in [68, Theorem 3.7] (successively taking relations $R = \emptyset$, $R = \vartheta^{-1}(1) \times \vartheta^{-1}(\mathrm{i}) \times \vartheta^{-1}(-1)$, and $R = \gamma \times \vartheta^{-1}(\mathrm{i}) \times \vartheta^{-1}(-1)$) the moduli spaces $\mathcal{M}$, $\mathcal{M}_{1,\mathrm{i},-1}$, and $\mathcal{M}_\gamma$ are smooth manifolds of dimension $n+2$, $n-1$, and 1, resp. By [68, Remark 3.6] we can assume that the generic perturbation of $\hat{J}$ is supported in the complement of the union of the negative end $(-\infty, 0] \times M_-$ of $\hat{W}$ and $U_B \cup U_{\partial N}$. This is because all local Bishop discs are Fredholm regular by [55, Proposition 9] and because all holomorphic discs that are contained completely inside $\big((-\infty, 0] \times M_-\big) \cup U_B \cup U_{\partial N}$ are the local Bishop discs.

Moreover, the boundary component of $\mathcal{M}_{1,\mathrm{i},-1}$ that corresponds to the local Bishop filling F has a collar neighbourhood in $\mathcal{M}_{1,\mathrm{i},-1}$ diffeomorphic to $[0,1) \times B$ via F. This results in a collar neighbourhood $[0,1) \times \{b_o\}$ in $\mathcal{M}_\gamma$, see Section 4.4. By the blocking property of $U_{\partial N}$ (see Section 3.5) the evaluation map $\mathcal{M}_\gamma \to \gamma$, $u \longmapsto u(1)$, is not surjective. Invariance of the mod-2 degree for proper maps, which counts the number of preimages modulo 2, implies that $\mathcal{M}_\gamma$ cannot be compact.

Proof of Theorem 5.1.2 part (i). Arguing by contradiction we suppose that there is a compact subset K of $\hat{W}$ such that the holomorphic discs $u(\mathbb{D})$ are contained in K for all $u \in \mathcal{M}_\gamma$. In this situation $\mathcal{M}_\gamma$ can be compactified in the sense of Gromov, see Remark 4.4.1. Observe that Gromov limiting stable holomorphic discs are contained in K also and have precisely one disc component that in addition must be simple. With [68] we can assume by an additional *a priori* perturbation of $\hat{J}$ that the moduli spaces of simple stable maps that cover the stable maps in the Gromov compactification of $\mathcal{M}$ are cut out transversally. This perturbation can be supported in the complement of the union of $(-\infty, 0] \times M_-$ and $U_B \cup U_{\partial N}$ because no holomorphic sphere can stay inside $\big((-\infty, 0] \times M_-\big) \cup U_B \cup U_{\partial N}$ completely by the maximum principle. But, similarly to the computations in [28, Section 6.3], the moduli spaces of the covering simple stable holomorphic discs are of negative dimensions, hence, empty. This argument uses the semi-positivity assumption. Consequently, there is no bubbling off for the moduli space $\mathcal{M}_\gamma$; in other words $\mathcal{M}_\gamma$ is compact. This is a contradiction and therefore a compact subset K of $\hat{W}$ that contains all holomorphic discs that belong to $\mathcal{M}_\gamma$ cannot exists.

Consequently, M_- is necessarily non-empty. Moreover, for any choice of ξ_--defining contact form α_- there exists a null-homologous Reeb link by the remarks made in Section 5.2. The relevant formulation of Gromov–Hofer convergence is obtained by combining the convergences statements in [36, 37] with [25]. □

6 A Deligne–Mumford type space

In Section 7 the proof of Theorem 5.1.2 under assumption (ii) will be given. In preparation we discuss moduli spaces of stable nodal boundary un-noded discs. We follow [43, Section 2.1], [18, 42] and indicate modifications necessary in the presence of boundaries.

6.1 Boundary un-noded nodal discs

Let S be an oriented surface that is equal to the disjoint union of one closed disc and a (possibly empty) finite collection of spheres. All connected components of S are provided with the standard orientation. Let j be an orientation preserving complex structure on S turning (S, j) into a Riemann surface with boundary, i.e. (S, j) admits a holomorphic atlas whose charts are given by open subsets of the closed upper half-plane.

We call a subset of $\mathrm{Int}(S)$ consisting of two distinct points a **nodal pair**. Each finite collection D of pair-wise disjoint nodal pairs defines an equivalence relation on S calling two points equivalent if and only if they from a nodal pair. The set S/D of equivalence classes is provided with the quotient space topology.

Let D be a finite collection of pair-wise disjoint nodal pairs such the quotient space S/D is simply connected. We call (S, j, D) a **boundary un-noded nodal disc**. A point of S that belongs to a nodal pair in D is called a **nodal point**. The set of nodal points is denoted by $|D|$. Observe that $|D|$ and ∂S are disjoint.

Let m_0, m_1, m_2 be pair-wise distinct **marked points** on the boundary ∂S ordered according to the boundary orientation of ∂S. We call $\big(S, j, D, \{m_0, m_1, m_2\}\big)$ a **marked boundary**

un-noded nodal disc. Two marked boundary un-noded nodal discs

$$\big(S, j, D, \{m_0, m_1, m_2\}\big) \quad \text{and} \quad \big(S', j', D', \{m'_0, m'_1, m'_2\}\big)$$

are **equivalent** if there exists a diffeomorphism $\varphi\colon\; S \to S'$ such that $\varphi^* j' = j$, the injection $D \to D'$ defined by $\{\varphi(x), \varphi(y)\} \in D'$ for all $\{x, y\} \in D$ is onto, and $\varphi(m_k) = m'_k$ for $k = 0, 1, 2$. Observe that φ necessarily preserves orientations.

6.2 Domain stabilisation

In Section 7.2 boundary un-noded nodal discs will appear as the domain of stable maps. If the domain nodal discs have sphere components, the nodal discs will be unstable. In order to obtain a natural groupoidal structure on the space of marked boundary un-noded nodal discs we have to stabilise these by adding marked points. A point that is a marked point or a nodal point is called **special**. We call a connected component C of S **stable** if the number of special points on C is greater or equal than 3. In particular the disc component of S is stable.

We consider equivalence classes of stable discs $\big[S, j, D, \{m_0, m_1, m_2\}, A\big]$ where $\big(S, j, D, \{m_0, m_1, m_2\}\big)$ is a marked boundary un-noded nodal disc as in Section 6.1 that we provide with an additional finite set of **auxiliary marked points** $A \subset S \setminus \partial S$ in the complement of $|D|$ so that $\#\big((A \cup |D|) \cap C\big) \geq 3$ for each sphere component C of S. In particular, all components C of S are stable. The **equivalence relation** is given by diffeomorphisms $\varphi\colon\; S \to S'$ as in Section 6.1 such that in addition φ maps A bijectively onto A'.

The set of all equivalence classes $\mathcal{R}$ is a **nodal Riemann moduli space**. Given a non-negative integer N we denote by $\mathcal{R}_N \subset \mathcal{R}$ the subset of stable nodal discs that are equipped with precisely $N = \#A$ auxiliary marked points so that $\mathcal{R}$ is the disjoint union over all $\mathcal{R}_N$, $N \geq 0$. The elements in $\mathcal{R}_N$ can be represented by stable nodal discs $\big(S, j, D, \{m_0, m_1, m_2\}, A\big)$ of different **stable nodal type** τ, which is an isomorphism class of weighted rooted trees. The vertices are given by the components of S, where the root corresponds to the disc component. All vertices are weighted by the number of (auxiliary) marked points on the corresponding component of S. The edge relation is given by the nodes in D. Observe that for given N the number of stable nodal types corresponding to stable discs $(S, j, D, \{m_0, m_1, m_2\}, A)$ with $N = \#A$ is finite so that $\mathcal{R}_N$ is a finite disjoint union of subsets of stable discs $\mathcal{R}_\tau \subset \mathcal{R}$ of the same stable nodal type τ.

6.3 Groupoid as an orbit space

Let τ be a stable nodal type. In order to rewrite $\mathcal{R}_\tau$ as an orbit space we choose natural representatives of the stable nodal marked discs $[S, j, D, \{m_0, m_1, m_2\}, A]$ in $\mathcal{R}_\tau$ as follows: We fix the oriented diffeomorphism types of $\mathbb{C}P^1$ and $\mathbb{D}$ for the components of S including the choice $\{1, \mathrm{i}, -1\}$ for the marked points $\{m_0, m_1, m_2\}$, i.e. $m_k = \mathrm{i}^k$, $k = 0, 1, 2$. We denote by σ the **area form** on S that is the Fubini–Study form on the $\mathbb{C}P^1$ components and equals $\mathrm{d}x \wedge \mathrm{d}y$ on $\mathbb{D}$ taking conformal coordinates $x + \mathrm{i}y$. Let $\mathcal{J} \equiv \mathcal{J}_S$ be the space of orientation preserving complex structures on S, which equals the space of almost complex structures on S tamed by σ, cf. [1, 63]. Furthermore we fix a selection of special points $|D|$ and A. By $\mathcal{G} \equiv \mathcal{G}(S, D, \{1, \mathrm{i}, -1\}, A)$ we denote the group of orientation preserving diffeomorphisms of S that preserve $\{1, \mathrm{i}, -1\}$ point-wise and inject A onto A and D onto D, resp.

We identify the nodal Riemann moduli space $\mathcal{R}_\tau$ with the orbit space

$$\mathcal{R}_\tau = \mathcal{J}/\mathcal{G} \qquad \text{via} \quad [S, j, D, \{1, \mathrm{i}, -1\}, A] \equiv [j]\,,$$

cf. [61, p. 612] or [67, Section 4.2]. The action

$$\mathcal{G} \times \mathcal{J} \longrightarrow \mathcal{J}\,, \quad (\varphi, j) \longmapsto \varphi^* j\,,$$

of $\mathcal{G}$ on $\mathcal{J}$ is given by the pull back

$$\varphi^* j := T_\varphi \varphi^{-1} \circ j_\varphi \circ T\varphi$$

for $\varphi \in \mathcal{G}$ and $j \in \mathcal{J}$. By [61, Lemma 7.5] or [67, Lemma 4.2.8] this action is proper, see also Remark 6.3.2 below. The action is free if and only if all isotropy subgroups $\mathcal{G}_j$, $j \in \mathcal{J}$, of j-holomorphic maps in $\mathcal{G}$ are trivial. The action is locally free because all isotropy subgroups $\mathcal{G}_j$ are finite by the following Remark 6.3.1:

Remark 6.3.1. Each connected component of S is provided with at least 3 special points, see Section 6.2. Hence, all **isotropy subgroups** $\mathcal{G}_j$ **are finite**:

To see this fix a biholomorphic identification of (S, j) with the surface given by the disjoint union of $(\mathbb{D}, \mathrm{i})$ and an at most finite number of copies of $(\mathbb{C}P^1, \mathrm{i})$. This is possible by uniformisation, cf. [1] and [54, Theorem C.5.1], [63, Satz 5.33] for boundary regularity. For the marked points we can assume that $m_k = \mathrm{i}^k$, $k = 0, 1, 2$. Then any automorphism in $\mathcal{G}_j$ conjugates to an i-holomorphic map that restricts to the identity on $(\mathbb{D}, \mathrm{i})$ and defines Möbius transformations of $(\mathbb{C}P^1, \mathrm{i})$ corresponding to the maps induced between the not necessarily identical sphere components. We get a finite number of possibilities to obtain those Möbius transformations each of which is permuting the set of special points that admits at least 3 points by the stability condition.

Remark 6.3.2. In order to describe the **topology for fixed stable nodal type** provide $\mathcal{J}$ with the C^∞-topology, which is metrisable, complete and locally compact by the Arzelà–Ascoli theorem.

We provide $\mathcal{R}_\tau = \mathcal{J}/\mathcal{G}$ with the **quotient topology** meaning that the open sets in $\mathcal{J}/\mathcal{G}$ are precisely those, whose preimage under the quotient map $[\,.\,]\colon\; \mathcal{J} \to \mathcal{J}/\mathcal{G}$ is open. In particular, $[\,.\,]$ is continuous by definition. The quotient map $[\,.\,]$ is open because for any open subset $\mathcal{K}$ of $\mathcal{J}$ the $[\,.\,]$-preimage of $[\mathcal{K}]$ is equal to the union of all $g\mathcal{K}$, $g \in \mathcal{G}$, which is open. Hence, a neighbourhood base of the topology on $\mathcal{J}/\mathcal{G}$ is given by the family of subsets whose elements $[j]$ can be represented by complex structures j belonging to an open subset of $\mathcal{J}$. Therefore, $\mathcal{J}/\mathcal{G}$ is a second countable locally compact and, hence, paracompact topological space.

In fact, $\mathcal{R}_\tau = \mathcal{J}/\mathcal{G}$ is Hausdorff. This follows from the **properness argument** as follows: Consider a sequence u_ν of equivalences

$$u_\nu\colon\; \big(S, j_\nu, D, \{1, \mathrm{i}, -1\}, A\big) \longrightarrow \big(S, k_\nu, D, \{1, \mathrm{i}, -1\}, A\big)$$

with $j_\nu \to j$ and $k_\nu \to k$ in $\mathcal{J}$. We claim that u_ν has a C^∞-convergent subsequence whose limit u will be an equivalence

$$u\colon\ \big(S,j,D,\{1,\mathrm{i},-1\},A\big)\longrightarrow\big(S,k,D,\{1,\mathrm{i},-1\},A\big)$$

also, cf. [42, Proposition 3.25].

It suffices to proof C^∞-convergence because the limit u will be automatically an equivalence as u restricts to a degree 1 map on each component of S. Restricting u_ν to the components of S we obtain sequences of k_ν-holomorphic diffeomorphisms v_ν one sequence for each component of (S,j_ν). As the degree of each v_ν equals 1 viewed as map onto its image, the area $\int_C v_\nu^*\sigma=\pi$, C being $\mathbb{D}$ or $\mathbb{C}P^1$, is uniformly bounded via the transformation formula. Hence, we find a subsequence of u_ν so that all corresponding sequences v_ν converge in the sense of Gromov, see [25]. Again using that the degree of all v_ν is 1 we see that only one of the potential bubbles of each Gromov limit can intersect $0\in C$; the remaining bubbles would be necessarily constant as their area vanish. In other words, the chosen subsequence of u_ν converges in C^∞, because the reparametrisations by j_ν-holomorphic diffeomorphisms are fixed, i.e. equal id, due to the stability condition.

We verify the Hausdorff property, cf. [42, Proposition 3.19]: Assume that any pair of neighbourhoods of given points $[j]$ and $[k]$, resp., intersects non-trivially. Taking shrinking neighbourhoods of the corresponding points j and k in $\mathcal{J}$ we find $j_\nu\to j$ and $k_\nu\to k$ in $\mathcal{J}$ such that $[j_\nu]=[k_\nu]$ for all ν. Then the above properness argument yields $[j]=[k]$.

6.4 Infinitesimal action

We assume the situation of Section 6.3. Denote by Ω^0 the Lie algebra given by the tangent space of $\mathcal{G}$ at the identity diffeomorphism, which is the space of vector fields on $(S,\partial S)$ that are stationary on the set of all special points. In other words, Ω^0 is the set of all smooth sections of TS that are tangent to ∂S and vanish on $|D|\cup\{1,\mathrm{i},-1\}\cup A$. We denote by $\Omega^{0,1}_j$ the space of all endomorphism fields of the tangent bundle of S that anti-commute with j. Linearising the equation $j^2=-1$ we see that $\Omega^{0,1}_j$ is the tangent space of $\mathcal{J}$ at j.

Viewing $\Omega^{0,1}_j$ as the set of all j-complex anti-linear TS-valued differential forms on S, the infinitesimal action is

$$\Omega^0 \times \Omega^{0,1}_j \longrightarrow \Omega^{0,1}_j\,, \quad (X, y) \longmapsto L_X j + y\,.$$

A **complex linear Cauchy–Riemann operator** D is a j-complex linear operator that maps smooth vector fields of (S, j) to $\Omega^{0,1}_j$ such that $D(fX) = \bar\partial f \cdot X + fDX$ for all smooth vector fields X and smooth functions f, where $\bar\partial$ is the composition of the exterior derivative with the projection of the space of smooth 1-forms onto those that anti-commute with j, see [54, Appendix C.1]. Complex linearity can be expressed via $D(jX) = jDX$ for all vector fields X of S.

In order to compute the Lie derivative $L_X j$ we denote by $\bar\partial_j$ the uniquely determined complex linear Cauchy–Riemann operator of the holomorphic line bundle (TS, j) (ignoring boundary points) that agrees with

$$\bar\partial X = \tfrac{1}{2}\big(TX + \mathrm{i}\circ TX\circ\mathrm{i}\big)$$

in local holomorphic coordinates $(\mathbb{C}, \mathrm{i})$, cf. [54, Remark C.1.1]. The local holomorphic representation $\bar\partial$ of $\bar\partial_j$ implies that $\bar\partial_j$ induces a **real linear Cauchy–Riemann operator** on the bundle pair $\big((TS, T\partial S), j\big)$ taking local holomorphic coordinates in the closed upper half-plane. Nevertheless the operator $\bar\partial_j\colon\ \Omega^0 \to \Omega^{0,1}_j$ **is not complex linear** because j does not induce a complex linear vector space structure on Ω^0. Indeed, jX is not tangent to ∂S for all boundary points of S for which $X \in \Omega^0$ does not vanish.

With this preparation we compute

$$L_X j = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \varphi_t^* j\,,$$

where φ_t is a smooth path in $\mathcal{G}$ through $\varphi_0 = \mathrm{id}$ with

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \varphi_t = X\,.$$

In local holomorphic coordinates $\varphi_t^* j$ reads as

$$T_{\varphi_t}\varphi_t^{-1}\circ\mathrm{i}\circ T\varphi_t\,.$$

Taking the time derivative at $t = 0$ and using $\varphi_t^{-1} = \varphi_{-t}$ yields

$$-TX\circ\mathrm{i} + \mathrm{i}\circ TX = \mathrm{i}\big(\mathrm{i}\circ TX\circ\mathrm{i} + TX\big)\,,$$

so that the Lie derivative is equal to

$$L_X j = 2j\bar{\partial}_j X\,.$$

Remark 6.4.1. Taking a path φ_t in the isotropy subgroup $\mathcal{G}_j$ through $\varphi_0 = \mathrm{id}$, meaning that the path φ_t in $\mathcal{G}$ satisfies $\varphi_t^* j = j$ for all t, we obtain $X \in \ker\bar{\partial}_j$ by taking time derivative. Conversely, assuming $\bar{\partial}_j X = 0$ for a path φ_t in $\mathcal{G}$ through $\varphi_0 = \mathrm{id}$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\varphi_t^* j\big) = \varphi_t^*\Big(2j\bar{\partial}_j X\Big) = 0\,,$$

i.e. that the path $\varphi_t^* j$ in $\mathcal{J}$ is constant, hence, equals j. In other words, the tangent space at the identity of the group of all j-holomorphic maps in $\mathcal{G}$

$$T_{\mathrm{id}}\mathcal{G}_j = \ker\bar{\partial}_j$$

equals the space of all j-holomorphic vector fields in Ω^0.

Remark 6.4.2. The Cauchy–Riemann operator is **conformally invariant**: Let $\varphi\colon\ (S,j) \to (S,k)$ be a holomorphic diffeomorphism. Then $\varphi^*\circ\bar{\partial}_k = \bar{\partial}_j\circ\varphi^*$. In particular, the Cauchy–Riemann operator $\bar{\partial}_j$ commutes with the automorphisms of the Riemann surface with boundary (S,j).

Remark 6.4.3. Define a Riemannian metric g_j on (S,j) by setting

$$g_j(v,w) = \tfrac{1}{2}\big(\sigma(v,jw) + \sigma(w,jv)\big)\,,\quad v,w \in TS\,,$$

where the area form σ is the one chosen in Section 6.3. In particular, the complex structure j is orthogonal, so that the bilinear form g_j is a **Hermitian metric**. As any 2-form on a surface is closed the 2-form

$$\sigma_j := \tfrac{1}{2}(\sigma + j^*\sigma)$$

is symplectic and compatible with j. The latter means that $g_j(v,w) = \sigma_j(v, jw)$ for all $v, w \in TS$, so that the symplectic form σ_j is compatible with j.

Denoting the **Levi-Civita connection** of g_j by $\nabla \equiv \nabla^{g_j}$, [52, Lemma 4.15] says that closedness of σ_j and integrability of j together are equivalent to $\nabla j = 0$, cf. Remark 6.4.4 below. In particular, $\nabla(jX) = j\nabla X$ for all vector fields X of S, so that ∇ is a **Hermitian connection**, i.e. ∇ is a complex linear metric connection. Hence,

$$\big(\nabla X\big)^{0,1} := \tfrac{1}{2}\big(\nabla X + j\circ\nabla X\circ j\big)$$

defines a complex linear Cauchy–Riemann operator on (TS, j), cf. [54, Remark C.1.2]. In local holomorphic coordinates we write $\nabla X = TX + \Gamma(\,.\,, X)$ with help of the Christoffel symbols, so that the map $X \longmapsto \Gamma(\,.\,, X)$ is i-complex linear. With symmetry of Γ we get $\mathrm{i}\circ\Gamma(\mathrm{i}\,.\,, X) = \mathrm{i}\circ\Gamma(X, \mathrm{i}\,.\,) = -\Gamma(\,.\,, X)$, so that $\big(\nabla X\big)^{0,1} = \bar\partial X$. Therefore,

$$\big(\nabla X\big)^{0,1} = \bar\partial_j X\,.$$

Furthermore we remark that the Hermitian connection ∇ is uniquely determined by this equation, see [54, Remark C.1.2]. Consequently, $L_X j = 2j\bar\partial_j X$ can be obtained with the computations on [54, Theorem C.5.1] or [61, p. 631] as well.

Remark 6.4.4. We give an alternative argument for the fact that j is parallel, which we used in Remark 6.4.3: For any non-vanishing tangent vectors v, w at any given point of S consider a curve c tangent to v and extend w to a parallel vector field X along c. As X and jX are orthogonal and as the length of jX is constant ∇ being metric implies that $\nabla_{\dot c}(jX)$ is perpendicular to the span of $\{X, jX\}$, and hence vanishes. Therefore, with the Leibnitz rule and parallelity of X we get $(\nabla_{\dot c} j)X = 0$. Consequently, $\nabla j = 0$.

Observe that the argument works for all metrics (and corresponding Levi-Civita connections) for which j is orthogonal.

Remark 6.4.5. All elements of $\mathcal{G}_j$ are **isometries** of g_j. Indeed, by finiteness of $\mathcal{G}_j$ one finds for each $\psi \in \mathcal{G}_j$ a natural number k such that $\psi^k = \mathrm{id}$. On the other hand ψ pulls g_j back to $f g_j$ for some positive function f on S by the description in Remark 6.4.3 and the fact that all positive area forms on a surface are positively proportional. Therefore, the conformal factor of the pull-back of g_j by ψ^k becomes f^k, which necessarily is 1. Hence, $f = 1$.

Remark 6.4.6. Let (φ_t, j_t) be a smooth path in $\mathcal{G} \times \mathcal{J}$ through $(\varphi_0, j_0) = (\psi, j)$ and denote the velocity vector field by

$$\frac{\mathrm{d}}{\mathrm{d}t}(\varphi_t, j_t) = \big(X_t, y_t\big) \in T_{\varphi_t}\mathcal{G} \times T_{j_t}\mathcal{J} \,.$$

The corresponding velocity vector field of $\varphi_t^* j_t$ equals

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\varphi_t^* j_t\big) = \varphi_t^*\Big(2 j_t \bar{\partial}_{j_t} X_t + y_t\Big) \in T_{\varphi_t^* j_t}\mathcal{J} \,.$$

If $\psi \in \mathcal{G}_j$ we obtain with $(X, y) = (X_0, y_0)$ that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\big(\varphi_t^* j_t\big) = \psi^*\Big(2 j \bar{\partial}_j X + y\Big) \in T_j\mathcal{J} \,,$$

which is equal to

$$2 j \bar{\partial}_j\big(\psi^* X\big) + \psi^* y \,.$$

Indeed, this follows with Remark 6.4.2 or with Gauß's *theorema egregium* which gives $\psi^* \circ \bar{\partial}_j = \bar{\partial}_j \circ \psi^*$ because all $\psi \in \mathcal{G}_j$ are isometries of g_j, see Remark 6.4.5. Consequently, the corresponding infinitesimal action reads as

$$T_\psi \mathcal{G} \times \Omega_j^{0,1} \longrightarrow \Omega_j^{0,1} \,, \quad (X, y) \longmapsto 2 j \bar{\partial}_j\big(\psi^* X\big) + \psi^* y$$

for all $\psi \in \mathcal{G}_j$. Observe that the infinitesimal action $2 j \bar{\partial}_j \oplus \mathbf{1}$ at (id, j) sends $\big(\psi^* X, \psi^* y\big) \in \Omega^0 \times \Omega_j^{0,1}$ to the same element in $\Omega_j^{0,1}$.

6.5 A Fredholm index

We compute the Fredholm index of the $W^{1,3}$-Sobolev completed Cauchy–Riemann operator $\bar{\partial}_j \colon\ W^{1,3} \to L^3$ induced by the Cauchy–Riemann operator from Section 6.4.

Ignoring zeros for the moment, for each component C of S the Fredholm index is given by the Riemann–Roch [54, Theorem C.1.10] applied to the j-complex 1-dimensional bundle pair $(TC, T\partial C)$. Namely, the Fredholm index is the sum of the Euler characteristic of C and the Maslov index of $(TC, T\partial C)$. The Maslov index for $C = \mathbb{D}$ is 2 by normalisation; for $C = \mathbb{C}P^1$ twice the first Chern number, i.e. twice the Euler characteristic, which gives 4, see [54, Chapter C.3]. Hence, the Fredholm index is 3 restricted to the disc component and 6 on the spheres.

Now we take the zeros into account. Component-wise, for each boundary zero we have to subtract 1 from the computed Fredholm index; for each interior zero we subtract 2. Adding up, we obtain that $\operatorname{ind}\bar{\partial}_j$ equals

$$3 - \#\{1, \mathrm{i}, -1\} - 2\#\big((|D| \cup A) \cap \mathbb{D}\big) + 2\sum_C \Big(3 - \#\big((|D| \cup A) \cap C\big)\Big),$$

where the sum is taken over all sphere components C of S. This gives

$$\operatorname{ind}\bar{\partial}_j = -2\#|D| - 2\#A + 6\big(\#\{C\} - 1\big),$$

where $\#\{C\} = \#D + 1$ is the number of all components of S. Using $\#|D| = 2\#D$ we finally obtain

$$\operatorname{ind}\bar{\partial}_j = 2\big(\#D - \#A\big).$$

On the other hand, by a boundary version of the argument principle (see [3, Theorem A.5.4]), the Maslov index of $(TC, T\partial C)$ for each component C of S is the weighted sum of the number of zeros counted multiplicities of a non-zero element in the kernel of $\bar{\partial}_j$, where interior zeros are counted twice. As the corresponding Maslov index is 2 on the disc component and 4 on the spheres, the kernel of $\bar{\partial}_j$ is trivial due to the stability condition that the number of special points on each component C of S is at least 3. This argument uses elliptic regularity saying that the vector fields in $\ker\bar{\partial}_j$ are smooth, see [54]. Therefore, one can show triviality of $\ker\bar{\partial}_j = T_{\mathrm{id}}\mathcal{G}_j$ (see Remark 6.4.1) alternatively using finiteness of $\mathcal{G}_j$ under the stability condition, see Remark 6.3.1.

Using elliptic regularity as in [54] we see that the Cauchy–Riemann operator $\bar{\partial}_j\colon\ \Omega^0 \to \Omega^{0,1}_j$ is injective. The image is closed and has codimension $2\big(\#A - \#D\big)$, cf. [43, Proposition 2.5].

6.6 Interlude: Cayley transformation

We provide $\Omega^{0,1}_j$ with the norm $|y|_j$ given by the maximum of point-wise operator norms of y w.r.t. the metric g_j. Because y is self-adjoint w.r.t. g_j the norm $|y|_j$ is equal to the maximum of the square-root of the point-wise eigenvalues of the j-complex linear endomorphism field $y\circ y$. In particular, the resolvent $(1-y)^{-1}$ of y at 1 is defined provided $|y|_j < 1$. We obtain a homeomorphism

$$\mathfrak{i}:\ \Omega^{0,1}_j \cap B_1(0) \longrightarrow \mathcal{J}\,, \quad 0 \longmapsto j\,,$$

defined on the open unit ball about zero via the conjugation

$$y \longmapsto (1-y)j(1-y)^{-1}\,,$$

whose inverse is given by the

$$k \longmapsto (k+j)^{-1}(k-j)\,,$$

cf. [7, Proposition 1.1.6], [53, Proposition 2.6.4] or [61, p. 634]. Because $\mathcal{G}_j$ acts by isometries of g_j (see Remark 6.4.5) the conjugation map $\mathfrak{i}$ is $\mathcal{G}_j$-equivariant.

We claim that $\mathcal{J}$ is a submanifold of the space Ω^1 of all endomorphism fields of the tangent bundle of S and that $\mathfrak{i}$ is a global chart. The case of almost complex structures compatible with σ essentially follows with the expositions in the above cited literature [7, 53, 61]. We will follow [31, Chapter I.7.3] taking the modifications for the case of almost complex structures only tamed by σ into account:

We call a not necessarily symmetric endomorphism field $x \in \Omega^1$ positive and write $x > 0$ provided that for all non-zero tangent vectors $v \in TS$ the quadratic form $g_j(v, xv)$ is positive. As the kernel of a positive endomorphism field $x \in \Omega^1$ is trivial we see that the inverse $x^{-1} \in \Omega^1$ exists. Therefore, for positive $x \in \Omega^1$ the endomorphism field $1 + x$ is positive as well such that the inverse $(1+x)^{-1}$ exists. In fact, the half space of all positive endomorphism fields $\Omega^1 \cap \{x > 0\}$ is an open cone in Ω^1 closed under taking inverses. As above we provide Ω^1 with the norm $|x|_j$ given by the maximum of point-wise operator norms of $x \in \Omega^1$ w.r.t. the metric g_j.

The **Cayley transform**

$$\mathcal{C}(x) := (1-x)(1+x)^{-1} =: \tilde{y}$$

defines a map

$$\Omega^1 \cap \{x > 0\} \longrightarrow \Omega^1 \cap B_1(0)\,.$$

Indeed, setting $v = (1+x)^{-1}w$ polarisation yields $4g_j(v, xv) = |w|_j^2 - |\tilde{y}w|_j^2$, so that x is positive if and only if $|\tilde{y}|_j < 1$. We remark that there is an alternative formula

$$\mathcal{C}(x) = (1+x)^{-1}(1-x)$$

for the Cayley transform because $1+x$ and $1-x$ commute. Setting

$$\mathcal{C}(\tilde{y}) := (1-\tilde{y})(1+\tilde{y})^{-1}$$

we obtain a map

$$\mathcal{C}\colon\ \Omega^1 \cap B_1(0) \longrightarrow \Omega^1 \cap \{x > 0\}$$

in the converse direction. Again, this uses polarisation and that $|\tilde{y}|_j < 1$ implies triviality of the kernel of $1+\tilde{y}$. Because the Cayley transform $\mathcal{C}$ is involutive the map $\mathcal{C}\colon\ \Omega^1 \cap \{x > 0\} \to \Omega^1 \cap B_1(0)$ is a diffeomorphism.

Extending the conjugation map $\mathfrak{i}\colon\ \tilde{y} \longmapsto (1-\tilde{y})j(1-\tilde{y})^{-1}$ to the open set $\Omega^1 \cap B_1(0)$ yields a smooth map $\Omega^1 \cap B_1(0) \to \Omega^1$. Observe, that $\mathfrak{i}(\tilde{y})$ is a complex structure potentially reversing orientation. Restricting $\mathfrak{i}$ to the j-complex anti-linear part $\Omega^{0,1}_j$ of Ω^1 we get $\mathfrak{i}(y) = j\mathcal{C}(-y)$ for all $y \in \Omega^{0,1}_j \cap B_1(0)$. This defines an injection

$$\mathfrak{i} = j\circ\mathcal{C}\circ(-1)\colon\ \Omega^{0,1}_j \cap B_1(0) \longrightarrow \Omega^1\,.$$

In order to describe the image we observe that

$$\mathcal{C}\circ(-1)\colon\ \Omega^{0,1}_j \cap B_1(0) \longrightarrow \Omega^1 \cap \{x > 0\} \cap \{jx = x^{-1}j\}$$

is a well defined homeomorphism with inverse $(-1)\circ\mathcal{C}$. Additionally, the multiplication map $j\colon\ \Omega^1 \to \Omega^1$ is a diffeomorphism with inverse $-j$ and restricts to the homeomorphism

$$j\colon\ \Omega^1 \cap \{x > 0\} \cap \{jx = x^{-1}j\} \longrightarrow \mathcal{J}\,.$$

Indeed, positivity of x is equivalent to the positivity of $\sigma_j(v, jxv)$ for all $v \in TS$. Moreover, we find a smooth function f on the surface S such that the 2-forms σ and σ_j satisfy $\sigma = f\sigma_j$. Because $\sigma(w, jw) = f|w|_j^2$ is positive for all non-zero $w \in TS$ the function f must be positive. Hence, positivity of x is equivalent to the positivity of $\sigma(v, jxv)$ for all $v \in TS$, as σ_j and σ are positively proportional. It follows that the complex structure jx is tamed by σ, i.e. preserves the orientation of the Riemann surface (S, j). In total, the conjugation map

$$\mathfrak{i} = j\circ\mathcal{C}\circ(-1): \ \Omega_j^{0,1} \cap B_1(0) \longrightarrow \mathcal{J}$$

is a well defined homeomorphism. The inverse is $k \longmapsto -\mathcal{C}(-jk) = (k+j)^{-1}(k-j)$, where equality follows with the above alternative commuted formula for the Cayley transform. Furthermore $\mathfrak{i} = j\circ\mathcal{C}\circ(-1)$ is the restriction of the diffeomorphism obtained as the composite of $\mathcal{C}\circ(-1): \ \Omega^1 \cap B_1(0) \to \Omega^1 \cap \{x > 0\}$ with $j: \ \Omega^1 \to \Omega^1$ where defined. In other words, the inverse $(-1)\circ\mathcal{C}\circ(-j)$ of $j\circ\mathcal{C}\circ(-1)$ serves as a global submanifold chart of $\mathcal{J} \subset \Omega^1$ the model being $\Omega_j^{0,1} \cap B_1(0) \subset \Omega^1$.

In fact, $\mathcal{J}$ is a complex manifold: A complex structure on the vector space $\Omega_j^{0,1}$ is given by $y \longmapsto jy$. An almost complex structure on $\mathcal{J}$ is given by

$$\mathfrak{i}(y) = (1-y)j(1-y)^{-1}$$

on the tangent space $T_{\mathfrak{i}(y)}\mathcal{J} = \Omega^{0,1}_{\mathfrak{i}(y)}$ for all $y \in \Omega_j^{0,1} \cap B_1(0)$. Taking derivative of the equation

$$\mathfrak{i}(y)(1-y) = (1-y)j$$

w.r.t. y we get $T_y\mathfrak{i}(\dot{y})(1-y) - \mathfrak{i}(y)\dot{y} = -\dot{y}j$ and, using $-\dot{y}j = j\dot{y}$, that

$$T_y\mathfrak{i}(\dot{y}) = \big(j + \mathfrak{i}(y)\big)\dot{y}(1-y)^{-1}$$

for the linearisation of $\mathfrak{i}$ at $y \in B_1(0)$ for all tangent vectors $\dot{y}$ of $B_1(0) \subset \Omega_j^{0,1}$. Using $\mathfrak{i}(y)\big(j + \mathfrak{i}(y)\big) = \big(j + \mathfrak{i}(y)\big)j$ we obtain

$$\mathfrak{i}(y)\circ T_y\mathfrak{i} = T_y\mathfrak{i}\circ j\,.$$

In other words, the linearisation $T_y\mathfrak{i}$ is complex linear for all $y \in \Omega_j^{0,1} \cap B_1(0)$. Therefore, $\mathfrak{i}: \ \Omega_j^{0,1} \cap B_1(0) \longrightarrow \mathcal{J}$ is biholomorphism.

Remark 6.6.1. With [5, Corollary 6.4] one finds a further action by conjugation via $y \longmapsto \mathrm{e}^{y} j \mathrm{e}^{-y}$ for all $y \in \Omega^{0,1}_{j}$ defining a $\mathcal{G}_j$-equivariant diffeomorphism $\Omega^{0,1}_{j} \to \mathcal{J}$: To see this it is enough to argue fibre-wise. Identifying any given tangent space of S with $\mathbb{R}^2$ we claim that the space of all orientation preserving complex multiplications $\mathcal{J}$ on $\mathbb{R}^2$ is the homogeneous space $\mathrm{PSl}_2(\mathbb{R})/\,\mathrm{PSO}_2$. Indeed, the group $\mathrm{Gl}^+_2(\mathbb{R})$ of orientation preserving invertible linear maps on $\mathbb{R}^2$ acts transitively on $\mathcal{J}$ by conjugation $A \longmapsto A\mathrm{i}A^{-1}$ and the isotropy subgroup at i is isomorphic to $\mathrm{Gl}_1(\mathbb{C})$, cf. [52, Proposition 2.48]. Normalising via the determinant and dividing out ± 1 this action descents to a transitive and faithful action of $\mathrm{PSl}_2(\mathbb{R})$ with isotropy subgroup PSO_2, as a conformal linear map that preserves the area necessarily preserves the metric. In particular, we see that $\mathcal{J}$ is the hyperbolic upper half-plane $\mathrm{PSl}_2(\mathbb{R})/\,\mathrm{PSO}_2$, cf. [52, Exercise 4.17].

The Lie algebra of $\mathrm{PSl}_2(\mathbb{R})$ decomposes as a vector space into the Lie algebra of PSO_2 and the set $\Omega^{0,1}_{\mathrm{i}}$ of all 2×2 matrices that anti-commute with i. Observe that the elements of $\Omega^{0,1}_{\mathrm{i}}$ are symmetric and trace-free. Therefore, the exponential map of the tangent space of $\mathrm{PSl}_2(\mathbb{R})/\,\mathrm{PSO}_2$ at $[1]$ can be written as

$$\Omega^{0,1}_{\mathrm{i}} \longrightarrow \mathrm{PSl}_2(\mathbb{R})/\,\mathrm{PSO}_2\,, \qquad Y \longmapsto \big[\mathrm{e}^{Y}\big]\,.$$

This map is a diffeomorphism because any $A \in \mathrm{Sl}_2(\mathbb{R})$ can be written uniquely as $A = \mathrm{e}^{S} R$ for $Y \in \Omega^{0,1}_{\mathrm{i}}$ and $R \in \mathrm{SO}_2$ by the polar form theorem. Therefore, the composition with $[A] \longmapsto A\mathrm{i}A^{-1}$

$$\Omega^{0,1}_{\mathrm{i}} \longrightarrow \mathcal{J}\,, \qquad Y \longmapsto \mathrm{e}^{Y}\mathrm{i}\mathrm{e}^{-Y}\,,$$

is the diffeomorphism we wanted.

We remark that the composition with $[A] \longmapsto A \cdot \mathrm{i}$, where $A \cdot \mathrm{i}$ denotes the action of $\mathrm{PSl}_2(\mathbb{R})$ on the upper half-plane by Möbius transformations (which preserve the hyperbolic metric) is the exponential map $Y \longmapsto \mathrm{e}^{Y} \cdot \mathrm{i}$ of the hyperbolic upper half-plane at i.

6.7 The Kodaira differential

We continue the considerations from Section 6.5. For all $j \in \mathcal{J}$ the operator $j\bar{\partial}_j : \Omega^0 \to \Omega^{0,1}_j$ is injective and

$$H^1_j := \Omega^{0,1}_j / \operatorname{Im}(j\bar{\partial}_j)$$

has real dimension $2(\#A - \#D)$. Moreover, the operator $j\bar{\partial}_j$ is $\mathcal{G}_j$-equivariant by Remark 6.4.2. Alternatively, argue with the *theorema egregium* and with Remark 6.4.5 as done in Remark 6.4.6. Therefore, conjugation

$$\psi^* y := T_\psi \psi^{-1} \circ y_\psi \circ T\psi$$

by elements $\psi \in \mathcal{G}_j$ leaves $\operatorname{Im}(j\bar{\partial}_j)$ invariant, so that $\mathcal{G}_j$ induces an action on H^1_j via $\psi^*[y] := [\psi^* y]$.

Observe, that the complex structure $y \longmapsto jy$ on $\Omega^{0,1}_j$ introduced in Section 6.6 does **not** descent to a complex structure on the quotient H^1_j. The reason is that the subspace $\operatorname{Im}(j\bar{\partial}_j)$ of $\Omega^{0,1}_j$ is **not** j-invariant as the operator $j\bar{\partial}_j : \Omega^0 \to \Omega^{0,1}_j$ is **not** complex linear, see Section 6.4.

Choose a $\mathcal{G}_j$-invariant complementary subspace $E_j \subset \Omega^{0,1}_j$ of $\operatorname{Im}(j\bar{\partial}_j)$ so that the quotient map $y \longmapsto [y]$ of $\Omega^{0,1}_j$ onto H^1_j restricts to a $\mathcal{G}_j$-equivariant linear isomorphism $E_j \to H^1_j$.

Example 6.7.1. A $\mathcal{G}_j$-invariant complementary subspace E_j can be defined as the orthogonal complement of $\operatorname{Im}(j\bar{\partial}_j)$ w.r.t. to the L^2-inner product on $\Omega^{0,1}_j$ induced by a $\mathcal{G}_j$-invariant metric on S. Such a metric can be obtained via averaging any given metric over the finite set $\mathcal{G}_j$. The obtained $\mathcal{G}_j$-invariant L^2-inner product on $\Omega^{0,1}_j$ can be symmetrised as in Remark 6.4.3 so that the action of j will be orthogonal in addition. Alternatively, a natural choice for $\mathcal{G}_j$-invariant metric on S would be the following: The area form σ_j and the metric g_j, both being $\mathcal{G}_j$-invariant, together determine a Hodge star operator on TS-valued differential forms on S, namely, via $y \longmapsto -y \circ j$. Restricting to elements $y \in \Omega^{0,1}_j$ this yields $y \longmapsto jy$ so that we obtain a $\mathcal{G}_j$-invariant L^2-inner product on $\Omega^{0,1}_j$ by

$$\langle y_1, y_2 \rangle_j := \frac{1}{2} \int_S y_1 \wedge j y_2 \, .$$

The wedge product of two TS-valued differential forms is defined component-wise w.r.t. local g_j-orthonormal frames. Using conformal coordinates one shows that the integrand $y_1 \wedge jy_2$ equals $2\mathfrak{R}(y_1 \circ y_2)\sigma_j$, where, point-wise, $\mathfrak{R}(y_1 \circ y_2)$ denotes the real part of the complex eigenvalue of the j-complex linear map $y_1 \circ y_2 \in \Omega^1$ between complex lines. This shows j- and $\mathcal{G}_j$-invariance of the metric

$$\langle y_1, y_2 \rangle_j = \int_S \mathfrak{R}(y_1 \circ y_2)\, \sigma_j\,,$$

whose induced norm $\|y\|_j$ is given by the square-root of the integral of the point-wise eigenvalues of the j-complex linear endomorphism field $y \circ y$ over S against σ_j. Nevertheless, the orthogonal complement E_j of $\operatorname{Im}(j\bar{\partial}_j)$ in $\Omega^{0,1}_j$ of any j-invariant metric will not be invariant under j as $\operatorname{Im}(j\bar{\partial}_j)$ is not invariant under j. In Section 6.9 we will construct a complex linear complement E_j.

Consider a so-called **deformation** of $\big(S, j, D, \{1, \mathrm{i}, -1\}, A\big)$, which is a map

$$\mathfrak{j}\colon\ (V_j, 0) \longrightarrow (\mathcal{J}, j)\,, \qquad y \longmapsto j(y)\,,$$

defined on an open neighbourhood $V_j \subset E_j$ of 0. If $\mathfrak{j}$ is an embedding whose image $\mathfrak{j}(V_j)$ is transverse to the orbits of $\mathcal{G}$, then the deformation $\mathfrak{j}$ is called **effective**, cf. the [67, Definition 4.2.13] of a **local slice** through j. Transversality can be expressed via the invertibility of the so-called **Kodaira differential**

$$[T_y\mathfrak{j}]\colon\ E_j \longrightarrow \Omega^{0,1}_{j(y)} \longrightarrow H^1_{j(y)}$$

for all $y \in V_j$, which is the composition of the linearisation $T_y\mathfrak{j}$ with the quotient map $[\,.\,]$. Furthermore we call the deformation $\mathfrak{j}$ **symmetric** if V_j is $\mathcal{G}_j$-invariant (e.g. taking metric balls about $0 \in E_j$ w.r.t. to the L^∞-norm $|\,.\,|_j$ from the beginning of Section 6.6 or the L^2-norm $\|\,.\,\|_j$ induced by $\langle\,.\,,\,.\,\rangle_j$ from Example 6.7.1) and $\mathfrak{j}$ is $\mathcal{G}_j$-equivariant. The latter means that for all $y \in V_j$ and $\psi \in \mathcal{G}_j$ we have that $\psi^*\big(\mathfrak{j}(y)\big) = \mathfrak{j}(\psi^* y)$ so that

$$\psi\colon\ \big(S, j(\psi^* y), D, \{1, \mathrm{i}, -1\}, A\big) \longrightarrow \big(S, j(y), D, \{1, \mathrm{i}, -1\}, A\big)$$

is a holomorphic isomorphism. Finally, the deformation $\mathfrak{j}$ is called **complex** provided that the complementary subspace $E_j \subset \Omega^{0,1}_j$ is invariant under j, i.e. is a complex linear subspace, and that $\mathfrak{j}$ is holomorphic in the sense that the differential $T\mathfrak{j}$ is complex linear, i.e. $\mathfrak{j}(y)\circ T_y\mathfrak{j} = T_y\mathfrak{j}\circ j$ for all $y \in V_j \subset E_j$.

Example 6.7.2. The Cayley transformation from Section 6.6 or the conjugation by the exponential map studied in Remark 6.6.1 yield symmetric effective deformations $\mathfrak{j}$ restricting

$$y \longmapsto (1+jy)j(1+jy)^{-1} \quad \text{or} \quad y \longmapsto \mathrm{e}^{jy} j \mathrm{e}^{-jy}$$

to $V_j = E_j \cap \Omega^{0,1}_j \cap B_1(0)$ or to $V_j = E_j$, resp. Indeed, the differentials at $0 \in V_j$ in direction of $\dot{y} \in E_j$ are given by $\dot{y} \longmapsto 2\dot{y}$. For complex linear E_j the corresponding deformation will be complex.

Remark 6.7.3. The **Kodaira differential is natural** in the following sense: Consider deformations $\mathfrak{j}\colon\ V_j \ni y \longmapsto j(y)$ and $\mathfrak{k}\colon\ V_k \ni z \longmapsto k(z)$ of $(S, j, D, \{1, \mathrm{i}, -1\}, A)$ and $(S, k, D, \{1, \mathrm{i}, -1\}, A)$, resp. Choose a linear isomorphism $\zeta\colon\ E_j \to E_k$, which we will write as $\zeta\colon\ y \longmapsto z(y)$. Let $y \longmapsto \varphi(y)$ be a smooth family of diffeomorphisms

$$\varphi(y)\colon\ \big(S, j(y), D, \{1, \mathrm{i}, -1\}, A\big) \longrightarrow \Big(S, k\big(z(y)\big), D, \{1, \mathrm{i}, -1\}, A\Big)$$

defined on $V_j \cap \zeta^{-1}(V_k)$ and deforming the equivalence

$$\varphi = \varphi(0)\colon\ \big(S, j, D, \{1, \mathrm{i}, -1\}, A\big) \longrightarrow \big(S, k, D, \{1, \mathrm{i}, -1\}, A\big)\,.$$

Then the Kodaira differentials at 0 satisfy

$$[T_0\mathfrak{j}] = \varphi^*\circ[T_0\mathfrak{k}]\circ\zeta\,,$$

cf. [42, Proposition 1.6]. In particular, $\mathfrak{j}$ is effective if and only if $\mathfrak{k}$ is.

Indeed, choose $\dot{y} \in E_j$ and write $\mathfrak{j}_t = \mathfrak{j}(t\dot{y})$, $\varphi_t = \varphi(t\dot{y})$, and $\mathfrak{k}_t = \mathfrak{k}\big(\zeta(t\dot{y})\big)$ for $t \in (-1,1)$ and take the Lie derivative of $\mathfrak{j}_t = \varphi_t^*\mathfrak{k}_t$. By conformal invariance of the Cauchy–Riemann operator $\varphi^*\circ k\bar{\partial}_k = j\bar{\partial}_j\circ\varphi^*$ (see Remark 6.4.2) we obtain

$$T_0\mathfrak{j}(\dot{y}) = 2j\bar{\partial}_j(\varphi^*X) + \varphi^*\Big(T_0\mathfrak{k}\big(\zeta(\dot{y})\big)\Big)$$

similarly to Remark 6.4.6, where X is the velocity vector of the path $t \longmapsto \varphi_t$ in $\mathcal{G}$ at 0. As the restriction of φ_t to the special points in $D \cup \{1, \mathrm{i}, -1\} \cup A$ is constant by continuity of $t \longmapsto \varphi_t$ we obtain that $X \in \Omega^0$. Hence, $\varphi^* X \in \Omega^0$ and therefore $2j\bar{\partial}_j(\varphi^* X)$ represents the zero class. The claim follows now because φ defines a well–defined isomorphism $\varphi^* : \; H^1_k \to H^1_j$ via $\varphi^*[y] := [\varphi^* y]$,

$$\varphi^* y := T_\varphi \varphi^{-1} \circ y_\varphi \circ T\varphi \,,$$

by the previous argument. Hence,

$$[T_0 \mathfrak{j}(\dot{y})] = \varphi^* \Big[T_0 \mathfrak{k}\big(\zeta(\dot{y})\big) \Big]$$

for all $\dot{y} \in E_j$.

If both deformations $\mathfrak{j}$ and $\mathfrak{k}$ are effective, then there is a *tautological* choice for an isomorphism ζ for the given diffeomorphism φ, namely $\zeta = [T_0 \mathfrak{k}]^{-1} \circ (\varphi^*)^{-1} \circ [T_0 \mathfrak{j}]$. On the other hand, in the situation of Example 6.7.1, where the local slices are constructed via orthogonal complements, we get $\varphi^* E_k = E_j$ because g_j and $\varphi^* g_k$ are conformally equivalent. This allows the choice $(\varphi^*)^{-1} : \; E_j \to E_k$ for ζ and yields

$$[T_0 \mathfrak{j}] = \varphi^* \circ [T_0 \mathfrak{k}] \circ (\varphi^*)^{-1} \,.$$

6.8 Orbifold structure – fixed stable nodal type

In the situation of Section 6.7 we assume for the moment that the isotropy group $\mathcal{G}_j$ is trivial. Invertibility of the Kodaira differential $[T_y \mathfrak{j}]$ implies that the linear subspace $T_y \mathfrak{j}(E_j)$ in $\Omega^{0,1}_{j(y)}$ is complementary to $\operatorname{Im}\big(j(y)\bar{\partial}_{j(y)}\big)$. With the arguments from [61, p. 634/5] and [67, Theorem 4.2.14] we obtain that

$$\mathcal{G} \times V_j \longrightarrow \mathcal{J} \,, \qquad (\varphi, y) \longmapsto \varphi^*\big(j(y)\big) \,,$$

is a diffeomorphism onto a neighbourhood of the $\mathcal{G}$-orbit of $j = \mathrm{id}^*\big(j(0)\big)$ for a sufficiently small open neighbourhood $V_j \subset E_j$ of 0. In other words, the map $(\mathrm{id}, y) \longmapsto j(y)$ induces a homeomorphism

$$(\mathrm{id}, y) \longmapsto \big[S, j(y), D, \{1, \mathrm{i}, -1\}, A\big] = \big[j(y)\big]$$

defined on V_j onto a neighbourhood of $\big[S,j,D,\{1,\mathrm{i},-1\},A\big]=[j]$ in $\mathcal{R}_\tau=\mathcal{J}/\mathcal{G}$. The inverse serves as a chart of a smooth manifold structure on $\mathcal{R}_\tau$.

For $\mathcal{G}_j$ non-trivial we give an equivariant version of the above construction: Assuming the situation of Section 6.7 we consider a symmetric effective deformation $\mathfrak{j}\colon (V_j,0)\to(\mathcal{J},j)$, which we also denote by $y\longmapsto j(y)$. We would like to find a $\mathcal{G}$-equivariant diffeomorphism

$$\mathcal{G}\times_{\mathcal{G}_j}V_j\longrightarrow\mathcal{J}\,,\qquad(\varphi,y)\longmapsto\varphi^*\big(j(y)\big)\,,$$

onto a neighbourhood of the $\mathcal{G}$-orbit of $j=\mathrm{id}^*\big(j(0)\big)$, where $\mathcal{G}\times_{\mathcal{G}_j}V_j$ is the quotient of $\mathcal{G}\times V_j$ by the action $(\varphi,y)\longmapsto\big(\psi^{-1}\circ\varphi,\psi^*y\big)$, $\psi\in\mathcal{G}_j$.

In order to do so we would like to find an isomorphism

$$\varphi\colon\;\big(S,j(y),D,\{1,\mathrm{i},-1\},A\big)\longrightarrow\big(S,k,D,\{1,\mathrm{i},-1\},A\big)$$

for given $k\in\mathcal{J}$ sufficiently close to $j\in\mathcal{J}$ and for some $y\in V_j$. Holomorphicity of φ translates to $\varphi^*k=j(y)$. In other words, (φ,y,k) is a zero of the **non-linear Cauchy-Riemann operator**

$$F(\varphi,y,k):=\frac{1}{2}\Big(T\varphi+k\circ T\varphi\circ j(y)\Big)\,,$$

which is a section into the bundle $\mathcal{E}\to\mathcal{G}\times V_j\times\mathcal{J}$, whose fibre $\mathcal{E}_{(\varphi,y,k)}$ is the vector space of sections of the bundle of complex anti-linear bundle homomorphisms from $\big(TS,j(y)\big)$ to (TS,k).

Setting $F_k:=F(\,.\,,\,.\,,k)$ we will write the solutions (φ,y) of $F_k(\varphi,y)=0$ as a function of k via the implicit function theorem. The linearisation

$$T_{(\mathrm{id},0)}F_j\,:\,\Omega^0\times E_j\longrightarrow\Omega^{0,1}_j$$

of $(\varphi,y)\longmapsto F_j(\varphi,y)$ at $(\mathrm{id},0)$ equals

$$(X,\dot{y})\longmapsto\bar{\partial}_jX+\tfrac{1}{2}j\cdot\big(T_0\mathfrak{j}\big)(\dot{y})\,.$$

With Sections 6.5 and 6.7 the operator

$$-2j\cdot T_{(\mathrm{id},0)}F_j=-2j\bar{\partial}_j\oplus T_0\mathfrak{j}$$

is an isomorphism.

The **implicit function theorem** combined with an intermediate Sobolev completion and a subsequent elliptic regularity argument as in [61, p. 634/5] or in [67, Theorem 4.2.14] implies: There exists an open neighbourhood $\mathcal{K} \subset \mathcal{J}$ of j, a possibly smaller $\mathcal{G}_j$-invariant open neighbourhood $V_j \subset E_j$ of 0, an open neighbourhood $\mathcal{H} \subset \mathcal{G}$ of id such that the sets $\psi^*\mathcal{H}$ are pair-wise disjoint for all $\psi \in \mathcal{G}_j$, and a unique map

$$\Phi : \ (\mathcal{K}, j) \longrightarrow \big(\mathcal{H} \times V_j, (\mathrm{id}, 0)\big)$$

such that

$$F_k(\varphi, y) = 0 \quad \Longleftrightarrow \quad (\varphi, y) = \Phi(k)\,,$$

whenever $(\varphi, y, k) \in \mathcal{H} \times V_j \times \mathcal{K}$. Notice, that uniqueness implies $\Phi\big(j(y)\big) = (\mathrm{id}, y)$ for all $y \in V_j$.

For all $\psi \in \mathcal{G}_j$ define a map

$$\Phi_\psi : \ (\mathcal{K}, j) \longrightarrow \big(\psi^*\mathcal{H} \times V_j, (\psi, 0)\big)$$

setting

$$\Phi_\psi(k) := \psi^*\big(\Phi(k)\big)\,.$$

Observe that $\Phi_\psi\big(j(y)\big) = \big(\psi, \psi^* y\big)$ for all $y \in V_j$. Moreover, by symmetry of the deformation $\mathfrak{j}$, which reads as $\psi^* \circ \mathfrak{j} = \mathfrak{j} \circ \psi^*$ for all $\psi \in \mathcal{G}_j$, we get $\psi^* \circ F_k = F_k \circ \psi^*$, and hence $F_k\big(\Phi_\psi(k)\big) = 0$ for all $k \in \mathcal{J}$ and $\psi \in \mathcal{G}_j$. In other words, for all $\psi \in \mathcal{G}_j$ there exists a unique map Φ_ψ with the above properties such that

$$F_k(\varphi, y) = 0 \quad \Longleftrightarrow \quad (\varphi, y) = \Phi_\psi(k)\,,$$

whenever $(\varphi, y, k) \in \psi^*\mathcal{H} \times V_j \times \mathcal{K}$.

We get the following **global uniqueness statement**: There exists potentially smaller neighbourhoods V_j and $\mathcal{K}$ such that for all solutions $(\varphi, y, k) \in \mathcal{G} \times V_j \times \mathcal{K}$ of $F(\varphi, y, k) = 0$ there exists a unique $\psi \in \mathcal{G}_j$ such that $\varphi \in \psi^*\mathcal{H}$ and $(\varphi, y) = \Phi_\psi(k)$. This follows arguing by contradiction with properness of the action $\mathcal{G} \times \mathcal{J} \to \mathcal{J}$, $(\phi, j) \longmapsto \phi^* j$, see Remark 6.3.2.

Based on the current results of the implicit function theorem an **orbifold chart** about $j \in \mathcal{J}$ can be obtained as follows: Let $\mathcal{U}_j$ be the image of $\mathcal{K}$ under the projection $[\,.\,]\colon\ \mathcal{J} \to \mathcal{R}_\tau = \mathcal{J}/\mathcal{G}$, i.e.

$$\mathcal{U}_j = [\mathcal{K}]\,,$$

so that, in particular, $\mathcal{U}_j$ is open according to the quotient topology described in Remark 6.3.2. With help of Φ we find $\varphi \in \mathcal{H}$ and $y \in V_j$ for each given $k \in \mathcal{K}$ such that $\varphi^* k = \mathfrak{j}(y)$. Hence, $[k] = \big[\mathfrak{j}(y)\big]$, so that $\mathcal{U}_j$ is the set of all isomorphism classes $\big[S, \mathfrak{j}(y), D, \{1, \mathrm{i}, -1\}, A\big] = \big[\mathfrak{j}(y)\big]$ with $y \in V_j$.

The isotropy group $\mathcal{G}_j$ acts linearly on V_j by conjugation. In view of the metric obtained by restriction of the metric described in Example 6.7.1 this action is orthogonal. Hence, the action is **effective**, i.e. only for $\mathrm{id} \in \mathcal{G}_j$ all points of V_j are fixed points. The map

$$\mathfrak{p}_j\colon\ V_j/\mathcal{G}_j \longrightarrow \mathcal{U}_j\,, \qquad [y] \longmapsto \big[\mathfrak{j}(y)\big]$$

is well-defined by symmetry of the deformation $\mathfrak{j}$; $\mathfrak{p}_j$ is continuous because the deformation $\mathfrak{j}$ is and the respective quotient maps are open and continuous as explained in Remark 6.3.2.

We claim that

$$\mathcal{U}_j \longrightarrow V_j/\mathcal{G}_j\,, \qquad [k] \longmapsto \big[\Phi^2(k)\big]\,,$$

is the inverse map of $\mathfrak{p}_j$, where $\Phi^2(k)$ denotes the second component of $\Phi(k)$. First of all the map is well-defined by the following **compatibility condition for uniformisers**: Write $\big[\mathfrak{j}(y_1)\big] = [k] = \big[\mathfrak{j}(y_2)\big]$ for $y_1, y_2 \in V_j$ and choose an isomorphism

$$\phi\colon\ \big(S, \mathfrak{j}(y_1), D, \{1, \mathrm{i}, -1\}, A\big) \longrightarrow \big(S, \mathfrak{j}(y_2), D, \{1, \mathrm{i}, -1\}, A\big)\,,$$

whose existence is guaranteed by the definition of the equivalence relation. We claim that

$$\phi \in \mathcal{G}_j \quad \text{and} \quad y_1 = \phi^* y_2\,.$$

Indeed, we get $F\big(\phi, y_1, \mathfrak{j}(y_2)\big) = 0$, so that we find a unique $\psi \in \mathcal{G}_j$ such that $\phi \in \psi^*\mathcal{H}$ and $(\phi, y_1) = \Phi_\phi\big(\mathfrak{j}(y_2)\big)$ by the above global uniqueness statement. Because of $\Phi_\phi\big(\mathfrak{j}(y_2)\big) =$

$(\psi, \psi^* y_2)$ we obtain $\phi = \psi \in \mathcal{G}_j$ and $y_1 = \psi^* y_2$. Being well-defined follows now with the equation $\Phi\big(j(y)\big) = (\mathrm{id}, y)$ for $y \in \{y_1, y_2\}$. Similarly, in order to verify the two-sided inverse property we obtain

$$[y] \longmapsto \big[j(y)\big] \longmapsto \Big[\Phi^2\big(j(y)\big)\Big] = [y]$$

and, writing $\varphi^* k = j(y)$ for $y \in V_j$,

$$[k] \longmapsto \big[\Phi^2(k)\big] = [y] \longmapsto \big[j(y)\big] = [k]\,.$$

Therefore, the assignment $[k] \longmapsto \big[\Phi^2(k)\big]$ is the inverse map $\mathfrak{p}_j^{-1}\colon\ \mathcal{U}_j \longrightarrow V_j/\mathcal{G}_j$. The inverse $\mathfrak{p}_j^{-1}$ is continuous as well. This follows because $k \longmapsto \Phi^2(k)$ is is continuous and the involved quotient maps are open and continuous, cf. Remark 6.3.2. Consequently, the $\mathcal{G}_j$-invariant map

$$[\mathfrak{j}]\colon\ V_j \longrightarrow \mathcal{U}_j$$

induces a homeomorphism $\mathfrak{p}_j$ form $V_j/\mathcal{G}_j$ onto $\mathcal{U}_j$. In other words, $\big(V_j, \mathcal{G}_j, \mathfrak{p}_j^{-1}\big)$ is an **orbifold chart for $\mathcal{R}_\tau$ about** $[j] = \big[S, j, D, \{1, \mathrm{i}, -1\}, A\big]$ and $\big(E_j, \mathcal{G}_j, V_j, \mathcal{U}_j, [\mathfrak{j}]\big)$ is a τ**-uniformiser** by definition.

Remark 6.8.1. By the above implicit function theorem we find k-holomorphic maps $\varphi \in \mathcal{G}$ defined on $\big(S, j(y)\big)$, where $k \in \mathcal{K}$ and $y \in V_j$, so that (φ, y) is a solution of $F_k = 0$. By global uniqueness and a potential precomposition with the inverse of an element in $\mathcal{G}_j$ making use of the symmetry property of $\mathfrak{j}$ we can assume that $\varphi \in \mathcal{H}$ and write $\Phi(k) = (\varphi, y)$. By local uniqueness and $\mathcal{G}_j$-invariance of the solution set we obtain that $\{F_k = 0\}$ is equal to the set of all $\big(\psi^*\varphi, \psi^* y\big)$, $\psi \in \mathcal{G}_j$. Hence, for $y = \Phi^2(k)$ the solution set $\big\{F_k(\,.\,, y) = 0\big\}$ is given by $\{\psi^*\varphi \mid \psi \in \mathcal{G}_{j,y}\}$, where $\mathcal{G}_{j,y}$ denotes the **stabiliser at** $y \in V_j$ of the induced action by $\mathcal{G}_j$ on E_j. For that we will also write

$$\mathcal{G}_{j,y} := \big(\mathcal{G}_j|_{E_j}\big)_y$$

Notice that $\mathcal{G}_{j,0} = \mathcal{G}_j$.

For all $\hat{\psi} \in \mathcal{G}_k$ we have that $F_k(\hat{\psi}\circ\varphi, y) = T\hat{\psi}\circ F_k(\varphi, y)$, so that the set $\big\{\hat{\psi}\circ\varphi \mid \hat{\psi} \in \mathcal{G}_k\big\}$ is contained in $\{\psi^*\varphi \mid \psi \in \mathcal{G}_{j,y}\}$. For the converse observe that $\varphi\circ\psi\circ\varphi^{-1}$, $\psi \in \mathcal{G}_{j,y}$, is an element of $\mathcal{G}_k$ and $(\varphi\circ\psi\circ\varphi^{-1})\circ\varphi = \psi^*\varphi$. Therefore, we obtain two descriptions

$$\left\{\hat{\psi}\circ\varphi \;\middle|\; \hat{\psi} \in \mathcal{G}_k\right\} = \{\psi^*\varphi \mid \psi \in \mathcal{G}_{j,y}\}$$

for the solution set $\{F_k(\,.\,,y) = 0\}$ and

$$\mathcal{G}_{j,y} \longrightarrow \mathcal{G}_k\,, \qquad \psi \longmapsto \varphi\circ\psi\circ\varphi^{-1}$$

is an **isomorphism of isotropy groups** with inverse $\hat{\psi} \longmapsto \varphi^{-1}\circ\hat{\psi}\circ\varphi$.

In order to describe the **transformation behaviour of orbifold charts** we consider symmetric effective deformations $\mathfrak{j}\colon\ V_j \ni y \longmapsto j(y)$ and $\mathfrak{k}\colon\ V_k \ni z \longmapsto k(z)$ of $(S, j, D, \{1, \mathrm{i}, -1\}, A)$ and $(S, k, D, \{1, \mathrm{i}, -1\}, A)$, resp. About the respective τ-uniformiser $(E_j, \mathcal{G}_j, V_j, \mathcal{U}_j, [\mathfrak{j}])$ and $(E_k, \mathcal{G}_k, V_k, \mathcal{U}_k, [\mathfrak{k}])$ we assume that $\mathcal{U}_j \cap \mathcal{U}_k = \emptyset$. Hence, we find $\varphi \in \mathcal{G}$ such that $\varphi^*k = j$.

We claim that we can assume that $j = j(0) = k(0) = k$. Indeed, define a deformation $\mathfrak{k}'\colon\ V_{k'} \ni z \longmapsto k'(z)$,

$$k'(z) = \varphi^*\Big(k\big((\varphi^{-1})^*z\big)\Big)\,,$$

of $(S, j, D, \{1, \mathrm{i}, -1\}, A)$, whose domain is the subset $V_{k'} := \varphi^*V_k$ of the complementary space $E_{k'} := \varphi^*E_k$. For the latter use that $g_{k'}$ and φ^*g_k are conformally equivalent. By the naturality of the Kodaira differential we have

$$[T_0\mathfrak{k}'] = \varphi^*\circ[T_0\mathfrak{k}]\circ(\varphi^{-1})^*\,,$$

so that $\mathfrak{k}'$ is effective, see Remark 6.7.3. In view of Remark 6.8.1 the deformation $\mathfrak{k}'$ is symmetric because for all $\psi \in \mathcal{G}_k$, for which we have $\psi^*\circ\mathfrak{k} = \mathfrak{k}\circ\psi^*$, it follows that $\big(\varphi^{-1}\circ\psi\circ\varphi\big)^*\circ\mathfrak{k}' = \mathfrak{k}'\circ\big(\varphi^{-1}\circ\psi\circ\varphi\big)^*$.

Therefore, we consider deformations $\mathfrak{j}$ and $\mathfrak{k}$ such that $j = j(0) = k(0) = k$. In view of the implicit function theorem above there exists a $k(z)$-holomorphic map $\varphi(z) \in \mathcal{G}$ close to $\mathrm{id} \in \mathcal{G}_j$ defined on $\big(S, j\big(y(z)\big)\big)$ via the locally unique smooth map $\Phi\big(k(z)\big) = \big(\varphi(z), y(z)\big)$ such that $\Phi(j) = (\mathrm{id}, 0)$. Switching the roles results into a $j(y)$-holomorphic map $\hat{\varphi}(y) \in \mathcal{G}$ close to $\mathrm{id} \in \mathcal{G}_j$ defined on $\big(S, k(z(y))\big)$ via the locally unique smooth map $\Phi\big(j(y)\big) = \big(\hat{\varphi}(y), z(y)\big)$ such that $\Phi(j) = (\mathrm{id}, 0)$. Comparing both solutions using uniqueness yields

$$\hat{\varphi}\big(y(z)\big) = \big(\varphi(z)\big)^{-1}$$

as well as $z = z\big(y(z)\big)$ and $y = y\big(z(y)\big)$. Therefore, after shrinking the domains according to the implicit function theorem if necessary, which results into $\mathcal{U}_j = \mathcal{U}_k$, we obtain maps

$$V_j \longrightarrow V_k\,, \quad y \longmapsto z(y) \qquad \text{and} \qquad V_k \longrightarrow V_j\,, \quad z \longmapsto y(z)\,,$$

which are smooth and inverse to each other such that $[\mathfrak{k}]\circ\big(y \longmapsto z(y)\big) = \hat{\varphi}(y)^*[\mathfrak{j}]$. This results into a coordinate change of an orbifold structure because the construction is done in a $\mathcal{G}_j$-equivariant fashion: For that use Remark 6.8.1 and observe that the solution set $\big\{F_{k(z)}(\,.\,,y) = 0\big\}$ is given by

$$\Big\{\hat{\psi}\circ\varphi(z) \;\Big|\; \hat{\psi} \in \mathcal{G}_{k(z)}\Big\} = \big\{\psi^*\varphi(z) \mid \psi \in \mathcal{G}_{j,y}\big\}\,.$$

As for manifolds we obtain:

Proposition 6.8.2. *The above constructed orbifold charts provide the nodal Riemann moduli space $\mathcal{R}_\tau$ for the stable nodal type τ with the structure of an orbifold of real dimension $2\big(\#A - \#D\big)$, whose isotropy groups at $[j] \in \mathcal{R}_\tau$ are given by $\mathcal{G}_j$ up to conjugation.*

Referring to the current situation we define

$$\mathbf{T}_{j,k} := \Big\{(\varphi, y, z) \;\Big|\; F_{k(z)}(\varphi, y) = 0\Big\} \subset \mathcal{G} \times V_j \times V_k$$

provided with the subspace topology and call the projection $s\colon\ \mathbf{T}_{j,k} \to V_j$ onto V_j the **source map**; the the projection $t\colon\ \mathbf{T}_{j,k} \to V_k$ onto V_k the **target map**. These maps come with inverses $y \longmapsto \big(\hat{\varphi}^{-1}(y), y, z(y)\big)$ and $z \longmapsto \big(\varphi(z), y(z), z\big)$, resp., where $\hat{\varphi}\big(y(z)\big) = \big(\varphi(z)\big)^{-1}$ as above. Hence, s and t are homeomorphisms providing $\mathbf{T}_{j,k}$ with the structure of a smooth manifold of dimension $2\big(\#A - \#D\big)$, and $t\circ s^{-1}$ and $s\circ t^{-1}$ correspond to the above transition maps $y \longmapsto z(y)$ and $z \longmapsto y(z)$, resp. By the properness argument in Remark 6.3.2 the map $s \times t\colon\ \mathbf{T}_{j,k} \to V_j \times V_k$ is proper.

In other words, we obtain an **étale proper Lie groupoid** $(R_\tau, \mathbf{R}_\tau)$, which means the following: Take a sequence of τ-uniformisers $\big(E_{j_\nu}, \mathcal{G}_{j_\nu}, V_{j_\nu}, \mathcal{U}_{j_\nu}, [\mathfrak{j}_\nu]\big)$ such that $\bigcup_\nu \mathcal{U}_{j_\nu}$ covers $\mathcal{R}_\tau$. The **objects** are given by

$$R_\tau := \bigsqcup_\nu V_{j_\nu}$$

and the **morphisms** are

$$\mathbf{R}_\tau := \bigsqcup_{\nu,\mu} \mathbf{T}_{j_\nu, j_\mu} .$$

Morphisms can be composed whenever the corresponding target and source coincide. The resulting composition is smooth, has a unit and each morphism admits a smooth inverse. Furthermore all mentioned structure maps are smooth. The nodal Riemann moduli space $\mathcal{R}_\tau$ for the stable nodal type τ appears now as **orbit space**

$$\mathcal{R}_\tau = R_\tau / \sim ,$$

where two objects between which there exists a morphism are considered to be equivalent.

6.9 Skyscraper deformation

A symmetric effective deformation $\mathfrak{j}$ is called a **skyscraper deformation** if there exists a $\mathcal{G}_j$-invariant neighbourhood $U \subset S$ of ∂S together with the special points $|D| \cup \{1, \mathrm{i}, -1\} \cup A$ on which the **deformation is stationary**, i.e. if $j(y) = j$ restricted to U for all $y \in V_j$. In view of the examples of symmetric effective deformations in Section 6.6 and Remark 6.6.1 skyscraper deformations can be obtained by restriction of symmetric effective deformations to a complementary $\mathcal{G}_j$-invariant vector space E_j of $\mathrm{Im}(j\bar{\partial}_j)$ whose elements vanish on U.

In order to construct such a vector space E_j we denote by $\mathcal{X}$ the space of all smooth vector fields on S that are tangent to the boundary along ∂S and admit 3 zeros on each connected component of S equal to special points in $|D| \cup \{1, \mathrm{i}, -1\} \cup A$; on the disc component the zeros are required to be $1, \mathrm{i}, -1$. The operator $j\bar{\partial}_j$ restricted to $\mathcal{X}$ induces an isomorphism $\mathcal{X} \to \Omega^{0,1}_j$.

We begin with the un-isotropic situation $\mathcal{G}_j = \{\mathrm{id}\}$. In order to construct a complement of Ω^0 in $\mathcal{X}$ we write z_k for the elements of $|D| \cup A$ and choose local holomorphic charts $(\mathbb{C}, \mathrm{i})$ for (S, j) about the special points z_k. We require that the chart domains are mutually disjoint and contained in $S \setminus \partial S$. Let f_k be smooth cut off functions on S that have their supports in the interior of r-disc neighbourhoods about z_k w.r.t. g_j contained in the chosen chart domains; the f_k are required to by constantly 1 on the $r/2$-disc neighbourhoods about z_k. Given $X \in \mathcal{X}$ we define vector fields X_k on S. We require that the X_k are given by $X(z_k)f_k$ in the chosen charts extended by zero to S. Observe that the X_k vanish for special points that correspond to the zeros defining $\mathcal{X}$. Moreover, the X_k are holomorphic on the $r/2$-discs.

Let $P\colon\ \mathcal{X} \to \mathcal{X}$ be the projector, i.e. $P^2 = P$, given by

$$P(X) := X - \sum_k X_k\,.$$

Observe that P restricts to the identity on $P(\mathcal{X}) = \Omega^0$. The desired complement of Ω^0 in $\mathcal{X}$ is $(1-P)(\mathcal{X})$ as $1-P$ is a projector as well. The elements of $j\bar{\partial}_j(1-P)(\mathcal{X})$ vanish on a neighbourhood of all special points $|D| \cup \{1, \mathrm{i}, -1\} \cup A$ and of ∂S. Moreover, the dimension of $(1-P)(\mathcal{X})$ equals

$$2\big(\#A - \#D\big)$$

by the result of the computation of Section 6.5 multiplied by -1.

Finally we set $E_j := j\bar{\partial}_j(1-P)(\mathcal{X})$. The elements of E_j vanish on a neighbourhood of the union of the special points $|D| \cup \{1, \mathrm{i}, -1\} \cup A$ and of the boundary ∂S. Furthermore the isomorphism $j\bar{\partial}_j\colon\ \mathcal{X} \to \Omega^{0,1}_j$ sends the splitting $(1-P)(\mathcal{X}) \oplus \Omega^0$ of $\mathcal{X}$ to the splitting $E_j \oplus \mathrm{Im}(j\bar{\partial}_j)$ of $\Omega^{0,1}_j$.

We treat the case of a non-vanishing isotropy group $\mathcal{G}_j$, which acts by permutations on $\{z_k\} = |D| \cup A$. It suffices to change the above projector P by replacing the vector fields X_k by $\mathcal{G}_j$-invariant vector fields $\hat{X}_k$. For that denote by $B_r(z_k)$, $r > 0$, the interior of the above r-discs. Observe that the disjoint union of the $B_r(z_k)$ is $\mathcal{G}_j$-invariant as $\mathcal{G}_j$ acts by isometries on (S, g_j).

For each z_k we assign a **partner point** $w_k \in B_{r/2}(z_k) \setminus \{z_k\}$ requiring that the $\psi(w_k)$ are pair-wise distinct for all $\psi \in \mathcal{G}_j$ and all k. Notice, that the distance between z_k and its partner w_k is $\mathcal{G}_j$-invariant for all k. We choose $\varepsilon > 0$ such that $B_\varepsilon(w_k) \subset B_{r/2}(z_k) \setminus \{z_k\}$ for all k. Furthermore, we reqiure that the $\psi\big(B_\varepsilon(w_k)\big)$ are pair-wise disjoint for all $\psi \in \mathcal{G}_j$ and for all k. Modify the cut off functions f_k so that f_k has support in

$$B_r(z_k) \ \setminus \bigsqcup_{\ell \neq k \text{ and } \psi \in \mathcal{G}_j} \psi\big(\overline{B_{\varepsilon/2}(w_\ell)}\big)$$

and is equal to 1 on

$$B_{r/2}(z_k) \ \setminus \bigsqcup_{\ell \neq k \text{ and } \psi \in \mathcal{G}_j} \psi\big(\overline{B_{\varepsilon}(w_\ell)}\big) .$$

In particular, for all k, we get $f_k(w_k) = 1$ and $f_k\big(\psi(w_\ell)\big) = 0$ for all $\ell \neq k$ and $\psi \in \mathcal{G}_j$. With the cut off functions f_k modified we define X_k for given $X \in \mathcal{X}$ as in the un-isotropic case.

We define the symmetrisations via

$$\hat{X}_k := \sum_{\psi \in \mathcal{G}_j} \psi^* X_k .$$

We have $\phi^* \hat{X}_k = \hat{X}_k$ for all $\phi \in \mathcal{G}_j$ and for all k because $\mathcal{G}_j$ acts on itself via composition permuting $\mathcal{G}_j$. The $\hat{X}_k$ that are assigned to the zeros z_k of $\mathcal{X}$ vanish; the remaining $\hat{X}_k$ span a $2\big(\#A - \#D\big)$-dimensional vector space because

$$\hat{X}_k(w_\ell) = X_\ell(w_\ell)$$

for all k, ℓ. A basis can be obtained by taking $X \in \mathcal{X}$ with $X(z_k)$ non-zero, so that the corresponding partners $X_k(w_k)$ do not vanish.

Remark 6.9.1. Observe that the elements of $(1 - P)(\mathcal{X})$, which are linear combinations of the vector fields $\hat{X}_k$ constructed above, are vector fields on S that vanish on the boundary ∂S. Therefore, the complex structure j on S preserves $(1 - P)(\mathcal{X})$ and defines a **complex structure** on $E_j = \bar{\partial}_j (1 - P)(\mathcal{X})$ as $\bar{\partial}_j$ commutes with j, so that E_j is a complex vector space of complex dimension $\#A - \#D$. Consequently taking the complex deformations form Example 6.7.2 w.r.t. E_j yields holomorphic skyscraper deformations.

Remark 6.9.2. In the above construction the radii r_k of the discs

$$D_{r_k/2}(z_k) := \overline{B_{r_k/2}(z_k)}$$

are necessarily constant on the orbits of the $\mathcal{G}_j$-action on the points $z_k \in |D| \cup A$ because $\mathcal{G}_j$ acts on the discs $D_{r_k/2}(z_k)$ by isometries of (S, g_j); but the radii are allowed to vary on distinct orbits $\mathcal{G}_j z_k$. For a selection of orbit-wise constant radii r_k denoted by $\mathbf{r}$ and the disjoint union

$$\mathbf{D}_{j,\mathbf{r}} := \coprod_{z_k \in |D| \cup A} D_{r_k/2}(z_k)$$

a skyscraper deformation $\mathfrak{j}$ that is stationary on $\mathbf{D}_{j,\mathbf{r}}$ can be constructed by the above arguments. Given a neighbourhood U of $|D| \cup A$ one can choose $\mathbf{r}$ so small such that $\mathbf{D}_{j,\mathbf{r}} \subset U$. This yields an example of a **small disc structure** $\mathbf{D}_j$, which by definition is a $\mathcal{G}_j$-invariant disjoint union of discs D_z, $z \in |D| \cup A$, contained in a given neighbourhood of $|D| \cup A$ such that $z \in D_z$ for all $z \in |D| \cup A$. Furthermore the D_z are the image of a smooth embedding of the closed unit disc $\mathbb{D}$ into S. Observe that a k-holomorphic map $\varphi \in \mathcal{G}$ defined on (S, j) sends $\mathbf{D}_{j,\mathbf{r}}$ diffeomorphically onto a small disc structure $\mathbf{D}_k$ on which the φ-push-forward skyscraper deformation $\mathfrak{k}$ of $\mathfrak{j}$ is stationary.

Using small disc structures $\mathbf{D}_j$ orbifold charts $\left(V_j, \mathcal{G}_j, \mathfrak{p}_j^{-1}\right)$ and τ-uniformiser $\left(E_j, \mathcal{G}_j, V_j, \mathcal{U}_j, [\mathfrak{j}]\right)$ for $\mathcal{R}_\tau$ about $[j] = \left[S, j, D, \{1, \mathrm{i}, -1\}, A\right]$ can be constructed as in Section 6.8 using skyscraper deformations $\mathfrak{j}$ that are stationary on $\mathbf{D}_j$ exclusively. The transformation behaviour encoded in the $\mathbf{T}_{j,k}$ is compatible with skyscraper deformations which are stationary on disc structures. Correspondingly, a neighbourhood base of the topology on $\mathcal{R}_\tau = \mathcal{J}/\mathcal{G}$ described in Remark 6.3.2 can be given by the family of subsets whose elements $[k]$ can be represented by complex structures k that belong to an open subset of $\mathcal{J}$ and that satisfy $k = j$ restricted to some small disc structure $\mathbf{D}_j$. This follows with the implicit function theorem formulated in Section 6.8.

Remark 6.9.3. In view of the proceeding Remarks 6.9.1 and 6.9.2 the orbifold structure on $\mathcal{R}_\tau$ ensured in Proposition 6.8.2 and the subsequently described étale proper Lie groupoid structure admit subatlases generated by complex skyscraper deformations, so that the respective substructures are complex. For that one needs to verify that the transition maps $t\circ s^{-1}$ and $s\circ t^{-1}$ are holomorphic. In terms of complex skyscraper deformations $\mathfrak{j}\colon\ V_j \ni y \longmapsto j(y)$ and $\mathfrak{k}\colon\ V_k \ni z \longmapsto k(z)$ of $\big(S, j, D, \{1, \mathrm{i}, -1\}, A\big)$ and $\big(S, k, D, \{1, \mathrm{i}, -1\}, A\big)$, resp., such that there exists $\varphi \in \mathcal{G}$ with $\varphi^*k = j$ the transition maps are given by $y \longmapsto z(y)$ and $z \longmapsto y(z)$, resp. With the description before Proposition 6.8.2 we get $[\mathfrak{k}]\circ\big(y \longmapsto z(y)\big) = \hat{\varphi}(y)^*[\mathfrak{j}]$ and $[\mathfrak{j}]\circ\big(z \longmapsto y(z)\big) = \varphi(z)^*[\mathfrak{k}]$ for smooth maps $V_k \to \mathcal{G}$, $z \longmapsto \varphi(z)$, and $V_j \to \mathcal{G}$, $y \longmapsto \hat{\varphi}(y)$, with $\varphi(0) = \varphi$ and $\hat{\varphi}(0) = \varphi^{-1}$. By symmetry it will be sufficient to verify holomorphicity in the first situation: Taking the derivative w.r.t. $y \in V_j$ in $k\big(z(y)\big) = \hat{\varphi}(y)^* j(y)$ we obtain

$$T_{z(y)}\mathfrak{k}\circ T_y z(\dot{y}) = 2k\big(z(y)\big)\cdot\bar{\partial}_{k(z(y))}\Big(\hat{\varphi}(y)^*\big(T_y\hat{\varphi}(\dot{y})\big)\Big) + \hat{\varphi}(y)^*\big(T_y\mathfrak{j}(\dot{y})\big)$$

as in Remark 6.7.3. Replacing $\dot{y}$ by $j\dot{y}$ and composing with $-k\big(z(y)\big)$ from the left yields

$$-T_{z(y)}\mathfrak{k}\circ k\circ T_y z(j\dot{y}) = 2\cdot\bar{\partial}_{k(z(y))}\Big(\hat{\varphi}(y)^*\big(T_y\hat{\varphi}(j\dot{y})\big)\Big) + \hat{\varphi}(y)^*\big(T_y\mathfrak{j}(\dot{y})\big)\,.$$

The second summand on the right stays the same because $k\big(z(y)\big) = \hat{\varphi}(y)^* j(y)$ and $\mathfrak{j}(y)\circ T_y\mathfrak{j} = T_y\mathfrak{j}\circ j$ by complexitiy, see Remark 6.9.1. Similarly, to deal with the left hand side use $\mathfrak{k}(z)\circ T_z\mathfrak{k} = T_z\mathfrak{k}\circ k$. On the right hand side, the first summend is an element in

$$\mathrm{Im}\Big(k\big(z(y)\big)\cdot\bar{\partial}_{k(z(y))}\Big)$$

because $j\dot{y}$ vanishes along the boundary ∂S, so that the vector field $\hat{\varphi}(y)^*\big(T_y\hat{\varphi}(j\dot{y})\big)$ vanishes along ∂S as well, and because on boundary vanishing vector fields on $(S, \partial S)$ the Cauchy–Riemann operator is complex linear, cf. Section 6.4. Effectivity of $\mathfrak{k}$ yields the algebraic splitting

$$\Omega^{0,1}_{k(z(y))} = T_{z(y)}\mathfrak{k}\,(E_k)\,\oplus\,\mathrm{Im}\Big(k\big(z(y)\big)\cdot\bar{\partial}_{k(z(y))}\Big)\,.$$

Modding out the contributions to the second summand the above two equations compare to

$$T_{z(y)}\mathfrak{k}\circ T_y z(\dot{y}) = -T_{z(y)}\mathfrak{k}\circ k\circ T_y z(j\dot{y})\,.$$

As $T_{z(y)}\mathfrak{k}$ is injective this yields

$$T_y z(\dot{y}) = -k\circ T_y z(j\dot{y})\,,$$

i.e. $k \circ T_y z = T_y z \circ j$ meaning that $y \longmapsto z(y)$ is holomorphic.

Consequently, we obtain a complex version of Proposition 6.8.2, so that, in particular, $\mathcal{R}_\tau$ is orientable.

Proposition 6.9.4. *The nodal Riemann moduli space $\mathcal{R}_\tau$ admits the structure of a complex orbifold of complex dimension $\#A - \#D$.*

Remark 6.9.5. In order to derive Proposition 6.9.4 we fixed in Section 6.3 the combinatorial data $(S, D, \{m_0, m_1, m_2\}, A)$ to represent stable nodal marked discs $[S, j, D, \{m_0, m_1, m_2\}, A]$ and used deformations of j. A direct way to obtain a complex orbifold structure would be to change the roles. Apply uniformisation as in Remark 6.3.1 in order to represent the classes $[S, j, D, \{m_0, m_1, m_2\}, A]$ by nodal discs whose disc component equals $(\mathbb{D}, \mathrm{i}, \{1, \mathrm{i}, -1\})$ and whose sphere components are given by $(\mathbb{C}P^1, \mathrm{i})$. The complex orbifold structure can be read off from variations of the configurations of the nodal points D and the marked points A as such, resp.

6.10 Varying the stable nodal type via desingularisation

For a complex number a of modulus $|a| \leq 1$ we consider the intersection of the planar algebraic curve $\{zw = a\} \subset \mathbb{C} \times \mathbb{C}$ with the polydisc $\mathbb{D} \times \mathbb{D}$. For $a = 0$ this curve is the union of the discs $\{z = 0\} = \{0\} \times \mathbb{D}$ and $\{w = 0\} = \mathbb{D} \times \{0\}$ that intersect in the singularity of the curve. For $a \neq 0$ the equation $zw = a$ can be solved by $w = a/z$ so that we obtain a cylinder that has no singularities: The restriction of the projection $(z, w) \longmapsto z$ to the curve $\{zw = a\}$ yields a biholomorphism onto the annulus $\{|a| \leq |z| \leq 1\}$ in the first coordinate plane. Interchanging z and w yields a biholomorphism onto $\{|a| \leq |w| \leq 1\}$. Both biholomorphisms constitute holomorphic charts of $\{zw = a\}$. The transition map from the first annulus to the second is

$$z \longmapsto \frac{a}{z}\,.$$

Taking **positive** and **negative holomorphic polar coordinates** $(z,w) \longmapsto -\ln z$ and $(z,w) \longmapsto \ln w$, resp., i.e. writing

$$z = \mathrm{e}^{-(s+\mathrm{i}t)} \qquad \text{and} \qquad w = \mathrm{e}^{u+\mathrm{i}v}\,,$$

the transition map gets

$$\big[0,-\ln|a|\big]\times S^1 \longrightarrow \big[\ln|a|,0\big]\times S^1\,, \qquad (s,t)\longmapsto \big(s+\ln|a|, t+\arg a\big)\,,$$

where $S^1 = \partial\mathbb{D}$. For the complex logarithm we use the main branch.

Observe that rotations $z \longmapsto \mathrm{e}^{-\mathrm{i}\theta_+}z$ and $w \longmapsto \mathrm{e}^{\mathrm{i}\theta_-}w$ for $\theta_+,\theta_- \in S^1$ of the coordinate planes, which correspond to

$$(s,t)\longmapsto \big(s,t+\theta_+\big) \qquad \text{and} \qquad (s,t)\longmapsto \big(s,t+\theta_-\big)$$

w.r.t. positive and negative holomorphic polar coordinates, resp., result into a change of the defining equation to

$$zw = \mathrm{e}^{-\mathrm{i}(\theta_+-\theta_-)}a$$

as the pull back along $(z,w)\longmapsto \big(\mathrm{e}^{-\mathrm{i}\theta_+}z, \mathrm{e}^{\mathrm{i}\theta_-}w\big)$ yields $zw=a$. The coresponding transition map is

$$(s,t)\longmapsto \Big(s+\ln|a|, t+\arg a-\big(\theta_+-\theta_-\big)\Big)\,.$$

A switch of the coordinates $(z,w)\longmapsto(w,z)$ does not effect the proceeding consideration.

Given $\big[S,j,D,\{1,\mathrm{i},-1\},A\big]\in\mathcal{R}_\tau$ we describe a similar desingularisation about a nodal pair $\{z_0,w_0\}\in D$ in terms of **parametrised connected sum**. Choose a small disc structure $\mathbf{D}_j$ on $\big(S,j,D,\{1,\mathrm{i},-1\},A\big)$. Denote the corresponding discs about the nodal points $z_0,w_0\in|D|$ by D_{z_0} and D_{w_0}, resp., and choose boundary points $z_\partial\in\partial D_{z_0}$ and $w_\partial\in\partial D_{w_0}$. We call the pair $\{z_\partial,w_\partial\}$ a **decoration** of the nodal pair $\{z_0,w_0\}$. By [54, Theorem C.5.1] there exists unique biholomorphic identifications of $\big((D_{z_0},z_0,z_\partial),j\big)$ and $\big((D_{w_0},w_0,w_\partial),j\big)$, resp., with $\big((\mathbb{D},0,1),\mathrm{i}\big)$.

For given **gluing parameter** $a\in\mathbb{D}$, $a\neq 0$, replace $-\ln|a|$ by the **modulus**

$$R = \mathrm{e}^{1/|a|}-\mathrm{e}$$

in the discussion about the planar algebraic curve $\{zw = a\}$. Identify the first annulus $\{\mathrm{e}^{-R} \le |z| \le 1\}$ with the second $\{\mathrm{e}^{-R} \le |w| \le 1\}$ via the transition map

$$z \longmapsto \frac{\mathrm{e}^{-R+\mathrm{i}\arg(a)}}{z}\,,$$

which w.r.t. positive and negative holomorphic polar coordinates reads as

$$[0,R]\times S^1 \longrightarrow [-R,0]\times S^1\,, \qquad (s,t) \longmapsto \big(s-R, t+\arg a\big)\,.$$

We obtain a surface S_a from $S \setminus \big(\operatorname{Int}(D_{z_0}) \cup \operatorname{Int}(D_{w_0})\big)$ by gluing the finite cylinder $Z_a := [0,R]\times S^1$, which is identified with $[-R,0]\times S^1$ via the above transition map, along the respective boundary circles via the restrictions of the biholomorphic identifications of D_{z_0} and D_{w_0}, resp, with $\mathbb{D}$.

The construction of the surface S_a defines a complex structure j_a that coincides with j on $S \setminus \big(\operatorname{Int}(D_{z_0}) \cup \operatorname{Int}(D_{w_0})\big)$ and with i on the cylinder Z_a of modulus R. This results into an element $\big[S_a, j_a, D_a, \{1,\mathrm{i},-1\}, A\big]$ of $\mathcal{R}_{\tau'}$ with stable nodal type τ', which necessarily differs from τ. The respective special points are given by

$$D_a := D \setminus \big\{\{z_0, w_0\}\big\}\,,$$

$\{1,\mathrm{i},-1\}$, and A under the inclusion of $S \setminus \big(\operatorname{Int}(D_{z_0}) \cup \operatorname{Int}(D_{w_0})\big)$ into S_a.

A change of biholomorphic identifications of D_{z_0} and D_{w_0} with $\mathbb{D}$ is given by a rotation of the boundary points z_∂ and w_∂, resp., which in coordinates reads as $z \longmapsto \mathrm{e}^{-\mathrm{i}\theta_+}z$ and $w \longmapsto \mathrm{e}^{\mathrm{i}\theta_-}w$, say. Gluing with the rotated identifications yields a biholomorphic copy $\big(S_b, j_b, D_b, \{1,\mathrm{i},-1\}, A\big)$ of $\big(S_a, j_a, D_a, \{1,\mathrm{i},-1\}, A\big)$, where

$$b = \mathrm{e}^{-\mathrm{i}(\theta_+ - \theta_-)}a\,.$$

To obtain a biholomorphic map take the identity map on $S \setminus \big(\operatorname{Int}(D_{z_0}) \cup \operatorname{Int}(D_{w_0})\big)$ and the rotated transition map $Z_a \to Z_b$ given by

$$(s,t) \longmapsto \Big(s-R, t+\arg a - \big(\theta_+ - \theta_-\big)\Big)$$

on Z_a.

Such rotations naturally appear when a automorphism $\psi \in \mathcal{G}_j$ for which the nodal points z_0 and w_0 are fixed-points, i.e. $\psi(z_0) = z_0$ and $\psi(w_0) = w_0$, acts on S. Indeed, ψ preserves the complement of $D_{z_0} \cup D_{w_0}$ in S and induces rotations on $D_{z_0} \cup D_{w_0}$. The rotations are measured by the change of decorations from $\{z_\partial, w_\partial\}$ to $\psi\big(\{z_\partial, w_\partial\}\big)$ in terms of angles $-\theta_+$ and θ_-, say. Therefore, we get a holomorphic diffeomorphism

$$\big(S_a, j_a, D_a, \{1, \mathrm{i}, -1\}, A\big) \longrightarrow \big(S_b, j_b, D_b, \{1, \mathrm{i}, -1\}, A\big)$$

as above which this time coincides with ψ on $S \setminus \big(\operatorname{Int}(D_{z_0}) \cup \operatorname{Int}(D_{w_0})\big)$.

We denote by

$$\mathbb{D}^D$$

the set of all maps from the set of nodal points D to the set $\mathbb{D}$ of complex numbers of modulus less than or equal to 1. Choose a small disc structure $\mathbf{D}_j$ on $\big(S, j, D, \{1, \mathrm{i}, -1\}, A\big)$ together with a decoration for each disc in $\mathbf{D}_j$. The choice of decorations determine holomorphic diffeomorphisms of all discs of the disc structure with $\mathbb{D}$ such that the nodal point is mapped to $0 \in \mathbb{D}$ and the decoration to $1 \in \mathbb{D}$. Given $\mathbf{a} \in \mathbb{D}^D$ we perform the described parametrised connected sum about each nodal pair $\{z, w\} \in D$ with gluing parameter

$$a_{\{z,w\}} := \mathbf{a}\big(\{z, w\}\big)\,.$$

This is done by replacing the node $\{z, w\}$ with the cylinder $Z^{\{z,w\}}_{a_{\{z,w\}}}$. In the case of a vanishing gluing parameter $a_{\{z,w\}}$ formally

$$Z^{\{z,w\}}_0 := D_z \sqcup D_w$$

is given by the disjoint union of half-infinite cylinders $[0, \infty) \times S^1$ and $(-\infty, 0] \times S^1$ of **infinite modulus** after removing the nodal points z and w. In other words, if $a_{\{z,w\}} = 0$ we do nothing and keep the nodal pair $\{z, w\} \in D_{\mathbf{a}}$, so that $D_{\mathbf{a}}$ arises from D by removing all nodal pairs $\{z, w\}$ with $a_{\{z,w\}} \neq 0$. The resulting surface is denoted by

$$\big(S_{\mathbf{a}}, j_{\mathbf{a}}, D_{\mathbf{a}}, \{1, \mathrm{i}, -1\}, A\big)\,.$$

Starting off with a skyscraper deformation $V_j \ni y \longmapsto j(y)$ and a small disc structure of sufficiently small discs we will get

$$\big(S_{\mathbf{a}}, j(y)_{\mathbf{a}}, D_{\mathbf{a}}, \{1, \mathrm{i}, -1\}, A\big)$$

by the same construction.

In order to describe the effect of the $\mathcal{G}_j$-action of $\big(S, j, D, \{1, \mathrm{i}, -1\}, A\big)$ on the desingularisation we denote by $\kappa_{z,w}$, $\{z, w\} \in D$, the complex anti-linear map $T_w S \to T_z S$ that conjugated with the linearisations of the biholomorphic identifications of the discs D_w and D_z with $\mathbb{D}$ is equal to the complex conjugation map $x + \mathrm{i}y \longmapsto x - \mathrm{i}y$ on $\mathbb{C}$. Interchanging the role of z and w replaces $\kappa_{z,w}$ by its inverse $\kappa_{w,z} = (\kappa_{z,w})^{-1}$. We call $\kappa_{z,w}$ a **compatible nodal identifier**. Given $\psi \in \mathcal{G}_j$ and $\{z, w\} \in D$ we define the **phase function**

$$\Theta_{\{z,w\}}(\psi) \colon\; T_z S \longrightarrow T_z S$$

by

$$\Theta_{\{z,w\}}(\psi) := \kappa_{z,w} \circ T_{\psi(w)}(\psi)^{-1} \circ \kappa_{\psi(w),\psi(z)} \circ T_z \psi \,.$$

Taking positive and negative holomorphic polar coordinates about z and w, resp., so that ψ acts in coordinates by multiplication with $\mathrm{e}^{-\mathrm{i}\theta_+}$ and $\mathrm{e}^{\mathrm{i}\theta_-}$, resp., we get

$$\big(\Theta_{\{z,w\}}(\psi)\big)(v) = \mathrm{e}^{-\mathrm{i}(\theta_+ - \theta_-)} v$$

for all $v \in T_z S$, which we simply declare to a multiplication operator

$$\Theta_{\{z,w\}}(\psi) \equiv \mathrm{e}^{-\mathrm{i}(\theta_+ - \theta_-)} \,.$$

This shows independence of the phase function

$$\Theta \colon\; D \times \mathcal{G}_j \longrightarrow S^1 \,, \qquad \big(\{z, w\}, \psi\big) \longmapsto \Theta_{\{z,w\}}(\psi) \,,$$

of the chosen ordering of $\{z, w\}$ in the definition of $\Theta_{\{z,w\}}(\psi)$ and of the chosen parity of the holomorphic polar coordinates about $\psi(z)$ and $\psi(w)$. This results in a $\mathcal{G}_j$-action on $\mathbb{D}^D$ defined by $\psi_* \mathbf{a} = \mathbf{b}$ via

$$b_{\{\psi(z),\psi(w)\}} := \Theta_{\{z,w\}}(\psi) \cdot a_{\{z,w\}}$$

for all $\{z,w\} \in D$. Consequently, for any skyscraper deformation $V_j \ni y \longmapsto j(y)$, a small disc structure of sufficiently small discs, and $\psi \in \mathcal{G}_j$ we get an isomorphism

$$\psi_{\mathbf{a}} : \left(S_{\mathbf{a}}, j(\psi^* y)_{\mathbf{a}}, D_{\mathbf{a}}, \{1, \mathrm{i}, -1\}, A\right) \longrightarrow \left(S_{\psi_* \mathbf{a}}, j(y)_{\psi_* \mathbf{a}}, D_{\psi_* \mathbf{a}}, \{1, \mathrm{i}, -1\}, A\right)$$

by the gluing construction and symmetry of $y \longmapsto j(y)$.

6.11 Topology and orbifold structure – variable stable nodal type

A neighbourhood base of a second countable paracompact Hausdorff topology on $\mathcal{R}_N$, $N \geq 0$, is given by the family of subsets of $\mathcal{R}_N$, whose elements are of the form

$$\left[S_{\mathbf{a}}, k_{\mathbf{a}}, D_{\mathbf{a}}, \{1, \mathrm{i}, -1\}, A\right]$$

with $N = \#A$, which are obtained from a nodal disc $\left(S, k, D, \{1, \mathrm{i}, -1\}, A\right)$ by the parametrised connected sum construction with given decorated small disc structure $\mathbf{D}_j$, with gluing parameter $\mathbf{a} \in \mathbb{D}^D$ with $|\mathbf{a}| < \varepsilon$ for some $\varepsilon \in (0,1)$, with complex structures k that belong to an open neighbourhood of j in $\mathcal{J}$ such that $k = j$ restricted to $\mathbf{D}_j$. This follows as in [43, Proposition 2.4] and [42, Theorem 2.15 and Theorem 5.13] because no extra argument for boundary un-noded nodal discs is needed caused by absence of boundary nodes. The Hausdorff property follows with Gromov compactness for stable holomorphic discs, see [25].

The induced topology on $\mathcal{R}_\tau$ in $\mathcal{R}_N$ agrees with the one on $\mathcal{R}_\tau$ previously defined in Remark 6.3.2. The induced notion of convergence of sequences in $\mathcal{R}_N$ coincides with Gromov convergence as described in [1, Chapter 1], [66, Appendix B] or in [15, Section 4], [46, Chapter IV] after Schwartz reflection along the boundary of the nodal discs for example.

In order to obtain an orbifold structure on $\mathcal{R}_N$ we consider desingularisations

$$\left(S_{\mathbf{a}}, j(y)_{\mathbf{a}}, D_{\mathbf{a}}, \{1, \mathrm{i}, -1\}, A\right)$$

of $\left(S, j, D, \{1, \mathrm{i}, -1\}, A\right)$ as described in Section 6.10. For $\mathbf{a}_0 \in \mathbb{D}^D$ consider the set $D \setminus D_{\mathbf{a}_0}$ of all nodal pairs $\{z,w\} \in D$ on which the map $\mathbf{a}_0$ is non-zero. Define a deformation

$$\mathfrak{j}_{\mathbf{a}_0} \colon\; V_{j_{\mathbf{a}_0}} \times \mathbb{D}^{D\setminus D_{\mathbf{a}_0}} \longrightarrow \mathcal{J}_{S_{\mathbf{a}_0}}\,, \qquad (y,\mathbf{b}) \longmapsto j(y)_{\mathbf{a}_0+\mathbf{b}}\,,$$

of

$$\big(S_{\mathbf{a}_0}, j(y)_{\mathbf{a}_0}, D_{\mathbf{a}_0}, \{1,\mathrm{i},-1\}, A\big)$$

by setting $\mathbf{a} = \mathbf{a}_0 + \mathbf{b}$. For small deformation parameter $\mathbf{b}$ the deformed family of surfaces equals

$$\Big(S_{\mathbf{a}_0+\mathbf{b}}, j(y)_{\mathbf{a}_0+\mathbf{b}}, D_{\mathbf{a}_0+\mathbf{b}}, \{1,\mathrm{i},-1\}, A\Big)\,.$$

The nodal discs family is isomorphic to

$$\Big(S_{\mathbf{a}_0}, j'(y)_{\mathbf{b}}, D_{\mathbf{a}_0}, \{1,\mathrm{i},-1\}, A\Big)$$

with corresponding deformation

$$\mathfrak{j}'_{\mathbf{b}} \colon\; V_{j_{\mathbf{a}_0}} \times \mathbb{D}^{D\setminus D_{\mathbf{a}_0}} \longrightarrow \mathcal{J}_{S_{\mathbf{a}_0}}\,, \qquad (y,\mathbf{b}) \longmapsto j'(y)_{\mathbf{b}}\,,$$

via an isomorphism that is the identification map on the complement of the respective small disc structure, so that the deformation is given by rotations and stretchings of the cylindrical neck regions that correspond to the nodes, on which $\mathbf{a}_0$ not vanishes. The **partial Kodaira differential** of $\mathfrak{j}'_{\mathbf{b}}$ at $(y,0)$ is

$$\big[T_{(y,0)}\mathfrak{j}'_{\mathbf{b}}\big] \colon\; E_{j_{\mathbf{a}_0}} \times \mathbb{C}^{D\setminus D_{\mathbf{a}_0}} \longrightarrow \Omega^{0,1}_{j(y)_{\mathbf{a}_0}} \longrightarrow H^1_{j(y)_{\mathbf{a}_0}}\,.$$

Similarly to [43, Theorem 2.13] one constructs uniformisers of an orbifold structure on $\mathcal{R}_N$ as the above desingularisations stay away from the boundary of the nodal discs in $\mathcal{R}_N$. For given $\big[S,j,D,\{1,\mathrm{i},-1\},A\big] \in \mathcal{R}_N$ such a **uniformiser** is a deformation

$$\mathcal{V} \ni (y,\mathbf{a}) \longmapsto \big(S_{\mathbf{a}}, j(y)_{\mathbf{a}}, D_{\mathbf{a}}, \{1,\mathrm{i},-1\}, A\big)$$

of $\big(S,j,D,\{1,\mathrm{i},-1\},A\big)$ for an open subset $\mathcal{V}$ of $V_j \times \mathbb{D}^D$ such that the following holds:

- The union of all equivalence classes $\big[S_{\mathbf{a}}, j(y)_{\mathbf{a}}, D_{\mathbf{a}}, \{1,\mathrm{i},-1\},A\big]$ over all $(y,\mathbf{a}) \in \mathcal{V}$ is an open subset of $\mathcal{R}_N$.

- The map $\mathcal{V} \to \mathcal{U}$ that assigns to $(y, \mathbf{a})$ the class $\big[S_{\mathbf{a}}, j(y)_{\mathbf{a}}, D_{\mathbf{a}}, \{1, \mathrm{i}, -1\}, A\big]$ descends to a homeomorphism $\mathcal{V}/\mathcal{G}_j \to \mathcal{U}$.

- An isomorphism between the classes belonging to $(y, \mathbf{a}), (z, \mathbf{b}) \in \mathcal{V}$ is given by $\psi_{\mathbf{a}}$ for $\psi \in \mathcal{G}_j$ and $(z, \mathbf{b}) = \big(\psi_* y, \psi_* \mathbf{a}\big)$.

- For all points in $\mathcal{V}$ the partial Kodaira differential is an isomorphism.

Compatibility of uniformisers is expressed via the sets

$$\mathbf{T}_{j,k} := \Big\{\big(\varphi, (y, \mathbf{a}), (z, \mathbf{b})\big)\Big\} \subset \mathcal{G} \times (V_j \times \mathbb{D}^D) \times (V_k \times \mathbb{D}^D)$$

corresponding to all isomorphisms

$$\varphi : \Big(S_{\mathbf{a}}, j(y)_{\mathbf{a}}, D_{\mathbf{a}}, \{1, \mathrm{i}, -1\}, A\Big) \longrightarrow \Big(S_{\mathbf{b}}, k(z)_{\mathbf{b}}, D_{\mathbf{b}}, \{1, \mathrm{i}, -1\}, A\Big),$$

which are smooth manifolds of dimension $2\#A$, so that $\mathcal{R}_N$ supports an étale proper Lie groupoid structure as formulated after Proposition 6.8.2. This follows with the (anti-)gluing construction ([43, Section 2.4]) for the non-linear Cauchy–Riemann operator along the nodes (which take place away from the boundary) known from Floer theory, cf. [43, Theorem 2.16] and [42, Theorem 2.24]. Similarly, the universal property of the construction stated in [43, Theorem 2.16] translates into the present situation. The involved variation of marked points can be treaded as in [42, Remark 3.17]. Finally, using convex interpolation between the exponential gluing profile $\mathrm{e}^{1/r} - \mathrm{e}$ we used in the gluing construction and the logarithmic gluing profile $-\ln r$ that appeared in the desingularisation of the complex algebraic curve at the beginning of Section 6.10 naturally yields an orientation on $\mathcal{R}_N$ that extends the complex orientation on $\mathcal{R}_\tau$ given in Proposition 6.9.4, see [42, Section 2.3.2].

Theorem 6.11.1. *The nodal Riemann moduli space $\mathcal{R}_N$ of stable nodal boundary un-noded discs with $N = \#A$ interior marked points admits a naturally oriented orbifold structure of dimension $2\#A$.*

7 Polyfold perturbations

We prove Theorem 5.1.2 under assumption (ii). For that we place ourselves into the situation of Section 5.3 and follow the line of reasoning of the proof of Theorem 5.1.2 part (i). As we will not assume semi-positivity this time regularity of relevant moduli spaces can only be achieved for simple nodal holomorphic discs via perturbing the almost complex structure, cf. Section 5.3. For non-simple nodal holomorphic discs we will use additional abstract polyfold perturbations as introduced in [43].

7.1 Boundary un-noded stable disc maps

We consider the tame almost complex manifold $(\hat{W}, \hat{\Omega}, \hat{J})$ defined in Section 5.2. For boundary un-noded nodal discs (S, j, D) as introduced in Section 6.1 we consider smooth maps

$$u\colon\ (S, \partial S) \longrightarrow (\hat{W}, N^*)$$

that descend to continuous maps on S/D. If D is empty we call u **un-noded**. If in addition $Tu \circ j = \hat{J}(u) \circ Tu$ we call u a **nodal holomorphic disc map**. Observe that we do not need to consider nodal points on the boundary due to the Gromov compactification described in Remark 4.4.1.

More generally, we consider continuous maps $u\colon\ (S, \partial S) \to (\hat{W}, N^*)$ defined on a marked boundary un-noded nodal disc $\big(S, j, D, \{m_0, m_1, m_2\}\big)$ (see Section 6.1) such that u descends to a continuous map on the quotient S/D and such that $u\big(\operatorname{Int} S\big) \subset \operatorname{Int} \hat{W}$. Moreover, we require that u is contained in the Sobolev space of square integrable maps

$$H^{3,\sigma}(S, j) \equiv H^{3,\sigma}\big(S, j, D, \{m_0, m_1, m_2\}\big)$$

following [43, Definition 1.1]: We require that u is of class $u \in H^3_{\mathrm{loc}}\big(S \setminus |D|\big)$ and that w.r.t. positive holomorphic polar coordinates $[0, \infty) \times S^1$, $S^1 = \mathbb{R}/2\pi\mathbb{Z}$, about the nodal points $|D|$ (see Section 6.10) the map u is of weighted Sobolev class $H^{3,\sigma}$. The weights are given by $\mathrm{e}^{\sigma s}$, $s \in [0, \infty)$, for some $\sigma \in (0, 1)$. In other words, u is contained in $H^{3,\sigma}$ precisely if all weak derivatives $D^\alpha u$, $|\alpha| \leq 3$, of u on $[0,\infty)\times S^1$ exist and all $D^\alpha u \cdot \mathrm{e}^{\sigma s}$, $|\alpha| \leq 3$, are square integrable on $[0,\infty)\times S^1$. The latter is equivalent to $u\,\mathrm{e}^{\sigma s} \in H^3$ on $[0,\infty)\times S^1$. In particular, by Sobolev embedding, u is C^1 (up to the boundary ∂S) restricted to $S \setminus |D|$. But in general u is not differentiable at the nodal points $|D|$ on S. Consider for example the continuous function $u(z) = |z|^{\frac{1+\sigma}{2}}$ in holomorphic coordinates $z \in \mathbb{C}$, which w.r.t. positive holomorphic polar coordinates reads as $(s,t) \longmapsto \mathrm{e}^{-\frac{1+\sigma}{2}s}$.

The space $H^{3,\sigma}(S,j)$ is well defined, i.e. invariant under coordinate changes after possibly shrinking the chart domains. Away from the nodes $|D|$ this follows as for $H^3_{\mathrm{loc}}\big(S \setminus |D|\big)$ via [4, Theorem 3.41]. Near the nodes we observe that the area form $\mathrm{e}^{2\sigma s}\mathrm{d}t \wedge \mathrm{d}s$ w.r.t. positive holomorphic polar coordinates corresponds to the singular area form $|z|^{-2(1+\sigma)}\frac{\mathrm{i}}{2}\mathrm{d}z \wedge \mathrm{d}\bar{z}$ in holomorphic coordinates about the nodal point $0 \in \mathbb{C}$. The area form $\frac{\mathrm{i}}{2}\mathrm{d}z \wedge \mathrm{d}\bar{z}$ transforms under biholomorphic coordinate changes via a conformal factor, which we can assume to be bounded above and away from zero by shrinking the chart domains if necessary. The coordinate change itself is of the form $z \longmapsto zh(z)$, where 0 corresponds to a nodal point. Here h is a holomorphic function, whose absolute value can be assumed to be bounded above and away from zero also. Consequently, the singular area form $|z|^{-2(1+\sigma)}\frac{\mathrm{i}}{2}\mathrm{d}z \wedge \mathrm{d}\bar{z}$ transforms via a bounded above and away from zero conformal factor also. Hence, the same holds true for $\mathrm{e}^{2\sigma s}\mathrm{d}t \wedge \mathrm{d}s$. In fact, the above coordinate change becomes

$$(s,t) \longmapsto (s,t) - \ln\Big(h\big(\mathrm{e}^{-(s+\mathrm{i}t)}\big)\Big),$$

whose derivatives are bounded above and whose first derivative is bounded away from zero. Therefore, invariance under coordinate changes near the nodes follows as in [4, Theorem 3.41]. By the same arguments we see that locally defined norms on H^3_{loc} and $H^{3,\sigma}$ transform via the respective coordinate changes to equivalent norms. This defines a topology on $H^{3,\sigma}(S,j)$; a neighbourhood base is given by the set of those maps that restricted to one of the above charts belong to an open set in H^3_{loc} and $H^{3,\sigma}$, resp.

To each $u \in H^{3,\sigma}(S,j)$ we assign the **symplectic energy integral**

$$\int_S u^*\hat{\Omega}$$

by approximating the continuous map u by a C^1-map v and defining the symplectic energy integral via $\int_S u^*\hat{\Omega} := \int_S v^*\hat{\Omega}$. This is well defined and, in fact, by Stokes theorem, independent of the choice of representative of the homology class $[u]$ in $\hat{W}$ relative $u(\partial S) \subset N^*$. This can be seen as follows: Taking approximations v of u that are equal to u restricted to the complement of disc like neighbourhoods $B_r(0)$ in S of the nodal points $0 \in |D|$ the symplectic energy integral is given by $\int_{S\setminus|D|} u^*\hat{\Omega}$. Indeed, take $r > 0$ so small such that the $B_r(0)$ are contained in pair-wise disjoint chart domains of S about the nodal points 0 in $|D|$ and such that the $u\big(B_r(0)\big)$ are contained in pair-wise disjoint ball like chart domains of $\hat{W}$. By Stokes theorem decomposing $B_r(0) = \big(B_r(0) \setminus B_\varepsilon(0)\big) \cup B_\varepsilon(0)$ it suffices to show that the integrals

$$\int_{B_\varepsilon(0)} u^*\hat{\Omega} \qquad \text{and} \qquad \int_{\partial B_\varepsilon(0)} u^*\lambda$$

converge to zero as $\varepsilon \in (0,r)$ tends to 0, where λ is a local primitive of $\hat{\Omega}$ defined on the ball like neighbourhoods of $u(|D|)$ in $\hat{W}$. By the transformation formula we can compute the integrals w.r.t. positive holomorphic polar coordinates via

$$\int_{(R,\infty)\times S^1} \hat{\Omega}(u_s,u_t)\,\mathrm{d}s\wedge\mathrm{d}t \qquad \text{and} \qquad \int_{\{R\}\times S^1} \lambda(u_t)\,\mathrm{d}t$$

for $R = -\ln\varepsilon$. By the Sobolev inequality the C^1-norm of $u\,\mathrm{e}^{\sigma s}$ on $[0,\infty)\times S^1$ is bounded by $\|u\|_{3,\sigma}$, so that up to a positive constant the absolute value of the integrals is bounded by

$$\|u\|^2_{3,\sigma}\int_R^\infty \mathrm{e}^{-2\sigma s}\mathrm{d}s \qquad \text{and} \qquad \|u\|_{3,\sigma}\,\mathrm{e}^{-\sigma R}\,,$$

resp. In both cases the first factor is bounded by assumption; the second tends to zero for $R \to \infty$ and the claim follows, namely, that $\int_S u^*\hat{\Omega}$ is well defined.

Remark 7.1.1. The above arguments show that $u \longmapsto \int_S u^*\hat{\Omega}$ is a continuous function on $H^{3,\sigma}(S,j)$.

We call $(S, j, D, \{m_0, m_1, m_2\}, u)$ a **nodal disc map** provided that the following conditions are satisfied (cf. Section 5.3):

1. $u \in H^{3,\sigma}(S, j)$,
2. the **symplectic energy integral** restricted to a connected component C of S
$$\int_C u^*\hat{\Omega} \geq 0$$
is non-negative for all spherical components C of S; positive on the disc component,
3. the continuous map on S/D induced by u is homologous to a local Bishop discs u_{ε,b_o} relative N^*, so that $[u(S)] = [u_{\varepsilon,b_o}(\mathbb{D})]$ in $H_2(\hat{W}, N^*)$, and
4. $u(m_0) \in \gamma$ and $\vartheta \circ u(m_k) = \mathrm{i}^k$ for $k = 1, 2$.

For given $(S, j, D, \{m_0, m_1, m_2\})$ the space
$$\mathcal{H}^{3,\sigma}(S, j)$$
of nodal disc maps is called the **space of admissible maps**.

It follows that the degree of the C^1-map $\vartheta \circ u\colon\ \partial S \to S^1$ equals 1 for all nodal disc maps $(S, j, D, \{m_0, m_1, m_2\}, u)$. With the properties of the symplectic energy integral discussed above we obtain as in item (2) of Section 4.4 that
$$\int_S u^*\hat{\Omega} = \int_{\partial S} u^* f \cdot (\vartheta \circ u)^* \mathrm{d}\theta\,,$$
where f is a smooth function on N that is positive on N^* and vanishes on $B \cup \partial N$. As u takes values in N^* along the boundary ∂S we get that
$$\int_S u^*\hat{\Omega} \in \big(0, 2\pi \max f\big]$$
for all nodal disc maps $(S, j, D, \{m_0, m_1, m_2\}, u)$. By non-negativity of the symplectic energy integral on each connected component C of S we get that $\int_C u^*\hat{\Omega}$ takes values in $\big[0, 2\pi \max f\big]$. Moreover, as $\int_C u^*\hat{\Omega}$ only depends on the homology class represented by

$u(C)$ for the spherical components C of S assumption (2) puts an open condition to the space defined via $H^{3,\sigma}(S,j)$ and the constraints given by (3) and (4), so that the space of admissible maps $\mathcal{H}^{3,\sigma}(S,j)$ is an open subset.

In fact, $\mathcal{H}^{3,\sigma}(S,j)$ is a Hilbert manifold whose tangent space $\mathcal{H}^{3,\sigma}(u^*T\hat{W})$ at $u \in \mathcal{H}^{3,\sigma}(S,j)$ is the space of $H^{3,\sigma}$-sections into $u^*T\hat{W}$ that descent to continuous sections on S/D, that are tangent to N^* along ∂S as well as tangent to γ at m_0 and to the page $\vartheta^{-1}(\mathrm{i}^k)$ at m_k for $k = 1, 2$. This follows with the exponential map taken w.r.t. a metric on $\hat{W}$ for which each of the submanifolds N, $\vartheta^{-1}(\mathrm{i}^k)$, $k = 1, 2$, and γ is totally geodesic. The requirement for the sections to be of class $H^{3,\sigma}$ is understood as in Section 7.1, so that a norm on $\mathcal{H}^{3,\sigma}(u^*T\hat{W})$ as on [43, p. 66] can be defined. This turns $\mathcal{H}^{3,\sigma}(S,j)$ into a Riemannian Hilbert manifold.

By removal of singularities (see [54]) a holomorphic $u \in H^{3,\sigma}(S,j)$, which is continuous and has finite symplectic energy by the above discussion, is holomorphic on S. Therefore, u is smooth up to the boundary including all nodal points $|D|$ so that holomorphicity coincides with the notion of holomorphicity from the beginning of this section.

Given a nodal disc map $(S, j, D, \{m_0, m_1, m_2\}, u)$ we call a connected component C of S with vanishing symplectic energy integral a **ghost bubble**. Observe that a holomorphic nodal disc map restricted to a ghost bubble is constant. If $(S, j, D, \{m_0, m_1, m_2\}, u)$ is any nodal disc map such that each ghost bubble admits at least 3 nodal points, then we call $(S, j, D, \{m_0, m_1, m_2\}, u)$ a **stable nodal disc map**.

7.2 Boundary un-noded stable discs

We call two stable nodal disc maps

$$(S, j, D, \{m_0, m_1, m_2\}, u) \quad \text{and} \quad (S', j', D', \{m_0', m_1', m_2'\}, u')$$

equivalent if there exists a diffeomorphism $\varphi\colon\ S \to S'$ such that $\varphi^* j' = j$, the injection $D \to D'$ defined by $\{\varphi(x), \varphi(y)\} \in D'$ for all $\{x, y\} \in D$ is surjective, $\varphi(m_k) = m_k'$ for $k = 0, 1, 2$, and $u' \circ \varphi = u$. The discussions in Section 7.1 about $H^{3,\sigma}(S,j)$ imply that this equivalence relation is well defined. The equivalence classes

$$\mathbf{u} = \big[S, j, D, \{m_0, m_1, m_2\}, u\big]$$

are called **stable nodal discs** in $(\hat{W}, \hat{\Omega})$ relative N^*. The space of all equivalence classes is denoted by $\mathcal{Z}$.

Fixing the diffeomorphism type of S and the combinatorial data $\big(D, \{1, \mathrm{i}, -1\}\big)$ of $\big(S, j, D, \{1, \mathrm{i}, -1\}, u\big)$ as at the beginning of Section 6.3 we can write

$$\mathcal{Z} = \Big\{\mathbf{u} = [j, u] \text{ stable} \,\Big|\, [u(S)] = [u_{\varepsilon, b_o}(\mathbb{D})]\,,\ u(1) \in \gamma\,,\ \vartheta \circ u(\mathrm{i}^k) = \mathrm{i}^k\,, k = 1, 2\Big\}$$

for the **space of all stable nodal discs**

$$\mathbf{u} = \big[S, j, D, \{1, \mathrm{i}, -1\}, u\big] \equiv [j, u]\,, \quad u \in \mathcal{H}^{3,\sigma}(S, j)\,,$$

in $(\hat{W}, \hat{\Omega})$ relative N^*.

We define the nodal type τ of $\big(S, D, \{1, \mathrm{i}, -1\}\big)$ as in Section 6.2. Namely, the nodal type is the isomorphism class of the rooted tree given as follows: The vertices correspond to the components of S. The root is given by the disc component. The edge relation is induced by the nodes in D. As this time there are no auxiliary marked points all vertices different from the root are not weighted; the root has weight 3. The induced nodal type τ is necessarily unstable provided that there is at least one sphere component. Indeed, in this case, any end of a branch admits only one special point.

We denote by $\mathcal{Z}_\tau$ the **space of all stable nodal discs of nodal type** τ, so that $\mathcal{Z}$ is the disjoint union of the $\mathcal{Z}_\tau$ where τ ranges over all nodal types just described. Each of the subspaces $\mathcal{Z}_\tau$ of $\mathcal{Z}$ is the quotient of the total space of the fibration over $\mathcal{J} \equiv \mathcal{J}(S)$ with fibre $\mathcal{H}^{3,\sigma}(S, j)$ over $j \in \mathcal{J}$ by the action $\varphi \longmapsto (\varphi^* j, u \circ \varphi)$ of the group of orientation preserving diffeomorphisms φ of S preserving $\big(D, \{1, \mathrm{i}, -1\}\big)$, cf. Section 6.3. This puts a topology to $\mathcal{Z}_\tau$ similarly to Remark 6.3.2.

Notice that the stabiliser of the action is finite by the stability condition formulated at the end of Section 7.1: Each automorphism of $\mathbf{u}$ acts via the identity map on the disc component due to the ordered boundary marked points $\{1, \mathrm{i}, -1\}$. If $\int_C u^* \hat{\Omega} = 0$ for a connected component C of S, then the number of nodal points $C \cap |D|$ on C is at least 3. Furthermore, the automorphisms of $\mathbf{u}$ preserve those ghost components due to the transformation formula. If $\int_C u^*\hat{\Omega} > 0$, one finds $z \in C \setminus |D|$ such that u is immersive on $C \cap u^{-1}\big(u(z)\big)$ defining finitely many local branches via $C \cap u^{-1}\big(B_r\big) \subset C \setminus |D|$ for a sufficiently small ball $B_r \subset \operatorname{Int} \hat{W}$ around $u(z)$. In fact, due to the positivity of the symplectic energy integral we can find $z \in C \setminus |D|$ and $r > 0$ sufficiently small such that $u^*\hat{\Omega}$ is a positive area form on the branch through z, which is oriented via j. Observe that $u^*\hat{\Omega}$ is a positive area form on all branches through the points of the orbit (of the automorphism group of $\mathbf{u}$) defined by z. Identifying the sphere components with $(\mathbb{C}P^1, \mathrm{i})$ as in Remark 6.3.1 the identity theorem yields that an automorphism of $\mathbf{u}$ acts by a permutation on the local branches. This proves finiteness of the stabiliser.

In the following we describe a polyfold structure on $\mathcal{Z}$ that glues the components $\mathcal{Z}_\tau$ together. For any $\mathbf{u} = \big[S, j, D, \{1, \mathrm{i}, -1\}, u\big]$ in $\mathcal{Z}$ one can choose a so-called **stabilisation**, which is a finite set of auxiliary marked points $A \subset S$ disjoint from the special points $D \cup \{1, \mathrm{i}, -1\}$ such that the nodal disc $\big(S, j, D, \{1, \mathrm{i}, -1\}, A\big)$ is stable in the sense of Section 6.2. Due to the stability condition there is no need to provide the ghost bubbles with an auxiliary markt point. In addition one can assume, that the automorphisms of $\mathbf{u}$ preserve A, $u(A)$ is disjoint from the u-image of $D \cup \{1, \mathrm{i}, -1\}$, and the following two conditions hold:

1. Whenever $z, w \in A$ are mapped to the same point $u(z) = u(w)$ in $\hat{W}$, then there exists an automorphism of $\mathbf{u}$ sending z to w.

2. For all $z \in A$ the 2-form $(u^*\hat{\Omega})_z$ is positive on (T_zS, j_z).

This follows with [43, Lemma 3.2] ignoring the disc component, which already is stable: Namely, successively select finite orbits of the action of the automorphism group of $\mathbf{u}$ on local branches similarly to the above finiteness argument until all components are stable. Consequently, the underlying stable nodal disc $\big[S, j, D, \{1, \mathrm{i}, -1\}, A\big]$ possesses a uniformiser as described in Section 6.11.

As in Section 5.3 we wish to achieve an index-1 Fredholm problem. In view of Theorem

6.11.1 we compensate the stabilising auxiliary marked points A index-wise as follows: We choose a finite collection of pairwise disjoint codimension-2 symplectic discs in $(\operatorname{Int}\hat{W},\hat{\Omega})$ that intersect $u(S)$ along $u(A)$ transversally. This is possible by condition (2) above. Namely, the image of Tu at each auxiliary marked point in A is a symplectic plane in $T\hat{W}$. Integrating the respective symplectic normal subspaces one finds symplectic embeddings of small discs of codimension 2 that are normal to $u(S)$ at the images of the auxiliary marked points $u(A)$. We call the union of the discs $H_{u,A}$ **local transversal constraints** if the intersection of $u(S)$ and $H_{u,A}$ equals $u(A)$ and if each component of $H_{u,A}$ intersects $u(S)$ in a single point.

We denote by

$$E_{u,A}\subset\mathcal{H}^{3,\sigma}\big(u^*T\hat{W}\big)$$

the subspace of sections that are tangent to $H_{u,A}$ at the stabilising auxiliary points in A, which is scale-linear w.r.t. to $(3+\nu,\sigma_\nu)$, $\nu\in\mathbb{N}_0$, for a strictly increasing sequence σ_ν in $(0,1)$ with $\sigma_0=\sigma$, see [43, Section 2.6]. Uniformiser about any $\mathbf{u}=\big[S,j,D,\{1,\mathrm{i},-1\},u\big]$ in $\mathcal{Z}$ of the desired polyfold structure are obtained as in [43, Section 3.1/3.2]. To adapt to our situation start off with uniformisers for the stabilised domain $\big[S,j,D,\{1,\mathrm{i},-1\},A\big]\in\mathcal{R}$ from Section 6.11 and consider the deformation

$$(y,\mathbf{a},\eta)\longmapsto\big(S_{\mathbf{a}},j(y)_{\mathbf{a}},D_{\mathbf{a}},\{1,\mathrm{i},-1\},\oplus_{\mathbf{a}}\exp_u(\eta)\big)\,,$$

where $(y,\mathbf{a})\in\mathcal{V}$ for an open subset $\mathcal{V}$ of $V_j\times\mathbb{D}^D$, $\eta\in E_{u,A}$ is a sufficiently small section that is a fixed point of the splicing projection $\pi_{\mathbf{a}}$ and $\oplus_{\mathbf{a}}\exp_u(\eta)$ denotes the gluing operation both introduced in [43, Section 2.4/2.5]. Choosing u to be a smooth approximation of an element in $H^{3,\sigma}(S,j)$ we obtain scale-smooth gluing maps w.r.t. to the scale $(3+\nu,\sigma_\nu)$, see [43, Section 2.2/2.6]. Using Remark 6.3.2, Section 6.11 and [43, Section 3.3/3.4] one obtains a natural second countable paracompact Hausdorff topology on $\mathcal{Z}$ similarly to [43, Theorem 1.6]. In the same way using this time modifications in [43, Section 3.5] the space $\mathcal{Z}$ carries the structure of a polyfold as formulated in [43, Theorem 1.7] with a scale-smooth evaluation map $\mathcal{Z}\to\gamma$ sending $\mathbf{u}$ to $u(1)$, cf. [43, Theorem 1.8].

7.3 A nodal moduli space

We call $\mathbf{u}$ a **stable nodal holomorphic disc** if $\mathbf{u}$ can be represented by a stable nodal holomorphic disc map u. Notice, that all stable nodal disc maps u that represent a stable nodal holomorphic disc $\mathbf{u}$ are holomorphic. Denote by

$$\mathcal{N} := \left\{\mathbf{u} \in \mathcal{Z} \,\middle|\, \mathbf{u} \text{ is holomorphic}\right\}$$

the **nodal moduli space** of all stable nodal holomorphic discs.

Using uniformisation it is convenient to represent the classes $\mathbf{u} \in \mathcal{N}$ by holomorphic maps $u \in \mathcal{H}^{3,\sigma}(S,j)$ whose disc component has domain $\big(\mathbb{D}, \mathrm{i}, \{1, \mathrm{i}, -1\}\big)$ and for which the sphere components are given by $(\mathbb{C}P^1, \mathrm{i})$, cf. Remark 6.3.1. If $\mathbf{u}$ is un-noded, then we obtain $\mathbf{u} = \big[\mathbb{D}, \mathrm{i}, \emptyset, \{1, \mathrm{i}, -1\}, u\big]$. We abbriviate the elements $\mathbf{u} = [\mathrm{i}, u] \in \mathcal{N}$ (noded or un-noded) simply by $[u]$ for the following discussion:

The boundary conditions for $\mathcal{H}^{3,\sigma}(S,j)$ formulated in Section 7.1 are the boundary conditions used in Sections 4.4 and 5.3. In particular, all properties formulated in the un-noded case for holomorphic discs in Section 4.4 continue to hold in the noded case, hence, for all $\mathbf{u} = [u] \in \mathcal{N}$ in the following sense:

1. The **winding number** of $\mathbf{u} = [u] \in \mathcal{N}$, which by definition is the degree of the map $\vartheta \circ u : \partial S \to S^1$, is equal to 1. In particular, $u(\partial S)$ is an embedded curve in N^* positively transverse to ξ and the restriction of u to the disc component of S is a simple holomorphic map.

2. The **symplectic energy** $\int_S u^*\hat{\Omega}$ of $\mathbf{u} = [u] \in \mathcal{N}$, which is well defined and positive by Section 7.1, is uniformly bounded.

3. The boundary circle $u(\partial S)$ of $\mathbf{u} = [u] \in \mathcal{N}$ is disjoint from $U_{\partial N}$ because the restriction of u to the disc component of S must be disjoint from $U_{\partial N}$ by Lemma 3.5.1. If $u(S)$ intersects U_B then $\mathbf{u}$ is un-noded and equivalent to a local Bishop disc u_{ε,b_o} by Lemma 3.3.1 combined with the final paragraph of Remark 4.4.1.

The local Bishop discs u_{ε,b_o}, $\varepsilon \in (0,\delta)$, represent elements

$$\mathbf{u}_{\varepsilon,b_o} = \left[\mathbb{D}, \mathrm{i}, \emptyset, \{1, \mathrm{i}, -1\}, u_{\varepsilon,b_o}\right]$$

in $\mathcal{N}$. The corresponding local Bishop filling can be identified with $(0,\delta)$. We truncate the nodal moduli space $\mathcal{N}$ via

$$\mathcal{N}_{\mathrm{cut}} = \mathcal{N} \setminus (0, \delta/2)\,.$$

Remark 7.3.1. If there exists a compact subset K of $\hat{W}$ such that $\mathbf{u}(S)$ is contained in K for all $\mathbf{u} \in \mathcal{N}$, then the Gromov compactification of $\mathcal{M}_\gamma$ can be identified with a subset of $\mathcal{N}$ by taking equivalence classes, see [25].

7.4 Cauchy–Riemann section

The moduli space $\mathcal{N}$ is the zero set

$$\mathcal{N} = \left\{\mathbf{u} \in \mathcal{Z} \,\middle|\, \bar{\partial}_{\hat{J}}\mathbf{u} = \mathbf{0}\right\}$$

of the Cauchy–Riemann operator $\bar{\partial}_{\hat{J}}$, which appears as a section into the bundle

$$p\colon\; \mathcal{W} \longrightarrow \mathcal{Z}$$

over $\mathcal{Z}$. The fibre of p over $\mathbf{u} = \left[S, j, D, \{1, \mathrm{i}, -1\}, u\right] \in \mathcal{Z}$ consists of equivalence classes $\boldsymbol{\xi} = \left[S, j, D, \{1, \mathrm{i}, -1\}, u, \xi\right]$ of continuous sections ξ of $\mathrm{Hom}\left(TS, u^*T\hat{W}\right)$ so that for each $z \in S$ the map $\xi(z)\colon\; T_zS \to T_{u(z)}\hat{W}$ is complex anti-linear with respect to $j(z)$ and $\hat{J}\big(u(z)\big)$. Moreover, ξ is of Sobolev class H^2_{loc} on $S \setminus |D|$ and of weighted Sobolev class $H^{2,\sigma}$ near $|D|$ similarly to the description at the beginning of Section 7.1, see [43, Section 1.2]. Two such sections $\left(S, j, D, \{m_0, m_1, m_2\}, u, \xi\right)$ and $\left(S', j', D', \{m_0', m_1', m_2'\}, u', \xi'\right)$ are **equivalent**, if there exists an equivalence φ of stable nodal disc maps $\left(S, j, D, \{m_0, m_1, m_2\}, u\right)$ and $\left(S', j', D', \{m_0', m_1', m_2'\}, u'\right)$ as described at the beginning of Section 7.2 such that $\xi' \circ T\varphi = \xi$. By adapting [43, Theorem 1.9] to the situation of the current Sections 6 and 7 we obtain a natural second countable paracompact Hausdorff topology on the total space $\mathcal{W}$ and the bundle projection $p\colon\; \mathcal{W} \to \mathcal{Z}$ that maps $\left[S, j, D, \{1, \mathrm{i}, -1\}, u, \xi\right]$ to $\left[S, j, D, \{1, \mathrm{i}, -1\}, u\right]$ is continuous. Furthermore $p: \mathcal{W} \to \mathcal{Z}$ constitutes a strong polyfold bundle in view of [43, Theorem 1.10].

The **Cauchy–Riemann operator** $\bar{\partial}_{\hat{J}}$ is the section of p given by

$$\bar{\partial}_{\hat{J}}\mathbf{u} := \left[S, j, D, \{1, \mathrm{i}, -1\}, u, \tfrac{1}{2}\big(Tu + \hat{J}(u)\circ Tu\circ j\big)\right]$$

for all $\mathbf{u} = \left[S, j, D, \{1, \mathrm{i}, -1\}, u\right] \in \mathcal{Z}$. For a representative we write $\bar{\partial}_{\hat{J}}u$ also. As in [43, Theorem 1.11] the Cauchy–Riemann operator $\bar{\partial}_{\hat{J}}: \mathcal{Z} \to \mathcal{W}$ is a scale-smooth component-proper Fredholm section that admits a natural orientation which we describe in Remark 7.4.3 below. The Fredholm index of $\bar{\partial}_{\hat{J}}: \mathcal{Z} \to \mathcal{W}$ is 1 by the index computation in Section 5.3 taking local transversal constraints from Section 7.2 in view of Theorem 6.11.1 into account. As in [64, Section 5.3] the vertical differential of a local representation of $\bar{\partial}_{\hat{J}}$ near the local Bishop discs $\mathbf{u}_{\varepsilon,b_o}$, $\varepsilon \in (0,\delta)$ has a right-inverse. The same holds true for all simple stable nodal holomorphic discs in $\mathcal{Z}$ due to the generic choice of $\hat{J}$, see Section 5.3.

Remark 7.4.1. Preparing the orientation considerations in Remark 7.4.3 we will establish **homotopically unique trivialisations** under the assumption that the second Stiefel–Whitney class of N^* vanishes. This approach requires to build up the spaces $\mathcal{H}^{3,\sigma}(S,j)$ with continuous maps on S/D *homotopic* in $(\hat{W}, N^*)$ to a local Bisphop disc, see item (3) in Section 7.1. As the relative homotopy class is preserved under Gromov convergence (see [25]) this is not a restriction.

Consider the space of continuous maps $(\mathbb{D}, \partial\mathbb{D}) \to \big(\hat{W}, N^*\big)$ sending the marked points $\{1\}$ and $\{\mathrm{i}^k\}$ into γ and $\vartheta^{-1}(\mathrm{i}^k)$, $k = 1, 2$, resp. Denote by $\mathcal{C}$ the connected component of the Bishop disc $u_0 = u_{\delta/2,b_o}$. We claim that for all $u \in \mathcal{C}$ the pull back bundle u^*TN^* has a canonical trivialisation.

In order to specify what is meant by this we describe the situation for u_0. By Section 3.1 the base point u_0 of $\mathcal{C}$ is the map

$$(\mathbb{D}, \partial\mathbb{D}) \longrightarrow \Big((-\infty, 0] \times \mathbb{R} \times \mathbb{C} \times T^*B, \{0\} \times \{0\} \times \mathbb{C}^* \times B\Big)$$

given by

$$u_0(z) = \left(\tfrac{\delta^2}{4}\big(|z|^2 - 1\big), 0, \delta \cdot z, b_o\right).$$

in the local model U_B. The embedded path γ corresponds to $\{0\}\times\{0\}\times\mathbb{R}^+\times\{b_o\}$ as oriented curve and the pages $\vartheta^{-1}(\mathrm{i}^k)$, $k=1,2$, correspond to $\{0\}\times\{0\}\times\mathbb{R}^+\mathrm{i}^k\times B$, resp. The co-orientation of the pages $\vartheta^{-1}(\mathrm{i}^k)$, $k=1,2$, given in Section 2.1 is represented by the normal vectors $(0,0,\mathrm{i}^{k+1},0)$, resp. The local model defines a local frame $\partial_s,\partial_t,\partial_x,\partial_y,\partial_{\mathbf{p}},\partial_{\mathbf{q}}$ of the tangent bundle $T\hat{W}$ near $(-\infty,0]\times\mathbb{R}\times\mathbb{C}\times\{b_o\}$ inducing a trivialisation $\Phi\colon\; u_0^*T\hat{W}\to\mathbb{D}\times\mathbb{R}^{2n}$ of the pull back bundle $u_0^*T\hat{W}$. The trivialisation Φ restricts to a trivialisation $\Phi\colon\; u_0^*TN^*\to\partial\mathbb{D}\times\mathbb{R}^n$ of $u_0^*TN^*$, which corresponds to the sub-frame $\partial_x,\partial_y,\partial_{\mathbf{p}}$. Further, Φ restricts to isomorphisms $\Phi\colon\; T_{u_0(1)}\gamma\to\{1\}\times\mathbb{R}$ via the vector field ∂_x and $\Phi\colon\; T_{u_0(\mathrm{i}^k)}\big(\vartheta^{-1}(\mathrm{i}^k)\big)\to\{\mathrm{i}^k\}\times\mathbb{R}^{n-1}$, $k=1,2$, via the sub-frames $\partial_y/-\partial_x,\partial_{\mathbf{p}}$. The co-orientations of the pages $\vartheta^{-1}(\mathrm{i}^k)$ correspond to $-\partial_x/-\partial_y$, resp. We remark that Φ is not a complex trivialisation of the complex bundle pair $\big(u_0^*T\hat{W},u_0^*TN^*\big)$ as used to compute the Maslov index to be 2, see [55, Proposition 8].

Given any $u\in\mathcal{C}$ we claim that the pull back bundle u^*TN^* admits a homotopically unique trivialisation Φ_u with the properties listed for $\Phi_{u_0}:=\Phi|_{u_0^*TN^*}$. To see this let u_τ, $\tau\in[0,1]$, be a path in $\mathcal{C}$ connecting u_0 with $u_1=u$ and define $U\colon\;[0,1]\times\partial\mathbb{D}\to\hat{W}$ by $U(\tau,z):=u_\tau(z)$. By [47, Corollary 3.4.5] there exists a trivialisation $\Phi_U\colon\; U^*TN^*\to\big([0,1]\times\partial\mathbb{D}\big)\times\mathbb{R}^n$ that extends Φ_{u_0}. As above we denote the coordinates of $\mathbb{R}^n$ by $(x,y,\mathbf{p})$. We can assume that Φ_U restricts to isomorphisms $\Phi_U\colon\;\big(U(\,.\,,1)\big)^*T\gamma\to\big([0,1]\times\{1\}\big)\times\mathbb{R}$ with $\mathbb{R}$ provided with the coordinate x as well as $\Phi_U\colon\;\big(U(\,.\,,\mathrm{i}^k)\big)^*T\big(\vartheta^{-1}(\mathrm{i}^k)\big)\to\big([0,1]\times\{\mathrm{i}^k\}\big)\times\mathbb{R}^{n-1}$, $k=1,2$, with $\mathbb{R}^{n-1}$ provided with coordinates $(y/x,\mathbf{p})$ and co-oriantations $-\partial_x/-\partial_y$, resp. The claimed trivialisation Φ_u is $\Phi_{u_1}=\Phi_U|_{\{1\}\times\partial\mathbb{D}}$.

It remains to show homotopic uniqueness of Φ_u, i.e. that Φ_u is independent of the chosen path u_τ up to homotopy: We consider a loop u_τ in $\mathcal{C}$ for $\tau\in T^1=\mathbb{R}/2\mathbb{Z}$ extending a path u_τ, $\tau\in[0,1]$, in $\mathcal{C}$ as above and define $\hat{U}\colon\; T^1\times\partial\mathbb{D}\to\hat{W}$ by $\hat{U}(\tau,z):=u_\tau(z)$. The claim will follow by constructing a trivialisation $\Phi_{\hat{U}}$ that shares the triviality properties established for Φ_U.

Restricted to $[0,1]\times\partial\mathbb{D}$ we define $\Phi_{\hat{U}}$ to be equal to Φ_U. As $\hat{U}(T^1\times\{1\})$ is a subset of the embedded interval $\gamma=[0,1]$ and the tangent bundle TN^* is trivialised by $\partial_x,\partial_y,\partial_{\mathbf{p}}$ along $\gamma\cap U_B$ there exists by [47, Corollary 3.4.8] a trivialisation $\Phi_\gamma\colon\; T_\gamma N^*\to[0,1]\times\mathbb{R}^n$

that extends the canonical trivialisation over $\gamma \cap U_B$ such that $(x, y, \mathbf{p})$ are coordinates on $\mathbb{R}^n$ and such that $\Phi_\gamma : T\gamma \to [0,1] \times \mathbb{R}$ is provided with the fibre coordinate x.

Gluing the trivialisations $\Phi_{\hat{U}}$ and Φ_γ via the identity along the overlap we obtain a trivialisation (still denoted by) $\Phi_{\hat{U}}$ over $\big([0,1] \times \partial\mathbb{D}\big) \cup \big(T^1 \times \{1\}\big)$. In other words, $\Phi_{\hat{U}}$ trivialises $\hat{U}^*TN^*$ over the boundary of the 2-disc

$$\big(T^1 \times \partial\mathbb{D}\big) \setminus \Big(\big([0,1] \times \partial\mathbb{D}\big) \cup \big(T^1 \times \{1\}\big)\Big).$$

By the assumption that the second Stiefel–Whitney class of N^* vanishes this trivialisation extends to a trivialisation of $\hat{U}^*TN^*$, see [34, p. 75 and Section 3.3]. Hence, $\Phi_{\hat{U}} : \hat{U}^*TN^* \to \big(T^1 \times \partial\mathbb{D}\big) \times \mathbb{R}^n$ is a trivialisation with fibre coordinates $(x, y, \mathbf{p})$. By construction we have a trivialisation $\Phi_{\hat{U}} : \big(\hat{U}(\,.\,, 1)\big)^* T\gamma \to \big(T^1 \times \{1\}\big) \times \mathbb{R}$ with fibre coordinate x. Further, because a co-oriented linear subspace of $\mathbb{R}^n$ of codimension 1 is determined by the normal vector and S^{n-1}, $n \geq 3$, is simply connected we can assume that we have trivialisations $\Phi_{\hat{U}} : \big(\hat{U}(\,.\,, \mathrm{i}^k)\big)^* T\big(\vartheta^{-1}(\mathrm{i}^k)\big) \to \big(T^1 \times \{\mathrm{i}^k\}\big) \times \mathbb{R}^{n-1}$, $k = 1, 2$, with fibre coordinates $(y/x, \mathbf{p})$ and co-oriantations $-\partial_x / -\partial_y$, resp.

Consequently, $u_1^*TN^*$ shares the same triviality properties as u_0 independently of the chosen path u_t such that Φ_{u_1} is homotopically unique as claimed.

Remark 7.4.2. If the second Stiefel–Whitney class $w_2(TN^*)$ of N^* is not trivial a variant of Remark 7.4.1 gives **homotopically unique stable trivialisations** assuming N^* to be orientable and that $w_2(TN^*)$ lifts to a class in $H^2(\hat{W}; \mathbb{Z}_2)$.

Following [26, Chapter 8.1] we choose a triangulation of $\hat{W}$ such that N will be a subcomplex and $B \cup \partial N$ a subcomplex of N. The assumptions made allow the choice of a **relative spin structure** on $(\hat{W}, N^*)$ which is a choice of orientation on N^*, an oriented vector bundle V over the 3-skeleton $\hat{W}_{[3]}$ of $\hat{W}$ such that $w_2(V)$ restricts to $w_2(TN^*)$, and a spin structure on the vector bundle $TN^* \oplus V$ over the 2-skeleton $N^*_{[2]}$ of N^*. Such a choice of a spin structure is possible because w_2 of $TN^* \oplus V$ over $N^*_{[2]}$ vanishes, see [12].

As in Remark 7.4.1 we consider the space of continuous maps $(\mathbb{D}, \partial\mathbb{D}) \to \big(\hat{W}, N^*\big)$ that map $\{1\}$ and $\{\mathrm{i}^k\}$ into γ and $\vartheta^{-1}(\mathrm{i}^k)$, $k = 1, 2$, resp. By simplicial approximation (see

[16, Theorem IV.22.10]) we can replace all maps and homotopies of maps by simplicial representatives u and u_t up to homotopy. Therefore, the proof of [26, Theorem 8.1.1] yields homotopically unique trivialisations of $u^*(TN^* \oplus V)$ and u^*V. Similarly to Remark 7.4.1 we can achieve that $T\gamma$ and $T\big(\vartheta^{-1}(\mathrm{i}^k)\big)$, $k = 1, 2$, correspond to $\{1\} \times \mathbb{R}$ and $\{\mathrm{i}^k\} \times \mathbb{R}^{n-1}$, resp., in the trivialisation $\mathbb{D} \times \mathbb{R}^{n+\upsilon}$ of $u^*(TN^* \oplus V)$, where υ denotes the rank of the vector bundle V. Moreover, by possibly changing the spin structure on $TN^* \oplus V$ over $N^*_{[2]}$ we can assume that the obtained trivialisation of $u_0^*(TN^* \oplus V)$ for the Bishop disc u_0 is homotopic to the canonical one induced by Φ_{u_0}, see Remark 7.4.1.

Remark 7.4.3. The canonical trivialisations of the involved pull back bundles in Remark 7.4.1 and 7.4.2 orient the Cauchy–Riemann section $\bar{\partial}_{\hat{J}} : \mathcal{Z} \to \mathcal{W}$ in a natural way. This is based on [26, Lemma 8.1.4].

Namely, given a complex bundle pair (E, F) over $(\mathbb{D}, \partial\mathbb{D})$ such that the real sub-bundle F is trivial over $\partial\mathbb{D}$ each trivialisation orients the associated linear Cauchy–Riemann operator. The complexification of the trivialisation extends to a complex trivialisation of E over an annulus neighbourhood of $\partial\mathbb{D}$. Collapsing the inner boundary component of a slightly smaller annulus neighbourhood of $\partial\mathbb{D}$ yields a complex bundle pair over a one-noded disc. Over the sphere component the Cauchy–Riemann operator admits the complex orientation, which is canonical. Over the disc component the Cauchy–Riemann operator is onto with kernel consisting of constant sections. Hence, the kernel is isomorphic to an Euclidean space canonically, so that the Cauchy–Riemann operator is canonically oriented over the disc component. Incorporating the matching condition of the bundles over the two components the functoriality properties of the determinant line bundle canonically determine an orientation of the Cauchy–Riemann operator on (E, F), see [26, Lemma 8.1.4] and cf. [43, Section 5.10].

Observe that this construction is compatible with point-wise boundary conditions and also allows to begin with a complex bundle pair (E, F) with matching conditions over a noded disc.

In order to orient the linearised Cauchy–Riemann operator at an un-noded element u

of $\mathcal{H}^{3,\sigma}(S,j)$ apply the above construction to the complex bundle pair $\left(u^*T\hat{W}, u^*TN^*\right)$ in the context of Remark 7.4.1 (restricting to the connected component of discs homotopic to a local Bishop disc), and to the complex bundle pairs $\left(u^*(T\hat{W}\oplus V_{\mathbb{C}}), u^*(TN^*\oplus V)\right)$ and $(u^*V_{\mathbb{C}}, u^*V)$, where $V_{\mathbb{C}} := V\otimes\mathbb{C}$, in the context of Remark 7.4.2, resp. For the latter use the arguments in the proof of [26, Theorem 8.1.1] and the observation that the noded discs in $\mathcal{H}^{3,\sigma}(S,j)$ are at least of codimension 2. In fact, we obtain canonical orientations of the linearised Cauchy–Riemann operator at noded elements of $\mathcal{H}^{3,\sigma}(S,j)$ also with the above construction.

With the proceeding remarks a canonical orientation of the Cauchy–Riemann section $\bar{\partial}_{\hat{J}}\colon\ \mathcal{Z}\to\mathcal{W}$ is obtained as in [43, Section 5.11]. Simply, replace the *complex orientation* of the sphere case by the canonical orientation induced by boundary trivialisations of pull back bundles in the arguments of [43, Section 5.11]. Furthermore observe that preservation of orientations of the partial Kodaira differentials on the Riemann moduli spaces is ensured by Theorem 6.11.1, $\mathcal{Z}_\tau$ is at least of codimension 2 for all non-trivial nodal types τ by Proposition 6.9.4, automorphisms of nodal discs in $\mathcal{Z}$ restrict to the identity on the disc component as well as that we can collapse the interior boundary component of a small collar annulus in [26, Lemma 8.1.4] such that auxiliary marked points are contained on the resulting sphere components exclusively.

Proof of Theorem 5.1.2 part (ii). We place ourselves in to the situation of Section 5.2 and 5.3; but this time we do not assume semi-positivity as in Theorem 5.1.2 part (i). Instead, we assume the vanishing of $w_2(TN^*)$ or the relative spin condition as formulated in Theorem 5.1.2 part (ii) so that Remark 7.4.3 applies. The aim is to derive a contradiction to the existence of a compact subset K of $\hat{W}$ such that $u(S)\subset K$ for all $\mathbf{u}=\left[S,j,D,\{1,\mathrm{i},-1\},u\right]\in\mathcal{N}$. Theorem 5.1.2 part (ii) will then follow as in the proof of part (i).

We argue by contradiction assuming that such a compact subset K as above exists. The arguments form Remark 4.4.1 under the assumed C^0-bounds on $\mathcal{N}$ combined with Section 7.3 show compactness of $\mathcal{N}_{\mathrm{cut}}$, see Remark 7.3.1. Let W_K be a relative compact open neighbourhood of K in $\hat{W}$. Using Sobolev embedding we choose a neighborhood

$\mathcal{U} \subset \mathcal{Z}$ of $\mathcal{N}_{\text{cut}}$ such that $u(S) \subset W_K \setminus \left(U'_B \cup U_{\partial N}\right)$ for all $\mathbf{u} = \left[S, j, D, \{1, \mathrm{i}, -1\}, u\right]$ in $\mathcal{U}$, where $U'_B \subset U_B$ is defined as U_B but with δ replaced by $\delta/2$ in the proof of Lemma 3.3.1.

Let $\lambda : \mathcal{W} \to \mathbb{Q} \cap [0, \infty)$ be a **scale⁺-multisection** of $p : \mathcal{W} \to \mathcal{Z}$, i.e. λ is a groupoidal functor which in a local presentation is given by finitely many weighted local scale⁺-sections (s_i, w_i), $w_i \in \mathbb{Q} \cap (0, \infty)$, of total weight $\sum w_i = 1$ such that $\lambda(\boldsymbol{\xi})$ is the sum of those weights w_i for which the corresponding sections s_i satisfy $s_i\big(p(\boldsymbol{\xi})\big) = \boldsymbol{\xi}$; we set $\lambda(\boldsymbol{\xi}) = 0$ if there is no such section among the s_i, cf. [40, Definition 3.34]. The **support** of λ is the smallest closed set in $\mathcal{Z}$ outside which λ is trivial in the sense that $\lambda(0_{\mathbf{u}}) = 1$ for these $\mathbf{u} \in \mathcal{Z}$, see [40, Definition 3.35]. The **solution set**

$$\mathcal{S} = \left\{\mathbf{u} \in \mathcal{Z} \,\middle|\, \lambda\big(\bar{\partial}_{\jmath}\mathbf{u}\big) > 0\right\}$$

of the pair $\big(\bar{\partial}_{\jmath}, \lambda\big)$ is the set of all $\mathbf{u} = \left[S, j, D, \{1, \mathrm{i}, -1\}, u\right] \in \mathcal{Z}$ for which in a local presentation of λ there exist at least one s_i such that $\bar{\partial}_{\jmath} u = s_i(u)$ and $\lambda\big(\bar{\partial}_{\jmath}\mathbf{u}\big)$ is the sum of all the weights w_i for which the corresponding s_i satisfy such an equation. The solution set $\mathcal{S}$ is equipped with the **weight function**

$$\lambda_{\bar{\partial}_{\jmath}} : \ \mathcal{Z} \longrightarrow \mathbb{Q} \cap (0, \infty)\,, \quad \mathbf{u} \longmapsto \lambda\big(\bar{\partial}_{\jmath}\mathbf{u}\big)\,,$$

see [40, Section 4.3].

With [40, Theorem 4.17] we choose λ such that the support of λ is contained in $\mathcal{U}$ and that $\big(\bar{\partial}_{\jmath}, \lambda\big)$ is **transverse**. The latter means that the vertical differentials

$$\big(\bar{\partial}_{\jmath}\big)'(u) - s_i'(u)$$

of local presentations $\bar{\partial}_{\jmath} u$ of $\bar{\partial}_{\jmath}\mathbf{u}$ and s_i of λ are surjective for all $\mathbf{u} \in \mathcal{S}$ and for all (the finitely many) i, see [40, Definition 4.7(1)]. If $\big(\bar{\partial}_{\jmath}\big)'(u)$ is onto for an un-noded $\mathbf{u} \in \mathcal{N}$, which is representable by a necessarily simple holomorphic disc map, we choose λ to be a single local section s_1 that is identically 0 in a neighbourhood of $\mathbf{u}$ in $\mathcal{Z}$. This is possible in view of the proof of [40, Theorem 4.17]. In particular, λ is trivial over those $\mathbf{u}$. As observed right before Remark 7.4.1 this applies to all local Bishop discs $\mathbf{u}_{\varepsilon, b_o}$, $\varepsilon \in [\delta/2, \delta)$, so that λ is trivial over all local Bishop discs $\mathbf{u}_{\varepsilon,b_o}$. Consequently, the truncated solution set

$$\mathcal{S}_{\mathrm{cut}} = \mathcal{S} \setminus (0, \delta/2)$$

of $(\bar{\partial}_{\hat{J}}, \lambda)$ is a 1-dimensional oriented compact branched *suborbifold* with boundary $\partial\mathcal{S}_{\mathrm{cut}}$ given by the single Bishop disc $\mathbf{u}_{\delta/2,b_o}$, see [40, Theorem 4.17] or [43, Section 1.4]. A collar neighbourhood of $\partial\mathcal{S}_{\mathrm{cut}}$ in $\mathcal{S}_{\mathrm{cut}}$ is equal to a collar neighbourhood of $\partial\mathcal{N}_{\mathrm{cut}}$ in $\mathcal{N}_{\mathrm{cut}}$ given by the local Bishop discs $\mathbf{u}_{\varepsilon,b_o}$, $\varepsilon \in [\delta/2, \delta)$.

Furthermore observe that by compactness of $\mathcal{S}_{\mathrm{cut}}$ the intersection $\mathcal{S}_{\mathrm{cut}} \cap \mathcal{Z}_\tau$ is not empty only for finitely many nodal types τ. Therefore, we choose $(\bar{\partial}_{\hat{J}}, \lambda)$ to be transverse along the subpolyfolds $\mathcal{Z}_\tau$ for these nodal types τ turning the subsets $\mathcal{S}_{\mathrm{cut}} \cap \mathcal{Z}_\tau$ into suborbifolds of $\mathcal{S}$. As the codimensions will be at least 2 whenever the nodal type τ is non-trivial, the resulting suborbifolds $\mathcal{S}_{\mathrm{cut}} \cap \mathcal{Z}_\tau$ have negative dimension, hence, are empty. Therefore, all elements in $\mathcal{S}_{\mathrm{cut}}$ are un-noded and have trivial isotropy as they can be represented by un-noded stable nodal disc maps with trivial automorphism group. In other words, $\mathcal{S}_{\mathrm{cut}}$ is a 1-dimensional oriented compact branched *manifold* with precisely one boundary point, which has weight 1. This contradicts the fact that by [62, Lemma 5.11] the oriented sum of the weights taken over all boundary points vanishes. □

Remark 7.4.4. We give an alternative argument to obtain a contradiction which does not use the classification of 1-dimensional oriented compact branched manifolds with boundary given in [62, Section 5.4]: We identify γ with the interval $[0, 3\delta]$ such that $(0, \delta)$ corresponds to the local Bishop family and $[2\delta, 3\delta]$ is not contained in the image of the evaluation map $\mathrm{ev}\colon\ \mathcal{S} \to \gamma$ that evaluates $\mathbf{u}$ at the first boundary marked point 1. Let f be a smooth function on $[0, 3\delta]$ with support in $(\delta/2, \delta)$ such that $\int_0^{3\delta} f(x)\mathrm{d}x = 1$. Because ev restricts to a degree 1 map on the local Bishop discs,

$$\int_{(\mathcal{S}_{\mathrm{cut}}, \lambda_{\bar{\partial}})} \mathrm{ev}^*(f\mathrm{d}x) = 1$$

writing $\lambda_{\bar{\partial}}$ for the weight function $\lambda_{\bar{\partial}_{\hat{J}}}$. Denote by f_1 the function obtained from f by shifting f by 2δ and observe that the closed 1-form $(f - f_1)\mathrm{d}x$ has a primitive $g(x) = \int_0^x \big(f(t) - f_1(t)\big)\mathrm{d}t$

with support in $(\delta/2, 3\delta)$. Hence, $\mathrm{ev}^*\big((f - f_1)\mathrm{d}x\big)$ has primitive $\mathrm{ev}^* g$ and

$$\int_{(\mathcal{S}_{\mathrm{cut}},\lambda_{\bar{\partial}})} \mathrm{ev}^*(f_1 \mathrm{d}x) = 0$$

as the support of f_1 is contained in $(5\delta/2, 3\delta)$. With Stokes theorem [41, Theorem 1.27] for weighted integrals

$$1 = \int_{(\mathcal{S}_{\mathrm{cut}},\lambda_{\bar{\partial}})} \mathrm{ev}^*\big((f - f_1)\mathrm{d}x\big) = \int_{(\partial\mathcal{S}_{\mathrm{cut}},\lambda_{\bar{\partial}})} \mathrm{ev}^* g = g(\delta/2)\,.$$

As $g(\delta/2) = 0$ we reach the desired contradiction.

Acknowledgments

This research is supported by the *Institute for Advanced Study* (Princeton, NJ) in 2015, by the *Mathematisches Forschungsinstitut Oberwolfach* in 2020, and is part of a project in the SFB/TRR 191 *Symplectic Structures in Geometry, Algebra and Dynamics*, funded by the DFG.

Part of the research of this article was initiated while S.S. and K.Z.'s stay at the *Institut for Advanced Studies* Princeton. We would like to thank Helmut Hofer for the invitation and for introducing us to polyfolds. K.Z. would like to thank the *Mathematisches Forschungsinstitut Oberwolfach* for the support. We would like to thank Katrin Wehrheim for her constant support and motivation. We thank Peter Albers, Matthew Strom Borman, Joel Fish, Hansjörg Geiges, Stefan Nemirovski, and Sven Prüfer for enlightening discussions.

References

[1] C. Abbas, *An introduction to compactness results in symplectic field theory*, Springer, Heidelberg, (2014), viii+252. 36, 37, 68

[2] C. Abbas, K. Cieliebak, H. Hofer, The Weinstein conjecture for planar contact structures in dimension three, *Comment. Math. Helv.* **80** (2005), 771–793. 5

[3] C. Abbas, H. Hofer, *Holomorphic curves and global questions in contact geometry*, Birkhäuser, (2019), xii+322. 43

[4] R. A. Adams, J. J. F. Fournier, *Sobolev spaces*, Pure and Applied Mathematics (Amsterdam), **140**, Second, Elsevier/Academic Press, Amsterdam, (2003), xiv+305. 72

[5] B. Aebischer, M. Borer, M. Kälin, Ch. Leuenberger, H. M. Reimann, *Symplectic geometry*, Progress in Mathematics, **124**, An introduction based on the seminar in Bern, 1992, Birkhäuser Verlag, Basel, (1994), xii+239. 47

[6] P. Albers, H. Hofer, On the Weinstein conjecture in higher dimensions, *Comment. Math. Helv.* **84** (2009), 429–436. 5

[7] M. Audin, Symplectic and almost complex manifolds, with an appendix by P. Gauduchon, Chapter II of [8], pp. 41–74. 44

[8] M. Audin and J. Lafontaine (eds.), *Holomorphic Curves in Symplectic Geometry*, Progr. Math. **117**, Birkhäuser Verlag, Basel (1994). 89

[9] M. Audin, J. Lafontaine and L. Polterovich, Symplectic rigidity: Lagrangian submanifolds, Chapter X of [8], pp. 271–321. 11

[10] K. Barth, J. Schneider, K. Zehmisch, Symplectic dynamics of contact isotropic torus complements, *Münster J. Math.* **12** (2019), 31–48. 20

[11] M. S. Borman, Y. Eliashberg, E. Murphy, Existence and classification of overtwisted contact structures in all dimensions, *Acta Math.* **215** (2015), 281–361. 2, 10

[12] A. Borel, F. Hirzebruch, Characteristic Classes and Homogeneous Spaces, I, *Amer. J. Math.* **80** (1958), 458–538. 83

[13] F. Bourgeois, A survey of contact homology, in: *New perspectives and challenges in symplectic field theory*, CRM Proc. Lecture Notes, 49, Amer. Math. Soc., Providence, RI, 2009, 45–71. 7

[14] F. Bourgeois, O. van Koert, Contact homology of left-handed stabilizations and plumbing of open books, *Commun. Contemp. Math.* **12** (2010), 223–263. 6, 7

[15] F. Bourgeois, Y. Eliashberg, H. Hofer, K. Wysocki, E. Zehnder, Compactness results in symplectic field theory, *Geom. Topol.* **7** (2003), 799–888. 32, 68

[16] G. E. Bredon, *Topology and Geometry*, Grad. Texts in Math. **139**, Springer-Verlag, New York (1993). 31, 84

[17] R. Casals, E. Murphy, F. Presas, Geometric criteria for overtwistedness, *J. Amer. Math. Soc.* **32** (2019), 563–604. 7, 11

[18] K. Cieliebak, Nonlinear Functional Analysis, lecture note, (2018). 34

[19] K. Cieliebak, Y. Eliashberg, *From Stein to Weinstein and back*, American Mathematical Society Colloquium Publications, 59, Symplectic geometry of affine complex manifolds, American Mathematical Society, Providence, RI, 2012, xii+364. 16

[20] F. Ding, H. Geiges, Contact structures on principal circle bundles, *Bull. Lond. Math. Soc.* **44** (2012), 1189–1202. 2

[21] M. Dörner, H. Geiges, K. Zehmisch, Open books and the Weinstein conjecture, *Q. J. Math.* **65** (2014), 869–885. 5

[22] Y. Eliashberg, Classification of overtwisted contact structures on 3-manifolds, *Invent. Math.* **98** (1989), 623–637. 1

[23] Y. Eliashberg, Filling by holomorphic discs and its applications, in: *Geometry of Low-Dimensional Manifolds, Vol. 2* (Durham, 1989), London Math. Soc. Lecture Note Ser. **151**, Cambridge University Press (1990), 45–67. 2, 3, 6, 15, 22

[24] Y. Eliashberg, A. Givental, H. Hofer, Introduction to symplectic field theory, *Geom. Funct. Anal.* **2000**, Special Volume, Part II, 560–673. 6, 29

[25] U. Frauenfelder, K. Zehmisch, Gromov compactness for holomorphic discs with totally real boundary conditions, *J. Fixed Point Theory Appl.* **17** (2015), 521–540. 27, 34, 38, 68, 80, 81

[26] K. Fukaya, Y.-G. Oh, H. Ohta, K. Ono, *Lagrangian intersection Floer theory: anomaly and obstruction. Part II*, AMS/IP Studies in Advanced Mathematics, **46**, American Mathematical Society, Providence, RI; International Press, Somerville, MA, (2009), i–xii and 397–805. 4, 83, 84, 85

[27] H. Geiges, K. Zehmisch, Eliashberg's proof of Cerf's theorem, *J. Topol. Anal.* **2** (2010), 543–579. 16, 32

[28] H. Geiges, K. Zehmisch, Symplectic cobordisms and the strong Weinstein conjecture, *Math. Proc. Cambridge Philos. Soc.* **153** (2012), 261–279. 13, 19, 29, 31, 32, 33

[29] H. Geiges, K. Zehmisch, How to recognise a 4-ball when you see one, *Münster J. Math.* **6** (2013), 525–554. 2, 29

[30] H. Geiges, K. Zehmisch, Reeb dynamics detects odd balls, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **15** (2016), 663–681. 13

[31] H. Geiges, K. Zehmisch, *A Course on Holomorphic Discs*, Birkhäuser Advanced Texts Basler Lehrbücher, Birkhäuser/Springer, Cham, 2023, xviii+189 pp. 44

[32] M. Gromov, Pseudoholomorphic curves in symplectic manifolds, *Invent. Math.* **82** (1985), 307–347. 2, 3

[33] M. Gromov, *Partial differential relations*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), **9** Springer-Verlag, Berlin, (1986), x+363. 2

[34] A. Hatcher, *Vector Bundles & K-Theory*, (2017), http://pi.math.cornell.edu/~hatcher/VBKT/VBpage.html 83

[35] H. Hofer, Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three, *Invent. Math.* **114** (1993), 515–563. 5

[36] H. Hofer, K. Wysocki, E. Zehnder, The dynamics on three-dimensional strictly convex energy surfaces, *Ann. of Math. (2)* **148** (1998), 197–289. 6, 34

[37] H. Hofer, K. Wysocki, E. Zehnder, Finite energy foliations of tight three-spheres and Hamiltonian dynamics, *Ann. of Math. (2)* **157** (2003), 125–255. 6, 34

[38] H. Hofer, K. Wysocki, E. Zehnder, A general Fredholm theory. I. A splicing-based differential geometry, *J. Eur. Math. Soc. (JEMS)* **9** (2007), 841–876. 6

[39] H. Hofer, K. Wysocki, E. Zehnder, A general Fredholm theory. II. Implicit function theorems, *Geom. Funct. Anal.* **19** (2009), 206–293. 6

[40] H. Hofer, K. Wysocki, E. Zehnder, A general Fredholm theory. III. Fredholm functors and polyfolds, *Geom. Topol.* **13** (2009), 2279–2387. 6, 86, 87

[41] H. Hofer, K. Wysocki, E. Zehnder, Integration theory on the zero sets of polyfold Fredholm sections, *Math. Ann.* **346** (2010), 139–198. 6, 88

[42] H. Hofer, K. Wysocki, E. Zehnder, Deligne–Mumford-type spaces with a view towards symplectic field theory, preprint, (2012) 6, 34, 38, 50, 68, 70

[43] H. Hofer, K. Wysocki, E. Zehnder, *Applications of polyfold theory I: The polyfolds of Gromov-Witten theory*, Mem. Amer. Math. Soc. **248** (2017), v+218 pp. 6, 34, 43, 68, 69, 70, 71, 75, 77, 78, 80, 81, 84, 85, 87

[44] Y. Huang, On Legendrian foliations in contact manifolds I: Singularities and neighborhood theorems, *Math. Res. Lett.* **22** (2015), 1373–1400. 9

[45] Y. Huang, On plastikstufe, bordered Legendrian open book and overtwisted contact structures, *J. Topol.* **10** (2017), 720–743. 11

[46] C. Hummel, *Gromov's compactness theorem for pseudo-holomorphic curves*, Progress in Mathematics, **151** Birkhäuser Verlag, Basel, (1997), viii+131. 68

[47] D. Husemoller, *Fibre bundles*, Graduate Texts in Mathematics **20**, Third, Springer-Verlag, New York, (1994), xx+353. 82

[48] M. Kwon, K. Wiegand, K. Zehmisch, Diffeomorphism type via aperiodicity in Reeb dynamics, *J. Fixed Point Theory Appl.* **24** (2022), Paper No. 21, 26 pp. 13, 19

[49] J. Latschev, C. Wendl, Algebraic torsion in contact manifolds (With an appendix by Michael Hutchings.), *Geom. Funct. Anal.* **21** (2011), 1144–1195. 6

[50] P. Massot, K. Niederkrüger, C. Wendl, Weak and strong fillability of higher dimensional contact manifolds, *Invent. Math.* **192** (2013), 287–373. 2, 3, 4, 5, 8, 9, 12, 20, 21, 23, 30, 32

[51] D. McDuff, Symplectic manifolds with contact type boundaries, *Invent. Math.* **103** (1991), 651–671. 29

[52] D. McDuff, D. Salamon, *Introduction to symplectic topology*, Oxford Mathematical Monographs, Second, The Clarendon Press Oxford University Press, New York, (1998), 1–486. 23, 24, 25, 41, 47

[53] D. McDuff, D. Salamon, *Introduction to symplectic topology*, Oxford Graduate Texts in Mathematics, Third, Oxford University Press, Oxford, (2017), xi+623. 44

[54] D. McDuff, D. Salamon, *J-holomorphic Curves and Symplectic Topology*, Amer. Math. Soc. Colloq. Publ. **52**, American Mathematical Society, Providence, RI (2004). 19, 30, 37, 39, 41, 42, 43, 64, 75

[55] K. Niederkrüger, The plastikstufe – a generalization of the overtwisted disk to higher dimensions, *Algebr. Geom. Topol.* **6** (2006), 2473–2508. 2, 3, 10, 11, 13, 14, 17, 32, 33, 82

[56] K. Niederkrüger, *On fillability of contact manifolds*, Habilitation à diriger des recherches, Université Paul Sabatier - Toulouse III, 2013, tel-00922320. 8, 9, 10

[57] K. Niederkrüger, A. Rechtman, The Weinstein conjecture in the presence of submanifolds having a Legendrian foliation, *J. Topol. Anal.* **3** (2011), 405–421. 5

[58] K. Niederkrüger, C. Wendl, Weak symplectic fillings and holomorphic curves, *Ann. Sci. Éc. Norm. Supér. (4)* **44** (2011), 801–853. 6

[59] J. Pardon, Contact homology and virtual fundamental cycles, *J. Amer. Math. Soc.* **32** (2019), 825–919. 3, 7

[60] L. Polterovich, The surgery of Lagrange submanifolds, *Geom. Funct. Anal.* **1** (1991), 198–210. 10

[61] J. Robbin, D. Salamon, A construction of the Deligne-Mumford orbifold, *J. Eur. Math. Soc. (JEMS)* **8** (2006), 611–699. 36, 41, 44, 51, 53

[62] D. Salamon, *Lectures on Floer homology*, in: Symplectic geometry and topology (Park City, UT, 1997), IAS/Park City Math. Ser. **7**, Amer. Math. Soc., Providence, RI, (1999), 143–229. 87

[63] D. Salamon, *Funktionentheorie*, Birkhäuser, Basel, (2011), viii + 218. 36, 37

[64] S. Suhr, K. Zehmisch, Polyfolds, cobordisms, and the strong Weinstein conjecture, *Adv. Math.* **305** (2017), 1250–1267. 6, 7, 29, 32, 81

[65] A. Weinstein, On the hypotheses of Rabinowitz' periodic orbit theorems, *J. Differential Equations* **33** (1979), 336–352. 5

[66] C. Wendl, *Finite energy foliations and surgery on transverse links*, Thesis (Ph.D.)–New York University, (2005), xii + 411. 68

[67] C. Wendl, *Lectures on Holomorphic Curves in Symplectic and Contact Geometry*, preprint, `arXiv: 1011.1690` 36, 49, 51, 53

[68] K. Zehmisch, Holomorphic jets in symplectic manifolds, *J. Fixed Point Theory Appl.* **17** (2015), 379–402. 32, 33

[69] K. Zehmisch, The Eliashberg–Gromov tightness theorem, Diplomarbeit, Universität Leipzig (2003). 3

AUTHORS

Wolfgang Schmaltz,
Fakultät für Mathematik,
Ruhr-Universität Bochum,
Universitätsstraße 150,
D-44801 Bochum, Germany
email: Wolfgang.Schmaltz@rub.de

Stefan Suhr,
Fakultät für Mathematik,
Ruhr-Universität Bochum,
Universitätsstraße 150,
D-44801 Bochum, Germany
email: Stefan.Suhr@rub.de

Kai Zehmisch,
Fakultät für Mathematik,
Ruhr-Universität Bochum,
Universitätsstraße 150,
D-44801 Bochum, Germany
email: Kai.Zehmisch@rub.de

Arnold Mathematical Journal
Volume 11, Issue 2, 2025

Kustaanheimo-Stiefel Transformation, Birkhoff-Waldvogel Transformation and Integrable Mechanical Billiards

Airi Takeuchi   Lei Zhao

Received 11 Jan 2024; Accepted 27 Nov 2024

Abstract:

The three-dimensional Kepler problem is related to the four-dimensional isotropic harmonic oscillators by the Kustaanheimo-Stiefel transformation. In the first part of this paper, we study how certain integrable mechanical billiards are related by this transformation. This in part illustrates the rotation-invariance of integrable reflection walls in the three-dimensional Kepler billiards found so far. The second part of this paper deals with the Birkhoff-Waldvogel Transformation of the three-dimensional problem with two Kepler centers. In particular, we establish an analogous theory of Levi-Civita planes for the Birkhoff-Waldvogel Transformation and show the integrability of certain three-dimensional two-center billiards via a different approach.

DOI: 10.56994/ARMJ.011.002.002

Airi Takeuchi, Lei Zhao

AMS Classification: 37C83, 70F99

Key words and phrases: Mechanical billiard systems, Kepler problem, Hooke problem, Integrability, Kustaanheimo-Stiefel transformation

1 Introduction

The Kustaanheimo-Stiefel transformation was introduced in [10, 11] using spinors as a way to regularize the three-dimensional Kepler problem. This transformation also admits formulation with quaternions [26, 20, 27], and can be regarded as an unfolding of the Levi-Civita regularization [13, 14] of the planar Kepler problem through the theory of Levi-Civita planes [19, 27].

The use of the conformal Levi-Civita transformation [15, 5, 14] to study planar integrable mechanical billiards defined with the Hooke and Kepler problems has been first pointed out in [16] and extended in [22].

In the first part of this note, we discuss some consequences of the K.S. (Kustaanheimo-Stiefel) transformation on integrable four-dimensional Hooke and integrable three-dimensional Kepler billiards. An n-dimensional mechanical billiard system is integrable if there exist n first integrals of the underlying natural mechanical system that satisfy the following conditions: they are functionally independent, in involution, and they remain invariant under reflections at the reflection wall (c.f. [24]). It is widely known that for the four-dimensional Hooke problem, a centered quadric reflection wall gives an integrable billiard system [6], [4]. We shall show that when this reflection wall is invariant under an S^1-symmetry of the K.S. transformation, then its image under the Hopf mapping is one of five special type of quadrics, with the Kepler center as a focus. This is consistent with the results of [23] and provides a partial explanation of why we conjecture that only these quadrics appear in three-dimensional integrable Kepler billiards. This conjecture is closely related to the analogue of Birkhoff-Poritsky conjecture [17] of the planar case in the setting of Kepler billiards. We can think that the restriction on the type of quadrics is forced by the S^1-invariance of the centered quadric reflection wall lying on the four-dimensional Hooke side. This generalizes the studies in the planar case [22], [16], [21], [28] to the spatial case of Kepler billiards based on the Levi-Civita transformation. (See also [7], [8], [9] for some other studies of integrable mechanical billiards in the plane) This way we obtain

Thoerem A. *Consider a surface of revolution in $\mathbb{R}^3$, by revolving a conic with a focus at the origin about its principal axis. Then reflecting spatial Kepler orbits (attracted to/repelled from the origin) off such a surface of revolution gives rise to an integrable mechanical billiard.*

This reproves the three-dimensional version of Theorem 14 in [23]. Moreover, though we shall not discuss this aspect in this article, the method provides a transformation that maps the orbits of one system to another, in such a way that the reflection law of the first system on an energy surface corresponds to that of the second system on its image. This method is not limited to integrable mechanical billiard systems.

The Kustaanheimo-Stiefel transformation has been extended to a transformation which simultaneously regularizes both double collisions in the spatial two-center problem first announced in a 1-page note of Stiefel-Waldvogel [18], which generalized the transformation of Birkhoff used in the planar case. The thesis of Waldvogel [25] provided a much more extensive geometrical study of this transformation. In particular, the relation between this transformation and the Kustaanheimo-Stiefel transformation has been clarified. Waldvogel later illustrated this theory again in [26] with the use of quaternions. In this article, we provide a quaternionic formulation of this Birkhoff-Waldvogel transformation in the spatial case, largely inspired by the studies of Waldvogel as well as combining the symplectic viewpoint of [27]. We investigate in part an analogous theory of Levi-Civita planes in this setting, consisting of planes and spheres in the space of quaternions $\mathbb{H} \cong \mathbb{R}^4$ and a reduction of this transformation to a dense open subset of $\mathbb{IH} \cong \mathbb{R}^3$, which already regularizes the double collisions without increasing the dimension of the space. With this we link integrable billiards on both sides, which illustrates some results in [23] with a different method.

Thoerem B. *Consider a surface of revolution in $\mathbb{R}^3$ by revolving a conic with foci at the two Kepler centers around the axis joining the centers. Then reflecting orbits of the spatial two-center problem off such a surface of revolution is an integrable mechanical billiard. Moreover, taking a finite combination of these surfaces does not destroy the integrability of the resulting two-center mechanical billiard systems.*

This provides an alternative proof of Theorem 14 in [23] for the spatial two-center case.

We organize this article as follows: In Section 2, we recall the theory of Kustaanheimo-Stiefel regularization, which largely follows [27]. Then we apply this transformation to link integrable mechanical billiards in Section 3. The theory of Birkhoff-Waldvogel transformation and the corresponding link on integrable mechanical billiards are discussed in Section 4.

2 The Kustaanheimo-Stiefel Transformation

In this section, we discuss the Kustaanheimo-Stiefel transformation. We follow the quaternionic formulation of [27].

A quaternion is represented as

$$z = z_0 + z_1 i + z_2 j + z_3 k, \quad z_0, z_1, z_2, z_3 \in \mathbb{R}$$

in which

$$i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j.$$

Addition and multiplication of quaternions are then naturally defined. With these operations, the quaternions form a non-commutative normed division algebra which we denote by $\mathbb{H}$. For a quaternion $z = z_0 + z_1 i + z_2 j + z_3 k$, its real part is given by

$$Re(z) = z_0$$

and its imaginary part is given by

$$Im(z) = z_1 i + z_2 j + z_3 k.$$

Furthermore, the conjugation of z is defined as

$$\bar{z} = z_0 - z_1 i - z_2 j - z_3 k.$$

The norm of z is defined as $|z| := \sqrt{z \cdot \bar{z}}$.

We denote the set of purely imaginary quaternions by

$$\mathbb{IH} = \{z \in \mathbb{H} \mid Re(z) = 0\}.$$

We identify $\mathbb{H}$ with $\mathbb{R}^4$ and denote by $\mathcal{S}^3$ the unit sphere

$$\{z \in \mathbb{H} \mid |z|^2 = z_0^2 + z_1^2 + z_2^2 + z_3^2 = 1\} \subset \mathbb{H}.$$

Also, we identify $\mathbb{IH}$ with $\mathbb{R}^3$. The unit sphere $\mathcal{S}^2$ therein is

$$\{z \in \mathbb{IH} \mid |z|^2 = z_1^2 + z_2^2 + z_3^2 = 1\} \subset \mathbb{IH}.$$

To introduce the Kustaanheimo-Stiefel transformation, we first recall the Levi-Civita transformation [14]

$$T^*(\mathbb{C} \setminus \{0\}) \to T^*(\mathbb{C} \setminus \{0\}), (z, w) \longmapsto \left(q = z \cdot z, p = \frac{z}{2|z|^2} \cdot w\right).$$

It is well-known that this transformation is canonical, and transforms the planar Kepler problem into the planar Hooke problem after making a proper time reparametrization on an energy level. To see this, we start with the shifted Hamiltonian of the Kepler problem and consider its zero-energy level:

$$\frac{|p^2|}{2} + \frac{m}{|q|} - f = 0.$$

The Levi-Civita transformation pulls this system back to

$$\frac{|w^2|}{8|z|^2} + \frac{m}{|z|^2} - f = 0.$$

We may now multiply this transformed Hamiltonian by $|z|^2$, which only reparametrizes the flow on this energy-level. We obtain

$$\frac{|w^2|}{8} + m - f|z|^2 = 0.$$

which is the restriction of the Hamiltonian of the planar Hooke problem $\frac{|w^2|}{8} - f|z|^2$ on its $(-m)$-energy hypersurface.

The whole construction is based on the complex square mapping

$$\mathbb{C} \setminus \{0\} \longmapsto \mathbb{C} \setminus \{0\}, \qquad z \longmapsto z^2,$$

which is a 2-to-1 conformal mapping.

A generalization of the complex square mapping with quaternions is the following *Hopf mapping*

$$\mathbb{H} \to \mathbb{IH}, \qquad z \longmapsto \bar{z}iz.$$

Note that this mapping is well-defined, since

$$Re(\bar{z}iz) = 0, \forall z \in \mathbb{H}.$$

This mapping is “S^1-to-1”, namely the circle

$$\{\exp(i\theta)z \mid, z \in \mathbb{H} \setminus \{0\},\ \theta \in \mathbb{R}/2\pi\mathbb{Z}\} \subset \mathbb{H}$$

is mapped under the Hopf mapping to the same point $\bar{z}iz \in \mathbb{IH}$.

Moreover, this mapping restricts to a mapping $\mathcal{S}^3 \to \mathcal{S}^2$. This is a mapping with S^1-fibres, and induces the non-trivial Hopf fibration

$$S^1 \hookrightarrow \mathcal{S}^3 \to \mathcal{S}^2.$$

Associated to the Hopf mapping, the Kustaanheimo-Stiefel mapping is defined as

$$T^*(\mathbb{H} \setminus \{0\}) \to \mathbb{IH} \times \mathbb{H}, (z, w) \longmapsto \left(Q = \bar{z}i \cdot z, P = \frac{\bar{z}i}{2|z|^2} \cdot w\right).$$

The fibers of the mapping are the circle orbits of the S^1-Hamiltonian action

$$\theta \cdot (z, w) \longmapsto (\exp(i\theta)z, \exp(i\theta)w)$$

on the cotangent bundle $T^*\mathbb{H}$. The bilinear function

$$BL(z, w) := Re(\bar{z}iw)$$

is the associated moment map.

We define

$$\Sigma := \{(z, w) | BL(z, w) = 0\} \subset T^*\mathbb{H} \cong \mathbb{H} \times \mathbb{H},$$

and

$$\Sigma^1 = \Sigma \setminus \{z = 0\}.$$

Both are invariant under this S^1-Hamiltonian action.

We define the restricted K.S. mapping as

$$KS := K.S.|_{\Sigma_1} : \Sigma^1 \to T^*(\mathbb{IH}\backslash\{0\}).$$

For the following lemma from [27], we present an alternative, simpler proof.

Lemma 1. *For the restricted Kustaanheimo-Stiefel mapping* $KS : \Sigma^1 \to T^*(\mathbb{IH}\backslash\{0\})$ *we have*

$$KS^*(Re(d\bar{P} \wedge dQ)) = Re(d\bar{w} \wedge dz)|_{\Sigma^1}.$$

Proof. We shall show

$$KS^*Re(\bar{P}dQ) = Re(\bar{w}dz)|_{\Sigma^1}. \tag{1}$$

which then implies the assertion of this lemma by taking differentials on both sides.

To see (1), we compute

$$\begin{aligned} \bar{P}dQ &= -\frac{\bar{w}i\bar{z}^{-1}}{2}((d\bar{z})iz + \bar{z}idz) \\ &= (-\bar{w}i\bar{z}^{-1}(d\bar{z})iz + \bar{w}dz)/2. \end{aligned} \tag{2}$$

The condition

$$BL(z, w) = Re(\bar{z}iw) = 0$$

is equivalent to

$$Re(z^{-1}iw) = 0.$$

Consequently, we also have

$$Re(\bar{w}i\bar{z}^{-1}) = 0.$$

This implies

$$Re(\bar{w}i\bar{z}^{-1}(d\bar{z})iz) = Re(\bar{w}i\bar{z}^{-1} \cdot Im((d\bar{z})iz)).$$

Since

$$Im((d\bar{z})iz) = -Im(\bar{z}(-i)dz),$$

we have

$$Re(\bar{w}i\bar{z}^{-1}(d\bar{z})iz) = -Re(\bar{w}i\bar{z}^{-1} \cdot Im(\bar{z}(-i)dz)) = -Re(\bar{w}i\bar{z}^{-1}\bar{z}(-i)dz) = -Re(\bar{w}dz), \tag{3}$$

where in the second equation, we have used

$$Re(\bar{w}i\bar{z}^{-1}) = 0.$$

The assertion (1) is thus obtained by combining the equations (2) and (3). □

On Σ^1, the orbits of the S^1-action mentioned above lie in the direction of the one-dimensional kernel distribution of the 2-form $Re(d\bar{w} \wedge dz)$. By the theory of symplectic reduction, the 2-form $Re(d\bar{w} \wedge dz)$ of Σ^1 gives rise to the reduced symplectic form ω_1 on the quotient space V^1 of Σ^1 by the S^1-action. Thus, the Kustaanheimo-Stiefel mapping induces a symplectomorphism

$$KS_{red} : (V_1, \omega_1) \to (T^*(\mathbb{IH}\backslash\{0\}), Re(d\bar{P} \wedge dQ)).$$

We have

$$KS = KS_{red} \circ \phi$$

in which $\phi : \Sigma^1 \to V_1$ is the quotient map.

Proposition 2. *Any zero-energy orbit of the four-dimensional Hooke problem with the shifted Hamiltonian*

$$\frac{||w||^2}{8} - f||z||^2 + m$$

in Σ^1 is sent via KS to a zero-energy orbit of the three-dimensional Kepler Problem in $T^*(\mathbb{IH}\backslash\{0\})$ with Hamiltonian

$$\frac{||P||^2}{2} + \frac{m}{||Q||} - f.$$

after a proper time reparametrization.

Proof. We first observe that the function BL is a first integral of the system

$$\left(T^*\mathbb{H}, Re(d\bar{w} \wedge dz), H = \frac{\|w\|^2}{8} - f\|z\|^2 + m\right).$$

This follows either from a direct verification, or alternatively from the invariance of H under the above mentioned (Hamiltonian) S^1-action. Consequently, the set Σ^1 is invariant under its flow.

We consider the restriction of this system on Σ^1. Any orbit of this restricted system descends to an orbit in the quotient system in (V_1, ω_1, H_1) so that

$$\phi^* H_1 = H,$$

which is consequently sent to an orbit via KS_{red} in the system

$$(T^*\mathbb{IH}, Re(d\bar{p} \wedge dq), K)$$

such that

$$KS_{red}^* K = H_1.$$

Applying ϕ^* to both sides of this identity, we get

$$H = \phi^* KS_{red}^* K = KS^* K.$$

From this we deduce

$$K = \frac{\|P\|^2\|Q\|}{2} + m - f\|Q\|.$$

Now we restrict the system to $\{K = 0\} = \{H = 0\}$. We observe that the restricted flow can now be time reparametrized (with factor $\|Q\|^{-1}$) into the restricted flow on the zero-energy hypersurface of the three-dimensional Kepler Hamiltonian

$$\frac{\|P\|^2}{2} + \frac{m}{\|Q\|} - f.$$

□

The link between Kustaanheimo-Stiefel transformation and the Levi-Civita transformation is given by the Levi-Civita planes. These are planes in $\mathbb{H}$ generated by two unit quaternions v_1, v_2 such that $v_1 \neq \pm v_2$ and satisfy

$$BL(v_1, v_2) = 0.$$

The key property of such a plane is that its image is a plane in $\mathbb{IH}$ and in relevant basis the restriction of the Hopf mapping is equivalent to the complex square mapping

$$\mathbb{C} \to \mathbb{C}, z \longmapsto z^2.$$

Therefore $K.S.$ is restricted to $L.C.$ on the tangent bundle of such a plane. We proceed with the details.

Definition 3. *A Levi-Civita plane is a plane in $\mathbb{H}$ spanned by two linearly independent unit quarternions $v_1, v_2 \in \mathbb{H}$ satisfying $BL(v_1, v_2) = 0$.*

Proposition 4. *The Hopf mapping*

$$\mathbb{H} \to \mathbb{IH}, \quad z \longmapsto \bar{z}iz$$

sends a Levi-Civita plane to a plane passing through the origin in $\mathbb{IH}$. On the other hand, any plane in $\mathbb{IH}$ passing through the origin is the image of a $\mathcal{S}^1$-family of Levi-Civita planes.

Proof. Let V be a Levi-Civita plane spanned by two unit, orthogonal quaternions v_1 and v_2 in $\mathbb{H}$: This means that we have

$$|v_1| = |v_2| = 1,\ BL(v_1, v_2) = 0 \text{ and } \langle v_1, v_2 \rangle = 0.$$

Then, we have

$$\bar{v}_1 i v_1 = -\bar{v}_2 i v_2,$$

which follows from the computation

$$\begin{aligned} 2\bar{v}_1 i v_1 + 2\bar{v}_2 i v_2 &= (\bar{v}_2 v_1 + \bar{v}_1 v_2)(\bar{v}_1 i v_2 + \bar{v}_2 i v_1) \\ &= 2\langle v_1, v_2 \rangle (\bar{v}_1 i v_2 + \bar{v}_2 i v_1) \qquad (4) \\ &= 0. \end{aligned}$$

For the first equation in (4) we used the following fact: The condition

$$BL(v_1, v_2) = 0$$

is equivalent to

$$\bar{v}_1 i v_2 - \bar{v}_2 i v_1 = 0.$$

Thus

$$(\bar{v}_1 v_2 - \bar{v}_2 v_1)(\bar{v}_1 i v_2 - \bar{v}_2 i v_1) = 0$$

which is equivalent to

$$\bar{v}_1 i v_1 + \bar{v}_2 i v_2 = \bar{v}_1 v_2 \bar{v}_1 i v_2 + \bar{v}_2 v_1 \bar{v}_2 i v_1.$$

Thus the quaternion $v_1 + v_2$ in V is sent via the Hopf mapping to the quaternion

$$\bar{v}_1 i v_1 + \bar{v}_1 i v_2 + \bar{v}_2 i v_1 + \bar{v}_1 i v_1 = 2\bar{v}_1 i v_2.$$

As a vector in $\mathbb{IH}$, it is linearly independent of the vector $\bar{v}_1 i v_1$, which follows from $\bar{v}_1 i \neq 0$ and the linear independency of v_1 and v_2.

As a consequence, the image of V is the plane passing through the origin, linearly spanned by $\bar{v}_1 i v_1$ and $\bar{v}_1 i v_2$.

On the other hand, for any unit quaternion $w \in \mathbb{IH}$, there exists a S^1-family of unit vectors $\{e^{i\theta} v\}$ in $\mathbb{H}$ whose image under the Hopf map is w. Take a plane W in $\mathbb{IH}$ passing through the origin spanned by two linearly independent unit vectors w_1 and w_2. We can choose the pre-images of v_1 and v_2 in $\mathbb{H}$ of w_1 and w_2 to be such that $BL(v_1, v_2) = 0$. Indeed, for $\bar{v}_1 i v_2 = z_0 + z_1 i + z_2 j + z_4 k$, we have

$$Re(e^{i\theta} \bar{v}_1 i v_2) = z_0 \cos\theta - z_1 \sin\theta,$$

thus we can take $e^{i\theta_1} v_1$ such that $z_0 \cos\theta_1 - z_1 \sin\theta_1 = 0$ in the place of v_1. Thus we get the family of Levi-Civita planes $\operatorname{span}\{e^{i\theta} v_1, e^{i\theta} v_2\}, \theta \in \mathbb{R}/2\pi\mathbb{Z}$ which are sent to W. □

Proposition 5. *There exists an identification with $\mathbb{C}$ of a Levi-Civita plane V together with its image under the Hopf mapping, such that under this identification, the restriction of the*

K.S. mapping to T^*V is given by

$$T^*\mathbb{C} \to T^*\mathbb{C}, \quad (z, w) \longmapsto (z^2, \frac{z}{2|z|^2} \cdot w)$$

which is the Levi-Civita transformation.

Proof. Let v_1 and v_2 be orthogonal unit vectors in V, which allows us to identify V with $\mathbb{C}$. We write $z = av_1 + bv_2$ and $w = cv_1 + dv_2$. Then K.S. sends (z, w) into

$$((a^2 - b^2)\bar{v}_1 iv_1 + 2ab\bar{v}_1 iv_2, \frac{(ac - bd)\bar{v}_1 iv_1 + (ad + bc)\bar{v}_1 iv_2}{2(a^2 + b^2)}).$$

From the orthogonality of v_1 and v_2, we obtain

$$\langle \bar{v}_1 iv_1, \bar{v}_1 iv_2 \rangle = \frac{\bar{v}_1 v_2 + \bar{v}_2 v_1}{2} = \langle v_1, v_2 \rangle = 0.$$

Hence we just need to identify $\bar{v}_1 iv_1$ and $\bar{v}_1 iv_2$ with the standard orthogonal basis of $\mathbb{C}$. The conclusion follows after both V and its image have been identified to $\mathbb{C}$. □

3 Application to integrable Hooke and Kepler billiards

We extend the correspondence shown above to the corresponding billiard systems. This generalizes the correspondence of Hooke and Kepler billiards in the plane [16], [22] to the spatial (Kepler) case.

A centered quadric in $\mathbb{H} \cong \mathbb{R}^4$ is called S^1-invariant, if it is invariant under the S^1-action

$$S^1 \curvearrowright \mathbb{H}, \quad \theta \cdot z \longmapsto \exp(i\theta)z.$$

Equivalently, these are quadrics which are pre-images of subsets in $\mathbb{IH}$ under the Hopf mapping.

A centered quadric in $\mathbb{H}$ is called non-singular if it does not contain the origin.

For an unbounded non-singular centered quadric in $\mathbb{R}^4$ given by

$$F(z_0, z_1, z_2, z_3) = 1$$

where F is a quadratic homogeneous function of $z = (z_0, z_1, z_2, z_3) \in \mathbb{H}$, we define its dual quadric by

$$-F(z_0, z_1, z_2, z_3) = 1.$$

In normal form, for the quadric

$$\sum_{i=0}^{3} \sigma_i \frac{\hat{z}_i^2}{a_i^2} = 1,$$

where $\sigma_i \in \{1, -1\}, a_i \in \mathbb{R}$ and $\{\hat{z}_0\}_{i=0}^3$ is an orthonormal basis in $\mathbb{R}^4$, its dual is

$$\sum_{i=0}^{3} -\sigma_i \frac{\hat{z}_i^2}{a_i^2} = 1.$$

Indeed for a quadric homogeneous function $F(z_0, z_1, z_2, z_3)$ there exists a real symmetric 4×4 matrix A and a real orthogonal matrix Q such that $z^T A z = F$ and $Q^T A Q$ is diagonal, thus its normal form is given by $(Qz)^T AQz = 1$. Clearly, we have $z^T(-A)z = -F$ and $(Qz)^T(-A)Qz = -(Qz)^T AQz$.

Lemma 6. *For an unbounded non-singular centered quadric $\mathcal{E}$ and its dual quadric $\tilde{\mathcal{E}}$ in $\mathbb{H}$, we denote their images in $\mathbb{IH}$ by the Hopf mapping by $\mathcal{F}$ and $\tilde{\mathcal{F}}$ respectively. Let $P \in \mathcal{F}$ be the point of $\mathcal{F}$ with the least distance from $O \in \mathbb{IH}$. Let $\tilde{P} \in \tilde{\mathcal{F}}$ be the point of $\tilde{\mathcal{F}}$ with the least distance from $O \in \mathbb{IH}$. Then the three points $O, P, \tilde{P}$ are collinear.*

Proof. Consider a plane contains $O, P, \tilde{P}$ such that the intersection of $\mathcal{F}$ is unbounded. Then the intersection of $\tilde{\mathcal{F}}$ is unbounded as well. If $\mathcal{E}$ is non-degenerate, then the intersections of $\mathcal{F}$ and $\tilde{\mathcal{F}}$ with this plane are either two centrally symmetric parallel lines or two branches of a focused hyperbola, since they are the images of a pair of dual hyperbolae in the corresponding Levi-Civita plane by the complex square mapping, see [22, Thm. 4]. In the case of parallel lines, these two lines are centrally symmetric, therefore the three points $O, P, \tilde{P}$ are collinear. In the case of hyperbola, the points P and $\tilde{P}$ lie on different branches of the hyperbola, and $P\tilde{P}$ is its major axis which necessarily contains O. When the quadric $\mathcal{E}$ is degenerate, we may have a parabola as an intersection of $\mathcal{F}$ with the plane as well. A parabola is obtained as the image of a line by the complex square mapping [22], and the dual line is sent to the same parabola. In this case, we have $P = \tilde{P}$. □

Proposition 7. *The image of any S^1-invariant, non-singular, centered quadric in $\mathbb{H}$ under the Hopf mapping is either a plane, or a centered sphere, or a spheroid, or a sheet of a two-sheeted circular hyperboloid, or a paraboloid in $\mathbb{IH}$, with always a focus at the origin $O \in \mathbb{IH}$ in the latter three cases. These surfaces correspond precisely to those obtained by revolving a Kepler orbit (a conic with focus at O) about its principal axis.*

Proof. We take an S^1-invariant, non-singular, centered quadric $\mathcal{E}$ in $\mathbb{H}$ and denote its image in $\mathbb{IH}$ by $\mathcal{F}$. The quadrics $\mathcal{E}$ and $\mathcal{F}$ are bounded away from the origin O. We intersect $\mathcal{F}$ with a plane trough $O \in \mathbb{IH}$. By the above theory of Levi-Civita planes, this plane is the image of an S^1-family of Levi-Civita planes on each of them the Hopf mapping restricts to the complex square mapping. The intersection of any of these Levi-Civita planes with the centered quadric in $\mathbb{H}$ is a centered conic section. The image of this centered conic section is thus a branch of a conic section in the plane through the origin O in $\mathbb{IH}$. In case that this branch is neither a line nor a circle, then O is a focus of it [22].

We first assume that $\mathcal{E}$ is bounded in $\mathbb{H}$. Then its image $\mathcal{F}$ is also bounded in $\mathbb{IH}$. If all points from $\mathcal{F}$ have the same distance to the origin, then $\mathcal{F}$ is a centered sphere in $\mathbb{IH}$. Otherwise, there exist a point P_1 with least distance, and another distinct point P_2 with most distance from O. We consider the line passing through these two points and take a plane in $\mathbb{IH}$ containing both this line and the origin. By the above discussion on Levi-Civita planes, the intersection of this plane with the image $\mathcal{F}$ is an ellipse focused at O. Consequently the indicated line passes through the origin, since for an ellipse this line is the major axis and passes through the foci. So the distance $|P_1P_2|$ is the major axis length of this ellipse.

We consider the family of planes passing through this line. If we take such a plane close to the plane we first took, then by continuity, the intersection of $\mathcal{F}$ on this plane is again an ellipse focused at O and the points P_1 and P_2 lie on the ellipse as pericenter and as apocenter respectively. Thus the ellipses obtained as intersection of $\mathcal{F}$ on nearby planes from the family are related by a rotation around the line P_1P_2. Consequently $\mathcal{E}$ is a spheroid with the line P_1P_2 as the symmetric axis.

This argument can be refined to the following local rigidity for (eccentric) ellipses, without assuming that $\mathcal{F}$ is bounded: Consider the line P_1O and a plane through this line such that the intersection of $\mathcal{F}$ with it is an ellipse. Then the intersection of $\mathcal{F}$ with nearby planes through P_1O are also ellipses, and these ellipses are obtained from each other by rotations along the axis P_1O. Indeed all these ellipses need to intersect P_1O at the same point P_2, which necessarily is the apocenter for all of them. This implies this local rigidity for ellipses.

Now we consider the case that $\mathcal{E}$ is not bounded, thus $\mathcal{F}$ is not bounded as well. We take a point $P_1 \in \mathcal{F}$ which has the least distance from O. Since the centered quadric $\mathcal{E}$ is not given by a positive-definite quadric form, its dual quadric $\tilde{\mathcal{E}}$ is non-empty in $\mathbb{H}$. The image in $\mathbb{IH}$ of the dual $\tilde{\mathcal{E}}$ is $\tilde{\mathcal{F}}$. We take the point $\tilde{P}_1 \in \tilde{\mathcal{F}}$ which has the least distance from $O \in \mathbb{IH}$. From Lemma 6, the three points $O, P_1, \tilde{P}_1$ lie on the same line. We consider the family of planes passing through this line. Since $\mathcal{F}$ is unbounded, there exists a plane in this family which has unbounded intersection with $\mathcal{F}$. Thus the intersection of $\mathcal{F} \cup \tilde{\mathcal{F}}$ with this plane is either a pair of two centrally symmetric parallel lines, a pair of branches of a hyperbola with its focus at O, or a parabola with its focus at O.

In the case of a hyperbola, note that we have the local rigidity just as in the elliptic case: In a nearby plane from this family, the intersection of $\mathcal{F} \cup \tilde{\mathcal{F}}$ is again a hyperbola focused at O, with P_1 and P_2 as vertices at each branch. We conclude that $\mathcal{F}$ is a branch of a circular two-sheeted hyperboloid with a focus at O.

In the case of parallel lines, this local rigidity implies that $\mathcal{F}$ intersects nearby planes in lines with P_1 being the closed point from these lines to O. We conclude $\mathcal{F}$ is a plane perpendicular to the line OP_1.

The only left case is when the intersection of $\mathcal{F}$ with a plane containing OP_1 is a parabola. This happens when the original quadric $\mathcal{E}$ is unbounded and degenerate. From the local rigidity of ellipses and hyperbolae, we conclude that if the intersection with a plane passing through OP_1 is a parabola, then the intersections of $\mathcal{F}$ with nearby planes passing through OP_1, we again obtain parabolae. These parabolae are focused at O and have P_1 as the vertex. Thus, $\mathcal{F}$ intersects the nearby planes from this family in parabolae with the focus and the vertex in common. Thus in this case the image $\mathcal{F}$ is a paraboloid with a focus at O.

□

Corollary 8. *Any combination of confocal $\mathcal{S}^1$-invariant centered spheroids or two-sheeted circular hyperboloids in $\mathbb{H}$ is sent to a combination of confocal spheroid or a sheet of a two-sheeted circular hyperboloid.*

Proof. This follows from the fact that any confocal family of conic sections on a plane is sent to a confocal family of conic sections by the complex square mapping ([22, Thm. 4]) and the rotational symmetry of the images of centered quadrics with respect to the symmetry axis shown in Proposition 7. □

Proposition 9. *If a centered quadric in $\mathbb{H}$ is invariant under the $\mathcal{S}^1$-action on $\mathbb{H}$ given by*

$$\theta \cdot z \longmapsto \exp(i\theta) z, \quad z \in \mathbb{H}, \tag{5}$$

then it is a centered quadric given in the non-degenerate case by the normal form equation

$$\frac{u_1^2}{A^2} \pm \frac{u_2^2}{B^2} + \frac{u_3^2}{A^2} \pm \frac{u_4^2}{B^2} = 1, \quad A, B > 0 \tag{6}$$

or in the degenerate case by the normal form equation

$$\frac{u_1^2}{A^2} + \frac{u_3^2}{A^2} = 1, \quad A > 0. \tag{7}$$

The image under the Hopf mapping $z \longmapsto Q = \bar{z} i \cdot z$ of such a centered quadric is a spheroid/a sheet of a circular hyperboloid in the non-degenerate case, including the sphere and plane as degeneracies, and a circular paraboloid in the degenerate case.

Proof. By Proposition 7, the image of an S^1-invariant centered quadric is either a spheroid, or a sheet of a two-sheeted circular hyperboloid, or a paraboloid, all with a focus at the origin, or otherwise a centered sphere or a plane. We shall only discuss the case that this image is a spheroid. The other cases are similar.

The proof is computational. Up to normalization, a spheroid in $\mathbb{IH}$ focused at the origin is given by an equation of the form

$$\frac{q_1^2 - \sqrt{C^2 - D^2}}{C^2} + \frac{q_2^2}{D^2} + \frac{q_3^2}{D^2} - 1 = 0, \quad C > D > 0.$$

The mapping $z \longmapsto Q = \bar{z}i \cdot z$ pulls this equation back to

$$G_1 \cdot G_2 = 0$$

where the factors are

$$G_1 := Cz_1^2 + Cz_2^2 + Cz_3^2 + Cz_4^2 - 2\sqrt{C^2 - D^2} z_1 z_3 - 2\sqrt{C^2 - D^2} z_2 z_4 - D^2$$

and

$$G_2 := Cz_1^2 + Cz_2^2 + Cz_3^2 + Cz_4^2 + 2\sqrt{C^2 - D^2} z_1 z_3 + 2\sqrt{C^2 - D^2} z_2 z_4 + D^2.$$

It is readily seen that the equation $G_2 = 0$ does not admit any real solutions.

In the rotated coordinates (u_1, u_2, u_3, u_4) defined as

$$z_1 = \frac{u_1 + u_2}{\sqrt{2}}, z_2 = \frac{u_3 + u_4}{\sqrt{2}}, z_3 = \frac{u_1 - u_2}{\sqrt{2}}, z_4 = \frac{u_3 - u_4}{\sqrt{2}},$$

we write

$$G_1 = (C - \sqrt{C^2 - D^2})u_1^2 + (C + \sqrt{C^2 - D^2})u_2^2 + (C - \sqrt{C^2 - D^2})u_3^2 + (C + \sqrt{C^2 - D^2})u_4^2 - D^2 = 0$$

and thus by a further normalization we get the desired form (6).

□

In [16], it is noticed that conformal transformations between mechanical systems preserves billiard trajectories. A generalization of this observation to our current situation is the following:

Proposition 10. *Let $\mathcal{R}$ be an S^1-invariant hypersurface in $\mathbb{H} \setminus O$ and $\tilde{\mathcal{R}} \subset \mathbb{IH}$ its image under the Hopf mapping. Let v_1 be an incoming vector at a point $z \in \mathcal{R}$ such that $(z, v_1) \in \Sigma^1$ with the outgoing vector v_2 after reflection. Then $(z, v_2) \in \Sigma^1$. Assume that the Hopf mapping pushes (v_1, v_2) into $(\tilde{v}_1, \tilde{v}_2)$. Then $\tilde{v}_1$ is reflected to $\tilde{v}_2$ by the reflection at $q = \bar{z}iz$ off $\tilde{\mathcal{R}}$.*

In the opposite direction, if $\tilde{v}_1$ is reflected to $\tilde{v}_2$ by the reflection at q off $\tilde{\mathcal{R}} \subset \mathbb{H} \setminus O$, then for any z such that $q = \bar{z}iz$, there exists based vectors v_1, v_2 at z such that $(z, v_1), (z, v_2) \in \Sigma^1$ which is pushed-forward into $(\tilde{v}_1, \tilde{v}_2)$ by the Hopf mapping, such that v_1 is reflected to v_2 at z off the pre-image $\mathcal{R}$ of $\tilde{\mathcal{R}}$.

Proof. By assumption, we have

$$BL(z, v_1) = 0.$$

Consider the normal vector N_z to $\mathcal{R}$ at z. Since $\mathcal{R}$ is S^1-invariant, we have that N_z is orthogonal to the S^1-symmetric direction, which is given by iz. Consequently, we have

$$BL(z, N_z) = Re(\bar{z}iN_z) = -\langle iz, N_z\rangle = 0.$$

Since BL is linear in its second variable, we conclude that

$$BL(z, v_2) = 0$$

as well.

The second assertion follows as long as we show that the push-forward of N_z is orthogonal to $\tilde{\mathcal{R}}$ at $q = \bar{z}iz$. The push-forward of a vector $v \in \Sigma$ is $2\bar{z}iv$. Thus

$$\langle \bar{z}iv, \bar{z}iN_v\rangle = |z|^2\langle v, N_v\rangle, \tag{8}$$

meaning that the angle between v and N_v is preserved. Applying this for any vector $v \in \Sigma \cap T_z\mathcal{R}$ we conclude that $\bar{z}iN_v$ is orthogonal to $\tilde{\mathcal{R}}$ at q.

For the opposite direction, if $\tilde{v}$ is a vector at $q \neq 0$ and $z \in \mathbb{H} \setminus O$ such that $q = \bar{z}iz$, then the vector v such that $\bar{z}iv = \tilde{v}$ is a vector at z which is pushed-forward to $\tilde{v}$. With this construction we get at each z a pair of vectors $\{v_1, v_2\}$ from the pair of vectors $\{\tilde{v}_1, \tilde{v}_2\}$ at q. There follows directly that

$$(z, v_1), (z, v_2) \in \Sigma^1.$$

Moreover it follows from the angle-preservation relationship (8) that if $\tilde{v}_1$ is reflected to $\tilde{v}_2$, then v_1 is reflected to v_2.

□

As part of the proof, we have shown that if an orbit of the four-dimensional Hooke problem satisfies the bilinear relation, then so is its reflection. Therefore we may say that a billiard orbit satisfies the bilinear relation. As only this type of orbits are related to the spatial Kepler problem, we propose the following definition.

Definition 11. *The subsystem of a four-dimensional Hooke billiard consisting only of orbits satisfying the bilinear relation is called the restricted four-dimensional Hooke billiard.*

Definition 12. *A spatial Kepler billiard and a four-dimensional Hooke billiard are called in correspondence, if the reflection wall of the Hooke problem in $\mathbb{H}$ is the pre-image of the reflection wall of the Kepler problem in $\mathbb{IH}$ by the Hopf map.*

With these definitions we get the following theorem, which generalizes the planar Hooke-Kepler billiard correspondence as has been investigated in [16] and [22].

Theorem 13. *Any billiard orbit of the spatial Kepler billiard is the image of an S^1-family of billiard orbits of the corresponding restricted four-dimensional Hooke billiard. In the opposite direction, the image of any orbit of the restricted four-dimensional Hooke billiard under the Hopf mapping is an orbit of the corresponding spatial Kepler billiard.*

This theorem is not limited to the integrable case and thus may be useful to understand the dynamics of non-integrable four-dimensional Hooke and three-dimensional Kepler billiards.

For the integrable case, we know that a four-dimensional Hooke billiard with a centered quadric reflection wall is integrable [6], [4]. We directly obtain the following result, established in [23] via a completely different approach.

Thoerem A. *Consider a surface of revolution in $\mathbb{R}^3$, by revolving a conic with a focus at the origin about its principal axis. Then reflecting spatial Kepler orbits (attracted to/repelled from the origin) off such a surface of revolution gives rise to an integrable mechanical billiard.*

The first integrals for the three-dimensional integrable Kepler billiards can be obtained from the first integrals of the four-dimensional Hooke billiards. On the other hand the explicit representations of the first integrals are already obtained in [23]. We here recall:

$$
\begin{aligned}
E &= \frac{\dot{x}^2+\dot{y}^2+\dot{z}^2}{2} + \frac{m}{\sqrt{x^2+y^2+z^2}},\\
L_{yz} &= \dot{y}z - \dot{z}y,\\
\tilde{E}_{sph} &= \frac{1}{2}\left((1+a^2)\dot{x}^2+\dot{y}^2+\dot{z}^2+((\sqrt{1+a^2}x+a)\dot{y}-\sqrt{1+a^2}y\dot{x})^2\right)\\
&\quad+\frac{1}{2}\left((y\dot{z}-z\dot{y})^2+(\sqrt{1+a^2}z\dot{x}-(\sqrt{1+a^2}x+a)\dot{z})^2\right)\\
&\quad+\frac{m(1+a^2+a\sqrt{1+a^2}x)}{\sqrt{x^2+y^2+z^2}},
\end{aligned}
$$

where a is the half distance between the two foci.

4 The two-center problem and integrable billiards

In this section, we consider the spatial two-center problem, which describes the motion of a particle in $\mathbb{R}^3$ moving under the gravitational attraction of two fixed centers. In the plane, this system is known to be integrable due to the works of Euler and Lagrange [3], [12]. The system is also integrable in $\mathbb{R}^3$. It is considered as a simplification of the planar or spatial circular restricted three-body problem with the Coriolis force and the centrifugal force ignored.

In [1], Birkhoff designed a way to simultaneously desingularize the two double collisions of the particle with the two centers in the planar problem. This has been subsequently generalized to the spatial problem as first announced in Stiefel and Waldvogel [18]. In [25], Waldvogel explained that the construction is analogous to the observation that on the Riemann sphere, Birkhoff's mapping is conjugate to the complex square mapping via a Möbius transformation. The approach was then subsequently applied to the spatial problem. The use of quaternions was introduced in [26].

The goal of this section is to discuss this transformation in the spatial case with the language of the quaternions and symplectic geometry, with the hope of clarifying the geometry of this transformation even further. Subsequently we apply this transformation to the problem of integrable billiards. The main fact we will use is that for a separable 2-degree of freedom Hamiltonian system of the form:

$$H = \frac{a(x)P_x^2 + b(x)P_y^2}{2} + A(x) + B(y) \tag{9}$$

any coordinate line as a reflection wall results in an integrable mechanical billiard. Indeed one has the independent integrals $\frac{a(x)P_x^2}{2} + A(x)$ and $\frac{b(x)P_y^2}{2} + B(y)$ constant along orbits and under reflections at a coordinate line. The same result can be obtained by considering the spatial two-center problem in spheroidal elliptic coordinates, as this approach leads to the same class of the separated system after reduction by rotations around the axis containing the centers. However, it is worth mentioning that the method used here does not require elliptic coordinates; instead, it utilizes spherical coordinates through Birkhoff-Waldvogel's Transformation.

We first recall Waldvogel's view of Birkhoff transformation of the planar two-center problem from [26]. See also [2] for a discussion on the geometry of this transformation.

Consider the mappings

$$\varphi_1 : \mathbb{C} \cup \{\infty\} \longmapsto \mathbb{C} \cup \{\infty\}, \quad z \longmapsto \alpha = 1 - \frac{2}{1-z},$$

$$L.C. : \mathbb{C} \cup \{\infty\} \longmapsto \mathbb{C} \cup \{\infty\}, \quad \alpha \longmapsto q = \alpha^2,$$

$$\varphi_2 : \mathbb{C} \cup \{\infty\} \longmapsto \mathbb{C} \cup \{\infty\}, \quad \longmapsto x = 1 - \frac{2}{1-q}.$$

The mappings φ_1 and φ_2 are Möbius transformations on the Riemann sphere $\mathbb{C} \cup \{\infty\}$. The mapping $L.C.$ is the complex square mapping, branched at $0, \infty$ on the Riemann sphere.

The composition of these mappings in the natural order gives rise to

$$\varphi_2 \circ L.C. \circ \varphi_1 : \mathbb{C} \cup \{\infty\} \longmapsto \mathbb{C} \cup \{\infty\}, \quad z \longmapsto x = \frac{z+z^{-1}}{2}.$$

This is Birkhoff's transformation, used to simultaneously regularize both double collisions with two Kepler centers placed at $-1, 1 \in \mathbb{C}$.

This suggests the following construction for the spatial two-center problem.

We define the *base Birkhoff-Waldvogel mapping* as the composition

$$
\begin{aligned}
\phi_2 \circ \mathrm{Hopf} \circ \phi_1 \, : \mathbb{H} \cup \{\infty\} &\longmapsto \mathbb{IH} \cup \{\infty\}, \\
z &\longmapsto x = i - 4\|z-i\|^4 \left((z - \bar{z} - 2i)\|z-i\|^2 + 2(z-i)i(\bar{z}+i) \right)^{-1}
\end{aligned}
$$

where

$$
\begin{aligned}
\phi_1 \, : \mathbb{H} \cup \{\infty\} &\longmapsto \mathbb{H} \cup \{\infty\}, \\
z &\longmapsto \alpha = i - \frac{2}{z-i},
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Hopf} \, : \mathbb{H} \cup \{\infty\} &\longmapsto \mathbb{IH} \cup \{\infty\} \\
\alpha &\longmapsto q = \bar{\alpha} i \alpha
\end{aligned}
$$

and

$$
\begin{aligned}
\phi_2 \, : \mathbb{IH} \cup \{\infty\} &\to \mathbb{IH} \cup \{\infty\}, \\
q &\longmapsto x = i - \frac{2}{q-i}.
\end{aligned}
$$

In coordinates, we have

$$
\begin{aligned}
x_1 &= \frac{1}{2}\left(z_1 + \frac{z_1(z_0^2+1)}{z_1^2+z_2^2+z_3^2} \right) \\
x_2 &= \frac{1}{2}\left(z_2 + \frac{z_2(z_0^2-1)+2z_0 z_3}{z_1^2+z_2^2+z_3^2} \right) \\
x_3 &= \frac{1}{2}\left(z_3 + \frac{z_3(z_0^2-1)-2z_0 z_2}{z_1^2+z_2^2+z_3^2} \right).
\end{aligned} \tag{10}
$$

By restriction and properly lifting the mappings to the cotangent bundles, we get the *unrestricted Birkhoff-Waldvogel mapping*

$$
\widetilde{B.W.} := \Phi_2 \circ K.S. \circ \Phi_1 \, : (\mathbb{H} \setminus \{i, -i\}) \times \mathbb{H} \to (\mathbb{H} \setminus \{i, -i\}) \times \mathbb{H}, (z, w) \longmapsto (x, y)
$$

where

$$
\begin{aligned}
\Phi_1 \, : (\mathbb{H} \setminus \{i, -i\}) \times \mathbb{H} &\to (\mathbb{H} \setminus \{0, i\}) \times \mathbb{H}, \\
(z, w) &\longmapsto \left(\alpha = i - \frac{2}{z-i}, \beta = \frac{\overline{(z-i)} w \overline{(z-i)}}{2} \right),
\end{aligned}
$$

$$\begin{aligned} K.S. :& (\mathbb{H} \setminus \{0, i\}) \times \mathbb{H} \to (\mathbb{H} \setminus \{0, i\}) \times \mathbb{H} \\ & (\alpha, \beta) \longmapsto \left(q = \bar{\alpha} i \alpha, p = \frac{\bar{\alpha} i \beta}{2|\alpha|^2} \right), \end{aligned}$$

$$\begin{aligned} \Phi_2 :& (\mathbb{H} \setminus \{0, i\}) \times \mathbb{H} \to (\mathbb{H} \setminus \{i, -i\}) \times \mathbb{H} \\ & (q, p) \longmapsto \left(x = i - \frac{2}{q - i}, y = \frac{\overline{(q-i)} p \overline{(q-i)}}{2} \right). \end{aligned}$$

Explicitly, the unrestricted Birkhoff-Waldvogel mapping $\widetilde{B.W.}$ is given by $(z, w) \longmapsto (x, y)$ with

$$\begin{aligned} x &= i - |z-i|^2 (2|z-i|^{-2}(zi\bar{z} + \bar{z} - z + i) - \bar{z} + z - 2i)^{-1} \\ y &= \frac{1}{|i - 2(z-i)^{-1}|^2} \\ &\quad \times ((z-i)^{-1}(\bar{z}+i)^{-1} i(\bar{z}+i) - i(\bar{z}+i)^{-1} + (z-i)^{-1} i + 2(z-i)^{-1}(\bar{z}+i)^{-1}) \\ &\quad \times w(1 - (\bar{z}+i)(z-i)^{-1} + 2i(z-i)^{-1}). \end{aligned}$$

The mappings Φ_1, Φ_2 are constructed in a way that the transformations on positions are natural generalizations of ϕ_1, ϕ_2, while the transformations on momenta are obtained as contragradients. The mapping $K.S.$ is the usual Kustaanheimo-Stiefel transformation.

In $(\mathbb{H} \setminus \{i\}) \times \mathbb{H}$ we define the subsets

$$\hat{\Lambda} := \{(z, w) \in (\mathbb{H} \setminus (\mathbb{R} \cup \{i\} \cup \{-i\})) \times \mathbb{H} \mid Re((\bar{z} - i) w (\bar{z} + i)) = 0\}$$

and

$$\hat{\Sigma} := \{(\alpha, \beta) \in (\mathbb{H} \setminus (\{e^{i\theta}\} \cup \{0\})) \times \mathbb{H} \mid BL(\alpha, \beta) = Re(\bar{\alpha} i \beta) = 0\}.$$

Then we have the following:

Lemma 14. *The image of the mapping Φ_1 with the restricted domain $\hat{\Lambda}$ is $\hat{\Sigma}$. Additionally, the image of the mapping Φ_2 restricted to $(\mathbb{IH} \setminus \{i, 0\}) \times \mathbb{IH}$ is $(\mathbb{IH} \setminus \{i, -i\}) \times \mathbb{IH}$.*

To show this we first show

Lemma 15. $\phi_1^{-1}(\{e^{i\theta}\}) = \mathbb{R}$.

Since $K.S.(\{e^{i\theta}\}) = i$ and $\phi_2(i) = \infty$, this shows in particular that $\mathbb{R} \subset \mathbb{H}$ represents physical infinity of the physical space $\mathbb{IH}$.

Proof. The pre-image of $\alpha = e^{i\theta}$ is

$$\begin{aligned} z = i - \frac{2}{\alpha - i} &= i - \frac{2}{\cos\theta - i(1-\sin\theta)} \\ &= i - \frac{2(\cos\theta + i(1-\sin\theta))}{(\cos\theta - i(1-\sin\theta))(\cos\theta + i(1-\sin\theta))} \\ &= i - \frac{\cos\theta + i(1-\sin\theta)}{1-\sin\theta} \\ &= \frac{\cos\theta}{\sin\theta - 1}. \end{aligned}$$

We thus have

$$\left\{ z = \frac{\cos\theta}{\sin\theta - 1} \mid \theta \in \mathbb{R}/2\pi\mathbb{Z} \right\} = \mathbb{R}.$$

□

Proof. (of Lemma 15) The image $\Phi_1(\hat{\Lambda})$ is contained in $\hat{\Sigma}$, since

$$\begin{aligned} Re(\bar{\alpha} i \beta) = 0 &\Leftrightarrow Re\left(-i - \frac{2}{\bar{z}+i}\right) i \left(\frac{(\bar{z}+i)w(\bar{z}+i)}{2}\right) = 0 \\ &\Leftrightarrow Re((1 - 2(\bar{z}+i)^{-1} i)(\bar{z}+i)w(\bar{z}+i)) = 0 \\ &\Leftrightarrow Re((\bar{z}+i)^{-1}(\bar{z}+i-2i)(\bar{z}+i)w(\bar{z}+i)) = 0 \\ &\Leftrightarrow Re((\bar{z}-i)w(\bar{z}+i)) = 0. \end{aligned} \tag{11}$$

On the other hand, for any $(\alpha, \beta) \in \hat{\Sigma}$, its pre-image $(z, w) \in \hat{\Lambda}$ by Φ_1 is given by the formulas

$$z = \frac{2}{i - \alpha} + i$$

and

$$w = 2(\bar{z}+i)^{-1}\beta(\bar{z}+i)^{-1} = 2\alpha\beta\alpha.$$

Thus, the first part of the lemma follows.

For $(q, p) \in (\mathbb{IH} \setminus \{i\}) \times \mathbb{IH}$, the conjugation of its image (x, y) by Φ_2 is obtained as

$$\bar{x} = -i - \frac{2}{-q+i} = -(i - \frac{2}{q-i}) = -x$$

and

$$\bar{y} = -\frac{(q-i)p(q-i)}{2} = -\frac{(-q+i)p(-q+i)}{2} = -\frac{(\overline{q-i})p(\overline{q-i})}{2} = -y,$$

thus $(x, y) \in (\mathbb{IH} \setminus \{i\}) \times \mathbb{IH}$.

On the other hand, for any $(x, y) \in (\mathbb{IH} \setminus \{i\}) \times \mathbb{IH}$, the conjugate of its pre-image (p, q) is obtained as

$$\bar{q} = \frac{2}{-i + x} + i = -\left(\frac{2}{i - x} + i\right) = -q$$

and

$$\bar{p} = -2xyx = -p,$$

thus the pre-image (p, q) belongs again to $(\mathbb{IH} \setminus \{i\}) \times \mathbb{IH}$.

$\square$

Lemma 16. *The image of the $K.S$ mapping restricted to $\hat{\Sigma}$ is $(\mathbb{IH} \setminus \{i, 0\}) \times \mathbb{IH}$.*

Proof. The image $K.S.(\hat{\Sigma})$ is included in $\mathbb{IH} \times \mathbb{IH}$ since $\bar{\alpha} i \alpha \in \mathbb{IH}$ for any $\alpha \in \mathbb{H}$ and

$$BL(\alpha, \beta) = 0 \Leftrightarrow Re(\bar{\alpha} i \beta) = Re(p) = 0.$$

On the other hand, for any $(q, p) \in (\mathbb{IH} \setminus \{i\}) \times \mathbb{IH}$, we can take $(\alpha, \beta) \in \hat{\Sigma}$ such that $K.S.(\alpha, \beta) = (q, p)$. Indeed,for any $(q, p) \in \mathbb{IH} \times \mathbb{IH}$, there exists an S^1-family $\{(e^{i\theta_1}\alpha, e^{i\theta_1}\beta)\}$ satisfying $BL(\alpha, \beta) = 0$. $\square$

From these lemmas, we obtain the following proposition:

Proposition 17. *The* restricted Birkhoff-Waldvogel mapping

$$B.W. : \hat{\Lambda} \to (\mathbb{IH} \setminus \{i, -i\}) \times \mathbb{IH} \quad (z, w) \longmapsto (x, y),$$

where

$$x = i - |z - i|^2 (2|z - i|^{-2}(zi\bar{z} + \bar{z} - z + i) - \bar{z} + z - 2i)^{-1}$$
$$y = \frac{1}{|i - 2(z - i)^{-1}|^2}((z - i)^{-1}(\bar{z} + i)^{-1} i(\bar{z} + i) - i(\bar{z} + i)^{-1} + (z - i)^{-1} i + 2(z - i)^{-1}(\bar{z} + i)^{-1})$$
$$\times w(1 - (\bar{z} + i)(z - i)^{-1} + 2i(z - i)^{-1})$$

is surjective.

The following proposition describes the symplectic property of the restricted Birkhoff-Waldvogel mapping:

Proposition 18. $B.W.^*(Re(d\bar{y} \wedge dx)) = Re(d\bar{w} \wedge dz)|_{\hat{\Lambda}}$.

Proof. We compute the 1-form:

$$\begin{aligned}\Phi_1^*(Re(\bar{\beta}d\alpha)) &= Re\left(\frac{(z-i)\bar{w}(z-i)}{2} \cdot (-2d(z-i)^{-1})\right)\\ &= Re\left(\frac{(z-i)\bar{w}(z-i)}{2} \cdot 2(z-i)^{-1}(d(z-i))(z-i)^{-1}\right)\\ &= Re(\bar{w}dz)\end{aligned}$$

Similarly, we get

$$\Phi_2^*(Re(\bar{y}dx)) = Re(\bar{p}dq).$$

We now recall the fact

$$K.S.|_{\Sigma}^*(Re(\bar{p}dq)) = Re(\bar{\beta}d\alpha).$$

Since $\hat{\Sigma} \subset \Sigma$, we have

$$K.S.|_{\hat{\Sigma}}^*(Re(\bar{p}dq)) = Re(\bar{\beta}d\alpha).$$

By combining these facts, we obtain

$$B.W.^*(Re(\bar{y}dx)) = Re(\bar{w}dz).$$

$\square$

We now apply this mapping to the two center problem in $\mathbb{R}^3 \cong \mathbb{IH}$, with the two centers at $\pm i \in \mathbb{IH}$. We start with the shifted-Hamiltonian of the two-center problem

$$H - f = \frac{|y|^2}{2} + \frac{m_1}{|x-i|} + \frac{m_2}{|x+i|} - f$$

and consider its 0-energy hypersurface. By multiplying the above equation by $|x-i||x+i|$, we obtain

$$|x-i||x+i|(H-f) = \frac{|y|^2|x-i||x+i|}{2} + m_1|x+i| + m_2|x-i| - f|x-i||x+i|.$$

With the following identities

$$B.W.^*(|x-i|) = \frac{|z-i|^2}{|\bar{z}-z|}$$
$$B.W.^*(|x+i|) = \frac{|z+i|^2}{|\bar{z}-z|}$$
$$B.W.^*(|y|^2) = \frac{|\bar{z}-z|^4|w|^2}{4|z-i|^2|z+i|^2}$$

we obtain

$$\tilde{K} = \frac{|w|^2|\bar{z}-z|^2}{8} + m_1\frac{|z+i|^2}{|\bar{z}-z|} + m_2\frac{|z-i|^2}{|\bar{z}-z|} - f\frac{|z-i|^2|z+i|^2}{|\bar{z}-z|^2} = 0, \tag{12}$$

which can be put in the standard form of a natural mechanical system in the plane by a further multiplication of $|\bar{z}-z|^{-2}$: In this way we get

$$K := \frac{|w|^2}{8} + m_1\frac{|z+i|^2}{|\bar{z}-z|^3} + m_2\frac{|z-i|^2}{|\bar{z}-z|^3} - f\frac{|z-i|^2|z+i|^2}{|\bar{z}-z|^4} = 0. \tag{13}$$

Note that the Hamiltonian (13) is regular at the physical double collisions $\{z = \pm i\}$. The physical collisions are therefore regularized. Its singular set $\{z \in \mathbb{R}\}$ corresponds to ∞ of the original system, and is not contained in any finite energy level (Lem. 15).

Proposition 19. *Consider a plane in $\mathbb{H}$ containing the i-axis given by the equation*

$$k_2x_2 + k_3x_3 = 0 \tag{14}$$

with $(k_1, k_2) \in \mathbb{R}^2 \setminus O$. The pre-image of this plane by the B.W. mapping is the family of two-dimensional spheres and planes given by

$$\begin{cases} (\sin\theta z_0 - \cos\theta)^2 + (z_1^2 + z_2^2 + z_3^2)\sin^2\theta = 1 \\ k_2(z_2\cos\theta + z_3\sin\theta) + k_3(z_3\cos\theta - z_2\sin\theta) = 0. \end{cases} \tag{15}$$

For each $\theta \not\equiv 0, \pi\,(mod\ 2\pi)$, Equation (15) describes a two-dimensional sphere as the intersection of a three-dimensional sphere with a hyperplane in $\mathbb{H}$. We call them Birkhoff spheres. *We denote them by $S_{\theta,\kappa}$ respectively.*

For $\theta \equiv 0, \pi \,(mod\ 2\pi)$, Equation (15) describes the plane

$$\begin{cases} z_0 = 0 \\ k_2 z_2 + k_3 z_3 = 0, \end{cases} \tag{16}$$

which we call a Birkhoff plane *and we denote it by π_κ. In both cases, the angle κ is the unique angle which satisfies*

$$\cos\kappa = z_2, \sin\kappa = z_3.$$

Moreover, the mapping $B.W.$ is restricted to the Birkhoff mapping on the cotangent bundle of a Birkhoff plane.

Proof. The pre-image of the plane (14) by the mapping ϕ_2 is a plane given by

$$k_1 q_2 + k_2 q_3 = 0$$

in $\mathbb{IH}$. The pre-image of this plane by the Hopf map is the family of Levi-Civita planes given by

$$\begin{cases} \alpha_0 \cos\theta + \alpha_1 \sin\theta = 0 \\ k_2(\alpha_2\cos\theta + \alpha_3\sin\theta) + k_3(\alpha_3\cos\theta - \alpha_2\sin\theta) = 0 \end{cases} \tag{17}$$

in which $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ is an angle parametrizing the S^1-symmetry of the Hopf mapping. The pre-image of this family of Levi-Civita planes by ϕ_1 is

$$\begin{cases} \cos\theta(-z_0) + \sin\theta\left(z_1 - 1 + (z_0^2 + (z_1-1)^2 + z_2^2 + z_3^2)^2/2\right) = 0, \\ k_2(\cos\theta z_2 + \sin\theta z_3) + k_3(\cos\theta z_3 - \sin\theta z_2) = 0, \end{cases} \tag{18}$$

which is equivalent to Eq. 15. For the last assertion, the restriction of the $B.W.$ mapping to a Birkhoff plane is the composition of planar mappings each of them can be identified with $\varphi_2, L.C., \varphi_1$ respectively. Indeed, the restriction of ϕ_1 to the ij-plane is obtained as

$$\phi_1(z_1 i + z_2 j) = (1 - 2((z_1-1)^2 + z_2)^{-1}(1 - z_1))i + 2((z_1-1)^2 + z_2)^{-1} z_2 j$$

which is equivalent to the Möbius transformation on $\mathbb{C} \cup \{\infty\}$ given by

$$\varphi_1(z_0 + z_1 i) = 1 - 2((z_0-1)^2 + z_1^2)^{-1}(1 - z_0) - 2(z_0-1)^2 + z_1^2)^{-1} z_1 i$$

up to some basis changes. One can generalize this identification to any planes in $\mathbb{IH}$ containing the i-axis by rotating the plane with respect to the i-axis. Analogously, we can identify the restriction of ϕ_2 to a plane in $\mathbb{IH}$ containing the i-axis. Finally, we recall the argument from Proposition 5 and use the equivalence between the restriction of the Hopf mapping to the $ij-$plane in $\mathbb{IH}$ and the complex square mapping. The conclusion follows.

$\square$

It is desirable to relate $S_{\theta,\kappa}$ and π_κ, as they are related by the symmetry of the Birkhoff-Waldvogel mapping. We also would like to introduce natural coordinates to analyze the transformed system. For this purpose, we have the following lemma:

Lemma 20. *Let $z \in \pi_\kappa$ be expressed as*

$$z = (r\cos\psi)\, i + (r\sin\psi\cos\kappa)\, j + (r\sin\psi\sin\kappa)\, k,$$

and $z_\theta \in S_{\theta,\kappa}$ be related to z by the action of the S^1-symmetry of the Birkhoff-Waldvogel mapping by shifting the corresponding angle by θ. Then we have

$$z_\theta = \frac{(1-r^2)\sin\theta + 2r\cos\psi i + 2r\sin\psi\cos(\theta+\kappa)j + 2r\sin\psi\sin(\theta+\kappa)k}{(r^2+1)-(r^2-1)\cos\theta}. \tag{19}$$

Proof. The mapping $\pi_\kappa \to S_{\theta,\kappa}, z \longmapsto z_\theta$ is computed as $z_\theta = \phi_1^{-1}(e^{i\theta}\phi_1(z))$. This leads to the formula above. $\square$

We may thus use (r,ψ,κ,θ) as coordinates for points in $\mathbb{H}\setminus O$ with the help of Eq. (19). The mapping $(r,\psi,\kappa,\theta) \longmapsto z := z_\theta$ is seen to be 2-to-1, as both (r,ψ,κ,θ) and $(r,\psi,\kappa,\theta+\pi)$ is sent to the same point $z \in \mathbb{H}$.

We compute $\tilde{K}$ in Eq. (12) with these coordinates. We denote by $(P_r, P_\psi, P_\kappa, P_\theta)$ the corresponding conjugate momenta. We set $P_\theta = 0$, which is equivalent to the condition $Re((\bar{z}+i)w(\bar{z}-i)) = 0$. This follows from Eq.(11). We then obtain after this restriction the formula

$$\tilde{K} = \frac{r^2 P_r^2}{2} + \frac{P_\psi^2}{2} + \frac{P_\kappa^2}{2\sin^2\psi} - \frac{2r^2 P_\kappa^2}{(r^2-1)^2} \tag{20}$$

$$+ 4f\cos^2\psi + (m_1 - m_2)\cos\psi + \frac{(fr^2 - (m_1+m_2)r/2 + f)(r^2+1)}{r^2}. \tag{21}$$

This can be considered as the reduced system with respect to the S^1-symmetry in the direction of θ.

We have

$$\tilde{K} = \tilde{K}_1 + \tilde{K}_2, \tag{22}$$

with

$$\begin{aligned} \tilde{K}_1(r, P_r, P_\kappa) &= \frac{r^2 P_r^2}{2} - \frac{2r^2 P_\kappa^2}{(r^2-1)^2} + \frac{(fr^2 - (m_1+m_2)r/2 + f)(r^2+1)}{r^2}; \\ \tilde{K}_2(\psi, P_\psi, P_\kappa) &= \frac{P_\psi^2}{2} + \frac{P_\kappa^2}{2\sin^2\psi} + 4f\cos^2\psi + (m_1 - m_2)\cos\psi. \end{aligned} \tag{23}$$

The angle κ does not appear in this formula, reflecting the rotational invariance of the system around the axis of centers in $\mathbb{H}$. We may thus fix $P_\kappa = C$. The further reduced Hamiltonian is

$$\tilde{K}_{red} = \tilde{K}_{red,1} + \tilde{K}_{red,2}, \tag{24}$$

with

$$\begin{aligned} \tilde{K}_{red,1}(r, P_r) &= \frac{r^2 P_r^2}{2} - \frac{2r^2 C^2}{(r^2-1)^2} + \frac{(fr^2 - (m_1+m_2)r/2 + f)(r^2+1)}{r^2}; \\ \tilde{K}_{red,2}(\psi, P_\psi) &= \frac{P_\psi^2}{2} + \frac{C^2}{2\sin^2\psi} + 4f\cos^2\psi + (m_1 - m_2)\cos\psi. \end{aligned} \tag{25}$$

Both $\tilde{K}_{red,1}(r, P_r)$, $\tilde{K}_{red,2}(\psi, P_\psi)$ are 1 degree of freedom systems. The theory of [22] applies. Any finite combination of coordinate lines $\{r = cst.\}$ and $\{\psi = cst.\}$ in the (r, ψ)-plane are integrable reflection walls.

It follows from Eq. (19) that each fibre of the $B.W.$-mapping intersects the subspace $\mathbb{H}$ in two points when $r \neq 1$, and lie completely in this subspace when $r = 1$. In this latter case, only the combination of the angles $\theta+\kappa$ appears in the formula, meaning that in this case the κ-orbit is the same as the θ-orbit. This is reflected in the formula (20) for $\tilde{K}$, which is singular at $\{r=1\}$ if $C \neq 0$. Indeed it is not hard to check that this set corresponds to the i-axis in the physical space $\mathbb{IH}$. This follows from Eq.(10). Otherwise, it is also singular at $\psi = 0, \pi \,(mod\, 2\pi)$, corresponding again to the $i-$axis. With this in mind, we consider the restriction of the system to the set

$$\tilde{D} = \{z = z_1 i + z_2 j + z_3 k \in \mathbb{IH} \backslash \,:\, |z| \neq 0, 1, (z_2, z_3) \neq (0,0)\}$$

with (orthogonal) spherical coordinates (r, ψ, κ), given by Eqs. (22), (23).

Proposition 21. *A mechanical billiard system in $\mathbb{IH}$, defined by the restriction of $\tilde{K}$ and any finite combination of concentric spheres and any cones symmetric around the i-axis with the vertex at the origin is integrable.*

Proof. The spherical coordinates (r, ψ, κ) are orthogonal. At a point of reflection we decompose the velocity as $v_r\,\vec{e}_r + v_\psi\,\vec{e}_\psi + v_\kappa\,\vec{e}_\kappa$. We consider a sphere centered at the origin O or a cone symmetric around the i-axis with a vertex at O as reflection wall. Due to the symmetry of the wall with respect to the i-axis, the κ-component of the velocity, thus the P_κ does not change under reflections. Also, the intersection of the wall and a plane containing the i-axis is a circle centered at the origin or a line passing through the origin, thus the argument in [22, Lemma 3] applies and we see that both P_r^2 and P_ψ^2 are conserved. Therefore, we conclude that $\tilde{K}_1$ and $\tilde{K}_2$ are conserved.

□

By a quadric of class R in $\mathbb{IH}$ we mean a spheroid or a circular hyperboloid of two sheets there-in with foci at the two Kepler centers $\pm i$ following [23]. Restricting the system in $\mathbb{IH}$ to the Birkhoff planes and making use of [22], we obtain that the above mentioned system is equivalent to the two-center billiards in $\mathbb{R}^3$ with any combinations of quadrics of class R as reflection walls.

Theorem 22. *The above-mentioned mechanical billiard system is equivalent to the two-center billiards in $\mathbb{R}^3$ with any combinations of quadrics of class R as reflection walls.*

Thoerem B. *Consider a surface of revolution in $\mathbb{R}^3$ by revolving a conic with foci at the two Kepler centers around the axis joining the centers. Then reflecting orbits of the spatial two-center problem on such a surface of revolution R is an integrable mechanical billiard. Moreover, taking a finite combination of these surfaces does not destroy the integrability of the resulting two-center mechanical billiard systems.*

This provides an alternative way to show the integrability of these two-center billiards, and generalizes [22] to dimension 3.

Remark 23. *The Lagrange problem in $\mathbb{R}^3$ is given by the Hamiltonian*

$$H = \frac{|w|^2}{2} + m_0|z|^2 + \frac{m_1}{|z-i|} + \frac{m_2}{|z+i|},$$

with $m_0, m_1, m_2 \in \mathbb{R}$ as parameters.

The same procedure shows that this system is separable after reduction in the same coordinates as above. Consequently, Thm. 22 also holds with the Lagrange problem as the underlying system, as well as for other similar systems separable after reduction in these coordinates.

We here recall the explicit representation of the first integrals for the integrable Lagrange billiard obtained in [23]:

$$\begin{aligned}
&E = \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{2} + m_0(x^2 + y^2 + z^2) + \frac{m_1}{\sqrt{(x-1) + y^2 + z^2}} + \frac{m_2}{\sqrt{(x+1) + y^2 + z^2}},\\
&L_{yz} = \dot{y}z - \dot{z}y,\\
&E_{sph} = \frac{1}{2}\Big(2\dot{x}^2 + \dot{y}^2 + \dot{z}^2 + ((\sqrt{2}x+1)\dot{y} - \sqrt{2}y\dot{x})^2 + (\sqrt{2}z\dot{x} - (\sqrt{2}x+1)\dot{z})^2 + (y\dot{z} - z\dot{y})^2\Big)\\
&\qquad + m_0(2x^2 + y^2 + z^2) + \frac{m_1(1+\sqrt{2}x)}{\sqrt{(x-1/\sqrt{2})^2 + y^2 + z^2}} + \frac{m_2(1-\sqrt{2}x)}{\sqrt{(x+1/\sqrt{2})^2 + y^2 + z^2}}.
\end{aligned}$$

Setting $m_0 = 0$ yields the first integrals of the two-center problem stated in Theorem B.

Acknowledgement We thank the referees and the editor for their suggestions and corrections. A.T. and L.Z. are supported by DFG ZH 605/1-1, ZH 605/1-2. A.T. is supported in part by DFG FR 2637/4-1. L.Z is also supported in part by the Heisenberg programme (ZH 605/4-1).

References

[1] G. D. Birkhoff, The restricted problem of three bodies, *Rend. Circ. Matem. Palermo*, 39: 265-334, (1915).116

[2] K. Cieliebak, U. Frauenfelder, and L. Zhao, J^+-invariants for planar two-center Stark-Zeeman systems, *Ergodic Theory Dynam. System*, 43(7): 2258-2292, (2023). 117

[3] L. Euler, De motu corporis ad duo centra virium fixa attracti, *Nov. Comm. Acad. Imp. Petropolitanae*, 10: 207-242, (1767). 116

[4] Y. N. Fedorov, An ellipsoidal billiard with a quadratic potential, *Funct. Anal. Appl.*, 35(3): 199-208, (2001). 98, 115

[5] E. Goursat, Les transformations isogonales en mécanique, *C. R. Math. Acad. Sci. Paris*, 108: 446-450, (1887). 98

[6] C. G. J. Jacobi, *Vorlesungen über Dynamik*, Verlag von Georg Reimer, Berlin (1866). 98, 115

[7] H. J. Korsch and J. Lang, A new integrable gravitational billiard, *J. Phys. A: Math. Gen.*, 24: 45-52, (1991). 99

[8] V. V. Kozlov and D. V. Treschev, *Billiards. A genetic introduction to the dynamics of systems with impacts. Translation of Mathematical Monographs, 89*, American Mathematical Society, Providence, (1991). 99

[9] V. V. Kozlov, Some integrable extensions of Jacobi's Problem of geodesics on an ellipsoid, *J. Appl. Math. Mech.*, 59(1): 1-7, (1995). 99

[10] P. Kustaanheimo, Spinor Regularization of Kepler Motion, *Ann. Univ. Turkuens.*, Ser. A, 73, (1964). 98

[11] P. Kustaanheimo and E. Stiefel, Perturbation Theory of Kepler Motion based on Spinor Regularisation, *J. reine angew. Math.*, 218: 204-219, (1965). 98

[12] J.-L. Lagrange, Recherches sur le mouvement d'un corps qui est attiré vers deux centres fixes, Second mémoire, VIII, *Miscellanea Taurinensia*, t. IV(1): 67-121, (1766-1769). 116

[13] T. Levi-Civita, Sur la résolution qualitative du problème restreint des trois corps, *Verhandlungen des dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904*, B. G. Teubner, Leipzig, (1904). 98

[14] T. Levi-Civita, Sur la régularisation du problème des trois corps, *Acta. Math.*, 42: 204-219, (1920). 98, 101

[15] C. MacLaurin, *Treatise of Fluxions. In two books.* Ruddimans, Edinburgh, (1742). 98

[16] A. A. Panov, An elliptic billiard system with Newtonian potential, *Mat. Zametki*, 55(3): 139-140, (1994). 98, 99, 108, 113, 115

[17] H. Poritsky, The Billard Ball Problem on a Table With a Convex Boundary–An Illustrative Dynamical Problem, *Ann. Math.*, 51(2): 446-470, (1950). 98

[18] E. Stiefel and J. Waldvogel, Problème restreint des trois corps. Généralisation de la régularisation de Birkhoff pour le mouvement du mobile dans l'espace à trois dimensions, *C.R. Acad. Sc. Paris*, 260: 805, (1965). 99, 116

[19] E. Stiefel and G. Scheifele, *Linear and Regular Celestial Mechanics*, Springer-Verlag Berlin-Heidelberg, (1971). 98

[20] P. Saha, Interpreting the Kustaanheimo Stiefel transform in gravitational dynamics, *Month. Not. Royal Astron. Soc.*, 400(1): 228-231, (2009). 98

[21] A. Takeuchi and L. Zhao, Projective Integrable Mechanical Billiards, *Nonlinearity*, 37(1): 015011, (2023). 99

[22] A. Takeuchi and L. Zhao, Conformal Transformations and Integrable Mechanical Billiards, *Adv. Math.*, 436: 109411, (2024). 98, 99, 108, 109, 110, 112, 115, 126, 127, 128

[23] A. Takeuchi and L. Zhao. Integrable Mechanical Billiards in Higher-Dimensional Space Forms, *Regul. Chaotic Dyn.*, 29(3): 405-434, (2024). 98, 99, 100, 115, 116, 127, 128

[24] A. P. Veselov, Confocal surfaces and integrable billiards on the sphere and in the Lobachevsky space, *J. Geom. Phys.*, 7(1): 81-107, (1990). 98

[25] J. Waldvogel, Die Verallgemeinerung der Birkhoff-Regularisierung für das räumliche Dreikörper-problem, *Bull. Astron.*, Série 3, Tome II, Fasc. 2, (1967), 295-341, (1965). 99, 116

[26] J. Waldvogel, Quaternions for regularizing celestial mechanics: the right way, *Cel. Mech. Dyn. Astron.*, 102(1): 149-162, (2008). 98, 99, 116, 117

[27] L. Zhao, Kustaanheimo-Stiefel regularization and the quadrupolar conjugacy, *Regul. Chaot. Dyn.*, 20: 19-36, (2008). 98, 99, 100, 103

[28] L. Zhao, Projective Dynamics and an Integrable Boltzmann Billiard Model, *Comm. Cont. Math.*, 24(10): 2150085, (2022). 99

AUTHORS

Airi Takeuchi
University of Augsburg
Universitätsstrasse,14
86159, Augsburg, Germany
email: airi1.takeuchi@uni-a.de

Lei Zhao,
University of Augsburg
Universitätsstrasse,14
86159, Augsburg, Germany
and
School of Mathematical Sciences,
Dalian University of Technology,
No.2 Linggong Road, Dalian, PR China
email: lei.zhao@math.uni-augsburg.de

Arnold Mathematical Journal
Volume 11, Issue 2, 2025

On fields of meromorphic functions on neighborhoods of rational curves

Serge Lvovski

Received 05 Sep 2024; Accepted 09 Feb 2025

Abstract: Suppose that F is a smooth and connected complex surface (not necessarily compact) containing a smooth rational curve with positive self-intersection. We prove that if there exists a non-constant meromorphic function on F, then the field of meromorphic functions on F is isomorphic to the field of rational functions in one or two variables over $\mathbb{C}$.

AMS Classification: 32H99, 32G13

Key words and phrases: Neighborhoods of rational curves, deformations of analytic subspaces, Enriques classification of surfaces

1 Introduction

It is well known (see Proposition 5.1 below for references) that the field of meromorphic functions on a 2-dimensional neighborhood of the Riemann sphere with positive self-intersection is a finitely generated extension of $\mathbb{C}$, of transcendence degree at most 2. In

DOI: 10.56994/ARMJ.011.002.003

recent papers [5, 6, 8] examples of such neighborhoods were constructed for which this transcendence degree assumes all values from 0 through 2 (in particular, examples of non-algebraizable neighborhoods with transcendence degree 2 were found).

Now it seems natural to ask what fields may occur as such fields of meromorphic functions (in the case of transcendence degree 1 or 2, of course). It turns out that the answer to this question is simple and somehow disappointing. To wit, the main results of the paper are as follows.

Proposition 1.1. *Suppose that F is a non-singular connected complex surface and that there exists a curve $C \subset F$, $C \cong \mathbb{P}^1$, such that $(C \cdot C) > 0$. Let $\mathcal{M}$ be the field of meromorphic functions on F.*

If the transcendence degree of $\mathcal{M}$ over $\mathbb{C}$ is at least 2, then $\mathcal{M} \cong \mathbb{C}(T_1, T_2)$ (the field of rational functions).

Proposition 1.2. *Suppose that F is a non-singular connected complex surface and that there exists a curve $C \subset F$, $C \cong \mathbb{P}^1$, such that $(C \cdot C) > 0$. Let $\mathcal{M}$ be the field of meromorphic functions on F.*

If the transcendence degree of $\mathcal{M}$ over $\mathbb{C}$ is 1, then $\mathcal{M} \cong \mathbb{C}(T)$ (the field of rational functions).

Summing up, if F is a smooth and connected complex surface contaning a copy of the Riemann sphere with positive self-intersetion, then the field of meromorphic functions on F is isomorphic to either $\mathbb{C}$ or $\mathbb{C}(T)$ or $\mathbb{C}(T_1, T_2)$.

Thus, the field of meromorphic functions without any additional structure cannot serve as an invariant that would help to classify neighborhoods of rational curves with positive self-intersection.

Proposition 1.2 agrees with the example from [6, Section 3.2].

The proofs of Propositions 1.1 and 1.2 are based on the study of (embedded) deformations of the curve $C \subset F$. Properties of such deformations are well known in the algebraic context; the classical paper [7] implies a complete description of deformations of rational curves on arbitrary smooth complex surfaces, but this paper does not contain a description of deformations of rational curves passing through given points; I prove the necessary facts (Propositions 3.1 and 3.2) in the ad hoc manner, using a result of Savelyev [11].

The paper is organized as follows. In Section 2 we recall, following Douady [4], general facts on deformations of compact analytic subspaces in a given analytic space. In Section 3 we prove some pretty natural results on deformations of smooth rational curves in smooth (and not necessarily compact) complex surfaces; the results of this section do not claim much novelty. In Section 4 we establish some more specific properties of deformations of rational curves on surfaces. Finally, in Section 5 (resp. 6) we prove Proposition 1.1 (resp. 1.2).

Acknowledgements

I am grateful to Ekaterina Amerik for discussions and to the anonymous referee for numerous useful suggestions.

Notation and conventions

All topological terms refer to the classical topology unless specified otherwise. By coherent sheaves we mean analytic coherent sheaves.

If X is a connected complex manifold, then $\mathcal{M}(X)$ is the field of meromorphic functions on X.

If Y is a complex submanifold of a complex manifold X, then the normal bundle to Y in X is denoted by $\mathcal{N}_{X|Y}$.

Our notation for the n-dimensional complex projective space is $\mathbb{P}^n$.

The projectivization $\mathbb{P}(E)$ of a linear space E is the set of lines in E, not of hyperplanes.

If C_1 and C_2 are compact Riemann surfaces embedded in a smooth complex surface F, then their intersection index is denoted by $(C_1 \cdot C_2)$.

If C is a Riemann surface isomorphic to $\mathbb{P}^1$ and $n \in \mathbb{Z}$, then $\mathcal{O}_C(n)$ stands for the line bundle aka invertible sheaf of degree n on C.

By *Veronese curve* $C_d \subset \mathbb{P}^d$ we mean the image of the mapping $\mathbb{P}^1 \to \mathbb{P}^d$ defined by the formula $(z_0 : z_1) \longmapsto (z_0^d : z_0^{d-1}z_1 : \cdots : z_1^d)$.

Analytic spaces are allowed to have nilpotents in their structure sheaves (however, analytic spaces with nilpotents will be acting mostly behind the scenes). If X is an analytic space, then the analytic space obtained from X by quotienting out the nilpotents is denoted by X_{red}.

If X is an analytic space and $x \in X$, then T_xX is the Zariski tangent space to X at x (i.e., $T_xX = (\mathfrak{m}_x/\mathfrak{m}_x^2)^*$, where $\mathfrak{m}_x$ is the maximal ideal of the local ring $\mathcal{O}_{X,x}$).

In the last two sections we use meromorphic mappings (which will be denoted by dashed arrows). For the general definition we refer the reader to [2, page 75] (one caveat: a meromorphic function on a smooth complex manifold X *is not*, in general, a meromorphic mapping from X to $\mathbb{C}$); for our purposes it suffices to keep in mind two facts concerning them. First, if $F\colon X \dashrightarrow Y$ is a meromorphic mapping, where X is a complex manifold, then the indeterminacy locus of F is an analytic subset in X of codimension at least 2. Second, if X is a connected complex manifold and $f_0, \dots, f_n$ are meromorphic functions on X of which not all are identically zero, then the formula $x \longmapsto (f_0(x) : \dots f_n(x))$ defines a meromorphic mapping from X to $\mathbb{P}^n$.

2 Deformations: generalities

In this section we recall (briefly and without proofs) the general theory (see [4] for details). Suppose that F is an analytic space (in the applications we have in mind F will be a smooth complex surface). Then there exists the *Douady space* $\boldsymbol{D}(F)$, which parametrizes all the compact analytic subspaces of F. This means the following.

For any analytic space B, a *family of compact analytic subspaces of* F *with the base* B is a closed analytic subspace $\mathfrak{H} \subset B \times F$ that is proper and flat over B. Now the Douady space $\boldsymbol{D}(F)$ comes equipped with the *universal family* $\mathfrak{H}(F) \subset \boldsymbol{D}(F) \times F$ of subspaces of F over $\boldsymbol{D}(F)$, which satisfies the following property: for any family over an analytic space B there exists a unique morphism $B \to \boldsymbol{D}(F)$ such that the family over B is induced, via this morphism, from the universal family over $\boldsymbol{D}(F)$. Applying this definition to the case in which B is a point (hence, a family over B is just an individual compact analytic subspace of F), one sees that there is a 1–1 correspondence between compact analytic subspaces of F and fibers of the projection $\mathfrak{H}(F) \to \boldsymbol{D}(F)$.

At this point one has to say that the Douady space is not an analytic space: it is a more general object, which Douady calls a Banach analytic space. However, every point $a \in \boldsymbol{D}(F)$ has a neighborhood $\Delta \ni a$ that is isomorphic to an analytic space in the usual sense.

This construction can be generalized as follows. If $\mathcal{E}$ is a coherent analytic sheaf on F, then there exists a Banach analytic space $\boldsymbol{Dou}(\mathcal{E})$ parametrizing coherent subsheaves $\mathcal{S} \subset \mathcal{E}$ such that the quotient $\mathcal{E}/\mathcal{S}$ has compact support. To be more precise, a family of subsheaves of $\mathcal{E}$ with base B is a coherent subsheaf $\Sigma \subset \operatorname{pr}_2^* \mathcal{E}$ on $B \times F$ such that $\operatorname{pr}_2^* \mathcal{E}/\Sigma$ is flat over B and $\operatorname{supp}(\operatorname{pr}_2^* \mathcal{E}/\Sigma)$ is proper over B, and there is a universal family of subsheaves of $\mathcal{E}$ over $\boldsymbol{Dou}(\mathcal{E})$.

The space $\boldsymbol{Dou}(\mathcal{E})$ is also locally isomorphic to an analytic space. If one puts $\mathcal{E} = \mathcal{O}_F$ in this construction, one obtains a canonical isomorphism $\boldsymbol{D}(F) \cong \boldsymbol{Dou}(\mathcal{O}_F)$.

If $a \in \boldsymbol{Dou}(\mathcal{E})$ is a point corresponding to the subsheaf $\mathcal{S} \subset \mathcal{E}$, then one can define the Zariski tangent space $T_a\boldsymbol{Dou}(\mathcal{E})$ to $\boldsymbol{Dou}(\mathcal{E})$ at a as $T_a\Delta$, where $\Delta \subset \boldsymbol{Dou}(\mathcal{E})$ is any neighborhood of a that is isomorphic to an analytic space. This Zariski tangent space is canonically isomorphic to $\operatorname{Hom}(\mathcal{S}, \mathcal{E}/\mathcal{S})$ (see [4, Section 9.1, Remarque 3]).

If a coherent sheaf $\mathcal{E}$ is a subsheaf of a coherent sheaf $\mathcal{F}$ and if $\mathcal{F}/\mathcal{E}$ has compact support, then $\boldsymbol{Dou}(\mathcal{E})$ is naturally embedded in $\boldsymbol{Dou}(\mathcal{F})$ (a subsheaf of $\mathcal{E}$ can be regarded as a subsheaf of $\mathcal{F}$). This embedding induces injective homomorphisms of Zariski tangent spaces. Indeed, let $\operatorname{Spec} \mathbb{C}[\varepsilon]/(\varepsilon^2)$ be the analytic space consisting of one point such that the ring of functions is $\mathbb{C}[\varepsilon]/(\varepsilon^2)$. Then $T_a\boldsymbol{Dou}(\mathcal{E})$, as a set, is canonically bijective to the set of families of subsheaves of $\mathcal{E}$ over the base $\operatorname{Spec}\mathbb{C}[\varepsilon]/(\varepsilon^2)$ (ibid.). If $\operatorname{supp}(\mathcal{F}/\mathcal{E})$ is compact, any family of subsheaves of $\mathcal{E}$ (over an arbitrary base) is automatically a family of subsheaves of $\mathcal{F}$, and different families of subsheaves of $\mathcal{E}$, being different subsheaves of $\operatorname{pr}_2^*\mathcal{E}$, are ipso facto different subsheaves of $\operatorname{pr}_2^*\mathcal{F}$.

In the sequel we will be using the following notation.

Notation 2.1. If F is a complex manifold and $a \in \boldsymbol{D}(F)$, then C_a stands for the analytic subspace of F corresponding to a.

Similarly, if $\mathcal{E}$ is a coherent sheaf on a complex manifold and $a \in \boldsymbol{Dou}(\mathcal{E})$, then $\mathcal{S}_a$ is the subsheaf in $\mathcal{E}$ corresponding to a.

3 Deformations of rational curves

In this section we state and prove two auxiliary results concerning deformations of smooth rational curves on complex surfaces. These results are well known for deformations of curves on which no restrictions are imposed. For example, Proposition 3.1 below follows immediately from the main result of [7], and its algebraic-geometric counterpart (for smooth algebraic surfaces over a field of characteristic zero) follows immediately from the theorem in Lecture 23 of [9]. However, I did not manage to find a suitable reference for deformations of curves passing through given points.

We will be using the general theory from Section 2 in the following setting. F will always be a smooth and connected complex surface, $C \subset F$ will be a complex submanifold isomorphic to $\mathbb{P}^1$ (the Riemann sphere), and we will always assume that the self-intersection index $d = (C \cdot C)$ is non-negative. By $\boldsymbol{D}(F, C)$ we will mean an unspecified open subset of $\boldsymbol{D}(F)$ that contains the point corresponding to $C \subset F$ and is isomorphic to an analytic space. The reader will check that this indeterminacy of definition does not affect the arguments that follow.

Moreover, suppose that $S = \{p_1, \dots, p_m\} \subset C$ is a subset of cardinality $m \le d = (C \cdot C)$. Let $\mathcal{I}_S \subset \mathcal{O}_F$ be the ideal sheaf of the analytic subset $S \subset F$. If $\mathcal{I} \subset \mathcal{I}_S$ is a coherent subsheaf, then it follows from the exact sequence

$$0 \to \mathcal{I}_S/I \to \mathcal{O}_F/\mathcal{I} \to \mathcal{O}_F/\mathcal{I}_S \to 0$$

that $\operatorname{supp}(\mathcal{O}_F/\mathcal{I}) = \operatorname{supp}(\mathcal{I}_S/I) \cup S$, so $\operatorname{supp}(\mathcal{I}_S/I)$ is compact if and only if $\operatorname{supp}(\mathcal{O}_F/\mathcal{I})$ is compact. Hence, the Douady space $\boldsymbol{Dou}(\mathcal{I}_S)$ parametrizes the (ideal sheaves of) compact analytic subspaces of F containing the subset S. An unspecified open subset of $\boldsymbol{Dou}(\mathcal{I}_S)$ containing the point corresponding to C (strictly speaking, to the ideal sheaf of C, which is a subsheaf of $\mathcal{I}_S$) and isomorphic to an analytic space, will be denoted by $\boldsymbol{D}(F, C, S)$. In view of the natural embedding of $\boldsymbol{Dou}(\mathcal{I}_S)$ into $\boldsymbol{Dou}(\mathcal{O}_F) = \boldsymbol{D}(F)$ we will always assume that $\boldsymbol{D}(F, C, S) \subset \boldsymbol{D}(F, C)$.

Extending Notation 2.1, we will denote by $C_a \subset F$ the analytic subspace of F corresponding to the point $a \in \boldsymbol{D}(F, C, S)$.

Let $a \in \boldsymbol{D}(F, C)$ be the point corresponding to $C \subset F$, and let $\mathcal{I}_C \cong \mathcal{O}_F(-C)$ be the ideal sheaf of $C \subset F$. According to the general theory, the Zariski tangent space to $\boldsymbol{D}(F, C)$ at a is

$$T_a\boldsymbol{D}(F, C) = \operatorname{Hom}_{\mathcal{O}_F}(\mathcal{I}_C, \mathcal{O}_F/\mathcal{I}_C) \cong \operatorname{Hom}_{\mathcal{O}_C}(\mathcal{I}_C/\mathcal{I}_C^2, \mathcal{O}_C) = \mathcal{N}_{F|C} \cong \mathcal{O}_C(d) \tag{1}$$

(here and below, Hom refers to the space of global homomorphisms, not to the Hom sheaf). Similarly, taking into account that $\mathcal{I}_C \subset \mathcal{I}_S$ and denoting by $b \in \boldsymbol{D}(F, C, S)$ the point corresponding to C, one has

$$T_b\boldsymbol{D}(F, C, S) \cong \operatorname{Hom}_{\mathcal{O}_F}(\mathcal{I}_C, \mathcal{I}_S/\mathcal{I}_C) \cong \operatorname{Hom}_{\mathcal{O}_C}(\mathcal{I}_C/\mathcal{I}_C^2, \mathcal{I}_S) \cong \mathcal{N}_{F|C} \otimes \mathcal{O}_C(-S) \cong \mathcal{O}_C(d - m).$$

The main results about deformations of $C \subset F$ that we need are as follows.

Proposition 3.1. *Suppose that F is a smooth and connected complex surface, $C \subset F$ is a complex submanifold isomorphic to $\mathbb{P}^1$, and $d = (C \cdot C) \geq 0$. Then there exists a neighborhood $\Delta \ni a$ of the point $a \in \boldsymbol{D}(F, C)$ corresponding to C such that the analytic space Δ is a smooth complex manifold of dimension $d + 1$ and, for any $b \in \Delta$, $C_b \cong \mathbb{P}^1$.*

A similar result, of which Proposition 3.1 is a particular case, holds for $\boldsymbol{D}(F, C, S)$.

Proposition 3.2. *In the above setting, suppose that* $S = \{p_1, \dots, p_m\} \subset C$ *is a subset of cardinality* $m \leq d$*. Then there exists a neighborhood* $\Delta \ni a$ *of the point* $a \in \boldsymbol{D}(F, C, S)$ *corresponding to* C *such that the analytic space* Δ *is a smooth complex manifold of dimension* $d - m + 1$ *and, for any* $b \in \Delta$*,* $C_b \cong \mathbb{P}^1$*.*

We begin with a particular case, which is essentially contained in [11] (and which follows from the main result of [7]).

Lemma 3.3. *Proposition 3.1 holds if* $d = 0$*.*

Proof. Let $a \in \boldsymbol{D}(F, C)$ be the point corresponding to $C \subset F$. Since $d = 0$, one has $\mathcal{N}_{F|C} \cong \mathcal{O}_C$, so it follows from (1) that $\dim T_a\boldsymbol{D}(F, C) = 1$. But, according to the main result of [11], there exist a neighborhood $W \supset C$ and an isomorphism $\varphi \colon W \to D \times C$, where D is the unit disk in $\mathbb{C}$, such that $\varphi(p) = (0, p)$ for any $p \in C$. If one puts

$$\mathfrak{G} = \{(z, x) \in D \times F \colon x \in W, \ \mathrm{pr}_1(\varphi(x)) = z\},$$

then the family $\mathfrak{G}$ induces a morphism $\Phi \colon D \to \boldsymbol{D}(F, C)$ such that $\Phi(0) = a$ (the point corresponding to C) and Φ is 1–1 onto its image. Hence, $\dim_a \boldsymbol{D}(F, C) \geq 1$. Since $\dim T_a\boldsymbol{D}(F, C) = 1$, one concludes that $\boldsymbol{D}(F, C)$ is a smooth 1-dimensional complex manifold in a neighborhood of a. □

To prove Proposition 3.2 in full generality, we will need two simple lemmas.

Lemma 3.4. *If* $p \colon \mathfrak{H} \to D$*, where* D *is the unit disk in* $\mathbb{C}$*, is a proper and flat morphism of analytic spaces, and if the fiber* $p^{-1}(0)$ *is reduced and isomorphic to* $\mathbb{P}^1$*, then there exists an* $\varepsilon \in (0; 1)$ *such that the fiber* $f^{-1}(a)$ *is also reduced and isomorphic to* $\mathbb{P}^1$ *whenever* $|a| < \varepsilon$*.*

Sketch of proof. It is easy to see that there exists an $\varepsilon > 0$ such that $p^{-1}(D_\varepsilon) \to D_\varepsilon$ is a proper submersion of complex manifolds. Hence, topologically it is a locally trivial bundle, so all the fibers are homeomorphic to S^2, whence the result. □

Lemma 3.5. *Suppose that* X *is an analytic space,* $a \in X$*, and* $\dim T_aX = n$*. Then the following two assertions are equivalent.*

(1) *X is a smooth n-dimensional complex manifold in a neighborhood of the point a.*

(2) *There exists a non-empty Zariski open subset $V \subset \mathbb{P}(T_aX)$ such that for any 1-dimensional linear subspace $\ell \subset T_aX$ corresponding to a point of V there exists a smooth 1-dimensional locally closed complex submanifold $Y \subset X$ such that $Y \ni a$ and $T_aY = \ell \subset T_aX$.*

Proof. Only the implication $(2) \Rightarrow (1)$ deserves a proof.

Observe that X is a complex manifold near a if and only if $\dim_a X = n = \dim T_aX$. Furthermore, the question being local, we may and will assume that X is a closed analytic subspace of a polydisc $D \subset \mathbb{C}^N$. Let $\bar{D}$ be the blowup of D at a, and let $\bar{X}$ be the strict transform of X_{red}. If $\sigma\colon \bar{D} \to D$ is the blowdown morphism and $E = \sigma^{-1}(a)$ is the exceptional divisor, then $\bar{X} \cap E$ is a projective submanifold of $E \cong \mathbb{P}^{N-1}$, $\dim \bar{X} \cap E = \dim_a X - 1$, and $\bar{X} \cap E \subset \mathbb{P}(T_aX) \subset E$.

Now if $Y \subset X$ is a locally closed 1-dimensional complex submanifold such that $Y \ni a$ and if $\ell = T_aY \subset T_aX$, then the point of $\mathbb{P}(T_a(X))$ corresponding to ℓ belongs to $\bar{X} \cap E$; thus, it follows from (2) that $\bar{X} \cap E$ contains a non-empty Zariski open subset of $\mathbb{P}(T_aX)$, hence $\bar{X} \cap E = \mathbb{P}(T_aX)$, hence $\dim_a X = n$, and we are done. □

Proof of Proposition 3.2. Choose $d - m$ distinct points $q_1, \dots, q_{d-m} \in C \setminus S$. Let $\bar{F}$ be the blowup of F at the points $p_1, \dots, p_m, q_1, \dots, q_{d-m}$, let $\sigma\colon \bar{F} \to F$ be the corresponding blowdown morphism, and let $\bar{C} \subset \bar{F}$ be the strict transform of C. One has $\bar{C} \cong \mathbb{P}^1$ and $(\bar{C}, \bar{C}) = 0$. Let $\bar{a} \in \boldsymbol{D}(\bar{F}, \bar{C})$ be the point corresponding to $\bar{C}$, and let $a \in \boldsymbol{D}(F, C, S)$ be the point corresponding to C.

Applying Lemma 3.3 to the pair $(\bar{F}, \bar{C})$, one concludes that there exists a family $\bar{\mathfrak{H}}_0 \subset D \times \bar{F}$, where D is the unit disk in the complex plane, such that its fiber over 0 is $\bar{C} \subset \bar{F}$ and, for the induced mapping $\bar{\varphi}\colon D \to \boldsymbol{D}(\bar{F}, \bar{C})$, its derivative $D\bar{\varphi}(0)\colon T_0D \to T_{\bar{a}}\boldsymbol{D}(\bar{F}, \bar{C})$ is non-degenerate.

If we put $\mathfrak{H}_0 = (\mathrm{id} \times \sigma)(\bar{\mathfrak{H}}_0) \subset D \times F$, then $\mathfrak{H}_0$ is a family of analytic subspaces in F containing S; its fiber over 0 is C. Let $\varphi\colon D \to \boldsymbol{D}(F, C, S)$ be the mapping induced by this family.

It is clear that the diagram

$$\begin{tikzcd}
& T_{\bar a}\boldsymbol{D}(\bar F,\bar C) \arrow[r, equal] & H^0(\mathcal{N}_{\bar F|\bar C}) \arrow[dd, "D\sigma"] \\
T_0D \arrow[ur, "D\bar\varphi(0)"] \arrow[urr, "\bar\alpha"'] \arrow[dr, "D\varphi(0)"'] \arrow[drr, "\alpha"] & & \\
& T_a\boldsymbol{D}(F,C,S) \arrow[r, equal] & H^0(\mathcal{N}_{F|C}(-S)),
\end{tikzcd}$$

where the vertical arrow is induced by the natural homomorphism $\mathcal{N}_{\bar F|\bar C} \to \sigma^*\mathcal{N}_{F|C}$, is commutative. It follows from (the proof of) Lemma 3.3 that $\bar\alpha(\partial/\partial z)$, where z is the coordinate on D, is a nowhere vanishing section of $\mathcal{N}_{\bar F|\bar C} \cong \mathcal{O}_{\bar C}$; since the derivative of the mapping σ is non-degenerate outside $\sigma^{-1}\{p_1,\dots,p_m,q_1,\dots,q_{d-m}\}$, the section $\alpha(\partial/\partial z) = D\sigma(\bar\alpha(\partial/\partial z))$ is not identically zero. Hence, φ induces an embedding of a possibly smaller disk $D_\varepsilon \subset D$ in $\boldsymbol{D}(F,C,S)$.

Moreover, since σ maps each of the curves $\sigma^{-1}(p_i)$, $\sigma^{-1}(q_j)$ to a point, and since each of these curves is transverse to $\bar C$, the section $\alpha(\partial/\partial z)$ vanishes at $p_1,\dots,p_m,q_1,\dots,q_{d-m}$, so $\alpha(\partial/\partial z)$ spans the 1-dimensional linear space

$$H^0(\mathcal{N}_{F|C}(-S)(-q_1-\dots-q_{d-m})) \subset H^0(\mathcal{N}_{F|C}(-S)) = T_a\boldsymbol{D}(F,C,S). \tag{2}$$

In the argument that follows we will assume that $d-m \geq 2$, so that the words about Veronese curves in $\mathbb{P}^d$ and $\mathbb{P}^{d-m}$ make sense; we leave it to the reader to modify the wording for the case $d-m=1$.

Keeping the above in mind, identify C with $\mathbb{P}^1$ and $\mathcal{N}_{F|C}$ with $\mathcal{O}_{\mathbb{P}^1}(d) = \mathcal{O}_C(d)$, embed C in $\mathbb{P}^d$ with the complete linear system $|\mathcal{O}_C(d)|$ to obtain a Veronese curve $C_d \subset \mathbb{P}^d$, and project C_d from $\mathbb{P}^d$ to $\mathbb{P}^{d-m}$, the center of projection being the linear span of the images of the points $p_1,\dots,p_m$. The image of this projection will be a Veronese curve $C_{d-m} \subset \mathbb{P}^{d-m}$; denote the resulting isomorphism between C and C_{d-m} by $\varphi\colon C \to C_{d-m}$.

One has

$$\mathbb{P}^{d-m} = \mathbb{P}((H^0(\mathcal{N}_{F|C}(-S))^*);$$

for any $d-m$ distinct points $q_1, \dots, q_{d-m} \in C \backslash S$ the points $\varphi(q_1), \dots, \varphi(q_{d-m})$ span a hyperplane in $\mathbb{P}^{d-m}$, and the linear span $\langle \varphi(q_1), \dots, \varphi(q_{d-m}) \rangle$ is the projectivization of

$$\mathrm{Ann}(H^0(\mathcal{N}_{F|C}(-S)(-q_1 - \cdots - q_{d-m}))) \subset H^0(\mathcal{N}_{F|C}(-S))^*.$$

Now the hyperplanes in $\mathbb{P}^{d-m}$ that are transverse to C_{d-m} form a non-empty Zariski open subset in $(\mathbb{P}^{d-m})^* = \mathbb{P}(H^0(\mathcal{N}_{F|C}(-S)))$ and any such hyperplane intersects C_{d-m} at $d-m$ distinct points that are of the form $\varphi(q_1), \dots, \varphi(q_{d-m})$, where $q_1, \dots, q_{d-m} \in C \setminus S$. Thus, the linear subspaces of the form (2) fill a non-empty Zariski open subset in $\mathbb{P}(T_a \boldsymbol{D}(F, C, S))$ as we vary $q_1, \dots, q_{d-m} \in C \setminus S$, so the hypotheses of Lemma 3.5 are satisfied if one puts $Y = \varphi(D)$, hence the smoothness is established. Now the assertion to the effect that the fibers are isomorphic to $\mathbb{P}^1$ follows from Lemma 3.4. □

4 Good neighborhoods

Suppose that F is a smooth and connected complex surface and $C \subset F$ is a curve that is isomorphic to $\mathbb{P}^1$ and $(C \cdot C) \geq 0$. In the definition below we use Notation 2.1.

Definition 4.1. We will say that an open subset $W \subset F$, $W \supset C$ is a *good neighborhood of* C if there exists a connected open subset $\Delta \subset \boldsymbol{D}(F, C)$ such that $W = \bigcup_{b \in \Delta} C_b$ and each C_b for $b \in \Delta$ is a smooth curve isomorphic to $\mathbb{P}^1$.

Proposition 4.2. *In the above setting, there exists a fundamental system of good neighborhoods of* C.

Proof. If $d = 0$, it follows immediately from Lemma 3.3. Suppose that $d > 0$.

Let $a \in \boldsymbol{D}(F)$ be the point corresponding to C, and let $\Delta \ni a$, $\Delta \subset \boldsymbol{D}(F, C)$ be the neighborhood whose existence is asserted by Proposition 3.1. We denote by $\mathfrak{H} \subset \Delta \times F$ the family of analytic subspaces of F induced by the embedding $\Delta \hookrightarrow \boldsymbol{D}(F, C)$ (informally speaking, $\mathfrak{H} = \{(b, x) \in \Delta \times F : x \in C_b\}$). Since all the C_b's are smooth 1-dimensional complex submanifolds of F and the base Δ is a smooth complex manifold as well, $\mathfrak{H}$ is a smooth complex manifold.

I claim that the projection $q\colon \mathfrak{H} \to F$ is a submersion; once we have established this fact, it will follow that $q(\mathfrak{H}) \subset F$ is a good neighborhood of C.

To check this submersivity, consider an arbitrary point $(b, x) \in \mathfrak{H}$ (i.e., $x \in C_b$); we are to show that the derivative $Dq(b, x)\colon T_{(b,x)}\mathfrak{H} \to T_xF$ is surjective. To that end, pick d distinct point $p_1, \dots, p_d \in C_b \setminus \{x\}$, where $d = (C \cdot C) = (C_b \cdot C_b)$, and put $\{p_1, \dots, p_d\} = S$.

Let $i\colon \boldsymbol{D}(F, C, S) \hookrightarrow \boldsymbol{D}(F, C)$ be the natural embedding, and let $\beta \in \boldsymbol{D}(F, C, S)$ be the point corresponding to the curve C_b (so C_β and C_b are the same curve in F, and $i(\beta) = b$).

Let $\Delta_0 \subset \boldsymbol{D}(F, C_b, S)$, $\Delta_0 \ni \beta$ be a neighborhood whose existence is asserted by Proposition 3.2; we may and will assume that $\Delta_0 \subset i^{-1}(\Delta)$. Finally, let $\mathfrak{H}_0 \subset \Delta_0 \times F$ be the family of analytic subspaces of F (containing S) induced by the inclusion $\Delta_0 \hookrightarrow \boldsymbol{D}(F, C, S)$, and let $q_0\colon \mathfrak{H}_0 \to F$ be the projection. The inclusion $\Delta_0 \subset \boldsymbol{D}(F, C)$ induces an inclusion $\mathfrak{H}_0 \to \mathfrak{H}$, and one has the following obvious commutative diagram:

$$\begin{tikzcd} & & F \\ \mathfrak{H}_0 \arrow[urr, "q_0"] \arrow[r, hook] \arrow[d] & \mathfrak{H} \arrow[ur, "q"'] \arrow[d] & \\ \Delta_0 \arrow[r, hook, "i"] & \Delta & \end{tikzcd} \tag{3}$$

Observe that if $c \in i(\Delta_0) \subset \Delta$ and $y \in C_c \setminus S$ then there exists a unique $\gamma \in \Delta_0$ such that $y \in C_\gamma$. Indeed, if $y \in C_\gamma \cap C_{\gamma_1}$, $\gamma, \gamma_1 \in \Delta_0$, then $C_\gamma \cap C_{\gamma_1} \supset \{y\} \cup S$, whence $(C_\gamma \cdot C_{\gamma_1}) \geqslant d + 1$, which contradicts the fact that

$$(C_\gamma \cdot C_{\gamma_1}) = (C \cdot C) = d.$$

Hence, there exists a neighborhood $\mathfrak{V} \ni (\beta, x)$ in $\mathfrak{H}_0$ such that the restriction of q_0 to $\mathfrak{V}$ is 1–1 onto its image.

Since, according to Proposition 3.2, the Douady space $\boldsymbol{D}(F, C_b, S)$ is 1-dimensional and smooth in a neighborhood of the point β corresponding to $C_b \subset F$, and since a holomorphic mapping of complex manifolds of the same dimension that is 1–1 onto its image is an open embedding, it follows now that $q_0(\mathfrak{V})$ is open in F and $q_0|_{\mathfrak{V}}\colon \mathfrak{V} \to q_0(\mathfrak{V})$

is a biholomorphism. In particular, $Dq_0(\beta, x)\colon T_{(\beta,x)}\mathfrak{H}_0 \to T_xF$ is an isomorphism. Now it follows from the diagram (3) that $Dq(b, x)\colon T_{(b,x)}\mathfrak{H} \to T_xF$ is a surjection.

This proves the submersivity of the projection $q\colon \mathfrak{H} \to F$, so $q(\mathfrak{H}) \subset F$ is a good neighborhood of C, and, for any open and connected $\Delta' \subset \Delta$, $\Delta' \ni a$, the set $q(p^{-1}(\Delta')) \supset C$, where $p\colon \mathfrak{H} \to \Delta$ is the projection, is a good neighborhood of C as well. The neighborhoods $q(p^{-1}(\Delta'))$, for various such Δ', form a fundamental system of good neighborhoods of C. □

Proposition 4.3. *Suppose that $C \subset F$, where F is a smooth complex surface, is a curve that is isomorphic to $\mathbb{P}^1$, and that $(C \cdot C) = d > 0$. If $W \subset F$ is a good neighborhood of C in the sense of Definition 4.1, then for any $x \in W$ there exist two curves $C_1, C_2 \ni x$, $C_1, C_2 \subset W$ such that $C_1 \cong C_2 \cong \mathbb{P}^1$ and the curves C_1 and C_2 are transverse at x.*

Proof. Since W is a good neighborhood, there exists a curve $C_b \ni x$, where $b \in \Delta$. Pick $d-1$ distinct points $p_1, \dots, p_{d-1} \in C_b \setminus \{x\}$ and put $S = \{x, p_1, \dots, p_{d-1}\}$. According to Proposition 3.2, one has $\dim_\beta \boldsymbol{D}(F, C_b, S) = 1$, where β is the point of $\boldsymbol{D}(F, C_b, S)$ corresponding to C_b. Let $\Delta_0 \ni \beta$, $\Delta_0 \subset \boldsymbol{D}(F, C_b, S)$ be the neighborhood whose existence is asserted by Proposition 3.2, and let $\gamma_1, \gamma_2 \in \Delta_0$ be two distinct points. I claim that the curves $C_1 := C_{\gamma_1}$ and $C_2 := C_{\gamma_2}$ are transverse at x. Indeed, if this not the case, then the local intersection index of C_1 and C_2 at x is at least 2, whence

$$(C \cdot C) = (C_1 \cdot C_2) \geq d - 1 + 2 \geq d + 1,$$

contrary to the fact that $(C \cdot C) = d$. This contradiction proves the required transversality. □

5 Transcendence degree 2

In this section we prove Proposition 1.1. We begin with two simple observations.

Proposition 5.1. *If F is a smooth connected complex surface that contains a curve $C \subset \mathbb{P}^1$ such that $(C \cdot C) > 0$, then the field of meromorphic functions $\mathcal{M}(F)$ is finitely generated over $\mathbb{C}$ and $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(F) \leq 2$.*

Proof. Theorem 2.1 from [10] asserts that there exists a pseudoconcave neighborhood $U \supset C$. According to [1, Théorème 5], the field $\mathcal{M}(U)$ is finitely generated over $\mathbb{C}$ and $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(U) \le 2$. Observe that $\mathcal{M}(F)$ embeds in $\mathcal{M}(U)$ as an extension of $\mathbb{C}$; since any sub-extension of a finitely generated extension of fields is also finitely generated and the transcendence degree is additive in towers, we are done. □

Lemma 5.2. *Suppose that F is a connected complex surface such that $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(F) \ge 2$. If there exists a connected open set $U \subset F$ such that $\mathcal{M}(U) \cong \mathbb{C}(T_1, T_2)$, then $\mathcal{M}(F) \cong \mathbb{C}(T_1, T_2)$.*

Proof. Since $\mathcal{M}(F)$ embeds into $\mathcal{M}(U)$, it follows immediately from the two-dimensional Lüroth theorem. □

(Recall that the two-dimensional Lüroth theorem asserts that if $K \subset \mathbb{C}(T_1, T_2)$ is a subfield containing $\mathbb{C}$ and $\operatorname{tr.deg}_{\mathbb{C}} K = 2$, then $K \cong \mathbb{C}(T_1, T_2)$; this fact follows immediately from the existence of a smooth projective model for any finitely generated extension of $\mathbb{C}$ and from Theorem 3.5 in [3, Chapter VI].)

Now we may begin the proof of Proposition 1.1. Thus, let F be a smooth connected complex surface such that $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(F) \ge 2$ and let $C \cong \mathbb{P}^1$ be a curve (one-dimensional complex submanifold) such that $(C \cdot C) = d > 0$. Let $W \subset F$ be a good neighborhood of C in the sense of Definition 4.1. We are to prove that $\mathcal{M}(F) \cong \mathbb{C}(T_1, T_2)$; by virtue of Lemma 5.2 it suffices to prove that $\mathcal{M}(W) \cong \mathbb{C}(T_1, T_2)$.

Since $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(F) \le 2$ by virtue of Proposition 5.1 and $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(F) \ge 2$ by hypothesis, one has $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(F) = 2$. Now $\mathcal{M}(F)$ is isomorphic to a subfield of $\mathcal{M}(W)$, so $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(W) \ge 2$. Since Proposition 5.1 implies that $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(W) \le 2$ and $\mathcal{M}(F)$ is finitely generated over $\mathbb{C}$, one concludes that $\mathcal{M}(W)$ is a finitely generated extension of $\mathbb{C}$, of transcendence degree 2. Hence, $\mathcal{M}(W) = \mathbb{C}(f, g, h)$, where the meromorphic functions f and g are algebraically independent over $\mathbb{C}$ and h is algebraic over $\mathbb{C}(f, g)$ (of course, if one may set $h = 0$, there is nothing to prove). Denote by P an irreducible polynomial in three independent variables F, G, and H such that $P(f, g, h)$ is identically zero.

Now let $Y \subset \mathbb{C}^3$ be the affine algebraic surface that is the zero locus of P, and let $X \subset \mathbb{P}^N$ be a smooth projective model of Y.

Denote by $V \subset W$ the open subset on which each of the meromorphic functions f, g, and h is well defined and consider the holomorphic mapping $\Phi\colon V \to Y$ defined by the formula $x \longmapsto (f(x), g(x), h(x))$. The mapping Φ extends to a meromorphic mapping from W to $\bar{Y} \subset \mathbb{P}^3$, where $\bar{Y}$ is the closure of Y; composing this meromorphic mapping with a birational mapping $\bar{Y} \dashrightarrow X$, one obtains a meromorphic mapping $\Phi_1\colon W \dashrightarrow X$.

Lemma 5.3. *There exists a non-empty open subset $O \subset W$ such that Φ_1 is defined on O and the derivative $D\Phi_1(x)$ is non-degenerate for any $x \in O$.*

Assuming this lemma for a while, let us finish the proof of Proposition 1.1.

Our construction of the surfaces X and Y implies that $\mathcal{M}(X) \cong \mathcal{M}(W)$; hence, to prove Proposition 1.1 it suffices to show that $\mathcal{M}(X) \cong \mathbb{C}(T_1, T_2)$. We will derive this fact from the Castelnuovo rationality criterion (see for example [3, Chapter VI, 3.4]), which may be stated as follows.

Theorem 5.4 (Castelnuovo)**.** *Suppose that X is a smooth projective surface over $\mathbb{C}$. Then $\mathcal{M}(X) \cong \mathbb{C}(T_1, T_2)$ if and only if $H^0(X, \Omega^1_X) = 0$ and $H^0(X, \omega_X^{\otimes 2}) = 0$.*

Here, $H^0(X, \Omega^1_X)$ is the space of holomorphic 1-forms on X and $H^0(X, \omega_X^{\otimes 2})$ is the space of holomorphic 2-forms of weight 2 on X; we are to check that, for our surface X, both these linear spaces do not contain non-zero elements.

To begin with, observe that if η is a holomorphic covariant tensor field on X that is not identically zero, then $\Phi_1^*\eta$ is a holomorphic tensor field on $W \setminus I$, where I is the indeterminacy locus of Φ_1, and, in view of Lemma 5.3, $\Phi_1^*\eta$ is not identically zero. Since I is a discrete subset of the complex surface W, $\Phi_1^*\eta$ extends to a tensor field on the entire W. Thus, to show that $H^0(X, \Omega^1_X) = 0$ and $H^0(X, \omega_X^{\otimes 2}) = 0$ it suffices to show that $H^0(W, \Omega^1_W) = 0$ and $H^0(W, \omega_W^{\otimes 2}) = 0$, that is, that there are no non-trivial holomorphic 1-forms or 2-forms of weight 2 on W. We deal with these two types of tensor fields separately.

The absence of holomorphic 1-forms. This is just the following lemma, which will be used in Section 6 as well.

Lemma 5.5. *Suppose that F is a non-singular complex surface, $C \subset F$ is a curve such that $C \cong \mathbb{P}^1$ and $(C \cdot C) > 0$, and W is a good neighborhood of C. Then any holomorphic 1-form on W is identically zero.*

Proof. Suppose that ω is such a form. Proposition 4.3 implies that, for any $x \in W$ there exist curves $C_1, C_2 \subset W$, $C_1 \cong C_2 \cong \mathbb{P}^1$ such that $C_1 \cap C_2 \ni x$ and C_1 and C_2 are transverse at X. Since there are no non-zero holomorphic 1-forms on the Riemann sphere, the restriction of ω to both C_1 and C_2 is identically zero. Hence, the linear functional ω_x that ω induces on T_xW is zero on $T_xC_1, T_xC_2 \subset T_xW$. Since C_1 and C_2 are transverse at x, these two linear spaces span the entire T_xW, so $\omega_x = 0$. Since $x \in W$ was arbitrary, $\omega = 0$ and we are done. □

The absence of holomorphic 2-forms of weight 2. Recall that differential 2-forms of weight 2 on a surface G have, in local coordinates (z, w), the form $f(z, w)(dz \wedge dw)^2$. If ω is such a form, then, for any point $x \in G$, ω_x is a mapping from $T_xG \times T_xG$ to $\mathbb{C}$; this mapping is uniquely determined by its value at a given pair of linearly independent tangent vectors.

Lemma 5.6. *If $G = U \times \mathbb{P}^1$, where U is an open subset of $\mathbb{C}$, then there is no non-trivial holomorphic 2-form of weight 2 on G.*

Proof. Suppose that ω is such a form. If z is the coordinate on $U \subset \mathbb{C}$, then there exists a nowhere vanishing holomorphic vector field $\partial/\partial z$ on G. For any $b \in U$, put $C_b = \{b\} \times \mathbb{P}^1$. To show that $\omega = 0$ it suffices to show that for any $b \in U$ and $x \in C_b$ one has $\omega_x(\partial/\partial z, v) = 0$, where $v \in T_xC_b$ is a nonzero tangent vector.

Consider the tensor field $\eta = i_{\partial/\partial z}\omega$ (the contraction of ω with $\partial/\partial z$), which is a family of functions $\eta_x : T_xG \to \mathbb{C}$ for all $x \in G$, $\eta_x(w) = \omega_x(\partial/\partial z, w)$ for $w \in T_xG$. The field ω is a holomorphic section of $\operatorname{Sym}^2 \Omega^1_G$, and its restriction to each C_b is a section of $\omega_{C_b}^{\otimes 2}$, i.e., a quadratic differential on C_b, i.e., a section of $\mathcal{O}_{C_b}(-4)$; such a holomorphic section must be identically zero, so, for any $x \in C_b$ and any $v \in T_xC_b$, $\omega_x(\partial/\partial z, v) = \eta(v) = 0$, and we are done. □

Now suppose that ω is a differential 2-form of weight 2 on W. To show that $\omega = 0$, pick d distinct points $p_1, \dots, p_d \in C$, where $d = (C \cdot C)$, and let $\bar{W}$ be the blowup of W at $p_1, \dots, p_d$ and $\bar{C} \subset \bar{W}$ be the strict transform of C. It suffices to show that $\sigma^*\omega = 0$, where $\sigma\colon\ \bar{W} \to W$ is the blowdown morphism, and it will suffice to show that $\sigma^*\omega = 0$ on a non-empty open subset of $\bar{W}$. Since $\bar{C} \cong \mathbb{P}^1$ and $(\bar{C} \cdot \bar{C}) = 0$, it follows from the main result of Savelyev's paper [11] that a neighborhood of $\bar{C}$ in $\bar{W}$ is isomorphic to $U \times \mathbb{P}^1$, where U is an open subset of $\mathbb{C}$. Now Lemma 5.6 applies.

This completes the proof of Proposition 1.1 modulo Lemma 5.3.

Proof of Lemma 5.3. It suffices to prove this assertion for $\Phi\colon\ V \to Y \subset \mathbb{C}^3$ instead of Φ_1. Moreover, if $\pi\colon\ Y \to \mathbb{C}^2$ is the projection defined by forgetting the third coordinate, then the derivative of π is non-degenerate on a non-empty Zariski open subset of the smooth locus of Y; hence, it suffices to establish the existence of such a set O for the mapping $\Psi = \pi\circ\Phi\colon\ V \to \mathbb{C}^2$, $\Psi\colon\ x \longmapsto (f(x), g(x))$.

The mapping Ψ extends to a meromorphic mapping $W \dashrightarrow \mathbb{P}^2$ defined, in the homogeneous coordinates, by $x \longmapsto (1 : f(x) : g(x))$; abusing the notation, we will denote this meromorphic mapping by Ψ as well. The indeterminacy locus of the meromorphic mapping Ψ is a discrete subset of W.

If there exists at least one point $x \in W$ where Ψ is defined and $D\Psi(x)$ is non-degenerate, we are done. Assume now that the derivative of Ψ is degenerate at any point where Ψ is determined; we will show that this assumption leads to a contradiction.

Let $\Delta \subset \boldsymbol{D}(F, C)$ be the open subset such that $W = q(p^{-1}(\Delta))$, where $p\colon\ \mathfrak{H} \to \boldsymbol{D}(F)$ and $q\colon\ \mathfrak{H} \to F$ are the canonical projections of the restriction of the universal family $\mathfrak{H}$; recall that the curve $q(p^{-1}(b)) \subset F$, where $b \in \Delta$, is denoted by C_b.

Observe that the restriction of Ψ to any C_b is a meromorphic, hence holomorphic, mapping from C_b to $\mathbb{P}^2$. For any $b_1, b_2 \in \Delta$, $(C_{b_1} \cdot C_{b_2}) = (C \cdot C) > 0$, hence $C_{b_1} \cap C_{b_2} \neq \varnothing$. Thus, if the restriction of Ψ to each C_b is constant, then Ψ is constant, which is nonsense. Hence, we may and will pick a $b \in \Delta$ such that the restriction of Ψ to C_b is not constant. Put $\Psi(C_b) = Z \subset \mathbb{P}^2$; it follows from the Chow theorem that Z is a projective algebraic curve.

Observe as well that the set of points $x \in W$ where Ψ is defined and $D\Psi(x) = 0$ must have empty interior (otherwise Ψ would be constant). Hence, there exists a closed analytic subset D with empty interior such that, for any $x \in W \setminus D$, Ψ is defined at x and $\operatorname{rank} D\Psi(x) = 1$. Hence, all the fibers of the restriction $\Psi|_{W\setminus D}$ are smooth analytic curves in $W \setminus D$.

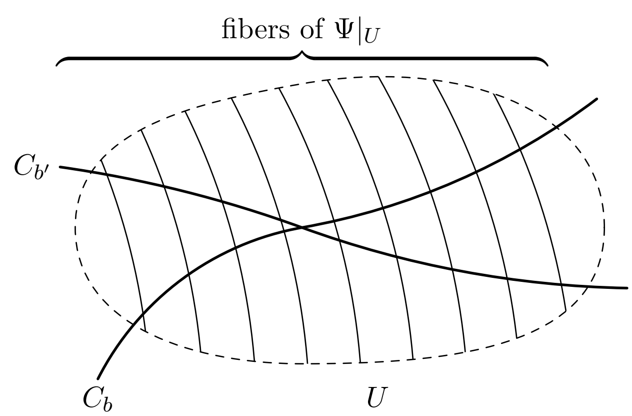


Figure 1: To the proof of Lemma 5.3.

Pick a point $x \in C_b \setminus D$ such that $T_x C_b$ is not contained in $\operatorname{Ker} D\Psi(x)$. There exists an open set $U \ni x$, $U \subset W \setminus D$ such that for any $y \in C_b \cap U$ the set $\Psi^{-1}(\Psi(y))$ is a smooth analytic curve transverse to C at y. Now for any $b' \in \Delta$ that is close enough to b there exists a non-empty open set $V \subset C_{b'} \cap U$ such that for any $x' \in V$ there exists a point $y' \in C_b \cap U$ such that $\Psi^{-1}(\Psi(y')) \cap C_{b'}$ contains x' (see Fig. 1).

Hence, $\Psi(V) \subset \Psi(C_b)$; since V is a non-empty open subset of $C_{b'}$, one concludes that $\Psi(C_{b'}) = \Psi(C_b) = Z \subset \mathbb{P}^2$. Since the curves $C_{b'}$, for all b' close enough to b, sweep, by virtue of Proposition 4.2, an open subset of W, one concludes that $\Psi(W) \subset Z$. Since Z is an algebraic curve in $\mathbb{P}^2$ and Ψ is defined by the formula $x \longmapsto (1 : f(x) : g(x))$, it follows that the meromorphic functions f and g are algebraically dependent, which yields the desired contradiction. $\square$

6 Transcendence degree 1

In this section we prove Proposition 1.2. Its proof is similar to that of Proposition 1.1, but simpler.

To wit, by virtue of Proposition 5.1 the field $\mathcal{M}(F)$ is finitely generated over $\mathbb{C}$. Since $\operatorname{tr.deg}_{\mathbb{C}} \mathcal{M}(F) = 1$, one has $\mathcal{M}(F) = \mathbb{C}(f, g)$, where the meromorphic functions f and g are algebraically dependent over $\mathbb{C}$ (if $\mathcal{M}(F)$ is generated by one function, there is nothing to prove). Denote by P an irreducible polynomial in two independent variables F and G such that $P(f, g) = 0$; let $Y \subset \mathbb{C}^2$ be the affine curve that is the zero locus of P, and let X be the smooth projective curve (aka compact Riemann surface) for which $\mathcal{M}(X) \cong \mathcal{M}(Y)$.

Denote by $V \subset F$ the open subset on which both f and g are well defined and consider the holomorphic mapping $\Phi \colon V \to Y$ defined by the formula $x \longmapsto (f(x), g(x))$. The mapping Φ extends to a meromorphic mapping from F to $\bar{Y} \subset \mathbb{P}^2$, where $\bar{Y}$ is the closure of Y; composing this meromorphic mapping with a birational mapping $\bar{Y} \dashrightarrow X$, one obtains a meromorphic mapping $\Phi_1 : F \dashrightarrow X$. Since, by our construction, $\mathcal{M}(F) \cong \mathcal{M}(X)$, it suffices to show that $X \cong \mathbb{P}^1$, or, equivalently, that there are no non-trivial holomorphic 1-forms on X.

To that end, let $I \subset F$ be the indeterminacy locus of Φ_1; it is a discrete subset of F. Choose a good neighborhood $W \supset C$; since Φ_1 is not constant, there exists a non-empty open subset $O \subset W \setminus I$ such that $\operatorname{rank} D\Phi_1(x) = 1$ for any $x \in O$. Now if $\omega \neq 0$ is a holomorphic form on X, then $(\Phi_1|_{W \setminus I})^* \omega$ is a holomorphic form such that its restriction to O is not identically zero. Extending it to W, one obtains a holomorphic 1-form on W which is not identically zero. This contradicts Lemma 5.5, and this contradiction completes the proof of Proposition 1.2.

Acknowledgments

This study was partially supported by the HSE University Basic Research Program and by the SRISA project FNEF-2024-0001 (Reg. No 1023032100070-3-1.2.1).

References

[1] Aldo Andreotti. Théorèmes de dépendance algébrique sur les espaces complexes pseudo-concaves. *Bull. Soc. Math. France*, 91:1–38, 1963. 146

[2] Aldo Andreotti and Wilhelm Stoll. *Analytic and algebraic dependence of meromorphic functions*. Lecture Notes in Mathematics, Vol. 234. Springer-Verlag, Berlin-New York, 1971. 136

[3] Wolf P. Barth, Klaus Hulek, Chris A. M. Peters, and Antonius Van de Ven. *Compact complex surfaces*, volume 4 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer-Verlag, Berlin, second edition, 2004. 146, 147

[4] Adrien Douady. Le problème des modules pour les sous-espaces analytiques compacts d'un espace analytique donné. *Ann. Inst. Fourier (Grenoble)*, 16(1):1–95, 1966. 135, 136, 137

[5] Maycol Falla Luza and Frank Loray. Neighborhoods of rational curves without functions. *Math. Ann.*, 382(3-4):1047–1058, 2022. 134

[6] Maycol Falla Luza, Frank Loray, and Paulo Sad. Non-algebraizable neighborhoods of curves. preprint arXiv:2407.20206 [math.AG]. 134

[7] K. Kodaira. A theorem of completeness of characteristic systems for analytic families of compact submanifolds of complex manifolds. *Ann. of Math. (2)*, 75:146–162, 1962. 134, 138, 140

[8] Serge Lvovski. On algebraic and non-algebraic neighborhoods of rational curves. *Mosc. Math. J.*, 24(3):391–405, 2024. 134

[9] David Mumford. *Lectures on curves on an algebraic surface*. Annals of Mathematics Studies, No. 59. Princeton University Press, Princeton, NJ, 1966. With a section by G. M. Bergman. 138

[10] S. Yu. Orevkov. An example in connection with the Jacobian conjecture. *Mat. Zametki*, 47(1):127–136, 173, 1990. English translation: Mathematical Notes, 1990, Vol. 47, Issue 1, 82–88; erratum: Mathematical Notes, 1991, Vol. 49, Issue 6, 659. 146

[11] V. I. Savelyev. Zero-type imbedding of a sphere into complex surfaces. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, (4):28–32, 85, 1982. English translation: Mosc. Univ. Math. Bull. **37**, 34–39 (1982). 135, 140, 149

AUTHOR

Serge Lvovski,
National Research University Higher School of Economics,
6, Usacheva street, Moscow, 119048
Russian Federation
Scientific Research Institute for System Analysis
of the National Research Centre “Kurchatov Institute”
email: lvovski@gmail.com