

# $G_2$ -instantons over generalised Kummer constructions

Thomas Walpuski

23rd August 2014

## Abstract

In this article we introduce a method to construct  $G_2$ -instantons on  $G_2$ -manifolds arising from Joyce’s generalised Kummer construction [Joy96a; Joy96b]. The method is based on gluing ASD instantons over ALE spaces to flat bundles on  $G_2$ -orbifolds of the form  $T^7/\Gamma$ . We use this construction to produce non-trivial examples of  $G_2$ -instantons.

**Changes to the published version** This is a revision of the article [Wal13a], which has been published in *Geometry & Topology*, volume 17, issue 4, in 2013. It corrects a few minor issues and includes some additional material. Dominic Joyce pointed out to me in my PhD viva that parts of the beginning of the discussion in Section 6 in [Wal13a] were not quite phrased correctly. This is rectified in the current version; in particular, Definition 6.3 of the notion of gluing data has been clarified. This version includes Theorem 1.6, which was already contained in my PhD thesis [Wal13b, Theorem 2.4]. This theorem provides a first small step towards understanding when Theorem 1.1 gives a complete description of all  $G_2$ -instantons on certain topologically simple bundles.

## 1 Introduction

The seminal paper [DT98] of Donaldson–Thomas has inspired a considerable amount of work related to gauge theory in higher dimensions. Tian [Tiao0] and Tao–Tian [TT04] made significant progress on important foundational analytical questions. Recent work of Donaldson–Segal [DS11] and Haydys [Hay12] shed some light on the shape of the theories to be expected.

In this article we will focus on the study of gauge theory on  $G_2$ -manifolds. These are 7-manifolds equipped with a torsion-free  $G_2$ -structure. The  $G_2$ -structure allows us to define a special class of connections, called  $G_2$ -instantons (see Definition 3.1). These share many formal properties with flat connections on 3-manifolds and it is expected that there are  $G_2$ -analogues of those 3-manifold invariants that are related to “counting flat connections”, that is, the Casson invariant, instanton Floer homology, etc.

So far non-trivial examples of  $G_2$ -instantons are rather rare. By exploiting the special geometry of the known  $G_2$ -manifolds some progress has been made recently. At the time of writing, there are

essentially two methods for constructing compact  $G_2$ -manifolds in the literature. Both yield  $G_2$ -manifolds close to degenerate limits. One is Kovalev's twisted connected sum construction [Kov03], which produces  $G_2$ -manifolds with "long necks" from certain pairs of Calabi–Yau 3-folds with asymptotically cylindrical ends. A technique for constructing  $G_2$ -instantons on Kovalev's  $G_2$ -manifolds has recently been proposed by Sá Earp [Sá15; Sá11]. The other (and historically the first) method for constructing  $G_2$ -manifolds is due to Joyce [Joy96a; Joy96b] and is based on desingularising  $G_2$ -orbifolds. In this article we introduce a method to construct  $G_2$ -instantons on  $G_2$ -manifolds arising from Joyce's construction.

To set up the framework for our construction, let us briefly review the geometry of Joyce's construction: Equip  $T^7$  with a flat  $G_2$ -structure  $\phi_0$  and let  $\Gamma$  be a finite group of diffeomorphisms of  $T^7$  preserving  $\phi_0$ . Then  $Y_0 := T^7/\Gamma$  is a flat  $G_2$ -orbifold. The singular set  $S$  of  $Y_0$  can, in general, be quite complicated. In this article we restrict to *admissible*  $G_2$ -orbifolds  $Y_0$ . That is, we assume that each of the connected components  $S_j$  of  $S$  has a neighbourhood modelled on  $(T^3 \times \mathbb{C}^2/G_j)/H_j$ . Here  $G_j$  is a non-trivial finite subgroup of  $SU(2)$  and  $H_j$  is a finite group acting by isometries on  $T^3$  as well as on  $\mathbb{C}^2/G_j$ . Suppose we are given *resolution data*  $\mathbf{r} = \{(X_j, \rho_j)\}$  for  $Y_0$ , that is, for each  $j$ , an ALE space  $X_j$  asymptotic to  $\mathbb{C}^2/G_j$  together with an isometric action  $\rho_j$  of  $H_j$  on  $X_j$  which is asymptotic to the action of  $H_j$  on  $\mathbb{C}^2/G_j$ . Then using Joyce's generalised Kummer construction [Joy96a; Joy96b] we can resolve the singularities in  $Y_0$  and produce a compact 7-manifold  $Y$  together with a family of torsion-free  $G_2$ -structures  $(\phi_t)_{t \in (0, T)}$ .

In this article we will construct  $G_2$ -instantons over  $(Y, \phi_t)$  given *gluing data*  $\mathbf{g}$  compatible with the resolution data  $\mathbf{r}$  for  $Y_0$ . The notion of gluing data will be defined carefully in Section 6. For now, it suffices to say that  $\mathbf{g}$  consists of

- a  $G$ -bundle  $E_0$  over  $Y_0$  together with a flat connection  $\theta$  and
- for each  $j$ , a  $G$ -bundle  $E_j$  over  $X_j$  together with a framed ASD instanton  $A_j$

as well as various auxiliary data satisfying a number of compatibility conditions. Here we take  $G$  to be a compact connected semi-simple Lie group, for example,  $G = SO(3)$ .

**Theorem 1.1.** *Let  $Y_0$  be an admissible flat  $G_2$ -orbifold, let  $\mathbf{r}$  be resolution data for  $Y_0$  and let  $\mathbf{g}$  be compatible gluing data. Suppose that the flat connection  $\theta$  is acyclic and that the ASD instantons  $A_j$  are infinitesimally rigid. Then there is a constant  $T' \in (0, T]$  and a  $G$ -bundle  $E$  over  $Y$  as well as for each  $t \in (0, T')$  a connection  $A_t$  on  $E$  that is an acyclic  $G_2$ -instanton over  $(Y, \phi_t)$ . Moreover, the adjoint bundle  $\mathfrak{g}_E$  associated with  $E$  satisfies*

$$(1.2) \quad p_1(\mathfrak{g}_E) = - \sum_j k_j \text{PD}[S_j] \quad \text{with } k_j := \frac{1}{8\pi^2} \int_{X_j} |F_{A_j}|^2,$$

$$(1.3) \quad \text{and} \quad \langle w_2(\mathfrak{g}_E), \Sigma \rangle = \langle w_2(\mathfrak{g}_{E_j}), \Sigma \rangle$$

for each  $\Sigma \in H_2(X_j)^{H_j} \subset H_2(Y)$ . Here  $[S_j] \in H_3(Y, \mathbb{Q})$  is the rational homology class arising from  $S_j$  and  $H_2(X_j)^{H_j}$  denotes the  $H_j$ -invariant part of  $H_2(X_j)$ ; see Remark 4.11.

*Remark 1.4.* We will specify in Definition 3.11 and Definition 5.19, respectively, what it means for a  $G_2$ -instanton, and thus for a flat connection, being a particular instance of a  $G_2$ -instanton, to be acyclic and for an ASD instanton to be infinitesimally rigid.

*Remark 1.5.* We equip the adjoint bundles  $\mathfrak{g}_{E_j}$  and  $\mathfrak{g}_E$  with the inner product arising from the negative of the Killing form on the Lie algebra  $\mathfrak{g}$  associated with  $G$ .

The proof of Theorem 1.1 is based on a gluing construction. The analysis involved is similar to work on Spin(7)-instantons in Lewis' DPhil thesis [Lew98], unpublished work of Brendle on the Yang-Mills equation in higher dimension [Bre03] and Pacard-Ritoré's work on the Allen-Cahn equation [PR03]. From a geometric perspective our result can be viewed as a higher-dimensional analogue of Kronheimer's work on ASD instantons on Kummer surfaces [Kro91].

It is not too unreasonable to expect that under certain topological assumptions all  $G_2$ -instantons on  $G_2$ -manifolds arising from Joyce's generalised Kummer construction close to the degenerate limit come from a suitable generalisation of our construction. Optimistically, one could hope that this will some day make the (so far conjectural)  $G_2$  Casson invariant accessible to computation. The following result is a first step in this direction.

**Theorem 1.6.** *Let  $Y_0$  be an admissible  $G_2$ -orbifold all of whose singularities  $S_j$  are modelled on  $(T^3 \times \mathbb{C}^2 / \mathbb{Z}_2) / H_j$  and let  $Y$  be the compact 7-manifold and let  $(\phi_t)_{t \in (0, T)}$  be the family of torsion-free  $G_2$ -structures obtained via Joyce's generalised Kummer construction from resolution data for  $Y_0$ . Suppose that  $E$  is a  $\mathrm{SO}(3)$ -bundle over  $Y$  with the property that*

$$p_1(E) = - \sum_j \frac{\varepsilon_j}{2} \mathrm{PD}[S_j]$$

where  $\varepsilon_j = 1$  if  $H_2(X_j)^{H_j} \subset H_2(Y)$  is non-trivial and  $w_2(E)$  pairs non-trivially with  $H_2(X_j)^{H_j}$  and  $\varepsilon_j = 0$  otherwise. Suppose that  $(A_t)_{t \in (0, T)}$  is a family of connections on  $E$  such that  $A_t$  is a  $G_2$ -instanton over  $(Y, \phi_t)$ . Then there exists a  $\mathrm{SO}(3)$ -bundle  $E_0$  over  $Y_0$  together with a flat connection  $\theta$ , such that away from the singular set of  $Y_0$  the family of connections  $(A_t)$  converges up to gauge transformations to  $\theta$  in  $C_{\mathrm{loc}}^\infty$  on  $Y_0 \setminus S$  as  $t$  tends to zero.

In order to prove that Theorem 1.1 gives a complete description of the moduli space of  $G_2$ -instantons under the hypothesis of Theorem 1.6 one additionally needs to control the behaviour of the family  $(A_t)$  on the resolution locus in  $Y_t$  and on the neck region between the resolution locus and the regular part of  $Y_0$ . This appears to be technically very challenging. Moreover, one would need to either understand under which conditions the flat connection  $\theta$  will be unobstructed or extend Theorem 1.1 to handle the case of obstructed  $\theta$  as well. Instead of directly producing  $G_2$ -instantons this latter generalisation of Theorem 1.6 would give a local description of the moduli space of  $G_2$ -instantons in terms of a Kuranishi model.

**Further directions** In this article we have only considered the simplest case of Joyce's generalised Kummer construction. Joyce's construction in [Joy96a] can also handle  $G_2$ -orbifolds with singular

sets of codimension 6 which are resolved using ALE Calabi–Yau 3–folds. Using the work of Anda Degeratu and myself on rigid HYM connections over ALE Calabi–Yau 3–folds arising as moduli spaces of  $G$ –constellations on  $\mathbb{C}^3$  [DW16] much of the work in this chapter can be extended to this situation. The latest version of Joyce’s construction [Joy00] can handle very complicated singular sets which require QALE Calabi–Yau 3–folds in order to be resolved. Generalising Theorem 1.1 to this setting is a daunting task.

Using ideas from [Wal17] it seems reasonable to extend Theorem 1.1 roughly as follows. Suppose that  $(Y_t, \phi_t)$  is a family of  $G_2$ –manifolds, which degenerates to a (not necessarily) flat  $G_2$ –orbifold  $(Y_0, \phi_0)$  by bubbling of a family of ALE spaces  $\mathfrak{X}$  along an associative singular set  $S$ . Given a  $G_2$ –instanton  $B$  over  $Y_0$ , one can construct a bundle  $\mathfrak{M}$  over  $S$ , whose fibre over  $x \in S$  is a moduli space of ASD instantons on the ALE space  $\mathfrak{X}_x$ . If  $\mathfrak{S}$  is an unobstructed Fueter section of  $\mathfrak{M}$ , i.e., the linearisation of the Fueter operator at  $\mathfrak{S}$  has trivial cokernel, then one should be able to glue  $\mathfrak{S}$  and  $B$  to produce a family of  $G_2$ –instantons over  $(Y_t, \phi_t)$  for  $0 < t \ll 1$ . In the situation considered in this article  $\mathfrak{M}$  has discrete fibres; hence, if a section  $\mathfrak{S}$  exists, it automatically satisfies the Fueter equation and is unobstructed. At the time of writing all known examples of degenerations of  $G_2$ –manifolds arise from Joyce’s generalised Kummer construction and thus have flat limits  $(Y_0, \phi_0)$ ; however, forthcoming work of Dominic Joyce and Spiro Karigiannis is expected to produce examples with non-flat limits.

**Outline of the article** Sections 2, 3, 4 and 5 contain some foundational material on  $G_2$ –manifolds and  $G_2$ –instantons as well as brief reviews of Joyce’s generalised Kummer construction and Kronheimer and Nakajima’s work on ASD instantons on ALE spaces. The proof of Theorem 1.1 begins in earnest in Section 6, where we construct approximate  $G_2$ –instantons from gluing data and introduce weighted Hölder spaces adapted to the problem at hand. In Section 7 we set up the analytical problem underlying the proof of Theorem 1.1 and discuss a model for the linearised problem. We complete the proof of Theorem 1.1 in Section 8. A number of concrete examples of  $G_2$ –instantons with  $G = \text{SO}(3)$  are constructed in Section 9. The proof of Theorem 1.6 is given in Section 10.

**Acknowledgements** This article is the outcome of work undertaken by the author for his PhD thesis at Imperial College London, supported by European Research Council Grant 247331. I am grateful to my PhD supervisor Simon Donaldson his encouragement and support. Moreover, I would like to thank the anonymous referee for helpful comments on an earlier version of this article.

## 2 Review of $G_2$ –manifolds

In this section we recall some basic definitions and results in  $G_2$ –geometry. For a more comprehensive treatment we refer the reader to Joyce’s book [Joy00], specifically Chapter 10.

The Lie group  $G_2$  can be defined as the subgroup of elements of  $GL(7)$  fixing the 3–form

$$(2.1) \quad \phi_0 := dx^{123} + dx^{145} + dx^{167} + dx^{246} - dx^{257} - dx^{347} - dx^{356}.$$

Here  $dx^{ijk}$  is a shorthand for  $dx^i \wedge dx^j \wedge dx^k$  and  $x_1, \dots, x_7$  are standard coordinates on  $\mathbf{R}^7$ . The particular choice of  $\phi_0$  is not important. Any non-degenerate 3–form  $\phi$  on  $\mathbf{R}^7$  is equivalent to  $\phi_0$  under a change of coordinates; see, for example, Salamon and Walpuski [SW17, Theorem 3.2]. Here we say that  $\phi$  is non-degenerate if for each non-zero vector  $u \in \mathbf{R}^7$  the 2–form  $i(u)\phi$  on  $\mathbf{R}^7/\langle u \rangle$  is symplectic. It follows from the identity

$$(2.2) \quad i(u)\phi_0 \wedge i(v)\phi_0 \wedge \phi_0 = 6g_{\mathbf{R}^7}(u, v)\text{vol}_{\mathbf{R}^7}$$

that any element of  $GL(7)$  which preserves  $\phi_0$  also preserves the standard inner product  $g_{\mathbf{R}^7}$  and the standard volume form  $\text{vol}_{\mathbf{R}^7}$  on  $\mathbf{R}^7$ . Therefore,  $G_2$  is a subgroup of  $SO(7)$ . In particular, every non-degenerate 3–form  $\phi$  on a 7–dimensional vector space induces an inner product and an orientation on this vector space. As an aside, we should point out here that non-degenerate 3–forms constitute one of two open orbits of  $GL(7)$  in  $\Lambda^3(\mathbf{R}^7)^*$ . For  $\phi$  in the other open orbit, the analogue of equation (2.2) yields an indefinite metric of signature (3, 4). In particular, if we take  $u = v$  to be a light-like vector, then  $i(u)\phi$  is not a symplectic form on  $\mathbf{R}^7/\langle u \rangle$ .

From the above discussion it is clear that a non-degenerate 3–form  $\phi$  on  $Y$  is equivalent to a reduction of the structure group of  $TY$  from  $GL(7)$  to  $G_2$ , that is, a  $G_2$ –structure. Moreover,  $\phi$  induces a Riemannian metric  $g_\phi$  and an orientation on  $Y$ . The intrinsic torsion of the  $G_2$ –structure corresponding to  $\phi$  can be identified with  $\nabla_{g_\phi}\phi$ .

**Definition 2.3.** A  $G_2$ –manifold is a 7–manifold  $Y$  equipped with a torsion-free  $G_2$ –structure  $\phi$ , that is,

$$\nabla_{g_\phi}\phi = 0.$$

*Remark 2.4.* Analogously, one can define the general notion of a  $G_2$ –orbifold. (For a thorough discussion of orbifolds we recommend the book of Adem–Leida–Ruan [ALR07].) In this article, however, we will only encounter very simple  $G_2$ –orbifolds of the form  $(Y/\Gamma, \phi)$  where  $(Y, \phi)$  is a  $G_2$ –manifold and  $\Gamma$  is a finite group of diffeomorphism of  $Y$  preserving  $\phi$ .

There is a plethora of reasons to be interested in  $G_2$ –manifolds.  $G_2$ –manifolds have holonomy group  $\text{Hol}(g_\phi) \subset G_2$  which appears as one of the exceptional cases in Berger’s classification of holonomy groups of irreducible non-symmetric Riemannian manifolds [Ber55, Theorem 3].  $G_2$ –manifolds are spin manifolds and carry (at least) one non-zero parallel spinor (see Joyce [Joy00, Proposition 10.1.6]) and, hence, are Ricci-flat and of relevance to theoretical physics. Moreover,  $G_2$ –manifolds carry a pair of calibrations in the sense of Harvey–Lawson [HL82]: the *associative calibration*  $\phi$  and the *coassociative calibration*  $\psi := *\phi$ . This makes their submanifold geometry very rich and interesting. Furthermore, it is very appealing to study gauge theory on  $G_2$ –manifolds as we will see in Section 3.

**Example 2.5.** The 7–torus  $T^7 = \mathbf{R}^7/\mathbf{Z}^7$  equipped with the  $G_2$ –structure  $\phi_0$  defined in (2.1) is a  $G_2$ –manifold.

**Definition 2.6.** A *hyperkähler manifold* is a Riemannian manifold  $(X, g)$  together with a triple  $(I_1, I_2, I_3)$  of parallel orthogonal complex structures satisfying  $I_1 I_2 = -I_2 I_1 = I_3$ .

*Remark 2.7.* If  $(X, g, I_1, I_2, I_3)$  is a hyperkähler manifold, then the metric  $g$  is Kähler with respect to each of complex structures  $a_1 I_1 + a_2 I_2 + a_3 I_3$  with  $(a_1, a_2, a_3) \in S^2 \subset \mathbf{R}^3$ .

**Example 2.8.** Let  $(X, g, I_1, I_2, I_3)$  be a hyperkähler 4-manifold. For  $i = 1, 2, 3$  denote by  $\omega_i := g(I_i \cdot, \cdot)$  the Kähler form associated with the complex structure  $I_i$ . Choose an orthonormal triple  $(\delta^1, \delta^2, \delta^3)$  of constant 1-forms on  $T^3$ . Then  $T^3 \times X$  is a  $G_2$ -manifold with torsion-free  $G_2$ -structure  $\phi$  defined by

$$\phi := \delta^1 \wedge \delta^2 \wedge \delta^3 + \delta^1 \wedge \omega_1 + \delta^2 \wedge \omega_2 - \delta^3 \wedge \omega_3.$$

The metric and the orientation on  $T^3 \times X$  induced by  $\phi$  coincide with the product metric and the product orientation. To see that, note that each cotangent space to  $X$  has a positive orthonormal basis  $(e^0, \dots, e^3)$  with  $e^i = I_i e^0$ , for  $i = 1, 2, 3$ , such that

$$(2.9) \quad \begin{aligned} \omega_1 &= e^0 \wedge e^1 + e^2 \wedge e^3, \\ \omega_2 &= e^0 \wedge e^2 - e^1 \wedge e^3, \\ \omega_3 &= e^0 \wedge e^3 + e^1 \wedge e^2. \end{aligned}$$

This immediately yields a orientation-preserving isometry  $T_x(T^3 \times X) \rightarrow \mathbf{R}^7$  identifying  $\phi$  with  $\phi_0$ . Note that in the current example the coassociative calibration  $\psi := *\phi$  is given by

$$(2.10) \quad \psi = \frac{1}{2} \omega_1 \wedge \omega_1 + \delta^2 \wedge \delta^3 \wedge \omega_1 + \delta^3 \wedge \delta^1 \wedge \omega_2 - \delta^1 \wedge \delta^2 \wedge \omega_3.$$

*Remark 2.11.* The above examples have holonomy strictly contained in  $G_2$ . This is clear from their construction, but can also be seen as a consequence of their topology since a compact  $G_2$ -manifold  $(Y, \phi)$  satisfies  $\text{Hol}(g_\phi) = G_2$  if and only if  $\pi_1(Y)$  is finite; see Joyce [Joy00, Proposition 10.2.2].

The following observation is central for the construction of  $G_2$ -manifolds.

**Theorem 2.12** (Fernández and Gray [FG82, Theorem 4.9]). *Let  $Y$  be a 7-manifold. Denote by  $\mathcal{P} \subset \Omega^3(Y)$  the subspace of all non-degenerate 3-forms on  $Y$  and define  $\Theta: \mathcal{P} \rightarrow \Omega^4(Y)$  by*

$$(2.13) \quad \Theta(\phi) := *\phi.$$

Here  $*_\phi$  is the Hodge  $*$ -operator associated with  $\phi$ . Then a  $G_2$ -structure  $\phi$  is torsion-free if and only if

$$d\phi = 0 \quad \text{and} \quad d\Theta(\phi) = 0.$$

The key difficulty in constructing  $G_2$ -manifolds comes from the fact that  $\Theta$  is non-linear. It is currently unknown which compact 7-manifolds do admit torsion-free  $G_2$ -structures. All known non-trivial compact examples arise by way of gluing constructions. One of those constructions will be described in more detail in Section 4.

Before we move on, let us recall a few facts, going back at least to the work of Fernández-Gray [FG82], that will be useful in the following. We refer the interested reader to Salamon and Walpuski [SW17, Theorem 8.4] for a detailed proof.

**Proposition 2.14.** *There is a  $G_2$ -invariant orthogonal splitting*

$$\Lambda^2(\mathbf{R}^7)^* = \Lambda_7^2 \oplus \Lambda_{14}^2,$$

where

$$\Lambda_7^2 := \{\omega : *(\omega \wedge \phi_0) = 2\omega\} \quad \text{and} \quad \Lambda_{14}^2 := \{\omega : *(\omega \wedge \phi_0) = -\omega\}.$$

Moreover,  $\Lambda_{14}^2$  is the kernel of the map  $\omega \mapsto \omega \wedge \psi_0$ , where  $\psi_0 := *\phi_0$ , and can be identified with  $\mathfrak{g}_2 \subset \mathfrak{so}(7) \cong \Lambda^2(\mathbf{R}^7)^*$ .

### 3 Gauge theory on $G_2$ -manifolds

Let  $(Y, \phi)$  be a compact  $G_2$ -manifold (or, more generally, a compact  $G_2$ -orbifold), let  $\psi := \Theta(\phi)$  and let  $E$  be a  $G$ -bundle over  $Y$ . Denote by  $\mathcal{A}(E)$  the space of connections on  $E$ .

**Definition 3.1.** A connection  $A \in \mathcal{A}(E)$  on  $E$  is called a  $G_2$ -instanton if it satisfies

$$(3.2) \quad *(F_A \wedge \phi) = -F_A.$$

These equations have first appeared in the physics literature (see Corrigan, Devchand, Fairlie, and Nuyts [CDFN83]) and were later brought to a wider attention by Donaldson and Thomas [DT98, Section 3]. (3.2) can be thought of as a 7-dimensional version of the anti-self-duality condition familiar from dimension four. As we will discuss shortly,  $G_2$ -instantons also have a striking similarity with flat connections over 3-manifolds.

**Example 3.3.** Flat connections are  $G_2$ -instantons.

**Example 3.4.** Let  $X$  be a hyperkähler manifold, let  $E$  be a  $G$ -bundle over  $X$  and let  $A$  be an ASD instanton on  $E$ , that is, a connection on  $E$  whose curvature  $F_A$  is anti-self-dual. Then the pullback of  $A$  to the  $G_2$ -manifold  $T^3 \times X$  from Example 2.8 is a  $G_2$ -instanton:

$$*(F_A \wedge \phi) = *(F_A \wedge \delta^1 \wedge \delta^2 \wedge \delta^3) = *_X F_A = -F_A.$$

Here we used that  $F_A \wedge \omega_i = 0$  and  $*_X$  denotes the Hodge  $*$ -operator on  $X$ .

**Example 3.5.** The Levi-Civita connection on a  $G_2$ -manifold is a  $G_2$ -instanton. To see that, observe that at each point we can think of the Riemannian curvature tensor  $R$  as an element of  $S^2 \mathfrak{g}_2 \subset \Lambda^2 \otimes \mathfrak{gl}(7)$ , since  $\text{Hol}(g_\phi) \subset G_2$ . But then it follows from Proposition 2.14 that  $*(R \wedge \phi) = -R$ .

Since  $\phi$  is closed, it follows from the Bianchi identity that  $G_2$ -instantons are Yang–Mills connections, that is,  $d_A^* F_A = 0$ . In fact, they are absolute minima of the Yang–Mills functional  $\text{YM}: \mathcal{A}(E) \rightarrow \mathbf{R}$ , since

$$(3.6) \quad \text{YM}(A) := \int_Y |F_A|^2 \text{vol} = \frac{1}{3} \int_Y |F_A + *(F_A \wedge \phi)|^2 \text{vol} - \int_Y \langle F_A \wedge F_A \rangle \wedge \phi$$

and, by Chern–Weil theory, the second term is a topological constant depending only on  $E$ . The energy identity (3.6) follows from a straight-forward computation using Proposition 2.14.

**Proposition 3.7.** *Let  $A \in \mathcal{A}(E)$  be a connection on  $E$ . The following are equivalent:*

1.  $A$  is  $G_2$ -instanton.
2.  $A$  satisfies  $F_A \wedge \psi = 0$ .
3. There is a  $\xi \in \Omega^0(Y, \mathfrak{g}_E)$  such that

$$(3.8) \quad *(F_A \wedge \psi) + d_A \xi = 0.$$

*Proof.* The equivalence of (1) and (2) follows immediately from Proposition 2.14. Obviously, (2) implies (3). By the Bianchi identity and since  $d\psi = 0$  it follows from (3) that  $d_A^* d_A \xi = 0$ . Hence, by integration by parts,

$$\int_Y |d_A \xi|^2 = \int_Y \langle d_A^* d_A \xi, \xi \rangle = 0.$$

Therefore  $d_A \xi = 0$  and (3) implies (2).  $\square$

From Proposition 3.7 it becomes apparent that  $G_2$ -instantons are rather similar to flat connections on 3-manifolds. In particular, if  $A_0$  is a  $G_2$ -instanton on  $E$ , then there is a  $G_2$  Chern–Simons functional  $CS^\psi : \mathcal{A}(E) \rightarrow \mathbf{R}$  defined by

$$CS^\psi(A_0 + a) := \int_Y \langle a \wedge d_{A_0} a + \frac{1}{3} a \wedge [a \wedge a] \rangle \wedge \psi$$

whose critical points are precisely the  $G_2$ -instantons on  $E$ . It is not entirely unreasonable to expect that some of the 3-manifold invariants arising from the Chern–Simons functional, like the Casson invariant and instanton Floer homology, have  $G_2$ -analogues. This idea goes back at least to the seminal paper of Donaldson–Thomas [DT98] and is one of the main motivations for studying  $G_2$ -instantons. Since (3.2) is invariant under the action of the group  $\mathcal{G}$  of gauge transformations of  $E$ , we can consider the *moduli space* of  $G_2$ -instantons on  $E$  over  $(Y, \phi)$ :

$$\mathcal{M}(E, \phi) := \{A \in \mathcal{A}(E) : F_A \wedge \psi = 0\} / \mathcal{G}.$$

Very roughly speaking, the conjectural  $G_2$  Casson invariant should be obtained by “counting”  $\mathcal{M}(E, \phi)$ . Whether there is a rigorous construction of such a  $G_2$  Casson invariant and whether it can, in fact, be arranged to be invariant under isotopies of the  $G_2$ -structure is an open question. A brief discussion of parts of this circle of ideas can be found in Donaldson and Segal [DS11, Section 6].

It is customary in gauge theory to work with local slices of the gauge group action. A particularly useful slicing condition is to require that  $B \in \mathcal{A}(E)$  be in *Coulomb gauge* with respect to a fixed reference connection  $A \in \mathcal{A}(E)$ , that is,  $d_A^*(B - A) = 0$ . (The importance of the Coulomb gauge stems from the foundational work of Uhlenbeck [Uhl82a]. For a careful discussion of how the Coulomb gauge is used in the construction moduli spaces we refer the reader to Donaldson and Kronheimer [DK90, Section 4.2].) For a fixed connection  $A \in \mathcal{A}(E)$  we consider the system of equations

$$(3.9) \quad *(F_{A+a} \wedge \psi) + d_{A+a} \xi = 0 \quad \text{and} \quad d_A^* a = 0$$

for  $\xi \in \Omega^0(Y, \mathfrak{g}_E)$  and  $a \in \Omega^1(Y, \mathfrak{g}_E)$ . This is simply (3.8) for  $A + a$  instead of  $A$  together with the condition that  $A + a$  be in Coulomb gauge with respect to  $A$ . The linearisation  $L_A: \Omega^0(Y, \mathfrak{g}_E) \oplus \Omega^1(Y, \mathfrak{g}_E) \rightarrow \Omega^0(Y, \mathfrak{g}_E) \oplus \Omega^1(Y, \mathfrak{g}_E)$  of (3.9) is given by

$$(3.10) \quad L_A := \begin{pmatrix} 0 & d_A^* \\ d_A & *(\psi \wedge d_A) \end{pmatrix}.$$

This is a self-adjoint elliptic operator. If  $A \in \mathcal{A}(E)$  is a  $G_2$ -instanton, then  $L_A$  controls the infinitesimal deformation theory of  $A$  as a  $G_2$ -instanton.

**Definition 3.11.** A  $G_2$ -instanton  $A$  is called *acyclic* if the operator  $L_A$  is invertible.

One can show that if every  $G_2$ -instanton  $A$  on  $E$  is acyclic, then  $\mathcal{M}(E, \phi)$  is, in fact, a smooth zero-dimensional manifold, that is, a discrete set.

## 4 Joyce's generalised Kummer construction

Equip  $T^7$  with a flat  $G_2$ -structure  $\phi_0$ , as in Example 2.5, and let  $\Gamma$  be a finite group of diffeomorphisms of  $T^7$  preserving  $\phi_0$ . Then  $Y_0 := T^7/\Gamma$  is a flat  $G_2$ -orbifold. Denote by  $S$  the singular set of  $Y_0$  and denote by  $S_1, \dots, S_k$  its connected components.

**Definition 4.1.**  $Y_0$  is called *admissible* if each  $S_j$  has a neighbourhood isometric to a neighbourhood of the singular set of  $(T^3 \times \mathbb{C}^2/G_j)/H_j$ . Here  $G_j$  is a non-trivial finite subgroup of  $SU(2)$  and  $H_j$  is a finite group acting by isometries on  $T^3$  as well as on  $\mathbb{C}^2/G_j$ .

Let  $Y_0$  be an admissible flat  $G_2$ -orbifold. Then there is a constant  $\zeta > 0$  such that if we denote by  $T$  the set of points at distance less than  $\zeta$  to  $S$ , then  $T$  decomposes into connected components  $T_1, \dots, T_k$  such that  $T_j$  contains  $S_j$  and is isometric to  $(T^3 \times B_\zeta^4/G_j)/H_j$ . On  $T_j$  we can write

$$\phi_0 = \delta^1 \wedge \delta^2 \wedge \delta^3 + \delta^1 \wedge \omega_1 + \delta^2 \wedge \omega_2 - \delta^3 \wedge \omega_3,$$

where  $(\delta^1, \delta^2, \delta^3)$  is an orthonormal triple of constant 1-forms on  $T^3$  and where  $(\omega_1, \omega_2, \omega_3)$  is the triple of Kähler forms associated with the standard hyperkähler structure  $(g, I_1, I_2, I_3)$  on  $\mathbb{C}^2 \cong \mathbb{H}$ .

**Definition 4.2.** Let  $G$  be a finite subgroup of  $SU(2)$ . Then an *ALE space asymptotic to  $\mathbb{C}^2/G$*  is a hyperkähler 4-manifold  $(X, \hat{g}, \hat{I}_1, \hat{I}_2, \hat{I}_3)$  together with a continuous map  $\pi: X \rightarrow \mathbb{C}^2/G$  inducing a diffeomorphism from  $X \setminus \pi^{-1}(0)$  to  $(\mathbb{C}^2 \setminus \{0\})/G$  such that

$$(4.3) \quad \nabla^k(\pi_*\hat{g} - g) = O(r^{-4-k}) \quad \text{and} \quad \nabla^k(\pi_*\hat{I}_i - I_i) = O(r^{-4-k})$$

as  $r \rightarrow \infty$  for  $i = 1, 2, 3$  and  $k \geq 0$ . Here  $r: \mathbb{C}^2/G_j \rightarrow [0, \infty)$  denotes the radius function.

We will remove the singularity in  $Y_0$  along  $S_j$  by, roughly speaking, replacing each  $\mathbb{C}^2/G_j$  with an ALE space asymptotic to  $\mathbb{C}^2/G_j$ . Due to work of Kronheimer [Kro89b; Kro89a], ALE spaces are very well understood.

**Theorem 4.4** (Kronheimer [Kro89a, Theorems 1.1, 1.2, and 1.3]). *Let  $G$  be a non-trivial finite subgroup of  $SU(2)$ . Denote by  $X$  the real 4-manifold underlying the crepant resolution  $\widetilde{X}/G$ . Then for each three cohomology classes  $\alpha_1, \alpha_2, \alpha_3 \in H^2(X, \mathbf{R})$  satisfying*

$$(4.5) \quad (\alpha_1(\Sigma), \alpha_2(\Sigma), \alpha_3(\Sigma)) \neq 0 \in \mathbf{R}^3$$

*for each  $\Sigma \in H_2(X, \mathbf{Z})$  with  $\Sigma \cdot \Sigma = -2$  there is a unique ALE hyperkähler structure on  $X$  for which the cohomology classes of the Kähler forms  $[\omega_i]$  are given by  $\alpha_i$ . Moreover, each ALE space asymptotic to  $\mathbf{C}^2/G$  is diffeomorphic to  $\widetilde{X}/G$  and its associated triple of Kähler classes satisfies (4.5).*

*Remark 4.6.* The crepant resolution  $\widetilde{X}/G$  can be obtained from  $\mathbf{C}^2/G$  by a sequence of blow-ups. The exceptional divisor  $E$  of  $X = \widetilde{X}/G$  has irreducible components  $\Sigma_1, \dots, \Sigma_k$ . By the McKay correspondence [McK80], these components form a basis of  $H_2(X, \mathbf{Z})$  and the matrix with coefficients  $C_{ij} = -[\Sigma_i] \cdot [\Sigma_j]$  is the Cartan matrix associated with the Dynkin diagram corresponding to  $G$  in the ADE classification of finite subgroups of  $SU(2)$ .

**Definition 4.7.** A collection  $\mathbf{r} = \{(X_j, \rho_j)\}$  consisting of, for each  $j$ , an ALE space  $X_j$  asymptotic to  $\mathbf{C}^2/G_j$  together with an isometric action  $\rho_j$  of  $H_j$  on  $X_j$  which is asymptotic to the action of  $H_j$  on  $\mathbf{C}^2/G_j$  is called *resolution data* for  $Y_0$ .

Suppose we are given resolution data  $\mathbf{r} = \{(X_j, \rho_j)\}$ . Denote by  $\pi_j: X_j \rightarrow \mathbf{C}^2/G_j$  the resolution map for  $X_j$ . For  $t > 0$  define

$$(4.8) \quad \pi_{j,t} := t\pi_j: X_j \rightarrow \mathbf{C}^2/G_j$$

and set

$$(4.9) \quad \tilde{T}_{j,t} := (T^3 \times \pi_{j,t}^{-1}(B_\zeta^4/G_j))/H_j \quad \text{and} \quad \tilde{T}_t := \bigcup_j \tilde{T}_{j,t}.$$

Using  $\pi_{j,t}$  we can replace each  $T_j$  in  $Y_0$  by  $\tilde{T}_{j,t}$  and thus obtain a compact 7-manifold  $Y_t$ .

*Remark 4.10.* The diffeomorphism type of  $Y_t$  is independent of  $t > 0$ . Hence, we will sometimes drop the label  $t$  and pretend to be working with a fixed 7-manifold  $Y$ . However, at various points it will be important to remember the precise way in which  $Y_t$  was constructed.

*Remark 4.11.* The (co)homology groups and the fundamental group of  $Y$  can relatively easily be computed from the above construction, the latter being especially important in view of Remark 2.11. In particular, it can be seen that every  $\Sigma \in H_2(X_j, \mathbf{Z})$  invariant under the action of  $H_j$  yields a cohomology class  $\Sigma \in H_2(Y, \mathbf{Z})$ . Also each component of singular set  $S_j$  gives rise to a *rational* homology class

$$(4.12) \quad [S_j] := \frac{1}{|H_j|} (\iota_{j,t})_*(T^3 \times \{x\}) \in H_3(Y, \mathbf{Q}),$$

where  $\iota_{j,t}: T^3 \times \pi_{j,t}^{-1}(B_\zeta^4/G_j) \rightarrow Y$  denotes the projection to  $\tilde{T}_{j,t}$  followed by the inclusion into  $Y$  and  $x$  denotes a point in  $\pi_{j,t}^{-1}(B_\zeta^4/G_j)$ .

On  $\tilde{T}_{j,t}$  there is a torsion-free  $G_2$ -structure given by

$$\hat{\phi}_{j,t} := \delta^1 \wedge \delta^2 \wedge \delta^3 + t^2 \delta^1 \wedge \hat{\omega}_{j,1} + t^2 \delta^2 \wedge \hat{\omega}_{j,2} - t^2 \delta^3 \wedge \hat{\omega}_{j,3}.$$

Near the boundary of  $\tilde{T}_{j,t}$  the 3-forms  $\hat{\phi}_{j,t}$  and  $\phi_0$  are close to each other. In order to patch them together note that there are 1-forms  $\varrho_{j,t,i}$  on  $(\mathbb{C}^2 \setminus \{0\})/G_j$  such that

$$t^2(\pi_{j,t})_* \hat{\omega}_{j,i} = \omega_i + d\varrho_{j,t,i}$$

with  $\nabla^k \varrho_{j,t,i} = t^4 O(r^{-3-k})$  for  $k \geq 0$ ; see [Joy00, Theorem 8.2.3]. Now, fix a smooth non-decreasing function  $\chi: [0, \zeta] \rightarrow [0, 1]$  such that  $\chi(s) = 0$  for  $s \leq \zeta/4$  and  $\chi(s) = 1$  for  $s \geq \zeta/2$  and set

$$\tilde{\omega}_{j,t,i} := t^2 \hat{\omega}_{j,i} - d(\chi(|\pi_{j,t}|) \cdot \pi_{j,t}^* \varrho_{j,t,i}).$$

Then  $(\pi_{j,t})_* \tilde{\omega}_{j,t,i}$  and  $\omega_i$  agree on  $r^{-1}[\zeta/2, \infty)$  and we can define a 3-form  $\tilde{\phi}_t \in \Omega^3(Y_t)$  by  $\tilde{\phi}_t := \phi_0$  on  $Y_0 \setminus T_t = Y_t \setminus \tilde{T}_t$  and by

$$\tilde{\phi}_t := \delta^1 \wedge \delta^2 \wedge \delta^3 + \delta^1 \wedge \tilde{\omega}_{j,t,1} + \delta^2 \wedge \tilde{\omega}_{j,t,2} - \delta^3 \wedge \tilde{\omega}_{j,t,3}$$

on  $\tilde{T}_{j,t}$ . Define the function  $r_t: Y_t \rightarrow [0, \zeta]$  by

$$(4.13) \quad r_t(p) := \begin{cases} |\pi_{j,t}(y)| & \text{for } p = [(x, y)] \in \tilde{T}_{j,t} \\ \zeta & \text{for } p \in Y_t \setminus \tilde{T}_t \end{cases}$$

and set

$$(4.14) \quad R_{j,t} := \tilde{T}_{j,t} \cap r_t^{-1}[\zeta/4, \zeta/2] \quad \text{and} \quad R_t := \bigcup_j R_{j,t} = r_t^{-1}[\zeta/4, \zeta/2].$$

Outside  $R_t$  the 3-form  $\tilde{\phi}_t$  defines a torsion-free  $G_2$ -structure, while on  $R_{j,t}$  it satisfies  $\nabla^k(\tilde{\phi}_t - \hat{\phi}_{j,t}) = O(t^4)$  for  $k \geq 0$  and similarly, for each fixed  $\varepsilon > 0$ , on  $r_t^{-1}[\varepsilon, \zeta]$  we have  $\nabla^k(\tilde{\phi}_t - \phi_0) = O(t^4)$  for  $k \geq 0$ . In particular,  $\tilde{\phi}_t$  defines a  $G_2$ -structure on  $Y_t$  provided  $t > 0$  is sufficiently small.

We equip  $Y_t$  with the Riemannian metric  $\tilde{g}_t := g_{\tilde{\phi}_t}$  associated with  $\tilde{\phi}_t$ .

*Remark 4.15.* Note that on the complement of  $\tilde{T}_t$  the metric  $\tilde{g}_t$  agrees with the flat metric  $g_0$  on  $(T^7/\Gamma) \setminus T$  and on  $\tilde{T}_{j,t} \setminus R_{j,t}$  it agrees with the metric

$$g_{\hat{\phi}_{j,t}} = g_{\mathbb{R}^3} \oplus t^2 g_{X_j}.$$

Here  $g_{\mathbb{R}^3}$  denotes the standard metric on  $\mathbb{R}^3$  and  $g_{X_j}$  denotes the metric on  $X_j$ . Moreover, since the map  $\phi \mapsto g_\phi$  is smooth, on  $R_{j,t}$  we have  $\nabla^k(\tilde{g}_t - g_{\mathbb{R}^3} \oplus t^2 g_{X_j}) = O(t^4)$  for  $k \geq 0$  and, for each fixed  $\varepsilon > 0$ , on  $r_t^{-1}[\varepsilon, \zeta]$  we have  $\nabla^k(\tilde{g}_t - g_0) = O(t^4)$  for  $k \geq 0$ .

**Theorem 4.16** (Joyce [Joy96a, Theorems A and B] and [Joy96b, Theorem 2.2.1]). *There are constants  $T, c > 0$  and for each  $t \in (0, T)$  a 2-form  $\eta_t$  on  $Y_t$  such that  $\phi_t := \tilde{\phi}_t + d\eta_t$  defines a torsion-free  $G_2$ -structure and*

$$(4.17) \quad \|d\eta_t\|_{L^\infty} \leq ct^{1/2}.$$

*Remark 4.18.* In view of Theorem 2.12 the above is tantamount to saying that one can solve the non-linear partial differential equation

$$(4.19) \quad d\Theta(\tilde{\phi}_t + d\eta_t) = 0$$

with estimates on  $d\eta_t$ . For small  $\eta_t$ , the dominant part of this equation is essentially the Laplacian on 2-forms. Now, as  $t > 0$  decreases the size of  $d\Theta(\tilde{\phi}_t)$  becomes smaller and smaller, but at the same time the mapping properties of the Laplacian degenerate. Solving (4.19) thus is a rather delicate balancing act.

For our application we need to slightly strengthen the estimate in Theorem 4.16. Let  $w_t(x, y) := t + \min\{r_t(x), r_t(y)\}$ . For a Hölder exponent  $\alpha \in (0, 1)$  define

$$[f]_{C_{0,t}^{0,\alpha}(U)} := \sup_{d(x,y) \leq w_t(x,y)} w_t(x,y)^\alpha \frac{|f(x) - f(y)|}{d(x,y)^\alpha},$$

$$\|f\|_{C_{0,t}^{0,\alpha}(U)} := \|f\|_{L^\infty(U)} + [f]_{C_{0,t}^{0,\alpha}(U)},$$

for a tensor field  $f$  over  $U \subset Y_t$ . Here we use parallel transport to compare the values of  $f$  at various points of  $U$ . If  $U$  is unspecified, then we take  $U = Y_t$ .

**Proposition 4.20.** *The constants  $T, c > 0$  in Theorem 4.16 can be chosen such that for all  $t \in (0, T)$  we have*

$$\|d\eta_t\|_{C_{0,t}^{0,\alpha}} \leq ct^{1/2} \quad \text{and} \quad \|\Theta(\phi_t) - \Theta(\hat{\phi}_{j,t})\|_{C_{0,t}^{0,\alpha}(\tilde{Y}_{j,t})} \leq ct^{1/2}.$$

For the proof of this result it will be helpful to note the following.

**Proposition 4.21.** *For each  $\mu > 0$  and  $K \in \mathbf{N}_0$  there exists a constant  $\varepsilon > 0$  such that the following holds for all  $t \in (0, T)$  and  $p \in Y_t$ :  $R := \varepsilon(t + r_t(p))$  is less than the injectivity radius of  $(Y_t, \tilde{g}_t)$  at  $p$  and if we identify  $T_p Y$  isometrically with  $\mathbf{R}^7$  and denote by  $s_R: B_1 \rightarrow B_R(p)$  the map obtained by multiplication with  $R$  followed by the exponential map, then*

$$(4.22) \quad |\partial^k (R^{-2} s_R^* \tilde{g}_t - g_{\mathbf{R}^7})| \leq \mu$$

for all  $k \in \{0, \dots, K\}$ . Here  $g_{\mathbf{R}^7}$  denotes the standard metric on  $\mathbf{R}^7$ .

*Proof.* From Remark 4.15 it is clear that we can find  $\varepsilon > 0$  such that the above statement holds for all  $p \in r_t^{-1}[\zeta/8, \zeta]$ . Moreover, for  $p \in r_t^{-1}[0, \zeta/8]$  inequality (4.22) is equivalent to

$$|\partial^k (\tilde{R}^{-2} s_{\tilde{R}}^* (g_{\mathbf{R}^3} \oplus g_{X_j}) - g_{\mathbf{R}^7})| \leq \mu,$$

where  $\tilde{R} := \varepsilon(1 + |\pi_j(y)|)$  and  $p = [(x, y)]$ . Because of (4.3) this holds for all  $\varepsilon \leq \frac{1}{2}$  as long as  $|\pi_j(y)|$  is sufficiently large, say,  $|\pi_j(y)| > N$ . For  $|\pi_j(y)| \leq N$  it can be arranged to hold by choosing  $\varepsilon > 0$  sufficiently small.  $\square$

*Proof of Proposition 4.20.* Note that the second part follows from the first and the construction of  $\tilde{\phi}_t$ , because  $\Theta$  is a smooth map. To obtain the estimate on  $d\eta_t$  recall from Joyce's construction that  $\eta_t$  solves a non-linear partial differential equation that can be written schematically as

$$(4.23) \quad d^*d\eta_t + P(d\eta_t, \nabla d\eta_t) = G(d\eta_t, \dots) \quad \text{and} \quad d^*\eta_t = 0;$$

see Joyce [Joy96a, Equation (33)]. The crucial points are that  $P(x, y)$  is a smooth function which depends linearly on  $y$  and satisfies  $P(0, y) = 0$  and that there is a constant  $c > 0$  such that

$$(4.24) \quad \|G(d\eta_t, \dots)\|_{L^\infty} \leq ct^{1/2}.$$

Now, define

$$D_t\sigma := (d^*\sigma + P(d\eta_t, \nabla\sigma), d\sigma).$$

Since  $d\eta_t$  is small provided  $T > 0$  is small, this is a small perturbation of the operator  $d^* \oplus d$ . We extend  $D_t$  to an operator from  $\Omega^*(Y_t)$  to itself by defining  $D_t\sigma = (d^* \oplus d)\sigma$  for  $\sigma \in \Omega^k(Y_t)$  with  $k \neq 3$ , so that it becomes an elliptic operator. We will now prove that there are constants  $c > 0$  and  $\varepsilon \in (0, \frac{1}{2})$  such that for all  $t \in (0, T)$  and each  $p \in Y_t$  the following holds:

$$(4.25) \quad R^\alpha[\sigma]_{C^{0,\alpha}(B_{R/2}(p))} \leq c(R\|D_t\sigma\|_{L^\infty(B_R(p))} + \|\sigma\|_{L^\infty(B_R(p))})$$

with  $R := \varepsilon(t + r_t(p))$ . From this the asserted bound on  $[d\eta_t]_{C^{0,\alpha}_{0,t}}$  follows at once using (4.17), (4.23) and (4.24), since on  $B_{R/2}(p)$  we have  $w_t \leq 2\varepsilon^{-1}R$ .

For  $\mu > 0$  choose  $\varepsilon > 0$  according to Proposition 4.21 with  $K = 1$ . Let  $s_R: B_1^7 \rightarrow B_R(p)$  be as in Proposition 4.21. We define a rescaled operator  $\tilde{D}_{t,p}: \Omega^*(B_1) \rightarrow \Omega^*(B_1)$  by

$$\tilde{D}_{t,p}\sigma := (R^2s_R^*\tau, s_R^*\theta)$$

for  $\sigma \in \Omega^k(B_1)$ , where  $(\tau, \theta) := D_t(s_R^{-1})^*\sigma \in \Omega^{k-1}(B_1) \oplus \Omega^{k+1}(B_1)$ . It follows from Theorem 4.16 and Proposition 4.21 that by choosing  $T, \mu > 0$  sufficiently small, we can arrange that for all  $t \in (0, T)$  and  $p \in Y_t$  the rescaled operator  $\tilde{D}_{t,p}$  is as close to  $d \oplus d^*: \Omega^*(B_1) \rightarrow \Omega^*(B_1)$  as we wish. In particular, we can arrange that the family of operators  $\tilde{D}_{t,p}$  is uniformly elliptic with coefficients uniformly bounded in  $C^1$ . Hence, by standard elliptic theory, we can find a constant  $c > 0$  independent of  $t \in (0, T)$  and  $p \in Y_t$  such that the following  $L^q$  estimate holds:

$$\|\sigma\|_{W^{1,q}(B_{1/2})} \leq c(\|\tilde{D}_{t,p}\sigma\|_{L^q(B_1)} + \|\sigma\|_{L^q(B_1)}).$$

Combined with the Sobolev embedding  $W^{1,q} \hookrightarrow C^{0,1-7/q}$  this yields

$$[\sigma]_{C^{0,\alpha}(B_{1/2})} \leq c(\|\tilde{D}_{t,p}\sigma\|_{L^\infty(B_1)} + \|\sigma\|_{L^\infty(B_1)})$$

with  $c > 0$  independent of  $t \in (0, T)$  and  $p \in Y_t$ . This, however, is equivalent to the estimate (4.25) for the unscaled operator  $D_t$ .  $\square$

*Remark 4.26.* Proposition 4.20 can be viewed as a quantification of Joyce’s proof of the fact that  $\eta_t$  is smooth. In a similar fashion, one can also obtain estimates on higher Hölder norms of  $d\eta_t$ .

*Remark 4.27.* The kind of argument we used above goes back to work of [NW73, Theorem 3.1]. We will encounter this line of reasoning again in the proofs of Proposition 5.10 and Proposition 7.11.

## 5 ASD instantons on ALE spaces

Let  $\Gamma$  be a finite subgroup of  $SU(2)$ , let  $X$  be an ALE space asymptotic to  $\mathbb{C}^2/\Gamma$  and let  $E$  be a  $G$ -bundle over  $X$ . We denote by  $\mathcal{A}(E)$  the space of connections on  $E$ .

**Definition 5.1.** A *framing at infinity* of  $E$  is a bundle isomorphism  $\Phi: E_\infty|_U \rightarrow \pi_*E|_U$  where  $E_\infty$  is a  $G$ -bundle over  $(\mathbb{C}^2 \setminus \{0\})/\Gamma$  and  $U$  is the complement of a compact neighbourhood of the singular point in  $\mathbb{C}^2/\Gamma$ .

Let  $\theta$  be a flat connection on a  $G$ -bundle  $E_\infty$  over  $(\mathbb{C}^2 \setminus \{0\})/\Gamma$ .

**Definition 5.2.** Let  $\Phi: E_\infty|_U \rightarrow \pi_*E|_U$  be a framing at infinity of  $E$ . Then a connection  $A \in \mathcal{A}(E)$  is called *asymptotic to  $\theta$  at rate  $\delta$  with respect to  $\Phi$*  if

$$(5.3) \quad \nabla^k(\Phi^*A - \theta) = O(r^{\delta-k})$$

for all  $k \geq 0$ . Here  $\nabla$  is the covariant derivative associated with  $\theta$ .

**Definition 5.4.** A *framed ASD instanton asymptotic to  $\theta$  (at rate  $\delta$ )* is an ASD instanton  $A \in \mathcal{A}(E)$  on  $E$  together with a framing at infinity  $\Phi$  of  $E$  such that  $A$  is asymptotic to  $\theta$  at rate  $\delta$  with respect to  $\Phi$ . If no rate  $\delta$  is specified, then we take  $\delta = -3$ .

**Proposition 5.5.** *Let  $A \in \mathcal{A}(E)$  be an ASD instanton on  $E$  with finite energy, that is,*

$$\int_X |F_A|^2 \text{vol} < \infty,$$

*then there is a  $G$ -bundle  $E_\infty$  over  $(\mathbb{C}^2 \setminus \{0\})/\Gamma$  together with a flat connection  $\theta$  and a framing  $\Phi: E_\infty|_U \rightarrow \pi_*E|_U$  such that (5.3) holds with  $\delta = -3$*

*Proof.* We extend the argument in Donaldson and Kronheimer [DK90, page 98]. The topological space  $\hat{X} := X \cup \{\infty\}$  can be given the structure of an orbifold whose atlas contains the charts of  $X$  as well as a uniformising chart at infinity  $\varphi: B_\varepsilon/\Gamma \rightarrow \hat{X}$  which is constructed as follows. Fix an orientation reversing linear isometry  $\sigma$  of  $\mathbb{R}^4$ . We let  $\Gamma$  act on  $B_\varepsilon$  by  $(g, x) \mapsto \sigma^{-1}(g \cdot \sigma(x))$  and define  $\varphi(0) := \infty$  and  $\varphi(x) = \pi^{-1}(\sigma(x)/|x|^2)$ . If  $g$  denotes the metric on  $X$ , then the conformally equivalent metric  $\hat{g} := (1 + |\pi|^2)^{-2}g$  extends to  $\hat{X}$  as an orbifold metric. The metric is not necessarily smooth, but only  $C^{3,\alpha}$ ; however, that does not cause any problems. One should think of  $\hat{X}$  as a conformal compactification of  $X$  in the same way that  $S^4$  is a conformal compactification of  $\mathbb{R}^4$ .

Since the equation  $F_A^+ = 0$  as well as the energy are conformally invariant, we can think of  $A$  as a finite energy ASD instanton on  $(\hat{X} \setminus \{\infty\}, \hat{g})$ . By Uhlenbeck's removable singularities theorem [Uhl82b, Theorem 4.1], the pullback of  $A$  to  $B_\varepsilon \setminus \{0\}$  extends to a  $\Gamma$ -invariant ASD instanton over all of  $B_\varepsilon$ . Hence,  $A$  extends to an ASD instanton  $\hat{A}$  on an orbifold  $G$ -bundle  $\hat{E}$  over  $\hat{X}$ . Using radial parallel transport from  $\infty$  we obtain a trivialisation of  $\hat{E}$  over  $\varphi(B_\varepsilon/\Gamma)$  in which the connection matrix representing  $\hat{A}$  vanishes at  $\infty = \varphi(0)$ . Denote by  $\rho: \Gamma \rightarrow G$  the monodromy representation associated with  $\hat{E}|_\infty$ . Associated with  $\rho$  there are a  $G$ -bundle  $E_\infty$  over  $\varphi((B_\varepsilon \setminus \{0\})/\Gamma)$  and a flat connection  $\theta$  on  $E_\infty$ . The above trivialisation of  $\hat{E}$  over  $\varphi(B_\varepsilon/\Gamma)$  amounts to a bundle isomorphism  $\Phi: E_\infty \rightarrow \hat{E}|_{\varphi(B_\varepsilon \setminus \{0\})/\Gamma}$  and the fact that the connection matrix representing  $\hat{A}$  vanishes at  $\infty = \varphi(0)$  implies that  $\nabla^k(\Phi^*(\hat{A}) - \theta) = O(|x|^{1-k})$  for all  $k \geq 0$ . By considering the action of the inversion  $x \mapsto \sigma(x)/|x|^2$  on  $k$ -fold derivatives of 1-forms one sees that  $\nabla^k(\Phi^*A - \theta) = O(r^{-3-k})$ .  $\square$

Let us briefly discuss moduli spaces of framed ASD instantons on  $E$  asymptotic to  $\theta$ . For a detailed discussion we refer the reader to Nakajima's beautiful article [Nak90]. Fix a framing at infinity  $\Phi$  of  $E$ , a rate  $\delta \in (-3, -1)$  and denote by  $\mathcal{A}(E, \theta)$  the space of all connections asymptotic to  $\theta$  at rate  $\delta$  with respect to  $\Phi$ . Similarly, define  $\mathcal{G}(E)$  to be the group of gauge transformations asymptotic to a constant element of  $G$  at infinity at rate  $\delta + 1$  with respect to  $\Phi$ . Denote by  $g_\infty: \mathcal{G}(E) \rightarrow G$  the homomorphism assigning to each gauge transformation its asymptotic value at infinity and let  $\mathcal{G}_0(E) := \ker g_\infty \subset \mathcal{G}(E)$  be the based gauge group consisting of gauge transformations asymptotic to the identity. Then the space

$$M(E, \theta) := \{A \in \mathcal{A}(E, \theta) : F_A^+ = 0\} / \mathcal{G}_0(E)$$

is called the *moduli space of framed ASD instantons on  $E$  asymptotic to  $\theta$* .

*Remark 5.6.* The space does not depend on the choice of  $\delta \in (-3, -1)$ . This is a consequence of Proposition 5.5.

*Remark 5.7.* If we denote by  $\rho: \Gamma \rightarrow G$  the monodromy representation associated with  $\theta$  and by  $G_\rho := \{g \in G : g\rho g^{-1} = \rho\}$  the stabiliser of  $\rho$ , then  $G_\rho \subset G \cong \mathcal{G}(E)/\mathcal{G}_0(E)$  acts on  $M(E, \theta)$ .

**Theorem 5.8** ([Nak90, Theorem 2.6 and Proposition 5.1]). *The moduli space  $M(E, \theta)$  is a smooth hyperkähler manifold.*

Formally, this can be seen as an infinite-dimensional instance of a hyperkähler reduction (see Hitchin, Karlhede, Lindström, and Roček [HKLR87]). The space  $\mathcal{A}(E, \theta)$  inherits a hyperkähler structure from  $X$  and the action of the based gauge group  $\mathcal{G}_0$  has a hyperkähler moment map given by  $\mu(A) = F_A^+$ . To make this rigorous one needs to set up a suitable Kuranishi model for  $M(E, \theta)$  along the lines of Donaldson and Kronheimer [DK90, Section 4.2.5]. This can be done using weighted Sobolev space completions of  $\mathcal{A}(E, \theta)$  and  $\mathcal{G}_0(E)$ ; see Nakajima [Nak90, Section 2] for a detailed discussion. An important role is played by the operator  $\delta_A: \Omega^1(X, \mathfrak{g}_E) \rightarrow \Omega^0(X, \mathfrak{g}_E) \oplus \Omega^+(X, \mathfrak{g}_E)$  defined by

$$(5.9) \quad \delta_A(a) := (d_A^*a, d_A^+a)$$

which governs the infinitesimal deformation theory of the ASD instanton  $A$ .

**Proposition 5.10.** *Let  $A \in \mathcal{A}(E)$  be a finite energy ASD instanton on  $E$ . Then the following holds.*

1. *If  $a \in \ker \delta_A$  decays to zero at infinity, then  $\nabla_A^k a = O(|\pi|^{-3-k})$  for all  $k \geq 0$ .*
2. *If  $(\xi, \omega) \in \ker \delta_A^*$  decays to zero at infinity, then  $(\xi, \omega) = 0$ .*

*Remark 5.11.* From the second part of this proposition one can deduce that the deformation theory of framed finite energy ASD instantons is always unobstructed; hence,  $M(E, \theta)$  is a smooth manifold (see also [Nak90, Proposition 5.1]). By the first part the tangent space of  $M(E, \theta)$  at  $[A]$  agrees with the  $L^2$  kernel of  $\delta_A$  and thus the formal hyperkähler structure is indeed well-defined.

The proof of Proposition 5.10 rests on the following refined Kato inequality.

**Proposition 5.12.** *Let  $A \in \mathcal{A}(E)$  be an ASD instanton on  $E$ . If  $a \in \Omega^1(X, \mathfrak{g}_E)$  satisfies  $\delta_A a = 0$ , then*

$$(5.13) \quad |d|a|| \leq \sqrt{\frac{3}{4}} |\nabla_A a|$$

*on the complement of the vanishing locus of  $a$ .*

*Proof.* Recall that the Kato inequality follows from the Cauchy–Schwarz inequality  $|\langle \nabla_A a, a \rangle| \leq |\nabla_A a| |a|$ . If  $\delta_A a = 0$ , then it is not hard to see that equality can only hold if  $\nabla_A a = 0$ . This shows that (5.13) holds with some constant  $\varepsilon < 1$  instead of  $\sqrt{3/4}$ .

To see that one can take  $\varepsilon = \sqrt{3/4}$  we follow an argument of Feehan [Fee01, Section 3]; however, also note that we could simply read off the value from the table given in Calderbank, Gauduchon, and Herzlich [CGH00, Appendix]. We can write  $\delta_A$  as a Dirac-type operator

$$\delta_A a = \sum_i \gamma(e_i) \nabla_{e_i}^A a.$$

Here  $(e_i)$  is a local orthonormal frame and the Clifford multiplication  $\gamma$  is defined by  $\gamma(v)a := (-i_v a, (v^* \wedge a)^+)$ , where  $v^*$  denotes the dual of  $v$  with respect to the metric on  $X$ . For  $x \in X$  with  $a(x) \neq 0$  and  $d|a|(x) \neq 0$  pick an orthonormal basis  $(e_i)$  of  $T_x X$  with  $e_1 := \nabla|a|/|\nabla|a||$ . Since  $\delta_A a = 0$  and  $|\gamma(v)a| = |v||a|$ , we have

$$\begin{aligned} |d|a||^2 &= |\nabla_{e_1} |a||^2 \leq |\nabla_{e_1}^A a|^2 \\ &= |\gamma(e_1) \nabla_{e_1}^A a|^2 = \left| \sum_{i \geq 2} \gamma(e_i) \nabla_{e_i}^A a \right|^2 \leq 3 \sum_{i \geq 2} |\nabla_{e_i}^A a|^2 \end{aligned}$$

and therefore

$$4|d|a||^2 = 4|\nabla_{e_1}^A a|^2 \leq 3 \sum_i |\nabla_{e_i}^A a|^2 = 3|\nabla_A a|^2.$$

This finishes the proof. □

*Proof of Proposition 5.10.* First of all note that (1) implies (2): if  $\delta_A^*(\xi, \omega) = 0$ , then  $d_A^* d_A \xi = 0$  and  $d_A^+ d_A \xi = [F_A^+, \xi] = 0$ ; therefore  $d_A \xi = O(|\pi|^{-3})$ . Thus integration by parts yields  $d_A \xi = 0$  and, hence,  $\xi = 0$ . Similarly, one shows that  $\omega = 0$ .

We will first explain why (1) for  $k = 0$  implies the asserted estimates for  $k > 0$  as well. The argument is similar to that in Proposition 4.20. For  $x \in X$  set  $R := \frac{1}{2}(1 + |\pi(x)|)$ . We claim that there is a constant  $c = c(k) > 0$  independent of  $x \in X$  such that

$$(5.14) \quad R^k \|\nabla_A^k a\|_{L^\infty(B_{R/2}(x))} \leq c \|a\|_{L^\infty(B_R(x))}$$

for all  $a \in \ker \delta_A$ . This clearly implies (1) for  $k > 0$  given the statement for  $k = 0$ . For  $|\pi(x)|$  sufficiently large, say  $|\pi(x)| > R_0$ , the restriction of  $A$  to  $B_R(x)$  is arbitrarily close to a flat connection by Proposition 5.5. We rescale to a ball of radius one and denote the rescaled connection by  $\tilde{A}$  and the rescaling of  $\delta_A$  by  $\tilde{D}_x$ . Then the family of operators  $\tilde{D}_x$  is uniformly elliptic with coefficients uniformly bounded in  $C^1$ . Therefore, there is a constant  $c > 0$  independent of  $x \in X$  such that the following Schauder estimates holds:

$$\|\nabla_{\tilde{A}}^k a\|_{L^\infty(B_{1/2})} \leq c (\|\tilde{D}_x a\|_{C^{k,\alpha}(B_1)} + \|a\|_{L^\infty(B_1)}).$$

If  $a$  is in the kernel of  $\tilde{D}_x$ , the first term vanishes. Rescaling this inequality yields (5.14) for  $a \in \ker \delta_A$  and  $|\pi(x)| > R$ . For  $1/2 \leq |\pi(x)| \leq R_0$ , (5.14) follows from standard Schauder estimates.

We now prove (1) for  $k = 0$ . Recall, for example, from Freed and Uhlenbeck [FU91, Equation (6.25)], that the operator  $\tilde{\delta}_A: \Omega^1(X, \mathfrak{g}_E) \rightarrow \Omega^0(X, \mathfrak{g}_E) \oplus \Omega^+(X, \mathfrak{g}_E)$  defined by  $\tilde{\delta}_A(a) := (d_A^* a, \sqrt{2} d_A^+ a)$  satisfies a Weitzenböck formula of the form

$$(5.15) \quad \tilde{\delta}_A^* \tilde{\delta}_A a = \nabla_A^* \nabla_A a + \{\text{Ric}, a\} + \{F_A^-, a\}.$$

Here  $\{\cdot, \cdot\}$  denote certain universal bilinear forms, whose precise form, however, is not important for our purposes and Ric denotes the Ricci tensor of  $X$ . In our situation, since  $X$  is hyperkähler and thus Ricci flat, the second term vanishes. Now, suppose that  $\delta_A a = 0$  and thus  $\tilde{\delta}_A a = 0$ . Then Proposition 5.12, the identity

$$\Delta |a|^2 + 2|\nabla_A a|^2 = 2 \langle a, \nabla_A^* \nabla_A a \rangle$$

(see [FU91, Equation (6.18)]) and the Weitzenböck formula (5.15) yield the following estimate on the complement of the vanishing locus of  $a$ :

$$\begin{aligned} 3\Delta |a|^{2/3} &\leq |a|^{-4/3} (\Delta |a|^2 + \frac{8}{3} |d|a|^2) \\ &\leq |a|^{-4/3} (\Delta |a|^2 + 2|\nabla_A a|^2) \\ &= 2|a|^{-4/3} \langle a, \nabla_A^* \nabla_A a \rangle \\ &= 2|a|^{-4/3} (\langle \tilde{\delta}_A^* \tilde{\delta}_A a, a \rangle + \langle \{F_A^-, a\}, a \rangle) \\ &\leq O(|\pi|^{-4}) |a|^{2/3}. \end{aligned}$$

In the last step we used  $\tilde{\delta}_A a = 0$  and  $|F_A^-| = O(|\pi|^{-4})$ , which is a consequence of Proposition 5.5.

Now, let  $U := \{x \in X : a(x) \neq 0\}$  and set  $f := |a|^{2/3}$ . We will show that  $f = O(|\pi|^{-2})$  which is equivalent to the desired decay estimate for  $a$ . It follows from the above that on  $U$ ,

$$\Delta f \leq \frac{cf}{1 + |\pi|^4}$$

for some constant  $c > 0$ . Since  $f$  is bounded, by [Joy00, Theorem 8.3.6(a)], there is a  $g = O(|\pi|^{-1})$  such that

$$\Delta g = \begin{cases} (\Delta f)^+ & \text{on } U \\ 0 & \text{on } X \setminus U. \end{cases}$$

Here  $(\cdot)^+$  denotes taking the positive part. Since  $g$  is superharmonic and decays to zero at infinity, the maximum principle implies that  $g$  is non-negative. The function  $f - g$  is a subharmonic on  $U$ , decays to zero at infinity and is non-positive on the boundary of  $U$ ; hence, by the maximum principle  $f \leq g$  and thus  $f \leq g = O(|\pi|^{-1})$ . Now,  $(\Delta f)^+ = O(|\pi|^{-5})$  on  $U$  and an application of [Joy00, Theorem 8.3.6(b)] shows that we could, in fact, have chosen  $g$  such that  $g = O(|\pi|^{-2})$ . It follows that  $f = O(|\pi|^{-2})$  as desired.  $\square$

The dimension of  $M(E, \theta)$  can be computed using the following index formula.

**Theorem 5.16** (Nakajima [Nak90, Theorem 2.7]). *Let  $A$  be a framed ASD instanton asymptotic to  $\theta$ . Then the dimension of the  $L^2$  kernel of  $\delta_A$  is given by*

$$(5.17) \quad \dim \ker \delta_A = -2 \int_X p_1(\mathfrak{g}_E) + \frac{2}{|\Gamma|} \sum_{g \in \Gamma \setminus \{e\}} \frac{\chi_{\mathfrak{g}}(g) - \dim \mathfrak{g}}{2 - \text{tr } g}.$$

Here  $p_1(\mathfrak{g}_E)$  is the Chern–Weil representative of the first Pontryagin class of  $E$  and  $\chi_{\mathfrak{g}}$  is the character of  $\Gamma$  acting on  $\mathfrak{g}$ , the Lie algebra associated with  $G$ , via the monodromy representation  $\rho: \Gamma \rightarrow G$  of  $\theta$ .

*Proof.* Let us briefly explain how to derive (5.17) from Nakajima’s formula, which can be written as

$$(5.18) \quad \dim \ker \delta_A = - \int_X (\dim \mathfrak{g} + p_1(\mathfrak{g}_E)) \text{ch}(S^+) \hat{A}(X) + \dim \mathfrak{g}^\Gamma + \frac{1}{|\Gamma|} \sum_{g \in \Gamma \setminus \{e\}} \chi_{\mathfrak{g}}(g) \frac{\text{tr } g}{2 - \text{tr } g}.$$

Here  $\mathfrak{g}^\Gamma$  denotes the  $\Gamma$ -invariant part of  $\mathfrak{g}$ ,  $S^+$  denotes the positive spin bundle on  $X$ , and  $\text{ch}(S^+)$  and  $\hat{A}(X)$  denote the Chern–Weil representatives of the Chern character of  $S^+$  and the  $\hat{A}$ -genus of  $X$ , respectively.

If  $A$  is the product connection on the trivial bundle rank 1 bundle and  $a$  lies in the  $L^2$  kernel of  $\delta_A$ , then it follows from the fact that  $X$  is Ricci-flat and the Weitzenböck formula (5.15) that  $\nabla^* \nabla a = 0$  and then by integration by parts, which is justified because of the decay asserted by

Proposition 5.10, that  $\nabla a = 0$ . Since  $a$  lies in  $L^2$ , it necessarily vanishes. Therefore  $\dim \ker \delta_A = 0$  and (5.18) yields

$$\int_X \text{ch}(S^+) \hat{A}(X) = 1 + \frac{1}{|\Gamma|} \sum_{g \in \Gamma \setminus \{e\}} \frac{\text{tr } g}{2 - \text{tr } g}.$$

By plugging this back into (5.18) we obtain

$$\begin{aligned} \dim \ker \delta_A &= -2 \int_X p_1(\mathfrak{g}_E) + \dim \mathfrak{g}^\Gamma - \dim \mathfrak{g} \\ &\quad + \frac{1}{|\Gamma|} \sum_{g \in \Gamma \setminus \{e\}} (\chi_g(\mathfrak{g}) - \dim \mathfrak{g}) \frac{\text{tr } g}{2 - \text{tr } g}. \end{aligned}$$

Since

$$\frac{1}{|\Gamma|} \sum_{g \in \Gamma} (\chi_g(\mathfrak{g}) - \dim \mathfrak{g}) = \dim \mathfrak{g}^\Gamma - \dim \mathfrak{g},$$

this leads to the index formula (5.17) given above.  $\square$

There is a very rich existence theory for ASD instantons on ALE spaces. Gocho–Nakajima [GN92] observed that for each representation  $\rho: \Gamma \rightarrow \text{U}(n)$  there is a bundle  $\mathcal{R}_\rho$  over  $X$  together with an ASD instanton  $A_\rho$  asymptotic to the flat connection determined by  $\rho$ , and if  $\sigma$  is a further representation of  $\Gamma$ , then  $A_{\rho \oplus \sigma} = A_\rho \oplus A_\sigma$ . Kronheimer–Nakajima [KN90] took this as the starting point for an ADHM construction of ASD instantons on ALE spaces. One important consequence of their work is the following rigidity result.

**Definition 5.19.** An ASD instanton  $A$  is called *infinitesimally rigid* if the  $L^2$  kernel of the linear operator  $\delta_A$  is trivial.

**Theorem 5.20** ([KN90, Lemma 7.1]). *For each  $\rho: \Gamma \rightarrow \text{U}(n)$  the ASD instanton  $A_\rho$  is infinitesimally rigid.*

By combining this result applied to the regular representation with the index formula Kronheimer–Nakajima derive a geometric version of the McKay correspondence [KN90, Appendix A]. Let  $\Delta(\Gamma)$  denote the Dynkin diagram associated with  $\Gamma$  in the ADE classification of the finite subgroups of  $\text{SU}(2)$ . Each vertex of  $\Delta(\Gamma)$  corresponds to a non-trivial irreducible representation. We label these by  $\rho_1, \dots, \rho_k$  and denote the associated bundles by  $\mathcal{R}_j$  and the associated ASD instantons by  $A_j$ .

**Theorem 5.21** ([KN90, Appendix A]). *The harmonic 2-forms  $c_1(\mathcal{R}_j) = \frac{i}{2\pi} \text{tr } F_{A_j}$  form a basis of  $L^2 \mathcal{H}^2(X) \cong H^2(X, \mathbf{R})$  and satisfy*

$$\int_X c_1(\mathcal{R}_i) \wedge c_1(\mathcal{R}_j) = -(C^{-1})_{ij},$$

where  $C$  is the Cartan matrix associated with  $\Delta(\Gamma)$ . Moreover, there is an isometry  $\kappa \in \text{Aut}(H_2(X, \mathbf{Z}), \cdot)$  such that  $\{c_1(\mathcal{R}_j)\}$  is dual to  $\{\kappa[\Sigma_j]\}$ , where  $\Sigma_j$  are the irreducible components of the exceptional divisor  $E$  of  $\widetilde{X/\Gamma}$ . If  $X$  is isomorphic to  $\widetilde{X/\Gamma}$  as a complex manifold, then  $\kappa = \text{id}$ .

This result is very useful for computing the index of  $\delta_A$  when  $A$  is constructed out of ASD instantons of the form  $A_\rho$  (by taking tensor products, direct sums, etc.).

**Proposition 5.22.** *Let  $X$  be an ALE space asymptotic to  $\mathbf{C}^2/\mathbf{Z}_k$ . Denote by  $\rho_j: \mathbf{Z}_k \rightarrow U(1)$  the irreducible representation defined by  $\rho_j(\ell) = \exp(\frac{2\pi i}{k} j\ell)$ . For  $n, m \in \mathbf{Z}_k$ , let  $E_{n,m}$  be the  $\mathrm{SO}(3)$ -bundle underlying  $\mathbf{R} \oplus (\mathcal{R}_n^* \otimes \mathcal{R}_{n+m})$  and denote by  $A_{n,m}$  the ASD instanton on  $E_{n,m}$  induced by  $A_n$  and  $A_{n+m}$ . Then  $A_{n,m}$  is infinitesimally rigid, asymptotic at infinity to the flat connection associated with  $\rho_m$  and*

$$\frac{1}{8\pi^2} \int_X |F_{A_{n,m}}|^2 = \frac{(k-m)m}{k}$$

as well as

$$w_2(\mathfrak{g}_{E_{n,m}}) = c_1(\mathcal{R}_{n+m}) - c_1(\mathcal{R}_n) \in H^2(X, \mathbf{Z}_2).$$

*Proof.* To see that  $A_{n,m}$  is infinitesimally rigid apply Theorem 5.20 to  $A_n \oplus A_{n+m}$  and observe that  $\mathfrak{g}_{E_{n,m}} = \mathbf{R} \oplus (\mathcal{R}_n^* \otimes \mathcal{R}_{n+m})$  is a parallel subbundle of  $\mathfrak{g}_{\mathcal{R}_n \oplus \mathcal{R}_{n+m}}$ .

The energy of  $A_{n,m}$  can be computed using Theorem 5.21 or by noting that the first term in the index formula (5.17) is precisely twice the energy and the second term is given by  $(-\frac{2}{k})$ -times

$$-\sum_{g \neq e} \frac{\chi_g(g) - \dim g}{2 - \mathrm{tr} g} = \sum_{j=1}^{k-1} \frac{1 - \cos(2\pi m j/k)}{1 - \cos(2\pi j/k)} = (k-m)m.$$

The statement about the second Stiefel–Whitney class is clear. □

## 6 Approximate $G_2$ -instantons

Throughout this section, let  $Y_0$  be an admissible  $G_2$ -orbifold, let  $\mathbf{r} = \{(X_j, \rho_j)\}$  be resolution data for  $Y_0$  and denote by  $(Y_t, \phi_t)_{t \in (0, T)}$  the family of  $G_2$ -manifolds obtained from  $\mathbf{r}$  via Theorem 4.16. Denote by  $\psi_t := \Theta(\phi_t)$  the coassociative calibration on  $Y_t$ . If  $\theta$  is a flat connection on a  $G$ -bundle  $E_0$  over  $Y_0$ , then the monodromy of  $\theta$  around  $S_j$  induces a representation  $\mu_j: \pi_1(T_j, x_j) \rightarrow G$  of the orbifold fundamental group of  $T_j$  based at  $x_j \in T_j \setminus S_j$ . We set

$$K_j := \pi_1(T_j, x_j)/G_j.$$

Both  $\pi_1(T_j)$  and  $K_j$  are extensions of  $H_j$  (by  $\mathbf{Z}^3 \times G_j$  and  $\mathbf{Z}^3$  respectively); however, they usually do not split.

*Remark 6.1.* For a general definition of orbifold fundamental group we refer the reader to Adem, Leida, and Ruan [ALR07, Definition 1.50 and Section 2.2]. All orbifold fundamental groups  $\pi_1(X)$  encountered in this article can be identified with the fundamental groups  $\pi_1(X^{\mathrm{reg}})$  of the regular part of the orbifold in question. This is a consequence of the next proposition which is a special case of [TT04, Lemma 5.2].

**Proposition 6.2.** *If  $B$  is an open ball in  $\mathbf{R}^n$  and  $S \subset B$  is a closed subset of Hausdorff codimension at least four, then  $\pi_1(B \setminus S) = 1$ .*

*Proof.* Note that for any given Lipschitz loop  $\gamma : S^1 \rightarrow B \setminus S$ , the set of  $x \in B \setminus S$  such that the cone  $C(\gamma, x) = \{\ell_{x, \gamma(s)}(t) : s, t \in [0, 1]\}$ , where  $\ell_{x, y}(t) = (1-t)x + ty$ , is contained in  $B \setminus S$  has Hausdorff codimension at least 2. Since any continuous loop is homotopic to a Lipschitz loop in  $B \setminus S$  it follows that any continuous loop is homotopic to a constant one.  $\square$

**Definition 6.3.** A collection  $\mathbf{g} = ((E_0, \theta), \{(x_j, f_j)\}, \{(E_j, A_j, \tilde{\rho}_j)\})$  consisting of  $E_0$  and  $\theta$  as above as well as, for each  $j$ , the choice of

- a point  $x_j \in T_j \setminus S_j$  together with a framing  $f_j : (E_0)_{x_j} \rightarrow G$  of  $E_0$  at  $x_j$ ,
- a  $G$ -bundle  $E_j$  over  $X_j$  together with a framed ASD instanton  $A_j$  asymptotic at infinity to the flat connection on the bundle  $E_{\infty, j}$  over  $(\mathbf{C}^2 \setminus \{0\})/G_j$  induced by the representation  $\mu_j|_{G_j}$  and
- an action  $\tilde{\rho}_j$  of  $K_j$  on  $E_j$  covering the action of  $K_j$  on  $X_j$  induced by  $\rho_j$

is called *gluing data compatible with  $\mathbf{r} = \{(X_j, \rho_j)\}$*  if

- the action  $\tilde{\rho}_j$  of  $K_j$  on  $E_j$  preserves  $A_j$  and is asymptotic at infinity, with respect to the framing associated with  $A_j$ , to the action of  $K_j$  on  $E_{\infty, j}$ . Note that the action of  $H_j$  on  $E_{\infty, j}$  is induced from the action of  $\pi_1(T_j, x_j)$  on the trivial bundle  $G \times (\mathbf{C}^2 \setminus \{0\})$  given by  $h \cdot (g, x) = ((\mu_j(h) \cdot g, h \cdot x))$ .

We should point out here that it is by far not always possible to extend a choice of  $(E_0, \theta)$  and  $\{(E_j, A_j)\}$  to compatible gluing data: the choice of  $\tilde{\rho}_j$  is essentially dictated by  $\theta$ ; however, we cannot always find an  $A_j$  which is preserved by  $\tilde{\rho}_j$ —in other words: the bundle  $\mathfrak{M}$  of moduli space of ASD instantons has no section at all. This will become clearer from the discussion in Section 9.

Before we proceed to construct approximate  $G_2$ -instantons, we introduce weighted Hölder norms. It will become more transparent over the course of the next two sections that these are well adapted to the problem at hand. We define weight functions by

$$w_t(x) := t + r_t(x) \quad \text{and} \quad w_t(x, y) := \min\{w_t(x), w_t(y)\}.$$

For  $t \in (0, T)$ , a Hölder exponent  $\alpha \in (0, 1)$  and a weight parameter  $\beta \in \mathbf{R}$  we define

$$\begin{aligned} [f]_{C_{\beta, t}^{0, \alpha}(U)} &:= \sup_{d(x, y) \leq w_t(x, y)} w_t(x, y)^{\alpha - \beta} \frac{|f(x) - f(y)|}{d(x, y)^\alpha}, \\ \|f\|_{L_{\beta, t}^\infty(U)} &:= \|\mathbf{w}_t^{-\beta} f\|_{L^\infty(U)}, \\ \|f\|_{C_{\beta, t}^{k, \alpha}(U)} &:= \sum_{j=0}^k \|\nabla^j f\|_{L_{\beta-j, t}^\infty(U)} + [\nabla^j f]_{C_{\beta-j, t}^{0, \alpha}(U)}. \end{aligned}$$

Here  $f$  is a section of a vector bundle over  $U \subset Y_t$  equipped with an inner product and a compatible connection. On tensor bundles associated with  $Y_t$  we use the metrics induced by  $\tilde{g}_t$ ; however, in view of Proposition 4.20, we could equivalently use those induced by  $\phi_t = \tilde{\phi}_t + d\eta_t$ . We use parallel transport to compare the value of  $f$  at different points in  $Y$ . If  $U$  is not specified, then we take  $U = Y_t$ . We denote by  $C_{\beta,t}^{k,\alpha}$  the Banach space  $C^{k,\alpha}$  equipped with the norm  $\|\cdot\|_{C_{\beta,t}^{k,\alpha}}$

*Remark 6.4.* For fixed  $t \in (0, T)$  and  $\beta \in \mathbf{R}$ , the norms  $\|\cdot\|_{C_{\beta,t}^{k,\alpha}}$  and  $\|\cdot\|_{C^{k,\alpha}}$  are equivalent, but not uniformly so as  $t > 0$  tends to zero.

Note that, if  $\beta = \beta_1 + \beta_2$ , then

$$(6.5) \quad \|f \cdot g\|_{C_{\beta,t}^{k,\alpha}} \leq \|f\|_{C_{\beta_1,t}^{k,\alpha}} \cdot \|g\|_{C_{\beta_2,t}^{k,\alpha}}.$$

Also for  $\beta > \gamma$  we have

$$(6.6) \quad \|f\|_{C_{\beta,t}^{k,\alpha}} \leq t^{\gamma-\beta} \|f\|_{C_{\gamma,t}^{k,\alpha}}.$$

**Proposition 6.7.** *Let  $\mathbf{g}$  be gluing data compatible with  $\mathbf{r}$ . Then there is a constant  $c > 0$  and for each  $t \in (0, T)$  a  $G$ -bundle  $E_t$  over  $Y_t$  together with a connection  $\tilde{A}_t$  satisfying*

$$(6.8) \quad \|F_{\tilde{A}_t} \wedge \psi_t\|_{C_{-2,t}^{0,\alpha}} \leq ct^{1/2}.$$

Moreover, the adjoint bundle  $\mathfrak{g}_{E_t}$  associated with  $E_t$  satisfies

$$(6.9) \quad p_1(\mathfrak{g}_{E_t}) = - \sum_j k_j \text{PD}[S_j] \quad \text{with } k_j := \frac{1}{8\pi^2} \int_{X_j} |F_{A_j}|^2$$

and

$$(6.10) \quad \langle w_2(\mathfrak{g}_{E_t}), [\Sigma] \rangle = \langle w_2(\mathfrak{g}_{E_j}), [\Sigma] \rangle$$

for each  $[\Sigma] \in H_2(X_j)^{H_j} \subset H_2(Y_t)$ .

*Proof.* The choices of  $\tilde{\rho}_j$  and  $m_j$  define a lift of the action of  $\mathbf{Z}^3 \rtimes H_j$  on  $\mathbf{R}^3 \times X_j$  to the pullback of  $E_j$  to  $\mathbf{R}^3 \times X_j$ . Passing to the quotient yields a  $G$ -bundle over  $(T^3 \times X_j)/H_j$  which we denote by  $E_j$ , by abuse of notation. It follows from the compatibility conditions that the pullback of  $A_j$  to  $\mathbf{R}^3 \times X_j$  passes to the quotient and induces a connection on  $E_j$  which we denote by  $A_j$ , again by abuse of notation.

Fix  $t \in (0, T)$ . Recall that in (4.14) we defined  $R_{j,t} := \tilde{T}_{j,t} \cap r_t^{-1}[\zeta/4, \zeta/2]$  with  $\tilde{T}_{j,t}$  and  $r_t$  as defined in (4.9) and (4.13), respectively. By the compatibility conditions the monodromy of  $A_j$  along  $S_j$  on the fibre at infinity matches up with the monodromy of  $\theta$  along  $E_0|_{S_j}$ . Thus, via parallel transport the framing of  $E_0$  at  $x_j$  and the framing of  $E_j$  yield an identification of  $E_0|_{R_{j,t}}$  with  $E_j|_{R_{j,t}}$ . Patching  $E_0$  and the  $E_j$  via this identification yields the bundle  $E_t$ .

Under the identification of  $E_0|_{R_{j,t}}$  with  $E_j|_{R_{j,t}}$ , we can write

$$(6.11) \quad A_j = \theta + a_j \quad \text{with} \quad \nabla^k a_j = t^{2+k} O(r_t^{-3-k}),$$

because of Remark 4.15 and Proposition 5.10. Fix a smooth non-increasing function  $\chi: [0, \zeta] \rightarrow [0, 1]$  such that  $\chi(s) = 1$  for  $s \leq \zeta/4$  and  $\chi(s) = 0$  for  $s \geq \zeta/2$ . Set  $\chi_t := \chi \circ r_t$ . After cutting off  $A_j$  to  $\theta + \chi_t \cdot a_j$  it can be matched with  $\theta$  and we obtain the connection  $\tilde{A}_t$  on the bundle  $E_t$ .

To estimate  $F_{\tilde{A}_t} \wedge \psi_t$  note that on  $Y_t \setminus \tilde{T}_t$  the connection  $\tilde{A}_t$  is flat. Thus we can focus our attention on  $\tilde{T}_{j,t}$ . By the definition of  $\tilde{A}_t$  we have

$$F_{\tilde{A}_t} = \chi_t F_{A_j} + d\chi_t \wedge a_j + \frac{\chi_t^2 - \chi_t}{2} [a_j \wedge a_j].$$

The last two terms in this expression are supported in  $R_{j,t}$  and of order  $t^2$  in  $C^{0,\alpha}$  by (6.11). By Example 3.4 and Proposition 4.20 we have

$$\|F_{A_j} \wedge \psi_t\|_{C_{-2,t}^{0,\alpha}(\tilde{T}_{j,t})} = \|F_{A_j} \wedge (\psi_t - \hat{\psi}_t)\|_{C_{-2,t}^{0,\alpha}(\tilde{T}_{j,t})} \leq ct^{1/2} \|F_{A_j}\|_{C_{-2,t}^{0,\alpha}}.$$

It follows from Proposition 5.5 and Remark 4.15 that

$$\nabla^k F_{A_j} = t^{2+k} O(r_t^{-4-k}).$$

This implies that

$$\|F_{A_j}\|_{C_{-4,t}^{0,\alpha}(\tilde{T}_{j,t})} \leq ct^2$$

and, hence,

$$\|F_{A_j}\|_{C_{-2,t}^{0,\alpha}(\tilde{T}_{j,t})} \leq c$$

by (6.6) with  $c > 0$  independent of  $t \in (0, T)$ . Now, putting everything together yields (6.8).

Let  $\iota_{j,t}: T^3 \times \pi_{j,t}^{-1}(B_\zeta^4/G_j) \rightarrow Y$  be as in Remark 4.11. Then  $\iota_{j,t}^* \mathfrak{g}_{E_t}$  is isomorphic to the pullback of  $\mathfrak{g}_{E_j}$  to  $T^3 \times \pi_{j,t}^{-1}(B_\zeta^4/G_j)$ . This implies (6.10) by naturality of Stiefel–Whitney classes. To compute  $p_1(\mathfrak{g}_{E_t})$  we use Chern–Weil theory to represent it as  $p_1(\mathfrak{g}_{E_t}) = -\frac{1}{8\pi^2} \text{tr}(F_{\tilde{A}_t} \wedge F_{\tilde{A}_t})$ . We can write this as  $p_1(\mathfrak{g}_{E_t}) = \sum_j p_j$ , where  $p_j$  are compactly supported 4-forms on  $\tilde{T}_{j,t}$ . Recalling the definition of  $[S_j]$  in (4.12) and considering the behaviour of Poincaré duality with respect to coverings we see that in order to prove (6.9) we have to show

$$\iota_{j,t}^* p_j = k_j \text{PD}[T^3 \times \{x\}] \in H_c^4(T^3 \times \pi_{j,t}^{-1}(B_\zeta^4/G_j), \mathbf{R}).$$

From our construction of  $\tilde{A}_t$  it follows that the form  $\iota_{j,t}^* p_j$  is the pullback of a compactly supported 4-form on  $X_j$ , which we can write as  $-\frac{1}{8\pi^2} \text{tr}(F_{\tilde{A}_j} \wedge F_{\tilde{A}_j})$  where  $\tilde{A}_j = A_j + \alpha$  and, by slight abuse of notation,  $\alpha = (1 - \chi_t)a_j$ . Consequently,  $\iota_{j,t}^* p_j$  is a multiple of  $\text{PD}[T^3 \times \{x\}]$ . To see that the multiplicity is precisely  $k_j$  we use the Chern–Simons 3-form (see Donaldson and Kronheimer [DK90, Equation (2.1.17)]) to write

$$\text{tr}(F_{\tilde{A}_j} \wedge F_{\tilde{A}_j}) - \text{tr}(F_{A_j} \wedge F_{A_j}) = d \text{tr}(\alpha \wedge d_A \alpha + \frac{1}{3} \alpha \wedge [\alpha \wedge \alpha]).$$

By Proposition 5.10 the 1-form  $\alpha$  decays sufficiently fast to conclude from Stokes' theorem that

$$-\frac{1}{8\pi^2} \int_{X_j} \text{tr} (F_{\tilde{A}_j} \wedge F_{\tilde{A}_j}) = -\frac{1}{8\pi^2} \int_{X_j} \text{tr} (F_{A_j} \wedge F_{A_j}) = \frac{1}{8\pi^2} \int_{X_j} |F_{A_j}|^2 = k_j.$$

This completes the proof.  $\square$

*Remark 6.12.* If we identify all  $Y_t$  with one fixed  $Y$ , then the isomorphism type of the bundles  $E_t$  does not depend on  $t \in (0, T)$ . We can therefore think of them as one fixed  $G$ -bundle  $E$  over  $Y$ .

## 7 A model operator on $\mathbf{R}^3 \times \text{ALE}$

In order to prove Theorem 1.1 we need to find  $(\xi_t, a_t) \in (\Omega^0 \oplus \Omega^1)(Y_t, \mathfrak{g}_{E_t})$  such that

$$(7.1) \quad *_t (F_{\tilde{A}_t + a_t} \wedge \psi_t) + d_{\tilde{A}_t + a_t} \xi_t = 0$$

for  $t \in (0, T')$  provided  $T' \in (0, T]$  is sufficiently small. Here  $*_t$  denotes the Hodge  $*$ -operator associated with  $\phi_t$ . Equation (7.1) together with the Coulomb gauge condition  $d_{\tilde{A}_t}^* a_t = 0$  can be written as

$$(7.2) \quad L_t \underline{a}_t + Q_t(\underline{a}_t) + *_t (F_{\tilde{A}_t} \wedge \psi_t) = 0.$$

Here we use the notation  $\underline{a}_t := (\xi_t, a_t)$ , the linear operator  $L_t := L_{\tilde{A}_t}$  is defined as in (3.10) with  $\psi = \psi_t := *_t \phi_t$  and  $Q_t$  is defined by

$$(7.3) \quad Q_t(\underline{a}) := \frac{1}{2} *_t ([a \wedge a] \wedge \psi_t) + [a, \xi].$$

The key to solving (7.2) is a good understanding of the linearisation  $L_t$ . In this section, we study a model for  $L_t$  on  $r_t^{-1}([0, \zeta])$ .

Let  $X$  be an ALE space, let  $A$  be a  $G$ -bundle over  $X$  and let  $A$  be a finite energy ASD instanton on  $E$ . Fix an orthonormal triple  $(\delta^1, \delta^2, \delta^3)$  of constant 1-forms on  $\mathbf{R}^3$  and denote by  $(\omega_1, \omega_2, \omega_3)$  the triple of Kähler forms associated with  $X$ . Consider  $\mathbf{R}^3 \times X$  as a  $G_2$ -manifold as in Example 2.8. Denote by  $p_{\mathbf{R}^3}: \mathbf{R}^3 \times X \rightarrow \mathbf{R}^3$  and  $p_X: \mathbf{R}^3 \times X \rightarrow X$  the projection onto the first and second factor, respectively. Slightly abusing notation, we denote the respective pullbacks of  $E$  and  $A$  to  $\mathbf{R}^3 \times X$  via  $p_X$  by  $E$  and  $A$  as well. As in (3.10) we define  $L_A: \Omega^0(\mathbf{R}^3 \times X, \mathfrak{g}_E) \oplus \Omega^1(\mathbf{R}^3 \times X, \mathfrak{g}_E) \rightarrow \Omega^0(\mathbf{R}^3 \times X, \mathfrak{g}_E) \oplus \Omega^1(\mathbf{R}^3 \times X, \mathfrak{g}_E)$  by

$$L_A = \begin{pmatrix} 0 & d_A^* \\ d_A & *(\psi \wedge d_A) \end{pmatrix}$$

with  $\psi$  as in (2.10).

**Proposition 7.4.** *If we identify  $p_{\mathbf{R}^3}^* T^* \mathbf{R}^3$  with  $p_X^* \Lambda^+ T^* X$  via  $\delta^1 \mapsto \omega_1$ ,  $\delta^2 \mapsto \omega_2$ ,  $\delta^3 \mapsto -\omega_3$  and accordingly*

$$\begin{aligned} \Omega^0(\mathbf{R}^3 \times X, \mathfrak{g}_E) \oplus \Omega^1(\mathbf{R}^3 \times X, \mathfrak{g}_E) \\ = \Omega^0(\mathbf{R}^3 \times X, p_X^* [(\mathbf{R} \oplus \Lambda^+ T^* X \oplus T^* X) \otimes \mathfrak{g}_E]), \end{aligned}$$

then the operator  $L_A$  can be written as  $L_A = F + D_A$ , where

$$F(\xi, \omega, a) = \sum_{i=1}^3 (-\langle \partial_i \omega, \omega_i \rangle, \partial_i \xi \cdot \omega_i, I_i \partial_i a) \quad \text{and} \quad D_A = \begin{pmatrix} 0 & \delta_A \\ \delta_A^* & 0 \end{pmatrix}.$$

Here  $\delta_A: \Omega^1(X, \mathfrak{g}_E) \rightarrow \Omega^0(X, \mathfrak{g}_E) \oplus \Omega^+(X, \mathfrak{g}_E)$  denotes the linear operator defined in (5.9). Moreover,

$$(7.5) \quad L_A^* L_A = \Delta_{\mathbf{R}^3} + \begin{pmatrix} \delta_A \delta_A^* & \\ & \delta_A^* \delta_A \end{pmatrix}$$

where  $\Delta_{\mathbf{R}^3} = -\sum_{i=1}^3 \partial_i^2$  and  $\partial_i$  denotes taking the derivative of a section of  $p_X^* [(\mathbf{R} \oplus \Lambda^+ T^* X \oplus T^* X) \otimes \mathfrak{g}_E]$  in the direction of the  $i^{\text{th}}$  coordinate on  $\mathbf{R}^3$ .

*Proof.* It is a straight-forward computation to verify that  $L_A = F + D_A$ . It is also easy to see that  $F^* F = \Delta_{\mathbf{R}^3}$  and that  $F^* D_A + D_A^* F = 0$ . This immediately implies (7.5).  $\square$

To understand the properties of  $L_A$  we work with weighted Hölder norms. We define weight functions by

$$w(x) := 1 + |\pi(p_X(x))| \quad \text{and} \quad w(x, y) := \min\{w(x), w(y)\}.$$

Here  $\pi: X \rightarrow \mathbf{C}^2/G$  denotes the resolution map associated with the ALE space  $X$ . For a Hölder exponent  $\alpha \in (0, 1)$  and a weight parameter  $\beta \in \mathbf{R}$  we define

$$\begin{aligned} [f]_{C_\beta^{0, \alpha}(U)} &:= \sup_{d(x, y) \leq w(x, y)} w(x, y)^{\alpha - \beta} \frac{|f(x) - f(y)|}{d(x, y)^\alpha}, \\ \|f\|_{L_\beta^\infty(U)} &:= \left\| w^{-\beta} f \right\|_{L^\infty(U)}, \\ \|f\|_{C_\beta^{k, \alpha}(U)} &:= \sum_{j=0}^k \|\nabla^j f\|_{L_{\beta-j}^\infty(U)} + [\nabla^j f]_{C_{\beta-j}^{0, \alpha}(U)}. \end{aligned}$$

Here  $f$  is a section of a vector bundle over  $U \subset \mathbf{R}^3 \times X$  equipped with an inner product and a compatible connection. We use parallel transport to compare the values of  $f$  at different points. If  $U$  is not specified, then we take  $U = Y_t$ . We denote by  $C_\beta^{k, \alpha}$  the subspace of elements  $f$  of the Banach space  $C^{k, \alpha}$  with  $\|f\|_{C^{k, \alpha}} < \infty$  and equip it with the norm  $\|\cdot\|_{C_\beta^{k, \alpha}}$ .

Under the assumptions of Section 6 and with  $\mathfrak{g}$  denoting compatible  $\beta$  gluing data suppose that  $X = X_j$  and that  $A = A_j$ . Define  $\tilde{t}_{j, t}: \mathbf{R}^3 \times \pi_{j, t}^{-1}(B_\zeta^4/G_j) \rightarrow \tilde{T}_{j, t}$  by

$$\tilde{t}_{j, t}(x, y) := [(tx, y)].$$

For a parameter  $\beta \in \mathbf{R}$  and  $\underline{a} = (\xi, a) \in \Omega^0(Y_t, \mathfrak{g}_{E_t}) \oplus \Omega^1(Y_t, \mathfrak{g}_{E_t})$  we define

$$(7.6) \quad s_{\beta,t}(\xi, a)(x, y) := t^{\beta-1} (t(\tilde{l}_{j,t})^* \xi, (\tilde{l}_{j,t})^* a).$$

**Proposition 7.7.** *There is a constant  $c > 0$  such that for  $t \in (0, T)$*

$$\begin{aligned} \frac{1}{c} \|\underline{a}\|_{C_{\beta,t}^{k,\alpha}(\tilde{T}_{j,t})} &\leq \|s_{\beta,t} \underline{a}\|_{C_{\beta}^{k,\alpha}(\mathbf{R}^3 \times \pi_{j,t}^{-1}(B_{\xi}^4/G_j))} \leq c \|\underline{a}\|_{C_{\beta,t}^{k,\alpha}(\tilde{T}_{j,t})}, \\ \|L_t \underline{a} - s_{\beta-1,t}^{-1} L_{A_j} s_{\beta,t} \underline{a}\|_{C_{\beta-1,t}^{0,\alpha}(\tilde{T}_{j,t})} &\leq ct^{1/2} \|\underline{a}\|_{C_{\beta,t}^{1,\alpha}(\tilde{T}_{j,t})}. \end{aligned}$$

*Proof.* The map  $\tilde{l}_{j,t}$  pulls back the metric on  $\tilde{T}_{j,t}$  associated with  $\hat{\phi}_t$ , that is  $g_{\hat{\phi}_t} = g_{\mathbf{R}^3} \oplus t^2 g_{X_j}$ , to  $t^2(g_{\mathbf{R}^3} \oplus g_{X_j})$ . This implies the first estimate in view of Remark 4.15. The second estimate is immediate from the construction of  $\tilde{A}_t$  and Proposition 4.20.  $\square$

**Proposition 7.8.** *Let  $\beta \in (-3, 0)$ . Then  $\underline{a} \in C_{\beta}^{1,\alpha}$  is in the kernel of  $L_A: C_{\beta}^{1,\alpha} \rightarrow C_{\beta-1}^{0,\alpha}$  if and only if it is given by the pullback of an element of the  $L^2$  kernel of  $\delta_A$  to  $\mathbf{R}^3 \times X$ .*

The proof of Proposition 7.8 relies on the following lemma which we will prove in the Appendix A.

**Definition 7.9.** A Riemannian manifold  $X$  is said to be of *bounded geometry* if it is complete, its Riemann curvature tensor is bounded from above and its injectivity radius is bounded from below. A vector bundle over  $X$  is said to be of *bounded geometry* if it has trivialisations over balls of a fixed radius such that the transitions functions and all of their derivatives are uniformly bounded. We say that a complete oriented Riemannian manifold  $X$  has *subexponential volume growth* if for each  $x \in X$  the function  $r \mapsto \text{vol}(B_r(x))$  grows subexponentially, that is,  $\text{vol}(B_r(x)) = o(\exp(cr))$  as  $r \rightarrow \infty$  for every  $c > 0$ .

**Lemma 7.10.** *Let  $E$  be a vector bundle of bounded geometry over a Riemannian manifold  $X$  of bounded geometry and with subexponential volume growth, and suppose that  $D: C^\infty(X, E) \rightarrow C^\infty(X, E)$  is a uniformly elliptic operator of second order whose coefficients and their first derivatives are uniformly bounded, that is non-negative, such that  $\langle Da, a \rangle \geq 0$  for all  $a \in W^{2,2}(X, E)$ , and formally self-adjoint. If  $a \in C^\infty(\mathbf{R}^n \times X, E)$  satisfies*

$$(\Delta_{\mathbf{R}^n} + D)a = 0$$

*and  $\|a\|_{L^\infty}$  is finite, then  $a$  is constant in the  $\mathbf{R}^n$ -direction, that is  $a(x, y) = a(y)$ . Here, by slight abuse of notation, we denote the pullback of  $E$  to  $\mathbf{R}^n \times X$  by  $E$  as well.*

*Proof of Proposition 7.8.* Suppose  $\underline{a} \in C_{\beta}^{1,\alpha}$  satisfies  $L_A \underline{a} = 0$ . Then  $\underline{a}$  is smooth by elliptic regularity and satisfies  $L_A^* L_A \underline{a} = 0$ . By Definition 4.2 and by Proposition 5.5 both  $\mathbf{R}^3 \times X$  and  $\mathfrak{g}_E$  have bounded geometry. Moreover, by Proposition 7.4,  $L_A^* L_A = \Delta_{\mathbf{R}^3} + D_A^* D_A$  and  $D_A^* D_A$  is non-negative, self-adjoint, uniformly elliptic of second order and its coefficients and their first derivatives are uniformly bounded as can be seen from Proposition 5.5. Therefore, we can apply Lemma 7.10 to conclude that  $\underline{a}$  is invariant under translations in the  $\mathbf{R}^3$ -direction and, hence, by Proposition 5.10 and Proposition 7.4 must be the pullback of an element in the  $L^2$  kernel of  $\delta_A$ .  $\square$

**Proposition 7.11.** *For  $\beta \in \mathbf{R}$  there is a constant  $c > 0$  such that*

$$\|\underline{a}\|_{C_\beta^{1,\alpha}} \leq c \left( \|L_A \underline{a}\|_{C_{\beta-1}^{0,\alpha}} + \|\underline{a}\|_{L_\beta^\infty} \right).$$

*Proof.* This is a standard result; see Remark 4.27.

The desired estimate is local in the sense that is enough to prove estimates of the form

$$\|\underline{a}\|_{C_\beta^{1,\alpha}(U_i)} \leq c \left( \|L_A \underline{a}\|_{C_{\beta-1}^{0,\alpha}} + \|\underline{a}\|_{L_\beta^\infty} \right)$$

with  $c > 0$  independent of  $i$ , where  $\{U_i\}$  is a suitable open cover of  $\mathbf{R}^3 \times X$ .

Fix  $R > 0$  suitably large and set  $U_0 := \{(x, y) \in \mathbf{R}^3 \times X : |\pi(x)| \leq R\}$ . Then there clearly is a constant  $c > 0$  such that the above estimate holds for  $U_i = U_0$ . Pick a sequence  $(x_i, y_i) \in \mathbf{R}^3 \times X$  such that  $r_i := |\pi(y_i)| \geq R$  and the balls  $U_i := B_{r_i/8}(x_i, y_i)$  cover the complement of  $U_0$ . On  $U_i$ , we have a Schauder estimate of the form

$$\begin{aligned} \|\underline{a}\|_{L^\infty(U_i)} + r_i^\alpha [\underline{a}]_{C^{0,\alpha}(U_i)} + r_i \|\nabla_A \underline{a}\|_{L^\infty(U_i)} + r_i^{1+\alpha} [\nabla_A \underline{a}]_{C^{0,\alpha}(U_i)} \\ \leq c \left( r_i \|L_A \underline{a}\|_{L^\infty(V_i)} + r_i^{1+\alpha} [L_A \underline{a}]_{C^{0,\alpha}(V_i)} + \|\underline{a}\|_{L^\infty(V_i)} \right) \end{aligned}$$

where  $V_i = B_{r_i/4}(x_i, y_i)$  and  $\underline{a} = (\xi, a)$ . By arguing as in Proposition 4.20 and Proposition 5.10 one shows that the constant  $c > 0$  can be chosen to work for all  $i$  simultaneously. Since on  $V_i$  we have  $\frac{1}{2}r_i \leq w \leq 2r_i$ , multiplying the above Schauder estimate by  $r_i^{-\beta}$  yields the desired local estimate.  $\square$

## 8 Deforming to genuine $G_2$ -instantons

We continue with the assumptions of Section 6 and we suppose that the connection  $\tilde{A}_t$  on  $G$ -bundle  $E_t$  over  $Y_t$  was constructed using Proposition 6.7 from a choice of compatible gluing data  $\mathbf{g}$ . In this section we will prove the following result which will complete the proof of Theorem 1.1.

**Proposition 8.1.** *Suppose that  $\theta$  is acyclic and that each  $A_j$  is infinitesimally rigid. Then there are constants  $T' \in (0, T]$  and  $c > 0$  as well as, for each  $t \in (0, T')$ ,  $\underline{a}_t = (\xi_t, a_t) \in \Omega^0(Y_t, \mathfrak{g}_{E_t}) \oplus \Omega^1(Y_t, \mathfrak{g}_{E_t})$  such that*

$$(8.2) \quad *_t (F_{\tilde{A}_t + a_t} \wedge \psi_t) + d_{\tilde{A}_t + a_t} \xi_t = 0$$

and  $\|\underline{a}_t\|_{C_{-1,t}^{1,\alpha}} \leq ct^{1/2}$ . Moreover, the  $G_2$ -instanton  $A_t := \tilde{A}_t + a_t$  is acyclic.

As discussed in Section 7 it is crucial to understand the properties of the linear operator  $L_t$ . The key to proving Proposition 8.1 is the following result.

**Proposition 8.3.** *Given  $\beta \in (-3, 0)$  there are constants  $T' \in (0, T]$  and  $c > 0$  such that for  $t \in (0, T')$  we have*

$$\|\underline{a}\|_{C_{\beta,t}^{1,\alpha}} \leq c \|L_t \underline{a}\|_{C_{\beta-1,t}^{0,\alpha}}.$$

Before we move on to prove this, let us quickly show how it is used to establish Proposition 8.1. Recall the following elementary consequence of Banach's fixed point theorem.

**Lemma 8.4** ([DK90, Lemma 7.2.23]). *Let  $X$  be a Banach space and let  $T : X \rightarrow X$  be a smooth map with  $T(0) = 0$ . Suppose there is a constant  $c > 0$  such that*

$$\|Tx - Ty\| \leq c(\|x\| + \|y\|)\|x - y\|.$$

*Then if  $y \in X$  satisfies  $\|y\| \leq \frac{1}{10c}$ , there exists a unique  $x \in X$  with  $\|x\| \leq \frac{1}{5c}$  solving*

$$x + Tx = y.$$

*Moreover, this  $x \in X$  satisfies  $\|x\| \leq 2\|y\|$ .*

*Proof of Proposition 8.1 assuming Proposition 8.3.* By Proposition 8.3 the operator  $L_t : C_{-1,t}^{1,\alpha} \rightarrow C_{-2,t}^{0,\alpha}$  is injective and has closed range. Therefore its cokernel is isomorphic to the kernel of the dual operator  $L_t^*$ . By elliptic regularity any element in the kernel of  $L_t^*$  is smooth and thus, since  $L_t$  is formally self-adjoint, an element in the kernel of  $L_t$ , which is trivial. This shows that  $L_t$  is invertible. Denote its inverse by  $R_t : C_{-2,t}^{0,\alpha} \rightarrow C_{-1,t}^{1,\alpha}$ .

If we set  $\underline{a}_t := R_t \underline{b}_t$ , then (8.2) becomes

$$(8.5) \quad \underline{b}_t + Q_t(R_t \underline{b}_t) = - *_t (F_{\tilde{A}_t} \wedge \psi_t).$$

It follows from Proposition 8.3 and (6.5) that

$$\|Q_t(R_t \underline{b}_1) - Q_t(R_t \underline{b}_2)\|_{C_{-2,t}^{0,\alpha}} \leq c \left( \|\underline{b}_1\|_{C_{-2,t}^{0,\alpha}} + \|\underline{b}_2\|_{C_{-2,t}^{0,\alpha}} \right) \|\underline{b}_1 - \underline{b}_2\|_{C_{-2,t}^{0,\alpha}}$$

with a constant  $c > 0$  independent of  $t \in (0, T)$ . Since by Proposition 6.7

$$\|F_{\tilde{A}_t} \wedge \psi_t\|_{C_{-2,t}^{0,\alpha}} \leq ct^{1/2},$$

Lemma 8.4 provides us with, for each  $t \in (0, T')$ , a solution  $\underline{b}_t$  of (8.5) satisfying  $\|\underline{b}_t\|_{C_{-2,t}^{0,\alpha}} \leq ct^{1/2}$  provided  $T' \in (0, T]$  was chosen sufficiently small. Then

$$\underline{a}_t = (\xi_t, a_t) = R_t \underline{b}_t \in C_{-1,t}^{1,\alpha}$$

is the desired solution of (8.2) and satisfies  $\|\underline{a}_t\|_{C_{-1,t}^{1,\alpha}} \leq ct^{1/2}$ .

It follows from elliptic regularity that  $a_t$  and thus  $A_t := \tilde{A}_t + a_t$  is smooth. To see that  $A_t$  is acyclic, that is,  $L_{A_t}$  is injective, note that  $\|R_t L_{A_t} - \text{id}\|_{C_{-1,t}^{1,\alpha}} \leq ct^{1/2}$  and thus  $L_{A_t}$  is invertible for  $t \in (0, T')$  provided  $T' \in (0, T]$  was chosen sufficiently small.  $\square$

Before embarking on the proof of Proposition 8.3, it will be helpful to make a few observations. On  $Y_t \setminus \tilde{T}_t$  the operators  $L_t$  and  $L_\theta$  agree. For fixed  $\varepsilon > 0$ , the norms  $\|\cdot\|_{C_{\beta,t}^{k,\alpha}(r_t^{-1}[\varepsilon, \infty))}$  are uniformly equivalent to the corresponding unweighted Hölder norms. Moreover, the restriction of  $L_t$  to  $r_t^{-1}[\varepsilon, \infty)$  becomes arbitrarily close to  $L_\theta$  restricted to  $\{x \in Y_0 : d(x, S) > \varepsilon\}$  as  $t$  goes to zero. These observations and standard Schauder estimates combined with Proposition 7.7 and Proposition 7.11 yield the following Schauder estimate.

**Proposition 8.6.** *Given  $\beta \in \mathbf{R}$  there is a constant  $c > 0$  such that for all  $t \in (0, T)$  we have*

$$\|\underline{a}\|_{C_{\beta,t}^{1,\alpha}} \leq c(\|L_t \underline{a}\|_{C_{\beta-1,t}^{0,\alpha}} + \|\underline{a}\|_{L_{\beta,t}^\infty}).$$

This reduces the proof of Proposition 8.3 to the following statement.

**Proposition 8.7.** *Given  $\beta \in (-3, 0)$  there are constants  $T' \in (0, T)$  and  $c > 0$  such that for all  $t \in (0, T')$  the following holds:*

$$\|\underline{a}\|_{L_{\beta,t}^\infty} \leq c\|L_t \underline{a}\|_{C_{\beta-1,t}^{0,\alpha}}.$$

*Proof.* Suppose not. Then there exists a sequence  $(\underline{a}_i)$  and a null-sequence  $(t_i)$  such that

$$\|\underline{a}_i\|_{L_{\beta,t_i}^\infty} = 1 \quad \text{and} \quad \|L_{t_i} \underline{a}_i\|_{C_{\beta-1,t_i}^{0,\alpha}} \leq \frac{1}{i}.$$

Hence, by Proposition 8.6, we have

$$(8.8) \quad \|\underline{a}_i\|_{C_{\beta,t_i}^{1,\alpha}} \leq 2c.$$

Pick  $x_i \in Y_{t_i}$  such that

$$w_{t_i}(x_i)^{-\beta} |\underline{a}_i(x_i)| = 1.$$

After passing to a subsequence we can assume that one of the following three cases occurs. We will rule out all of them, thus proving the proposition.

**Case 1.** *The sequence  $(x_i)$  accumulates on the regular part of  $Y_0$ :  $\lim r_{t_i}(x_i) > 0$ .*

Let  $K$  be a compact subset of  $Y_0 \setminus S$ . We can view  $K$  as a subset of  $Y_t$ . As  $t$  goes to zero, the metric on  $K$  induced from the metric on  $Y_t$  converges to the metric on  $Y_0$ , similarly we can identify  $E_0|_K$  with  $E_t|_K$  and via this identification  $\hat{A}_t$  converges to  $\theta$  on  $K$ . By (8.8) the sequence  $(\underline{a}_i|_K)$  is uniformly bounded in  $C^{1,\alpha}$ . We can thus extract a convergent subsequence using Arzelà–Ascoli. Using a diagonal sequence argument over a sequence of compact sets  $(K_i)$  exhausting  $Y_0 \setminus S$  we can pass to a further subsequence which converges in  $C_{\text{loc}}^{1,\alpha/2}$  to a limit  $\underline{a} \in \Omega^0(Y_0 \setminus S, \mathfrak{g}_{E_0}) \oplus \Omega^1(Y_0 \setminus S, \mathfrak{g}_{E_0})$ . This limit satisfies

$$(8.9) \quad |\underline{a}| < c \cdot d(\cdot, S)^\beta$$

as well as

$$L_\theta \underline{a} = 0.$$

Since  $\beta > -3$ , it follows from (8.9) that  $\underline{a}$  satisfies  $L_\theta \underline{a} = 0$  in the sense of distributions on all of  $Y_0$  and, therefore, is smooth by elliptic regularity. Because  $\theta$  is assumed to be acyclic,  $\underline{a}$  must be zero. However, by passing to a further subsequence we can arrange that  $(x_i)$  converges to some point  $x \in Y_0 \setminus S$ . At this point we have  $|\underline{a}|(x) = d(x, S)^\beta \neq 0$ . This is a contradiction.

**Case 2.** *The sequence  $(x_i)$  accumulates on one of the ALE spaces:  $\lim r_{t_i}(x_i)/t_i < \infty$ .*

There is no loss in assuming that each  $x_i$  lies in  $\tilde{T}_{j,t_i}$  for some fixed  $j$ . With  $s_{\beta,t_i}$  as in (7.6) we define  $\tilde{a}_i := s_{\beta,t_i} a_i$  and denote by  $\tilde{x}_i$  a lift of  $x_i$  to  $\mathbf{R}^3 \times \pi_{j,t_i}^{-1}(B_\zeta^4/G_j)$ . This rescaled sequence satisfies, in the notation of Section 7,

$$\|\tilde{a}_i\|_{C_\beta^{1,\alpha}} \leq 4c \quad \text{and} \quad (1 + |\pi_j(\tilde{x}_i)|)^{-\beta} |\tilde{a}_i(\tilde{x}_i)| \geq \frac{1}{2}$$

as well as

$$(8.10) \quad \|L_{A_j} \tilde{a}_i\|_{C_{\beta-1}^{0,\alpha}} \leq 2/i.$$

Arguing as in the previous case, we can extract a subsequence of  $(\tilde{a}_i)$  which converges to a limit  $\tilde{a} \in C_\beta^{1,\alpha/2}$  in  $C_{\text{loc}}^{1,\alpha/2}$  on  $\mathbf{R}^3 \times X_j$ . It follows from (8.10) that  $\tilde{a}$  satisfies

$$L_{A_j} \tilde{a} = 0.$$

By Proposition 7.8,  $\tilde{a}$  must be zero since  $\beta \in (-3, 0)$  and  $A_j$  is infinitesimally rigid. However, by translation we can arrange that the  $\mathbf{R}^3$ -component of  $\tilde{x}_i$  is zero and thus we can view  $\tilde{x}_i$  as a point in  $X_j$ . Then the condition  $\lim d_{t_i}(x_i)/t_i < \infty$  translates to  $\lim |\pi_j(\tilde{x}_i)| < \infty$ . Therefore, we can assume without loss of generality that  $\tilde{x}_i$  converges to some point  $\tilde{x} \in X_j$ . But then  $|\tilde{a}(\tilde{x})| \geq \frac{1}{2}(1 + |\pi_j(\tilde{x})|)^\beta > 0$ , which contradicts  $\tilde{a} = 0$ .

**Case 3.** *The sequence  $(x_i)$  accumulates on one of the necks:  $\lim r_{t_i}(x_i) = 0$  and  $\lim r_{t_i}(x_i)/t_i = \infty$ .*

As in the previous case, we rescale to obtain  $(\tilde{a}_i)$  and  $(\tilde{x}_i)$ , and we arrange it so that the  $\mathbf{R}^3$ -component of  $\tilde{x}_i$  is zero. Since  $\lim d_{t_i}(x_i)/t_i = \infty$ , we have  $\lim |\pi_j(\tilde{x}_i)| = \infty$ . Fix a sequence  $(R_i)$  tending to infinity such that  $\varepsilon_i := R_i/|\pi_j(\tilde{x}_i)|$  goes to zero. Using  $\pi_j: X \rightarrow \mathbf{C}^2/G$ , we can think of the sets  $\mathbf{R}^3 \times (\mathbf{C}^2 \setminus B_{R_i}^4)/G_j$  as subsets of  $\mathbf{R}^3 \times X_j$ . Restricting to these sets and rescaling everything by  $1/|\pi_j(\tilde{x}_i)|$  we obtain, without changing notation,  $\tilde{a}_i \in \Omega^0(\mathbf{R}^3 \times (\mathbf{C}^2 \setminus B_{\varepsilon_i}^4)/G_j) \oplus \Omega^1(\mathbf{R}^3 \times (\mathbf{C}^2 \setminus B_{\varepsilon_i}^4)/G_j)$  and  $\tilde{x}_i \in \mathbf{C}^2 \setminus B_{\varepsilon_i}^4$  satisfying

$$\|\tilde{a}_i\|_{C_\beta^{1,\alpha}} \leq 8c \quad \text{and} \quad |\tilde{x}_i|^{-\beta} |\tilde{a}_i(\tilde{x}_i)| \geq \frac{1}{4}$$

as well as

$$\|L \tilde{a}_i\|_{C_{\beta-1}^{0,\alpha}} \leq 4/i.$$

Here the norms  $\|\cdot\|_{C_\beta^{k,\alpha}}$  are defined like those in Section 7 except with the weight function now defined by  $w(x, y) := |y|$  for  $(x, y) \in \mathbf{R}^3 \times \mathbf{C}^2/G_j$ . The operator  $L$  is defined by

$$L(\xi, a) := (d^* a, d\xi + *(\psi_0 \wedge da))$$

with  $\psi_0 := \frac{1}{2}\omega_1 \wedge \omega_1 + \delta^2 \wedge \delta^3 \wedge \omega_1 + \delta^3 \wedge \delta^1 \wedge \omega_2 - \delta^1 \wedge \delta^2 \wedge \omega_3$  and  $\omega_i \in \Omega^2(\mathbf{C}^2)$  as in Section 4.

As before, we can extract a subsequence converging in  $C_{\text{loc}}^{1,\alpha/2}$  to a limit

$$\tilde{a} \in \Omega^0(\mathbf{R}^3 \times (\mathbf{C}^2 \setminus \{0\})/G_j) \oplus \Omega^1(\mathbf{R}^3 \times (\mathbf{C}^2 \setminus \{0\})/G_j)$$

satisfying

$$(8.11) \quad |\tilde{a}| < cw^\beta$$

as well as

$$L\tilde{a} = 0.$$

Since  $\beta > -3$ , it follows from (8.11) that  $\tilde{a}$  satisfies  $L\tilde{a} = 0$  in the sense of distributions on all of  $\mathbf{R}^3 \times \mathbf{C}^2/G_j$  and therefore  $\tilde{a}$  is smooth by elliptic regularity. It also follows from (8.11) that both  $\tilde{a}$  and  $\nabla\tilde{a}$  are uniformly bounded: This is clear outside a tubular neighbourhood of  $\mathbf{R}^3 \times \{0\}$ . If  $B_1$  is a ball of radius one centred at some point in  $\mathbf{R}^3 \times \{0\}$ , then (8.11) gives a uniform bound on  $\|\tilde{a}\|_{L^p(B_1)}$ , for some fixed  $p \in (1, \infty)$ . Using elliptic estimates this yields a uniform  $W^{k,p}$  estimate on the ball of radius one-half; hence, using Sobolev embedding, uniform bounds on  $\tilde{a}$  and  $\nabla\tilde{a}$ . Because  $L^*L = \Delta_{\mathbf{R}^3} + \Delta_{\mathbf{C}^2}$ , it follows from Lemma 7.10 that  $\tilde{a}$  is invariant under translations in the  $\mathbf{R}^3$ -direction. Thus we can think of the components of  $\tilde{a}$  as harmonic functions on  $\mathbf{C}^2$ . Since  $\beta < 0$ , they decay to zero at infinity and thus vanish identically. However, we know that  $|\tilde{x}_i| = 1$  and thus a subsequence of  $(\tilde{x}_i)$  converges to a point  $\tilde{x} \in \mathbf{C}^2/G_j$  with  $|\tilde{x}| = 1$  at which  $|\tilde{a}|(\tilde{x}) \geq \frac{1}{4}$ , contradicting  $\tilde{a} = 0$ .  $\square$

## 9 Examples with $G = \text{SO}(3)$

We will now explain how to use Theorem 1.1 to construct a few concrete examples of  $G_2$ -instantons on the  $G_2$ -manifolds from [Joy00, Sections 12.3 and 12.4]. The flat  $G_2$ -structure  $\phi_0$  on  $T^7$  given by (2.1) is preserved by  $\alpha, \beta, \gamma \in \text{Diff}(T^7)$  defined by

$$\begin{aligned} \alpha(x_1, \dots, x_7) &:= (x_1, x_2, x_3, -x_4, -x_5, -x_6, -x_7), \\ \beta(x_1, \dots, x_7) &:= (x_1, -x_2, -x_3, x_4, x_5, \frac{1}{2} - x_6, -x_7), \\ \gamma(x_1, \dots, x_7) &:= (-x_1, x_2, -x_3, x_4, -x_5, x_6, \frac{1}{2} - x_7). \end{aligned}$$

It is easy to see that  $\Gamma := \langle \alpha, \beta, \gamma \rangle \cong \mathbf{Z}_2^3$ .

To understand the singular set  $S$  of  $T^7/\Gamma$  note that the only elements of  $\Gamma$  having fixed points are  $\alpha, \beta$  and  $\gamma$ . The fixed point set of each of these elements consists of 16 copies of  $T^3$ . The group  $\langle \beta, \gamma \rangle$  acts freely on the set of  $T^3$  fixed by  $\alpha$  and  $\langle \alpha, \gamma \rangle$  acts freely on the set of  $T^3$  fixed by  $\beta$ , while  $\alpha\beta \in \langle \alpha, \beta \rangle$  acts trivially on the set of  $T^3$  fixed by  $\gamma$ . It follows that  $S$  consists of 8 copies of  $T^3$  coming from the fixed points of  $\alpha$  and  $\beta$  and 8 copies of  $T^3/\mathbf{Z}_2$ . Near the copies of  $T^3$  the singular set is modelled on  $T^3 \times \mathbf{C}^2/\mathbf{Z}_2$  while near the copies of  $T^3/\mathbf{Z}_2$  it is modelled on  $(T^3 \times \mathbf{C}^2/\mathbf{Z}_2)/\mathbf{Z}_2$  where the action of  $\mathbf{Z}_2$  on  $T^3 \times \mathbf{C}^2/\mathbf{Z}_2$  is given by

$$(x_1, x_2, x_3, \pm(z_1, z_2)) \mapsto (x_1, x_2, x_3 + \frac{1}{2}, \pm(z_1, -z_2)).$$

The 8 copies of  $T^3$  can be desingularised by any choice of 8 ALE spaces asymptotic to  $\mathbf{C}^2/\mathbf{Z}_2$ . To desingularise the copies of  $T^3/\mathbf{Z}_2$  we need to choose ALE spaces which admit an isometric action

of  $\mathbf{Z}_2$  asymptotic to the action  $\mathbf{Z}_2$  on  $\mathbf{C}^2/\mathbf{Z}_2$  given by  $\pm(z_1, z_2) \mapsto \pm(z_1, -z_2)$ . Two possible choices are the resolution of  $\mathbf{C}^2/\mathbf{Z}_2$  or a smoothing of  $\mathbf{C}^2/\mathbf{Z}_2$ . See Joyce [Joy00, pages 313–314] for details.

We construct our examples on desingularisations of quotients of  $T^7/\Gamma$ . To this end we define  $\sigma_1, \sigma_2, \sigma_3 \in \text{Diff}(T^7)$  by

$$\begin{aligned}\sigma_1(x_1, \dots, x_7) &:= (x_1, x_2, \frac{1}{2} + x_3, \frac{1}{2} + x_4, \frac{1}{2} + x_5, x_6, x_7), \\ \sigma_2(x_1, \dots, x_7) &:= (x_1, \frac{1}{2} + x_2, x_3, \frac{1}{2} + x_4, x_5, x_6, x_7), \\ \sigma_3(x_1, \dots, x_7) &:= (\frac{1}{2} + x_1, x_2, x_3, x_4, \frac{1}{2} + x_5, \frac{1}{2} + x_6, x_7).\end{aligned}$$

The elements  $\sigma_j$  commute with all elements of  $\Gamma$  and thus act on  $T^7/\Gamma$ . Moreover, this action is free.

**Example 9.1.** Let  $A := \langle \sigma_2, \sigma_3 \rangle$ . By analysing how  $A$  acts on the singular set of  $T^7/\Gamma$  one can see that the singular set of  $Y_0 := T^7/(\Gamma \times A)$  consists of one copy of  $T^3$ , denoted by  $S_1$ , and 6 copies of  $T^3/\mathbf{Z}_2$ , denoted by  $S_2, \dots, S_7$ .  $S_1$  has a neighbourhood modelled on  $T^3 \times \mathbf{C}^2/\mathbf{Z}_2$ , while  $S_2, \dots, S_6$  have neighbourhoods modelled on  $(T^3 \times \mathbf{C}^2/\mathbf{Z}_2)/\mathbf{Z}_2$  where  $\mathbf{Z}_2$  acts by  $\pm(z_1, z_2) \mapsto \pm(z_1, -z_2)$  on  $\mathbf{C}^2/\mathbf{Z}_2$ . As before,  $S_1$  can be desingularised by any choice of an ALE space asymptotic to  $\mathbf{C}^2/\mathbf{Z}_2$ .  $S_2, \dots, S_6$  can be desingularised by the resolution of  $\mathbf{C}^2/\mathbf{Z}_2$  or a smoothing of  $\mathbf{C}^2/\mathbf{Z}_2$ .

To compute the orbifold fundamental group  $\pi_1(Y_0)$ , note that it is isomorphic to the fundamental group  $\pi_1(Y_0 \setminus S)$  of the regular part of  $Y_0$ . Denote by  $p: \mathbf{R}^7 \rightarrow Y_0$  the canonical projection. Then  $p: p^{-1}(Y_0 \setminus S) \rightarrow Y_0 \setminus S$  is a universal cover. Up to conjugation we can therefore identify  $\pi_1(Y_0)$  with the group of deck transformations

$$\pi_1(Y_0) = \langle \alpha, \beta, \gamma, \sigma_2, \sigma_3, \tau_1, \dots, \tau_7 \rangle \subset \text{Aff}(7) = \text{GL}(7) \ltimes \mathbf{R}^7.$$

Here we think of  $\alpha, \beta, \gamma, \sigma_2, \sigma_3$  as elements of  $\text{Aff}(7)$  defined by the formulae above and  $\tau_i$  translates the  $i^{\text{th}}$  coordinate of  $\mathbf{R}^7$  by one. The group  $\pi_1(Y_0)$  is a non-split extension

$$0 \rightarrow \mathbf{Z}^7 \rightarrow \pi_1(Y_0) \rightarrow \Gamma \times A \rightarrow 0.$$

To work out the orbifold fundamental group  $\pi_1(T_j)$  of  $T_j$ , again up to conjugation, one simply has to understand the subgroup of deck transformations preserving a fixed component of  $p^{-1}(T_j) \subset p^{-1}(Y_0 \setminus S)$ . In this way one can compute

$$\begin{aligned}\pi_1(T_1) &= \langle \alpha, \tau_1, \tau_2, \tau_3 \rangle, \\ \pi_1(T_2) &= \langle \beta, \sigma_3 \alpha, \tau_1, \tau_4, \tau_5 \rangle, & \pi_1(T_3) &= \langle \tau_3 \beta, \sigma_3 \alpha, \tau_1, \tau_4, \tau_5 \rangle, \\ \pi_1(T_4) &= \langle \gamma, \alpha \beta, \sigma_2, \tau_4, \tau_6 \rangle, & \pi_1(T_5) &= \langle \tau_3 \gamma, \tau_3 \alpha \beta, \sigma_2, \tau_4, \tau_6 \rangle, \\ \pi_1(T_6) &= \langle \tau_5 \gamma, \tau_5 \alpha \beta, \sigma_2, \tau_4, \tau_6 \rangle, & \pi_1(T_7) &= \langle \tau_3 \tau_5 \gamma, \tau_3 \tau_5 \alpha \beta, \sigma_2, \tau_4, \tau_6 \rangle.\end{aligned}$$

Here  $\tau_2$  does not appear explicitly in  $\pi_1(T_j)$ , for  $j = 4, \dots, 7$ , because  $\sigma_2^2 = \tau_2 \tau_4$ .

Denote by  $V := \langle a, b, c \mid a^2 = b^2 = c^2 = 1, ab = c \rangle \cong \mathbf{Z}_2^2$  the Klein four-group.  $V$  can be thought of as a subgroup of  $\text{SO}(3)$ :  $a = \text{diag}(1, -1, -1)$ ,  $b = \text{diag}(-1, 1, -1)$  and  $c = \text{diag}(-1, -1, 1)$ . We

define  $\rho: \pi_1(Y_0) \rightarrow V \subset \text{SO}(3)$  by

$$\begin{aligned} \beta, \gamma, \tau_1, \dots, \tau_7 &\mapsto 1, & \alpha &\mapsto a, \\ \sigma_2 &\mapsto a, & \sigma_3 &\mapsto b. \end{aligned}$$

To see that the flat connection  $\theta$  induced by  $\rho$  is acyclic we use the following observation.

**Proposition 9.2.** *A flat connection  $\theta$  on a  $G$ -bundle  $E_0$  over a flat  $G_2$ -orbifold  $Y_0$  corresponding to a representation  $\rho: \pi_1(Y_0) \rightarrow G$  is acyclic if and only if the induced representation of  $\pi_1(Y_0)$  on  $\mathfrak{g} \oplus (\mathbf{R}^7 \otimes \mathfrak{g})$  has no non-zero fixed vectors.*

*Proof.* Since  $Y_0$  is flat as a Riemannian orbifold and  $\theta$  is a flat connection

$$L_\theta^* L_\theta = \nabla_\theta^* \nabla_\theta.$$

Therefore, all elements in the kernel of  $L_\theta$  are actually parallel sections of the bundle  $\mathfrak{g}_{E_0} \oplus (T^*Y_0 \otimes \mathfrak{g}_{E_0})$  and these are in one-to-one correspondence with fixed vectors of the representation of  $\pi_1(Y_0)$  on  $\mathfrak{g} \oplus (\mathbf{R}^7 \otimes \mathfrak{g})$ .  $\square$

The elements  $\sigma_2$  and  $\sigma_3$  act trivially on  $\mathbf{R}^7$  and their action on  $\mathfrak{so}(3)$  has no common non-zero fixed vectors. Therefore the action of  $\pi_1(Y_0)$  on  $\mathfrak{g} \oplus (\mathbf{R}^7 \otimes \mathfrak{g})$  has no non-zero fixed vector and thus  $\theta$  is acyclic.

The monodromy representation  $\mu_j|_{G_j}: G_j = \mathbf{Z}_2 \rightarrow \text{SO}(3)$  associated with the flat connection  $\theta$  is non-trivial only for  $j = 1$ . Let  $A_1 := A_{0,1}$  be the infinitesimally rigid ASD instanton on  $E_1 := E_{0,1}$  given in Proposition 5.22. For  $j = 2, \dots, 6$  we choose  $A_j$  to be the product connection on the trivial  $\text{SO}(3)$ -bundle  $E_j$ . We take  $\tilde{\rho}_1$  to be trivial. For  $j = 2, \dots, 6$  we can choose  $\tilde{\rho}_j$  accordingly to satisfy the compatibility conditions. Thus we obtain examples of  $G_2$ -instantons on each of the desingularisations of  $Y_0$  by appealing to Theorem 1.1.

Note that any choice of resolution data for  $T^7/(\Gamma \times A)$  lifts to an  $A$ -invariant choice of resolution data for  $T^7/\Gamma$ . We can then carry out Joyce's generalised Kummer construction in a  $A$ -invariant way and lift up the  $G_2$ -instanton constructed above. However, we could not have constructed this  $G_2$ -instanton directly using Theorem 1.1, since the lift of  $\theta$  to  $T^7/\Gamma$  is not acyclic.

**Example 9.3.** Here is a more complicated example. Let  $Y_0 := T^7/(\Gamma \times A)$  be as before. Define  $\rho: \pi_1(Y_0) \rightarrow V \subset \text{SO}(3)$  by

$$\begin{aligned} \gamma, \tau_1, \dots, \tau_7 &\mapsto 1, & \alpha &\mapsto a, & \beta &\mapsto b, \\ \sigma_2 &\mapsto b, & \sigma_3 &\mapsto a. \end{aligned}$$

Again, the resulting flat connection  $\theta$  is acyclic. For  $j = 1, 2, 3$  let  $A_j := A_{0,1}$  be the rigid ASD instanton on  $E_j := E_{0,1}$ . By adapting the framings of  $E_2$  and  $E_3$ , we can arrange that  $A_2$  and  $A_3$  are asymptotic at infinity to the flat connection with monodromy given by  $b \in V$ . For  $j = 4, \dots, 7$  let  $A_j$  be the product connection on the trivial bundle  $E_j$ . To be able to extend this to compatible

gluing data we need a lift  $\tilde{\rho}_j$  of the action of  $\mathbf{Z}_2$  on  $X_j$  to  $E_j$  preserving  $A_j$  and acting trivially on the framing at infinity for  $j = 2, 3$ . If  $X_j$  is a smoothing of  $\mathbf{C}^2/\mathbf{Z}_2$ , then the  $\mathbf{Z}_2$  action on  $X_j$  does lift to  $E_j$  preserving  $A_j$ . However, the action does not lift if  $X_j$  is the resolution of  $\mathbf{C}^2/\mathbf{Z}_2$ . The reason for this is that in the first case the action of  $\mathbf{Z}_2$  on  $H^2(X, \mathbf{R})$  is given by the identity, while in the second case it acts via multiplication by  $-1$ ; see Joyce [Joy00, pages 313–314]. Thus we can only find compatible gluing data if we resolve both  $S_2$  and  $S_3$  using a smoothing of  $\mathbf{C}^2/\mathbf{Z}_2$ .

Here is a small modification of this example. Define  $\rho: \pi_1(Y_0) \rightarrow V \subset \mathrm{SO}(3)$  by

$$\begin{aligned} \gamma, \tau_1, \dots, \tau_7 &\mapsto 1, & \alpha &\mapsto a, & \beta &\mapsto b, \\ \sigma_2 &\mapsto b, & \sigma_3 &\mapsto c. \end{aligned}$$

To find compatible gluing data, one simply has to compose  $\tilde{\rho}_j$  as above with multiplication by  $b \in \mathcal{G}(E_j)$ , for  $j = 2, 3$ .

**Example 9.4.** Let  $B := \langle \sigma_1, \sigma_2, \sigma_3 \rangle$  and  $Y_0 := T^7/(\Gamma \times B)$ . Then the singular set of  $Y_0$  consists of 4 copies of  $T^3/\mathbf{Z}_2$ , denoted by  $S_1, \dots, S_4$ , each of which has a neighbourhood modelled on  $(T^3 \times \mathbf{C}^2/\mathbf{Z}_2)/\mathbf{Z}_2$  where  $\mathbf{Z}_2$  acts on  $\mathbf{C}^2/\mathbf{Z}_2$  by  $\pm(z_1, z_2) \mapsto \pm(z_1, -z_2)$ . The orbifold fundamental group  $\pi_1(Y_0)$  is given by

$$\pi_1(Y_0) = \langle \alpha, \beta, \gamma, \sigma_1, \sigma_2, \sigma_3, \tau_1, \dots, \tau_7 \rangle \subset \mathrm{Aff}(7).$$

Up to conjugation the fundamental groups of the neighbourhoods  $T_j$  of  $S_j$  are given by

$$\begin{aligned} \pi_1(T_1) &= \langle \alpha, \tau_4^{-1} \tau_5^{-1} \beta \sigma_1 \sigma_2 \sigma_3, \tau_1, \tau_2, \tau_3 \rangle, & \pi_1(T_2) &= \langle \beta, \sigma_3 \alpha, \tau_1, \tau_4, \tau_5 \rangle, \\ \pi_1(T_3) &= \langle \gamma, \alpha \beta, \sigma_2, \tau_4, \tau_6 \rangle, & \pi_1(T_4) &= \langle \tau_3 \gamma, \tau_3 \alpha \beta, \sigma_2, \tau_4, \tau_6 \rangle. \end{aligned}$$

Define  $\rho: \pi_1(Y_0) \rightarrow V \subset \mathrm{SO}(3)$  by

$$\begin{aligned} \alpha, \beta, \sigma_3, \tau_1, \dots, \tau_7 &\mapsto 1, & \gamma &\mapsto b \\ \sigma_1 &\mapsto a, & \sigma_2 &\mapsto b. \end{aligned}$$

The induced flat connection  $\theta$  is clearly acyclic. As before, for  $j = 3, 4$ , we require  $S_j$  to be desingularised using a resolution of  $\mathbf{C}^2/\mathbf{Z}_2$  in order to be able to find a lift  $\tilde{\rho}_j$ . Also note that, for  $j = 3, 4$ , now we have to make a non-trivial choice for  $\tilde{\rho}_j$ , but this causes no problem since  $b \in V$  lies in  $\mathcal{G}(E_j)$  and preserves  $A_j$ .

Again, the resulting  $G_2$ -instanton can be lifted to appropriate  $\sigma_1$ -invariant desingularisations of  $T^7/(\Gamma \times A)$ ; however we could not have constructed the lifted  $G_2$ -instanton directly, since the lift of  $\theta$  to  $T^7/(\Gamma \times A)$  is not acyclic.

This list of examples is not exhaustive. The reader will have no difficulty finding more examples by modifying the ones given above.

## 10 Proof of Theorem 1.6

The key to proving Theorem 1.6 is the following observation.

**Proposition 10.1.** *Let  $X$  denote an ALE space asymptotic to  $\mathbb{C}^2/\mathbb{Z}_2$ , let  $E$  be a  $\mathrm{SO}(3)$ -bundle with non-trivial  $w_2(E)$  and let  $D \subset X$  be a discrete subset. Then every connection  $A$  on  $E|_{X \setminus D}$  satisfies*

$$\mathrm{YM}(A) = \int_{X \setminus D} |F_A|^2 \geq 4\pi^2.$$

*Equality is achieved if and only if  $A$  is gauge equivalent to the rigid ASD instanton  $A_{0,1}$  from Proposition 5.22.*

The proof requires some knowledge of Sedlacek's work on weak compactness in 4-dimensional gauge theory [Sed82]. For the reader's benefit I will briefly recall his main results.

**Theorem 10.2** (Sedlacek [Sed82, Theorem 3.1]). *Let  $M$  be a Riemannian 4-manifold and let  $E$  be a  $G$ -bundle over  $M$  where  $G$  is a compact Lie group together with a fixed embedding  $G \subset U(n) \subset \mathbb{C}^{n \times n}$ . Suppose that  $(A_i)$  is a sequence of connections on  $E$  whose energy is uniformly bounded, i.e.,  $\mathrm{YM}(A_i) = \|F_{A_i}\|_{L^2} \leq c$  for some fixed finite constant  $c$ . Then, after passing to a subsequence, there exists a finite subset  $D \subset M$  and a cover  $\{U_\alpha\}$  of  $M \setminus D$  consisting of small geodesic balls over which  $E$  is trivialised, such that on each  $U_\alpha$  the sequence  $(A_i|_{U_\alpha})$  converges up to gauge transformations to a connection  $A_\alpha \in W^{1,2}(U_\alpha, T^*U_\alpha \otimes \mathfrak{g}_E)$  weakly in  $W^{1,2}$  and on  $U_{\alpha\beta} := U_\alpha \cap U_\beta$  the connections  $A_\alpha$  and  $A_\beta$  are related by a transition function  $g_{\alpha\beta} \in W^{2,2}(U_{\alpha\beta}, G)$ .*

Sedlacek derived this result in a rather straight-forward manner from the work of Uhlenbeck [Uhl82a]. A proof of the same result using a different method can be found in Kessel's PhD thesis [Keso8, Chapter 3].

**Definition 10.3.** In the situation of Theorem 10.2 we say that the sequence  $(A_i)$  converges weakly to the  $W^{1,2}$  connection  $A = (A_\alpha)$  on the  $W^{2,2}$   $G$ -bundle  $E = (g_{\alpha\beta})$ .

In the case  $G = \mathrm{SO}(3)$  the main result of [Sed82] can be stated as follows.

**Theorem 10.4** (Sedlacek). *Let  $M$  be a Riemannian 4-manifold. Then for each  $w \in H^2(M, \mathbb{Z}_2)$  the infimum*

$$\inf\{\mathrm{YM}(A) : A \text{ is a connection on a } \mathrm{SO}(3)\text{-bundle } E \text{ with } w_2(E) = w\}$$

*is attained. Denote this infimum by  $m(w)$ .*

*Proof Sketch.* Let  $(A_i)$  be a sequence of connections on  $\mathrm{SO}(3)$ -bundles  $E_i$  with  $w_2(E_i) = w$  with a limiting  $W^{1,2}$  connection  $\tilde{A} = (\tilde{A}_\alpha)$  on a  $W^{2,2}$   $\mathrm{SO}(3)$ -bundle  $\tilde{E} = (\tilde{g}_{\alpha\beta})$ . Then each  $\tilde{A}_\alpha$  is weakly Yang-Mills [Sed82, Theorem 4.1]; hence, each  $\tilde{A}_\alpha$  and thus also each  $\tilde{g}_{\alpha\beta}$  is smooth [Sed82, Proposition 4.2]. By Uhlenbeck's removable singularities theorem [Uhl82b, Theorem 4.1]  $\tilde{A}$  extends to a smooth connection  $A$  on a smooth bundle  $E$  over  $M$ . By lower semi-continuity of the Yang-Mills functional under weak convergence,  $\mathrm{YM}(A) \leq m(w)$ . According to [Sed82, Theorem 5.5] the second Stiefel-Whitney class is preserved in this limit procedure, i.e.,  $w_2(E) = w$  and thus  $\mathrm{YM}(A) = m(w)$ .  $\square$

*Proof of Proposition 10.1.* We already know that the asserted minimum is achieved for  $A = A_{0,1}$ . Now, let  $A$  be any connection on a  $\mathrm{SO}(3)$ -bundle  $E$  with non-trivial  $w_2(E)$ . By Theorem 10.4 we may assume that the energy of  $A$  realises the infimum  $m(w)$  where  $w \in H^2(X, \mathbf{Z}_2)$  is the unique non-trivial element. By Proposition 5.5 it follows that in some framing  $A$  is asymptotic to a flat connection  $\theta$  over  $(\mathbf{C}^2 \setminus \{0\})/\Gamma$ . By Theorem 5.16 the  $L^2$  index of  $\delta_A$  is given by

$$\mathrm{index} \delta_A = -2 \int_X p_1(A) + \frac{\chi_{\mathfrak{so}(3)}(-1) - 3}{4}$$

Here  $\chi_{\mathfrak{so}(3)}$  is the character of the representation of  $\mathbf{Z}_2$  on  $\mathfrak{so}(3)$  associated with  $\theta$ . Depending on whether  $\theta$  is trivial or not the second summand, which we denote by  $\varepsilon$  in the following, is either 0 or  $-1$ . Since  $\delta_A$  is quaternionic linear, the index of  $\delta_A$  is divisible by four. We can write the energy of  $A$  as

$$\begin{aligned} \mathrm{YM}(A) &= 2 \int_X |F_A^\pm|^2 \mp 8\pi^2 \int_X p_1(A) \\ &= 2 \int_X |F_A^\pm|^2 \pm 4\pi^2(\mathrm{index} \delta_A - \varepsilon). \end{aligned}$$

If  $\mathrm{index} \delta_A > 0$ , then  $\mathrm{index} \delta_A \geq 4$  and thus  $\mathrm{YM}(A) \geq 16\pi^2$ . If  $\mathrm{index} \delta_A < 0$ , then  $\mathrm{index} \delta_A \leq -4$  and thus  $\mathrm{YM}(A) \geq 12\pi^2$ . In both cases  $A$  cannot be a minimiser for the energy; therefore,  $\mathrm{index} \delta_A = 0$ . If  $\varepsilon = 0$ , then  $\int_X p_1(A) = 0$  and  $\theta$  is trivial; however, this is impossible: The whole situation is conformally invariant and we may thus think of  $X$  as having a cylindrical end. Now, we can double  $X$  to obtain a compact manifold  $W$  and construct a  $\mathrm{SO}(3)$ -bundle  $F$  over this manifold which coincides with  $E$  over one half of  $W$  and is trivial over the second half of  $W$ . This bundle will have  $p_1(F) = 0$  and  $w_2(F)^2 = 2 \bmod 4$  which contradicts the identity  $p_1(F) = w_2(F)^2 \bmod 4$ , see [DK90, Equation (2.1.36)]. We therefore have  $\varepsilon = -1$  which implies that  $\mathrm{YM}(A) \geq 4\pi^2$ . Moreover, it follows that  $\theta$  has precisely one non-trivial parallel section,  $A$  is anti-self-dual and, hence,  $\dim \ker \delta_A = \mathrm{index} \delta_A = 0$  and  $\delta_A$  is invertible, say, from  $W^{1,2}$  to  $L^2$ . The fact that  $\delta_A$  is invertible can be used to extend the non-trivial parallel section of  $\theta$  to a non-trivial parallel section of  $E$ . Therefore  $E$  is of the form  $\mathbf{R} \oplus \mathcal{L}$  and  $A$  comes from a  $U(1)$  ASD instanton on  $\mathcal{L}$ ; hence, it must be gauge equivalent as a  $\mathrm{SO}(3)$  ASD instanton to  $A_{0,1}$ .  $\square$

The above combined with a contradiction argument based on Sedlacek's work yields the following.

**Proposition 10.5.** *In the situation of Proposition 10.1 suppose that  $(U_i)$  is an exhaustion of  $X$  by increasing compact sets. Then for each  $\varepsilon > 0$ , there is an  $i_0 > 0$  such that for  $i \geq i_0$  the following holds: If  $A$  is a smooth connection on  $E|_{U_i}$ , then*

$$\mathrm{YM}(A) = \int_{U_i} |F_A|^2 \geq 4\pi^2 - \varepsilon.$$

Now, as in the hypothesis of Theorem 1.6, let  $Y_0$  be an admissible  $G_2$ -orbifold all of whose singularities  $S_j$  are modelled on  $(T^3 \times \mathbb{C}^2 / \mathbb{Z}_2) / H_j$  and let  $Y$  be compact 7-manifold and  $(\phi_t)_{t \in (0, T)}$  be a family of torsion-free  $G_2$ -structures on  $Y$  obtained via Joyce's generalised Kummer construction from resolution data for  $Y_0$ . Suppose that  $E$  is a  $SO(3)$ -bundle over  $Y$  with the property that

$$(10.6) \quad p_1(E) = - \sum_j \frac{\varepsilon_j}{2} \text{PD}[S_j]$$

where  $\varepsilon_j = 1$  if  $H_2(X_j)^{H_j} \subset H_2(Y)$  is non-trivial and  $w_2(E)$  pairs non-trivially with  $H_2(X_j)^{H_j}$  and  $\varepsilon_j = 0$  otherwise. Suppose that  $(A_t)_{t \in (0, T)}$  is a family of connections on  $E$  such that  $A_t$  is a  $G_2$ -instanton over  $(Y, \phi_t)$ .

By (3.6), (10.6) and Theorem 4.16 we have

$$\text{YM}(A_t) = 4\pi^2 \sum_j \varepsilon_j \text{vol}(S_j) + O(t^{1/2}).$$

Moreover, it follows from Proposition 10.5 that

$$\text{YM}(A_t|_{r_t^{-1}[0, \delta\zeta]}) \geq 4\pi^2 \sum_j \varepsilon_j \text{vol}(S_j) - o(1)$$

with  $r_t$  as in (4.13) and  $\delta \in (0, 1/4]$  fixed, that is, the energy of  $A_t$  is mostly concentrated inside the resolution locus. Consequently, the energy outside the resolution locus goes to zero

$$(10.7) \quad \text{YM}(A_t|_{r_t^{-1}(\delta\zeta, \infty)}) = o(1).$$

**Theorem 10.8** (Nakajima [Nak88, Lemma 3.1]). *Let  $M$  be a Riemannian  $n$ -manifold and let  $E$  be a  $G$ -bundle over  $M$ . Then there are constants  $c, r_0, \varepsilon_0 > 0$ , such that the following holds. Let  $A$  be a Yang–Mills connection on  $E$ . If  $x \in M$  and  $0 < r \leq r_0$  are such that*

$$(10.9) \quad \varepsilon := r^{4-n} \int_{B_r(x)} |F_A|^2 \text{vol} \leq \varepsilon_0,$$

then

$$\sup_{y \in B_{\frac{r}{4}}(x)} |F_A|^2(y) \leq c\varepsilon r^{-4}.$$

*Remark 10.10.* This result is one of the cornerstones of compactness theory for the Yang–Mills equation in higher dimensions. Its proof heavily relies on Price's monotonicity formula [Pri83] which says that the quantity  $\varepsilon = \varepsilon(r)$  defined in (10.9) is (essentially) monotonically increasing in  $r$ .

It follows from (10.7) and Theorem 10.8 that on every fixed closed ball in  $Y_0 \setminus S$  the  $L^\infty$ -norm of  $F_{A_t}$  will be arbitrarily small provided  $t > 0$  is sufficiently small; hence, we can put  $A_t$  into Uhlenbeck gauge [Uhl82a, Theorem 1.3]. In this gauge the  $G_2$ -instanton equation becomes elliptic

and, using standard elliptic theory, one can show that on every closed ball in  $Y_0 \setminus S$  there exist gauge transformations  $(\tilde{g}_t)$  such that  $(\tilde{g}_t^* A_t)$  converges to a flat connection on this ball. As in [DK90, Section 4.4.2] these local gauge transformations can be patched to yield global gauge transformations  $(g_t)$  such that  $(g_t^* A_t)$  converges to a flat connection  $\theta$  in  $C_{\text{loc}}^\infty(Y_0 \setminus S)$ .

From Proposition 6.2 it follows that  $\theta$  extends to an orbifold flat connection on  $Y_0$ . This completes the proof of Theorem 1.6.  $\square$

## A An infinite-dimensional Liouville-type theorem

The following result is an abstraction of various results that have appeared in the literature, for example, in Pacard–Ritoré’s work on the Allen–Cahn equation [PR03, Corollary 7.5] and in Brendle’s unpublished work on the Yang–Mills equation in higher dimension [Bre03, Proposition 3.3].

**Lemma A.1.** *Let  $E$  be a vector bundle of bounded geometry over a Riemannian manifold  $X$  of bounded geometry and with subexponential volume growth, and suppose that  $D: C^\infty(X, E) \rightarrow C^\infty(X, E)$  is a uniformly elliptic operator of second order whose coefficients and their first derivatives are uniformly bounded, that is non-negative, such that  $\langle Da, a \rangle \geq 0$  for all  $a \in W^{2,2}(X, E)$ , and formally self-adjoint. If  $a \in C^\infty(\mathbf{R}^n \times X, E)$  satisfies*

$$(\Delta_{\mathbf{R}^n} + D)a = 0$$

and  $\|a\|_{L^\infty}$  is finite, then  $a$  is constant in the  $\mathbf{R}^n$ -direction, that is  $a(x, y) = a(y)$ . Here, by slight abuse of notation, we denote the pullback of  $E$  to  $\mathbf{R}^n \times X$  by  $E$  as well.

Here is a heuristic argument. Denote by  $\hat{a}$  the partial Fourier transform of  $a$  in the  $\mathbf{R}^n$ -direction. Then  $\hat{a}$  solves  $(D + |k|^2)\hat{a} = 0$ . But  $D + |k|^2$  is invertible for  $k \neq 0$ . Thus  $\hat{a}$  is supported on  $\{0\} \times X$  and hence must be a linear combination of derivatives of various orders of  $\Gamma(E)$ -valued  $\delta$ -functions. Reversing the Fourier transform shows that  $a$  must be a polynomial in  $\mathbf{R}^n$ . But then it follows from the assumptions that  $a$  is constant in the  $\mathbf{R}^n$ -direction. The actual proof will be slightly more pedestrian.

First we need to set-up some notation. We fix a point  $p \in X$  and denote by  $\rho: X \rightarrow [0, \infty)$  a smoothing of the distance from  $p$ , as in Kordyukov [Kor91, Proposition 4.1]. For  $\delta \in \mathbf{R}$  we introduce a weight function  $w_\delta := e^{-\delta\rho}$  and weighted Hilbert spaces  $W_\delta^{s,2}(X, E)$  consisting of locally integrable sections  $f$  such that  $w_\delta \cdot f$  lies in  $W^{s,2}(X, E)$  with inner product defined by  $\langle \cdot, \cdot \rangle_{W_\delta^{s,2}} := \langle w_\delta \cdot, w_\delta \cdot \rangle_{W^{s,2}}$ . As usual we set  $L_\delta^2(X, E) := W_\delta^{0,2}(X, E)$ .

**Proposition A.2.** *For each  $k_0 > 0$  there is a constant  $\varepsilon = \varepsilon(k_0) > 0$  such that for all  $\delta \in (-\varepsilon, \varepsilon)$  and  $k \in [k_0, \infty)$  the operator  $D + k^2: W_\delta^{2,2}(X, E) \rightarrow L_\delta^2(X, E)$  is an isomorphism. Moreover, for  $\ell \geq 0$  there is a constant  $c_\ell = c_\ell(k_0) > 0$  such that*

$$(A.3) \quad \left\| \partial_k^\ell (D + k^2)^{-1} a \right\|_{W_\delta^{2,2}} \leq c_\ell (1 + k)^\ell \|a\|_{L_\delta^2}$$

for all  $k \in [k_0, \infty)$  and  $a \in L_\delta^2(X, E)$ .

*Proof.* By standard elliptic theory we have

$$\|a\|_{W^{2,2}} \leq c(\|Da\|_{L^2} + \|a\|_{L^2}).$$

Since  $D$  is non-negative, we have

$$\|Da\|_{L^2} \leq \|(D + k^2)a\|_{L^2} \quad \text{and} \quad k^2\|a\|_{L^2} \leq \|(D + k^2)a\|_{L^2}.$$

Putting everything together yields

$$\|a\|_{W^{2,2}} \leq c(1 + 1/k_0^2)\|(D + k^2)a\|_{L^2}$$

for  $k \in [k_0, \infty)$ . This implies that  $D + k^2: W^{2,2} \rightarrow L^2$  is an injective operator with closed range. It is also surjective, since its co-kernel can be identified with the  $L^2$  kernel of  $D + k^2$  which is trivial.

We now argue as in [Kor91, Proposition 4.4]. Via the Hilbert space isomorphism  $W_\delta^{s,2} \cong W^{s,2}$  defined by multiplication with  $w_\delta$  the operator  $D + k^2: W_\delta^{2,2} \rightarrow L_\delta^2$  is equivalent to  $D_\delta + k^2: W^{2,2} \rightarrow L^2$  where  $D_\delta := w_\delta D w_\delta^{-1}$ . We can write  $D_\delta$  as

$$D_\delta = D + \delta P_\delta$$

with  $P_\delta: W^{2,2} \rightarrow L^2$  bounded independent of  $\delta$ . Therefore,

$$\|((D + k^2) - (D_\delta + k^2))(D + k^2)^{-1}a\|_{L^2} \leq |\delta|c(1 + 1/k_0^2)\|a\|_{L^2}.$$

If we choose  $\varepsilon = \varepsilon(k_0) > 0$  sufficiently small, then for  $\delta \in (-\varepsilon, \varepsilon)$  the factor on the right-hand side is less than  $\frac{1}{2}$ ; thus, the series

$$(D_\delta + k^2)^{-1} := (D + k^2)^{-1} \sum_{i \geq 0} [((D + k^2) - (D_\delta + k^2))(D + k^2)^{-1}]^i$$

converges and the operator norm of  $(D_\delta + k^2)^{-1}$  is bounded by  $2c(1 + 1/k_0^2)$ . This establishes (A.3) for  $\ell = 0$ . For  $\ell > 0$ , we have

$$\partial_k^\ell (D + k^2)^{-1} = \sum_{i=0}^{\ell} \sum_{j=2}^{\ell+1} c_{i,j,\ell} \cdot k^i [(D + k^2)^{-1}]^j$$

for universal constants  $c_{i,j,\ell}$ . Thus (A.3) for  $\ell > 0$  can be reduced to the case  $\ell = 0$ .  $\square$

Lemma A.1 can now be proved using an argument similar to the one used by Brendle [Bre03, Proposition 3.3]. This is essentially the proof of the ingredients from classical distribution theory used in the heuristic proof adapted to our infinite-dimensional setting.

*Proof of Lemma A.1.* We proceed in 3 steps.

**Step 1.** Let  $\chi \in \mathcal{S}(\mathbf{R}^n)$  be a fast decaying function whose Fourier transform  $\hat{\chi}$  vanishes in  $B_{k_0}(0)$  and let  $b \in L^2_\delta(X, E)$  for some  $\delta \in (-\varepsilon, \varepsilon)$  with  $\varepsilon = \varepsilon(k_0)$ . Then there exists  $a \in \mathcal{S}(\mathbf{R}^n, W_\delta^{2,2}(X, E))$  such that  $(\Delta_{\mathbf{R}^n} + D)a = \chi b$ .

We construct  $a \in \mathcal{S}(\mathbf{R}^n, W_\delta^{2,2}(X, E))$  using Fourier synthesis. By assumption  $\hat{\chi}(k) = 0$  for  $|k| \leq k_0$ . For  $|k| > k_0$  set

$$\hat{a}_k := (D + |k|^2)^{-1} b.$$

and define

$$a(x, y) := \int_{\mathbf{R}^n} e^{i\langle x, k \rangle} \hat{a}_k(y) \hat{\chi}(k) \, d\mathcal{L}^n(k).$$

Here  $\mathcal{L}^n$  denotes the  $n$ -dimensional Lebesgue measure on  $\mathbf{R}^n$ . Then

$$(\Delta_{\mathbf{R}^n} + D)a(x, y) = b\chi.$$

Moreover, one can verify that  $x \mapsto \|a(x, \cdot)\|_{W_\delta^{2,2}}$  is in  $\mathcal{S}(\mathbf{R}^n)$  using a slight variation of the proof that the Fourier transform maps fast decaying functions to fast decaying functions and the estimate  $\|\partial_k^\ell \hat{a}_k\|_{W_\delta^{2,2}} \leq c_\ell (1 + |k|)^\ell \|b\|_{L^2_\delta}$ .

**Step 2.** Let  $\chi \in \mathcal{S}(\mathbf{R}^n)$  with  $\hat{\chi}(0) = 0$ . Then there is a family  $(\chi_\varepsilon)_{\varepsilon > 0}$  of fast decaying functions such that  $\hat{\chi}_\varepsilon$  vanishes on  $B_\varepsilon(0)$  and  $\lim_{\varepsilon \rightarrow 0} \|\chi_\varepsilon - \chi\|_{L^1} = 0$ .

Pick a smooth function  $\rho: \mathbf{R} \rightarrow [0, 1]$  such that  $\rho(k) = 0$  for  $|k| \leq 1$  and  $\rho(k) = 1$  for  $|k| \geq 2$ . Set  $\hat{\chi}_\varepsilon(k) := \rho(|k|/\varepsilon) \hat{\chi}(k)$  and denote its inverse Fourier transform by  $\chi_\varepsilon$ . Then  $\chi_\varepsilon$  clearly satisfies the first part of the conclusion. To see that the second part also holds, note that from  $\hat{\chi}(0) = 0$  it follows that

$$\|\nabla^n(\hat{\chi}_\varepsilon - \hat{\chi})\|_{L^{2n/(2n-1)}} = O(\varepsilon^{1/2})$$

and therefore

$$\begin{aligned} \|\chi_\varepsilon - \chi\|_{L^1} &\leq \|(1 + |x|)^{-n}\|_{L^{2n/(2n-1)}} \cdot \|(1 + |x|)^n(\chi_\varepsilon - \chi)\|_{L^{2n}} \\ &\leq c(\|\hat{\chi}_\varepsilon - \hat{\chi}\|_{L^{2n/(2n-1)}} + \|\nabla^n(\hat{\chi}_\varepsilon - \hat{\chi})\|_{L^{2n/(2n-1)}}) = O(\varepsilon^{1/2}), \end{aligned}$$

where  $c > 0$  is a constant depending only on  $n$ . Here we used that the inverse Fourier transform is a bounded linear map from  $L^{2n/(2n-1)}$  to  $L^{2n}$  and the Fourier transform's behaviour with respect to derivatives.

**Step 3.** Suppose that  $(\Delta_{\mathbf{R}^n} + D)a = 0$ . Then for  $\sigma \in \mathcal{S}^n(\mathbf{R}^n)$ ,  $\delta \in \mathbf{R}^n$  and  $b \in C_c^\infty(X, E)$  we have

$$\int_{\mathbf{R}^n} \langle a(x, \cdot), b \rangle_{L^\infty, L^1} (\sigma(x + \delta) - \sigma(x)) \, d\mathcal{L}^n(x) = 0.$$

In particular, the conclusion of the lemma holds.

Set  $\chi(x) := \sigma(x + \delta) - \sigma(x)$ . Then  $\hat{\chi}(0) = 0$ . Let  $\chi_\varepsilon$  be as in Step 2. According to Step 1, for each  $\varepsilon > 0$  there is some small  $\delta > 0$  and  $c_\varepsilon \in \mathcal{S}(\mathbf{R}^n, W_{-\delta}^{2,2}(X, E))$  such that  $(\Delta_{\mathbf{R}^n} + D)c_\varepsilon = \chi_\varepsilon b$ . By the assumptions on  $a$  and since  $X$  has subexponential volume growth we have

$$\begin{aligned} & \int_{\mathbf{R}^n} \langle a(x, \cdot), b \rangle \chi(x) \, d\mathcal{L}^n(x) \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbf{R}^n} \langle a(x, \cdot), b \rangle \chi_\varepsilon(x) \, d\mathcal{L}^n(x) \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbf{R}^n} \int_X \langle a(x, y), (\Delta_{\mathbf{R}^n} + D)c_\varepsilon \rangle \, d\mathcal{L}^n(x) \, d\text{vol}(y) \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbf{R}^n} \int_X \langle (\Delta_{\mathbf{R}^n} + D)a(x, y), c_\varepsilon \rangle \, d\mathcal{L}^n(x) \, d\text{vol}(y) \\ &= 0. \end{aligned}$$

Since  $\sigma$ ,  $\delta$  and  $b$  are arbitrary, it follows that  $a$  is invariant in the  $\mathbf{R}^n$ -direction. This finishes the proof.  $\square$

*Remark A.4.* It is clear from the proof that in Lemma 7.10 one can replace the assumptions that  $X$  has subexponential volume growth and that  $\|a\|_{L^\infty}$  is finite by the assumption that  $\|a(x, \cdot)\|_{L^2_\delta}$  is bounded independent of  $x \in \mathbf{R}^n$  for all  $\delta > 0$ .

## References

- [ALR07] A. Adem, J. Leida, and Y. Ruan. *Orbifolds and stringy topology*. Vol. 171. Cambridge Tracts in Mathematics. Cambridge, 2007, pp. xii+149. DOI: 10.1017/CBO9780511543081. MR: 2359514 (cit. on pp. 5, 20).
- [Ber55] M. Berger. *Sur les groupes d'holonomie homogène des variétés à connexion affine et des variétés riemanniennes*. *Bulletin de la Société Mathématique de France* 83 (1955), pp. 279–330. MR: 0079806. Zbl: 0068.36002 (cit. on p. 5).
- [Bre03] S. Brendle. *On the construction of solutions to the Yang–Mills equations in higher dimensions*. 2003. arXiv: math/0302093v3 (cit. on pp. 3, 38, 39).
- [CDFN83] E. Corrigan, C. Devchand, D. B. Fairlie, and J. Nuyts. *First-order equations for gauge fields in spaces of dimension greater than four*. *Nuclear Physics. B* 214.3 (1983), pp. 452–464. DOI: 10.1016/0550-3213(83)90244-4. MR: 698892 (cit. on p. 7).
- [CGH00] D. M. J. Calderbank, P. Gauduchon, and M. Herzlich. *Refined Kato inequalities and conformal weights in Riemannian geometry*. *Journal of Functional Analysis* 173.1 (2000), pp. 214–255. DOI: 10.1006/jfan.2000.3563. MR: 1760284 (cit. on p. 16).
- [DK90] S. K. Donaldson and P. B. Kronheimer. *The geometry of four-manifolds*. Oxford Mathematical Monographs. Oxford Science Publications. New York, 1990, pp. x+440. MR: MR1079726. Zbl: 0904.57001 (cit. on pp. 8, 14, 15, 23, 28, 36, 38).

- [DS11] S. K. Donaldson and E. P. Segal. *Gauge theory in higher dimensions, II. Surveys in differential geometry. Volume XVI. Geometry of special holonomy and related topics*. Vol. 16. 2011, pp. 1–41. arXiv: 0902.3239. MR: 2893675. Zbl: 1256.53038 (cit. on pp. 1, 8).
- [DT98] S. K. Donaldson and R. P. Thomas. *Gauge theory in higher dimensions. The geometric universe (Oxford, 1996)*. Oxford, 1998, pp. 31–47. MR: MR1634503. Zbl: 0926.58003. URL: <http://www.ma.ic.ac.uk/~rpwt/skd.pdf> (cit. on pp. 1, 7, 8).
- [DW16] A. Degeratu and T. Walpuski. *Rigid HYM connections on tautological bundles over ALE crepant resolutions in dimension three. Symmetry, Integrability and Geometry: Methods and Applications (SIGMA)* 12.17 (2016). arXiv: 1207.6938 (cit. on p. 4).
- [Fee01] P. M. N. Feehan. *A Kato–Yau inequality and decay estimate for eigenspinors*. *J. Geom. Anal.* 11.3 (2001), pp. 469–489. DOI: 10.1007/BF02922015. MR: 1857853 (cit. on p. 16).
- [FG82] M. Fernández and A. Gray. *Riemannian manifolds with structure group  $G_2$* . *Ann. Mat. Pura Appl. (4)* 132 (1982), 19–45 (1983). DOI: 10.1007/BF01760975. MR: 696037. Zbl: 0524.53023 (cit. on p. 6).
- [FU91] D. S. Freed and K. K. Uhlenbeck. *Instantons and four-manifolds*. Second. Vol. 1. Mathematical Sciences Research Institute Publications. New York, 1991, pp. xxii+194. MR: MR1081321 (cit. on p. 17).
- [GN92] T. Gocho and H. Nakajima. *Einstein–Hermitian connections on hyper-Kähler quotients*. *J. Math. Soc. Japan* 44.1 (1992), pp. 43–51. DOI: 10.2969/jmsj/04410043. MR: 1139657 (cit. on p. 19).
- [Hay12] A. Haydys. *Gauge theory, calibrated geometry and harmonic spinors*. *Journal of the London Mathematical Society* 86.2 (2012), pp. 482–498. DOI: 10.1112/jlms/jds008. arXiv: 0902.3738. MR: 2980921. Zbl: 1256.81080 (cit. on p. 1).
- [HKLR87] N. J. Hitchin, A. Karlhede, U. Lindström, and M. Roček. *Hyper-Kähler metrics and supersymmetry*. *Communications in Mathematical Physics* 108.4 (1987), pp. 535–589. DOI: 10.1007/BF01214418. MR: 877637. Zbl: 0632.53073 (cit. on p. 15).
- [HL82] R. Harvey and H. B. Lawson Jr. *Calibrated geometries*. *Acta Math.* 148 (1982), pp. 47–157. DOI: 10.1007/BF02392726. MR: MR666108. Zbl: 0584.53021 (cit. on p. 5).
- [Joy00] D. D. Joyce. *Compact manifolds with special holonomy*. Oxford Mathematical Monographs. Oxford, 2000, pp. xii+436. MR: 1787733. Zbl: 1027.53052 (cit. on pp. 4–6, 11, 18, 31, 32, 34).
- [Joy96a] D. D. Joyce. *Compact Riemannian 7-manifolds with holonomy  $G_2$ . I*. *Journal of Differential Geometry* 43.2 (1996), pp. 291–328. DOI: 10.4310/jdg/1214458109. MR: MR1424428. Zbl: 0861.53022 (cit. on pp. 1–3, 12, 13).
- [Joy96b] D. D. Joyce. *Compact Riemannian 7-manifolds with holonomy  $G_2$ . II*. *Journal of Differential Geometry* 43.2 (1996), pp. 329–375. DOI: 10.4310/jdg/1214458110. MR: MR1424428. Zbl: 0861.53023 (cit. on pp. 1, 2, 12).

- [Keso8] T. M. Kessel. “Singular Bundles with Bounded  $L^2$ -curvatures”. PhD thesis. ETH Zürich, 2008 (cit. on p. 35).
- [KN90] P. B. Kronheimer and H. Nakajima. *Yang–Mills instantons on ALE gravitational instantons*. *Math. Ann.* 288.2 (1990), pp. 263–307. DOI: 10.1007/BF01444534. MR: 1075769 (cit. on p. 19).
- [Kor91] Yu. A. Kordyukov.  *$L^p$ -theory of elliptic differential operators on manifolds of bounded geometry*. *Acta Appl. Math.* 23.3 (1991), pp. 223–260. MR: 1120831 (cit. on pp. 38, 39).
- [Kovo3] A. Kovalev. *Twisted connected sums and special Riemannian holonomy*. *Journal für die Reine und Angewandte Mathematik* 565 (2003), pp. 125–160. DOI: 10.1515/crll.2003.097. MR: MR2024648. Zbl: 1043.53041 (cit. on p. 2).
- [Kro89a] P. B. Kronheimer. *A Torelli-type theorem for gravitational instantons*. *Journal of Differential Geometry* 29.3 (1989), pp. 685–697. MR: 992335. URL: <http://projecteuclid.org/getRecord?id=euclid.jdg/1214443067> (cit. on pp. 9, 10).
- [Kro89b] P. B. Kronheimer. *The construction of ALE spaces as hyper-Kähler quotients*. *Journal of Differential Geometry* 29.3 (1989), pp. 665–683. MR: 992334. URL: <http://projecteuclid.org/getRecord?id=euclid.jdg/1214443066> (cit. on p. 9).
- [Kro91] P. B. Kronheimer. *Instanton invariants and flat connections on the Kummer surface*. *Duke Math. J.* 64.2 (1991), pp. 229–241. DOI: 10.1215/S0012-7094-91-06411-2. MR: 1136374 (cit. on p. 3).
- [Lew98] C. Lewis. “Spin(7) instantons”. PhD thesis. 1998 (cit. on p. 3).
- [McK80] J. McKay. *Graphs, singularities, and finite groups*. *Proc. Sympos. Pure Math.* 37 (1980), pp. 183–186 (cit. on p. 10).
- [Nak88] H. Nakajima. *Compactness of the moduli space of Yang–Mills connections in higher dimensions*. *Journal of the Mathematical Society of Japan* 40.3 (1988), pp. 383–392. DOI: 10.2969/jmsj/04030383. MR: 945342 (cit. on p. 37).
- [Nak90] H. Nakajima. *Moduli spaces of anti-self-dual connections on ALE gravitational instantons*. *Invent. Math.* 102.2 (1990), pp. 267–303. DOI: 10.1007/BF01233429. MR: 1074476 (cit. on pp. 15, 16, 18).
- [NW73] L. Nirenberg and H. F. Walker. *The null spaces of elliptic partial differential operators in  $\mathbb{R}^n$* . *J. Math. Anal. Appl.* 42 (1973). Collection of articles dedicated to Salomon Bochner, pp. 271–301. MR: MR0320821 (cit. on p. 14).
- [PR03] F. Pacard and M. Ritoré. *From constant mean curvature hypersurfaces to the gradient theory of phase transitions*. *Journal of Differential Geometry* 64.3 (2003), pp. 359–423. MR: 2032110. URL: <http://projecteuclid.org/getRecord?id=euclid.jdg/1090426999> (cit. on pp. 3, 38).
- [Pri83] P. Price. *A monotonicity formula for Yang–Mills fields*. *Manuscripta Math.* 43.2-3 (1983), pp. 131–166. DOI: 10.1007/BF01165828. MR: MR707042 (cit. on p. 37).

- [Sá 11] H. N. Sá Earp.  *$G_2$ -instantons on Kovalev manifolds II*. 2011. arXiv: 1103.2138v1 (cit. on p. 2).
- [Sá 15] H. N. Sá Earp.  *$G_2$ -instantons over asymptotically cylindrical manifolds*. *Geometry and Topology* 19.1 (2015), pp. 61–111. DOI: 10.2140/gt.2015.19.61. MR: 3318748. Zbl: 1312.53044 (cit. on p. 2).
- [Sed82] S. Sedlacek. *A direct method for minimizing the Yang–Mills functional over 4-manifolds*. *Communications in Mathematical Physics* 86.4 (1982), pp. 515–527. MR: 679200. URL: <http://projecteuclid.org/getRecord?id=euclid.cmp/1103921842> (cit. on p. 35).
- [SW17] D. A. Salamon and T. Walpuski. *Notes on the octonions*. *Proceedings of the 23rd Gökova Geometry–Topology Conference*. 2017, pp. 1–85. arXiv: 1005.2820. MR: 3676083. Zbl: 06810387 (cit. on pp. 5, 6).
- [Tiao0] G. Tian. *Gauge theory and calibrated geometry. I*. *Annals of Mathematics* 151.1 (2000), pp. 193–268. DOI: 10.2307/121116. arXiv: math/0010015. MR: MR1745014. Zbl: 0957.58013 (cit. on p. 1).
- [TT04] T. Tao and G. Tian. *A singularity removal theorem for Yang–Mills fields in higher dimensions*. *Journal of the American Mathematical Society* 17.3 (2004), pp. 557–593. DOI: 10.1090/S0894-0347-04-00457-6. arXiv: math/0209352. MR: 2053951. Zbl: 1086.53043 (cit. on pp. 1, 20).
- [Uhl82a] K. K. Uhlenbeck. *Connections with  $L^p$  bounds on curvature*. *Communications in Mathematical Physics* 83.1 (1982), pp. 31–42. MR: MR648356. URL: <http://projecteuclid.org/getRecord?id=euclid.cmp/1103920743> (cit. on pp. 8, 35, 37).
- [Uhl82b] K. K. Uhlenbeck. *Removable singularities in Yang–Mills fields*. *Communications in Mathematical Physics* 83.1 (1982), pp. 11–29. MR: MR648355. URL: <http://projecteuclid.org/getRecord?id=euclid.cmp/1103920742> (cit. on pp. 15, 35).
- [Wal13a] T. Walpuski.  *$G_2$ -instantons on generalised Kummer constructions*. *Geometry and Topology* 17.4 (2013), pp. 2345–2388. DOI: 10.2140/gt.2013.17.2345. arXiv: 1109.6609. MR: 3110581. Zbl: 1278.53051 (cit. on p. 1).
- [Wal13b] T. Walpuski. “Gauge theory on  $G_2$ -manifolds”. PhD thesis. Imperial College London, 2013. URL: <https://spiral.imperial.ac.uk/bitstream/10044/1/14365/1/Walpuski-T-2013-PhD-Thesis.pdf> (cit. on p. 1).
- [Wal17] T. Walpuski.  *$G_2$ -instantons, associative submanifolds and Fueter sections*. *Communications in Analysis and Geometry* 25.4 (2017), pp. 847–893. DOI: 10.4310/CAG.2017.v25.n4.a4. arXiv: 1205.5350. MR: 3731643. Zbl: 06823232 (cit. on p. 4).