Geometry & Topology 17 (2013) 2345–2388

G_2–instantons on generalised Kummer constructions

THOMAS WALPUSKI

In this article we introduce a method to construct G_2–instantons on G_2–manifolds arising from Joyce's generalised Kummer construction [16; 17]. The method is based on gluing ASD instantons over ALE spaces to flat bundles on G_2–orbifolds of the form T^7/Γ. We use this construction to produce non-trivial examples of G_2–instantons.

53C07; 53C25, 53C38

1 Introduction

The seminal paper [8] of Donaldson–Thomas has inspired a considerable amount of work related to gauge theory in higher dimensions. Tian [34] and Tao–Tian [33] made significant progress on important foundational analytical questions. Recent work of Donaldson–Segal [7] and Haydys [14] 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 [20], 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 [30; 31]. The other (and historically the first) method for constructing G_2–manifolds is due to Joyce [16; 17] and is based on desingularising

Published: 22 August 2013 DOI: 10.2140/gt.2013.17.2345

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 ϕ_0 and let Γ be a finite group of diffeomorphisms of T^7 preserving ϕ_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 $\mathbb{C}^2/G_j$; moreover, the action of H_j on $T^3 = \mathbb{R}^3/\mathbb{Z}^3$ is induced by a free affine action on $\mathbb{R}^3$ normalising the action of $\mathbb{Z}^3$. 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 ρ_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 [16; 17] 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, ϕ_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 θ 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, $G = \mathrm{SO}(3)$.

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 θ 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, ϕ_t). Moreover, the adjoint bundle $\mathfrak{g}_E$ associated with E satisfies*

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

$$\textit{and} \quad \langle w_2(\mathfrak{g}_E), \Sigma\rangle = \langle w_2(\mathfrak{g}_{E_j}), \Sigma\rangle \tag{1-2}$$

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 4.7.

Remark 1.2 We will specify in Definition 3.6 and Definition 5.12, 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 1.3 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.

It is not unreasonable to expect that under certain topological assumptions all G_2–instantons on G_2–manifolds arsing 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 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 [25], unpublished work of Brendle on the Yang–Mills equation in higher dimension [3] and Pacard–Ritoré's work on the Allen–Cahn equation [29]. From a geometric perspective our result can be viewed as a higher-dimensional analogue of Kronheimer's work on ASD instantons on Kummer surfaces [23].

Here is an 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 $G = \mathrm{SO}(3)$ are constructed in Section 9.

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 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 [18], specifically Chapter 10.

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

$$\phi_0 := dx^{123} + dx^{145} + dx^{167} + dx^{246} - dx^{257} - dx^{347} - dx^{356}. \tag{2-1}$$

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 $\mathbb{R}^7$. The particular choice of ϕ_0 is not important. Any non-degenerate 3–form ϕ on $\mathbb{R}^7$ is equivalent to ϕ_0 under a change of coordinates; see, for example, Salamon–Walpuski [32, Theorem 3.2]. Here we say that ϕ is non-degenerate if for each non-zero vector $u \in \mathbb{R}^7$ the 2–form $i(u)\phi$ on $\mathbb{R}^7/\langle u \rangle$ is symplectic. It follows from the identity

$$i(u)\phi_0 \wedge i(v)\phi_0 \wedge \phi_0 = 6 g_{\mathbb{R}^7}(u,v)\mathrm{vol}_{\mathbb{R}^7} \tag{2-2}$$

that any element of GL(7) which preserves ϕ_0 also preserves the standard inner product $g_{\mathbb{R}^7}$ and the standard volume form $\mathrm{vol}_{\mathbb{R}^7}$ on $\mathbb{R}^7$. Therefore, G_2 is a subgroup of $SO(7)$. In particular, every non-degenerate 3–form ϕ 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 $\Lambda^3(\mathbb{R}^7)^*$. For ϕ 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 $\mathbb{R}^7/\langle u \rangle$.

From the above discussion it is clear that a non-degenerate 3–form ϕ 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, ϕ induces a Riemannian metric g_ϕ and an orientation on Y. The intrinsic torsion of the G_2–structure corresponding to ϕ can be identified with $\nabla_{g_\phi}\phi$.

Definition 2.1 A G_2*–manifold* is a 7–manifold Y equipped with a torsion-free G_2–structure ϕ, that is,

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

Remark 2.2 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 [1].) In this article, however, we will only encounter very simple G_2–orbifolds of the form $(Y/\Gamma, \phi)$ where (Y, ϕ) is a G_2–manifold and Γ is a finite group of diffeomorphism of Y preserving ϕ.

There is a plethora of reasons to be interested in G_2–manifolds. G_2–manifold have holonomy group $\mathrm{Hol}(g_\phi) \subset G_2$ which appears as one of the exceptional cases in Berger's classification of holonomy groups of irreducible non-symmetric Riemannian manifolds [2, Theorem 3]. G_2–manifolds are spin manifolds and carry (at least) one non-zero parallel spinor (see Joyce [18, 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 [13]: the *associative calibration* ϕ 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 2.3 The 7–torus $T^7 = \mathbb{R}^7/\mathbb{Z}^7$ equipped with the G_2–structure ϕ_0 defined in (2-1) is a G_2–manifold.

Definition 2.4 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 2.5 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 $(a_1, a_2, a_3) \in S^2 \subset \mathbb{R}^3$.

Example 2.6 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 ϕ 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 ϕ 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

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

This immediately yields a orientation-preserving isometry $T_x(T^3 \times X) \to \mathbb{R}^7$ identifying ϕ with ϕ_0. Note that in the current example the coassociative calibration $\psi := *\phi$ is given by

$$\psi = \tfrac{1}{2}\omega_1 \wedge \omega_1 + \delta^2 \wedge \delta^3 \wedge \omega_1 + \delta^3 \wedge \delta^1 \wedge \omega_2 - \delta^1 \wedge \delta^2 \wedge \omega_3. \tag{2-4}$$

Remark 2.7 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, ϕ) satisfies $\mathrm{Hol}(g_\phi) = G_2$ if and only if $\pi_1(Y)$ is finite; see Joyce [18, Proposition 10.2.2].

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

Theorem 2.8 (Fernández–Gray [10, 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 $\Theta: \mathcal{P} \to \Omega^4(Y)$ by*

$$\Theta(\phi) := *_\phi \phi. \tag{2-5}$$

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

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

The key difficulty in constructing G_2–manifolds comes from the fact that Θ 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 [10], that will be useful in the following. We refer the interested reader to Salamon–Walpuski [32, Theorem 8.4] for a detailed proof.

Proposition 2.9 *There is a G_2–invariant orthogonal splitting*

$$\Lambda^2(\mathbb{R}^7)^* = \Lambda^2_7 \oplus \Lambda^2_{14},$$

where

$$\Lambda^2_7 := \{\omega : *(\omega \wedge \phi_0) = 2\omega\} \quad \textit{and} \quad \Lambda^2_{14} := \{\omega : *(\omega \wedge \phi_0) = -\omega\}.$$

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

3 Gauge theory on G_2–manifolds

Let (Y, ϕ) 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

$$*(F_A \wedge \phi) = -F_A. \tag{3-1}$$

These equations have first appeared in the physics literature (see Corrigan–Devchand–Fairlie–Nuyts [5]) and were later brought to a wider attention by Donaldson–Thomas [8, Section 3]. Equation (3-1) 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 3.2 Flat connections are G_2–instantons.

Example 3.3 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 2.6 is a G_2–instanton:

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

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

Example 3.4 The Levi-Civita connection on a G_2–manifold (Y, ϕ) 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 $\mathrm{Hol}(g_\phi) \subset G_2$. But then it follows from Proposition 2.9 that $*(R \wedge \phi) = -R$.

Since ϕ 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 $\mathrm{YM}\colon \mathscr{A}(E) \to \mathbb{R}$, since

$$\mathrm{YM}(A) := \int_Y |F_A|^2 \mathrm{dvol} = \tfrac{1}{3} \int_Y |F_A + *(F_A \wedge \phi)|^2 \mathrm{dvol} - \int_Y \langle F_A \wedge F_A \rangle \wedge \phi \tag{3-2}$$

and, by Chern–Weil theory, the second term is a topological constant depending only on E. The energy identity (3-2) follows from a straight-forward computation using Proposition 2.9.

Proposition 3.5 *Let $A \in \mathscr{A}(E)$ 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*

$$*(F_A \wedge \psi) + d_A \xi = 0. \tag{3-3}$$

Proof The equivalence of (1) and (2) follows immediately from Proposition 2.9. 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). □

From Proposition 3.5 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 $CS^\psi \colon \mathcal{A}(E) \to \mathbb{R}$ defined by

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

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 [8] and is one of the main motivations for studying G_2–instantons. Since Equation (3-1) 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, ϕ):

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

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–Segal [7, 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 [35]. For a careful discussion of how the Coulomb gauge is used in the construction moduli spaces we refer the reader to Donaldson–Kronheimer [6, Section 4.2].) For a fixed connection $A \in \mathcal{A}(E)$ we consider the system of equations

$$*(F_{A+a} \wedge \psi) + d_{A+a} \xi = 0 \quad \text{and} \quad d_A^* a = 0 \tag{3-4}$$

for $\xi \in \Omega^0(Y, \mathfrak{g}_E)$ and $a \in \Omega^1(Y, \mathfrak{g}_E)$. This is simply (3-3) for $A + a$ instead of A together with the condition that $A + a$ be in Coulomb gauge with respect to A. The linearisation $L_A\colon \Omega^0(Y, \mathfrak{g}_E) \oplus \Omega^1(Y, \mathfrak{g}_E) \to \Omega^0(Y, \mathfrak{g}_E) \oplus \Omega^1(Y, \mathfrak{g}_E)$ of (3-4) is given by

$$L_A := \begin{pmatrix} 0 & \mathrm{d}_A^* \\ \mathrm{d}_A & *(\psi \wedge \mathrm{d}_A) \end{pmatrix}. \tag{3-5}$$

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

Definition 3.6 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 $\mathscr{M}(E, \phi)$ 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 ϕ_0, as in Example 2.3, and let Γ be a finite group of diffeomorphisms of T^7 preserving ϕ_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 $\mathbb{C}^2/G_j$; moreover, the action of H_j on $T^3 = \mathbb{R}^3/\mathbb{Z}^3$ is induced by a free affine action on $\mathbb{R}^3$ normalising the action of $\mathbb{Z}^3$.

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 that ζ to S, then T decomposes into connected components $T_1, \dots, T_k$ such that T_j contains S_j and is isometric to $(T^3 \times B^4_\zeta/G_j)/H_j$. 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 $\mathrm{SU}(2)$. 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 $\pi\colon X \to \mathbb{C}^2/G$ inducing a diffeomorphism from $X \setminus \pi^{-1}(0)$ to $(\mathbb{C}^2 \setminus \{0\})/G$ such that

$$\nabla^k(\pi_*\hat{g} - g) = O\big(r^{-4-k}\big) \quad \text{and} \quad \nabla^k(\pi_*\hat{I}_i - I_i) = O\big(r^{-4-k}\big) \tag{4-1}$$

as $r \to \infty$ for $i = 1, 2, 3$ and $k \geq 0$. Here $r\colon \mathbb{C}^2/G_j \to [0, \infty)$ 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 [21; 22], ALE spaces are very well understood.

Theorem 4.3 (Kronheimer [22, Theorems 1.1, 1.2 and 1.3]) *Let G be a non-trivial finite subgroup of $\mathrm{SU}(2)$. Denote by X the real 4–manifold underlying the crepant resolution $\widetilde{\mathbb{C}^2/G}$. Then for each three cohomology classes $\alpha_1, \alpha_2, \alpha_3 \in H^2(X, \mathbb{R})$ satisfying*

$$(\alpha_1(\Sigma), \alpha_2(\Sigma), \alpha_3(\Sigma)) \neq 0 \in \mathbb{R}^3 \tag{4-2}$$

for each $\Sigma \in H_2(X, \mathbb{Z})$ 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 α_i. Moreover, each ALE space asymptotic to $\mathbb{C}^2/G$ is diffeomorphic to $\widetilde{\mathbb{C}^2/G}$ and its associated triple of Kähler classes satisfies (4-2).

Remark 4.4 The crepant resolution $\widetilde{\mathbb{C}^2/G}$ can be obtained from $\mathbb{C}^2/G$ by a sequence of blow-ups. The exceptional divisor E of $X = \widetilde{\mathbb{C}^2/G}$ has irreducible components $\Sigma_1, \dots, \Sigma_k$. By the McKay correspondence [26], these components form a basis of $H_2(X, \mathbb{Z})$ 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 $\mathrm{SU}(2)$.

Definition 4.5 A collection $\mathbf{r} = \{(X_j, \rho_j)\}$ consisting of, for each j, an ALE space X_j asymptotic to $\mathbb{C}^2/G_j$ together with an isometric action ρ_j of H_j on X_j which is asymptotic to the action of H_j on $\mathbb{C}^2/G_j$ is called *resolution data for Y_0*.

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

$$\pi_{j,t} := t\pi_j : X_j \to \mathbb{C}^2/G_j \tag{4-3}$$

and set

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

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 4.6 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 4.7 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 2.7. In particular, it can be seen that every $\Sigma \in H_2(X_j, \mathbb{Z})$ invariant under the action of H_j yields a cohomology class $\Sigma \in H_2(Y, \mathbb{Z})$. Also each component of singular set S_j gives rise to a *rational* homology class

$$[S_j] := \frac{1}{|H_j|}(\iota_{j,t})_*\big(T^3 \times \{x\}\big) \in H_3(Y, \mathbb{Q}), \tag{4-5}$$

where $\iota_{j,t}\colon T^3 \times \pi_{j,t}^{-1}(B_\zeta^4/G_j) \to Y$ 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 ϕ_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

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

with $\nabla^k \varrho_{j,t,i} = t^4 O(r^{-3-k})$ for $k \geq 0$; see Joyce [18, Theorem 8.2.3]. Now, fix a smooth non-decreasing function $\chi\colon [0,\zeta] \to [0,1]$ such that $\chi(s) = 0$ for $s \leq \zeta/4$ and $\chi(s) = 1$ for $s \geq \zeta/2$ and set

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

Then $(\pi_{j,t})_*\tilde{\omega}_{j,t,i}$ and ω_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 $\widetilde{T}_{j,t}$. Define the function $r_t\colon Y_t \to [0,\zeta]$ by

$$r_t(p) := \begin{cases} |\pi_{j,t}(y)| & \text{for } p = [(x,y)] \in \widetilde{T}_{j,t} \\ \zeta & \text{for } p \in Y_t \setminus \widetilde{T}_t \end{cases} \tag{4-6}$$

and set

$$R_{j,t} := \widetilde{T}_{j,t} \cap r_t^{-1}[\zeta/4,\zeta/2] \quad \text{and} \quad R_t := \bigcup_j R_{j,t} = r_t^{-1}[\zeta/4,\zeta/2]. \tag{4-7}$$

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

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

Remark 4.8 Note that on the complement of $\widetilde{T}_t$ the metric $\widetilde{g}_t$ agrees with the flat metric g_0 on $(T^7/\Gamma) \setminus T$ and on $\widetilde{T}_{j,t} \setminus R_{j,t}$ 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 $\nabla^k(\widetilde{g}_t - g_{\mathbb{R}^3} \oplus t^2 g_{X_j}) = O(t^4)$ for $k \geq 0$ and, for each fixed $\epsilon > 0$, on $r_t^{-1}[\epsilon,\zeta]$ we have $\nabla^k(\widetilde{g}_t - g_0) = O(t^4)$ for $k \geq 0$.

Theorem 4.9 (Joyce [16, Theorems A and B]; [17, Theorem 2.2.1]) *There are constants $T, c > 0$ and for each $t \in (0,T)$ a 2–form η_t on Y_t such that $\phi_t := \widetilde{\phi}_t + \mathrm{d}\eta_t$ defines a torsion-free G_2–structure and*

$$\|\mathrm{d}\eta_t\|_{L^\infty} \leq c t^{1/2}. \tag{4-8}$$

Remark 4.10 In view of Theorem 2.8 the above is tantamount to saying that one can solve the non-linear partial differential equation

$$\mathrm{d}\Theta\big(\widetilde{\phi}_t + \mathrm{d}\eta_t\big) = 0 \tag{4-9}$$

with estimates on $\mathrm{d}\eta_t$. For small η_t, the dominant part of this equation is essentially the Laplacian on 2–forms. Now, as $t > 0$ decreases the size of $\mathrm{d}\Theta(\widetilde{\phi}_t)$ becomes smaller and smaller, but at the same time the mapping properties of the Laplacian degenerate. Solving (4-9) thus is a rather delicate balancing act.

For our application we need to slightly strengthen the estimate in Theorem 4.9. Let $w_t(x,y) := t + \min\{r_t(x), r_t(y)\}$. For a Hölder exponent $\alpha \in (0,1)$ define

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

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

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 4.11 *The constants $T, c > 0$ in Theorem 4.9 can be chosen such that for all $t \in (0,T)$ we have*

$$\|d\eta_t\|_{C^{0,\alpha}_{0,t}} \le ct^{1/2} \quad \textit{and} \quad \big\|\Theta(\phi_t) - \Theta\big(\hat{\phi}_{j,t}\big)\big\|_{C^{0,\alpha}_{0,t}(\tilde{T}_{j,t})} \le ct^{1/2}.$$

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

Proposition 4.12 *For each $\mu > 0$ and $K \in \mathbb{N}_0$ there exists a constant $\epsilon > 0$ such that the following holds for all $t \in (0,T)$ and $p \in Y_t$: $R := \epsilon(t + r_t(p))$ is less than the injectivity radius of $(Y_t, \tilde{g}_t)$ at p and if we identify T_pY isometrically with $\mathbb{R}^7$ and denote by $s_R\colon B_1 \to B_R(p)$ the map obtained by multiplication with R followed by the exponential map, then*

$$\big|\partial^k\big(R^{-2}s_R^*\tilde{g}_t - g_{\mathbb{R}^7}\big)\big| \le \mu \tag{4-10}$$

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

Proof From Remark 4.8 it is clear that we can find $\epsilon > 0$ 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 (4-10) is equivalent to

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

where $\tilde{R} := \epsilon(1 + |\pi_j(y)|)$ and $p = [(x,y)]$. Because of (4-1) this holds for all $\epsilon \le \frac{1}{2}$ as long as $|\pi_j(y)|$ is sufficiently large, say, $|\pi_j(y)| > N$. For $|\pi_j(y)| \le N$ it can be arranged to hold by choosing $\epsilon > 0$ sufficiently small. □

Proof of Proposition 4.11 Note that the second part follows from the first and the construction of $\tilde{\phi}_t$, because Θ is a smooth map. To obtain the estimate on $d\eta_t$ recall from Joyce's construction that η_t solves a non-linear partial differential equation that can be written schematically as

$$d^*d\eta_t + P(d\eta_t, \nabla d\eta_t) = G(d\eta_t, \dots) \quad \text{and} \quad d^*\eta_t = 0; \tag{4-11}$$

see Joyce [16, 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

$$\|G(\mathrm{d}\eta_t,\ldots)\|_{L^\infty} \le ct^{1/2}. \tag{4-12}$$

Now, define

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

Since $\mathrm{d}\eta_t$ is small provided $T>0$ is small, this a small perturbation of the operator $\mathrm{d}^*\oplus\mathrm{d}$. We extend D_t to an operator from $\Omega^*(Y_t)$ to itself by defining $D_t\sigma=(\mathrm{d}^*\oplus\mathrm{d})\sigma$ 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 $\epsilon\in(0,\frac{1}{2})$ such that for all $t\in(0,T)$ and each $p\in Y_t$ the following holds:

$$R^\alpha[\sigma]_{C^{0,\alpha}(B_{R/2}(p))} \le c\big(R\|D_t\sigma\|_{L^\infty(B_R(p))} + \|\sigma\|_{L^\infty(B_R(p))}\big) \tag{4-13}$$

with $R:=\epsilon(t+r_t(p))$. From this the asserted bound on $[\mathrm{d}\eta_t]_{C^{0,\alpha}_{0,t}}$ follows at once using (4-8), (4-11) and (4-12), since on $B_{R/2}(p)$ we have $w_t\le 2\epsilon^{-1}R$.

For $\mu>0$ choose $\epsilon>0$ according to Proposition 4.12 with $K=1$. Let $s_R\colon B_1^7\to B_R(p)$ be as in Proposition 4.12. We define a rescaled operator $\widetilde{D}_{t,p}\colon\Omega^*(B_1)\to\Omega^*(B_1)$ by

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

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 4.9 and Proposition 4.12 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 $\widetilde{D}_{t,p}$ is as close to $\mathrm{d}\oplus\mathrm{d}^*\colon\Omega^*(B_1)\to\Omega^*(B_1)$ as we wish. In particular, we can arrange that the family of operators $\widetilde{D}_{t,p}$ 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:

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

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

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

with $c>0$ independent of $t\in(0,T)$ and $p\in Y_t$. This, however, is equivalent to the estimate (4-13) for the unscaled operator D_t. □

Remark 4.13 Proposition 4.11 can be viewed as a quantification of Joyce's proof of the fact that η_t is smooth. In a similar fashion, one can also obtain estimates on higher Hölder norms of $\mathrm{d}\eta_t$.

Remark 4.14 The kind of argument we used above goes back to work of Nirenberg–Walker [28, Theorem 3.1]. We will encounter this line of reasoning again in the proofs of Propositions 5.8 and 7.6.

5 ASD instantons on ALE spaces

Let Γ 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 $\mathscr{A}(E)$ the space of connections on E.

Definition 5.1 A *framing at infinity* of E is a bundle isomorphism $\Phi\colon E_\infty|_U \to \pi_* E|_U$ where E_∞ 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 θ be a flat connection on a G–bundle E_∞ over $(\mathbb{C}^2 \setminus \{0\})/\Gamma$.

Definition 5.2 Let $\Phi\colon E_\infty|_U \to \pi_* E|_U$ be a framing at infinity of E. Then a connection $A \in \mathscr{A}(E)$ is called *asymptotic to θ at rate δ with respect to Φ* if

$$\nabla^k(\Phi^* A - \theta) = O\big(r^{\delta-k}\big) \tag{5-1}$$

for all $k \geq 0$. Here ∇ is the covariant derivative associated with θ.

Definition 5.3 A *framed ASD instanton asymptotic to θ (at rate δ)* is an ASD instanton $A \in \mathscr{A}(E)$ on E together with a framing at infinity Φ of E such that A is asymptotic to θ at rate δ with respect to Φ. If no rate δ is specified, then we take $\delta = -3$.

Proposition 5.4 *Let $A \in \mathscr{A}(E)$ be an ASD instanton on E with finite energy, that is,*

$$\int_X |F_A|^2 \mathrm{dvol} < \infty,$$

then there is a G–bundle E_∞ over $(\mathbb{C}^2 \setminus \{0\})/\Gamma$ together with a flat connection θ and a framing $\Phi\colon E_\infty|_U \to \pi_* E|_U$ such that (5-1) *holds with $\delta = -3$*

Proof We extend the argument in Donaldson–Kronheimer [6, 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 $\varphi\colon B_\epsilon/\Gamma \to \hat{X}$ which is constructed as follows. Fix an orientation reversing linear isometry σ of $\mathbb{R}^4$. We let Γ act on B_ϵ 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 [36, Theorem 4.1], the pullback of A to $B_\epsilon \setminus \{0\}$ extends to a Γ–invariant ASD instanton over all of B_ϵ. Hence, A extends to an ASD instanton $\hat{A}$ on an orbifold G–bundle $\hat{E}$ over $\hat{X}$. Using radial parallel transport from ∞ we obtain a trivialisation of $\hat{E}$ over $\varphi(B_\epsilon/\Gamma)$ in which the connection matrix representing $\hat{A}$ vanishes at $\infty = \varphi(0)$. Denote by $\rho\colon \Gamma \to G$ the monodromy representation associated with $\hat{E}|_\infty$. Associated with ρ there are a G–bundle E_∞ over $\varphi((B_\epsilon \setminus \{0\})/\Gamma)$ and a flat connection θ on E_∞. The above trivialisation of $\hat{E}$ over $\varphi(B_\epsilon/\Gamma)$ amounts to a bundle isomorphism $\Phi\colon E_\infty \to \hat{E}|_{\varphi(B_\epsilon \setminus \{0\}/\Gamma)}$ and the fact that the connection matrix representing $\hat{A}$ vanishes at $\infty = \varphi(0)$ implies that $\nabla^k(\varphi^*(\Phi^*(\hat{A}) - \theta)) = O(x^{1-k})$ 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 $\nabla^k(\Phi^* A - \theta) = O(r^{-3-k})$. □

Let us briefly discuss moduli spaces of framed ASD instantons on E asymptotic to θ. For a detailed discussion we refer the reader to Nakajima's beautiful article [27]. Fix a framing at infinity Φ of E, a rate $\delta \in (-3,-1)$ and denote by $\mathscr{A}(E,\theta)$ the space of all connections asymptotic to θ at rate δ with respect to Φ. Similarly, define $\mathscr{G}(E)$ to be the group of gauge transformations asymptotic to a constant element of G at infinity at rate $\delta + 1$ with respect to Φ. Denote by $g_\infty\colon \mathscr{G}(E) \to G$ the homomorphism assigning to each gauge transformation its asymptotic value at infinity and let $\mathscr{G}_0(E) := \ker g_\infty \subset \mathscr{G}(E)$ be the based gauge group consisting of gauge transformations asymptotic to the identity. Then the space

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

is called the *moduli space of framed ASD instantons on E asymptotic to θ*.

Remark 5.5 The space does not depend on the choice of $\delta \in (-3,-1)$. This is a consequence of Proposition 5.4.

Remark 5.6 If we denote by $\rho\colon \Gamma \to G$ the monodromy representation associated with θ and by $G_\rho := \{g \in G : g\rho g^{-1} = \rho\}$ the stabiliser of ρ, then $G_\rho \subset G \cong \mathscr{G}(E)/\mathscr{G}_0(E)$ acts on $M(E,\theta)$.

Theorem 5.7 (Nakajima [27, 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–Roček [15]). The space $\mathscr{A}(E,\theta)$ inherits a hyperkähler structure from X and the action of the based gauge group $\mathscr{G}_0$ 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–Kronheimer [6, Section 4.2.5]. This can be done using weighted Sobolev space completions of $\mathscr{A}(E,\theta)$ and $\mathscr{G}_0(E)$; see Nakajima [27, Section 2] for a detailed discussion. An important role is played by the operator $\delta_A\colon \Omega^1(X,\mathfrak{g}_E) \to \Omega^0(X,\mathfrak{g}_E) \oplus \Omega^+(X,\mathfrak{g}_E)$ defined by

$$\delta_A(a) := \left(\mathrm{d}_A^* a, \mathrm{d}_A^+ a\right) \tag{5-2}$$

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

Proposition 5.8 *Let $A \in \mathscr{A}(E)$ 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 5.9 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 [27, Proposition 5.1]). By the first part the tangent space of $M(E,\theta)$ at $[A]$ agrees with the L^2 kernel of δ_A and thus the formal hyperkähler structure is indeed well-defined.

The proof of Proposition 5.8 rests on the following refined Kato inequality.

Proposition 5.10 *Let $A \in \mathscr{A}(E)$ be an ASD instanton on E. If $a \in \Omega^1(X,\mathfrak{g}_E)$ satisfies $\delta_A a = 0$, then*

$$|\mathrm{d}|a|| \leq \sqrt{\tfrac{3}{4}}\,|\nabla_A a| \tag{5-3}$$

on the complement of the vanishing locus of a.

Proof Recall that the Kato inequality follows from the Cauchy–Schwarz inequality $|\langle \nabla_A a, a\rangle| \leq |\nabla_A \alpha||\alpha|$. 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 (5-3) holds with some constant $\epsilon < 1$ instead of $\sqrt{3/4}$.

To see that one can take $\epsilon = \sqrt{3/4}$ we follow an argument of Feehan [9, Section 3]; however, also note that we could simply read off the value from the table given in Calderbank [4, Appendix]. We can write δ_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 γ 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 $\mathrm{d}|a|(x) \neq 0$ 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

$$\big|\mathrm{d}|a|\big|^2 = \big|\nabla_{e_1}|a|\big|^2 \le \big|\nabla^A_{e_1} a\big|^2 = \big|\gamma(e_1)\nabla^A_{e_1} a\big|^2 = \Big|\textstyle\sum_{i\ge 2} \gamma(e_i)\nabla^A_{e_i} a\Big|^2 \le 3\sum_{i\ge 2} \big|\nabla^A_{e_i} a\big|^2$$

and therefore

$$4\big|\mathrm{d}|a|\big|^2 = 4\big|\nabla^A_{e_1} a\big|^2 \le 3\sum_i \big|\nabla^A_{e_i} a\big|^2 = 3\big|\nabla_A a\big|^2.$$

This finishes the proof. □

Proof of Proposition 5.8 First of all note that (1) implies (2), because if $\delta_A^*(\xi,\omega) = 0$, then $\mathrm{d}_A^*\mathrm{d}_A\xi = 0$ and $\mathrm{d}_A^+\mathrm{d}_A\xi = [F_A^+, \xi] = 0$; therefore $\mathrm{d}_A\xi = O(|\pi|^{-3})$. Thus integration by parts yields $\mathrm{d}_A\xi = 0$ 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 4.11. 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

$$R^k \|\nabla_A^k a\|_{L^\infty(B_{R/2}(x))} \le c\|a\|_{L^\infty(B_R(x))} \tag{5-4}$$

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 5.4. We rescale to a ball of radius one and denote the rescaled connection by $\widetilde{A}$ and the rescaling of δ_A by $\widetilde{D}_x$. Then the family of operators $\widetilde{D}_x$ 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:

$$\big\|\nabla^k_{\widetilde{A}} a\big\|_{L^\infty(B_{1/2})} \le c\big(\|\widetilde{D}_x a\|_{C^{k,\alpha}(B_1)} + \|a\|_{L^\infty(B_1)}\big).$$

If a is in the kernel of $\widetilde{D}_x$, the first term vanishes. Rescaling this inequality yields (5-4) for $a \in \ker\delta_A$ and $|\pi(x)| > R$. For $1/2 \le |\pi(x)| \le R_0$, (5-4) follows from standard Schauder estimates.

Let us now prove (1) for $k = 0$. Recall, for example, from Freed–Uhlenbeck [11, Equation (6.25)], that the operator $\widetilde{\delta}_A\colon \Omega^1(X, \mathfrak{g}_E) \to \Omega^0(X, \mathfrak{g}_E) \oplus \Omega^+(X, \mathfrak{g}_E)$ defined by $\widetilde{\delta}_A(a) := (\mathrm{d}_A^* a, \sqrt{2}\mathrm{d}_A^+ a)$ satisfies a Weitzenböck formula of the form

$$\widetilde{\delta}_A^*\widetilde{\delta}_A a = \nabla_A^*\nabla_A a + \{\mathrm{Ric}, a\} + \{F_A^-, a\}. \tag{5-5}$$

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 $\widetilde{\delta}_A a = 0$. Then Proposition 5.10, the identity

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

(see [11, Equation (6.18)]) and the Weitzenböck formula (5-5) yield the following estimate on the complement of the vanishing locus of a:

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

In the last step we used $\widetilde{\delta}_A a = 0$ and $|F_A^-| = O(|\pi|^{-4})$, which is a consequence of Proposition 5.4.

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,

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

for some constant $c > 0$. Since f is bounded, by Joyce [18, 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 $f \le g$ and thus $f \le g = O(|\pi|^{-1})$. Now, $(\Delta f)^+ = O(|\pi|^{-5})$ on U and an application of [18, 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. □

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

Theorem 5.11 (Nakajima [27, Theorem 2.7]) *Let A be a framed ASD instanton asymptotic to θ. Then the dimension of the L^2 kernel of δ_A is given by*

$$\dim\ker\delta_A = -2\int_X p_1(\mathfrak{g}_E) + \frac{2}{|\Gamma|}\sum_{g\in\Gamma\setminus\{e\}} \frac{\chi_{\mathfrak{g}}(g)-\dim\mathfrak{g}}{2-\operatorname{tr} g}. \tag{5-6}$$

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 Γ acting on $\mathfrak{g}$, the Lie algebra associated with G, via the monodromy representation $\rho\colon \Gamma \to G$ of θ.

Proof Let us briefly explain how to derive (5-6) from Nakajima's formula, which can be written as

$$\begin{aligned}\dim\ker\delta_A = -\int_X (\dim\mathfrak{g} + p_1(\mathfrak{g}_E))\operatorname{ch}\bigl(S^+\bigr)\hat{A}(X)\\ + \dim\mathfrak{g}^\Gamma + \frac{1}{|\Gamma|}\sum_{g\in\Gamma\setminus\{e\}}\chi_{\mathfrak{g}}(g)\frac{\operatorname{tr} g}{2-\operatorname{tr} g}.\end{aligned} \tag{5-7}$$

Here $\mathfrak{g}^\Gamma$ denotes the Γ–invariant part of $\mathfrak{g}$, S^+ denotes the positive spin bundle on X, and $\operatorname{ch}(S^+)$ 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 δ_A, then it follows from the fact that X is Ricci-flat and the Weitzenböck formula (5-5) that $\nabla^*\nabla a = 0$ and then by integration by parts, which is justified because of the decay asserted by Proposition 5.8, that $\nabla a = 0$. Since a lies in L^2, it necessarily vanishes. Therefore $\dim\ker\delta_A = 0$ and (5-7) yields

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

By plugging this back into (5-7) we obtain

$$\dim\ker\delta_A = -2\int_X p_1(\mathfrak{g}_E) + \dim\mathfrak{g}^\Gamma - \dim\mathfrak{g} + \frac{1}{|\Gamma|}\sum_{g\in\Gamma\setminus\{e\}}(\chi_{\mathfrak{g}}(g)-\dim\mathfrak{g})\frac{\operatorname{tr} g}{2-\operatorname{tr} g}.$$

Since

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

this leads to the index formula (5-6) given above. □

There is a very rich existence theory for ASD instantons on ALE spaces. Gocho–Nakajima [12] observed that for each representation $\rho\colon \Gamma \to U(n)$ there is a bundle $\mathcal{R}_\rho$ over X together with an ASD instanton A_ρ asymptotic to the flat connection determined by ρ, and if σ is a further representation of Γ, then $A_{\rho\oplus\sigma} = A_\rho \oplus A_\sigma$. Kronheimer–Nakajima [24] 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 5.12 An ASD instanton A is called *infinitesimally rigid* if the L^2 kernel of the linear operator δ_A is trivial.

Theorem 5.13 (Kronheimer–Nakajima [24, Lemma 7.1]) *For each $\rho\colon \Gamma \to U(n)$ the ASD instanton A_ρ 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 [24, Appendix A]. Let $\Delta(\Gamma)$ denote the Dynkin diagram associated with Γ in the ADE classification of the finite subgroups of $SU(2)$. 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 5.14 (Kronheimer–Nakajima [24, Appendix A]) *The harmonic 2–forms $c_1(\mathcal{R}_j) = \frac{i}{2\pi} \operatorname{tr} F_{A_j}$ form a basis of $L^2\mathcal{H}^2(X) \cong H^2(X,\mathbb{R})$ 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 $\kappa \in \mathrm{Aut}(H_2(X,\mathbb{Z}),\cdot)$ such that $\{c_1(\mathcal{R}_j)\}$ is dual to $\{\kappa[\Sigma_j]\}$, where Σ_j are the irreducible components of the exceptional divisor E of $\widetilde{\mathbb{C}^2/\Gamma}$. If X is isomorphic to $\widetilde{\mathbb{C}^2/\Gamma}$ as a complex manifold, then $\kappa = \mathrm{id}$.

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

Proposition 5.15 *Let X be an ALE space asymptotic to $\mathbb{C}^2/\mathbb{Z}_k$. Denote by $\rho_j\colon \mathbb{Z}_k \to U(1)$ the irreducible representation defined by $\rho_j(\ell) = \exp(\frac{2\pi i}{k} j\ell)$. For $n, m \in \mathbb{Z}_k$, let $E_{n,m}$ be the $SO(3)$–bundle underlying $\mathbb{R} \oplus (\mathcal{R}_n^* \otimes \mathcal{R}_{n+m})$ 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 ρ_m and*

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

as well as

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

Proof To see that $A_{n,m}$ is infinitesimally rigid apply Theorem 5.13 to $A_n \oplus A_{n+m}$ and observe that $\mathfrak{g}_{E_{n,m}} = \mathbb{R} \oplus (\mathcal{R}_n^* \otimes \mathcal{R}_{n+m})$ 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 5.14 or by noting that the first term in the index formula (5-6) is precisely twice the energy and the second term is given by $(-\frac{2}{k})$–times

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

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 4.9. Denote by $\psi_t := \Theta(\phi_t)$ the coassociative calibration on Y_t. If θ is a flat connection on a G–bundle E_0 over Y_0, then the monodromy of θ around S_j induces a representation $\mu_j \colon \pi_1(T_j, x_j) \cong (\mathbb{Z}^3 \times G_j) \rtimes H_j \to G$ of the orbifold fundamental group of T_j based at $x_j \in T_j \setminus S_j$.

Remark 6.1 For a general definition of orbifold fundamental group we refer the reader to Adem–Leida–Ruan [1, 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, since the singular sets have sufficiently large codimension.

Definition 6.2 A collection $\mathbf{g} = ((E_0, \theta), \{(x_j, f_j)\}, \{(E_j, A_j, \tilde{\rho}_j, m_j)\})$ consisting of E_0 and θ 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 $f_j \colon (E_0)_{x_j} \to G$ 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 $(\mathbb{C}^2 \setminus \{0\})/G_j$ induced by the representation $\mu_j|_{G_j}$,
- a lift $\tilde{\rho}_j$ of the action ρ_j of H_j on X_j to E_j and
- a homomorphism $m_j \colon \mathbb{Z}^3 \to \mathcal{G}(E_j)$

is called *gluing data compatible with* $\mathbf{r} = \{(X_j, \rho_j)\}$ if the following compatibility conditions are satisfied:

- The action $\tilde{\rho}_j$ of H_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 H_j on $E_{\infty,j}$. Note that the lift of the action of H_j on $E_{\infty,j}$ to the trivial bundle $G \times (\mathbb{C}^2 \setminus \{0\})$ is given by $h \cdot (g, x) = (\mu_j(h) \cdot g, h \cdot x)$.
- The action of $\mathbb{Z}^3$ on E_j given by m_j preserves A_j and m_j is asymptotic at infinity to $\mu_j|_{\mathbb{Z}^3}$, that is, $g_\infty \circ m_j = \mu_j|_{\mathbb{Z}^3}$ with $g_\infty \colon \mathcal{G}(E_j) \to G$ as in the paragraph following the proof of Proposition 5.4.
- For all $h \in H_j$ and $g \in \mathbb{Z}^3$ we have $\tilde{\rho}_j(h) m_j(g) \tilde{\rho}_j(h)^{-1} = m_j(hgh^{-1})$.

We should point out here that it is by far not always possible to extend a choice of (E_0, θ) and $\{(E_j, A_j)\}$ to compatible gluing data. This will become clear 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 $\beta \in \mathbb{R}$ we define

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

$$\|f\|_{L^\infty_{\beta,t}(U)} := \|w_t^{-\beta} f\|_{L^\infty(U)},$$

$$\|f\|_{C^{k,\alpha}_{\beta,t}(U)} := \sum_{j=0}^{k} \|\nabla^j f\|_{L^\infty_{\beta-j,t}(U)} + [\nabla^j f]_{C^{0,\alpha}_{\beta-j,t}(U)}.$$

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 4.11, 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 $C^{k,\alpha}_{\beta,t}$ the Banach space $C^{k,\alpha}$ equipped with the norm $\|\cdot\|_{C^{k,\alpha}_{\beta,t}}$

Remark 6.3 For fixed $t \in (0, T)$ and $\beta \in \mathbb{R}$, the norms $\|\cdot\|_{C^{k,\alpha}_{\beta,t}}$ and are $\|\cdot\|_{C^{k,\alpha}}$ equivalent, but not uniformly so as $t > 0$ tends to zero.

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

(6-1) $$\|f \cdot g\|_{C^{k,\alpha}_{\beta,t}} \le \|f\|_{C^{k,\alpha}_{\beta_1,t}} \cdot \|g\|_{C^{k,\alpha}_{\beta_2,t}}.$$

Also for $\beta > \gamma$ we have

(6-2) $$\|f\|_{C^{k,\alpha}_{\beta,t}} \le t^{\gamma-\beta}\|f\|_{C^{k,\alpha}_{\gamma,t}}.$$

Proposition 6.4 *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* $\widetilde{A}_t$ *satisfying*

(6-3) $$\left\|F_{\widetilde{A}_t} \wedge \psi_t\right\|_{C^{0,\alpha}_{-2,t}} \le ct^{1/2}.$$

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

(6-4) $$p_1(\mathfrak{g}_{E_t}) = -\sum_j k_j \,\mathrm{PD}[S_j] \quad \text{with } k_j := \frac{1}{8\pi^2}\int_{X_j} |F_{A_j}|^2$$

and

(6-5) $$\langle w_2(\mathfrak{g}_{E_t}), [\Sigma]\rangle = \langle w_2(\mathfrak{g}_{E_j}), [\Sigma]\rangle$$

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

Proof The choices of $\widetilde{\rho}_j$ and m_j define a lift of the action of $\mathbb{Z}^3 \rtimes H_j$ on $\mathbb{R}^3 \times X_j$ to the pullback of E_j to $\mathbb{R}^3 \times X_j$. 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 $\mathbb{R}^3 \times X_j$ 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 (4-7) we defined $R_{j,t} := \widetilde{T}_{j,t} \cap r_t^{-1}[\zeta/4, \zeta/2]$ with $\widetilde{T}_{j,t}$ and r_t as defined in (4-4) and (4-6), respectively. By the compatibility conditions the monodromy of A_j along S_j on the fibre at infinity matches up with the monodromy of θ 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

(6-6) $$A_j = \theta + a_j \quad \text{with } \nabla^k a_j = t^{2+k} O(r_t^{-3-k}),$$

because of Remark 4.8 and Proposition 5.8. Fix a smooth non-increasing function $\chi\colon [0, \zeta] \to [0, 1]$ such that $\chi(s) = 1$ for $s \le \zeta/4$ and $\chi(s) = 0$ for $s \ge \zeta/2$. Set

$\chi_t := \chi \circ r_t$. After cutting off A_j to $\theta + \chi_t \cdot a_j$ it can be matched with θ and we obtain the connection $\widetilde{A}_t$ on the bundle E_t.

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

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

The last two terms in this expression are supported in $R_{j,t}$ and of order t^2 in $C^{0,\alpha}$ by (6-6). By Example 3.3 and Proposition 4.11 we have

$$\| F_{A_j} \wedge \psi_t \|_{C^{0,\alpha}_{-2,t}(\widetilde{T}_{j,t})} = \| F_{A_j} \wedge \big(\psi_t - \widehat{\psi}_t\big) \|_{C^{0,\alpha}_{-2,t}(\widetilde{T}_{j,t})} \le ct^{1/2} \| F_{A_j} \|_{C^{0,\alpha}_{-2,t}}.$$

It follows from Proposition 5.4 and Remark 4.8 that

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

This implies that

$$\| F_{A_j} \|_{C^{0,\alpha}_{-4,t}(\widetilde{T}_{j,t})} \le ct^2$$

and, hence,

$$\| F_{A_j} \|_{C^{0,\alpha}_{-2,t}(\widetilde{T}_{j,t})} \le c$$

by (6-2) with $c > 0$ independent of $t \in (0, T)$. Now, putting everything together yields (6-3).

Let $\iota_{j,t}\colon T^3 \times \pi_{j,t}^{-1}(B^4_\zeta/G_j) \to Y$ be as in Remark 4.7. Then $\iota_{j,t}^* \mathfrak{g}_{E_t}$ is isomorphic to the pullback of $\mathfrak{g}_{E_j}$ to $T^3 \times \pi_{j,t}^{-1}(B^4_\zeta/G_j)$. This implies (6-5) by naturality of Stiefel–Whitney classes. To compute $p_1(\mathfrak{g}_{E_t})$ we use Chern–Weil theory to represent it as $p_1(\mathfrak{g}_{E_t}) = -\frac{1}{8\pi^2}\operatorname{tr}(F_{\widetilde{A}_t} \wedge F_{\widetilde{A}_t})$. We can write this as $p_1(\mathfrak{g}_{E_t}) = \sum_j p_j$, where p_j are compactly supported 4–forms on $\widetilde{T}_{j,t}$. Recalling the definition of $[S_j]$ in (4-5) and considering the behaviour of Poincaré duality with respect to coverings we see that in order to prove (6-4) we have to show

$$\iota_{j,t}^* p_j = k_j \,\mathrm{PD}\big[T^3 \times \{x\}\big] \in H^4_c\big(T^3 \times \pi_{j,t}^{-1}\big(B^4_\zeta/G_j\big), \mathbb{R}\big).$$

From our construction of $\widetilde{A}_t$ 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 $-\frac{1}{8\pi^2}\operatorname{tr}(F_{\widetilde{A}_j} \wedge F_{\widetilde{A}_j})$ where $\widetilde{A}_j = A_j + \alpha$ and, by slight abuse of notation, $\alpha = (1-\chi_t)a_j$. Consequently, $\iota_{j,t}^* p_j$ is a multiple of $\mathrm{PD}[T^3 \times \{x\}]$. To see that the multiplicity is precisely k_j we use the Chern–Simons 3–form (see Donaldson–Kronheimer [6, Equation (2.1.17)]) to write

$$\operatorname{tr}\big(F_{\widetilde{A}_j} \wedge F_{\widetilde{A}_j}\big) - \operatorname{tr}\big(F_{A_j} \wedge F_{A_j}\big) = \mathrm{d}\operatorname{tr}\big(\alpha \wedge \mathrm{d}_{A_j}\alpha + \tfrac{1}{3}\alpha \wedge [\alpha \wedge \alpha]\big).$$

By Proposition 5.8 the 1–form α decays sufficiently fast to conclude from Stokes' theorem that

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

This completes the proof. □

Remark 6.5 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 $\mathbb{R}^3 \times$ ALE

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

$$*_t\left(F_{\tilde{A}_t + a_t} \wedge \psi_t\right) + \mathrm{d}_{\tilde{A}_t}\xi_t = 0 \tag{7-1}$$

for $t \in (0, T')$ provided $T' \in (0, T]$ is sufficiently small. Here $*_t$ denotes the Hodge $*$–operator associated with ϕ_t. Equation (7-1) together with the Coulomb gauge condition $\mathrm{d}^*_{\tilde{A}_t} a_t = 0$ can be written as

$$L_t \underline{a}_t + Q_t(\underline{a}_t) + *_t\left(F_{\tilde{A}_t} \wedge \psi_t\right) = 0. \tag{7-2}$$

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 (3-5) with $\psi = \psi_t := *_t\phi_t$ and Q_t is defined by

$$Q_t(\underline{a}) := \tfrac{1}{2} *_t ([a \wedge a] \wedge \psi_t) + [a, \xi]. \tag{7-3}$$

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 $\mathbb{R}^3$ and denote by $(\omega_1, \omega_2, \omega_3)$ the triple of Kähler forms associated with X. Consider $\mathbb{R}^3 \times X$ as a G_2–manifold as in Example 2.6. Denote by $p_{\mathbb{R}^3}\colon \mathbb{R}^3 \times X \to \mathbb{R}^3$ and $p_X\colon \mathbb{R}^3 \times X \to X$ the projection onto the first and second factor, respectively. Slightly abusing notation, we denote the respective pullbacks of E and A to $\mathbb{R}^3 \times X$ via p_X by E and A as well. As in (3-5) we define $L_A\colon \Omega^0(\mathbb{R}^3 \times X, \mathfrak{g}_E) \oplus \Omega^1(\mathbb{R}^3 \times X, \mathfrak{g}_E) \to \Omega^0(\mathbb{R}^3 \times X, \mathfrak{g}_E) \oplus \Omega^1(\mathbb{R}^3 \times X, \mathfrak{g}_E)$ by

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

with ψ as in (2-4).

Proposition 7.1 *If we identify* $p_{\mathbb{R}^3}^* T^*\mathbb{R}^3$ *with* $p_X^*\Lambda^+T^*X$ *via* $\delta^1 \mapsto \omega_1$, $\delta^2 \mapsto \omega_2$, $\delta^3 \mapsto -\omega_3$ *and accordingly*

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

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

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

Here $\delta_A\colon \Omega^1(X,\mathfrak{g}_E)\to\Omega^0(X,\mathfrak{g}_E)\oplus\Omega^+(X,\mathfrak{g}_E)$ *denotes the linear operator defined in* (5-2). *Moreover,*

$$L_A^*L_A = \Delta_{\mathbb{R}^3} + \begin{pmatrix}\delta_A\delta_A^* & \\ & \delta_A^*\delta_A\end{pmatrix} \tag{7-4}$$

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

Proof It is a straight-forward computation to verify that $L_A = F + D_A$. It is also easy to see that $F^*F = \Delta_{\mathbb{R}^3}$ and that $F^*D_A + D_A^*F = 0$. This immediately implies (7-4). □

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

$$[f]_{C^{0,\alpha}_\beta(U)} := \sup_{d(x,y)\le w(x,y)} w(x,y)^{\alpha-\beta}\frac{|f(x)-f(y)|}{d(x,y)^\alpha},$$

$$\|f\|_{L^\infty_\beta(U)} := \big\|w^{-\beta}f\big\|_{L^\infty(U)},$$

$$\|f\|_{C^{k,\alpha}_\beta(U)} := \sum_{j=0}^k \big\|\nabla^j f\big\|_{L^\infty_{\beta-j}(U)} + \big[\nabla^j f\big]_{C^{0,\alpha}_{\beta-j}(U)}.$$

Here f is a section of a vector bundle over $U\subset\mathbb{R}^3\times X$ 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 $C^{k,\alpha}_\beta$ the subspace of elements f of the Banach space $C^{k,\alpha}$ with $\|f\|_{C^{k,\alpha}_\beta} < \infty$ and equip it with the norm $\|\cdot\|_{C^{k,\alpha}_\beta}$.

Under the assumptions of Section 6 and with $\mathbf{g}$ denoting compatible gluing data suppose that $X = X_j$ and that $A = A_j$. Define $\tilde\iota_{j,t}\colon \mathbb{R}^3 \times \pi_{j,t}^{-1}(B^4_\zeta/G_j) \to \tilde T_{j,t}$ by

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

For a parameter $\beta \in \mathbb{R}$ 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

$$s_{\beta,t}(\xi,a)(x,y) := t^{\beta-1}\big(t(\tilde\iota_{j,t})^*\xi, (\tilde\iota_{j,t})^* a\big). \tag{7-5}$$

Proposition 7.2 *There is a constant $c > 0$ such that for $t \in (0,T)$*

$$\frac{1}{c}\|\underline{a}\|_{C^{k,\alpha}_{\beta,t}(\tilde T_{j,t})} \le \|s_{\beta,t}\underline{a}\|_{C^{k,\alpha}_\beta(\mathbb{R}^3\times\pi_{j,t}^{-1}(B^4_\zeta/G_j))} \le c\|\underline{a}\|_{C^{k,\alpha}_{\beta,t}(\tilde T_{j,t})},$$

$$\|L_t\underline{a} - s_{\beta-1,t}^{-1} L_{A_j} s_{\beta,t}\underline{a}\|_{C^{0,\alpha}_{\beta-1,t}(\tilde T_{j,t})} \le ct^{1/2}\|\underline{a}\|_{C^{1,\alpha}_{\beta,t}(\tilde T_{j,t})}.$$

Proof The map $\tilde\iota_{j,t}$ pulls back the metric on $\tilde T_{j,t}$ associated with $\hat\phi_t$, that is $g_{\hat\phi_t} = g_{\mathbb{R}^3} \oplus t^2 g_{X_j}$, to $t^2(g_{\mathbb{R}^3} \oplus g_{X_j})$. This implies the first estimate in view of Remark 4.8. The second estimate is immediate from the construction of $\tilde A_t$ and Proposition 4.11. □

Proposition 7.3 *Let $\beta \in (-3,0)$. Then $\underline{a} \in C^{1,\alpha}_\beta$ is in the kernel of $L_A\colon C^{1,\alpha}_\beta \to C^{0,\alpha}_{\beta-1}$ if and only if it is given by the pullback of an element of the L^2 kernel of δ_A to $\mathbb{R}^3 \times X$.*

The proof of Proposition 7.3 relies on the following lemma which we will prove in the Appendix.

Definition 7.4 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 $r \mapsto \mathrm{vol}(B_r(x))$ grows subexponentially, that is, $\mathrm{vol}(B_r(x)) = o(\exp(cr))$ as $r \to \infty$ for every $c > 0$.

Lemma 7.5 *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 $D\colon C^\infty(X,E) \to C^\infty(X,E)$ 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* $a \in C^{\infty}(\mathbb{R}^n \times X, E)$ *satisfies*

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

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

Proof of Proposition 7.3 Suppose $\underline{a} \in C^{1,\alpha}_{\beta}$ 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 5.4 both $\mathbb{R}^3 \times X$ and $\mathfrak{g}_E$ have bounded geometry. Moreover, by Proposition 7.1, $L_A^* L_A = \Delta_{\mathbb{R}^3} + D_A^* D_A$ 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 5.4. Therefore, we can apply Lemma 7.5 to conclude that $\underline{a}$ is invariant under translations in the $\mathbb{R}^3$–direction and, hence, by Propositions 5.8 and 7.1 must be the pullback of an element in the L^2 kernel of δ_A. □

Proposition 7.6 *For* $\beta \in \mathbb{R}$ *there is a constant* $c > 0$ *such that*

$$\|\underline{a}\|_{C^{1,\alpha}_{\beta}} \leq c\Big(\|L_A\underline{a}\|_{C^{0,\alpha}_{\beta-1}} + \|\underline{a}\|_{L^\infty_\beta}\Big).$$

Proof This is a standard result; see Remark 4.14.

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

$$\|\underline{a}\|_{C^{1,\alpha}_{\beta}(U_i)} \leq c\Big(\|L_A\underline{a}\|_{C^{0,\alpha}_{\beta-1}} + \|\underline{a}\|_{L^\infty_\beta}\Big)$$

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

Fix $R > 0$ suitably large and set $U_0 := \{(x,y) \in \mathbb{R}^3 \times X : |\pi(x)| \leq R\}$. Then there clearly is a constant $c > 0$ such that the above estimate holds for $U_i = U_0$. Pick a sequence $(x_i, y_i) \in \mathbb{R}^3 \times X$ 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

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

where $V_i = B_{r_i/4}(x_i, y_i)$ and $\underline{a} = (\xi, a)$. By arguing as in Propositions 4.11 and 5.8 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. □

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 6.4 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 θ 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*

$$*_t\big(F_{\tilde{A}_t + a_t} \wedge \psi_t\big) + \mathrm{d}_{\tilde{A}_t}\xi_t = 0 \tag{8-1}$$

and $\|\underline{a}_t\|_{C^{1,\alpha}_{-1,t}} \le ct^{1/2}$. 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 8.2 *Given $\beta \in (-3, 0)$ there are constants $T' \in (0, T]$ and $c > 0$ such that for $t \in (0, T')$ we have*

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

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 8.3 (Donaldson–Kronheimer [6, Lemma 7.2.23]) *Let X be a Banach space and let $T\colon X \to X$ be a smooth map with $T(0) = 0$. Suppose there is a constant $c > 0$ such that*

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

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

$$x + Tx = y.$$

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

Proof of Proposition 8.1 assuming Proposition 8.2 By Proposition 8.2 the operator $L_t\colon C^{1,\alpha}_{-1,t} \to C^{0,\alpha}_{-2,t}$ 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 $R_t\colon C^{0,\alpha}_{-2,t} \to C^{1,\alpha}_{-1,t}$.

If we set $\underline{a}_t := R_t \underline{b}_t$, then (8-1) becomes

$$\underline{b}_t + Q_t(R_t \underline{b}_t) = - *_t \big(F_{\tilde{A}_t} \wedge \psi_t\big). \tag{8-2}$$

It follows from Proposition 8.2 and (6-1) that

$$\|Q_t(R_t\underline{b}_1) - Q_t(R_t\underline{b}_2)\|_{C^{0,\alpha}_{-2,t}} \le c\Big(\|\underline{b}_1\|_{C^{0,\alpha}_{-2,t}} + \|\underline{b}_2\|_{C^{0,\alpha}_{-2,t}}\Big)\|\underline{b}_1 - \underline{b}_2\|_{C^{0,\alpha}_{-2,t}}$$

with a constant $c > 0$ independent of $t \in (0, T)$. Since by Proposition 6.4

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

Lemma 8.3 provides us with, for each $t \in (0, T')$, a solution $\underline{b}_t$ of (8-2) satisfying $\|\underline{b}_t\|_{C^{0,\alpha}_{-2,t}} \le ct^{1/2}$ provided $T' \in (0, T]$ was chosen sufficiently small. Then

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

is the desired solution of (8-1) and satisfies $\|\underline{a}_t\|_{C^{1,\alpha}_{-1,t}} \le ct^{1/2}$.

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 $\|R_t L_{A_t} - \mathrm{id}\|_{C^{1,\alpha}_{-1,t}} \le ct^{1/2}$ and thus L_{A_t} is invertible for $t \in (0, T')$ provided $T' \in (0, T]$ was chosen sufficiently small. □

Before embarking on the proof of Proposition 8.2, it will be helpful to make a few observations. On $Y_t \setminus \tilde{T}_t$ the operators L_t and L_θ agree. For fixed $\epsilon > 0$, the norms $\|\cdot\|_{C^{k,\alpha}_{\beta,t}(r_t^{-1}[\epsilon,\infty))}$ are uniformly equivalent to the corresponding unweighted Hölder norms. Moreover, the restriction of L_t to $r_t^{-1}[\epsilon, \infty)$ becomes arbitrarily close to L_θ restricted to $\{x \in Y_0 : d(x, S) > \epsilon\}$ as t goes to zero. These observations and standard Schauder estimates combined with Propositions 7.2 and 7.6 yield the following Schauder estimate.

Proposition 8.4 *Given $\beta \in \mathbb{R}$ there is a constant $c > 0$ such that for all $t \in (0, T)$ we have*

$$\|\underline{a}\|_{C^{1,\alpha}_{\beta,t}} \le c\big(\|L_t \underline{a}\|_{C^{0,\alpha}_{\beta-1,t}} + \|\underline{a}\|_{L^\infty_{\beta,t}}\big).$$

This reduces the proof of Proposition 8.2 to the following statement.

Proposition 8.5 *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:*

$$\|\underline{a}\|_{L^\infty_{\beta,t}} \le c\|L_t\underline{a}\|_{C^{0,\alpha}_{\beta-1,t}}.$$

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

$$\|\underline{a}_i\|_{L^\infty_{\beta,t_i}} = 1 \quad \text{and} \quad \|L_{t_i}\underline{a}_i\|_{C^{0,\alpha}_{\beta-1,t_i}} \le \frac{1}{i}.$$

Hence, by Proposition 8.4, we have

$$\|\underline{a}_i\|_{C^{1,\alpha}_{\beta,t_i}} \le 2c. \tag{8-3}$$

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 $\widetilde{A}_t$ converges to θ on K. By (8-3) 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 $C^{1,\alpha/2}_{\mathrm{loc}}$ 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

$$|\underline{a}| < c \cdot d(\cdot, S)^\beta \tag{8-4}$$

as well as

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

Since $\beta > -3$, it follows from (8-4) 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 θ 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 $\widetilde{T}_{j,t_i}$ for some fixed j. With s_{β,t_i} as in (7-5) we define $\underline{\widetilde{a}}_i := s_{\beta,t_i}\underline{a}_i$ and denote by $\widetilde{x}_i$ a lift of x_i to $\mathbb{R}^3 \times \pi_{j,t}^{-1}(B^4_\zeta/G_j)$. This rescaled sequence satisfies, in the notation of Section 7,

$$\|\underline{\widetilde{a}}_i\|_{C^{1,\alpha}_\beta} \le 4c \quad \text{and} \quad (1+|\pi_j(\widetilde{x}_i)|)^{-\beta}|\underline{\widetilde{a}}(\widetilde{x}_i)| \ge \tfrac{1}{2}$$

as well as

$$\|L_{A_j}\underline{\widetilde{a}}_i\|_{C^{0,\alpha}_{\beta-1}} \le 2/i. \tag{8-5}$$

Arguing as in the previous case, we can extract a subsequence of $(\underline{\widetilde{a}}_i)$ which converges to a limit $\underline{\widetilde{a}} \in C^{1,\alpha/2}_\beta$ in $C^{1,\alpha/2}_{\mathrm{loc}}$ on $\mathbb{R}^3 \times X_j$. It follows from (8-5) that $\underline{\widetilde{a}}$ satisfies

$$L_{A_j}\underline{\widetilde{a}} = 0.$$

By Proposition 7.3, $\underline{\widetilde{a}}$ must be zero since $\beta \in (-3,0)$ and A_j is infinitesimally rigid. However, by translation we can arrange that the $\mathbb{R}^3$–component of $\widetilde{x}_i$ is zero and thus we can view $\widetilde{x}_i$ as a point in X_j. Then the condition $\lim d_{t_i}(x_i)/t_i < \infty$ translates to $\lim|\pi_j(\widetilde{x}_i)| < \infty$. Therefore, we can assume without loss of generality that $\widetilde{x}_i$ converges to some point $\widetilde{x} \in X_j$. But then $|\underline{\widetilde{a}}(\widetilde{x})| \ge \frac{1}{2}(1+|\pi_j(\widetilde{x})|)^\beta > 0$, which contradicts $\underline{\widetilde{a}} = 0$.

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

$$\|\underline{\widetilde{a}}_i\|_{C^{1,\alpha}_\beta} \le 8c \quad \text{and} \quad |\widetilde{x}_j|^{-\beta}|\underline{\widetilde{a}}_i(\widetilde{x}_i)| \ge \tfrac{1}{4}$$

as well as

$$\|L\underline{\widetilde{a}}_i\|_{C^{0,\alpha}_{\beta-1}} \le 4/i.$$

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

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

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 $\omega_i \in \Omega^2(\mathbb{C}^2)$ as in Section 4.

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

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

satisfying

$$|\underline{\tilde{a}}| < cw^{\beta} \tag{8-6}$$

as well as

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

Since $\beta > -3$, it follows from (8-6) that $\underline{\tilde{a}}$ satisfies $L\underline{\tilde{a}} = 0$ in the sense of distributions on all of $\mathbb{R}^3 \times \mathbb{C}^2/G_j$ and therefore $\underline{\tilde{a}}$ is smooth by elliptic regularity. It also follows from (8-6) that both $\underline{\tilde{a}}$ and $\nabla\underline{\tilde{a}}$ are uniformly bounded: This is clear outside a tubular neighbourhood of $\mathbb{R}^3 \times \{0\}$. If B_1 is a ball of radius one centred at some point in $\mathbb{R}^3 \times \{0\}$, then (8-6) gives a uniform bound on $\|\underline{\tilde{a}}\|_{L^p(B_1)}$, 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 $\underline{\tilde{a}}$ and $\nabla\underline{\tilde{a}}$. Because $L^*L = \Delta_{\mathbb{R}^3} + \Delta_{\mathbb{C}^2}$, if follows from Lemma 7.5 that $\underline{\tilde{a}}$ is invariant under translations in the $\mathbb{R}^3$–direction. Thus we can think of the components of $\underline{\tilde{a}}$ as harmonic functions on $\mathbb{C}^2$. 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 $\tilde{x} \in \mathbb{C}^2/G_j$ with $|\tilde{x}| = 1$ at which $|\underline{\tilde{a}}|(\tilde{x}) \geq \frac{1}{4}$, contradicting $\underline{\tilde{a}} = 0$. □

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

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 [18, Sections 12.3 and 12.4]. The flat G_2–structure ϕ_0 on T^7 given by (2-1) is preserved by $\alpha, \beta, \gamma \in \mathrm{Diff}(T^7)$ defined by

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

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

To understand the singular set S of T^7/Γ note that the only elements of Γ having fixed points are α, β and γ. 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 α and

$\langle \alpha, \gamma \rangle$ acts freely on the set of T^3 fixed by β, while $\alpha\beta \in \langle \alpha, \beta \rangle$ acts trivially on the set of T^3 fixed by γ. It follows that S consists of 8 copies of T^3 coming from the fixed points of α and β and 8 copies of $T^3/\mathbb{Z}_2$. Near the copies of T^3 the singular set is modelled on $T^3 \times \mathbb{C}^2/\mathbb{Z}_2$ while near the copies of $T^3/\mathbb{Z}_2$ it is modelled on $(T^3 \times \mathbb{C}^2/\mathbb{Z}_2)/\mathbb{Z}_2$ where the action of $\mathbb{Z}_2$ on $T^3 \times \mathbb{C}^2/\mathbb{Z}_2$ is given by

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

The 8 copies of T^3 can be desingularised by any choice of 8 ALE spaces asymptotic to $\mathbb{C}^2/\mathbb{Z}_2$. To desingularise the copies of $T^3/\mathbb{Z}_2$ we need to chose ALE spaces which admit an isometric action of $\mathbb{Z}_2$ asymptotic to the action $\mathbb{Z}_2$ on $\mathbb{C}^2/\mathbb{Z}_2$ given by $\pm(z_1, z_2) \mapsto \pm(z_1, -z_2)$. Two possible choices are the resolution of $\mathbb{C}^2/\mathbb{Z}_2$ or a smoothing of $\mathbb{C}^2/\mathbb{Z}_2$. See Joyce [18, pages 313–314] for details.

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

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

The elements σ_j commute with all elements of Γ and thus act on T^7/Γ. 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/Γ 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 $T^3/\mathbb{Z}_2$, denoted by $S_2, \dots, S_7$. S_1 has a neighbourhood modelled on $T^3 \times \mathbb{C}^2/\mathbb{Z}_2$, while $S_2, \dots, S_6$ have neighbourhoods modelled on $(T^3 \times \mathbb{C}^2/\mathbb{Z}_2)/\mathbb{Z}_2$ where $\mathbb{Z}_2$ acts by $\pm(z_1, z_2) \mapsto \pm(z_1, -z_2)$ on $\mathbb{C}^2/\mathbb{Z}_2$. As before, S_1 can be desingularised by any choice of an ALE space asymptotic to $\mathbb{C}^2/\mathbb{Z}_2$. $S_2, \dots, S_6$ can be desingularised by the resolution of $\mathbb{C}^2/\mathbb{Z}_2$ or a smoothing of $\mathbb{C}^2/\mathbb{Z}_2$.

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 $p\colon \mathbb{R}^7 \to Y_0$ the canonical projection. Then $p\colon p^{-1}(Y_0 \setminus S) \to Y_0 \setminus S$ is a universal cover. Up to conjugation we can therefore identify $\pi_1(Y_0)$ with the group of deck transformations

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

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

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

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 τ_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 $V := \langle a, b, c \mid a^2 = b^2 = c^2 = 1, ab = c \rangle \cong \mathbb{Z}_2^2$ the Klein four-group. V can be thought of as a subgroup of $\mathrm{SO}(3)$: $a = \mathrm{diag}(1, -1, -1)$, $b = \mathrm{diag}(-1, 1, -1)$ and $c = \mathrm{diag}(-1, -1, 1)$. We define $\rho\colon \pi_1(Y_0) \to V \subset \mathrm{SO}(3)$ 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 θ induced by ρ is acyclic we use the following observation.

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

Proof Since Y_0 is flat as a Riemannian orbifold and θ is a flat connection

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

Therefore, all elements in the kernel of L_θ 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 $\mathfrak{g} \oplus (\mathbb{R}^7 \otimes \mathfrak{g})$. □

The elements σ_2 and σ_3 act trivially on $\mathbb{R}^7$ and their action on $\mathfrak{so}(3)$ has no common non-zero fixed vectors. Therefore the action of $\pi_1(Y_0)$ on $\mathfrak{g} \oplus (\mathbb{R}^7 \otimes \mathfrak{g})$ has no non-zero fixed vector and thus θ is acyclic.

The monodromy representation $\mu_j|_{G_j}\colon G_j = \mathbb{Z}_2 \to \mathrm{SO}(3)$ associated with the flat connection θ 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 5.15. For $j = 2, \dots, 6$ we choose A_j to be the product connection on the trivial SO(3)–bundle E_j. We take m_1 and $\widetilde{\rho}_1$ to be trivial. For $j = 2, \dots, 6$ we can choose m_j and $\widetilde{\rho}_j$ 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/Γ. 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 θ to T^7/Γ is not acyclic.

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

$$\gamma, \tau_1, \dots, \tau_7 \mapsto 1, \qquad \alpha \mapsto a, \qquad \beta \mapsto b,$$
$$\sigma_2 \mapsto b, \qquad \sigma_3 \mapsto a.$$

Again, the resulting flat connection θ 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 $\widetilde{\rho}_j$ of the action of $\mathbb{Z}_2$ 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 $\mathbb{C}^2/\mathbb{Z}_2$, then the $\mathbb{Z}_2$ 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 $\mathbb{C}^2/\mathbb{Z}_2$. The reason for this is that in the first case the action of $\mathbb{Z}_2$ on $H^2(X, \mathbb{R})$ is given by the identity, while in the second case it acts via multiplication by -1; see Joyce [18, pages 313–314]. Thus we can only find compatible gluing data if we resolve both S_2 and S_3 using a smoothing of $\mathbb{C}^2/\mathbb{Z}_2$.

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

$$\gamma, \tau_1, \dots, \tau_7 \mapsto 1, \qquad \alpha \mapsto a, \qquad \beta \mapsto b,$$
$$\sigma_2 \mapsto b, \qquad \sigma_3 \mapsto c.$$

To find compatible gluing data, one simply has to compose $\widetilde{\rho}_j$ 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 $T^3/\mathbb{Z}_2$, denoted by $S_1, \dots, S_4$, each of which has a neighbourhood modelled on $(T^3 \times \mathbb{C}^2/\mathbb{Z}_2)/\mathbb{Z}_2$ where $\mathbb{Z}_2$ acts on $\mathbb{C}^2/\mathbb{Z}_2$ 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

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

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

$$\alpha, \beta, \sigma_3, \tau_1, \dots, \tau_7 \mapsto 1, \qquad \gamma \mapsto b$$
$$\sigma_1 \mapsto a, \qquad \sigma_2 \mapsto b.$$

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

Again, the resulting G_2–instanton can be lifted to appropriate σ_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 θ to $T^7/(\Gamma \times A)$ it 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.

Appendix: 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 [29, Corollary 7.5] and in Brendle's unpublished work on the Yang–Mills equation in higher dimension [3, 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 $D\colon C^\infty(X, E) \to C^\infty(X, E)$ 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 $a \in C^\infty(\mathbb{R}^n \times X, E)$ satisfies*

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

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

Here is a heuristic argument. Denote by $\hat{a}$ the partial Fourier transform of a in the $\mathbb{R}^n$–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 δ–functions. Reversing the Fourier transform shows that a must be a polynomial in $\mathbb{R}^n$. But then it follows from the assumptions that a is constant in the $\mathbb{R}^n$–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 $\rho\colon X\to[0,\infty)$ a smoothing of the distance from p, as in Kordyukov [19, Proposition 4.1]. For $\delta\in\mathbb{R}$ we introduce a weight function $w_\delta:=e^{-\delta\rho}$ and weighted Hilbert spaces $W^{s,2}_\delta(X,E)$ consisting of locally integrable sections f such that $w_\delta\cdot f$ lies in $W^{s,2}(X,E)$ with inner product defined by $\langle\cdot,\cdot\rangle_{W^{s,2}_\delta}:=\langle w_\delta\cdot,w_\delta\cdot\rangle_{W^{s,2}}$. As usual we set $L^2_\delta(X,E):=W^{0,2}_\delta(X,E)$.

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

$$\left\|\partial_k^\ell\big(D+k^2\big)^{-1}a\right\|_{W^{2,2}_\delta}\leq c_\ell(1+k)^\ell\|a\|_{L^2_\delta} \tag{A-1}$$

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

Proof By standard elliptic theory we have

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

Since D is non-negative, we have

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

Putting everything together yields

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

for $k\in[k_0,\infty)$. This implies that $D+k^2\colon W^{2,2}\to L^2$ 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 [19, Proposition 4.4]. Via the Hilbert space isomorphism $W^{s,2}_\delta \cong W^{s,2}$ defined by multiplication with w_δ the operator $D+k^2\colon W^{2,2}_\delta \to L^2_\delta$ is equivalent to $D_\delta + k^2\colon W^{2,2} \to L^2$ where $D_\delta := w_\delta D w_\delta^{-1}$. We can write D_δ as

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

with $P_\delta\colon W^{2,2} \to L^2$ bounded independent of δ. Therefore,

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

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

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

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

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

for universal constants $c_{i,j,\ell}$. Thus (A-1) for $\ell > 0$ can be reduced to the case $\ell = 0$. □

Lemma A.1 can now be proved using an argument similar to the one used by Brendle in [3, 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 $\chi \in \mathscr{S}(\mathbb{R}^n)$ 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 $\delta \in (-\epsilon, \epsilon)$ with $\epsilon = \epsilon(k_0)$. Then there exists $a \in \mathscr{S}(\mathbb{R}^n, W^{2,2}_\delta(X, E))$ such that $(\Delta_{\mathbb{R}^n} + D)a = \chi b$.*

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

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

and define

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

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

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

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

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

Pick a smooth function $\rho\colon \mathbb{R} \to [0,1]$ such that $\rho(k) = 0$ for $|k| \le 1$ and $\rho(k) = 1$ for $|k| \ge 2$. Set $\hat{\chi}_\epsilon(k) := \rho(|k|/\epsilon)\hat{\chi}(k)$ and denote its inverse Fourier transform by χ_ϵ. Then χ_ϵ 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

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

and therefore

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

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 $(\Delta_{\mathbb{R}^n} + D)a = 0$. Then for $\sigma \in \mathcal{S}^n(\mathbb{R}^n)$, $\delta \in \mathbb{R}^n$ and $b \in C^\infty_c(X,E)$ we have*

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

In particular, the conclusion of the lemma holds.

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

$$\begin{aligned}\int_{\mathbb{R}^n} \langle a(x,\cdot), b\rangle \chi(x)\, \mathrm{d}\mathcal{L}^n(x) &= \lim_{\epsilon\to 0} \int_{\mathbb{R}^n} \langle a(x,\cdot), b\rangle \chi_\epsilon(x)\, \mathrm{d}\mathcal{L}^n(x) \\ &= \lim_{\epsilon\to 0} \int_{\mathbb{R}^n} \int_X \langle a(x,y), (\Delta_{\mathbb{R}^n} + D)c_\epsilon\rangle\, \mathrm{d}\mathcal{L}^n(x)\, \mathrm{dvol}(y) \\ &= \lim_{\epsilon\to 0} \int_{\mathbb{R}^n} \int_X \langle (\Delta_{\mathbb{R}^n} + D)a(x,y), c_\epsilon\rangle\, \mathrm{d}\mathcal{L}^n(x)\, \mathrm{dvol}(y) \\ &= 0.\end{aligned}$$

Since σ, δ and b are arbitrary, it follows that a is invariant in the $\mathbb{R}^n$–direction. This finishes the proof. □

Remark A.3 It is clear from the proof that in Lemma 7.5 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 $x \in \mathbb{R}^n$ for all $\delta > 0$.

References

[1] **A Adem**, **J Leida**, **Y Ruan**, *Orbifolds and stringy topology*, Cambridge Tracts in Mathematics 171, Cambridge Univ. Press (2007) MR2359514

[2] **M Berger**, *Sur les groupes d'holonomie homogène des variétés à connexion affine et des variétés riemanniennes*, Bull. Soc. Math. France 83 (1955) 279–330 MR0079806

[3] **S Brendle**, *On the construction of solutions to the Yang–Mills equations in higher dimensions* arXiv:math/0302093

[4] **D M J Calderbank**, **P Gauduchon**, **M Herzlich**, *Refined Kato inequalities and conformal weights in Riemannian geometry*, J. Funct. Anal. 173 (2000) 214–255 MR1760284

[5] **E Corrigan**, **C Devchand**, **D B Fairlie**, **J Nuyts**, *First-order equations for gauge fields in spaces of dimension greater than four*, Nuclear Phys. B 214 (1983) 452–464 MR698892

[6] **S K Donaldson**, **P B Kronheimer**, *The geometry of four-manifolds*, Oxford Mathematical Monographs, Oxford Univ. Press (1990) MR1079726

[7] **S K Donaldson**, **E Segal**, *Gauge theory in higher dimensions II*, from: "Geometry of special holonomy and related topics", (N C Leung, S-T Yau, editors), Surv. Differ. Geom. 16, International Press (2011) 1–41 MR2893675

[8] **S K Donaldson**, **R P Thomas**, *Gauge theory in higher dimensions*, from: "The Geometric Universe", (S A Huggett, L J Mason, K P Tod, S Tsou, N M J Woodhouse, editors), Oxford Univ. Press (1998) 31–47 MR1634503

[9] **P M N Feehan**, *A Kato–Yau inequality and decay estimate for eigenspinors*, J. Geom. Anal. 11 (2001) 469–489 MR1857853

[10] **M Fernández**, **A Gray**, *Riemannian manifolds with structure group G_2*, Ann. Mat. Pura Appl. 132 (1982) 19–45 MR696037

[11] **D S Freed**, **K K Uhlenbeck**, *Instantons and four-manifolds*, 2nd edition, Mathematical Sciences Research Institute Publications 1, Springer, New York (1991) MR1081321

[12] **T Gocho**, **H Nakajima**, *Einstein–Hermitian connections on hyper–Kähler quotients*, J. Math. Soc. Japan 44 (1992) 43–51 MR1139657

[13] **R Harvey**, **H B Lawson, Jr**, *Calibrated geometries*, Acta Math. 148 (1982) 47–157 MR666108

[14] **A Haydys**, *Gauge theory, calibrated geometry and harmonic spinors*, J. Lond. Math. Soc. 86 (2012) 482–498 MR2980921

[15] **N J Hitchin**, **A Karlhede**, **U Lindström**, **M Roček**, *Hyper–Kähler metrics and supersymmetry*, Comm. Math. Phys. 108 (1987) 535–589 MR877637

[16] **D D Joyce**, *Compact Riemannian 7–manifolds with holonomy G_2 I*, J. Differential Geom. 43 (1996) 291–328 MR1424428

[17] **D D Joyce**, *Compact Riemannian 7–manifolds with holonomy G_2 II*, J. Differential Geom. 43 (1996) 329–375 MR1424428

[18] **D D Joyce**, *Compact manifolds with special holonomy*, Oxford Mathematical Monographs, Oxford Univ. Press (2000) MR1787733

[19] **Y A Kordyukov**, *L^p–theory of elliptic differential operators on manifolds of bounded geometry*, Acta Appl. Math. 23 (1991) 223–260 MR1120831

[20] **A Kovalev**, *Twisted connected sums and special Riemannian holonomy*, J. Reine Angew. Math. 565 (2003) 125–160 MR2024648

[21] **P B Kronheimer**, *The construction of ALE spaces as hyper–Kähler quotients*, J. Differential Geom. 29 (1989) 665–683 MR992334

[22] **P B Kronheimer**, *A Torelli-type theorem for gravitational instantons*, J. Differential Geom. 29 (1989) 685–697 MR992335

[23] **P B Kronheimer**, *Instanton invariants and flat connections on the Kummer surface*, Duke Math. J. 64 (1991) 229–241 MR1136374

[24] **P B Kronheimer**, **H Nakajima**, *Yang–Mills instantons on ALE gravitational instantons*, Math. Ann. 288 (1990) 263–307 MR1075769

[25] **C Lewis**, Spin(7) *instantons*, DPhil thesis, University of Oxford (1998)

[26] **J McKay**, *Graphs, singularities, and finite groups*, from: "The Santa Cruz Conference on Finite Groups", (B Cooperstein, G Mason, editors), Proc. Sympos. Pure Math. 37, Amer. Math. Soc. (1980) 183–186 MR604577

[27] **H Nakajima**, *Moduli spaces of anti-self-dual connections on ALE gravitational instantons*, Invent. Math. 102 (1990) 267–303 MR1074476

[28] **L Nirenberg**, **H F Walker**, *The null spaces of elliptic partial differential operators in* $\mathbb{R}^n$, J. Math. Anal. Appl. 42 (1973) 271–301 MR0320821

[29] **F Pacard**, **M Ritoré**, *From constant mean curvature hypersurfaces to the gradient theory of phase transitions*, J. Differential Geom. 64 (2003) 359–423 MR2032110

[30] **H N Sá Earp**, G_2*–instantons on Kovalev manifolds* arXiv:1101.0880

[31] **H N Sá Earp**, G_2*–instantons on Kovalev manifolds II* arXiv:1101.0880

[32] **D A Salamon**, **T Walpuski**, *Notes on the octonions* arXiv:1005.2820

[33] **T Tao**, **G Tian**, *A singularity removal theorem for Yang–Mills fields in higher dimensions*, J. Amer. Math. Soc. 17 (2004) 557–593 MR2053951

[34] **G Tian**, *Gauge theory and calibrated geometry I*, Ann. of Math. 151 (2000) 193–268 MR1745014

[35] **K K Uhlenbeck**, *Connections with* L^p *bounds on curvature*, Comm. Math. Phys. 83 (1982) 31–42 MR648356

[36] **K K Uhlenbeck**, *Removable singularities in Yang–Mills fields*, Comm. Math. Phys. 83 (1982) 11–29 MR648355

Department of Mathematics, Imperial College London
180 Queen's Gate, London SW7 2AZ, UK

tw809@ic.ac.uk

Proposed: Ronald Stern Received: 26 January 2012
Seconded: Richard Thomas, Simon Donaldson Revised: 17 December 2012