Four-Dimensional Einstein Manifolds, and Beyond

Claude LeBrun

1. Introduction

The aim of modern Riemannian geometry is to understand the relationship between topology and curvature. A case in point is that one would like to know when a given smooth compact n-manifold $M$ admits an Einstein metric — that is, a Riemannian metric $g$ such that

$$r = \lambda g,$$

where $r$ is the Ricci tensor of $g$ and $\lambda$ is some real constant. When such a metric exists, moreover, it is natural to ask to what extent it is unique; in other words, one would like to understand the Einstein moduli space of $M$ — i.e. the set of unit-volume Einstein metrics on $M$, modulo the action of the diffeomorphism group.

These existence and uniqueness questions are easily answered in dimensions 2 and 3, because a Riemannian manifold of dimension $n < 4$ is Einstein iff it has constant sectional curvature. In low dimensions, the sign of $\lambda$ is therefore completely determined by the topology of $M$ — indeed, by the size of $\pi_1(M)$. Moreover, the moduli space of Einstein metrics on a 2- or 3-manifold is always connected [43], so the value of $\lambda$, for unit-volume Einstein metrics $g$, is actually an invariant of $M$. In dimension 2, the moduli space is never empty, and has positive dimension if $\lambda \leq 0$. By contrast, the moduli space of a 3-manifold is [43] a single point if $\lambda < 0$. On the other hand, the Einstein moduli space is empty [8, 61] for any 3-manifolds with $\pi_2 \neq 0$; cf. [3].

In dimension $n \geq 4$, the curvature tensor of $g$ is no longer determined by the Einstein condition in a point-wise manner, and Einstein metrics are no longer describable in terms of universal local models. While this, of course, is precisely what gives the subject its interest, the existence and uniqueness problems are commensurately harder when $n \geq 4$. Indeed, there are, to date, no non-existence or uniqueness results known when $n > 4$. Fortunately, however, a constellation of low-dimensional accidents makes the borderline case of $n = 4$ comparatively tractable. My aim here is to survey the current state of our knowledge regarding the existence and uniqueness of Einstein metrics on 4-manifolds, and point out some hints these give us regarding higher dimensions.

Supported in part by NSF grant DMS-9802722.
2. The Hitchin-Thorpe Inequality

Four-dimensional Riemannian geometry displays many features which have no adequate analogues in other dimensions. These are largely attributable to a single Lie-group-theoretic fluke: the rotation group $SO(4)$ isn't simple. Indeed,

$$so(4) \cong so(3) \oplus so(3),$$

so the adjoint action of $SO(4)$ on its Lie algebra preserves a decomposition into two 3-dimensional subspaces. Now $so(n)$ and $\Lambda^2(\mathbb{R}^n)$ are isomorphic as $SO(n)$-modules. Thus the rank-6 bundle of 2-forms on an oriented Riemannian 4-manifold decomposes invariantly into two rank-3 bundles:

$$\Lambda^2 = \Lambda^+ \oplus \Lambda^-.$$  

Schur's lemma tells us that these bundles must coincide with the eigenspaces of the Hodge duality operator

$$\star : \Lambda^2 \to \Lambda^2.$$  

With appropriately chosen conventions, sections of $\Lambda^+$ are thus characterized by $\star \varphi = \varphi$, and sections of $\Lambda^-$ satisfy $\star \varphi = -\varphi$.

**Definition 2.1.** On any smooth oriented 4-manifold, sections of $\Lambda^+$ are called **self-dual 2-forms**, whereas sections of $\Lambda^-$ are called **anti-self-dual 2-forms**.

Now let us suppose that $(M, g)$ is a **compact** oriented Riemannian 4-manifold. The Hodge theorem then tells us that every de Rham class on $M$ has a unique harmonic representative; in particular, there is a canonical identification

$$H^2(M, \mathbb{R}) = \{ \varphi \in \Gamma(\Lambda^2) \mid d\varphi = 0, \ d\star \varphi = 0 \}.$$  

But the Hodge star operator $\star$ defines an involution of the right-hand side. We therefore have a direct sum decomposition

$$H^2(M, \mathbb{R}) = \mathcal{H}_g^+ \oplus \mathcal{H}_g^-,$$

where

$$\mathcal{H}_g^\pm = \{ \varphi \in \Gamma(\Lambda^\pm) \mid d\varphi = 0 \}$$

are the spaces of self-dual and anti-self-dual harmonic forms. Notice that $\star$ is conformally invariant in the middle dimension, so the decomposition (2) remains unchanged if the metric $g$ is multiplied by a smooth positive function.

The intersection form

$$\sim : H^2(M, \mathbb{R}) \times H^2(M, \mathbb{R}) \to \mathbb{R}$$

$$( [\varphi], [\psi] ) \mapsto \int_M \varphi \wedge \psi$$

becomes positive-definite when restricted to $\mathcal{H}_g^+$, and negative-definite when restricted to $\mathcal{H}_g^-$; and the two are mutually orthogonal with respect to $\sim$. Thus, combining an $L^2$-orthonormal basis for $\mathcal{H}_g^+$ with an $L^2$-orthonormal basis for $\mathcal{H}_g^-$ gives us a basis for $H^2(M, \mathbb{R})$ in which the intersection form is represented by the
diagonal matrix

\[
\begin{pmatrix}
1 \\
\vdots \\
1 \\
\frac{1}{b_+(M)} \\
-1 \\
\frac{-1}{b_-(M)} \\
\end{pmatrix}
\]

The numbers \( b_\pm(M) = \dim \mathcal{H}^\pm \) are therefore oriented homotopy invariants of \( M \); namely, \( b_+ \) (respectively, \( b_- \)) is the dimension of any maximal linear subspace of \( H^2(M, \mathbb{R}) \) on which the restriction of \( \sim \) is positive (respectively, negative) definite.

The intersection form described above is a bilinear form over \( \mathbb{R} \). But of course, the cup product is also defined on integer cohomology, and one should therefore think of the intersection form over \( \mathbb{R} \) as a mere shadow of a more fundamental object

\[\sim : H^2(M, \mathbb{Z}) \times H^2(M, \mathbb{Z}) \to \mathbb{Z},\]

concretely representable as a \( b_2 \times b_2 \) integer matrix of determinant \( \pm 1 \). While such an integer quadratic form can of course be diagonalized over the reals, the analogous assertion fails over the integers. For example, the intersection form

\[
\begin{pmatrix}
0 & 1 \\
1 & 0 \\
\end{pmatrix}
\]

of \( S^2 \times S^2 \) is an \textit{even} form, meaning that \( \alpha \sim \beta \equiv 0 \mod 2 \) for all \( \alpha, \beta \in H^2(M, \mathbb{Z}) \).

By contrast, of course, the diagonal form

\[
\begin{pmatrix}
1 & 0 \\
0 & -1 \\
\end{pmatrix}
\]

is \textit{odd} — which, by definition, just means that it is \textit{not even}!

Manifolds with any specified values of \( b_\pm \) can easily be constructed by the following operation:

\textbf{DEFINITION 2.2.} Let \( M_1 \) and \( M_2 \) be connected compact oriented 4-manifolds. Their connected sum \( M_1 \# M_2 \) is then the oriented 4-manifold obtained by deleting a small ball from each manifold and gluing together the resulting \( S^3 \) boundaries via a reflection.

For example, the \( 2 \times 2 \) diagonal form considered above can be realized as the intersection form of \( \mathbb{CP}_2 \# \mathbb{CP}_2 \), where \( \mathbb{CP}_2 \) is the complex projective plane with its standard orientation, and \( \overline{\mathbb{CP}_2} \) is the same smooth 4-manifold with the \textit{opposite} orientation. Similarly, the iterated connected sum

\[k\mathbb{CP}_2 \# \ell\overline{\mathbb{CP}_2} = \overbrace{\mathbb{CP}_2 \# \cdots \# \mathbb{CP}_2}^k \# \overbrace{\overline{\mathbb{CP}_2} \# \cdots \# \overline{\mathbb{CP}_2}}^\ell\]

has diagonal intersection form, with \( b_+ = k \) and \( b_- = \ell \). Notice that \( n(S^2 \times S^2) \) and \( n\mathbb{CP}_2 \# n\overline{\mathbb{CP}_2} \) are simply connected 4-manifolds with the same invariants \( b_\pm \), but are not homotopy equivalent because one has even intersection form and one has odd intersection form. This distinction can be restated by saying that one is \textit{spin} and the other is \textit{non-spin}. An oriented manifold is called \textit{spin} iff it satisfies
\[ w_2 = 0, \text{ where } w_2 \in H^2(\mathbb{Z}_2) \text{ denotes the second Stiefel-Whitney class of the tangent bundle. In dimension 4, this is equivalent to the statement that every } \alpha \in H^2(\mathbb{Z}_2) \text{ satisfies } \alpha \sim \alpha = 0 \in \mathbb{Z}_2, \text{ as a consequence of the Wu relation} \]

\[ w_2 \sim \alpha = \alpha \sim \alpha \in \mathbb{Z}_2. \]

In particular, a simply connected 4-manifold is spin iff its intersection form on \( H^2(\mathbb{Z}) \) is even.

Once this distinction between spin and non-spin 4-manifolds is understood, the topological classification of smooth simply connected 4-manifolds is easily stated.

**Theorem 2.1 (Freedman).** Two smooth simply connected oriented 4-manifolds are orientedly homeomorphic iff

- they have the same invariants \( b_+ \) and \( b_- \); and
- both are spin, or both are non-spin.

Freedman’s result was originally stated [19] in terms of the equivalence of intersection forms; but Donaldson’s celebrated theorem [17] on the diagonalizability of definite intersection forms and the Minkowski-Hasse classification of indefinite forms [28] allow one to make the simplified statement given here. On the other hand, the reader should immediately be warned that the classification of 4-manifolds up to diffeomorphism, while still poorly understood, is at least known to be much more complicated. In particular, the Seiberg-Witten invariants discussed in §4 allow one to show that some of the homeotypes treated by Theorem 2.1 can be realized by infinitely many distinct diffeotypes.

The difference \( \tau(M) = b^+(M) - b^-(M) \) is called the *signature* of \( M \). It is precisely the index of an elliptic operator

\[ d^- d^* : \Gamma(\Lambda^+) \to \Gamma(\Lambda^-), \]

and the Atiyah-Singer index theorem therefore predicts [4] that it must be calculable by integrating an invariant polynomial in curvature; and indeed, this has been discovered much earlier by Hirzebruch [25], using a less general argument. Of course, the same is also true of the Euler characteristic \( \chi(M) = 2 - 2b_1(M) + b_2(M) \), which is the index of

\[ d + d^* : \Gamma(\Lambda^{\text{even}}) \to \Gamma(\Lambda^{\text{odd}}); \]

in this case, the corresponding Gauss-Bonnet formula was first proved by Allen-Doofer and Weil [1]. In both cases, the integrand is quadratic in curvature, as is forced on one by invariance under rescalings \( g \to cg \), where \( c > 0 \) is any real constant.

Now let \( g \) be an arbitrary Riemannian metric on an oriented 4-manifold \( M \), and, by raising an index, identify its curvature tensor with the *curvature operator* \( \mathcal{R} : \Lambda^2 \to \Lambda^2 \). Decomposing the 2-forms as in (1), this linear endomorphism of \( \Lambda^2 \) can then be decomposed into primitive pieces

\[
\mathcal{R} = \begin{pmatrix}
W^+ + \frac{\epsilon}{12} & \frac{\partial}{\partial t} \\
\frac{\partial}{\partial t} & W^- + \frac{\epsilon}{12}
\end{pmatrix}.
\]
Here $W_\pm$ are the trace-free pieces of the appropriate blocks, and are called the self-dual and anti-self-dual Weyl curvatures, respectively. The scalar curvature $s$ is understood to act by scalar multiplication, whereas the trace-free Ricci curvature $\tilde{\tau} = r - \frac{s}{4} g$ acts on 2-forms by

$$\varphi_{ab} \mapsto \tilde{\varphi}_{bc} \tilde{\varphi}^c - \tau_{bc} \varphi^c.$$ 

Each of these curvatures corresponds to a different irreducible representation of $SO(4)$, and so any invariant quadratic polynomial in curvature must be a linear decomposition of $s^2$, $|\tilde{\tau}|^2$, $|W^+|^2$ and $|W^-|^2$, and the signature and Euler characteristic are thus expressible as a linear combination of their integrals. The coefficients, of course, may then be deduced by inspecting a handful of well-chosen examples. Thus the 4-dimensional Gauss-Bonnet formula may explicitly be written as

$$\chi(M) = \frac{1}{8\pi^2} \int_M \left[ |W^+|^2 + |W^-|^2 + \frac{s^2}{24} - \frac{|\tilde{\tau}|^2}{2} \right] d\mu,$$

whereas the Hirzebruch signature theorem takes the form

$$\tau(M) = \frac{1}{12\pi^2} \int_M [ |W^+|^2 - |W^-|^2 ] d\mu.$$

Here the curvatures, norms $|\cdot|$, and volume form $d\mu$ are, of course, those of our chosen Riemannian metric $g$.

In particular, it follows that

$$(2\chi \pm 3\tau)(M) = \frac{1}{4\pi^2} \int_M \left[ 2|W_\pm|^2 + \frac{s^2}{24} - \frac{|\tilde{\tau}|^2}{2} \right] d\mu. \tag{4}$$

Since the above integrand is non-negative for any Einstein metric, we therefore have the following celebrated result of Thorpe [60] and Hitchin [26]:

**Theorem 2.2 (Hitchin-Thorpe Inequality).** If the smooth compact oriented 4-manifold $M$ admits an Einstein metric $g$, then

$$2\chi(M) \geq 3|\tau(M)|,$$

with equality iff the $g$-induced connection on one of the bundles $\Lambda^\pm$ is flat.

The last statement follows from the observation [53] that the self-dual and anti-self-dual parts of the curvature of $\Lambda^+$ are precisely represented by the two left-hand blocks of (3), whereas the two right-hand blocks represent the self-dual and anti-self-dual parts of the curvature of $\Lambda^-$. An oriented Riemannian 4-manifold is called *locally hyper-Kähler* if $\Lambda^+$ is flat; and $\Lambda^-$ is therefore flat iff the orientation-reverse of the manifold is locally hyper-Kähler. We will discuss the classification of locally hyper-Kähler manifolds in the next section. For now, suffice it observe that the bundle $\Lambda^+$ becomes trivial when pulled back to the universal cover of any locally hyper-Kähler manifold, so that the universal cover must, in particular, be spin.

**Example 2.1.** The simply connected non-spin 4-manifold $k\mathbb{CP}^2 \# (4k - 4)\mathbb{CP}^2$ has $\chi = 2 + k + \ell$ and $\tau = k - \ell$, and so cannot admit an Einstein metric unless $4 + 5k > \ell > (k - 4)/5$.

It is worth pointing out that the invariant $2\chi + 3\tau$ has an intrinsic importance: it is the first Pontrjagin number of the bundle $\Lambda^+$. Indeed, the above description of the curvature of $\Lambda^+$ tells us that our integral formula for $2\chi + 3\tau$ thus coincides with
the usual integral formula for $p_1(\Lambda^+)$. Notice that the Riemannian connection on $\Lambda^+$ is self-dual iff $g$ is Einstein, so the Hitchin-Thorpe inequality is a special case of the celebrated fact that a bundle with self-dual connection must have non-negative instanton number [18].

Note that the Hitchin-Thorpe inequality only involves homotopy invariants of the 4-manifold in question. Thus, for instance, we could have reached precisely the same conclusion in the above example if $M$ were merely homeomorphic to one of the connected sums $k\mathbb{CP}_2 \# l\mathbb{CP}_2$ considered in the above example. On the other hand, the scalar and Weyl terms have effectively been treated as junk terms. My primary aim in this essay will be to describe some interesting new estimates on these terms which allow one to improve on the Hitchin-Thorpe result. At times, however, this will be done at the price of sacrificing the homotopy invariance of the obstruction.

Let me conclude this section by mentioning an amusing elementary interpretation [53] of the 4-dimensional Einstein equations. By (3), one sees that a 4-manifold is Einstein iff the curvature operator $\mathcal{R}$ commutes with the Hodge star operator $\ast$. But this is clearly the same as asking that the sectional curvature assigned to any 2-plane be the same as that assigned to its orthogonal complement:

$$(M^4, g) \text{ Einstein } \iff K(P) = K(P^\perp) \ \forall \ 2\text{-plane } P \subset TM.$$ 

Of course, this can also be proved in a completely elementary manner. Indeed, the definition of the Ricci tensor and the symmetries of the Riemann tensor tell one that

$$(r_{11} + r_{22}) - (r_{33} + r_{44}) = 2(R_{1212} - R_{3434})$$

in any orthonormal frame on a 4-manifold. But the left-hand side obviously vanishes for every orthonormal frame iff the eigenvalues of $r$ are all equal.

### 3. Complex and Almost-Complex Structures

In order to give our discussion some substance, we need to have some examples. The simplest examples of Einstein manifolds are of course the spaces of constant curvature. A much richer and more illuminating family of examples, however, is provided by the Kähler-Einstein manifolds. Let us begin our description of these by first recalling the notion of an almost-complex structure.

An almost-complex structure on a smooth $n$-manifold $M$ is by definition an endomorphism $J : TM \to TM$ of the tangent bundle such that $J^2 = -1$. Such an object may be thought of as scalar multiplication by $\sqrt{-1}$, and so makes $TM$ into a complex vector bundle, denoted by $T^{1,0}$; in particular, such a structure can exist only if $M$ has even dimension $n = 2m$. Sections of the dual $\Lambda^{1,0}$ of $T^{1,0}$ may concretely be identified with those complex-valued 1-forms on $M$ which convert $J$ into multiplication by $i$:

$$\phi \in \Lambda^{1,0} \iff \phi(Jv) = i\phi(v) \ \forall v \in TM.$$ 

The sections of the rank-$m$ complex vector bundle $\Lambda^{1,0} \to M^{2m}$ are therefore called $(1,0)$-forms. More generally, a complex-valued $(p + q)$-form on $M$ is called a $(p, q)$-form (with respect to $J$) if it is a section of

$$\Lambda^{p,q} = \Lambda^p(\Lambda^{1,0}) \otimes \Lambda^q(\Lambda^{1,0}).$$
Definition 3.1. Let $(M, J)$ be an almost-complex manifold of real dimension $2m$. The rank 1 bundle

$$K = \Lambda^{m,0} \rightarrow M^{2m}$$

is called the canonical line bundle of $(M, J)$. Its dual

$$K^{-1} = \Lambda^m T^{1,0}$$

is called the anti-canonical line bundle.

Notice that we thus have a number of equivalent expressions for the first Chern class of $(M, J)$:

$$c_1(M, J) := c_1(T^{1,0}) = c_1(K^{-1}) = -c_1(K) = -c_1(\Lambda^{1,0}).$$

A Riemannian metric $g$ and an almost-complex structure $J$ on $M$ are said to be compatible iff $J$ is an orthogonal transformation with respect to $g$:

$$g(\cdot, \cdot) = g(J\cdot, J\cdot).$$

This is the same as requiring that the tensor field

$$\omega(\cdot, \cdot) = g(J\cdot, \cdot)$$

be skew-symmetric. When this happens, $\omega$ will be called the associated 2-form of $(g, J)$. Notice that $\omega$ is automatically $J$-invariant, in the sense that

$$\omega(J\cdot, J\cdot) = \omega(\cdot, \cdot),$$

which is to say that $\omega$ is a (real) $(1,1)$-form with respect to $J$. If $J$ is an almost-complex structure, and if $\omega$ is a real $(1,1)$-form, then we may, conversely define a symmetric tensor field $g$ by

$$g(\cdot, \cdot) = \omega(\cdot, J\cdot);$$

if $g$ is positive-definite, it is then a $J$-compatible metric for which $\omega$ is the associated 2-form.

If $g$ is any Riemannian metric on $M$, and if $J$ is any almost complex structure, then we can produce a $J$-compatible metric $h$ by setting $h = [g + g(J\cdot, J\cdot)]/2$. But any metric $h$ on $M$ may be uniquely written as $h = g(H\cdot, H\cdot)$, where the symmetric endomorphism $H$ of $TM$ corresponds to yet another Riemannian metric $g(H\cdot, \cdot)$. Since the set of such $H$'s is convex, this provides us with a deformation-retraction $J \mapsto HJH^{-1}$ of the space of almost-complex structures $J$ onto the space of $g$-compatible almost-complex structures on $M$. In particular, $M$ admits an almost-complex structure iff it admits some $J$ compatible with any given metric $g$.

Now if $(M, g)$ is an oriented Riemannian 4-manifold, and if $J$ is a compatible almost-complex structure, then $J$ has matrix

$$\begin{pmatrix}
1 & -1 \\
-1 & 1
\end{pmatrix} \quad \text{or} \quad \begin{pmatrix}
1 & -1 \\
-1 & 1
\end{pmatrix}$$

in an appropriate oriented orthonormal frame $e_1, \ldots, e_4$. The associated 2-form

$$\omega = e^1 \wedge e^2 \pm e^3 \wedge e^4$$

is therefore always either self-dual or anti-self-dual, and has norm $\sqrt{2}$. The self-dual/anti-self-dual distinction amounts to whether or not $J$ determines the given orientation on $M$. Conversely, every self-dual or anti-self-dual 2-form of norm $\sqrt{2}$
arises from a $g$-compatible $J$. Thus a smooth compact oriented 4-manifold $M$ admits an orientation-compatible almost-complex structure iff $\Lambda^+$ admits a nowhere-zero section.

In fact, the specification of an almost-complex structure $J$ compatible with $g$ and the orientation gives us a concrete alternate description of $\Lambda^+$. Indeed, if $e_1, \ldots, e_4$ is an oriented orthonormal frame in which $J$ is given by (5a), then $K$ is spanned by

$$(e^1 + ie^2) \wedge (e^3 + ie^4) = (e^1 \wedge e^3 - e^2 \wedge e^4) + i(e^1 \wedge e^4 + e^2 \wedge e^3),$$

the real and imaginary parts of which are self-dual 2-forms. Thus

$$\Lambda^+ = \mathbb{R}\omega \oplus \mathbb{R}eK,$$

and

$$\Lambda^+ \otimes \mathbb{C} = \mathbb{C}\omega \oplus K \oplus \bar{K}.$$

In particular,

$$(2\chi + 3\tau)(M) = p_1(\Lambda^+) = -c_2(\mathbb{C} \oplus K \oplus K^{-1}) = [c_1(K^{-1})]^2 = c_1^2(M, J).$$

We also see that the first Chern class satisfies the constraint

$$w_2(M) = w_2(\Lambda^+) = w_2(\mathbb{R}eK) \equiv c_1(K) \equiv c_1(M, J) \mod 2.$$ 

Conversely, if $\alpha \in H^2(M, \mathbb{Z})$ is any element satisfying

$$(6) \quad \alpha^2 = 2\chi + 3\tau$$

(7) \quad $\alpha \equiv w_2 \mod 2$

we may take $K$ to be a complex line bundle with $c_1(K) = -\alpha$, and notice that $\mathbb{R} \oplus \mathbb{R}eK$ then has the same characteristic classes $p_1$ and $w_2$ as $\Lambda^+$. Since these characteristic classes completely classify $SO(3)$-bundles over any 4-manifold [16], it follows that $\Lambda^+$ has a non-zero section, and that $M$ admits an orientation-compatible almost-complex structure, iff equations (6) and (7) have a solution $\alpha \in H^2(M, \mathbb{Z})$.

**Example 3.1.** The 4-sphere $S^4$ does not admit an almost-complex structure, since $H^2(S^4) = 0$, whereas $(2\chi + 3\tau)(S^4) = 4 \neq 0$. Notice, by the way, that the rank-3 bundle $\Lambda^+ \to S^4$ therefore does not admit a nowhere-zero section, even though its Euler class $e(\Lambda^+) \in H^3(S^4)$ is of course zero.

An almost-complex structure $J$ on a $2m$-manifold $M$ is said to be *integrable* if there is an atlas of charts on $M$ in which $J$ becomes the standard, constant-coefficient almost-complex structure on $\mathbb{R}^{2m} = \mathbb{C}^m$. For such an atlas, the transition functions are biholomorphisms, and $M$ acquires the structure of a complex $m$-manifold. In this case, we will therefore say that $J$ is a *complex structure* on $M$. If $\nabla$ is any torsion-free connection on $TM$, the Newlander-Nirenberg theorem asserts that the obstruction to integrability is precisely the $(\Lambda^{2,0} \oplus \Lambda^{0,2}) \otimes TM$ component of $\nabla J$; the latter is usually called the *Nijenhuis tensor* or the *Fröhlicher torsion*. An easy partition-of-unity argument therefore shows that $J$ is integrable iff there is a torsion-free connection $\nabla$ such that $\nabla J = 0$.

A Riemannian metric $g$ is said to be Kähler with respect to a compatible almost-complex structure $J$ iff $\nabla J = 0$, where $\nabla$ is now the Riemannian (Levi-Civita) connection. When this happens, $(M, J)$ is a complex manifold, per the above discussion. Moreover, the 2-form $\omega$, which is now known as the *Kähler form*, satisfies $\nabla \omega = 0$, and so is closed. Conversely, $g$ is Kähler with respect to $J$ iff $J$ is integrable and $\omega$ is closed. Since $g$ is completely determined by $J$ and $\omega$, this
allows one to construct all Kähler manifolds as complex manifolds equipped with closed, real, non-degenerate \((1,1)\)-forms.

The Kähler concept may be further clarified by a discussion of holonomy. On any Riemannian manifold \((M,g)\), parallel transport around a piece-wise smooth loop \(\gamma\) based at \(x \in M\) gives rise to a so-called holonomy transformation \(L_\gamma: T_x M \to T_x M\). Of course, \(L_\gamma\) is automatically an orthogonal transformation, since Riemannian parallel transport preserves \(g\). The Kähler condition may now be re-stated as requiring that every \(L_\gamma\) be a unitary transformation. (When this happens, the relevant complex structure on \(T_x M\) can be declared to be \(J|_x\), and this can be uniquely extended to an almost-complex structure \(J\) on \(M\) by Riemannian parallel transport.) Since curvature just represents parallel transport around infinitesimal loops, it follows the curvature tensor of a Kähler manifold is a 2-form with values in the skew-Hermitian endomorphisms of the tangent space. But index-lowering with \(g\) identifies the skew-Hermitian endomorphisms of \(TM\) with the bundle \(\Lambda^{1,1}_R\) of real \((1,1)\)-forms. This tells us that the curvature operator \(\mathcal{R}\) of a Kähler manifold is just an endomorphism of \(\Lambda^{1,1}_R\), since the first Bianchi identity always tells one that \(\mathcal{R}\) is self-adjoint.

In particular, the 2-form \(\rho = \mathcal{R}(\omega/2)\) is of type \((1,1)\) on any Kähler manifold. Now one can use the first Bianchi identity and the fact that \(\nabla\nabla J = 0\) to show that

\[ \rho(\cdot, \cdot) = r(J\cdot, \cdot), \]

and the \((1,1)\)-form \(\rho\) is therefore called the Ricci form. On the other hand, \(\rho\) represents the half the real trace of the infinitesimal holonomy composed with \(J\), and so is \(-i\) times the curvature of the canonical line bundle \(K\) with its induced connection. The latter connection is called the Chern connection, and can be characterized by the fact that it preserves the induced inner product, and that its \((0,1)\) component is

\[
\begin{align*}
\Gamma(K) & \xrightarrow{\delta} \Gamma(\Lambda^{0,1} \otimes K) \\
\| & \| \\
\Gamma(\Lambda^{m,0}) & \xrightarrow{d} \Gamma(\Lambda^{m,1}).
\end{align*}
\]

Because the Ricci tensor and Ricci form are related in exactly the same way as are the metric and Kähler form, a Kähler manifold is Einstein iff

\[ \rho = \lambda \omega. \]

When this happens, \(g\) is called a compatible Kähler-Einstein metric on the complex manifold \((M,J)\), and \((M,g,J)\) is called a Kähler-Einstein manifold. If \(\lambda < 0\), this says that \(K\) is a ‘positive’ holomorphic line bundle, and the Kodaira embedding theorem tells us that \(K\) is ample, meaning that there is a holomorphic embedding of \((M,J)\) in complex projective space defined by the holomorphic sections of \(K^{\otimes \ell}\) for any sufficiently large \(\ell\). If \(\lambda = 0\), one instead concludes that \(K^{\otimes \ell}\) is holomorphically trivial for some \(\ell \neq 0\). Finally, \(\lambda > 0\) would imply that \(K^{-1}\) is ample.

In the \(\lambda > 0\) case, however, the ampleness of \(K^{-1}\) is not enough to guarantee the existence of a Kähler-Einstein metric. Indeed, if there were such a metric, it would follow [41] that the identity component of the biholomorphism group would be a complexification of the identity component of the isometry group. Since the latter group is compact, this constrains the Lie algebra of holomorphic vector fields to be a reductive Lie algebra. Thus we have one extra necessary condition for the existence of a Kähler-Einstein metric in the \(\lambda > 0\) case. But amazingly, the
necessary conditions we have described also turn out to be sufficient \([5, 68, 55, 63, 62]\) in real dimension 4:

**Theorem 3.1** (Aubin/Yau). A compact complex manifold \((M, J)\) admits a compatible Kähler-Einstein metric with \(\lambda < 0\) iff its canonical line bundle \(K\) is ample.

**Theorem 3.2** (Yau). A compact complex manifold \((M, J)\) admits a compatible Kähler-Einstein metric with \(\lambda = 0\) iff \((M, J)\) admits a Kähler metric and \(K^{\otimes \ell}\) is trivial for some positive integer \(\ell\).

**Theorem 3.3** (Tian). A compact complex surface \((M^4, J)\) admits a compatible Kähler-Einstein metric with \(\lambda > 0\) iff its Lie algebra of holomorphic vector fields is reductive and its anti-canonical line bundle \(K^{-1}\) is ample.

For further discussion, see the essays by Tian and Yau in this volume.

**Example 3.2.** Consider the Fermat hypersurface

\[ \{ [u : v : w : z] \in \mathbb{CP}^3 \mid u^k + v^k + w^k + z^k = 0 \} \]

of degree \(k\) in complex projective 3-space. The canonical line bundle \(K\) of such a surface is the restriction of the hyperplane line bundle raised to the power \(k - 4\). Moreover, the Lie algebra of holomorphic vector fields is trivial, except for \(k = 1\), where it is the reductive Lie algebra \(\mathfrak{so}(3, \mathbb{C})\), and \(k = 2\), where it is the reductive Lie algebra \(\mathfrak{so}(4, \mathbb{C})\). Thus these complex algebraic surfaces all admit compatible Kähler-Einstein metrics.

Notice that \(\lambda\) has the same sign as \(4 - k\). All of these surfaces are simply connected (by the Lefschetz theorem), so we see that knowing the fundamental group alone cannot allow one to predict the sign of the Einstein constant \(\lambda\).

The first two of these surfaces are just \(\mathbb{CP}^2\) and \(\mathbb{CP}^1 \times \mathbb{CP}^1\), and their Kähler-Einstein metrics are just the obvious homogeneous ones. The cubic surface \(k = 3\) is much more interesting; it is diffeomorphic to \(\mathbb{CP}^2 \# 6\mathbb{CP}^2\), and its \(\lambda > 0\) Kähler-Einstein metric is not known explicitly.

The quartic \((k = 4)\) surface has trivial canonical line bundle, and carries Ricci-flat Kähler metrics. Notice that this manifold has \(2\chi + 3\tau = c_1^2 = 0\), and so exactly saturates the Hitchin-Thorpe inequality of Theorem 2.2. Generalizations of this quartic, called \(K3\) surfaces, will be discussed at length below.

Finally, notice that most of the Einstein manifolds under consideration have \(\lambda < 0\). As we let \(k \to \infty\), we run through infinitely many different homeotypes. As we will see in a moment, these \(k > 4\) surfaces are examples of surfaces of general type.

The quartic in \(\mathbb{CP}^3\) provides us with the prototypical example of a \(K3\) surface. By the usual definition \([7]\), a compact complex surface is called a \(K3\) iff it is simply connected with \(c_1 = 0\). (As it turns out, however, a compact complex surface is a \(K3\) iff it is diffeomorphic to our quartic prototype.) Every \(K3\) admits Kähler metrics \([54]\), and in light of Theorem 3.2, therefore admits Ricci-flat Kähler metrics. Now recall that \(\Lambda^+ = \mathbb{R} \oplus K\) for a Kähler surface, and \(K\) is flat iff the Ricci curvature vanishes. Thus any Ricci-flat Kähler surface is locally hyper-Kähler in the sense of Theorem 2.2. In fact \([26]\), this is essentially the general case.

**Proposition 3.4** (Hitchin). Let \((M, g)\) be a compact oriented Einstein 4-manifold with \(2\chi + 3\tau = 0\). Then the pull-back of \(g\) to some finite cover of \(M\)
is either a Ricci-flat Kähler metric on a K3 surface, or else a flat metric on a 4-torus.

**Proof.** The proof of Theorem 2.2 tells us that \( g \) is Ricci-flat, and induces a flat connection on \( \Lambda^+ \). But the Cheeger-Gromoll splitting theorem asserts that any compact Ricci-flat manifold has universal cover equal to the Riemannian product of a compact, simply connected Ricci-flat manifold with a Euclidean space. Since any Ricci-flat manifold of dimension < 4 is necessarily flat, this tells us that the universal cover \( \tilde{M} \) of our 4-manifold \( M \) must either be compact, or else is Euclidean. In the latter case, Bieberbach’s theorem [10, 67] asserts that \( \tilde{M} \) is finitely covered by a flat torus.

We are left with the case in which \( \tilde{M} \) is compact. But the pulled-back metric \( \tilde{g} \) induces a flat connection on \( \Lambda^+ \), and the simple-connectivity of \( \tilde{M} \) then guarantees that \( \Lambda^+ \) is then spanned by parallel 2-forms. An arbitrary such form \( \omega \) of norm \( \sqrt{2} \) corresponds to a parallel almost-complex structure \( J \) on \( \tilde{M} \), and makes \( (\tilde{M}, \tilde{g}) \) into a Kähler manifold. We then have \( \Lambda^+ = \mathbb{R}\omega \oplus K \), and since \( \Lambda^+ \) is flat and trivial, so is \( K \). Thus \( (\tilde{M}, J) \) is a K3 surface, and \( \tilde{g} \) is a compatible Ricci-flat Kähler metric on this K3.

Let us now consider how the Kähler-Einstein complex surfaces fit into Kodaira’s general scheme of surface classification. The single most important invariant of a compact complex surface is its *Kodaira dimension*. Let \((M^4, J)\) be a compact complex 2-manifold, and let \( K = \Lambda^2(T^{1,0}M)^* \) be its canonical line bundle. For each positive integer \( \ell \), we have a tautological map \( K^{-\ell} \to [\Gamma(M, \mathcal{O}(K^\ell))]^* \) defined by evaluation of a global holomorphic section of \( K^\ell \) on an element of its dual line bundle. This map descends to a holomorphic map \( M - B_\ell \to \mathbb{P}([\Gamma(M, \mathcal{O}(K^\ell))]^*) \) with values in a projective space, but at the price of throwing out the base locus \( B_\ell \) where all the holomorphic sections of \( K^\ell \) vanish. The Kodaira dimension is defined to be the maximal complex dimension of the image of \( M - B_\ell \) as \( \ell \) ranges over the positive integers. Here \( 0 \) is assigned dimension \( -\infty \), so the Kodaira dimension is an element of \( \{-\infty, 0, 1, 2\} \). The classification of complex surfaces with Kodaira dimension \( < 2 \) and \( b_1 \) even is thoroughly understood. A complex surface is said to be of *general type* if its Kodaira dimension is 2.

A following procedure [7] provides a simple, beautiful way of modifying a complex surface without changing its Kodaira dimension.

**Definition 3.2.** Let \((M, J)\) be a compact complex surface, and let \( x \in M \) be any point. The **blow-up** of \( M \) at \( x \) is the unique compact complex surface \((\tilde{M}, \tilde{J})\) obtained by replacing \( x \) with a complex projective line \( \mathbb{C}P_1 \).

The introduced \( \mathbb{C}P_1 \) has self-intersection \( -1 \), and so is called a \((-1)\)-curve. The blow-up can be explicitly constructed by replacing a small ball around \( x \) with a tubular neighborhood of the zero section in the Chern class \(-1\) line bundle over \( \mathbb{C}P_1 \). Since the one-point compactification of this line bundle is diffeomorphic to \( \mathbb{C}P_2 \) in an orientation-reversing manner, the blow-up \( \tilde{M} \) is diffeomorphic to the connected sum \( M \# \mathbb{C}F_2 \). Notice that the blow-up procedure can be iterated as many times as we like, and so gives us complex structures on \( M \# k\mathbb{C}F_2 \) for each positive integer \( k \).

There is an inverse process, called **blowing down.** Indeed, if a complex surface \((\tilde{M}, \tilde{J})\) contains a \( \mathbb{C}P_1 \) of self-intersection \(-1\), it is necessarily the blow-up of some other surface. Moreover, one can iterate this procedure until one finally produces
a surface without \((-1)\)-curves. (The process must terminate after a finite number of steps because each blow-down reduces \(b_2\) by 1.) A complex surface \(X\) without \((-1)\)-curves is called a minimal surface. If \(M\) is obtained from \(X\) by some sequence of blow-ups, we say that \(X\) is a minimal model for \(M\). If \(M\) has Kodaira dimension \(\geq 0\), moreover, its minimal model is unique.

Using Nakai's criterion, the Kodaira-Enriques classification [7] and a result of Siu [54], the previous criteria for the existence of Kähler-Einstein metrics can be restated as follows:

**Corollary 3.5.** Let \((M, J)\) be a compact complex surface. Then the following are equivalent:

- \((M, J)\) admits a compatible Kähler-Einstein metric with \(\lambda < 0\);
- \((M, J)\) has ample canonical line bundle;
- \((2\chi + 3\tau)(M) > 0\), and every \(\mathbb{CP}^1 \subset (M, J)\) has self-intersection \(\leq -3\);
- \((M, J)\) is minimal, of general type, and contains no \((-2)\)-curves.

Here a \((-2)\)-curve means a \(\mathbb{CP}^1\) of self-intersection \(-2\). If a minimal complex surface of general type contains such curves, we can collapse them all to obtain a complex orbifold which has \(K\) ample in the orbifold sense. The Aubin/Yau proof then constructs [29, 64] a Kähler-Einstein orbifold metric on this so-called pluricanonical model. This shows that a complex surface is of general type iff it can be obtained from a Kähler-Einstein orbifold with \(\lambda < 0\) by resolving the singularities and blowing up.

**Corollary 3.6.** Let \((M, J)\) be a compact complex surface. Then the following are equivalent:

- \((M, J)\) admits a compatible Kähler-Einstein metric with \(\lambda = 0\);
- \((M, J)\) is finitely covered by a K3 surface or complex torus;
- \((M, J)\) is minimal, of Kodaira dimension 0, and has \(b_1\) even.

**Corollary 3.7.** Let \((M, J)\) be a compact complex surface. Then the following are equivalent:

- \((M, J)\) admits a compatible Kähler-Einstein metric with \(\lambda > 0\);
- \((M, J)\) has ample anti-canonical line bundle and reductive automorphism algebra;
- \((M, J)\) is \(\mathbb{CP}^2\), \(\mathbb{CP}^1 \times \mathbb{CP}^1\), or the blow-up of \(\mathbb{CP}^2\) at \(k\) distinct points, \(3 \leq k \leq 8\), with no three on a line and no six on a conic.

While there is no Kähler-Einstein metric on the blow-up of \(\mathbb{CP}^2\) at one or two points, there is [8, 46] an Einstein metric on the one-point blow-up which is conformally Kähler. There is reason to hope that this so-called Page metric on \(\mathbb{CP}^2 \# \mathbb{CP}^2\) has a companion on the two-point blow-up \(\mathbb{CP}^2 \# 2\mathbb{CP}^2\). On the other hand, one can show [37] that the only compact complex surfaces which might admit Hermitian but non-Kähler Einstein metrics are the blow-ups of \(\mathbb{CP}^2\) at one, two, or three points in general position.

### 4. Seiberg-Witten Estimates

The Hitchin-Thorpe argument treats the \(L^2\) norms of \(s\) and \(W^+\) as 'junk terms,' about which one knows nothing except that they are non-negative. Seiberg-Witten theory [33, 66], however, provides remarkable information about both these terms...
In this section, we will develop the rudiments of Seiberg-Witten theory, and explore some of its ramifications regarding the scalar curvature.

Let \((M, g)\) be a compact oriented Riemannian 4-manifold, and suppose that \(M\) admits an almost-complex structure. As we saw in §3, we can then find almost complex structures \(J\) which are compatible with \(g\) in the sense that \(J^*g = g\). Choose such a \(J\), and consider the rank-2 complex vector bundles

\[
\begin{align*}
V_+ &= \Lambda^{0,0} \oplus \Lambda^{0,2} \\
V_- &= \Lambda^{0,1}.
\end{align*}
\]

Notice that \(g\) induces canonical Hermitian inner products on these bundles.

As described, these bundles depend on the choice of a particular almost-complex structure, but they have a deeper meaning [26] that is invariant under deformations of \(J\). Indeed, on any contractible open subset of \(M\) one can define Hermitian vector bundles

\[
\mathbb{C}^2 \to S_\pm \\
\downarrow \\
M
\]

called spin bundles, characterized by the fact that their determinant line bundles \(\Lambda^2 S_\pm\) are canonically trivial and that their projectivizations

\[
\mathbb{C}P_1 \to \mathbb{P}(S_\pm) \\
\downarrow \\
M
\]

are exactly the unit 2-sphere bundles \(S(\Lambda^\pm)\). On the other hand, one cannot generally define the bundles \(S_\pm\) globally on \(M\); manifolds on which this can be done are called spin, and are characterized by the vanishing of the Stiefel-Whitney class \(w_2 = w_2(TM)\). However, our bundles \(V_\pm\) still satisfy

\[
\mathbb{P}(V_\pm) = S(\Lambda^\pm),
\]

and we formally have

\[
V_\pm = S_\pm \otimes L^{1/2},
\]

where the Hermitian complex line bundle \(L = \Lambda^2 V_\pm\) is just the anti-canonical line-bundle \(K^{-1}\) associated with \(J\).

The isomorphism class \(c\) of such a choice of \(V_\pm\) is called a spin\(^c\) structure on \(M\). The cohomology group \(H^2(M, \mathbb{Z})\) acts freely and transitively on the spin\(^c\) structures by tensoring \(V_\pm\) with complex line bundles. Each spin\(^c\) structure has a first Chern class \(c_1 := c_1(L) = c_1(V_\pm) \in H^2(M, \mathbb{Z})\) such that

\[
c_1 \equiv w_2 \mod 2,
\]

and the previously mentioned \(H^2(M, \mathbb{Z})\)-action induces the action \(c_1 \mapsto c_1 + 2\alpha, \alpha \in H^2(M, \mathbb{Z})\), on first Chern classes. Thus, if \(H^2(M, \mathbb{Z})\) has trivial 2-torsion — as can always be arranged by replacing \(M\) with a finite cover — the spin\(^c\) structures are precisely in one-to-one correspondence with the set of cohomology classes \(c_1 \in H^2(M, \mathbb{Z})\) satisfying (10). A spin\(^c\) structures \(c\) arises from some almost-complex structure \(J\) iff its first Chern class satisfies the additional constraint

\[
c_1^2 = 2\chi + 3r.
\]

It is with these spin\(^c\) structures of almost-complex type we will concern ourselves here.
The Levi-Civita connection $\nabla$ of $g$ naturally induces Hermitian connections on the locally defined bundles $S_{\pm}$. Given a spin$^c$ structure $c$ and a Hermitian connection $A$ on the anti-canonical line bundle $L$, we therefore have induced Hermitian connections $\nabla_A$ on $V_{\pm}$. On the other hand, there is a canonical isomorphism $\Lambda^1 \otimes \mathbb{C} = \text{Hom}(S_+, S_-)$, so that $\Lambda^1 \otimes \mathbb{C} \cong \text{Hom}(V_+, V_-)$ for any spin$^c$ structure, and this induces a canonical homomorphism

$$\cdot : \Lambda^1 \otimes V_+ \rightarrow V_-$$

called Clifford multiplication. Composing these operations allows us to define a so-called twisted Dirac operator

$$D_A : \Gamma(V_+) \rightarrow \Gamma(V_-)$$

by $D_A \Phi = \nabla_A \cdot \Phi$. This is an elliptic operator of index

$$\text{ind} CD_A = \dim \text{Cker } D_A - \dim \text{Cker } D_A^* = \frac{c_2^2 - \tau(M)}{8}.$$  

If $c$ is of almost-complex type, this number becomes the Todd genus $(\chi + \tau)/4$ of the almost-complex manifold $(M, J)$.

**Example 4.1.** Let $(M, g, J)$ be a Kähler manifold of complex dimension 2. Let $c$ be the spin$^c$ structure induced by $J$, and let $A$ be the usual (Chern) connection on the anti-canonical line bundle $L = K^{-1}$. Then

$$D_A = \sqrt{2}(\overline{\partial} + \partial)^* : \Gamma(\Lambda^{0,0} \oplus \Lambda^{0,2}) \rightarrow \Gamma(\Lambda^{0,1})$$

where $\overline{\partial}$ is the Dolbeault operator and $\partial^*$ is its formal adjoint. In particular, the index of $D_A$ is just the alternating sum of the dimensions of the Dolbeault cohomology groups $H^{0,q}(M)$, and our (Noether) formula for its index (the Todd genus) is an elementary consequence of the Hodge decomposition.

For any spin$^c$ structure, we have already noted that there is a canonical diffeomorphism $\mathbb{P}(V_+) \cong S(\Lambda^+).$ In polar coordinates, we now use this to define the angular part of a unique continuous map

$$\sigma : V_+ \rightarrow \Lambda^+$$

with

$$|\sigma(\Phi)| = \frac{1}{2\sqrt{2}}|\Phi|^2.$$  

This map is actually real-quadratic on each fiber of $V_+; \text{ indeed, assuming our spin}^c \text{ structure is induced by a complex structure } J, \text{ then, in terms of (8), } \sigma \text{ is explicitly given by}$$

$$\sigma(f, \phi) = (|f|^2 - |\phi|^2)\frac{\omega}{4} + \Im (\overline{f}\phi),$$

where $f \in \Lambda^{0,0}$, $\phi \in \Lambda^{0,2}$, and where $\omega(\cdot, \cdot) = g(J\cdot, \cdot)$ is the associated 2-form of $(M, g, J)$.

We are now in a position to introduce the Seiberg-Witten equations

$$D_A \Phi = 0$$

(11)

(12)

$$F_A^+ = i\sigma(\Phi),$$

where the unknowns are a Hermitian connection $A$ on $L$ and a section $\Phi$ of $V_+$. Here $F_A^+$ is the self-dual part of the curvature of $A$, and so is a purely imaginary 2-form. A solution of (11–12) is said to be **reducible** if $\Phi \equiv 0$; otherwise, it is called **irreducible**.
Example 4.2. Let \((M, g, J)\) be a Kähler surface of scalar curvature \(s \equiv -1\). Let \(\Phi = (1, 0) \in \Lambda^{0,0} \oplus \Lambda^{0,2}\), and let \(\Lambda\) be the Chern connection on \(L = K^{-1}\).

Since \(F_A = -i \rho = -i r(J, \cdot)\), its self-dual part corresponds to the trace piece of the Ricci tensor, and so is given by \(F_A^+ = -is \omega/4 = i \omega/4\). On the other hand, \(D_A \Phi = \delta(1) + \delta^*(0) = 0\), and \(\sigma(\Phi) = \sigma(1, 0) = \omega/4\). Thus \((\Phi, A)\) is an irreducible solution of the Seiberg-Witten equations (11–12).

The geometric character of the Seiberg-Witten equations makes them invariant under automorphisms of \(L\). Thus the so-called gauge group of smooth maps \(f : M \to S^1\) acts on the space of smooth solutions of the Seiberg-Witten equations (11–12) by

\[
(A, \Phi) \mapsto (A - 2f^{-1}df, f\Phi).
\]

Notice that this action is free on the set of irreducible solutions, whereas the stabilizer is precisely \(S^1\) if \(\Phi \equiv 0\). In particular, the solution space is always either infinite-dimensional or empty. However, one can compensate for the action of the gauge group by choosing some background connection \(A_0\) and then imposing the gauge-fixing condition

\[
d^* (A - A_0) = 0.
\]

The system (11–13) is then elliptic, and the solution space is finite dimensional. There is still a residual part of the action of the gauge group; the constant \(S^1\)-valued functions still act, and after modding out by these there is still an action of the discrete group \(H^1(M, \mathbb{Z})\) of homotopy classes of maps \(M \to S^1\). After dividing out by these, however, we obtain the moduli space \(\mathcal{M}_{c,g}\), which is by definition the set solutions of (11–12) modulo the action of the gauge group.

Not only is this moduli space finite-dimensional — it is also compact. This is because (11–12) imply the Weitzenböck formula

\[
0 = 4 \nabla_A^* \nabla_A \Phi + s\Phi + |\Phi|^2 \Phi.
\]

Taking the inner product with \(\Phi\), we have

\[0 = 2\Delta |\Phi|^2 + 4|\nabla_A \Phi|^2 + s|\Phi|^2 + |\Phi|^4,\]

and at the maximum of \(|\Phi|^2\) we therefore have

\[0 \geq 4|\nabla_A \Phi|^2 + |\Phi|^2(s + |\Phi|^2),\]

so that any irreducible solution must satisfy the \(C^0\) estimate

\[
|\Phi|^2 \leq -\min s.
\]

Moreover, equality can only occur at points where \(\nabla_A \Phi = 0\). In particular, one has uniform \(L^p\)-bounds on \(\Phi\) for all solutions, and compactness therefore follows \([33, 42]\) via the \(L^p\) versions of the Gårding inequality for (11–13) and the Rellich lemma.

Now consider a perturbed versions of the Seiberg-Witten equations, obtained by replacing (12) with

\[
iF_A^+ + \sigma(\Phi) = \epsilon,
\]

where \(\epsilon\) is some self-dual 2-form. For generic \(\epsilon\), Smale’s infinite-dimensional version of Sard’s theorem implies that the corresponding ‘perturbed’ moduli space \(\mathcal{M}_{c,g,\epsilon}\) is
a smooth manifold whose dimension is given by the (real) index\textsuperscript{1} of the linearization of (11–13), which is to say that

\begin{equation}
\dim \mathcal{M}_{c,g,\varepsilon} = \left(b_1 - 1 - b^+\right) + 2 \text{ind} \mathcal{C}D_A \\
= -\frac{\chi + \tau}{2} + 2 \left(\frac{c_1^2 - \tau}{8}\right) \\
= \frac{c_1^2 - (2\chi + 3\tau)}{4}.
\end{equation}

If our spin\textsuperscript{c} structure \(c\) is of almost-complex type, the moduli space is therefore discrete. Moreover, a slight variation on the previous Weitzenböck argument shows that these moduli spaces are compact. Again assuming that our spin\textsuperscript{c} structure is of almost-complex type, the moduli spaces \(\mathcal{M}_{c,g,\varepsilon}\) are thus finite for generic \(\varepsilon\).

As we vary \(g\) and \(\varepsilon\), the moduli spaces remain cobordant as long as one can avoid hitting reducible solutions. Now a reducible solution can only occur when the self-dual part \(2\pi c_1^+\) of the harmonic representative of \(2\pi c_1 = [iF_A]\) agrees with the harmonic part of \(\varepsilon\). Since \((c_1^+)^2 \geq c_1^2\), it follows that we can avoid reducible solutions if we assume that

\begin{equation}
2\chi + 3\tau = c_1^2 > 0,
\end{equation}

and if, for each metric, we only consider \(\varepsilon\) with \(L^2\) norm smaller than \(2\pi \sqrt{c_1^2}\). Thus (18) is enough to guarantee that we have a cobordism class of \(\mathcal{M}_{c,g,\varepsilon}\) determined by the smooth structure of \(M\) and the spin\textsuperscript{c} structure \(c\) alone.

**Definition 4.1.** Let \((M,c)\) be a smooth compact 4-manifold, equipped with the spin\textsuperscript{c} structure and orientation determined by some almost-complex structure \(J\). Assume that (18) holds. Then the (mod 2) Seiberg-Witten invariant \(n_c(M) \in \mathbb{Z}_2\) is defined to be

\[ n_c(M) = \#\mathcal{M}_{c,g,\varepsilon} \mod 2, \]

where \(g\) is any Riemannian metric on \(M\) and \(\varepsilon\) is a generic self-dual form of small \(L^2\) norm on \((M,g)\).

Notice that (18) implies that \(b_+(M) \geq 1\). On the other hand, if \(b_+(M) \geq 2\), the set of \(\varepsilon\) for which there is a reducible solution has codimension \(\geq 2\); it is then easy to see that the generic moduli spaces \(\mathcal{M}_{c,g,\varepsilon}\) are all cobordant, and one can thus define the Seiberg-Witten invariant even if (18) fails. However, the Hitchin-Thorpe inequality makes (18) a very natural hypothesis for investigations concerning Einstein manifolds, and adopting it here will enable us to treat the \(b^+ = 1\) and \(b^+ \geq 2\) cases simultaneously.

We have now defined an elegant invariant of a smooth compact 4-manifold by counting solutions of a non-linear system of partial differential equations. But have we merely given a complicated definition of zero? Fortunately not!

**Theorem 4.1 (Witten/Kronheimer).** Let \((M,J)\) be a complex surface of general type for which (18) holds. Then \(n_c(M) \neq 0\), where \(c\) is the spin\textsuperscript{c} structure induced by \(J\).

\textsuperscript{1}This dimension count actually involves a subtle cancellation which is often overlooked. Namely, the contribution due to the 1-dimensional cokernel of \(d^* : \Gamma(\Lambda^1) \rightarrow \Gamma(\Lambda^0)\) is canceled out by the action of the 1-dimensional group \(S^1\) of constant gauge transformations.
For simplicity, let us just sketch a proof assuming that \((M, J)\) satisfies any of the equivalent conditions catalogued by Corollary 3.5. There is then a \(J\)-compatible Kähler-Einstein metric \(g\) on \(M\) of scalar curvature \(s \equiv -1\), and hence there is an irreducible solution of the Seiberg-Witten equations obtained by taking \(\Phi = (1, 0) \in \Gamma(\Lambda^{0,0} \oplus \Lambda^{0,2})\) and letting \(A\) be the Chern connection on \(L = K^{-1}\). It is not hard to see that the linearization of (11–12) is surjective at this solution, so it suffices to show that any other solution is gauge-equivalent to the constructed one. But the \(C^0\) estimate (15) implies that

\[
8 \int_M |F^+_A|^2 d\mu = \int_M |\Phi|^4 d\mu \leq \int_M s^2 d\mu = 32\pi^2 c_1^2(M),
\]

and equality can only hold if \(|\Phi|^2 \equiv -s = 1\) and \(\nabla_A \Phi = 0\). But since \(A\) is a connection on \(L = K^{-1}\), we have

\[
4\pi^2 c_1^2 = \int_M [|F^+_A|^2 - |F^-_A|^2] d\mu,
\]

and the reverse inequality also holds. Thus \(\Phi\) has unit length, and is parallel with respect to \(\nabla_A\). It follows that \((A, \Phi)\) is gauge-equivalent to our explicit solution.

If (18) holds and the Seiberg-Witten invariant \(n_e(M)\) is non-zero, it follows that the unperturbed equations (11–12) have an irreducible solution for each metric. Via the Weitzenböck formula (14), this severely constrains the geometry of Riemannian metrics on the given manifold:

**Theorem 4.2.** Let \((M, \sigma)\) be a smooth compact oriented 4-manifold with a spin\(e\) structure of almost-complex type. Assume that (18) holds, and that \(n_e(M) \neq 0\). Let \(g\) be any Riemannian metric on \(M\). Then the scalar curvature \(s_g\) of \(g\) is negative somewhere, and

\[
\int_M s_g^2 d\mu \geq 32\pi^2 (c_1^+)^2,
\]

with equality only if \(g\) has constant negative scalar curvature, and is Kähler with respect to a \(c\)-compatible complex structure.

**Proof.** The \(C^0\) estimate (15) for any irreducible solution of (11–12) forces any metric to satisfy \(\min s < 0\).

Now any conformal class of metrics \([g] = \{u^2 g | u : M \to \mathbb{R}^+\}\) contains a metric of constant scalar curvature \([40, 51]\), and when the constant is negative, such a metric is unique up to scale. Moreover, such a metric minimizes the functional \(\int s^2 d\mu\) within its conformal class. It thus suffices to prove the stated lower bound for \(\int s^2 d\mu\) assuming that \(g\) has constant scalar curvature.

Now when \(s\) is constant, the \(C^0\) estimate (15) takes the form

\[
|\Phi|^2 \leq -s,
\]

and if equality holds identically we have \(\nabla_A \Phi \equiv 0\). Thus

\[
\int s^2 d\mu \geq \int |\Phi|^4 d\mu = 8 \int |F^+_A|^2 d\mu \geq 32\pi^2 (c_1^+)^2.
\]

If equality holds, \(\nabla\sigma(\Phi) \equiv 0\), and the metric is Kähler with respect to a complex structure \(J\) for which \(\sigma(\Phi)\) is a constant positive multiple of the Kähler form \(\omega\). Now

\[
V_+ = (\Lambda^{0,0} \oplus \Lambda^{0,2}) \otimes E
\]
for some Hermitian line bundle $E$ with connection, and $\Phi \neq 0$ is a parallel section of $\Lambda^{0,0} \otimes E \subset V_+$. Thus $E$ is trivial, and $c$ is exactly the spin$^c$ structure determined by $J$. \hfill \square

This has an immediate application to the sign-of-the-Einstein-constant problem. Recall that the Hitchin-Thorpe inequality implies that if a 4-manifold admits a Kähler-Einstein metric with $\lambda = 0$, any other Einstein metric on $M$ also has $\lambda = 0$. Seiberg-Witten theory implies the analogous conclusion in the negative case:

**Corollary 4.3.** Let $M$ be a smooth compact 4-manifold which admits a Kähler-Einstein metric $g$ with $\lambda < 0$. Then any other Einstein metric $\tilde{g}$ on $M$ also has $\lambda < 0$.

**Proof.** Give $M$ the orientation and spin$^c$ structure $c$ induced by the complex structure. Then (18) holds by the Hitchin-Thorpe inequality, and has $n_c(M) \neq 0$ by Theorem 4.1. Theorem 4.2 thus tells us that $\tilde{g}$ cannot have $s \geq 0$, and so must have $\lambda < 0$. \hfill \square

But we also have the following remarkable estimate:

**Corollary 4.4.** Let $M$ be an oriented smooth compact 4-manifold equipped with a spin$^c$ structure $c$ of almost-complex type. Assume that (18) holds, and that $n_c(M) \neq 0$. Then any Riemannian metric $g$ on $M$ satisfies

$$\frac{1}{32\pi^2} \int_M s_\phi^2 d\mu_g \geq (2\chi + 3\tau)(M),$$

with equality iff $g$ is Kähler-Einstein.

**Proof.** One has

$$(c_1^+)^2 \geq c_1^2 = 2\chi + 3\tau,$$

with equality iff the harmonic representative of $c_1$ is self-dual. But for a Kähler metric of constant scalar curvature, the Ricci form $\rho$ is the harmonic representative of $2\pi c_1$, and $\rho$ is self-dual iff the metric is Einstein. The claim therefore follows from Theorem 4.2. \hfill \square

This implies a Riemannian version [35] of the so-called Miyaoka-Yau [7, 68] inequality:

**Theorem 4.5 (LeBrun).** Let $(M, g)$ be a non-flat compact Einstein 4-manifold which admits an almost-complex structure. Give $M$ the induced orientation and spin$^c$ structure $c$, and assume that $n_c(M) \neq 0$. Then Euler characteristic $\chi$ and signature $\tau$ of $M$ satisfy

$$\chi \geq 3\tau,$$

with equality only if the universal cover of $(M, g)$ is complex-hyperbolic 2-space $\mathbb{C}H_2 := SU(2,1)/U(2)$, with a constant multiple of its standard metric.

**Proof.** By Corollary 4.4 and the Gauss-Bonnet formula, we have

$$3(2\chi - 3\tau)(M) = \frac{3}{4\pi^2} \int_M \left( 2|W^-|^2 + \frac{s^2}{24} \right) d\mu \\
\geq \frac{1}{32\pi^2} \int_M s^2 d\mu \\
\geq (2\chi + 3\tau)(M),$$
with equality only if the metric is Kähler-Einstein and $W^- \equiv 0$.

If the latter happens, $\Re K$, $\Re \omega$, and $\Lambda^-$ are eigenspaces of the curvature operator $\mathcal{R}$, with respective eigenvalues $0$, $s/4$, and $s/12$. Since $s$ is constant and $\omega$ is parallel, this implies that $\nabla \mathcal{R} \equiv 0$, and $g$ is therefore locally symmetric. Since $2\chi + 3\tau > 0$ by the Hitchin-Thorpe inequality, the assumption that $n_\epsilon \neq 0$ forces $s$ to be negative, and the point-wise form of $\mathcal{R}$ then tells us that the universal cover is isometric to a rescaled version of the symmetric space $\mathcal{CH}_2$. \qed

In particular, we get a uniqueness result [35]:

**Corollary 4.6 (LeBrun).** Let $M = \mathcal{CH}_2/\Gamma$ be a compact complex-hyperbolic 4-manifold, and let $g_0$ be its tautological metric. Then every Einstein metric $g$ on $M$ is of the form $g = \varphi^* c g_0$, where $\varphi : M \to M$ is a diffeomorphism and $c > 0$ is a constant.

**Proof.** Because $M$ carries a tautological Kähler-Einstein metric with $\lambda < 0$, Theorem 4.1 guarantees that $M$ has a non-trivial Seiberg-Witten invariant $n_\epsilon$. Up to rescaling, any Einstein metric $g$ on $M$ is therefore complex-hyperbolic by Theorem 4.5. But Mostow rigidity [43] tells us that the fundamental group of a complex hyperbolic 4-manifold determines the manifold up to isometry, so the result follows. \qed

For constructions of compact complex-hyperbolic hyperbolic 4-manifolds, see [13, 43]. For a generalization of the above result to the non-compact, finite-volume setting, see [11].

The analogous result for real-hyperbolic 4-manifolds is also true, and indeed was proved by Besson-Courtois-Gallot [9] several months before the discovery of the above result [35]. However, as will explained in §9, the proof of the real-hyperbolic result proceeds upon completely different lines.

Seiberg-Witten theory also allows one to prove non-existence results for Einstein metrics [36]. To see this, first notice that Theorem 4.4 can be sharpened in the following direction:

**Theorem 4.7.** Let $M$ be a 4-manifold which has a smooth connected sum decomposition

$$M = X \# \underbrace{\mathbb{CP}_2 \# \cdots \# \mathbb{CP}_2}_k$$

for some positive integer $k$. Assume that (18) holds, and suppose that $M$ has a spin$^c$ structure of almost-complex type for which $n_\epsilon(M) \neq 0$. Then every Riemannian metric $g$ on $M$ satisfies

$$\frac{1}{32\pi^2} \int_M s_g^2 d\mu_g > (2\chi + 3\tau)(X) = (2\chi + 3\tau)(M) + k.$$

**Proof.** The first Chern class $c_1$ of $\epsilon$ can be uniquely expressed as

$$c_1 = c_1(X) + \sum_{j=1}^k a_j E_j,$$

where $c_1(X) \in H^2(X, \mathbb{Z})$ and the $E_j$ are generators for the $k$ relevant copies of $H^2(\mathbb{CP}_2, \mathbb{Z})$. Since $c_1 \equiv w_2 \mod 2$ and $E_j \cdot E_j = -1$, the integers $a_j$ are all necessarily odd. Let $g$ be the arbitrary metric on $M$ which we wish to examine, and let us fix our choice of generators $E_j$ so that $c_1(X)^+ \cdot E_j \geq 0$ for all $j$. 

Now notice that complex conjugation $[z^1 : z^2 : z^3] \mapsto [\bar{z}^1 : \bar{z}^2 : \bar{z}^3]$ defines a self-diffeomorphism of $\mathbb{CP}^2$ with non-empty fixed-point set which acts by $-1$ on $H^2$. Using this as a model, one can construct self-diffeomorphisms of $M$ which act trivially on $H^2(X)$ and reverse the sign of exactly one $E_j$. Moving $\epsilon$ by a sequence of such diffeomorphisms, one can thus obtain a spin$^c$ structure $\tilde{\epsilon}$ with $n_{\tilde{\epsilon}} \neq 0$ and

$$\tilde{c}_1 = c_1(X) + \sum_{j=1}^{k} |a_j|E_j.$$ 

Thus

$$(\tilde{c}_1^+)^2 = [c_1(X)^+]^2 + (\sum |a_j|E_j^+)^2 + 2(\sum |a_j|(c_1(X)^+ \cdot E_j)
\geq [c_1(X)^+]^2
\geq [c_1(X)]^2
= \tilde{c}_1^2 + \sum a_j^2
\geq (2\chi + 3\tau)(M) + k.$$ 

Theorem 4.2 therefore tells us that

$$\frac{1}{32\pi^2} \int_M s^2 d\mu \geq (2\chi + 3\tau)(M) + k = (2\chi + 3\tau)(X).$$

If equality held, $g$ would be necessarily be Kähler with respect to a $\tilde{\epsilon}$-compatible complex structure. Also, our specification of $\tilde{\epsilon}$ would have not been unique, since we would also have $[c_1(X)^+] \cdot E_j = 0$ for all $j$. The same reasoning could thus be applied to $2^k$ different spin$^c$ structures to give us $2^{k+1}$ different parallel complex structures, and $\Lambda^+$ would therefore have to be flat and trivial. But any two parallel sections would then by deformation-equivalent, and hence determine the same spin$^c$ structure, contradicting the construction. Hence the inequality is necessarily strict.

In particular, we get the following non-existence result:

**Theorem 4.8 (LeBrun).** Let $X$ be a minimal complex algebraic surface of general type, and let $M = X \# k\mathbb{CP}^2$ be obtained from $X$ by blowing up $k > 0$ points. If $k \geq \frac{2}{3}(2\chi + 3\tau)(X)$, then $M$ does not admit Einstein metrics.

**Proof.** If $M$ admits an Einstein metric $g$, it satisfies (18) by the Hitchin-Thorpe inequality, and has a non-trivial Seiberg-Witten invariant by Theorem 4.1. Using the scalar curvature estimate of Theorem 4.7 and the Gauss-Bonnet formula (4), we have

$$(2\chi + 3\tau)(X) - k = (2\chi + 3\tau)(M)$$

$$= \frac{1}{4\pi^2} \int_M \left(2|W^+|^2 + \frac{s^2}{24}\right) d\mu$$
$$\geq \frac{1}{3 \cdot 32\pi^2} \int_M s^2 d\mu$$
$$> \frac{1}{3}(2\chi + 3\tau)(X),$$

so that

$$\frac{2}{3}(2\chi + 3\tau)(X) > k.$$
Hence $M$ cannot admit an Einstein metric if $k \geq \frac{2}{3}(2\chi + 3\tau)(X)$.

**Corollary 4.9** (LeBrun). *Even up to homeomorphism, there are infinitely many smooth compact simply connected 4-manifolds which satisfy the strict Hitchin-Thorpe inequality $2\chi > 3|\tau|$, but nevertheless do not admit Einstein metrics.*

**Proof.** If $X$ is any minimal complex surface of general type with $2\chi + 3\tau \geq 3$, there is then at least one integer $k$ satisfying $(2\chi + 3\tau)(X) > k \geq \frac{2}{3}(2\chi + 3\tau)(X)$. The complex surface $M = X \# k\overline{\mathbb{C}P_2}$ then satisfies $2\chi > 3|\tau|$, but does not admit Einstein metrics by the above result. Theorem 4.9 therefore follows by considering the sequence of $X$'s given by the hypersurfaces of degree $> 4$ in $\mathbb{C}P_3$.

It should be pointed out that, even when $b^+ = 1$, Seiberg-Witten theory can be used to prove results along the lines of Theorem 4.7 without assuming (18). However, the proofs are complicated by the metric-dependence of the moduli spaces, and one is saved only by considering different spin$^c$ structures for different metrics. For details, see [20, 36].

It should also be observed that the results of this section really depend only on the existence of solutions of the Seiberg-Witten equations for each metric on the given manifold. This may occur even when $n_c \in \mathbb{Z}_2$ vanishes. In particular, it turns out that the moduli spaces $\mathcal{M}_{c,g}$ can be oriented in a natural way, and this gives rise to an invariant $SW_c \in \mathbb{Z}$ whose mod 2 reduction is $n_c$. One can also generalize the definition of $SW_c$ so as to allow [59] for spin$^c$ structures which are not of almost-complex type. Some of these invariants turn out, moreover, to be non-trivial [58] on any symplectic 4-manifold with $b_+ \geq 2$.

On the other hand, Kronheimer [34] recently showed that certain 4-manifolds with $SW \equiv 0$ nonetheless admit solutions of the Seiberg-Witten equations for each and every metric. A different construction of such examples, with direct relevance to the theory of Einstein manifolds, is described in §5 below. In any case, it would seem that the Seiberg-Witten equations have ramifications for the theory of Einstein manifolds which in contexts beyond the scope of the invariants which have been explored to date.

**5. Surgery and Scalar Curvature**

We have already observed that lower bounds for the $L^2$ norm of the scalar curvature have natural applications to the theory of Einstein metrics on 4-manifolds. Let us consider such bounds in a broader context. If $M$ is a smooth compact $n$-manifold, consider the diffeomorphism invariant

$$\mathcal{I}(M^n) = \inf_g \int_M |s_g|^{n/2} d\mu_g,$$

where the infimum is taken over all metrics on $M$. Notice that choice of the power $n/2$ is dictated by scale invariance; for any other power, the analogous infimum would be zero.

The invariant $\mathcal{I}$ is well behaved with respect to under surgeries in high codimension [49]; this fact is essentially a quantitative refinement of results of Gromov-Lawson [22] and Schoen-Yau [52] concerning metrics of positive scalar curvature. Recall that if $M$ is any smooth compact $n$-manifold, and if $S^q \subset M$ is a smoothly embedded $q$-sphere with trivial normal bundle, we may construct a new $n$-manifold
by replacing a tubular neighborhood $S^q \times \mathbb{R}^{n-q}$ of $S^q$ with $S^{n-q-1} \times \mathbb{R}^{q+1}$. One then says $\hat{M}$ is obtained from $M$ by performing a surgery in codimension $n - q$ (or dimension $q$). This operation precisely describes the way that level sets of a Morse function change as one passes a critical point of index $q + 1$, and two manifolds are therefore cobordant iff one can be obtained from the other by such a sequence of surgeries.

**Proposition 5.1 (Petean-Yun).** Let $M$ be any smooth compact $n$-manifold, and let $\hat{M}$ be obtained from $M$ by performing a surgery in codimension $\geq 3$. Then

$$I(\hat{M}) \leq I(M).$$

**Proof.** We may assume that $n \geq 3$, as otherwise there is nothing to prove. But with this assumption, $I$ can be rewritten as

$$I(M) = \inf_g \int_M |s_{-g}|^{n/2} d\mu_g,$$

where

$$s_- = \begin{cases} 
0 & s \geq 0 \\
-1 & s \leq 0.
\end{cases}$$

Indeed, if $M$ admits a metric of positive scalar curvature, it also [5] admits a metric with $s \equiv 0$, so both infima vanish. If, on the other hand, $M$ does not admit a metric of positive scalar curvature, both functionals [9] are minimized in each conformal class by a metric of constant scalar curvature $\leq 0$, and the claim is then an immediate consequence.

Now let $g$ be a metric on $M$ such that

$$\int_M |s_{-g}|^{n/2} d\mu_g < I(M) + \frac{\epsilon}{2},$$

and suppose that $S^q \subset M$ is any embedded sphere of codimension $n - q \geq 3$. By making a conformal change which is trivial outside a small tubular neighborhood of the sphere, one may produce a conformally related metric $\tilde{g} = ug$ which has positive scalar curvature along $S^q$, but still satisfies

$$\int_M |s_{-\tilde{g}}|^{n/2} d\mu_{\tilde{g}} < I(M) + \epsilon.$$ 

But on the manifold $\hat{M}$ obtained by surgery on $S^q$, a celebrated local construction of Gromov-Lawson [22] then gives us a metric $\hat{g}$ which has positive scalar curvature in the surgered region, and agrees with $\tilde{g}$ on the set where $s_{\tilde{g}} \leq 0$. Thus

$$\int_{\hat{M}} |s_{-\hat{g}}|^{n/2} d\mu_{\hat{g}} = \int_M |s_{-\tilde{g}}|^{n/2} d\mu_{\tilde{g}} < I(M) + \epsilon,$$

so that

$$I(\hat{M}) = \inf_{\hat{g}} \int_{\hat{M}} |s_{-\hat{g}}|^{n/2} d\mu_{\hat{g}} \leq I(M),$$

as claimed. \(\square\)

Because any surgery can be undone by another surgery, this implies [47]

**Corollary 5.2 (Petean).** If $M$ is any smooth compact $4$-manifold, then

$$I(M \# [S^1 \times S^3]) = I(M).$$
Proof. One may obtain $S^1 \times S^3$ from $S^4$ by a surgery in codimension 4, and $S^4$ may be obtained from $S^1 \times S^3$ by a surgery in codimension 3. Taking connected sums with $M$, we see that $M = M \# S^4$ and $M \# [S^1 \times S^3]$ can each be obtained from the other by a surgery in codimension $\geq 3$. By Theorem 5.1, it follows that

$$\mathcal{I}(M) \leq \mathcal{I}(M \# [S^1 \times S^3]) \leq \mathcal{I}(M),$$

and the result follows. \hfill \square

This and Theorem 4.7 now imply [47]

Theorem 5.3. Let $X = CH_2/\Gamma$ be any compact complex-hyperbolic 4-manifold. Then $X \# \ell(S^1 \times S^3)$ does not admit Einstein metrics for any $\ell > 0$. Moreover, $X \# kCP_2 \# \ell(S^1 \times S^3)$ does not admit Einstein metrics for any $\ell \geq \frac{5}{4}k > 0$.

Proof. Let $M = X \# kCP_2$ and $\hat{M} = M \# \ell(S^1 \times S^3) = X \# kCP_2 \# \ell(S^1 \times S^3)$ for some $k \geq 0$ and $\ell > 0$. If $\hat{M}$ admitted an Einstein metric, then the Hitchin-Thorpe inequality would then tell us that

$$(2\chi + 3\tau)(M) = (2\chi + 3\tau)(\hat{M}) + 4\ell > 0,$$

and $M$ would thus satisfy (18). But $M$ is the underlying 4-manifold of a complex surface of general type, and hence has a non-zero Seiberg-Witten invariant by Theorem 4.1. Thus Theorem 4.2 tells us that

$$\mathcal{I}(M) = \inf g \int_M s^2 d\mu \geq 32\pi^2 (2\chi + 3\tau)(X).$$

Hence

$$\mathcal{I}(\hat{M}) = \mathcal{I}(M) \geq 32\pi^2 (2\chi + 3\tau)(X)$$

by Theorem 5.1. Now suppose that $\hat{g}$ is an Einstein metric on $\hat{M}$. Then the Gauss-Bonnet formula (4) tells us that

$$(2\chi - 3\tau)(\hat{M}) = \frac{1}{4\pi^2} \int_{\hat{M}} \left(2|W^-|^2 + \frac{s^2}{24}\right) d\mu_{\hat{g}}$$

$$\geq \frac{1}{3 \cdot 32\pi^2} \mathcal{I}(\hat{M})$$

$$\geq \frac{1}{3}(2\chi + 3\tau)(X)$$

$$= (2\chi - 3\tau)(X),$$

where in the last step we have used the fact that $\chi(X) = 3\tau(X)$ for any complex-hyperbolic 4-manifold $X$. But since

$$(2\chi - 3\tau)(\hat{M}) = (2\chi - 3\tau)(X) + 5k - 4\ell,$$

this tells us that

$$5k - 4\ell \geq 0$$

if $\hat{M} = X \# kCP_2 \# \ell(S^1 \times S^3)$ admits an Einstein metric. The claim now follows by contraposition. \hfill \square

In particular, this yields a new proof of a beautiful result of Sambusetti [50], whose own proof will be discussed in §9 below.

Corollary 5.4 (Sambusetti). Let $(a, b)$ be any pair of integers with $a \equiv b \mod 2$. Then there is a smooth compact oriented 4-manifold $M$ which does not admit Einstein metrics, such that $\chi(M) = a$, $\tau(M) = b$. 
PROOF. Let $X_0$ be any compact complex-hyperbolic 4-manifold $CH_2/\Gamma$. Since $\Gamma \subset SU(1,2)$ is a finitely generated matrix group, Mal’tsev’s theorem [69, p. 151] asserts that it is residually finite, and in particular has a non-trivial homomorphism to a finite group. The kernel $\Gamma_1$ of such a homomorphism then defines a finite cover $X_1 = CH_2/\Gamma_1$ of $X_0$. Iterating this procedure, we obtain an infinite tower

$$\cdots \rightarrow X_i \rightarrow \cdots \rightarrow X_1 \rightarrow X_0$$

of finite covers of $X_0$. Thus there is a sequence of integers $j_i \rightarrow \infty$ such that each $(3j_i, j_i) = (\chi(X_i), \tau(X_i))$ for some complex-hyperbolic 4-manifold $X_i$. Now choose $i$ to be large enough so that

$$j_i > \max(b, \frac{a+b}{4}, \frac{2a}{3} - b).$$

The positive integers

$$k = j_i - b$$

$$\ell = 2j_i - \frac{a+b}{2}$$

then satisfy $\ell > \frac{5}{4} k$. Thus $M = X_i # k\overline{CP}^2 \# \ell(S^1 \times S^3)$ has

$$\chi(M) = 3j_i + k - 2\ell = a$$

$$\tau(M) = j_i - k = b,$$

and yet, by Theorem 5.3, does not admit Einstein metrics. \hfill \Box

The reader should note that the $\chi$ and $\tau$ are necessarily congruent mod 2, so that the above result is optimal.

Probing the scalar curvature estimates provided by Seiberg-Witten theory with concrete sequences of test metrics actually allows one to compute $I$ for any complex surface of general type [36]. Applying Corollary 5.2 then yields

PROPOSITION 5.5. If $X$ is any complex surface of general type, then

$$I(X # k\overline{CP}^2 \# \ell(S^1 \times S^3)) = 32\pi^2 (2\chi + 3\tau)(X).$$

This of course immediately implies non-existence results for Einstein metrics for certain values of $k$ and $\ell$. The curious thing about this argument, however, is that it ultimately exploits the existence of solutions of the Seiberg-Witten equations (11–12) on one manifold to obtain scalar curvature estimates on another! Might it not be more satisfying to show that there are actually solutions of the Seiberg-Witten equations on the manifold in question? Fortunately, as was recently proved by Ozsváth and Szabó [45], and noticed independently by the present author and various others, such Seiberg-Witten solutions do in fact exist.

To this end, suppose that we have a spin$^c$ structure $c$ on a 4-manifold $M$ for which (17) predicts that the moduli space $\mathcal{M}_{c,g,e}$ generically has dimension $\ell > 0$. Fix $\ell$ loops $\gamma_1, \ldots, \gamma_\ell$ in $M$, and define a map $\mathcal{M}_{c,g,e} \rightarrow T^\ell$ by sending $(\Phi, A)$ to the holonomies of $A$ around the $\ell$ given loops. For a fixed $(c, g, e)$, the homotopy class of this map only depends on the homology classes $[\gamma_i] \in H_1(M, \mathbb{Z})$, and we may therefore define $n_c(M, [\gamma_1], \ldots, [\gamma_\ell]) \in \mathbb{Z}_2$ to be the degree mod 2 of this map. If this invariant is non-zero for some choice of $[\gamma_i]$, it of course follows in particular that there must be a Seiberg-Witten solution for every metric $g$ on $M$. 
THEOREM 5.6. Let $N$ be a complex surface of general type, let $M = N \# \ell [S^1 \times S^3]$, and let $\gamma_1, \ldots, \gamma_\ell$ be $S^1$ factors of the $\ell$ relevant copies of $S^1 \times S^3$. Assume, for simplicity, that $M$ satisfies (18), and let $\mathfrak{c}$ be the spin$^c$ structure on $M$ obtained by pulling back the canonical spin$^c$ structure $\mathfrak{c}$ from the complex surface $N$. Then $n_\mathfrak{c}(M,[\gamma_1],\ldots,[\gamma_\ell]) \neq 0$.

One way of proving this is to consider metrics on $M$ which approximate standard product metrics on each $S^1 \times S^3$, where the $S^1$ factor is taken to be extremely long. By cutting out an $S^3$ and capping off, each such metric can be approximated by a metric on $N$ containing two long cylinders $[0,b] \times S^3$ for each $S^1 \times S^3$. On the other hand, the Weitzenböck formula (14) forces $|\Phi|^2$ to fall off exponentially along such a cylinder because of the positivity of the scalar curvature. Hence one can use a cut-off function to pass from a solution of any small perturbation of the Seiberg-Witten equations on $N$ to a solution of a small perturbation of the Seiberg-Witten equations on $M$ which has any specified holonomy around the $\gamma_i$; conversely, solutions on $M$ can be pasted back onto $N$. This allows one to conclude that $n_\mathfrak{c}(M,[\gamma_1],\ldots,[\gamma_\ell]) = n_\mathfrak{c}(N) = 1$. For a different argument, see [45].

Thus we see that the 4-dimensional scalar curvature estimates obtainable by Theorem 5.1 can, in practice, actually be deduced directly from the theory of the Seiberg-Witten equations. However, the most striking consequence of Theorem 5.1 is to be found in dimensions bigger than four. Indeed, this surgical argument implies [48]

THEOREM 5.7 (Petean). Let $M^n$ be any simply connected smooth compact $n$-manifold, where $n \geq 5$. Then $\mathcal{I}(M) = 0$.

The proof builds on a circle of ideas due to Gromov and Lawson [22], using Theorem 5.1 to reduce the problem to that of finding a suitable system of generators for the spin-cobordism ring.

It follows that any simply connected $n$-manifold, $n \geq 5$, has unit volume metrics of scalar curvature $-\epsilon$, for any $\epsilon > 0$. If the manifold is also non-spin, one can even find unit-volume metrics of constant scalar curvature $> 0$ by the earlier result of Gromov-Lawson [22]. Thus, while Seiberg-Witten theory tells us that a Kähler-Einstein metric with $\lambda < 0$ maximizes the scalar curvature among constant-scalarcurvature metrics of fixed volume, the analogous assertion is dramatically false on simply connected manifolds of higher dimension. Thus, one might suspect that the sign of the Einstein constant is not determined by the smooth topology in high dimensions. In the next section, we shall see that the facts show that this suspicion is completely justified.

6. The Sign of the Einstein Constant

We have already seen that the fundamental group alone does not contain enough information to determine the sign of the Einstein constant. However, one might still hope [8] that the sign of $\lambda$ is somehow determined by the topology of $M$. Indeed, Corollary 4.3 seems to support such a hope in dimension 4. In higher dimensions, however, the theory of Kähler-Einstein manifolds allow one to actually construct counter-examples to such a conjecture [14, 31]. The first step is to observe that Corollary 4.3 becomes false if the rules are altered so as to allow one to change not only the metric, but also the differentiable structure, on a fixed topological 4-manifold.
THEOREM 6.1. There is a homeomorphic pair of 4-manifolds \((M_1, M_2)\) such that \(M_1\) admits a Kähler-Einstein metric \(g_1\) with \(\lambda < 0\), and such that \(M_2\) admits a Kähler-Einstein metric \(g_2\) with \(\lambda > 0\).

In higher dimensions, it therefore turns out that the sign of \(\lambda\) cannot be deduced from the smooth topology.

THEOREM 6.2 (Catanese-LeBrun). There is a smooth 8-manifold \(M\) which admits a pair of Einstein metrics for which the Einstein constants \(\lambda\) have opposite signs. Moreover, one may arrange for both of these Einstein metrics to be Kähler, albeit with respect to wildly unrelated complex structures.

Indeed, one may take the 4-manifold \(M_2\) to be \(\mathbb{CP}_2 \# 8 \mathbb{CP}_2\), which, as we saw in Theorem 3.3, admits Kähler-Einstein metrics with \(\lambda > 0\). On the other hand, \(M_1\) may be taken to be the underlying smooth 4-manifold of the Barlow surface. The Barlow surface [6] is a simply connected minimal complex surface of general type with the same \(b_+\) as \(\mathbb{CP}_2 \# 8 \mathbb{CP}_2\). With Barlow’s complex structure, \(M_1\) contains four \((-2)\)-curves, and so does not have \(K\) ample, but one can deform this complex structure [14] so as to destroy these \((-2)\)-curves. Thus \(M_1\) admits other complex structures for which \(K\) is ample, and so admits Kähler-Einstein metrics with \(\lambda < 0\) by Theorem 3.1. In particular, by taking the product metrics, it follows that \(M_1 \times M_1\) and \(M_2 \times M_2\) admit Kähler-Einstein metrics with Einstein constants \(\lambda\) of opposite signs.

On the other hand, the intersection forms

\[ \sim \colon H^2(\mathbb{Z}) \times H^2(\mathbb{Z}) \to \mathbb{Z} \]

of \(M_1\) and \(M_2\) are isomorphic because the Minkowski-Hasse classification [28] asserts there is only one isomorphism class when \(b_+\) and \(b_-\) are both non-zero and \(\tau = b_+ - b_- \neq 0 \mod 8\). A theorem of Wall [65] therefore shows that \(M_1\) and \(M_2\) are h-cobordant; that is, there is a 5-manifold-with-boundary \(V\) with \(\partial V = M_1 \sqcup M_2\), such that the inclusions \(M_1, M_2 \hookrightarrow W\) are both homotopy equivalences. Hence \(M_1 \times M_1\) is h-cobordant to \(M_2 \times M_2\), via \((M_1 \times W) \cup (W \times M_2)\). But Smale’s h-cobordism theorem [56] asserts that simply connected h-cobordant smooth manifolds of dimension \(\geq 5\) are necessarily diffeomorphic. Thus \(M_1 \times M_1\) is diffeomorphic to \(M_2 \times M_2\), and the Kähler-Einstein metrics under discussion may therefore be considered to live on the same manifold \(M = M_1 \times M_1\).

On the other hand, Corollary 4.3 makes it painfully obvious that \(M_1\) and \(M_2\) are not diffeomorphic — a fact which was first proved [30] using Donaldson theory [18]; cf. [44]. In other words, the h-cobordism theorem breaks down in dimension 4. However, Freedman did manage to salvage the topological part of Smale’s proof in dimension 4, and Theorem 2.1 thus allows one to still conclude that \(M_1\) and \(M_2\) are homeomorphic.

7. Weyl Estimates

So far, we have seen that the Seiberg-Witten equations give rise to scalar-curvature estimates on 4-manifolds. We will now see that that also give rise [38] to estimates concerning the Weyl curvature.

LEMMA 7.1. Let \((M, g)\) be an oriented Riemannian 4-manifold, and let \(c\) be a \(\text{spin}^c\) structure on \(M\). Let \(\tilde{g}\) be a Yamabe metric conformal to \(g\). If there is an irreducible solution \((\Phi, A)\) of the Seiberg-Witten equations (11–12) on \((M, \tilde{g}, c)\),
then the $L^2$-norms of the self-dual Weyl curvature and scalar curvature of $g$ must satisfy

$$\frac{1}{\sqrt{3}}\|W^+\|_2 + \frac{1}{2\sqrt{2}}\|s\|_2 \geq \frac{8\pi}{3} |c_1^+|.$$ 

Moreover, equality occurs if $g$ is Yamabe and also Kähler, with respect to some $c$-compatible complex structure.

**Proof.** By conformal rescaling, we may assume that the scalar curvature $s$ is a negative constant. Now consider the Weitzenböck formula

$$(d + d^*)^2 \phi = \nabla^* \nabla \phi - 2W^+(\phi, \cdot) + \frac{s}{3} \phi,$$

which holds for any self-dual 2-form $\phi$. Assuming that $\phi \neq 0$, this formula implies that

$$\|W^+\|_2 \geq \frac{|s|}{2\sqrt{2}} \|\phi\|_2 \left[ \frac{1}{\sqrt[3]{3}}\|\phi\|_2 - \frac{\sqrt{3}}{|s|}\frac{\|\nabla \phi\|_2^2}{\|\phi\|_2} \right],$$

where again we have assumed that the scalar curvature $s$ is a negative constant. We now apply this to the particular 2-form $\phi = \sigma(\Phi) = -iF^+_A$ associated with a solution of the Seiberg-Witten equations. To do so, first notice that (14) and the Cauchy-Schwarz inequality tell us that

$$\left(\|s\|_2 - \sqrt{8}\|\phi\|_2\right) \sqrt{8}\|\phi\|_2 \geq \int [(-s)|\Phi|^2 - |\Phi|^4] \, d\mu = \int 4|\nabla A \Phi|^2 d\mu,$$

since $|\Phi|^4 = 8|\phi|^2$. On the other hand, $|\nabla \phi|^2 \leq \frac{1}{2}|\Phi|^2|\nabla A \Phi|^2$, and (15) tells us that $|\Phi|^2 \leq |s|$. Since harmonic theory tells us that

$$\|\phi\|_2 \geq 2\pi|c_1^+|,$$

we therefore have

$$\frac{\|\nabla \phi\|_2^2}{|s|\|\phi\|_2^2} \leq \frac{|s|_2}{2\sqrt{2}} - 2\pi|c_1^+|.$$ 

Finally, another application of (15) gives us

$$\frac{|s|}{2\sqrt{2}\|\phi\|_2^2} \geq 1.$$

Plugging (20–22) into (19) then proves the lemma. \(\square\)

This lemma can be usefully exploited by interpreting the left-hand side as a dot product in $\mathbb{R}^2$:

$$\frac{1}{\sqrt{3}}\|W^+\|_2 + \frac{1}{2\sqrt{2}}\|s\|_2 = \left(\frac{1}{\sqrt{6}}, \sqrt{3}\right) \cdot \left(\sqrt{2}\|W^+\|_2, \frac{|s|_2}{2\sqrt{6}}\right).$$

The Cauchy-Schwarz inequality therefore tells us that

$$\left(\frac{1}{6} + 3\right) \left(2\|W^+\|_2^2 + \frac{|s|_2^2}{24}\right) \geq \left(\frac{1}{\sqrt{3}}\|W^+\|_2 + \frac{1}{2\sqrt{2}}\|s\|_2\right)^2 \geq \frac{64\pi^2}{9}(c_1^+)^2,$$

or in other words that

$$\frac{1}{4\pi^2} \int_M \left(2\|W^+\|_g^2 + \frac{s_g^2}{24}\right) \, d\mu_g \geq \frac{32}{57}(c_1^+)^2.$$
Now there is no reason to believe that the constant $\frac{32}{3\pi^2}$ is sharp, so there is little to lose if we replace it here with $\frac{5}{9}$, which is only 1% smaller, and much more easily remembered. Doing so yields

**Theorem 7.2.** Let $(M, g)$ be a compact oriented Riemannian 4-manifold with a non-trivial Seiberg-Witten invariant. Let $c_1(L) \in H^2(M, \mathbb{R})$ be the first Chern class of the corresponding spin$^c$ structure on $M$, and let $c_1^+ \neq 0$ denote its projection into the space of $g$-self-dual harmonic 2-forms. Then

$$\frac{1}{4\pi^2} \int_M \left( 2|W^+|^2_g + \frac{s_g^2}{24} \right) \, d\mu_g > \frac{5}{9} (c_1^+)^2.$$ 

This leads to yet more obstructions to the existence of Einstein metrics. Indeed, the Gauss-Bonnet formula (4) tells us that the left-hand side of the inequality in Theorem 7.2 is just $(2\chi + 3\tau)(M)$ if $g$ is Einstein. This then gives us the following improvement of Theorem 4.8:

**Theorem 7.3 (LeBrun).** Let $X$ be a minimal complex algebraic surface of general type, and let $M = X \# k\mathbb{CP}^2$ be obtained from $X$ by blowing up $k > 0$ points. If $k \geq \frac{4}{9}(2\chi + 3\tau)(X)$, then $M$ does not admit Einstein metrics.

Again, the constant of $\frac{4}{9}$ is not sharp, but suffices for our present purposes.

**Example 7.1.** Let $X_\ell$ be the Fermat surface of degree $\ell \geq 8$ in $\mathbb{CP}^3$, and let $M_\ell = X_\ell \# k\mathbb{CP}^2$ be obtained from $X_\ell$ by blowing up $k = \ell(\ell - 4)^2 - 2\binom{\ell - 1}{3} + 4$ points. Since $c_1^2(X) = \ell(\ell - 4)^2 \sim \ell^3$, whereas $k \sim \frac{2}{3}\ell^3$, we must have $k > \frac{4}{9}c_1^2(X)$ for sufficiently large $\ell$; and indeed, closer inspection shows that this actually happens for all $\ell \geq 8$. Thus Theorem 7.3 implies that none of these 4-manifolds $M_\ell$ admits an Einstein metric.

Now assume, for simplicity, that $\ell$ is odd, so that $\binom{\ell - 1}{3} \equiv 0 \mod 4$, and notice that $M_\ell$ has

$$b_+ = 2\binom{\ell - 1}{3} + 1,$$
$$b_- = 8\binom{\ell - 1}{3} + 13,$$

exactly like the surface $N_\ell$ gotten by taking the double branched cover of $\mathbb{CP}^1 \times \mathbb{CP}^1$ ramified over a smooth holomorphic curve of bidegree $(6, \binom{\ell - 1}{3} + 2)$. (The latter is an example of a Horikawa surface [27].) Since the simply connected complex surfaces $M_\ell$ and $N_\ell$ both have $\tau = -6\binom{\ell - 1}{3} + 2 \equiv 4 \mod 8$, both are non-spin, so Theorem 2.1 tells us that $M_\ell$ and $N_\ell$ are homeomorphic. But $N_\ell$ is a minimal surface of general type, and contains no $(-2)$-curves. Corollary 3.5 therefore tells us that $N_\ell$ carries an Einstein metric, even though it is homeomorphic to $M_\ell$, which does not.

Thus Theorem 7.3 gives us a simple proof of a result originally deduced by Kotschick [31], who instead applied Theorem 4.8 to some rather more exotic algebraic-geometric examples.

**Theorem 7.4 (Kotschick).** For infinitely many homeotypes of compact simply connected non-spin 4-manifolds, there are some choices of smooth structure which admit Einstein metrics, and others which do not.
Presumably this also occurs in the spin case. However, $K3$ provides the only spin homeotype where this phenomenon has been observed to date.

Notice that the holonomy-modified Seiberg-Witten invariants of Theorem 5.6 also give rise to obstructions to the existence of Kähler-Einstein metrics. For example, one has

**Theorem 7.5.** Let $X$ be a minimal surface of general type. Then $M = X \# k \overline{CP}_2 \# \ell [S^1 \times S^2]$ does not admit Einstein metrics if $k + 4\ell \geq \frac{3}{2}(2\chi + 3\tau)(X)$.

The proof imitates that of Theorem 4.7, but uses Theorem 7.2 in place of Theorem 4.2. Details are left to the reader.

## 8. Minimal Volumes

If $M$ is a compact $n$-manifold, multiplying any given metric on $M$ by a large enough positive constant will yield a new metric on $M$ of sectional curvature $> -1$. This rescaling process, however, will also typically make the volume of $M$ enormous. Gromov [21] thus realized that it is natural to define a a diffeomorphism invariant, called the *minimal volume*, by setting

$$\text{Vol}_K(M) = \inf\{\text{Vol}(M, g) \mid g \text{ has } K \geq -1\}.$$  

But it is equally natural to consider minimal volumes with respect to lower bounds on the Ricci or scalar curvatures:

$$\text{Vol}_r(M^n) = \inf\{\text{Vol}(M, g) \mid g \text{ satisfies } r \geq -(n-1)g\}$$

$$\text{Vol}_s(M^n) = \inf\{\text{Vol}(M, g) \mid g \text{ has } s \geq -(n-1)g\}.$$  

Notice that our conventions have been chosen so that

$$\text{Vol}_K(M) \geq \text{Vol}_r(M^n) \geq \text{Vol}_s(M^n)$$

tautologically.

For any manifold of dimension $n \geq 3$, one can show, by first considering one conformal class at a time, that the minimal volume for $s$ is given by

$$\text{Vol}_s(M^n) = \frac{\mathcal{I}(M)}{n^2(n-1)^2},$$

where the invariant $\mathcal{I}$ was defined in §5. Inspection of the Gauss-Bonnet formula (4) therefore shows that an oriented 4-manifold $M$ can admit an Einstein metric $g$ only if

$$2\chi(M) - 3|\tau(M)| \geq \frac{1}{96\pi^2}\mathcal{I}(M) = \frac{3}{2\pi^2}\text{Vol}_s(M),$$

with equality iff $g$ is half-conformally flat and $\text{Vol}_s(M)$ is realized by a suitable rescaling of $g$. Much of what we have done so far simply consists of making this inequality effective by introducing non-trivial estimates for $\text{Vol}_s(M)$.

In a sense, however, this inequality is quite wasteful; after all, if $g$ is an Einstein metric, its Ricci curvature is determined by its scalar curvature. Thus, the same argument actually proves the following:

**Lemma 8.1.** Let $(M, g)$ be a 4-dimensional Einstein manifold. Then

$$2\chi(M) \geq 3|\tau(M)| + \frac{3}{2\pi^2}\text{Vol}_r(M),$$

with equality iff $g$ is half-conformally flat and can be rescaled so as to realize the minimal Ricci volume.
Of course, such an inequality only acquires content in conjunction with an effective method for estimating the invariant \( \text{Vol}_r(M) \). The first result in this direction was discovered by Gromov [21], and involves an invariant \( \|M\| \) of a compact topological \( n \)-manifold known as its \textit{simplicial volume}, and defined as the infimum of expressions of the form \( \sum |c_j| \), where \( \sum c_j \sigma_j \) is any singular homology cycle with \textit{real coefficients} \( c_j \) representing the fundamental cycle \( [M] \in H_n(M, \mathbb{R}) \).

\textbf{Proposition 8.2 (Gromov).} For every smooth compact \( n \)-manifold \( M \),

\[ \text{Vol}_r(M) > \frac{1}{(n-1)^{n-1}n!} \|M\|. \]

We will say a bit about the proof of this result in the next section. For the moment, let us merely notice that, with Lemma 8.1, it immediately implies

\textbf{Theorem 8.3 (Gromov/Kotschick).} Let \( (M, g) \) be a 4-dimensional Einstein manifold. Then

\[ 2\chi(M) \geq 3|\tau(M)| + \frac{\|M\|}{1296\pi^2}. \]

Curiously, Gromov only derived the weaker inequality obtained from this by dropping the \( \tau \) term. The fact that Gromov’s results actually imply an improved version of the Hitchin-Thorpe inequality was only recently brought to light by Kotschick [32]. Notice that, in contrast to results derived by Seiberg-Witten methods, the Gromov/Kotschick inequality only involves terms depending on the homotopy type of \( M \). However, the simplicial volume \( \|M\| \) turns out to vanish for any simply connected manifold, so the inequality only improves upon the Hitchin-Thorpe inequality in cases where the fundamental group is infinite.

On the other hand, Theorem 8.3 does represent an honest improvement over the Hitchin-Thorpe inequality. For example, let \( X \) is a hyperbolic 4-manifold, and recall that Mal’tsev’s Theorem [69] predicts that there are \( \ell \)-fold covers \( X_\ell \) of \( X \) for arbitrarily large \( \ell \). If \( M = X_\ell \# m\mathbb{C}P^2 \), then \( \|M\| \geq \ell\|X\| \), and \( \|X\| \) in turn is positive — in fact, \( \|X\| = \frac{4\pi^2}{3v_4} \chi(X) \), where \( v_4 \) is the volume of a regular ideal hyperbolic 4-simplex. Since the Gromov-Kotschick inequality requires that

\[ \left(2 - \frac{1}{972v_4}\right) \chi(X) > \frac{m}{\ell}, \]

whereas the Hitchin-Thorpe inequality would merely stipulate that

\[ 2\chi(X) > \frac{m}{\ell}, \]

Theorem 8.3 actually predicts non-existence in a certain range of \( m \) missed by Hitchin-Thorpe, provided that \( \ell \) is sufficiently large. But in the next section, we will see that one can do a great deal better: \textit{none} of these manifolds admits an Einstein metric!

Even without the signature term, Gromov was able to predict non-existence in cases missed by the Hitchin-Thorpe inequality by considering examples of the form \( M = 2(\Sigma \times \Sigma) \# k[S^1 \times S^3] \), where \( \Sigma \) is a Riemann surface of large genus. It is for this reason that simple connectivity was emphasized in Corollary 4.9.

\section{Entropy and Ricci Curvature}

Let \( (M, g) \) be a compact Riemannian manifold, and let \( (\hat{M}, g) \) be its universal cover. Let \( x \in \hat{M} \), and let \( B_g(x) \subset \hat{M} \) denote the open distance ball, consisting of
of points of distance $< \rho$ from $x$; let $\text{Vol}(B_\rho(x))$ denote the Riemannian volume of this distance-ball. Then the \textit{volume entropy} of $(M, g)$ is defined to be
\[
\hat{h}_{vol}(M, g) = \lim_{\rho \to \infty} \frac{\log \text{Vol}(B_\rho(x))}{\rho}.
\]
This is independent of the base-point $x$, but of course can be non-zero only if the fundamental group of $M$ is infinite.

An easy calculation shows that an $n$-manifold of constant sectional curvature $K \leq 0$ has entropy $\hat{h}_{vol} = (n-1)\sqrt{|K|}$. After a bit of of pure thought, we therefore get the following:

\textbf{Lemma 9.1.} \textit{Any compact Riemannian manifold $(M, g)$ with $r \geq -(n-1)g$ has volume entropy}
\[
\hat{h}_{vol}(M, g) \leq n - 1.
\]

Indeed, this is an immediate consequence of Bishop’s inequality [12, 8], which, in light of our assumption that the Ricci curvature of $(\tilde{M}, g)$ is no smaller than that of hyperbolic space $\mathcal{H}^n$, says a ball of radius $\rho$ in $(\tilde{M}, g)$ must have volume no bigger than that of the corresponding ball in $\mathcal{H}^n$.

Of course, the entropy $\hat{h}_{vol}(M, g)$ is not invariant under rescalings, and indeed it is easy to show that
\[
\hat{h}_{vol}(M, cg) = c^{-1/2} \hat{h}_{vol}(M, g).
\]
Fortunately, this is easily remedied by instead considering the scale-invariant quantity
\[
\mathcal{E}(M^n, g) = [\hat{h}_{vol}(M, g)]^n \text{Vol}(M, g).
\]
This invariant was already considered by Gromov [21], who showed that any metric on any compact $n$-manifold $M$ satisfies
\[
\mathcal{E}(M, g) \geq \frac{1}{n!} ||M||.
\]
With Lemma 9.1, this then implies Proposition 8.2.

While Gromov’s lower bound on $\mathcal{E}(M, g)$ opened up several new frontiers of mathematical research, it is, in practice, far from sharp. It was therefore a development of the greatest significance when Besson, Courtois, and Gallot [9] were able to prove that locally symmetric metrics of strictly negative curvature actually minimize this functional:

\textbf{Theorem 9.2 (Besson-Courtois-Gallot).} \textit{Let $M$ be any compact quotient of a real, complex, quaternionic, or octonionic hyperbolic space, and let $g_0$ be the standard metric on $M$. Then any other metric $g$ on $M$ satisfies}
\[
\mathcal{E}(M, g) \geq \mathcal{E}(M, g_0),
\]
\textit{with equality iff $g$ is locally symmetric.}

\textbf{Sketch of Proof.} Let $S^\infty$ denote the unit sphere in the real Hilbert space $L^2(\partial M)$ of square-integrable half-densities on the sphere-at-infinity of $M$, and let $S^\infty_+ \subset S^\infty$ denote its intersection with the open cone of \textit{positive} half-densities. We will consider smooth $\pi_1(M)$-equivariant maps $\Phi : M \to S^\infty_+$. Each such map induces a (possibly degenerate) metric $g_\Phi$ on $\tilde{M}$ which is $\pi_1(M)$-invariant, and so
descends to \( M \). The volume \( \text{Vol}(M, g_\Phi) \) may then be viewed as a sort of equivariant volume of the image of \( \Phi \), and it is natural to ask whether some choice of \( \Phi \) minimizes this volume.

Such a minimizer actually does exist. Indeed, for each \( x \in \hat{M} \), consider the map from the unit sphere in \( T_x \hat{M} \) to \( \partial \hat{M} \) obtained by following the geodesics of \( g_0 \) all the way to infinity. The Poisson measure \( d\varphi_x \) on \( \partial \hat{M} \) is the push-forward of the usual probability measure on the unit sphere in \( T_x \hat{M} \) via this map. The promised minimizer is then given by \( \Phi_0(x) = \sqrt{\varphi_x} \). The fact that this is a minimizer is proved by a calibrated geometry argument. Namely, there is a canonical baricenter map \( \varpi : S^\infty_+ \to \hat{M} \), and one may therefore define a closed \( n \)-form \( \omega \) on \( S^\infty_+ \) by

\[
\omega = \left( \frac{h_{\text{vol}}(g_0)}{2\sqrt{n}} \right)^n \varpi^*d\mu_{g_0},
\]

where \( d\mu_{g_0} \) is (traditional bad notation for) the volume \( n \)-form of \( g_0 \). The integral of \( \omega \) on any \( n \)-manifold is then less than or equal to the manifold’s volume, and equality holds for the image of \( \Phi_0 \). Since any other \( \Phi \) is smoothly, equivariantly homotopic to \( \Phi_0 \), Stokes’ theorem tells us that

\[
\text{Vol}(M, g_\Phi) \geq \int_M \Phi^*\omega = \int_M \Phi_0^*\omega = \text{Vol}(M, g_{\Phi_0}) = (4n)^{-n/2} \mathcal{E}(M, g_0).
\]

On the other hand, given any metric \( g \) on \( M \) and any constant \( c > h_{\text{vol}}(g) \), one may define a smooth equivariant map by

\[
\Phi_{g, c}(x) = \left[ \frac{\int_{S^\infty_+} e^{-c\delta_g(x, y)} d\varphi_y d\mu_{g, g}}{\int_{S^\infty_+} \int_M e^{-c\delta_g(x, y)} d\varphi_y d\mu_{g, g}} \right]^{1/2}
\]

where \( \delta_g \) is the Riemannian distance. One is then able to show that

\[
\frac{c^2}{4n} g \geq g_{\Phi_{g, c}},
\]

so that

\[
\mathcal{E}(M, g) \geq (4n)^{n/2} \inf_c \text{Vol}(M, g_{\Phi_{g, c}}).
\]

Thus \( \mathcal{E}(M, g) \geq \mathcal{E}(M, g_0) \), as claimed. \( \Box \)

In light of Lemma 9.1, Theorem 9.2 implies

**Corollary 9.3.** Let \( M \) be any compact hyperbolic 4-manifold. Then

\[
\text{Vol}_r(M) = \frac{4\pi^2}{3} \chi(M).
\]

Of course, we also get similar results in other dimensions; for example, if \( M^{2m} \) is any even-dimensional hyperbolic manifold, \( \text{Vol}_r(M) = (-4\pi)^m \frac{m!}{(2m)!} \chi(M) \). In high dimensions, alas, this tells us essentially nothing about Einstein metrics. But in dimension 4, we find that any Einstein metric on a hyperbolic 4-manifold saturates the inequality

\[
2\chi \geq 3|\tau| + \frac{3}{2\pi^2} \text{Vol}_r
\]

of Lemma 8.1, is therefore conformally flat, and hence has constant curvature. With Mostow rigidity, this yields:
Theorem 9.4 (Besson-Courtois-Gallot). Let $M^4$ be a smooth compact quotient of hyperbolic 4-space $H^4 = SO(4,1)/SO(4)$, and let $g_0$ be its standard metric of constant sectional curvature. Then every Einstein metric $g$ on $M$ is of the form $g = \lambda \varphi^* g_0$, where $\varphi : M \to M$ is a diffeomorphism and $\lambda > 0$ is a constant.

Notice that Theorem 9.2 also makes an assertion about complex-hyperbolic 4-manifolds, so one might expect to also be able to prove Theorem 4.6 by this method. Unfortunately, however, because Lemma 9.1 uses the Bishop comparison theorem to compare metrics with real hyperbolic space, the associated lower bound for $\text{Vol}_r$ turns out to be too small by a factor of $\frac{64}{81}$. It would of course be of fundamental interest to find some over-arching point of view which could explain both results simultaneously.

By imitating the proof of Theorem 9.2, one can also prove the following:

Theorem 9.5 (Besson-Courtois-Gallot). Let $(X, g_0)$ be a compact oriented locally symmetric space of negative curvature, and let $M$ be a compact manifold of the same dimension. Let $f : M \to X$ be any smooth map. Then any metric $g$ on $M$ satisfies

$$\mathcal{E}(M, g) \geq |\deg(f)|\mathcal{E}(X, g_0),$$

where $\deg(f)$ denotes the degree of $f$.

Corollary 9.6. Let $X$ be a compact oriented hyperbolic 4-manifold, and suppose $M$ is a compact oriented 4-manifold which admits a map $f : M \to X$ of degree $q$. Then

$$\text{Vol}_r(M) \geq q \frac{4\pi^2}{3} \chi(X).$$

Example 9.1. Let $X$ be a compact oriented hyperbolic 4-manifold, and let

$$M = X \# \ell(S^1 \times S^3) \# m\mathbb{C}P_2.$$ 

Then $M$ admits a degree-1 map to $X$, and hence

$$\text{Vol}_r(M) \geq \frac{4\pi^2}{3} \chi(X).$$

On the other hand, we have

$$\chi(M) = \chi(X) - 2\ell + m,$$

$$\tau(M) = m.$$ 

Plugging these numbers into the inequality

$$2\chi \geq 3\tau + \frac{3}{2\pi^2} \text{Vol}_r$$

of Lemma 8.1, we conclude that $M$ can admit an Einstein metric only if

$$2\chi(X) - 4\ell + 2m \geq 3m + 2\chi(X).$$

If either $\ell$ or $m$ is positive, such a manifold therefore never admits an Einstein metric. Moreover, since our estimate of $\text{Vol}_r$ only depends on the existence of a map of a degree-1 map to $X$, the same conclusion applies to any 4-manifold which is even homotopy equivalent to one of these examples.
Now for any compact oriented hyperbolic 4-manifold \(X\), the Euler characteristic \(\chi(X)\) is even and positive, and we may arrange for it to be as large as we like by passing to finite covers — which exist in abundance by Mal'ctev’s theorem [69]. By choosing \(X\) and \(\ell\) appropriately, we may, for any \(m \geq 0\), therefore construct manifolds \(M\) as above such that \(\chi(M)\) is any given integer \(\equiv m \mod 2\). Since one also has \(\tau(M) = m\), this family of manifolds \(M\), together with their orientation reversed versions \(\overline{M}\), suffices to prove a stronger version of Corollary 5.4:

**Theorem 9.7 (Sambusetti).** Any integer pair \((\chi, \tau)\) with \(\tau \equiv \chi \mod 2\) can be realized as the Euler characteristic and signature of a smooth compact oriented 4-manifold \(M\) (with infinite fundamental group) which is not homotopy equivalent to any 4-dimensional Einstein manifold.

Once again, notice how these entropy arguments involve the gigantic size of the fundamental group in an essential way. These beautiful results can therefore shed no light at all on the simply connected case.

### 10. The Positive Case

This essay has focused almost exclusively on recent results concerning 4-dimensional Einstein manifolds with \(\lambda < 0\). However, recent years have also witnessed remarkable progress in our knowledge of the \(\lambda > 0\) case. The most striking result in this direction is the weak compactness theorem of Anderson [2], which shows that, while the moduli space of \(\lambda > 0\) Einstein metrics on a 4-manifold \(M\) is generally non-compact, it can always be compactified by adding points representing orbifold Einstein metrics on spaces obtained from \(M\) by collapsing chains of 2-spheres. For a description of results in this direction, see the essay by Petersen in this volume.

In the negative case, we have seen that the presence of a non-vanishing Seiberg-Witten invariant is enough to guarantee that any Riemannian metric on a manifold satisfies \(\int_M s^2 d\mu \geq 32\pi^2 (2\chi + 2\tau)\). For an Einstein metric, one may use (4) to rewrite this in the form

\[
\int_M \frac{s^2}{24} d\mu \geq \int_M |W^+|^2 d\mu.
\]

It may therefore come as something of a surprise to learn that Gursky [23, 24] has proved that any Einstein 4-manifold with \(\lambda > 0\) satisfies exactly the opposite inequality, unless it is anti-self-dual:

**Theorem 10.1 (Gursky).** Let \((M, g)\) be a compact oriented Einstein 4-manifold with \(s > 0\) and \(W^+ \neq 0\). Then

\[
\int_M |W^+_g|^2 d\mu_g \geq \int_M \frac{s_g^2}{24} d\mu_g,
\]

with equality iff \(\nabla W^+ \equiv 0\).

**Sketch of Proof.** The Einstein equations and the second Bianchi identity imply that \(W^+\) is divergence-free. This implies that

\[
0 \geq \Box |W^+|^{1/3},
\]

where \(\Box = 6\Delta + s - 2\sqrt{6}|W^+|\). However, the quantity \(\mathcal{G} = s - 2\sqrt{6}|W^+|\) transforms under conformal transformations according to the rule

\[
\mathcal{G} u^2 g = u^{-3} \mathcal{G} \cdot u.
\]
Thus, assuming that $W^+ \neq 0$, we may take $u$ to be a smooth positive approximation of $|W^+|^{1/3}$, and thereby construct a conformally rescaled metric $\tilde{g} = u^2 g$ such that

$$\int_M \left[ s_{\tilde{g}} - 2\sqrt{6}|W^+_{\tilde{g}}| \right] d\mu_{\tilde{g}} \leq 0.$$

Now $g$ minimizes $\int s \, d\mu$ among all metrics of fixed volume in its conformal class, so that

$$\left( \int_M s_{\tilde{g}}^2 d\mu_{\tilde{g}} \right)^{1/2} = \frac{\int_M s_{\tilde{g}} d\mu_{\tilde{g}}}{\sqrt{\int_M d\mu_{\tilde{g}}}} \leq \frac{\int_M s_{\tilde{g}} d\mu_{\tilde{g}}}{\sqrt{\int_M d\mu_{\tilde{g}}}}$$

$$\leq \frac{\int_M 2\sqrt{6}|W^+_{\tilde{g}}| d\mu_{\tilde{g}}}{\sqrt{\int_M d\mu_{\tilde{g}}}}$$

$$\leq \left( 24 \int |W^+_g|^2 d\mu_g \right)^{1/2}$$

$$= \left( 24 \int |W^+_g|^2 d\mu_g \right)^{1/2},$$

by the positivity of $s_{\tilde{g}}$ and the conformal invariance of $|W^+|^2 d\mu$.

The resulting estimate $\int s^2 d\mu \leq 32\pi^2(2\chi + 3\tau)$ is certainly interesting, but not particularly powerful in the absence of other geometric assumptions. If one assumes, however, that the sectional curvature of $g$ is non-negative, one also has the point-wise estimate

$$\frac{s}{\sqrt{6}} \geq |W^+| + |W^-|,$$

and the Gauss-Bonnet formula therefore tells one that $\chi < \frac{5}{8\pi^2} \int_M s^2 d\mu$. Putting these two inequalities together, reversing the orientation if necessary, and using Bishop’s inequality, one thus obtains

**Proposition 10.2.** Let $(M, g)$ be a smooth compact oriented Einstein 4-manifold with non-negative sectional curvature. Assume, moreover, that $g$ is neither self-dual nor anti-self-dual. Then the Euler characteristic $\chi$ and the signature $\tau$ of $M$ satisfy

$$9 \geq \chi > \frac{15}{4} |\tau|.$$ 

In particular, if a 4-manifold $M$ has $b_+ = 0$ and $b_- \neq 0$, any Einstein metric of non-negative sectional curvature on $M$ must be self-dual. Since Hitchin has proved [8] that the only self-dual Einstein manifolds with positive scalar curvature are the symmetric spaces $S^4$ and $\mathbb{CP}_2$, this gives us a clean characterization of the Fubini-Study metric [24]:

**Theorem 10.3** (Gursky-LeBrun). Let $M$ be a smooth compact oriented 4-manifold with strictly positive intersection form. Suppose that $g$ is an Einstein metric on $M$ which has non-negative sectional curvature. Then $(M, g)$ is isometric to $\mathbb{CP}_2$, equipped with a constant multiple of its standard Fubini-Study metric.
One might thus hope that any Einstein 4-manifold with non-negative sectional curvature is actually locally symmetric. One piece of evidence in favor of such a conjecture is the fact [57] that any Einstein manifold of positive curvature operator $\mathcal{R} : \Lambda^2 \rightarrow \Lambda^2$ is locally symmetric. For a further discussion of the (extremely strong) condition $\mathcal{R} \geq 0$, see the essay by Chow in this volume.

In any case, there are, up to diffeomorphism, only finitely many compact 4-manifolds with Einstein metrics of non-negative sectional curvature. The flat 4-manifolds, of course, nominally form a subclass of the the manifolds under discussion, but Bieberbach’s theorem [10] in any case tells us that there are finitely many diffeomorphism types of these. For the others, the Ricci curvature must be positive, and we may thus rescale the metric so that, for example, $r = 3g$. The definition of the Ricci curvature then tells us that the sectional curvatures all satisfy $0 \leq K(P) \leq 3$. Gauss-Bonnet therefore tells us that the volume is $\geq 8\pi^2 / 15$. On the other hand, Myers’ theorem predicts that the diameter is $\leq \pi$. Given such bounds, Cheeger’s finiteness theorem [15] then predicts that there are only finitely many diffeomorphism types — although, of course, the actual number could still be astronomical. By contrast, Proposition 10.2 and Freedman’s classification [19] tell us that there are at most twelve homeotypes of simply connected compact Einstein 4-manifolds with non-negative sectional curvature.

**Remark 10.1.** If an Einstein manifold instead has non-positive sectional curvature, one still has the inequality

$$\frac{|s|}{\sqrt{6}} \geq |W^+| + |W^-|,$$

and it is straightforward to show that consequently

$$\chi > \frac{15}{8}|\tau|.$$

This is actually a minor improvement on a result of Hitchin [26], who observed that such an inequality holds for the somewhat smaller coefficient of $(\frac{5}{2})^{3/2}$. In all likelihood, however, the present constant of $\frac{15}{8}$ isn’t sharp, either. In any case, it would be extremely interesting to construct some non-locally-symmetric examples, and give this discussion some substance!

**11. Concluding Remarks**

In this essay, we have explored several recent streams of thought which bear upon the existence and uniqueness of Einstein metrics on 4-dimensional manifolds. For example, we have seen that Seiberg-Witten theory gives one control of the $L^2$-norms of scalar and Weyl curvature when certain diffeomorphism invariants are non-zero. Entropy estimates instead allow one to control the Ricci curvature under certain homotopy-theoretic assumptions. The mystery is that, while these techniques sometimes lead to analogous results, they seem completely unrelated. One might hope for a deeper, unified explanation of these results involving principles which remain to be discovered.

On the other hand, it could be that the striking parallels between these two sets of results are merely ephemeral. For example, the parallel formulations of Theorem 9.4 and Corollary 4.6 hide an important technical distinction. The proof of Theorem 9.4 actually shows that any Einstein 4-manifold which is homotopy equivalent to a hyperbolic manifold must itself be hyperbolic, whereas the proof of
Corollary 4.6 yields no such conclusion in the complex-hyperbolic case. Does this merely illustrate a limitation of the methods of proof, or does it capture a factual difference between the real- and complex-hyperbolic cases?

In the same vein, it is interesting to compare the information that these very different sets of techniques provide concerning blow-ups of complex-hyperbolic manifolds. Theorem 7.3 tells us that blowing up such a space at, say, $4\tau$ points will result in a smooth manifold without Einstein metrics. Theorem 9.5 is less efficient, but it does reach a similar conclusion [50] if something over $6\tau$ points are blown up. However, the entropy argument yields non-existence for every smooth structure on the manifold.

While we have described a number of techniques for showing that Einstein metrics do not exist on certain 4-manifolds, a direct variational approach to the existence problem [3] might suggest that one should instead try to construct sequences of metrics on a given 4-manifold which geometrically converge to a disjoint union of Einstein pieces. For example [36], while complex surfaces of general type do not generally admit Einstein metrics, they do always admit minimizing sequences for the functional $\int s^2 d\mu$ which converge to orbifold Einstein metrics on their pluricanonical models, at the price of ‘blopping off’ some topology. However, there are circumstances [3, 39] in which such minimizing sequences instead ‘collapse’ to a lower-dimensional object. At any rate, while most 4-manifolds do not admit Einstein metrics, one might still hope that unions of special Einstein manifolds will eventually play a rôle in 4-dimensional smooth topology similar to that played by minimal models in complex surface theory.

In a different direction, we have seen that the sign of the Einstein constant is definitely not a diffeomorphism invariant in high dimensions. On the other hand, we have seen some weak indications that just the opposite may hold in dimension 4. Further exploration of this issue would seem to be one of the most compelling potential directions for future research.

Finally, it is worth comparing the general state of our knowledge concerning the positive and negative cases. For example, we now know that there aren’t any non-standard Einstein metrics on compact quotients of $H^4$ or $\mathbb{C}H_2$. What about non-standard Einstein metrics on $S^4$ or $CP^2$? The question seems fair enough. Yet the only results currently available in this direction pertain to metrics of positive sectional curvature. The need for such an extraneous hypothesis should serve as a clear indication of the depth of our present ignorance.

References


[63] G. Tian and S. Yau, Kähler-Einstein metrics on complex surfaces with c1 > 0, Communications in Mathematical Physics, 112 (1987), pp. 175–203.

Department of Mathematics, SUNY at Stony Brook, Stony Brook, NY 11794-3651
E-mail address: claude@math.sunysb.edu