In this article we discuss the work of Karen Uhlenbeck, mainly from the 1980s, focused on variational problems in differential geometry.

The calculus of variations goes back to the 18th century. In the simplest setting we have a functional

\[ F(u) = \int \Phi(u, u') \, dx, \]

defined on functions \( u \) of one variable \( x \). Then the condition that \( F \) is stationary with respect to compactly supported variations of \( u \) is a second order differential equation—the Euler–Lagrange equation associated to the functional. One writes

\[ \delta F = \int \delta u \, \tau(u) \, dx, \]

where

\[ \tau(u) = \frac{\partial \Phi}{\partial u} - \frac{d}{dx} \frac{\partial \Phi}{\partial u'}. \]  

(1)

The Euler–Lagrange equation is \( \tau(u) = 0 \). Similarly for vector-valued functions of a variable \( x \in \mathbb{R}^n \). Depending on the context, the functions would be required to satisfy suitable boundary conditions or, as in most of this article, might be defined on a compact manifold rather than a domain in \( \mathbb{R}^n \), and \( u \) might not exactly be a function but...
We take

\[ \mathcal{F}(\phi(v)) \]

\[ \Gamma_{jk} \]

\[ y'' - \sum_{j,k} \Gamma_{jk} y_j y_k = 0, \]

with \( y(0) = p, y(1) = q \), and the energy functional

\[ \mathcal{F}(\gamma) = \int_0^1 |\nabla \gamma|^2, \]

where the norm of the “velocity vector” \( \nabla \gamma \) is computed using the Riemannian metric on \( N \). The Euler–Lagrange equation is the geodesic equation, in local co-ordinates,

\[ y'' - \sum_{j,k} \Gamma_{jk} y_j y_k = 0, \]

where the “Christoffel symbols” \( \Gamma_{jk} \) are given by well-known formulae in terms of the metric tensor and its derivatives. In this case the variational picture works as well as one could possibly wish. There is a geodesic from \( p \) to \( q \) minimising the energy. More generally one can use minimax arguments and (at least if \( p \) and \( q \) are taken in general position) the Morse theory asserts that the homology of the path space \( \mathcal{X} \) can be computed from a chain complex with generators corresponding to the geodesics from \( p \) to \( q \). This can be used in both directions: facts from algebraic topology about the homology of the path space give existence results for geodesics, and, conversely, knowledge of the geodesics can feed into algebraic topology, as in Bott’s proof of his periodicity theorem.

The existence of a minimising geodesic between two points can be proved in an elementary way and the original approach of Morse avoided the infinite dimensional path space \( \mathcal{X} \), working instead with finite dimensional approximations, but the infinite-dimensional picture gives the best starting point for the discussion to follow. The basic point is a compactness property: any sequence \( y_1, y_2, \ldots \) in \( \mathcal{X} \) with bounded energy has a subsequence which converges in \( C^0 \) to some continuous path from \( p \) to \( q \). In fact for a path \( y \) in \( \mathcal{X} \) and \( 0 \leq t_1 < t_2 \leq 1 \) we have

\[ d(y(t_1), y(t_2)) \leq \int_{t_1}^{t_2} |\nabla y| \leq \mathcal{F}(y)^{1/2} |t_1 - t_2|^{1/2}, \]

where the last step uses the Cauchy–Schwartz inequality. Thus a bound on the energy gives a \( 1/2 \)-Hölder bound on \( y \) and the compactness property follows from the Ascoli–Arzelà theorem.

In the same vein as the compactness principle, one can extend the energy functional \( \mathcal{F} \) to a completion \( \mathcal{X} \) of \( \mathcal{X} \) which is an infinite dimensional Hilbert manifold, and elements of \( \mathcal{X} \) are still continuous (in fact \( 1/2 \)-Hölder continuous) paths in \( N \). In this abstract setting, Palais and Smale introduced a general “Condition C” for functionals on Hilbert manifolds, which yields a straightforward variational theory. (This was extended to Banach manifolds in early work of Uhlenbeck [24].) The drawback is that, beyond the geodesic equations, most problems of interest in differential geometry do not satisfy this Palais–Smale condition, as illustrated by the case of harmonic maps.

**Harmonic Maps in Dimension 2**

We begin in dimension 1 where geodesics in a Riemannian manifold are classical examples of solutions to a variational problem. Here we take \( N \) to be a compact, connected, Riemannian manifold and fix two points \( p, q \) in \( N \). We take \( X \) to be the space of smooth paths \( y : [0, 1] \to N \)
The harmonic map equations were first studied systematically by Eells and Sampson [5]. We now take \( M, N \) to be a pair of Riemannian manifolds (say compact) and \( X = \text{Maps}(M, N) \) the space of smooth maps. The energy of a map \( u : M \to N \) is given by the same formula

\[
\mathcal{F}(u) = \int_M |\nabla u|^2,
\]

where at each point \( x \in M \) the quantity \( |\nabla u| \) is the standard norm defined by the metrics on \( TM_x \) and \( TN_{u(x)} \). In local co-ordinates the Euler Lagrange equations have the form

\[
\Delta_M u_i - \sum_{jk} \Gamma^j_{ik} \nabla_j u_k = 0,
\]

(2)

where \( \Delta_M \) is the Laplacian on \( M \). This is a quasi-linear elliptic system, with a nonlinear term which is quadratic in first derivatives. The equation is the natural common generalisation of the geodesic equation in \( N \) and the linear Laplace equation on \( M \).

The key point now is that when \( \dim M > 1 \) the energy functional does \emph{not} have the same compactness property. This is bound up with Sobolev inequalities and, most fundamentally, with the \emph{scaling behaviour} of the functional. To explain, in part, the latter consider varying the metric \( g_M \) on \( M \) by a conformal factor \( \lambda \). So \( \lambda \) is a strictly positive function on \( M \) and we have a new metric \( \tilde{g}_M = \lambda^2 g_M \). Then one finds that the energy \( \tilde{F} \) defined by this new metric is

\[
\tilde{F}(u) = \int_M \lambda^{2-n} |\nabla u|^2,
\]

where \( n = \dim M \). In particular if \( n = 2 \) we have \( \tilde{F} = F \).

Now take \( M = S^2 \) with its standard round metric and \( \phi : S^2 \to S^2 \) a Möbius map. This is a conformal map and it follows from the above that for any \( u : S^2 \to N \) we have \( \tilde{F}(u \circ \phi) = F(u) \). Since the space of Möbius maps is not compact we can construct a sequence of maps \( u \circ \phi_i \) with the same energy but with no convergent subsequence.

We now recall the Sobolev inequalities. Let \( f \) be a smooth real valued function on \( \mathbb{R}^n \), supported in the unit ball. We take polar co-ordinates \((r, \theta)\) in \( \mathbb{R}^n \), with \( \theta \in S^{n-1} \). For any fixed \( \theta \) we have

\[
f(0) = \int_{r=0}^1 \frac{\partial f}{\partial r} dr.
\]

So, integrating over the sphere,

\[
f(0) = \frac{1}{\omega_n} \int_{S^{n-1}} \int_{r=0}^1 \frac{\partial f}{\partial r} dr d\theta,
\]

where \( \omega_n \) is the volume of \( S^{n-1} \). Since the Euclidean volume form is \( d^n x = r^{n-1} dr d\theta \) we can write this as

\[
f(0) = \frac{1}{\omega_n} \int_{B^n} |x|^{1-n} \frac{\partial f}{\partial r} d^n x.
\]

The function \( x \mapsto |x|^{1-n} \) is in \( L^q \) over the ball for any \( q < n/n-1 \). Let \( p \) be the conjugate exponent, with \( p^{-1} + q^{-1} = 1 \), so \( p > n \). Then Hölder’s inequality gives

\[
|f(0)| \leq C_p \|f\|_{L^p}
\]

where \( C_p \) is \( \omega_n \) times the \( L^q(B^n) \) norm of \( x \mapsto |x|^{1-n} \).

The upshot is that for \( p > n \) there is a continuous embedding of the Sobolev space \( L^p \)—obtained by completing in the norm \( \|\nabla f\|_{L^p} \)—into the continuous functions on the ball. In a similar fashion, if \( p < n \) there is a continuous embedding \( L^p \to L^r \) for the exponent range \( r \leq np/(n-p) \), which is bound up with the isoperimetric inequality in \( \mathbb{R}^n \). The arithmetic relating the exponents and the dimension \( n \) reflects the scaling behaviour of the norms. If we define \( f_\mu(x) = f(\mu x) \), for \( \mu \geq 1 \), then

\[
\|f_\mu\|_{C^0} = \|f\|_{C^0},
\]

\[
\|f_\mu\|_{L^q} = \mu^{-n/q} \|f\|_{L^q},
\]

\[
\|f_\mu\|_{L^r} = \mu^{1-n/r} \|f\|_{L^r}.
\]

It follows immediately that there can be no continuous embedding \( L^p \to C^0 \) for \( p < n \) or \( L^p \to L^r \) for \( r > np/(n-p) \).

The salient part of this discussion for the harmonic map theory is that the embedding \( L^p \to C^0 \) fails at the critical exponent \( p = n \). (To see this, consider the function \( \log \log r^{-1} \).) Taking \( n = 2 \) this means that the energy of a map from a 2-manifold does not control the continuity of the map and the whole picture in the 1-dimensional case breaks down. This was the fundamental difficulty addressed in the landmark paper [12] of Sacks and Uhlenbeck which showed that, with a deeper analysis, variational arguments can still be used to give general existence results.

Rather than working directly with minimising sequences, Sacks and Uhlenbeck introduced perturbed functionals on \( X = \text{Maps}(M, N) \) (with \( M \) a compact 2-manifold):

\[
\mathcal{F}_\alpha(u) = \int_M (1 + |\nabla u|^2)^\alpha.
\]

For \( \alpha > 1 \) we are in the good Sobolev range, just as in the geodesic problem. Fix a connected component \( X_0 \) of \( X \) (i.e. a homotopy class of maps from \( M \) to \( N \)). For \( \alpha > 1 \) there is a smooth map \( u_\alpha \) realising the minimum of \( \mathcal{F}_\alpha \) on \( X_0 \). This map \( u_\alpha \) satisfies the corresponding Euler-Lagrange equation, which is an elliptic PDE given by a vari-ant of (2). The strategy is to study the convergence of \( u_\alpha \) as \( \alpha \) tends to 1. The main result can be outlined as follows. To simplify notation, we understand that \( \alpha \) runs over a suitable sequence decreasing to 1.

- There is a finite set \( S \subset M \) such that the \( u_\alpha \) converge in \( C^\infty \) over \( M \setminus S \).
The limit \( u \) of the maps \( u_\alpha \) extends to a smooth harmonic map from \( M \) to \( N \) (which could be a constant map).

If \( x \) is a point in \( S \) such that the \( u_\alpha \) do not converge to \( u \) over a neighbourhood of \( x \) then there is a non-trivial harmonic map \( \nu : S^2 \to N \) such that a suitable sequence of rescalings of the \( u_\alpha \) near \( x \) converge to \( \nu \).

In brief, the only way that the sequence \( u_\alpha \) may fail to converge is by forming “bubbles,” in which small discs in \( M \) are blown up into harmonic spheres in \( N \). We illustrate the meaning of this bubbling through the example of rational maps of the 2-sphere. (See also the expository article [11].) For distinct points \( z_1, \ldots, z_d \) in \( \mathbb{C} \) and non-zero coefficients \( a_i \) consider the map

\[
u(z) = \sum_{i=1}^{d} \frac{a_i}{z - z_i},
\]

which extends to a degree \( d \) holomorphic map \( \nu : S^2 \to S^2 \) with \( \nu(\infty) = 0 \). These are in fact harmonic maps, with the same energy \( 8\pi d \). Take \( z_1 = 0, a_1 = \epsilon \). If we make \( \epsilon \) tend to 0, with the other \( a_i \) fixed, then away from 0 the maps converge to the degree \( (d-1) \) map \( \sum_{i=2}^{d} a_i (z - z_i)^{-1} \).

On the other hand if we rescale about 0 by setting

\[
\tilde{u}(z) = u(\epsilon z) = \frac{1}{z} + \sum_{i=2}^{d} \frac{a_i}{\epsilon z - z_i},
\]

the rescaled maps converge (on compact subsets of \( \mathbb{C} \)) to the degree 1 map

\[
\nu(z) = \frac{1}{z} - c
\]

with \( c = \sum_{i=2}^{d} a_i / z_i \).

A key step in the Sacks and Uhlenbeck analysis is a “small energy” statement (related to earlier results of Morrey). This says that there is some \( \epsilon > 0 \) such that if the energy of a map \( u_\alpha \) on a small disc \( D \subset N \) is less than \( \epsilon \) then there are uniform estimates of all derivatives of \( u_\alpha \) over the half-sized disc. The convergence result then follows from a covering argument. Roughly speaking, if the energy of the map on \( M \) is at most \( \tilde{E} \) then there can be at most a fixed number \( E/\epsilon \) of small discs on which the map is not controlled. The crucial point is that \( \epsilon \) does not depend on the size of the disc, due to the scale invariance of the energy. To sketch the proof of the small energy result, consider a simpler model equation

\[
\Delta f = |\nabla f|^2,
\]

for a function \( f \) on the unit disc in \( \mathbb{C} \). Linear elliptic theory, applied to the Laplace operator, gives estimates of the schematic form

\[
\|\nabla f\|_{L^q} \leq C\|\Delta f\|_{L^q} + \text{LOT},
\]

where LOT stands for “lower order terms” in which (for this sketch) we include the fact that one will have to restrict to an interior region. Take for example \( q = 4/3 \). Then substituting into the equation (3) we have

\[
\|\nabla f\|_{L^{4/3}} \leq C\|\nabla f\|_{L^{4/3}}^2 + \text{LOT} \leq C\|\nabla f\|_{L^{8/3}}^2 + \text{LOT}.
\]

Now in dimension 2 we have a Sobolev embedding \( L^{4/3} \to L^4 \) which yields

\[
\|\nabla f\|_{L^4} \leq C\|\nabla f\|_{L^{8/3}}^2 + \text{LOT}.
\]

On the other hand, Hölder’s inequality gives the interpolation

\[
\|\nabla f\|_{L^{8/3}} \leq \|\nabla f\|_{L^2}^{1/2} \|\nabla f\|_{L^4}^{1/2}.
\]

So, putting everything together, one has

\[
\|\nabla f\|_{L^4} \leq C\|\nabla f\|_{L^4} \|\nabla f\|_{L^2} + \text{LOT}.
\]

If \( \|\nabla f\|_{L^2} \leq 1/2C \) we can re-arrange this to get

\[
\|\nabla f\|_{L^4} \leq \text{LOT}.
\]

In other words, in the small energy regime (with \( \sqrt{\epsilon} = 1/2C \)) we can bootstrap using the equation to gain an estimate on a slightly stronger norm \( L^4 \) rather than \( L^2 \) and one continues in similar fashion to get interior estimates on all higher derivatives.

This breakthrough work of Sacks and Uhlenbeck ties in with many other developments from the same era, some of which we discuss in the next section and some of which we mention briefly here.

- In minimal submanifold theory: when \( M \) is a 2-sphere the image of a harmonic map is a minimal surface in \( N \) (or more precisely a branched immersed submanifold). In this way, Sacks and Uhlenbeck obtained an important existence result for minimal surfaces.
- In symplectic topology the pseudoholomorphic curves, introduced by Gromov in 1986, are examples of harmonic maps and a variant of the Sacks-
Uhlenbeck theory is the foundation for all the ensuing developments (see, for example, [7]).

- In PDE theory other “critical exponent” variational problems, in which similar bubbling phenomena arise, were studied intensively (see for example the work of Brezis and Nirenberg [4]).
- In Riemannian geometry the Yamabe problem of finding a metric of constant scalar curvature in a given conformal class (on a manifold of dimension 3 or more) is a critical exponent variational problem for the Einstein-Hilbert functional (the integral of the scalar curvature), restricted to metrics of volume 1. Schoen proved the existence of a minimiser, completing the solution of the Yamabe problem, using a deep analysis to rule out the relevant bubbling [14].

A beautiful application of the Sacks–Uhlenbeck theory was obtained in 1988 by Micallef and Moore [8]. The argument is in the spirit of classical applications of geodesics in Riemannian geometry. Micallef and Moore considered a curvature condition on a compact Riemannian manifold \( N \) (of dimension at least 4) of having “positive curvature on isotropic 2-planes.” They proved that if \( N \) satisfies this condition and is simply connected then it is a homotopy sphere (and thus, by the solution of the Poincaré conjecture, is homeomorphic to a sphere). The basic point is that a non-trivial homotopy class in \( \pi_k(N) \) gives a non-trivial element of \( \pi_{k-2}(X) \), where \( X = \text{Maps}(S^2, N) \), which gives a starting point for a minimax argument. If \( N \) is not a homotopy sphere then by standard algebraic topology there is some \( k \) with \( 2 \leq k \leq \frac{1}{2}\dim N \) such that \( \pi_k(N) \neq 0 \), which implies that \( \pi_{k-2}(X) \) is non-trivial. By developing mini-max arguments with the Sacks–Uhlenbeck theory, using the perturbed energy functional, Micallef and Moore were able to show that this leads to a non-trivial harmonic map \( u : S^2 \to N \) of index at most \( k - 2 \). (Here the index is the dimension of the space on which the second variation is strictly negative.) On the other hand the Levi–Civita connection of \( N \) defines a holomorphic structure on the pull-back \( u^*(TN \otimes \mathbb{C}) \) of the complexified tangent bundle. By combining results about holomorphic bundles over \( S^2 \) and a Weitzenböck formula, in which the curvature tensor of \( N \) enters, they show that the index must be at least \( \frac{1}{2}\dim N - \frac{3}{2} \) and thus derive a contradiction.

If the sectional curvature of \( N \) is \( \frac{3}{4} \)-pinched (i.e. lies between \( \frac{1}{4} \) and 1 everywhere) then \( N \) has positive curvature on isotropic 2-planes. Thus the Micallef and Moore result implies the classical sphere theorem of Berger and Klingenberg, whose proof was quite different. In turn, much more recently, Brendle and Schoen [3] proved that a (simply connected) manifold satisfying this isotropic curvature condition is in fact diffeomorphic to a sphere. Their proof was again quite different, using Ricci flow.

**Gauge Theory in Dimension 4**

From the late 1970s, mathematics was enriched by questions inspired by physics, involving gauge fields and the Yang-Mills equations. These developments were many-faceted and here we will focus on aspects related to variational theory. In this set-up one considers a fixed Riemannian manifold \( M \) and a \( G \)-bundle \( P \to M \) where \( G \) is a compact Lie group. The distinctive feature, compared to most previous work in differential geometry, is that \( P \) is an auxiliary bundle not directly tied to the geometry of \( M \). The basic objects of study are connections on \( P \). In a local trivialisation \( \tau \) of \( P \) a connection \( A \) is given by a \( \text{Lie}(G) \)-valued 1-form \( A^\tau \). For simplicity we take \( G \) to be a matrix group, so \( A^\tau \) is a matrix of 1-forms. The fundamental invariant of a connection is its curvature \( F(A) \) which in the local trivialisation is given by the formula

\[
F^\tau = dA^\tau + A^\tau \wedge A^\tau.
\]

The Yang-Mills functional is

\[
\mathcal{F}(A) = \int_M |F(A)|^2,
\]

and the Euler–Lagrange equation is \( d_A^\tau F = 0 \) where \( d_A^\tau \) is an extension of the usual operator \( d^\tau \) from 2-forms to 1-forms, defined using \( A \). This Yang-Mills equation is a non-linear generalisation of Maxwell’s equations of electromagnetism (which one obtains taking \( G = U(1) \) and passing to Lorentzian signature).

In the early 1980s, Uhlenbeck proved fundamental analytical results which underpin most subsequent work in this area. The main case of interest is when the manifold \( M \) has dimension 4 and the problem is then of critical exponent type. In this dimension the Yang-Mills functional is conformally invariant and there are many analogies with the harmonic maps of surfaces discussed above. A new aspect involves gauge invariance, which does not have an analogy in the harmonic maps setting. That is, the infinite dimensional group \( G \) of automorphisms of the bundle \( P \) acts on the space \( \mathcal{A} \) of connections, preserving the Yang-Mills functional, so the natural setting for the variational theory is the quotient space \( \mathcal{A}/G \). Locally we are free to change a trivialisation \( \tau_0 \) by the action of a \( G \)-valued function \( g \) which will change the local representation of the connection to

\[
A^\tau g \tau_0 = gd(g^{-1}) + gA^\tau g^{-1}.
\]

While this action of the gauge group \( G \) may seem unusual, within the context of PDEs, it represents a fundamental phenomenon in differential geometry. In studying Riemannian metrics, or any other kind of structure, on
a manifold one has to take account of the action of the
infinite-dimensional group of diffeomorphisms; for ex-
ample the round metric on the sphere is only unique up
to this action. Similarly, the explicit local representation of
a metric depends on a choice of local co-ordinates. In fact
diffeomorphism groups are much more complicated than
the gauge group $G$. In another direction one can have in
mind the case of electromagnetism, where the connection
1-form $A^T$ is equivalent to the classical electric and mag-
netic potentials on space-time. The $G$-action corresponds
to the fact that these potentials are not unique.

Two papers of Uhlenbeck [25], [26] addressed both of
these aspects (critical exponent and gauge choice). The pa-
per [25] bears on the choice of an “optimal” local trivialisa-
tion $\tau$ of the bundle over a ball $B \subset M$ given a connection
$A$. The criterion that Uhlenbeck considers is the Coulomb
gauge fixing condition: $d^*A^T = 0$, supplemented with
the boundary condition that the pairing of $A^T$ with the
normal vector vanishes. Taking $\tau = gT_0$, for some arbi-
trary trivialisation $T_0$, this becomes an equation for the $G$-
valued function $g$ which is a variant of the harmonic map
equation, with Neumann boundary conditions. In fact the
equation is the Euler–Lagrange equation associated to the
functional $\|A^T\|_{L^2}$, on local trivialisations $\tau$. The
Yang-Mills equations in such a Coulomb gauge form an elliptic
system. (Following the remarks in the previous paragraph;
analogous discussion for Riemannian metrics involves
harmonic local co-ordinates, in which the Einstein equa-
tions, for example, form an elliptic system.)

The result proved by Uhlenbeck in [25] is of “small en-
ergy” type. Specialising to dimension 4 for simplicity, she
shows that there is an $\epsilon > 0$ and a constant $C$ such that if
$\|F\|_{L^2(B)} < \epsilon$ there is a Coulomb gauge $\tau$ over $B$ in which

$$\|\nabla A^T\|_{L^2} + \|A^T\|_{L^4} \leq C\|F\|_{L^2}.$$ 

The strategy of proof uses the continuity method, applied
to the family of connections given by restricting to smaller
balls with the same centre, and the key point is to obtain a
priori estimates in this family. The PDE arguments deriving
these estimates have some similarity with those sketched in
Section “Harmonic maps in dimension 2” above. An
important subtlety arises from the critical nature of the
Sobolev exponents involved. If $\tau = gT_0$ then an $L^2$ bound
on $\nabla A^T$ gives an $L^2$ bound on the second derivative of $g$
but in dimension 4 this is the borderline exponent where we
do not get control over the continuity of $g$. That makes
the nonlinear operations such as $g \rightarrow g^{-1}$ problematic.
Uhlenbeck overcomes this problem by working with $L^p$
for $p > 2$ and using a limiting argument.

In the companion paper [26], Uhlenbeck proves a
renowned “removal of singularities” result. The statement
is that a solution $A$ of the Yang-Mills equations over the
punctured ball $B^4 \setminus \{0\}$ with finite energy (i.e. with curva-
ture $F(A)$ in $L^2$) extends smoothly over 0 in a suitable local
trivialisation. One important application of this is that
finite-energy Yang-Mills connections over $\mathbb{R}^4$ extend to the
conformal compactification $S^4$. We will only attempt to
give the flavour of the proof. Given our finite-energy solu-
tion $A$ over the punctured ball let

$$f(r) = \int_{|x| < r} |F(A)|^2,$$

for $r < 1$. Then the derivative is

$$\frac{df}{dr} = \int_{|x| = r} |F(A)|^2.$$

The strategy is to express $f(r)$ also as a boundary integral,
plus lower order terms. To give a hint of this, consider
the case of an abelian group $G = U(1)$, so the connection
form $A^T$ is an ordinary 1-form, the curvature is simply $F =
dA^T$, and the Yang-Mills equation is $d^*F = 0$. Fix small
$\epsilon < r$ and work on the annular region $W$ where $\epsilon < |x| <
r$. We can integrate by parts to write

$$\int_W |F|^2 = \int_W (dA^T, F) = \int_W \langle A^T, d^*F \rangle + \int_{\partial W} A^T \wedge *F.$$ 

Since $d^*F = 0$ the first term on the right hand side van-
ishes. If one can show that the contribution from the inner
boundary $|x| = \epsilon$ tends to 0 with $\epsilon$ then one concludes
that

$$f(r) = \int_{|x| = r} A^T \wedge *F.$$ 

In the nonabelian case the same discussion applies up to
the addition of lower-order terms, involving $A^T \wedge A^T$. The
strategy is then to obtain a differential inequality of the shape

$$f(r) \leq \frac{1}{4r} \frac{df}{dr} + \text{LOT}, \quad (4)$$

by comparing the boundary terms over the 3-sphere. This
differential inequality integrates to give $f(r) \leq Cr^4$ and
from there it is relatively straightforward to obtain an $L^\infty$ bound on the curvature and to see that the connection
can be extended over 0. The factor $\frac{1}{4}$ in (4) is obtained from
an inequality over the 3-sphere. That is, any closed 2-form $\omega$ on $S^3$ can be expressed as $\omega = da$ where

$$\|a\|^2_{L^2(S^3)} \leq \frac{1}{4} \|\omega\|^2_{L^2(S^3)}.$$ 

The main work in implementing this strategy is to con-
struct suitable gauges over annuli in which the lower order,
nonlinear terms $A^T \wedge A^T$ are controlled.

These results of Uhlenbeck lead to a Yang-Mills analogue
of the Sacks–Uhlenbeck picture discussed in the previous
section. This was not developed explicitly in Uhlenbeck’s
1983 papers [25], [26] but results along those lines were
obtained by her doctoral student S. Sedlacek [16]. Let $c$
be the infimum of the Yang-Mills functional on connections on $P \to X$, where $X$ is a compact 4-manifold. Let $A_i$ be a minimising sequence. Then there is a (possibly different) $G$-bundle $\tilde{P} \to X$, a Yang-Mills connection $A_\infty$ on $\tilde{P}$, and a finite set $S \subset X$ such that, after perhaps passing to a subsequence $i'$, the $A_{i'}$ converge to $A_\infty$ over $X \setminus S$. 

(More precisely, this convergence is in $L^2_{1,\text{loc}}$ and implicitly involves a sequence of bundle isomorphisms of $P$ and $\tilde{P}$ over $X \setminus S$.) If $x$ is a point in $S$ such that the $A_{i'}$ do not converge to $A_\infty$ over a neighbourhood of $x$ then one obtains a non-trivial solution to the Yang-Mills equations over $S^4$ by a rescaling procedure similar to that in the harmonic map case. Similar statements apply to sequences of solutions to the Yang-Mills equations over $X$ and in particular to sequences of Yang-Mills “instantons.” These special solutions solve the first order equation $F = \pm \ast F$ and are closely analogous to the pseudoholomorphic curves in the harmonic map setting. Uhlenbeck’s analytical results underpinned the applications of instanton moduli spaces to 4-manifold topology which were developed vigorously throughout the 1980s and 1990s—just as for pseudoholomorphic curves and symplectic topology. But we will concentrate here on the variational aspects.

For simplicity fix the group $G = SU(2)$; the $SU(2)$-bundles $P$ over $X$ are classified by an integer $k = c_2(P)$ and for each $k$ we have a moduli space $\mathcal{M}_k$ (possibly empty) of instantons (where the sign in $F = \pm \ast F$ depends on the sign of $k$). Recall that the natural domain for the Yang-Mills functional is the infinite-dimensional quotient space $X_k = \mathcal{A}_k/G_k$ of connections modulo equivalence. The moduli space $\mathcal{M}_k$ is a subset of $X_k$ and (if non-empty) realises the absolute minimum of the Yang-Mills functional on $X_k$. In this general setting one could, optimistically, hope for a variational theory which would relate:

1. The topology of the ambient space $X_k$.
2. The topology of $\mathcal{M}_k$.
3. The non-minimal critical points: i.e. the solutions of the Yang-Mills equation which are not instantons.

A serious technical complication here is that the group $G_k$ does not usually act freely on $\mathcal{A}_k$, so the quotient space is not a manifold. But we will not go into that further here and just say that there are suitable homology groups $H_i(X_k)$, which can be studied by standard algebraic topology techniques and which have a rich and interesting structure.

Much of the work in this area in the late 1980s was driven by two specific questions.

- The Atiyah-Jones conjecture [1]. They considered the manifold $M = S^4$ where (roughly speaking) the space $X_k$ has the homotopy type of the degree $k$ mapping space $\text{Maps}_k(S^3, S^3)$, which is in fact independent of $k$. The conjecture was that the inclusion $\mathcal{M}_k \to X_k$ induces an isomorphism on homology groups $H_i$ for $i$ in a range $i \leq i(k)$, where $i(k)$ tends to infinity with $k$. One motivation for this idea came from results of Segal in the analogous case of rational maps [17].

- Again focusing on $M = S^4$: are there any non-minimal solutions of the Yang-Mills equations?

A series of papers of Taubes [20], [22] developed a variational approach to the Atiyah-Jones conjecture (and generalisations to other 4-manifolds). In [20] Taubes established a lower bound on the index of any non-minimal solution over the 4-sphere. If the problem satisfied the Palais–Smale condition this index bound would imply the Atiyah-Jones conjecture (with $i(k)$ roughly $2k$) but the whole point is that this condition is not satisfied, due to the bubbling phenomenon for mini-max sequences. Nevertheless, Taubes was able to obtain many partial results through a detailed analysis of this bubbling. The Atiyah-Jones conjecture was confirmed in 1993 by Boyer, Hurtubise, Mann, and Milgram [2] but their proof worked with geometric constructions of the instanton moduli spaces, rather than variational arguments.

The second question was answered, using variational methods, by Sibner, Sibner, and Uhlenbeck in 1989 [18], showing that indeed such solutions do exist. In their proof they considered a standard $S^1$ action on $S^4$ with fixed point set a 2-sphere, an $S^1$-equivariant bundle $P$ over $S^4$ and $S^1$-invariant connections on $P$. This invariance forces the “bubbling points” arising in variational arguments to lie on the 2-sphere $S^2 \subset S^4$ and there is a dimensional reduction of the problem to “monopoles” in 3-dimensions which has independent interest.

A connection over $\mathbb{R}^3$ which is invariant under the action of translations in one direction can be encoded as a pair $(A, \phi)$ of a connection $A$ over $\mathbb{R}^3$ and an additional Higgs field $\phi$ which is a section of the adjoint vector bundle $\text{ad}P$ whose fibres are copies of $\text{Lie}(G)$. The Yang-Mills functional induces a Yang-Mills-Higgs functional

$$\mathcal{F}(A, \phi) = \int_{\mathbb{R}^3} |F(A)|^2 + |\nabla A\phi|^2$$

on these pairs over $\mathbb{R}^3$. One also fixes an asymptotic condition $|\phi|$ tends to 1 at $\infty$ in $\mathbb{R}^3$. In 3 dimensions we are below the critical dimension for the functional, but the noncompactness of $\mathbb{R}^3$ prevents a straightforward verification of the Palais–Smale condition. Nonetheless, in a series of papers [19], [21] Taubes developed a far-reaching variational theory in this setting. By a detailed analysis, Taubes showed that, roughly speaking, a minimax sequence can always be chosen to have energy density concentrated...
in a fixed large ball in \( \mathbb{R}^3 \) and thus obtained the necessary convergence results. In particular, using this analysis, Taubes established the existence of non-minimal critical points for the functional \( J(A, \phi) \).

The critical points of the Yang-Mills-Higgs functional on \( \mathbb{R}^3 \) yield Yang-Mills solutions over \( \mathbb{R}^4 \), but these do not have finite energy. However the same ideas can be applied to the \( S^1 \)-action. The quotient of \( S^4 \setminus S^2 \) by the \( S^1 \)-action can naturally be identified with the hyperbolic 3-space \( H^3 \), and \( S^1 \)-invariant connections correspond to pairs \( (A, \phi) \) over \( H^3 \). There is a crucial parameter \( L \) in the theory which from one point of view is the weight of the \( S^1 \) action on the fibres of \( P \) over \( S^2 \). From another point of view the curvature of the hyperbolic space, after suitable normalisation, is \( -L^{-2} \). The fixed set \( S^2 \) can be identified with the sphere at infinity of hyperbolic space and bubbling of connections over a point in \( S^2 \subset S^4 \) corresponds, in the Yang-Mills-Higgs picture, to some contribution to the energy density of \( (A, \phi) \) moving off to the corresponding point at infinity.

The key idea of Sibner, Sibner, and Uhlenbeck was to make the parameter \( L \) very large. This means that the curvature of the hyperbolic space is very small and, on sets of fixed diameter, the hyperbolic space is well-approximated by \( \mathbb{R}^3 \). Then they show that Taubes’ arguments on \( \mathbb{R}^4 \) go over to this setting and are able to produce the desired non-minimal solution of the Yang-Mills equations over \( S^4 \). Later, imposing more symmetry, other solutions were found using comparatively elementary arguments [13], but the approach of Taubes, Sibner, Sibner, and Uhlenbeck is a paradigm of the way that variational arguments can be used “beyond Palais–Smale,” via a delicate analysis of the behaviour of minimax sequences.

We conclude this section with a short digression from the main theme of this article. This brings in other relations between harmonic mappings of surfaces and 4-dimensional gauge theory, and touches on another very important line of work by Karen Uhlenbeck, represented by papers such as [27], [28]. In this setting the target space \( N \) is a symmetric space and the emphasis is on explicit solutions and connections with integrable systems. There is a huge literature on this subject, stretching back to work of Calabi and Chern in the 1960s, and distantly connected with the Weierstrass representation of minimal surfaces in \( \mathbb{R}^3 \). From around 1980 there were many contributions from theoretical physicists and any kind of proper treatment would require a separate article, so we just include a few remarks here.

As we outlined above, the dimension reduction of Yang-Mills theory on \( \mathbb{R}^4 \) obtained by imposing translation-invariance in one variable leads to equations for a pair \( (A, \phi) \) on \( \mathbb{R}^3 \). Now reduce further by imposing translation-invariance in two directions. More precisely, write \( \mathbb{R}^4 = \mathbb{R}_1^2 \times \mathbb{R}_2^2 \), fix a simply-connected domain \( \Omega \subset \mathbb{R}_1^2 \), and consider connections on a bundle over \( \Omega \times \mathbb{R}_2^2 \) which are invariant under translations in \( \mathbb{R}_1^2 \). These correspond to pairs \( (A, \phi) \) where \( A \) is a connection on a bundle \( P \) over \( \Omega \) and \( \phi \) can be viewed as a 1-form on \( \Omega \) with values in the bundle \( \text{ad}P \). Now \( A + i\phi \) is a connection over \( \Omega \) for a bundle with structure group the complexification \( G^c \); for example if \( G = U(r) \) the complexified group is \( G^c = GL(r, \mathbb{C}) \). The Yang-Mills instanton equations on \( \mathbb{R}_1^2 \) imply that \( A + i\phi \) is a flat connection. By the fundamental property of curvature, since \( \Omega \) is simply-connected, this flat connection can be trivialised. The original data \( (A, \phi) \) is encoded in the reduction of the trivial \( G^c \)-bundle to the subgroup \( G \), which amounts to a map \( u \) from \( \Omega \) to the non-compact symmetric space \( G^c/G \). For example, when \( G = U(r) \) the extra data needed to recover \( (A, \phi) \) is a Hermitian metric on the fibres of the complex vector bundle, and \( GL(r, \mathbb{C})/U(r) \) is the space of Hermitian metrics on \( C^r \). The the remaining part of the instanton equations in four dimensions is precisely the harmonic map equation for \( u \). This is one starting point for Hitchin’s theory of “stable pairs” over compact Riemann surfaces [6].

One is more interested in harmonic maps to compact symmetric spaces and, as Uhlenbeck explained in [28], this can be achieved by a modification of the set-up above. She takes \( \mathbb{R}^4 \) with an indefinite quadratic form of signature \((2, 2)\) and a splitting \( \mathbb{R}^4 = \mathbb{R}_1^2 \times \mathbb{R}_2^2 \) into positive and negative subspaces. Then the invariant instantons correspond to harmonic maps from \( \Omega \) to the compact Lie group \( G \). Other symmetric spaces can be realised as totally geodesic submanifolds in the Lie group, for example complex Grassmann manifolds in \( U(r) \), and the theory can be specialised to suit. This builds a bridge between the “integrable” nature of the 2-dimensional harmonic map equations and the Penrose-Ward twistor description of Yang-Mills instantons over \( \mathbb{R}^4 \), although as we have indicated above much of the work on the former predates twistor theory. In her highly influential paper [28], Uhlenbeck found an action of the loop group on the space of harmonic maps from \( \Omega \) to \( G \), introduced an integer invariant “uniton number,” and obtained a complete description of all harmonic maps from the Riemann sphere to \( G \).

**Higher Dimensions**

In a variational theory with a critical dimension \( v \) certain characteristic features appear when studying questions in dimensions greater than \( v \). In the harmonic mapping theory, for maps \( u : M \to N \), the dimension in question is \( n = \dim M \) and, as we saw above, the critical dimension is \( v = 2 \). A breakthrough in the higher dimensional theory was obtained by Schoen and Uhlenbeck in [12]. Suppose for simplicity that \( N \) is isometrically embedded in
some Euclidean space $\mathbb{R}^k$ and define $L^2_1(M, N)$ to be the set of $L^2$ functions on $M$ with values in the vector space $\mathbb{R}^k$ which map to $N$ almost everywhere on $M$. The energy functional $\mathcal{F}$ is defined on $L^2_1(M, N)$ and Schoen and Uhlenbeck considered an energy minimising map $u \in L^2_1(M, N)$. The main points of the theory are:

- $u$ is smooth outside a singular set $\Sigma \subset M$ which has Hausdorff dimension at most $n - 3$;
- at each point $x$ in the singular set $\Sigma$ there is a tangent map to $u$.

The second item means that there is a sequence of real numbers $\sigma_i \to 0$ such that the rescaled maps 

$$u_i(\xi) = u(\exp(x)(\sigma_i \xi))$$

converge to a map $v : \mathbb{R}^n \to N$ which is radially invariant, and hence corresponds to a map from the sphere $S^{n-1}$ to $N$. (Here $\exp_x$ is the Riemannian exponential map and we have chosen a frame to identify $T_xM$ with $\mathbb{R}^n$.)

To relate this to the case $n = 2$ discussed above, the general picture is that a $f$-minimising sequence in Maps $(M, N)$ can be taken to converge outside a bubbling set of dimension at most $n - 2$ and the limit extends smoothly over the $(n-2)$-dimensional part of the bubbling set. The new feature in higher dimensions is that the limit can have a singular set of codimension 3 or more.

Two fundamental facts which underpin these results are energy monotonicity and $\epsilon$-regularity. To explain the first, consider a smooth harmonic map $U : B^n \to N$, where $B^n$ is the unit ball in $\mathbb{R}^n$. For $r < 1$ set 

$$E(r) = \frac{1}{r^{n-2}} \int_{|x| < r} |\nabla U|^2.$$

Then one has an identity, for $r_1 < r_2$:

$$E(r_2) - E(r_1) = 2 \int_{r_1 < |x| < r_2} |x|^{2-n} |\nabla U|^2, \quad (5)$$

where $\nabla_r$ is the radial component of the derivative. In particular, $E$ is an increasing function of $r$. The point of this is that $E(r)$ is a scale-invariant quantity. If we define $U_r(x) = U(rx)$ then $E(r)$ is the energy of the map $U_r$ on the unit ball. The monotonicity property means that $U$ “looks better” on a small scale, in the sense of this rescaled energy. The identity (5) follows from a very general argument, applying the stationary condition to the infinitemal variation of $U$ given by radial dilation. (One way of expressing this is through the theory of the stress-energy tensor.) Note that equality $E(r_2) = E(r_1)$ holds if and only if $U$ is radially-invariant in the corresponding annulus. This is what ultimately leads to the existence of radially-invariant tangent maps.

The monotonicity identity is a feature of maps from $\mathbb{R}^n$, but a similar result holds for small balls in a general Riemannian $n$-manifold $M$. For $x \in M$ and small $r > 0$ we define 

$$E_x(r) = \frac{1}{r^{n-2}} \int_{B_r(x)} |\nabla U|^2,$$

where $B_r(x)$ is the $r$-ball about $x$. Then if $U$ is a smooth harmonic map and $x$ is fixed the function $E_x(r)$ is increasing in $r$, up to harmless lower-order terms.

The $\epsilon$-regularity theorem of Schoen and Uhlenbeck states that there is an $\epsilon > 0$ such that if $u$ is an energy minimiser then $u$ is smooth in a neighbourhood of $x$ if and only if $E_x(r) < \epsilon$ for some $r$. An easier, related result is that if $u$ is known to be smooth then once $E_x(r) < \epsilon$ one has a priori estimates (depending on $r$) on all derivatives in the interior ball $B_x(r/2)$. The extension to general minimising maps is one of the main technical difficulties overcome by Schoen and Uhlenbeck.

We turn now to corresponding developments in gauge theory, where the critical dimension $\nu$ is 4. A prominent achievement of Uhlenbeck in this direction is her work with Yau on the existence of Hermitian-Yang-Mills connections [29]. The setting here involves a rank $r$ holomorphic vector bundle $E$ over a compact complex manifold $M$ with a Kähler metric. Any choice of Hermitian metric $h$ on the fibres of $E$ defines a principle $U(r)$ bundle of orthonormal frames in $E$ and a basic lemma in complex differential geometry asserts that there is a preferred connection on this bundle, compatible with the holomorphic structure. The curvature $F = F(h)$ of this connection is a bundle-valued 2-form of type $(1, 1)$ with respect to the complex structure, and we write $\Lambda F$ for the inner product with the $(1, 1)$ form defined by the Kähler metric. Then $\Lambda F$ is a section of the bundle of endomorphisms of $E$. The Hermitian-Yang-Mills equation is a constant multiple of the identity:

$$\Lambda F = \kappa \lambda$$

(where the constant $\kappa$ is determined by topology). As the name suggests, these are special solutions of the Yang-Mills equations. The result proved by Uhlenbeck and Yau is that a “stable” holomorphic vector bundle admits such a Hermitian-Yang-Mills connection. Here stability is a numerical condition on holomorphic sub-bundles, or more generally sub-sheaves, of $E$ which was introduced by algebraic geometers studying moduli theory of holomorphic bundles. The result of Uhlenbeck and Yau confirmed conjectures made a few years before by Kobayashi and Hitchin. These extend older results of Narasimhan and Seshadri, for bundles over Riemann surfaces, and fit into a large development over the past 40 years, connecting various stability conditions in algebraic geometry with differential geometry. We will not say more about this background here but focus on the proof of Uhlenbeck and Yau.

The problem is to solve the equation $\Lambda F(h) = \kappa \lambda$ for a Hermitian metric $h$ on $E$. This boils down to a second order, nonlinear, partial differential equation for $h$. 
While this problem does not fit directly into the variational framework we have emphasised in this article, the same compactness considerations apply. Uhlenbeck and Yau use a continuity method, extending to a 1-parameter family of equations for \( t \in [0, 1] \) which we write schematically as \( \Delta F(h_t) = K_t \), where \( K_t \) is prescribed and \( K_1 = 1 \). They set this up so that there is a solution \( h_0 \) for \( t = 0 \) and the set \( T \subset [0, 1] \) for which a solution \( h_t \) exists is open, by an application of the implicit function theorem. The essential problem is to prove that if \( E \) is a stable holomorphic bundle then \( T \) is closed, hence equal to the whole of \([0, 1]\) and in particular there is a Hermitian-Yang-Mills connection \( h_1 \).

The paper of Uhlenbeck and Yau gave two independent treatments of the core problem, one emphasising complex analysis and the other gauge theory. We will concentrate here on the latter. For a sequence \( t(i) \in T \) we have connections \( A_i \) defined by the hermitian metrics \( h_t(i) \) and the question is whether one can take a limit of the \( A_i \). The deformation of the equations by the term \( K_t \) is rather harmless here so the situation is essentially the same as if the \( A_i \) were Yang-Mills connections. In addition, an integral identity using Chern-Weil theory shows that the Yang-Mills energy \( \|F(A_i)\|_{L_2}^2 \) is bounded. Then Uhlenbeck and Yau introduced a small energy result, for connections over a ball \( B_n(r) \subset M \). Since the critical dimension \( v \) is 4, the relevant normalised energy in this Yang-Mills setting is

\[
E_n(r) = \frac{1}{r^{n-4}} \int_{B_n(r)} |F|^2,
\]

where \( n \) is the real dimension of \( M \). If \( E_n(r) \) is below a suitable threshold there are interior bounds on all derivatives of the connection, in a suitable gauge. Then the global energy bound implies that after perhaps taking a subsequence, the \( A_i \) converge outside a closed set \( S \subset M \) of Hausdorff codimension at least 4. Uhlenbeck and Yau show that if the metrics \( h_t(i) \) do not converge then a suitable rescaled limit produces a holomorphic subbundle of \( E \) over \( M \setminus S \). A key technical step is to show that this subbundle corresponds locally to a meromorphic map to a Grassmann manifold, which implies that the subbundle extends as a coherent sheaf over all of \( M \). The differential geometric representation of the first Chern class of this subsheaf, via curvature, shows that it violates the stability hypothesis.

The higher-dimensional discussion in Yang-Mills theory follows the pattern of that for harmonic maps above. The corresponding monotonicity formula was proved by Price [10] and a treatment of the small energy result was given by Nakajima [9]. Some years later, the theory was developed much further by Tian [23], including the existence of “tangent cones” at singular points.

This whole circle of ideas and techniques involving the dimension of singular sets, monotonicity, “small energy” results, tangent cones, etc. has had a wide-ranging impact in many branches of differential geometry over the past few decades and forms the focus of much current research activity. Apart from the cases of harmonic maps and Yang-Mills fields discussed above, prominent examples are minimal submanifold theory, where many of the ideas appeared first, and the convergence theory of Riemannian metrics with Ricci curvature bounds.

References


Credits
Photo of Karen Uhlenbeck is by Andrea Kane, ©IAS. Author photo is courtesy of the author.