SELFDUALITY IN GEOMETRY: YANG-MILLS CONNECTIONS AND SELFDUAL LAGRANGIANS

by

JASON RODERICK DONALDSON

B.Sc., Simon Fraser University, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

April 2007

Abstract

The convex theory of selfdual Lagrangians recently developed by Ghoussoub analyses functionals rooted in an expanse of partial differential equations and finds their minima not variationally but rather by realizing that they assume a prescribed lower bound. This is exactly the circumstance in the selfdual and anti-selfdual Yang-Mills equations that arise in the physical field theory and the study of the geometric and topological structure of four-dimensional manifolds.

I expose the Yang-Mills equations, building up the geometry from student-level and subsequently outline the setting of selfdual Lagrangians. The theories are clearly analogous and the last section feints at the exact link.

Table of Contents

Chapter 1: Introduction

What follows are two narrative introductions to loosely connected fields in modern analysis intended to be fully self-contained for a graduate student approaching the subjects after first graduate course sequences in differential geometry and real analysis. Given the basic understanding of finite-dimensional abstract manifolds—atlases, vector bundles, and differentiation—a student should experience the first sections as a solidification of notation and a reminder of definitions before learning the next batch of geometry—Lie and exterior algebras and integration—with the express purpose of understanding the Yang-Mills functional and the so-called selfdual and anti-selfdual equations for connections in four dimensions. Throughout I try to take special care to clarify the domains of operators and state clearly their alternative representations, since in learning geometry I have been and am still stumped particularly often by formulations switch and many spaces present themselves together here.

The exposition of the Yang-Mills theory should illustrate some geometric concepts, namely cohomology, while building up to some nice fundamental results. I compute the expression for the functional's critical points variationally, but the point is the four-dimensional case where the attainment of the lower bound, not the first order condition, gives the essential equations. This is the essential analogy with the convex theory of chapter 3. Jürgen Jost's book *Riemannian Geometry and Geomemtric Analysis* was my main reference and I largely follow its exposition.

Finally the third chapter outlines Ghoussoub's theory of selfdual Lagrangians. It strongly resembles the four-dimensional Yang-Mills theory in that it proposes to reformulate partial differential equations as minimization problems whose solutions determined depending on the realization that they attain their lower bounds. The key tool here is the Fenchel inequality of convex analysis. The framework's main strengths are its simplicity—it relies only on basic functional analysis and very basic convex analysis—and its overarchingness—it applies to an expanse of partial differential equations. I sketch some of the theory with a story-telling, pedagogical bent and formally align it with the Yang-Mills case in the final section. While Yang-Mills fits effortlessly into this framework the formal computations carried out at the end have yet to be fully justified mathematically. I hope (and suspect) they will be soon, especially because I believe too that they will extend to still richer circumstances.

Chapter 2: Geometric Buildup

To begin we clarify the definitions and notations, that are often malleable in the literature, for the objects essential to differential geometry that constitute the basal elements of what follows. The covariant derivative and its induced curvature are the essential objects of Yang-Mills theory. They sit on top of gauge theory and relate algebra and topology to real analysis and especially, as is our true subject, to modern partial differential equations which we view here through the application convex analysis in the calculus of variations.

The Covariant Derivative and Curvature Form

A covariant derivative[1] on a smooth, compact, finite dimensional manifold $\mathcal{M}$ is a map

$$D : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}),$$

$$D : (V, X) \mapsto D_V X$$

where $\mathfrak{X}(\mathcal{M})$ denotes the space of smooth vector fields on the manifold,

$$\mathfrak{X}(\mathcal{M}) = \{X \in C^\infty(M; T\mathcal{M}) \ ; \ X_p \in T_p\mathcal{M} \text{ for each } p \in \mathcal{M}\}.$$

[1]A covariant derivative is often called a connection. I will employ another convention and refer to the connection as the operator A in the decomposition $D = d+A$. Note though that while they are not equal in this formulation, there is still a one-to-one correspondence between covariant derivatives and connections.

In general we assume that the manifold has no boundary, $\partial\mathcal{M} = \emptyset$. Since in fact D depends on the first argument only pointwise—as opposed to locally—we may replace V with its image, $V_p = u$, and say instead, for each $p \in \mathcal{M}$,

$$D : T_p\mathcal{M} \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}),$$

$$D : (u, X) \mapsto D_u X,$$

or, finally but importantly, equivalently exchange the tangent space in the domain for its dual in the range, as

$$D : \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) \times T_p^*(\mathcal{M})$$

$$D : X \mapsto DX.$$

The covariant derivative is a derivative in the sense of parallel transport: Given a curve $\sigma \in C^1(\mathbb{R}; \mathcal{M})$ and a linear isomorphism

$$P_{t,\sigma} : T_{\sigma(0)}\mathcal{M} \to T_{\sigma(t)}\mathcal{M}$$

satisfying

$$\langle P_{t,\sigma}(u), P_{t,\sigma}(v)\rangle_{\sigma(t)} = \langle u, v\rangle_{\sigma(0)}$$

for all $u,\ v \in T_{\sigma(0)}\mathcal{M}$ and $t \in \mathbb{R}$ (where $\langle\cdot,\cdot\rangle_p$ denotes the chosen metric on

$T_p\mathcal{M}$), the equation

$$D_{\dot{\sigma}(0)}X = \frac{d}{dt}\bigg|_{t=0} P_{t,\sigma}^{-1} X(\sigma(t)) \tag{1}$$

defines the action of covariant derivative in the direction of any $u = \sigma(0) := \partial\sigma/\partial t \in T_p\mathcal{M},\ p = \sigma(0)$. But only up to the choice of the parallel transport map P—although this doesn't depend on the curve σ so long as it satisfies the initial condition. This definition clearly depends on the specification of the metric on $T\mathcal{M}$, but a more abstract equivalent formulation does not. The selection of covariant derivative to employ, however, typically depends on the metric. For example, the canonical Levi-Civita covariant derivative is the unique covariant derivative ∇ satisfying the torsion-free condition

$$\nabla_X Y - \nabla_Y X - [X, Y] = 0 \tag{2}$$

and the defining formula

$$X(\langle Y, Z\rangle) = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$$

for every $X,\ Y,\ Z \in \mathfrak{X}(\mathcal{M})$, where the Lie Bracket by the commutator $[\,\cdot\,,\cdot\,] : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$ acts as

$$[X, Y] := XY - YX$$

which means

$$[X,Y] = X^j \frac{\partial Y^i}{\partial x^j} \frac{\partial}{\partial x^i} - Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial}{\partial x^i}. \tag{3}$$

for

$$X = X^i \frac{\partial}{\partial x^i}, \quad Y = Y^i \frac{\partial}{\partial x^i}.$$

The definition of Yang-Mills covariant derivatives depends explicitly on the manifold's metric, but relies further on the global minimization of the manifold's curvature.

The curvature of a covariant derivative is its composition with itself in the second argument:

$$F_D := D \circ D : T_p\mathcal{M} \times T_p\mathcal{M} \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}). \tag{4}$$

Here, again consider the alternative formulation:

$$F_D : \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) \times T_p^*\mathcal{M} \times T_p^*\mathcal{M},$$

and introduce the notation

$$\Omega^p(T\mathcal{M}) := \mathfrak{X}(\mathcal{M}) \times \Omega^p(\mathcal{M}) \tag{5}$$

understanding that

$$\Omega^0(T\mathcal{M}) = \mathfrak{X}(\mathcal{M}).$$

Here $\Omega^s(\mathcal{M})$ is space of sections of the total space of p-differential forms. To be explicit (and perhaps, with apologies, verbose), the background notation is as follows: For $x \in \mathcal{M}$, p-forms at x are

$$\bigwedge^p T_x^*\mathcal{M} := \{\omega : T_x\mathcal{M} \times ... \times T_x\mathcal{M} \to \mathbb{R} \ ; \ \omega \text{ is linear}\}$$

and the corresponding total space is

$$\bigwedge^p T^*\mathcal{M} := \bigcup_{x \in \mathcal{M}} \bigwedge^p T_x^*\mathcal{M}.$$

As with any vector bundle, there is a projection $\pi : \bigwedge^p T^*\mathcal{M} \to \mathcal{M}$ such that for each $x \in \mathcal{M}$,

$$\pi^{-1}(x) = \bigwedge^p T_x^*\mathcal{M}$$

which allows the definition of the sections of the p-forms (and by analogy sections generally[2]) as

$$\begin{aligned} \Omega^p(\mathcal{M}) &:= \Gamma\left(\bigwedge^p T^*\mathcal{M}\right) \\ &:= \left\{ s : \bigwedge^p T^*\mathcal{M} \to \mathcal{M} \ ; \ \pi \circ s = \mathrm{Id} \right\}. \end{aligned} \tag{6}$$

Return to our examination of the curvature: In this language it is the

[2]Of course the vector fields are simply the sections of the tangent bundle, $\mathfrak{X}(\mathcal{M}) := \Gamma(T\mathcal{M})$. A later section addresses this definition as necessary for the expansion of the theory, but for now I attempt to sick as much as possible to the simplest case and spell out the required basics, because, as is already apparent, there is a lot of theory and notation to bog us down despite that my intended focus is examples.

mapping

$$F_D : \Omega^0(T\mathcal{M}) \to \Omega^2(T\mathcal{M}).$$

In the case of the Levi-Civita covariant derivative the curvature, $R := F_\nabla$, takes a simple form: For $u,\ v \in T_p\mathcal{M}$

$$R(u,v) : \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$$

as

$$R(u,v)X = \nabla_u\nabla_v X - \nabla_v\nabla_u X - \nabla_{[u,v]}X$$

(here it suffices to consider the bracket $[u,v] = [Y,Z]_p$ where the vector fields Y and Z are such that $Y_p = u$ and $Z_p = v$). While the Levi-Civita covariant derivatives come easily and give us a curvature form, the formulation required for the Yang-Mills functional and hence Yang-Mills covariant derivatives (which, in a manner opposite to the Levi-Civita case, result inversely from the curvature) will remain elusive since they depend on the development of further algebraic and analytic—geometric—structures on the manifold.

Lie Structure

Consider a group G which is itself a manifold ($G = \mathcal{M}$) where the group actions

$$G \times G \to G, \quad (x,y) \mapsto xy \tag{7}$$

and

$$G \to G, \quad x \mapsto x^{-1} \tag{8}$$

are smooth. G is called a Lie group. The diffeomorphic left translation $L_x : G \to G$,

$$L_x y := xy, \tag{9}$$

produces the essential mapping $L_{x*} : \mathfrak{X}(G) \to \mathfrak{X}(G)$,

$$L_{x*}X = dL_x X = X \circ L_x, \tag{10}$$

for exterior (total) derivative d, that defines the subspace $\mathfrak{g}$ of $\mathfrak{X}(G)$ called the Lie algebra of G by

$$\mathfrak{g} := \{X \in \mathfrak{X}(G)\ ; L_{x*}X_y = X_{xy} \text{ for all } x,\ y \in G\}. \tag{11}$$

(In this restricted context the above equation for $(L_x)_*$ suffices to define the push forward operator denoted by subscript star.) Observe computationally that for $X,\ Y \in \mathfrak{g}$ and $x \in G$,

$$L_{x*}[X, Y] = [L_{x*}X, L_{x*}Y], \tag{12}$$

so $[X, Y]$ is itself left-invarient and hence belongs to $\mathfrak{g}$. And note that the existence of such a bracket map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$ satisfying

$$[X, X] = 0 \tag{13}$$

and

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \tag{14}$$

whenever $X,\ Y,\ Z \in \mathfrak{g}$ defines the Lie algebra in a general setting. Thus the vector fields on $\mathcal{M}$ themselves are a Lie algebra.

For $x \in G$, the inner automorphism $A_x : G \longrightarrow G$ defined by

$$A_x y = xyx^{-1}, \tag{15}$$

induces an automorphism on $\mathfrak{g}$ via differentiation. For identity element e of G, $A_x e = e$, so the exterior derivative at e maps the tangent space at e into itself, $A_{x*}(e) : T_e G \hookrightarrow T_e G$, and the so-called adjoint representation of G, the map ad : $\mathfrak{g} \longrightarrow \mathfrak{g}$ given as

$$\mathrm{ad} : x \longmapsto A_{x*}(e), \tag{16}$$

yields an automorphism on $\mathfrak{g}$.

The first example of the imposition of Lie structure on a manifold comes in the form of the general linear group over the real numbers, the totality of nonsingular, real matrices of a given dimension,

$$\mathrm{GL}(n,\mathbb{R}) := \{X \in M(n,\mathbb{R})\ ; \det X \neq 0\}, \tag{17}$$

where, of course, $M(n,\mathbb{R})$ denotes the set of $n \times n$ real matrices. The group actions are in fact C^∞. And the associated Lie algebra $\mathfrak{gl}(n,\mathbb{R})$ comes from firstly attaching a bracket with the usual form, $[X,Y] = XY - YX$ for $X,\ Y \in M(n,\mathbb{R})$, and then equating the matrices with left-invariant vector fields through a mapping, $M(n,\mathbb{R}) \to \mathfrak{X}$, $X \mapsto \tilde{X}$ bijectively via the triple sum,

$$\tilde{X}_a := \sum_{i,\ j,\ k=1}^{n} a_{ik} X_{kj} \left(\frac{\partial}{\partial x_{ij}} \right)_a , \tag{18}$$

for each $a = (a_{ij}) \in \mathrm{GL}(n,\mathbb{R})$ and where $\{(\partial/\partial x_{ij})_a\,; 1 \leq i,j \leq n\}$ is a basis for $T_a\mathcal{M}$ which has dimension n^2[3]

The next example is the orthogonal group,

$$\mathrm{O}(n) = \{X \in M(n,\mathbb{R})\,; X^T X = I\}, \tag{19}$$

which has Lie algebra

$$\mathfrak{o}(n) := \{X \in \mathfrak{gl}(n,\mathbb{R})\,; X^T + X = O\}, \tag{20}$$

[3]Commonly the Lie algebra is studied separately from its sections, above. Following, for example, Urakawa, here we term these sections themselves the Lie algebra as it is illustrative in the present case, but since the definitions differ only by an isomorphism they are flexible in the literature and we will sometimes change conventions, assuming the context will make the definition clear.

for zero matrix $O = (0)_{i,j}$ and the special orthogonal group,

$$\mathrm{SO}(n) := \{X \in \mathrm{O}(n)\,; \det X = 1\}, \tag{21}$$

which is equipped with the same algebra as the orthogonal group,

$$\mathfrak{so}(n) := \mathfrak{o}(n) \tag{22}$$

But the structures which will provide the neatest examples in the following context of self-dual Yang-Mills theory are the unitary group,

$$\mathrm{U}(n) := \{Z \in M(n, \mathbb{C})\,; Z^*Z = I\}, \tag{23}$$

where as usual star denotes the adjoint matrix, and the special unitary group,

$$\mathrm{SU}(n) := \{Z \in \mathrm{U}(n)\,; \det Z = 1\}. \tag{24}$$

Their Lie algebras are

$$\mathfrak{u}(n) := \{Z \in M(n, \mathbb{C})\,; Z^* + Z = O\}, \tag{25}$$

and

$$\mathfrak{su}(n) := \{Z \in \mathfrak{u}(n)\,; \mathrm{tr} Z = 0\}, \tag{26}$$

respectively. Here the association of matrices to left-invariant vector fields must be carried out as in the above case of the general linear group. The adjoint representation assists in selecting a metric for on the manifold. Call the inner product, $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, is $\mathrm{ad}(G)$-invariant (and hence further bi-invariant) if whenever $X,\ Y \in \mathfrak{g}$ and $x \in G$

$$\langle \mathrm{ad}(x)X, \mathrm{ad}(x)Y \rangle = \langle X, Y \rangle . \tag{27}$$

Now for $X,\ Y \in \mathfrak{g} = \mathfrak{o}(n), \mathfrak{g} = \mathfrak{u}(n)$, or $\mathfrak{g} = \mathfrak{su}(n)$ the negative trace of the product defines $\mathrm{ad}(G)$-invariant metric

$$\langle X, Y \rangle = -\mathrm{tr} XY. \tag{28}$$

Connections

Having established a metric demand that the covariant derivative respect it via the relationship

$$d \langle X, Y \rangle = \langle DX, Y \rangle + \langle X, DY \rangle \tag{29}$$

for vector fields X and Y. Call such covariant derivatives metric. Henceforth metrics will be assumed $\mathrm{Ad}(G)$-invariant and connections assumed metric. The covariant derivative restricted from the vector fields to a Lie subalgebra has image contained in the same algebra,

$$D : \mathfrak{g} \to \mathfrak{g} \times \Omega^1(\mathcal{M}),$$

since it commutes with the commutator. Immediately we have the same quality for the curvature,

$$F_D : \mathfrak{g} \to \mathfrak{g} \times \Omega^2(\mathcal{M}),$$

or

$$F_D : \Omega^0(T\mathcal{M}; \mathfrak{g}) \to \Omega^2(T\mathcal{M}; \mathfrak{g}).$$

To investigate the action of the connection further recall its local representation in terms of the Christoffel symbols,

$$\Gamma^k_{ij} \frac{\partial}{\partial x_k} := D_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} \tag{30}$$

(summation convention applied here on the left and throughout),

$$D_{\dot{\sigma}(t)} X = \left(\dot{\xi}^i(t) + \Gamma^i_{jk}(\sigma(t)) \dot{\sigma}^j(t) \xi^k(t) \right) \frac{\partial}{\partial x^i} \tag{31}$$

for $X = \xi^i \, \partial/\partial x^i$ and $\sigma = \sigma^i \, \partial/\partial x^i$. View the Christoffel symbols as a map from the tangent space into the general linear group, $A := (\Gamma^i_{jk})_{i,j,k}$,

$$A : T\mathcal{M} \to \mathfrak{gl}(n, \mathbb{R}),$$

$$A : \dot{\sigma} \mapsto \left(\Gamma^i_{jk}(\sigma)\dot{\sigma}^j\right)_{i,k}, \tag{32}$$

or, perhaps more clearly rendered on basis vectors,

$$A\left(\frac{\partial}{\partial x^j}\right) = (\Gamma^i_{jk})_{i,k}, \tag{33}$$

where $n = \dim \mathcal{M}$ $(i,\ k = 1, ..., n)$ and $A(\dot{\sigma})$ now multiplies $X \in \mathfrak{X}$ as a matrix,

$$(A(\dot{\sigma})X)_i = (A(\dot{\sigma}))_{i,k}\,\xi^k. \tag{34}$$

But, moreover, interpret the operator A is itself (locally) a $\mathfrak{gl}$-valued one-form,

$$A \in \Omega^1(T\mathcal{M}; \mathfrak{gl}).$$

In the decomposition of the covariant derivative the first term is independent of the direction $\dot{\sigma}$ and in fact is just the exterior derivative of X at t. As such write simply,

$$D = d + A, \tag{35}$$

and call the operator A a *connection* on the tangent bundle $T\mathcal{M}$. In this formulation, the expression for the covariant derivative applied to a section of the tangent bundle, again parameterized via a local curve in $\mathcal{M}$ so that $X = \xi^i \partial/\partial x^i \in T_x\mathcal{M}$, as expressed in terms of the Christoffel symbols becomes

$$DX = (d\xi^i)\frac{\partial}{\partial x^i} + \xi^i A\left(\frac{\partial}{\partial x^i}\right). \tag{36}$$

The connection represents the curvature via the decomposition

$$\begin{aligned} F_{d+A}(X) &= (d+A)\circ(d+A)X \\ &= (d+A)(dX+AX) \\ &= d^2X + d(AX) + AdX + A(AX) \\ &= (dA)X - AdX + AdX + A\wedge AX, \end{aligned}$$

since the first term vanishes because $d^2 = 0$ and the minus sign comes from distributing the derivative via the product rule because A is a one-form. Thus

$$F_{d+A} = dA + A \wedge A. \tag{37}$$

The covariant derivative's being metric implies further that the corresponding connection is skew-symmetric,

$$A(X) \in \mathfrak{o}(n) \tag{38}$$

for every $X \in T\mathcal{M}$.

To see the result consider an orthonormal basis $\{E_1, ..., E_n\}$ for the fibres $T_p\mathcal{M}$ for each $p \in \mathcal{M}$ generated by inverting the bundle's projection and charts from an orthonomal basis of $\mathbb{R}^n$ [4], so

$$\langle E_i(p), E_j(p)\rangle = \delta_{ij}.$$

[4] For details see Jost page 38.

Observe now that since a vector field applied to any constant is zero for $X \in T_p\mathcal{M}$

$$X \langle E_i, E_j \rangle = 0.$$

Realize that the definition of a metric connection says that

$$X \langle E_i, E_j \rangle = \langle D_X E_i, E_j \rangle + \langle E_i, D_X E_j \rangle$$

or here

$$X \langle E_i, E_j \rangle = \langle A(X) E_i, E_j \rangle + \langle E_i, A(X) E_j \rangle$$

having used the fact that within a bundle chart the basis vectors are constant so for the exterior derivatived defined there

$$dE_i \equiv 0$$

for $i = 1, ..., n$, to eliminate the exterior derivative parts of the covariant derivative. Understanding that

$$A(X) = (A(X))^{i,j} : T_x\mathcal{M} \to T_x\mathcal{M}$$

is just and $n \times n$ matrix, formalize matrix multiplication in the current notation as

$$A(X) E_i = (A(X))^{i,j} E_j,$$

and now just compute

$$\begin{aligned}
0 = X\langle E_i, E_j\rangle &= \left\langle (A(X))^{i,k} E_k, E_j\right\rangle + \left\langle E_i, (A(X))^{j,k} E_k\right\rangle \\
&= (A(X))^{i,k}\langle E_k, E_j\rangle + (A(X))^{j,k}\langle E_i, E_k\rangle \\
&= (A(X))^{i,k}\,\delta_{kj} + (A(X))^{j,k}\,\delta_{ik} \\
&= (A(X))^{i,j} + (A(X))^{j,i}
\end{aligned}$$

or, for every $X \in T\mathcal{M}$,

$$A(X) = -A^T(X)$$

with the superscript denoting transposition, $(A^T)^{i,j} = (A)^{j,i}$ which is of course the definition of skew symmetry is the group structure is real and the transpose of a matrix thus corresponds with its adjoint,

$$\langle Au, v\rangle = \left\langle u, A^T v\right\rangle,$$

Thus we write in general

$$A \in \Omega^1(T\mathcal{M}; \mathfrak{o}). \tag{39}$$

The Hodge Star and Volume

Given a d-dimensional Riemannian manifold M, define an innerproduct on the p-fold exterior product $\bigwedge^p T_x^*\mathcal{M}$: For μ, $\omega \in T_x^*\mathcal{M}$,

$$\langle \mu_1 \wedge ... \wedge \mu_p, \omega_1 \wedge ... \wedge \omega_p \rangle := \det \left(\langle \mu_i, \omega_j \rangle \right)_{i,j} , \tag{40}$$

where the metric $\langle \cdot, \cdot \rangle_x = \langle \cdot, \cdot \rangle : T_x\mathcal{M} \to \mathbb{R}$ on the tangent fibre at each $x \in \mathcal{M}$, induces the innerproduct on the cotangent space, $T_x^*\mathcal{M}$. For the local-coordinate basis $\{\partial^i/\partial x^i\}_{i=1}^n$ of $T_x\mathcal{M}$ the equations

$$dx^i \left(\frac{\partial^j}{\partial x^j} \right) = \delta_{ij} \tag{41}$$

define the corresponding basis $\{dx^i\}_{i=1}^n$ for $T_x^*\mathcal{M}$ and hence an innerproduct by

$$\langle \mu, \omega \rangle = g^{ij} \mu_i \omega_j \tag{42}$$

for $\mu = \mu_i dx^i$, $\omega = \omega_i dx^i$, and where g^{ij} are the entries in the matrix inverse to that locally representing the metric,

$$g = g_{ij}\, dx^i \otimes dx^j. \tag{43}$$

The basis

$$\{dx_{i_1} \wedge ... \wedge dx_{i_p} \ ; \ 1 \leq i_1 < i_2 < ... < i_p \leq n\}$$

defines $\bigwedge^p T_x^*\mathcal{M}$.

The Hodge "star" operater

$$*: \bigwedge^{p} T_x^*\mathcal{M} \rightarrow \bigwedge^{n-p} T_x^*\mathcal{M}$$

for $0 \leq p \leq d$ is defined uniquely by the requirement that it be linear and that

$$*(e_1 \wedge ... \wedge e_p) = e_{p+1} \wedge ... \wedge e_n, \tag{44}$$

whenever $\{e_1, ..., e_n\}$ is a positive orthonormal basis of $T_x^*\mathcal{M}$ and

$$*(e_1 \wedge ... \wedge e_p) = -e_{p+1} \wedge ... \wedge e_n, \tag{45}$$

whenever $\{e_1, ..., e_d\}$ is a negative orthonormal basis of $T_x^*\mathcal{M}$. The definition of a positive basis comes from prescribing that a basis $\mathcal{B}$ be positive and defining that a basis $\mathcal{B}'$ be positive exactly if the change of basis matrix,

$$A: \mathcal{B} \rightarrow \mathcal{B}'$$

has positive determinant, $\det A > 0$.

The Hodge operator provides a reformulation of the inner product on p-forms:

$$\langle \cdot, \cdot \rangle : \bigwedge^{p} T^*\mathcal{M} \rightarrow \mathbb{R}$$

$$\langle v, w \rangle = *(w \wedge *v) = *(v \wedge *w). \tag{46}$$

To see that it truly is an innerproduct, realize first that it is bi-linear because both the wedge and star are linear. Now, given an orthornomal basis $\{e_1, ..., e_n\}$ for the cotangent bundle,

$$\mathcal{B}_p := \{e_{i_1} \wedge ... \wedge e_{i_p} \ ; 1 \leq i \leq n, \ i_1 < ... < i_p\}$$

is a basis for $\bigwedge^p T^*\mathcal{M}$. Realize that product vanishes for distinct basis vectors,

$$\begin{aligned} \langle e_{i_1} \wedge ... \wedge e_{i_p}, e_{j_1} \wedge ... \wedge e_{j_p} \rangle &= (e_{i_1} \wedge ... \wedge e_{i_p}) \wedge *(e_{j_1} \wedge ... \wedge e_{j_p}) \\ &= e_{i_1} \wedge ... \wedge e_{i_p} \wedge e_{j_{p+1}} \wedge ... \wedge e_{j_n} \\ &= \delta_{ij} * 1 \end{aligned}$$

since unless the vectors coincide the last wedge product will have a repeated element and thus be zero. This fact and bi-linearity show reflexivity, positive definiteness and symmetry and thus demonstrate that the given formula defines an innerproduct on the exterior algebra.

From multilinear algebra we know that for a p-by-p matrix A

$$Av_1 \wedge ... \wedge Av_p = \det A \, (v_1 \wedge .... \wedge v_p)$$

for any vectors $\omega_1, ..., \omega_p \in T_x^*\mathcal{M}$. Consider in particular the case where $p = n$ and $\{\omega_1, ..., \omega_n\}$ constitutes a basis for the cotangent space (so $p = d$) and A is the change-of-basis matrix relating it to an orthonormal basis $\{e_1, ..., e_p\}$

as

$$A\omega_i = e_i.$$

We can relate the n-fold exterior products of the bases vectors in terms of a kind of volume measure:

$$\begin{aligned} e_1 \wedge ... \wedge e_d &= A\omega_1 \wedge .... \wedge A\omega_n \\ &= (\det A)\, \omega_1 \wedge ... \wedge \omega_n \\ &= \sqrt{\det(\langle \omega_i, \omega_j \rangle)}\, \omega_1 \wedge ... \wedge \omega_n \end{aligned}$$

since

$$AA^T = \langle \omega_i, \omega_j \rangle.$$

Observe the immediate consequence of the definition of the operator $*$:

$$*(e_1 \wedge .... \wedge e_n) = 1$$

to write the above as

$$*1 = \frac{\omega_1 \wedge ... \wedge \omega_n}{\sqrt{\det(\langle \omega_i, \omega_j \rangle)}}.$$

Hence if $\omega_i = dx^i$ then

$$*1 = \frac{dx^1 \wedge ... \wedge dx^n}{\sqrt{\det(g^{ij})}}.$$

Recall that $g^{ij} = g_{ij}^{-1}$ and again mind the rules for determinants to uncover the so-called volume form

$$*1 = \sqrt{g}\, dx^1 \wedge ... \wedge dx^n, \tag{47}$$

defining the standard shorthand,

$$\sqrt{g} := \sqrt{\det g_{ij}}. \tag{48}$$

As such, for $\Omega \subset M$

$$\mathrm{vol}(\Omega) := \int_\Omega \sqrt{g}\, dx^1 \wedge ... \wedge dx^d = \int_\Omega *1,$$

where the integral of any continuous function $f : \mathcal{M} \to \mathbb{R}$ on the manifold is naturally defined by summing over the coordinate neighbourhoods of the system $\{(U_\alpha, \alpha); \alpha \in A\}$ as

$$\int_{U_\alpha} f * 1 := \int_{\alpha(U_\alpha)} (f \circ \alpha^{-1})\sqrt{g} dx^\alpha. \tag{49}$$

The right-hand integral is well defined over the neighbourhood $\alpha(U_\alpha) \subset \mathbb{R}^n$ with volume element $dx^\alpha = dx_1^\alpha ... dx_n^\alpha$. The extension to an arbitrary subset Ω on the manifold follows from the existence of a partition of unity $\phi_\alpha \in C^\infty(\mathcal{M}; \mathbb{R})$ by

$$\int_\Omega f * 1 := \sum_{\alpha \in A} \int_\Omega \phi_\alpha f * 1 \tag{50}$$

which is well defined since we insist that $\mathrm{supp}(\phi_\alpha) \subset U_\alpha$.

Note here the formula for the autocomposition of the Hodge operator, $** : \bigwedge^p T_x^*\mathcal{M} \to \bigwedge^p T_x^*\mathcal{M}$,

$$** = (-1)^{p(n-p)}. \tag{51}$$

We know that for an orthonormal vectors $\{e_1, ..., e_p\}$ in $T_x^*\mathcal{M}$

$$*(e_1 \wedge ... \wedge e_p) = e_{p+1} \wedge ... \wedge e_n.$$

such that the orthonormal basis $\{e_1, ..., e_p, e_{p+1}, e_n\}$ is (defined to be) positive. And for the orthonormal vectors $\{e_{p+1}, ..., e_n\}$,

$$*(e_{p+1} \wedge ... \wedge e_n) = (\det A)\, e_1 \wedge ... \wedge e_p$$

where $A : T_x^*\mathcal{M} \to T_x^*\mathcal{M}$ is the change-of-basis matrix that reorders the basis as

$$A : \{e_1, ..., e_p, e_{p+1}, ..., e_n\} \mapsto \{e_{p+1}, ..., e_n, e_1, ..., e_p\}.$$

The sign of the determinant by definition determines whether $\{e_{p+1}, ..., e_n, e_1, ..., e_p\}$ is positive (with respect to the convention established by asserting $\{e_1, ..., e_p, e_{p+1}, ..., e_n\}$ to be positive). If $\{e'_1, ..., e'_n\}$ is a negative basis—the image of the positive basis under a transformation with negative determinant—then in general we have

$$*(e'_1 \wedge ... \wedge e'_p) = -e_{p+1} \wedge ... \wedge e_n.$$

But reordering the wedge product exchanging the place of the p-form and the $(n-p)$-form gives

$$e_1 \wedge ... \wedge e_p \wedge e_{p-1} \wedge ... \wedge e_n = (-1)^{p(n-p)} e_{p+1} \wedge ... e_n \wedge e_1 \wedge ... \wedge e_p, \tag{52}$$

by the standard formula for such rearrangements. And since

$$\begin{aligned} *(e_1 \wedge ... \wedge e_p \wedge e_{p-1} \wedge ... \wedge e_n) &= *(Ae_{p+1} \wedge ... Ae_n \wedge Ae_1 \wedge ... \wedge Ae_p) \\ &= \det A\,(e_{p+1} \wedge ... \wedge e_n \wedge e_1 ... \wedge e_p), \end{aligned}$$

we conclude that $\det A = (-1)^{p(n-p)}$ and thus

$$\begin{aligned} *(e_1 \wedge ... \wedge e_p) &= (\det A)\, e_1 \wedge ... \wedge e_p \\ &= (-1)^{p(n-p)}\, e_1 \wedge ... \wedge e_p, \end{aligned}$$

which says $** = (-1)^{p(n-p)}$ as hoped.

Now define the global L^2-innerproduct of $\mu, \omega \in T_x^*\mathcal{M}$ as

$$(\mu, \omega) := \int_M \langle \mu, \omega \rangle * 1$$

and use the star formulation of the innerproduct to note that if $p = n/2$—so

$* : \bigwedge^p T^*\mathcal{M} \to \bigwedge^p T^*\mathcal{M}$—then

$$\begin{aligned}\langle *(e_{i_1} \wedge ... \wedge e_{i_p}), *(e_{j_1} \wedge ... \wedge e_{j_p})\rangle &= *\big(*(e_{i_1} \wedge ... \wedge e_{i_p}) \wedge *(*(e_{j_1} \wedge ... \wedge e_{j_p}))\big) \\ &= *\big(*(e_{i_1} \wedge ... \wedge e_{i_p}) \wedge (e_{j_1} \wedge ... \wedge e_{j_p})\big),\end{aligned}$$

having exploited the fact that $**$ is the identity in this case, here again observe that if the basis vectors are distinct then the product vanishes.

Contrariwise if vectors above are the same then

$$\begin{aligned}*(e_{i_1} \wedge ... \wedge e_{i_p}) \wedge (e_{i_1} \wedge ... \wedge e_{i_p}) &= e_{i_1} \wedge ... \wedge e_{i_n} \\ &= (e_{i_1} \wedge ... \wedge e_{i_p}) \wedge *(e_{i_1} \wedge ... \wedge e_{i_p})\end{aligned}$$

so

$$\begin{aligned}\langle *(e_{i_1} \wedge ... \wedge e_{i_p}), *(e_{j_1} \wedge ... \wedge e_{j_p})\rangle &= *\big((e_{i_1} \wedge ... \wedge e_{i_p}) \wedge *(e_{i_1} \wedge ... \wedge e_{i_p})\big) \\ &= \langle e_{i_1} \wedge ... \wedge e_{i_p}, e_{i_1} \wedge ... \wedge e_{i_p}\rangle.\end{aligned}$$

Linearity extends this to

$$\langle *v, *w\rangle = \langle v, w\rangle.$$

Likewise

$$(*v, *w) = \int_M \langle *v, *w\rangle * 1 = \int_M \langle v, w\rangle * 1 = (v, w),$$

so we can add as a corollary that $*$ is an L^2-isometry:

$$||*v||_{L^2} := (*v, *v) = (v, v) = ||v||_{L^2}.$$

A Gernalized Domain for Covariant Derivatives and the Second Bianchi Identity

While here we will stick mostly to covariant derivatives oporating on vector fields—sections of the tangent bundle—the operator as defined functions identically on the space of sections of any vector bundle $(E, \pi, \mathcal{M})$, where E is a vector space and π its projection onto $\mathcal{M}$,

$$\Gamma(E) := \{s \in C^1(\mathcal{M}; E) \ ; \ \pi \circ s = \mathrm{Id}_{\mathcal{M}}\}. \tag{53}$$

And hereafter we write

$$\Omega^p(E) := \Gamma(E) \times \Omega^p(\mathcal{M}), \tag{54}$$

($p \leq \dim \mathcal{M}$) which agrees with our current definition of $\Omega^p(T\mathcal{M})$ since $\Gamma(T\mathcal{M}) = \mathfrak{X}(\mathcal{M})$.

Now we wish to extend the covariant derivative to this space $\Omega^p(E)$ of sections crossed with forms. Motivated by the requirement the the covariant derivative satisfy a Liebnitz product and expoiting the established exeritor derivative for forms, take for $X \in \Gamma(E)$ and $\omega \in \Omega^p(\mathcal{M})$ that

$$D : \Gamma(E) \times \Omega^p(\mathcal{M}) \to \Gamma(E) \times \Omega^{p+1}(\mathcal{M}),$$

$$D(X \otimes \omega) := DX \wedge \omega + X \otimes d\omega \tag{55}$$

with the understanding that

$$(X \otimes \omega_1) \wedge \omega_2 := X \otimes (\omega_1 \wedge \omega_2),$$

whenever $X \in \Gamma(E)$ and $\omega_1,\ \omega_2 \in \Omega^p(\mathcal{M})$.

Furthermore, for distinct bundles E_1 and E_2 with assosiated covariant derivatives D_1 and D_2 respectively, define the covariant derivative on $E_1 \times E_2$ via

$$D(X \otimes Y) := D_1 X \otimes Y + X \otimes D_2 Y \tag{56}$$

whenever $X \in E_1$ and $Y \in E_2$. In particular this defines a covariant derivative D on the space $E \otimes E^*$: For $X := \xi^i_j v_i \otimes \omega^j \in \Gamma(E \otimes E^*)$

$$\begin{aligned} DX = (d + A)X &= dX + A(\xi^i_j\, v_i \otimes \omega^j) \\ &= dX + \left(A^i_j \xi^j_k\, v_i \otimes \omega^k - A^i_j \xi^k_i\, v_k \otimes \omega^j\right). \end{aligned}$$

But the last term is just the Lie bracket of the connection with th section, so

$$DX = dX + [A, X]. \tag{57}$$

Call $E \otimes E^* =: \mathrm{End}E$, motivated by the fact that each $\omega \in E^*$ is an endomorphism on the bundle E. Define further

$$\mathrm{Ad}E := \{T : E \hookrightarrow E\ ;\ T \text{ is linear, } T^* = -T\} \tag{58}$$

or $\mathrm{Ad}E$ is the subset of $\mathrm{End}E$ for which the endomorphism for each fibre is skew symmetric. From the previous section we know a connection A is skew symmetric in the sense that $A(X) \in \mathfrak{o}$ so for $D = d + A$ we have

$$A \in \Omega^1(\mathrm{Ad}E). \tag{59}$$

Now view the curvature tensor,

$$F_D : \Gamma(E) \to \Omega^2(E) = (\Gamma(E))^* \times \Omega^2(\mathcal{M}),$$

as a two-form assuming values in $E \otimes E^*$,

$$F_D \in \Gamma(E) \otimes (\Gamma(E))^* \otimes \Omega^2(\mathcal{M}),$$

and apply the result to commute the connection of the curvarture:

$$\begin{aligned} DF_D &= DF_D + [A, F] \\ &= d^2 + dA \wedge A - A \wedge dA + [A, dA + A \wedge A] \\ &= dA \wedge A - A \wedge dA + A \wedge dA - dA \wedge A + [A, A \wedge A] \\ &= [A, A \wedge A] \\ &= [A_i dx^i, A_j dx^j \wedge A_k dx^k] \\ &= A_i A_j A_k (dx^i \wedge dx^j \wedge dx^k - dx^j \wedge dx^k \wedge dx^i) = 0. \end{aligned}$$

Revealing the formula

$$DF_D \equiv 0, \tag{60}$$

termed the Second Bianchi Identity, which will prove usuful in our initial demonstration of the selfduality of the Yang-Mills equations in four dimensions.

Chapter 3: The Yang-Mills Functional

Define the here our central subject, the Yang-Mills functional, which is the norm of curvature over a manifold viewed as a function of connections with a given Lie structure. Physically, the Yang-Mills connections are stationary points of the field strength.

This second chapter examines the functional and its critical points, eventually examining our main focus of the functional's absolute minimizers in four dimensions and the corresponding equations for the, the selfdual and anti-selfdual equations, $F_D = *F_D$ and $F_D = - * F_D$.

The Functional Defined and the Yang-Mills Connections

Define an inner product for $A,\ B \in \Omega^p(T\mathcal{M};\mathfrak{g})$, where $A = X \otimes \omega$ and $B = Y \otimes \nu$ for $X,\ Y \in \mathfrak{g}$ and $\omega,\ \nu \in \Omega^p(\mathcal{M})$, via

$$\begin{aligned}\langle A, B\rangle &= \langle X \otimes \omega, Y \otimes \nu\rangle \\ &= \langle X, Y\rangle_{\mathfrak{g}} \langle \omega, \nu\rangle_{\Lambda^p T_x^* \mathcal{M}} \\ &= \langle X, Y\rangle_{\mathfrak{g}} * (\omega \wedge *\nu).\end{aligned}$$

So in the cases outlined (the algebras $\mathfrak{o}$, $\mathfrak{u}$, and $\mathfrak{su}$),

$$\langle A, B\rangle = -\mathrm{tr} XY * (\omega \wedge *\nu).$$

And analogously to above the L^2 scalar product is

$$(A, B)_{L^2} := \int_{\mathcal{M}} \langle A, B \rangle * 1.$$

Thus finally define the Yang-Mills functional as the L^2-norm of the curvature—$|\cdot|^2 := \langle \cdot, \cdot \rangle$—as

$$\mathcal{YM} : \Omega^1(T^*\mathcal{M}; \mathfrak{o}) \to \mathbb{R},$$

$$\mathcal{YM}(A) = \int_{\mathcal{M}} |F_{d+A}|^2 * 1 = \int_{\mathcal{M}} \langle F_{d+A}, F_{d+A} \rangle * 1. \tag{61}$$

Sometimes (as below) it will be easier to view this as a function of the covariant derivative rather than the connection, and when no ambiguity will result we will employ the same notation—$\mathcal{YM}(d + A) := \mathcal{YM}(A)$.

The objective is to choose a connection that is stationary with respect to this square energy. The traditional approach is variational. For covariant derivatives D and $\tilde{D} \in \Omega^1(T\mathcal{M}; \mathfrak{g})$ and a vector field $X \in \mathfrak{g}$ consider

$$\begin{aligned} F_{D+t\tilde{D}}X &= (D + t\tilde{D}) \circ (D + t\tilde{D})X \\ &= D^2X + tD(\tilde{D}X) + t\tilde{D} + t\tilde{D} \wedge DX + t^2\tilde{D} \wedge \tilde{D}X \\ &= (F_D + tD\tilde{D} + t^2\tilde{D} \wedge \tilde{D})X, \end{aligned}$$

having employed the fact that $D(\tilde{D}X) = (D\tilde{D})X - \tilde{D} \wedge DX$. Now take the variational derivative of the functional in order to find the conditions of the stationary points:

$$\begin{aligned}\delta\mathcal{YM}(D) &= \frac{d}{dt}\Big|_{t=0}\mathcal{YM}(D+t\tilde{D}) \\ &= \frac{d}{dt}\Big|_{t=0}\int_{\mathcal{M}}\langle F_{D+t\tilde{D}}, F_{D+t\tilde{D}}\rangle * 1 \\ &= 2\int_{\mathcal{M}}\left\langle D\tilde{D}, F_D\right\rangle * 1,\end{aligned}$$

since $\langle F_D, tD\tilde{D}\rangle$ and $\langle tD\tilde{D}, F_D\rangle$ are the only first-order terms in the expansion of the scalar product. Thus setting $\delta\mathcal{YM}(D) = 0$ yields the equation

$$(D\tilde{D}, F_D)_{L^2} = 0,$$

for all covariant derivatives $\tilde{D}$, for the functional's critical points. Given an arbitrary covariant derivative $D : \Omega^0 \to \Omega^1(T_x\mathcal{M})$ introduce the operator $D^* : \Omega^1(T_x\mathcal{M}) \to \Omega^0$, termed the dual covariant derivative to D, defined by the L^2 relationship

$$(D^*X, Y)_{L^2(\Omega^0)} = (X, DY)_{L^2(\Omega^1)}, \tag{62}$$

for every $X \in \Omega^1(T\mathcal{M})$ and $Y \in \Omega^0(T\mathcal{M})$, in order to rewrite the characterization of the functional's stationary points as

$$(\tilde{D}, D^*F_D)_{L^2} = 0$$

for every $\tilde{D}$, or better,

$$D^* F_D = 0. \tag{63}$$

Covariant derivatives satisfying this equation are called Yang-Mills covariant derivatives. (Likewise A is a Yang-Mills connection if $d+A$ solves the above.)

A Representation of Yang-Mills Connections

Returning to the decomposed representation of the covariant derivative, $D = d + A$, express the connection in components as $A = A_i dx^i$, where

$$A(X) = A_i dx^i(X),$$

$$A_i \in \mathfrak{gl}(n).$$

The connection acts on a vector field Y via exterior differentiation and exterior product with the connection,

$$DY = (d + A)Y = dY + A_i dx^i \wedge Y.$$

Utilise the skew symmetry of the connection to write the rewrite the dual covariant derivative simply in terms of the adjoint to the exterior derivative:

$$(X, DY) = (X, dY + A_i dx^i \wedge Y) = (d^* X, Y) - (A_i X, dx^i \wedge Y). \tag{64}$$

Now represent the curvature as

$$F = F_{ij} dx^i \wedge dx^j$$

in normal coordinates, i.e. the Kronecker delta represents the metric and the Christoffel symbols vanish,

$$g_{ij} = \delta_{ij},$$

$$\Gamma^i_{jk} \equiv 0.$$

And here the one-form d^*F is

$$d^*F = d^*(F_{ij} dx^i \wedge dx^j) = -\frac{\partial F_{ij}}{\partial x^i} dx^j, \tag{65}$$

understanding that summation is taken over i as well as j despite the indices residing "on the same level"—i.e. both being formally contravariant.

Substitute this formula into the representation of the L^2 innerproduct involving the covariant derivative in terms of its canonical decomposition to get

$$\begin{aligned}
(F, DY) &= (d^*F, Y) - (A_i F, dx^i \wedge Y) \\
&= \big(-\frac{\partial F_{ij}}{\partial x^i} dx^j, Y\big) - \big(A_k F_{ij} dx^i \wedge dx^j, dx^k \wedge Y\big) \\
&= \Big(-\frac{\partial F_{ij}}{\partial x^i} dx^j, Y\Big) - \Big(A_i (F_{ij} - F_{ji}) dx^j, Y\Big) \\
&= \Big(-\frac{\partial F_{ij}}{\partial x^i} dx^j, Y\Big) - \Big([A_i, F_{ij}]\, dx^j, Y\Big),
\end{aligned}$$

where having executed the summation in normal coordinates eliminated the k-index and reduced the right-hand innerproduct to agree with the right-hand one in domain and here the Lie bracket denotes exactly the symmetricness of the curvature as

$$[A_i, F_{ij}] = A_i F_{ij} - F_{ij} A_j.$$

So

$$D^* F = -\left(\frac{\partial F_{ij}}{\partial x^i} + [A_i, F_{ij}]\right) dx^j$$

and A is a Yang-Mills connection if

$$\frac{\partial F_{ij}}{\partial x^i} + [A_i, F_{ij}] = 0 \tag{66}$$

for each $j = 1, \ldots, n$ (of course F_{ij} corresponds to A as $F_{d+A} = F_{ij} dx^i \wedge dx^j$).

The Example of Two Dimensions

If $n = \dim \mathcal{M} = 2$ then the orthogonal group is Abelian, if $A,\ B \in \mathfrak{o}(2, \mathbb{R})$ then

$$AB = BA.$$

The determinant of an orthogonal matrix is always plus or minus one,

$$|\det A| = 1$$

for all $A \in \mathfrak{o}(n)$ since, by definition,

$$AA^T = I,$$

so

$$\det AA^T = \det A \det A^T = \det I = 1$$

and

$$\det A = \det A^T$$

implies

$$(\det A)^2 = 1$$

which gives the desired fact. Thus, in two dimensions, where,

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix},$$

the equation $A^{-1} = A^T$ implies

$$\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \pm \begin{pmatrix} a_{11} & a_{21} \\ -a_{12} & a_{22} \end{pmatrix}$$

which says every $A \in \mathfrak{o}(2)$ has either the form

$$\begin{pmatrix} s & t \\ -t & s \end{pmatrix} \text{ or } \begin{pmatrix} s & t \\ t & -s \end{pmatrix}.$$

Having established this there just for cases to compute directly to see that the group is Abelian. Of course that the Lie bracket vanishes identically is an immediate and trivial consequence,

$$AB - BA = AB - AB = 0.$$

We call such a Lie algebra trivial. And thus the skew-symmetric bundle $\mathrm{Ad}(T\mathcal{M}) \subset T\mathcal{M} \times T^*\mathcal{M}$ is also trivial, which means it is isomorphic to the direct product of the manifold with the real numbers, we say (writing equality for short)

$$\mathrm{Ad}(T\mathcal{M}) = \mathcal{M} \times \mathbb{R}.$$

In this representation the covariant derivative coincides with the exterior derivative,

$$D = d \tag{67}$$

or the connection vanishes locally—its derivatives do not vanish, since the representation

$$D_\alpha = d + A_\alpha$$

depends on the coordinate system $\{(U_\alpha, \alpha); \alpha \in A\}$ (mind that this set of charts A is unrelated to the connection A), and this decomposition is not in general global. In this setting the fundamental equations all simplify greatly. $A \wedge A = 0$ reduces the curvature to

$$F_{d+A} = dA + A \wedge A = dA \tag{68}$$

and thus here the Bianchi identity follows immediately from the fact that the autocomposition of the exterior derivative vanishes identically

$$DF_D = dF_D = d(dA) = d^2 A = 0 \tag{69}$$

Lastly, the local absence the connection reduces dual covariant derivative reduces to the adjoint exterior derivative,

$$D^* = d^*$$

so the Yang-Mills equations read

$$d^* F = 0.$$

Employing above identity for the curvature expands this to

$$d^* dA = 0.$$

This low-dimensional context unifies Yang-Mills theory with the study of harmonic forms since the Laplacian on forms $\Delta : \Omega^p(\mathcal{M}) \to \Omega^p(\mathcal{M})$ is defined as

$$\Delta := dd^* + d^*d. \tag{70}$$

So trivially the curvature is harmonic in this context,

$$\Delta F_D = 0.$$

Furthermore, if, without motivation, we assume the so-called gauge condition,

$$d^*A = 0, \tag{71}$$

and we have immediately

$$dd^*A = 0.$$

Hence the the above sequence of equations shows that the connection A is also automatically harmonic

$$\Delta A = (dd^* + d^*d)A = dd^*A + d^*dA = 0.$$

A Preliminary Fact

Recall Stokes's Theorem for forms, for any smooth $(n-1)$-form ω with compact support,

$$\int_{\mathcal{M}} d\omega * 1 = \oint_{\partial\mathcal{M}} \omega * 1,$$

(the Hodge operators of course correspond the respective cotangent bundles, $T * \mathcal{M}$ and $T^*\partial\mathcal{M}$, to generate the appropriate volume elements) and, moreover, since we are considering a manifold without boundary, $\partial\mathcal{M} = \emptyset$,

$$\int_{\mathcal{M}} d\omega * 1 = 0.$$

Now, for $\alpha \in \bigwedge^{p-1} T^*\mathcal{M}$ and $\beta \in \bigwedge^{p} T^*\mathcal{M}$, apply the formula to the $(n-1)$-form that comes from taking the wedge product of a $(p-1)$-form α and $(n-p)$-form $*\beta$:

$$\int_{\mathcal{M}} d(\alpha \wedge *\beta) * 1 = 0.$$

Minding the product rule for forms and keeping in mind that since $d * \beta \in \bigwedge^{n-p+1} T^*\mathcal{M}$

$$** (d * \beta) = (-1)^{(p-1)(n-p+1)} d * \beta,$$

we can compute

$$\begin{aligned} d(\alpha \wedge *\beta) &= d\alpha \wedge *\beta + (-1)^{p-1} \alpha \wedge d * \beta \\ &= d\alpha \wedge *\beta + (-1)^{p-1}(-1)^{(p-1)(n-p+1)} \alpha \wedge ** d * \beta. \end{aligned}$$

The exponent expands to $n(p-1) - p^2 + 3p - 2$ and because p and p^2 always have the same parity, we can cancel most of the terms and write

$$(-1)^{p-1}(-1)^{(p-1)(n-p+1)} = (-1)^{n(p-1)}.$$

The autocomposition of the Hodge star on the space of n-forms is the identity (assuming a positive basis as throughout), so carry on by employing this along with just linearity and the definition of the innerproduct

$$\begin{aligned}
d\alpha \wedge *\beta + (-1)^{n(p-1)}\alpha \wedge **d*\beta &= **\left(d\alpha \wedge *\beta + (-1)^{n(p-1)}\alpha \wedge **d*\beta\right) \\
&= *\big(*(d\alpha \wedge *\beta) + (-1)^{n(p-1)} *(\alpha \wedge **d*\beta)\big) \\
&= *\big(\langle d\alpha, \beta\rangle + (-1)^{n(p-1)} \langle \alpha, *d*\beta\rangle \big).
\end{aligned}$$

But by Stokes's Theorem we can say that this integrates to zero,

$$\int_{\mathcal{M}} d(\alpha \wedge *\beta) * 1 = \int_{\mathcal{M}} *\big(\langle d\alpha, \beta\rangle - \big\langle \alpha, (-1)^{n(p-1)+1} * d * \beta \big\rangle \big) * 1 = 0.$$

Since this holds for any $(p-1)$-form α and p-form β, the integrand must be zero and the above is in fact a statement about the relationship between the exterior derivative and the Hodge star,

$$\langle d\alpha, \beta\rangle + (-1)^{n(p-1)} \langle \alpha, *d*\beta\rangle = 0$$

or

$$\langle d\alpha, \beta\rangle = \big\langle \alpha, (-1)^{n(p-1)+1} * d * \beta \big\rangle,$$

which reads like the definition of the adjoint exterior derivative and says exactly that

$$d^* = (-1)^{n(p-1)+1} * d * . \tag{72}$$

The result extends to covariant derivatives. The Hodge operator acts on elements of $X \otimes \omega \in \Omega^p(T\mathcal{M})$, with $X \in \mathfrak{X}$ and $\omega \in \bigwedge^p(T^*\mathcal{M})$, as

$$*(X \otimes \omega) = X \otimes *\omega, \tag{73}$$

which is to say it is defined to act normally on the form but leave the vector field alone. In contrast recall that the image of connection A belongs to the orthogonal group, $A(X) \in \mathfrak{o}(n)$ for each $X \in T\mathcal{M}$, in the sense that for $A = A_i dx^i$ the map

$$A_i : T_x\mathcal{M} \to T_x\mathcal{M}$$

is skew-symeetric. Thus the matrix A_i acts only on the vector field but leaves the form alone,

$$A_i(X \otimes \omega) = A_i X \otimes \omega. \tag{74}$$

As such the operators commute, $*A_i = A_i$ folloing from the simple computation,

$$\begin{aligned}(*A_i)(X \otimes \omega) &= *(A_i X \otimes \omega) \\ &= A_i X \otimes *\omega \\ &= A_i(X \otimes *\omega) \\ &= (A_i *)(X \otimes \omega)\end{aligned}$$

and in particular, again since $** = (-1)^{p(n-p)}$, or $(-1)^{p(n-p)+1} ** = \mathrm{Id}$,

$$A_i = (-1)^{p(n-p)+1} * * A_i = (-1)^{p(n-p)+1} * A_i * .$$

Return to the formula for the dual covariant derivative D^*,

$$(D^* X, Y) = (d^* X, Y) - (A_i X, dx^i \wedge Y) \tag{75}$$

and manipulate to get

$$(D^* X, Y) = (-1)^{n(p-1)+1}(*d * X, Y) - (-1)^{p(n-p)+1}(*A_i * X, dx^i \wedge Y).$$

Now suppose that the manifold is of even dimension and look for the dual covariant derivative of a form of even order—that is n and p above are even. In the next, especially pertinent section, we will be narrow our focus to the situation when $n = 4$ and want to apply D^* to the curvature, which is a two-form. So, to resume, write

$$(D^*X, Y) = (- * d * X, Y) + (- * A_i * X, dx^i \wedge Y) \tag{76}$$

and we can say

$$D^*X = - * (d + A)* = - * D*, \tag{77}$$

as long as we understand that for $X \otimes \omega \in \Omega^p(T\mathcal{M})$ such that $X \in \mathfrak{X}$ and $\omega \in \Omega^p(\mathcal{M})$ here the connection A applies to the "form part" ω via contraction, not multiplication, that is to say exactly that

$$(A(X \otimes \omega), Y) = (A_i X \otimes \omega, dx^i \wedge Y)$$

for every $Y \in \Omega^{p-1}(T\mathcal{M})$; that the dual covariant derivative must decrease the order of the form part by one motivates this view.

Yang-Mills in Four Dimensions: The Selfdual Equations

Recall that in four dimensions the Hodge star is an L^2 isometry on the space of two-forms,

$$* : \bigwedge^2 T^*\mathcal{M} \to \bigwedge^2 T^*\mathcal{M}.$$

Given normal coodinates about x such that $\{dx^1,\ dx^2,\ dx^3,\ dx^4\}$ describes a basis for the fibre, define

$$\Lambda^+ := \mathrm{span}\{dx^1 \wedge dx^2 + dx^3 \wedge dx^4,\ dx^1 \wedge dx^4 + dx^2 \wedge dx^3,\ dx^1 \wedge dx^3 - dx^2 \wedge dx^4\}$$

and

$$\Lambda^- := \mathrm{span}\{dx^1 \wedge dx^3 + dx^2 \wedge dx^4,\ dx^1 \wedge dx^2 - dx^3 \wedge dx^4,\ dx^1 \wedge dx^4 - dx^2 \wedge dx^3\},$$

and realize that since the six vectors described above are independent and $\bigwedge^2 T^*\mathcal{M}$ has dimension six,

$$\dim\left(\bigwedge^p T^*\mathcal{M}\right) = \binom{\dim(\mathcal{M})}{p} = \binom{4}{2} = 6,$$

in our case of two-forms. Thus:

$$\Lambda^+ \oplus \Lambda^- = \bigwedge^2 T^*\mathcal{M}. \tag{78}$$

The division of the fibre into the prescribed subspaces speaks to the Hodge star. Computation minding the behaviour of the operator reveals

$$*\alpha = \pm\alpha \tag{79}$$

for $\alpha \in \Lambda^+$ or $\alpha \in \Lambda^-$—that is to say that the decomposition is into the eigenspaces of $*$ corresponding to the eigenvalues 1 and -1.

Interestingly, this splitting corresponds to the splitting of the special orthogonal group in the sense of isomophism

$$\Lambda^+ \cong \Lambda^- \cong \mathfrak{so}(3)$$

and

$$\bigwedge^2 T^*\mathcal{M} \cong \mathfrak{so}(4) = \mathfrak{so}(3) \oplus \mathfrak{so}(3).$$

Here we will turn to covariant derivatives D with curvature tensors $F_D = *F_D \in \Lambda^+$ called selfdual and $F_D = - * F_D \in \Lambda^-$ called anti-selfdual. Furthermore, call a connection inducing selfdual curvature an instanton and one inducing anti-selfdual curvature an anti-instanton.

The Bianchi formula

$$DF_D = 0$$

implies that

$$D * F_D = 0 \tag{80}$$

for $F_D \in \Lambda^+$ or $F_D \in \Lambda^-$. And hence

$$*D * F_D = 0. \tag{81}$$

But the main result of the previous section says that

$$*D* = -D^*$$

in this case, and we can now say

$$D^* F_D = 0. \tag{82}$$

Or *all instantons and anti-instantons are Yang-Mills connections.*

Cohomology and Chern Classes

Our nicest examples of Yang-Mills theory will pertain to the covariant derivatives on the compact subgroups $\mathrm{U}(m)$ and $\mathrm{SU}(m)$ of the complex general linear group $\mathrm{GL}(m, \mathbb{C})$, so return to the abstract setting define it on an arbitrary vector bundle E with rank m. In this context we will be able to rewrite the Yang-Mills functional elegantly and read off its minimizers.

To begin we need the concept of Chern classes and thus of cohomology. We say that two p-forms, $\alpha,\ \beta \in \Omega^p(\mathcal{M})$, are cohomologeous if their difference is exact, that is, there exists a $(p-1)$-form, $\gamma \in \Omega^{p-1}(\mathcal{M})$, such that

$$\alpha - \beta = d\gamma.$$

This cohomology relation is equivalence relation that partitions the space $\{\alpha \in \Omega^p(\mathcal{M})\ ;\ d\alpha = 0\}$ of closed forms in $\Omega^p(\mathcal{M})$. The set of all equivalence classes,

$$[\alpha] := \{\beta \in \Omega^p(\mathcal{M})\ ;\ \alpha - \beta \text{ is exact, } d\beta = 0\},$$

itself defines a vector space,

$$H^p(\mathcal{M}) := \{[\alpha] \ ; \ \alpha \in \Omega, \ d\alpha = 0\}, \tag{83}$$

called the p-th de Rham cohomology group.

The Chern classes are such equivalence classes belonging to such a group that depend, for our definition, on the elementary symmetric polynomials,

$$p^j(\lambda_1, ..., \lambda_m) := \sum_{1 \leq a_1 < ... < a_j \leq m} \lambda_{a_1} \cdots \lambda_{a_j}, \tag{84}$$

or, more precisely, on the matrix polynomials,

$$P^j : M(m, \mathbb{C}) \to \mathbb{C}$$

where for $B \in M(m, \mathbb{C})$

$$P^j(B) := p^j(\lambda_1, ..., \lambda_m) \tag{85}$$

and $\lambda_1, ..., \lambda_m \in \mathbb{C}$ are the (ordered) eigenvalues of B. We have the essential property that the polynomials represent in the expansion over the product of a first-order monomials, for $t \in \mathbb{C}$,

$$\prod_{j=1}^{m} (t - \lambda_j) = p^j t^{m-j},$$

having employed the shorthand $p^j := p^j(\lambda_1, ..., \lambda_m)$. The identity looks the same carried over to the matrix case,

$$\prod_{j=1}^{m}(t-\lambda_j) = P^j(B)t^{m-j}. \tag{86}$$

But here realize that of course

$$\prod_{j=1}^{m}(t-\lambda_j) = 0$$

exactly when t is an eigenvalue of B. But for t to be an eigenvalue by definition

$$B\alpha = t\alpha,$$

for each $\alpha \in \mathbb{C}^m$ which reformulates into the familiar equation

$$\det(B - t\mathrm{Id}) = 0. \tag{87}$$

Since these polynomials have the same roots and both have leading coefficient one they must be equal,

$$P^j(B)t^{m-j} = \det(B - t\mathrm{Id}), \tag{88}$$

and we have a tool, the determinant, with which to compute elementary symmetric polynomials.

These polynomials are homogenous with degree j and thus take map

p-forms to jp-forms. In particular we here consider the curvature as a homomorphism,

$$F_D \in \mathrm{Ad}E = \mathrm{End}E = \mathrm{Hom}_{\mathbb{C}}(E; E) = M(m, \mathbb{C}),$$

so,

$$\begin{aligned} P^j : \Omega^2(\mathcal{M}) &\rightarrow \Omega^{2j}(\mathcal{M}), \\ P^j(F_D) &\in \Omega^{2j}(\mathcal{M}). \end{aligned} \tag{89}$$

These polynomials are in fact exact,

$$dP^j(F_D) = 0 \tag{90}$$

and further are independent of the covariant derivative[5],

$$P^j(F_{D_1}) = P^j(F_{D_2}) \tag{91}$$

for any covariant derivatives D_1, $D_2 : \Omega \rightarrow \Omega^1(E)$.

So, independently of the covariant derivative—$F := F_D$ for arbitrary D—we can define the elements of the $2j$-th cohomology group,

$$c_j(E) := \left[P^j\left(\frac{i}{2\pi}F\right)\right] \in H^{2j}(\mathcal{M}), \tag{92}$$

[5]See Jost, pages 125-126 for a proof.

called the Chern classes of the bundle E.

Computation of Chern Classes

Now, to compute the classes, exploit the formula

$$\det\left(\frac{i}{2\pi}F_D + t\mathrm{Id}\right) = \sum_{j=0}^{m} P^j\left(\frac{i}{2\pi}F_D\right)t^{m-j},$$

or, in terms the curvature of an arbitrary covariant derivative,

$$\sum_{j=0}^{m} c_j(E)t^{m-j}\det\left(\frac{i}{2\pi}F_D + t\mathrm{Id}\right). \tag{93}$$

To simplify, divide by t^m and remember that the determinant is m-homogeneous,

$$\begin{aligned}\frac{1}{t^m}\sum_{j=0}^{m} c_j(E)t^{m-j} &= \sum_{j=0}^{m} c_j(E)t^{-j}\\ &= \frac{1}{t^m}\det\left(\frac{i}{2\pi}F_D + t\mathrm{Id}\right)\\ &= \det\left(\frac{i}{2\pi t}F_D + \mathrm{Id}\right).\end{aligned}$$

Realize that the curvature's eigenvalues are two-forms since $F_D : \Omega^0 \to \Omega^2(E)$. Λ_j is an eigenvalue of F_D when for $X \in \Gamma(E)$,

$$F_D(X) = \Lambda_j X \in \Omega^2(E). \tag{94}$$

This extends naturally since the roots $t \in \mathbb{C}$ and leading coefficient of the equation

$$\prod_{j=1}^{m}\left(t\mathrm{Id} - \frac{i}{2\pi}\Lambda_j\right) = 0 \tag{95}$$

again coincide with those of the determinant above. Thus, replacing Id with 1 in the product to agree with the formalism,

$$\det\left(\frac{i}{2\pi t}F_D + \mathrm{Id}\right) = \prod_{j=1}^{m}\left(t - \frac{i}{2\pi}\Lambda_j\right). \tag{96}$$

Moreover we will exploit the fact from linear algebra that the trace is the sum of the eigenvalues,

$$\mathrm{tr}B = \sum_{j=1}^{m}\lambda_j$$

to compute the Chern classes via the expression

$$\sum_{j=0}^{m} c_j\tau^j = \prod_{j=1}^{m}\left(1 - \frac{i}{2\pi}\tau\Lambda_j\right), \tag{97}$$

where we have substituted $\tau := 1/t$.

Viz, for $m = \mathrm{rank}(E) = 1$,

$$\begin{aligned}\sum_{j=0}^{m} c_j(E)\tau^j &= c_0(E) + c_1(E)\tau \\ &= \mathrm{Id} - \frac{i}{2\pi}\Lambda_1\tau.\end{aligned}$$

But the zeroth-order symmetric polynomial is one, here $c_0(E) = 1$. Thus

$$c_1(E) = \frac{i}{2\pi}\Lambda_1$$

since all other terms above cancel. Furthermore, since with $\mathrm{rank}(E) = 1$, $B : E \to E$, and hence likewise $F_D : \Omega^0 \to \Omega^2(E)$, has only (exactly) one eigenvalue must be equal to its trace,

$$\mathrm{tr}F_D = \Lambda_1,$$

or, moreover,

$$c_1(E) = \frac{i}{2\pi}\mathrm{tr}F. \tag{98}$$

To find $c_2(E)$ consider the case in which $m = 2$ so

$$\begin{aligned} c_0(E) + c_1(E)\tau + c_2(E)\tau^2 &= \left(1 + \frac{i}{2\pi}\Lambda_1\tau\right)\cdot\left(1 + \frac{i}{2\pi}\Lambda_2\tau\right) \\ &= 1 + \frac{i}{2\pi}(\Lambda_1 + \Lambda_2)\tau - \frac{1}{4\pi}\Lambda_1 \wedge \Lambda_2\tau^2. \end{aligned}$$

And simply matching coefficients gives

$$c_1(E) = \frac{i}{2\pi}(\Lambda_1 + \Lambda_2) = \frac{i}{2\pi}\mathrm{tr}F,$$

and

$$c_2(E) = -\frac{1}{4\pi}\Lambda_1 \wedge \Lambda_2.$$

Jost gives the general formula for the second Chern class of tangent bundle E with rank m as[6]

$$c_2(E) - \frac{m-1}{2m} c_1(E) \wedge c_1(E) = \frac{1}{8\pi^2} \mathrm{tr} F_0 \wedge F_0, \tag{99}$$

for the "trace free part" of F,

$$F_0 := F - \frac{1}{m} \mathrm{tr} F \cdot \mathrm{Id}_E.$$

But for our discussion focus on two examples, the cases when D is a $\mathfrak{u}(1)$-covariant derivative for illustration and, more importantly, when D is a $\mathfrak{su}(2)$-covariant derivative.

[6] See page 127.

Examples and the Topological Charge

Firstly, in the case when E is a complex line bundle with structure group $\mathrm{U}(1)$ the curvature is just a two-form,

$$F_{d+A} = dA =: f,$$

for an arbitrary $\mathfrak{u}(1)$-connection A. This is analogous with the two-dimentional real case discussed already. Thus the trace-eigenvalue approach is here very simple, or alternatively, we get the first Chern class trivially from the determinant definition,

$$c_0(E) + c_1(E)\tau = \det\left(\frac{i}{2\pi} f\tau + \mathrm{Id}\right) = 1 + \frac{i}{2\pi} f\tau$$

and thus

$$c_1(E) = \frac{i}{2\pi} f. \tag{100}$$

There are only two Chern classes here (the above and $c_0(E) = 1$) so this wholly defines the topological structure of the bundle E.

Moreover, secondly let E have structure group $\mathrm{U}(2)$ so that the curvature,

$$F_D : \Omega^0(E; \mathfrak{su}(2)) \to \Omega^2(E; \mathfrak{su}(2)),$$

is represented by a matrix of two-forms

$$F_D = \begin{pmatrix} f_1^1 & f_2^1 \\ f_1^2 & f_2^2 \end{pmatrix} \in \mathfrak{su}(2) \tag{101}$$

where, for $j,\ k = 1,\ 2$,

$$f_k^j \in \Omega^2(\mathcal{M}),$$

and

$$\mathrm{tr} F_D = f_1^1 + f_2^2 = 0$$

as required in the definition of SU. Thus right away the first Chern class vanishes,

$$c_1(E) = \mathrm{tr} F = 0.$$

And now in this formulation we can compute the second Chern class from the determinant,

$$\begin{aligned}
c_0(E) + c_1(E)\tau + c_2(E)\tau^2 &= \det\left(\mathrm{Id} + \frac{i}{2\pi}F_D\tau\right) \\
&= \det \begin{pmatrix} \frac{i}{2\pi} f_1^1 \tau + 1 & \frac{i}{2\pi} f_2^1 \tau \\ \frac{i}{2\pi} f_1^2 \tau & \frac{i}{2\pi} f_2^2 \tau + 1 \end{pmatrix} \\
&= \left(\frac{i}{2\pi} f_1^1 \tau + 1\right) \wedge \left(\frac{i}{2\pi} f_2^2 \tau + 1\right) - \frac{i}{2\pi} f_2^1 \tau \wedge \frac{i}{2\pi} f_1^2 \tau \\
&= -\frac{1}{4\pi^2}\left(f_1^1 \wedge f_2^2 - f_2^1 \wedge f_1^2\right)\tau^2 + \frac{i}{2\pi}\left(f_1^1 + f_2^2\right)\tau + 1 \\
&= -\frac{1}{4\pi^2}\left(f_1^1 \wedge f_2^2 - f_2^1 \wedge f_1^2\right)\tau^2 + 1,
\end{aligned}$$

since $\mathrm{tr}F_D = 0$. Clearly now we can write the second Chern class,

$$\begin{aligned} c_2(E) &= -\frac{1}{4\pi^2}\left(f_1^1 \wedge f_2^2 - f_2^1 \wedge f_1^2\right) \\ &= \frac{1}{8\pi^2}\mathrm{tr}(F_D \wedge F_D). \end{aligned} \quad (102)$$

And we have defined the topological structure of E[7].

Integrate over the second Chern class over manifold to obtain the second Chern number known as the *topological charge* and written,

$$-k = -k(\mathcal{M}, \mathfrak{su}(2)) := -c_2(E)[\mathcal{M}] = -\frac{1}{8\pi}\int_{\mathcal{M}} \mathrm{tr}(F \wedge F) * 1 \quad (103)$$

(In fact this is a constant over the fundamental class $[\mathcal{M}]$ of oriented, four-dimensional, compact manifolds.)

A First Lower Bound for the Yang-Mills Functional

Now recall that on $\mathfrak{su}$ the ad-invariant innerproduct is given by minus the trace, and the specified innerproduct on $\Omega^2(E; \mathfrak{su})$ is, for $X \otimes \omega,\ Y \otimes \nu \in \Omega^2(E; \mathfrak{su})$, with $X,\ Y \in \Gamma(E)$ and $\omega,\ \nu \in \Omega^2(\mathcal{M})$,

$$\langle X \otimes \omega, Y \otimes \nu \rangle_{\Omega^2(E;\mathfrak{su})} = -\mathrm{tr}(XY)\ \omega \wedge *\nu,$$

[7] Freed and Uhlenbeck say, "The characteristic class $[c_2(E)]$ classifies SU(2) bundles over compact 4-manifolds, but this classification fails in higher dimensions." See page 33 for references.

so that

$$\int_{\mathcal{M}} \langle F, *F\rangle_{\Omega^2(E;\mathfrak{su})} = -\int_{\mathcal{M}} \mathrm{tr}(F \wedge F) * 1 = -8\pi^2 k. \tag{104}$$

Recall that the Hodge operator is an L^2-isometry, so

$$\langle F, F\rangle = \langle *F, *F\rangle ,$$

and (dropping the implied $\mathfrak{su}$-subscript) employ this to reformulate the Yang-Mills functional via

$$\begin{aligned}
\mathcal{YM}(D) &= \int_{\mathcal{M}} \langle F_D, F_D\rangle * 1 \\
&= \frac{1}{2}\int_{\mathcal{M}} \Big(\langle F_D, F_D\rangle + \langle *F_D, *F_D\rangle\Big) * 1 \\
&= \frac{1}{2}\int_{\mathcal{M}} \Big(\langle F_D - *F_D, F_D - *F_D\rangle\Big) * 1 + \int_{\mathcal{M}} \langle F_D, *F_D\rangle * 1 \\
&= \frac{1}{2}\int_{\mathcal{M}} |F_D - *F_D|^2 * 1 - 8\pi^2 k.
\end{aligned}$$

Since the square is positive we can bound the functional from below by the topological charge,

$$\mathcal{YM}(D) \geq -8\pi^2 k$$

for every $\mathfrak{su}(2)$-covariant derivative D. Remember we have said nothing about the sign of $-k$ and since the functional is a norm write

$$\mathcal{YM} \geq \max\{-8\pi^2 k, 0\}. \tag{105}$$

Instantons as Absolute Minima

No matter the covariant derivative write the curvature into its selfdual and anti-selfdual components as

$$F_{d+A} = F = F^+ + F^-$$

for

$$F^+ \in \Lambda^+ \text{ and } F^- \in \Lambda^-.$$

From here express the Yang-Mills functional as

$$\begin{aligned}
\mathcal{YM}(A) = \int_{\mathcal{M}} |F|^2 * 1 &= \int_{\mathcal{M}} \langle F, F \rangle * 1 \\
&= \int_{\mathcal{M}} \left\langle F^+ + F^-, F^+ + F^- \right\rangle * 1 \\
&= \int_{\mathcal{M}} \Big(\left\langle F^+, F^+ \right\rangle + 2 \left\langle F^+, F^- \right\rangle + \left\langle F^-, F^- \right\rangle \Big) * 1 \\
&= \int_{\mathcal{M}} \Big(\left\langle F^+, F^+ \right\rangle + \left\langle F^-, F^- \right\rangle \Big) * 1 = \int_{\mathcal{M}} \left(|F^+|^2 + |F^-|^2 \right) * 1,
\end{aligned}$$

since $\langle F^+, F^- \rangle = 0$ as a result of the orthogonality of Λ^+ and Λ^-. Mind that, by assumption,

$$F^+ = *F^+ \text{ and } F^- = - * F^-,$$

and break down the second Chern class:

$$\begin{aligned}
8\pi^2 c_2(E) = \mathrm{tr}(F \wedge F) &= \mathrm{tr}\big((F^+ + F^-) \wedge (F^+ + F^-)\big) \\
&= \mathrm{tr}\big(F^+ \wedge F^+ + F^- \wedge F^-\big) \\
&= \mathrm{tr}(F^+ \wedge F^+) + \mathrm{tr}(F^- \wedge F^-) \\
&= \mathrm{tr}(F^+ \wedge *F^+) - \mathrm{tr}(F^- \wedge *F^-) \\
&= -|F^+|^2 + |F^-|^2,
\end{aligned}$$

since the cross terms $F^+ \wedge F^-$ and $F^- \wedge F^+$ cancel each other out and (of course) the trace is linear. Integrating returns us to the topological charge and looks like

$$8\pi^2 k = \int_{\mathcal{M}} \big(|F^+|^2 - |F^-|^2\big) * 1, \tag{106}$$

which looks remarkably like our current expression for the Yang-Mills functional—it differs only in the sign of one term. Comparing the two we see

$$\int_{\mathcal{M}} \big(|F^+|^2 + |F^-|^2\big) * 1 \geq \left| \int_{\mathcal{M}} \big(|F^+|^2 - |F^-|^2\big) * 1 \right|$$

which implies

$$\mathcal{YM} \geq 8\pi^2 |k|, \tag{107}$$

and strengthens our bound form the previous section. We see that equality is attained,

$$\mathcal{YM}(D) = 8\pi^2 |k|,$$

i.e. the Yang-Mills functional has an absolute minimum, exactly when either the selfdual or the anti-selfdual part of the curvature vanishes,

$$F^+ = 0 \text{ or } F^- = 0,$$

which is to say that the covariant derivative D is an instanton or anti-instanton. And which one depends on the sign of (minus) the Chern number k. Minimize the difference $\mathcal{YM} - 8\pi^2 k$, firstly for $k \geq 0$:

$$\begin{aligned}
\mathcal{YM} - \left| \int_{\mathcal{M}} \left(|F^+|^2 - |F^-|^2 \right) * 1 \right| &= \int_{\mathcal{M}} \left(|F^+|^2 + |F^-|^2 \right) * 1 - \int_{\mathcal{M}} \left(|F^+|^2 - |F^-|^2 \right) * 1 \\
&= 2 \int_{\mathcal{M}} |F^-|^2 * 1,
\end{aligned}$$

which is minimized at zero when $F_D = F^+ \in \Lambda+$ is selfdual, D is an instanton, by definition. Likewise, if $k < 0$ corresponds to anti-selfdual curvature,

$$\begin{aligned}
\mathcal{YM} - \left| \int_{\mathcal{M}} \left(|F^+|^2 - |F^-|^2 \right) * 1 \right| &= \int_{\mathcal{M}} \left(|F^+|^2 + |F^-|^2 \right) * 1 + \int_{\mathcal{M}} \left(|F^+|^2 - |F^-|^2 \right) * 1 \\
&= 2 \int_{\mathcal{M}} |F^+|^2 * 1,
\end{aligned}$$

which again is minimized at zero, this time when $F_D = F^- \in \Lambda^-$ and D is an anti-instanton.

We have not, remember, discussed the existence of such solutions to the self-dual equations. While Cliff Taubes has constructed examples of solutions, there is no theoretical guarantee that the topological minimum will be attained for any covariant derivative. On the other hand, in some settings, for example in the case of line bundles over the two-spheres $\mathbb{S}^2 \times \mathbb{S}^2$ [8] there exist covariant derivatives that are neither instantons or anti-instantons but minimize the Yang-Mills functional by assuming the bound.

[8] According to Freed and Uhlenbeck, page 37.

Chapter 4: The Convex Selfdual Framework

I hope that the resurgence of convex methods recently incited by Ghoussoub with lead to facile investigation of the absolute minima of the Yang-Mills functional in four dimensions because (in some contexts) basic analysis will displace geometric, topological, and algebraic methods and alleviate the challenge of working at the intersection of these fields. The selfdaul Yang-Mills equations are poised to become among the canonical of examples of classical equations reinterpreted as representatives from more general classes that will provide diverse extensions of and offshoots from known work.

The Convex Setup:
The Legendre Transform, Subdifferential, and Fenchel Inequality

Firstly define the Legendre transform $\varphi^* : X^* \to \mathbb{R} \cup \{+\infty\}$ of convex, lower semicontinuous functional $\varphi : X \to \mathbb{R}\cup\{+\infty\}$ defined on Banach space X,

$$\varphi^*(y) := \sup\{\langle x, y\rangle - \varphi(x) \ ; \ x \in X\}. \tag{108}$$

Directly from the definition of the supremum, write simply

$$\varphi(x) + \varphi^*(y) \geq \langle x, y\rangle \tag{109}$$

for every $x \in X$ and $y \in X^*$, called the Fenchel Inequality. Note here that φ^* will itself always be convex, which direct manipulation of convex combinations reveals.

Now define independently the subdifferential of φ, $\partial\varphi : X \to 2^{X^*}$, as the set-valued map

$$\partial\varphi : x \mapsto \{y \in X^* \ ; \ \varphi(z) \geq \varphi(x) + \langle z - x, y\rangle \ \text{ for all } z \in X\}. \tag{110}$$

Here realize that if $\varphi = f$ is a smooth function on $X = X^* = \mathbb{R}^n$ then the subdifferential reduces to the singleton of the gradient, $\partial f = \{\nabla f\}$, by supposing

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0) \bullet \mathbf{y}$$

for every $\mathbf{x} \in \mathbb{R}^n$, and in particular $\mathbf{x} = \mathbf{x}_0 + h\mathbf{e}_i$ for $h \in \mathbb{R}$ and standard orthonormal basis vectors $\mathbf{e}_1, ..., \mathbf{e}_n$, so

$$f(\mathbf{x}_0 + h\mathbf{e}_i) \geq f(\mathbf{x}_0) + h\mathbf{e}_i \bullet \mathbf{y}$$

and thus

$$\lim_{h\to 0^-} \frac{f(\mathbf{x}_0 + h\mathbf{e}_i) - f(\mathbf{x}_0)}{h} \leq \mathbf{e}_i \bullet \mathbf{y} \leq \lim_{h\to 0+} \frac{f(\mathbf{x}_0 + h\mathbf{e}_i) - f(\mathbf{x}_0)}{h},$$

or, component-wise,

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = y_i.$$

So the subdifferential contains at most one element,

$$\mathbf{y} = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), ..., \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right) = \nabla f(\mathbf{x}_0),$$

which is of course in fact in the set as the tangent plane at any point $\mathbf{x}_0$ lies below the a convex function:

$$f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \bullet (\mathbf{x} - \mathbf{x}_0) \leq f(\mathbf{x}).$$

This intuition from finite-dimensional calculus contextualizes the subdifferential as a derivative that does not depend explicitly on a limit and as such freely applies to functionals at their points of nonsmoothness, taking on multiple values in this situation. The connection between the derivative with the Legendre Transform is the crux of the elementary selfdual theory, and comes when the Fenchel Inequality is attained:

$$\varphi(x) + \varphi^*(y) = \langle x, y \rangle \quad \text{if and only if} \quad y \in \partial\varphi(x). \tag{111}$$

Trivial manipulation of the supremum proves the implication,

$$\langle x, y \rangle - \varphi(x) = \sup_{x \in X} \{ \langle x, y \rangle - \varphi(x) \} \geq \langle z, y \rangle - \varphi(z)$$

for every $z \in X$, which reads $y \in \partial\varphi(x)$, and the converse is only one line longer: If for every $z \in X$

$$\varphi(z) \geq \varphi(x) + \langle z - x, y \rangle$$

then

$$\langle x, y \rangle - \varphi(x) \geq \langle z, y \rangle - \varphi(z).$$

Since the inequality holds for all z, take the supremum of the right-hand side,

$$\langle x, y \rangle - \varphi(x) = \sup\{\langle z, y \rangle - \varphi(z) \ ; \ z \in X\} = \varphi^*(y).$$

Now suppose further that the space is reflexive, $X^{**} = X$, so that $\varphi^{**} : X \to \mathbb{R}$, and

$$\begin{aligned} \varphi^{**}(x) &= \sup_{y \in X^*} \{\langle x, y \rangle - \varphi^*(y)\} \\ &= \sup_{y \in X^*} \Big\{ \langle x, y \rangle - \sup_{z \in X}\{\langle z, y \rangle - \varphi(z)\}\Big\} \\ &\leq \varphi(x), \end{aligned}$$

since

$$\langle x, y \rangle - \sup\{\langle z, y \rangle - \varphi(z) \ ; \ z \in X\} \leq \varphi(x).$$

That is,

$$\varphi^{**} \leq \varphi.$$

When the functional is convex in fact the reverse inequality holds, $\varphi \geq \varphi^{**}$, so we have equality:

$$\varphi^{**} = \varphi. \tag{112}$$

Apply this identity in the Fenchel inequality to recover its dual equivalent in terms of the subdifferential:

$$\varphi^{**}(x) + \varphi^{*}(y) = \varphi(x) + \varphi^{*}(y) = \langle x, y \rangle ,$$

automatically gives the third equivalent proposition, thus

$$y \in \partial\varphi(x) \quad \text{if and only if} \quad x \in \partial\varphi^{*}(y). \tag{113}$$

remembering of course that this is also equivalent to attainment in the Fenchel inequality.

The Basic Example for our Application

Now, poised to address the theory of Langragians on $X \times X^*$ that constitutes the essence of the chapter, turn firstly to a concrete example that is illustrative here and to be fruitful in the sequel. If $\varphi(x) = \frac{1}{p}|x|^p$ then

$$\varphi^*(y) = \sup\left\{ \langle x, y\rangle - \frac{1}{p}|x|^p \ ; \ x \in X \right\}$$

is calculable variationally. At maximizing x,

$$\langle x, y\rangle - \frac{1}{p}|x|^p \leq \langle x + tw, y\rangle - \frac{1}{p}|x + tw|^p$$

for any $w \in X$ and every $t \in \mathbb{R}$. Thus one-dimensional calculus gives

$$\frac{d}{dt}\Big|_{t=0} \left(\langle x + tw, y\rangle - \frac{1}{p}|x + tw|^p \right) = 0,$$

which implies

$$\langle w, y\rangle - |x + tw|^{p-2} \langle x + tw, w\rangle \Big|_{t=0} = 0$$

or

$$\langle w, y\rangle - |x|^{p-2} \langle x, w\rangle = 0.$$

Thus

$$|x|^{p-2} x = y$$

and

$$|x| = |y|^{\frac{1}{p-2}},$$

provided the found functional does belong to the dual space. Perhaps note quickly that for a function $f \in X = L^p(\mathbb{R}^n)$, the maximizer in fact is a member of the dual, $|f|^{p-2} f \in X^* = L^q(\mathbb{R}^n)$, where q is the exponent dual to p, i.e. $\frac{1}{p} + \frac{1}{q} = 1$, as

$$\int_{\mathbb{R}^n} \left| |f|^{p-2} f \right|^q = \int_{\mathbb{R}^n} |f|^p < \infty.$$

Returning to the general case, plugging in for x yields the Legendre transform φ^* to be

$$\left(\frac{1}{p}|\cdot|^p\right)^* (y) = |y|^{\frac{1}{p-1}+1} - \frac{1}{p}|y|^{\frac{p}{p-1}} = \left(1 - \frac{1}{p}\right)|y|^{\frac{p}{p-1}} = \frac{1}{q}|y|^q,$$

having noted that here the inner product does not complicate computations and reduces simply by first inserting x for y,

$$\langle x, y \rangle = \left\langle x, |x|^{p-2} x \right\rangle = |x|^p = |y|^{\frac{p}{p-2}}.$$

In this context the Fenchel Inequality reads

$$\frac{1}{p}|x|^p + \frac{1}{q}|y|^q \geq \langle x, y \rangle,$$

whenever $x \in X,\ y \in X^*$.

Lagrangians

Call a functional $L : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ a Lagrangian if it is lower semicontinuous and convex in both variables and not identically $+\infty$. Star denotes the Legendre transform in both variables,

$$L^*(q, y) = \sup\{\langle q, x\rangle + \langle y, p\rangle - L(x, p)\ ;\ x \in X, p \in X^*\}. \tag{114}$$

If

$$L^*(p, x) = L(x, p) \text{ for all } (p, x) \in X^* \times X \tag{115}$$

then the Lagragian L is called anti-selfdual. The first (general) example of such a functional is the sum of a convex functional on X and its Legendre transform on X^*:

$$L(x, p) = \varphi(x) + \varphi^*(p). \tag{116}$$

Convexity and lower-semicontinuity follow immediately from the same assumptions on φ and consequently φ^*, namely for every fixed p, $\varphi(x) \geq \langle x, p\rangle$ and that the sum of lower-semicontinuous functions is lower-semicontinuous. Computation of the Legendre transform is simple since the variables separate:

$$\begin{aligned}
L^*(q, y) &= \sup\{\langle q, x\rangle + \langle y, p\rangle - \varphi(x) - \varphi^*(p)\ ;\ x \in X, p \in X^*\} \\
&= \sup\{\langle q, x\rangle - \varphi(x)\ ;\ x \in X\} + \sup\{\langle y, p\rangle - \varphi^*(p)\ ;\ p \in X^*\} \\
&= \varphi^*(q) + \varphi^{**}(y) \\
&= \varphi(y) + \varphi^*(q) = L(y, q),
\end{aligned}$$

where the linearity of X, X^*, and $\langle \cdot, \cdot \rangle$ allowed us to reposition the minus signs in the arguments of the supremum operators. Now, remarkably, the functional $I : X \to \mathbb{R}$, defined by

$$I(x) := L(x, \Lambda x) - \langle x, \Lambda x \rangle = \varphi(x) + \varphi^*(\Lambda x) - \langle x, \Lambda x \rangle , \tag{117}$$

for any linear $\Lambda : X \to X^*$ (which cleary preserves the necessary convexity), satisfies

$$\inf\{I(x) \; ; x \in X\} = 0 = I(\bar{x}), \tag{118}$$

again supposing that φ is coercive to ensure the existence of such a $\bar{x} \in X$. Recognize this minimum from the main inequality—$I(x) = 0$ when

$$\varphi(x) + \varphi^*(\Lambda x) = \langle x, \Lambda x \rangle \tag{119}$$

or, moreover,

$$\Lambda x \in \partial\varphi(x). \tag{120}$$

Thus, replacing the second-order Euler-Lagrange equations for critical points, self-duality gives a *first-order* (in a convex sense) equation to find the minimizer.

An Example from Partial Differential Equations

To reify the overarching setting and demonstrate the power of this case in particular, consider one of the simplest examples from Ghoussoub's series of papers on the subject, the non-symmetric Dirichlet problem: Given functions $f : \Omega \to \mathbb{R}$ and smooth $\mathbf{a} : \Omega \to \mathbb{R}^n$, for $\Omega \subset \mathbb{R}^n$ bounded,

$$\begin{aligned} \sum_{i=1}^{n} a_i \frac{\partial u}{\partial x_i} &= -\Delta u + |u|^{p-2}u + f \text{ on } \Omega, \\ u &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{121}$$

for $u \in H_0^1(\Omega)$. To reformulate this equation as a selfdual, convex minimization problem construct the functional,

$$\Psi(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{p}\int_\Omega |u|^p + \int_\Omega fu, \tag{122}$$

whose form is motivated by the fact that its Fréchet derivative $D\Psi$ is equal to the right hand side of the above sample application integrated. Where for $u \in X$, the Fréchet derivative $D\Psi$ at u is the solution of

$$\lim_{t\to 0} \frac{\left|\Psi(u+tv) - \Psi(u) + D\Psi(u)\right|}{\left|\left|tv\right|\right|_X} = 0 \tag{123}$$

if the limit exists and agrees for every $v \in X$. Since the norm in the numerator is just absolute value, we can rewrite this explicitly as,

$$D\Psi(u) = \lim_{t\to 0} \frac{\Psi(u+t\hat{v}) - \Psi(u) + D\Psi(u)}{t} \tag{124}$$

where $\hat{v} \in X$ is any vector with norm one. For completeness write explicitly the Lagrangian $L : H_0^1(\Omega) \times H^{-1}(\Omega) \to \mathbb{R}$, as $L(u, v) = \Psi(u) + \Psi^*(v)$, and realize the interesting and essential quality that throughout the following the Legendre conjugate Ψ^* remains uncalculated.

Build further the operator $A : H_0^1(\Omega) \to L_0^2(\Omega)$,

$$A(u) = \mathbf{a} \bullet \nabla u = \sum_{i=1}^{n} a_i \frac{\partial u}{\partial x_i},$$

and now finally the functional $I : H_0^1(\Omega) \to \mathbb{R}$,

$$I(u) = L(u, Au) - (u, Au),$$

or, replacing A and L by their definitions and integrating the product by parts,

$$I(u) := \Psi(u) + \Psi^*(\mathbf{a} \bullet \nabla u) + \int_\Omega \operatorname{div}(\mathbf{a})|u|^2.$$

Convexity and lower-semicontinuity follow immediately from their definitions and restrict the exponent so that $p > 1$ to ensure a growth condition sufficient for the attainment of the infimum. Hence look for $\bar{u}$ such that

$$I(\bar{u}) = \inf\{I(u)\ ; u \in H_0^1(\Omega)\} = 0, \tag{125}$$

or, moreover,

$$\mathbf{a} \bullet \nabla \bar{u} \in \partial \Psi(\bar{u}).$$

Relying on the established proof of the analogous fact in finite dimensions, realize that if a functional I has Fréchet derivative DI at $\bar{u} \in X$, then $\partial I(\bar{u}) = \{DI(\bar{u})\}$. Thus the calculation of the one-dimensional limit, bearing in mind the zero boundary condition imposed by $H_0^1(\Omega)$ when integrating by parts, yields the desired equation:

$$\mathbf{a} \bullet \nabla \bar{u} = \Delta \bar{u} + |\bar{u}|^{p-2}\bar{u} + f.$$

The Link with the Instantons

This final section, which indicates how the first and second chapters should be unified, will unfortunately be both cursory and formal. Contrariwise it be viewed positively how easy and powerful the newer, analytic formalism is and also how, once resolved in the present context it should extend to other still richer ones. Namely the generalization of Yang-Mills theory to address p-energies.

Ghoussoub has extended the theory as outlined to minimize functionals

$$I : X \to \mathbb{R}$$

of the form

$$I(x) := L(\Lambda x, \Gamma x) - \langle \Lambda x, \Gamma x \rangle \tag{126}$$

for selfdual Lagrangian $L : X \times X^* \to \mathbb{R}$ where either one of the two operators

$$\Lambda : X \to X$$

or

$$\Gamma : X \to X^*,$$

is nonlinear. Here,

$$I \geq 0$$

and the functional attains its minimum. We care about the basic case when, for a convex functional $\Psi : X \to \mathbb{R}$,

$$L(x,p) = \Psi(x) + \Psi^*(p).$$

In this case we obtain

$$I(x) = \Psi(\Lambda x) + \Psi^*(\Gamma x) - \langle \Lambda x, \Gamma x\rangle$$

and the minimizer $\bar{x}$ such that

$$I(\bar{x}) = 0$$

satisfies

$$\Gamma\bar{x} \in \partial\Psi(\Lambda\bar{x}) \tag{127}$$

as a result of the key relationship between the subdifferential and the Fenchel inequality.

Now reassign notation to match the geometric setting and (formally) apply this result. On the second special unitary group, the Yang-Mills functional aligns with this setting from completing the square as explained at the end of the second chapter.

$$\begin{aligned} \mathcal{YM}(A) &= \int_{\mathcal{M}} \left|F_{d+A}\right|^2 * 1 & (128)\\ &= \frac{1}{2}\int_{\mathcal{M}} \left|F_{d+A}\right|^2 * 1 + \frac{1}{2}\int_{\mathcal{M}} \left| * F_{d+A}\right|^2 * 1 & (129)\end{aligned}$$

Now define replace the general Banach space with the two-forms which constitute a Hilbert space and thus their own dual space by the Riesz theorem. Label the elements of the convex theory explicitly for a vector bundle E over the manifold $\mathcal{M}$: Begin with the convex functional, which here takes the form of a second power,

$$\Psi : \Omega^2\big(\mathrm{Ad}E; \mathfrak{su}(2)\big) \longrightarrow \mathbb{R},$$

$$\Psi : \omega \mapsto \frac{1}{2}\big|\omega\big|^2$$

and we already know it is its own Legendre transform,

$$\Psi^* = \Psi = \frac{1}{2}\big|\ \cdot\ \big|^2 : \Omega^2\big(\mathrm{Ad}E; \mathfrak{su}(2)\big) \longrightarrow \mathbb{R}.$$

Now, on the underlying algebra, define the operators

$$\Lambda\ , \Gamma : \mathfrak{su}(2) \longrightarrow \Omega^2\big(\mathrm{Ad}E; \mathfrak{su}(2)\big)$$

whose actions are defined by generating the curvature and its Hodge star,

$$\Lambda : A \mapsto F_{d+A} = dA + A \wedge A$$

and

$$\Gamma : A \mapsto *F_{d+A}.$$

Thus

$$\mathcal{YM}(A) = \int_{\mathcal{M}} \Psi(\Lambda A) * 1 + \int_{\mathcal{M}} \Psi^*(\Gamma A) * 1. \tag{130}$$

But since the inner product of the curvature with its Hodge star is constant,

$$\langle F_D, *F_D \rangle = -8\pi^2 k$$

so the Yang-Mills functional has the same minimizers as the functional

$$\begin{aligned} I_{\mathcal{YM}}(A) &:= \mathcal{YM}(A) + 8\pi^2 k \\ &= \int_{\mathcal{M}} \Psi(\Lambda A) * 1 + \int_{\mathcal{M}} \Psi^*(\Gamma A) * 1 - \int_{\mathcal{M}} \langle \Lambda A, \Gamma A \rangle * 1, \end{aligned} \tag{131}$$

which fits in with the theory of selfdual Lagrangians. Thus we have transported the Yang-Mills functional into a partial differential equations context where hopefully if can be addressed with simple convex techniques and the more direct language of real analysis. Since

$$\partial \Psi(x) = \partial \left(\frac{1}{2} |x|^2 \right) = \{x\},$$

the equation

$$\Gamma \bar{A} \in \partial \Psi(\Lambda \bar{A})$$

for the minimizer $\bar{A}$ such that

$$I_{\mathcal{YM}}(\bar{A}) = 0,$$

becomes

$$\Gamma\bar{A} \in \{\Lambda\bar{A}\},$$

or, rather, replacing the operators by their values,

$$*F_{d+\bar{A}} \in \{F_{d+\bar{A}}\}.$$

This is trivially equivalent to the pair of equation,

$$F_{d+\bar{A}} = *F_{d+\bar{A}},$$

which, of course, is the selfdual Yang-Mills equation requiring that the curvature $F_{d+\bar{A}}$ be selfdual thus that the covariant derivative $d+\bar{A}$ be an instanton.

To recover the anti-selfdual equation, simply exploit that $|*F_{d+A}|^2 = |-*F_{d+A}|^2$ and write the Yang-Mills functional as

$$\mathcal{YM}(A) = \frac{1}{2}\int_{\mathcal{M}} |F_{d+A}|^2 * 1 + \frac{1}{2}\int_{\mathcal{M}} |-*F_{d+A}|^2 * 1. \tag{132}$$

Thus the above programme applies identically with the operator Γ replaced by

$$\Gamma' : A \mapsto -*F_{d+A}$$

and all else identical. This gives way to

$$\Gamma'\bar{A} \in \{\Lambda\bar{A}\}$$

or

$$- * F_{d+\bar{A}} \in \{F_{d+\bar{A}}\},$$

which, exactly in parallel with the selfdual case above, is exactly the equation

$$F_{d+\bar{A}} = - * F_{d+\bar{A}},$$

the anit-self dual equation for an anti-instanton $d + \bar{A}$.

The Future

There is a dense literature surrounding instantons and their role in gauge theory—we did not talk about gauge groups, but know that all Yang-Mills connections have gauge equivalent curvatures—and the insights they provide into topology. Alas, we have hardly touched on the real geometry of four-dimensional manifolds.

We have seen, though, the facile application of a new and surprisingly easy theory in our complex setting. As a suggestion for further investigation then, focus on exploiting this framework and realize that here we can naturally extend the theory to p-energies:

$$\mathcal{YM}^p(A) := \frac{1}{p}\int_{\mathcal{M}} \left|F_{d+A}\right|^p * 1 + \frac{1}{q}\int_{\mathcal{M}} \left| * F_{d+A}\right|^q * 1 - \int_{\mathcal{M}} \langle F_{d+A}, *F_{d+A}\rangle * 1, \tag{133}$$

where $1/p + 1/q = 1$. By the same argument as in the $p = 2$ case ($\mathcal{YM}^2 =$

$\mathcal{YM}$), but with the functionals defined as

$$\Psi = \frac{1}{p}|\cdot|^p, \text{ and } \Psi^* = \frac{1}{q}|\cdot|^q,$$

we get the subdifferential,

$$\partial\Psi(x) = \partial\left(\frac{1}{p}|x|^p\right) = \left\{x|x|^{p-2}\right\}.$$

Thus the minimum of $\mathcal{YM}^p$ is attained at $\bar{A}$ when

$$*F_{d+\bar{A}} = F_{d+\bar{A}}\left|F_{d+\bar{A}}\right|^{p-2}. \tag{134}$$

References

Freed, D., and K. Uhlenbeck, 1991, *Instantons and Four-Manifolds, Second Ed.*, Springer-Verlag, New York.

N. Ghoussoub: "Anti-selfdual Lagrangians: Variational resolutions of non self-adjoint equations and dissipative evolutions", AIHP-Analyse non linaire (In press) 35 pp

N. Ghoussoub: 2005, "A class of self-dual differential equations and its variational principles".

Jaffe, A., and C. Taubes, 1980, *Vortices and Monopoles: Structure of Static Gauge Theories*, Birkhaüser, Boston.

Jost, J., 2005, *Riemannian Geometry and Geometric Maps, Fourth Ed.*, Springer, Heidelberg.

Urakawa, H., 1990, *The Calculus of Variations and Harmonic Maps*, AMS Translations of Mathematical Mongraphs, Providence.

K. Wehrheim: "Anti-self-dual instantons with Lagrangian boundary conditions I: Elliptic theory", Comm.Math.Phys. 254 (2005), no.1, 45-89.

Wehrheim, K., 2004, *Uhlenbeck Compactness*, EMS Series of Lectures in Mathematics.