CHAPTER 5: YANG-MILLS THEORY

5.1 Yang-Mills equations.

Let M be a Riemann manifold with Riemann metric g. In local coordinates the metric is represented as a symmetric nondegenerate tensor field $g_{ij}(x)$ with $i, j = 1, 2, \ldots, n$, where $n = \dim M$. Let $\pi : P \to M$ be a principal G bundle over M. Let $\rho : G \to Aut(V)$ be a unitary finite-dimensional representation of G in V. This defines an associated vector bundle $E = P \times_\rho V$ and the curvature tensor F of a connection in P is represented (locally) by matrix functions $F_{ij}(x) = \partial_i A_j - \partial_j A_i + [A_i, A_j]$ acting on vectors in V.

We shall define raising and lowering of space-time indices (i.e., coordinate indices in M) as usual, $A^i = g^{ij} A_j, B_i = g_{ij} B^j$, where the matrix (g^{ij}) is the inverse of (g_{ij}). Recall also that the metric g defines a volume form on M, $d(vol_M) = \sqrt{\det(g)} dx_1 \wedge dx_2 \cdots \wedge dx_n$. We define the *Yang-Mills functional*

$$Y(A) = \frac{1}{4} \int_M \operatorname{tr} F_{\mu\nu} F^{\mu\nu} d(vol_M).$$

The Yang-Mills action is invariant under gauge transformations $F' = g^{-1} F g$. There is an alternative way to write the YM action as

$$Y(A) = -\frac{1}{2} \int_M \operatorname{tr} F \wedge *F.$$

The action leads to field equations through Euler-Lagrange variational priciple. Let $A + tB$ be a 1-parameter family of vector potentials:

$$\frac{d}{dt} Y(A + tB)|_{t=0} = \frac{1}{2} \int_M \operatorname{tr} F^{\mu\nu} (\partial_\mu B_\nu - \partial_\nu B_\mu + [A_\mu, B_\nu] + [B_\mu, A_\nu]) d(vol_M).$$

When M is a manifold without boundary, we can integrate by parts and we get

$$\delta Y(A) = -\int_M \operatorname{tr} B_\nu (\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}]) d(vol_M).$$

If A is an extremal the YM action then we obtain *the Yang-Mills equations*

$$D_\mu F^{\mu\nu} = \partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}] = 0.$$

When G is abelian this gives the Maxwell's equations $\partial_\mu F^{\mu\nu} = 0$ in vacuum. In addition, we have the Bianchi identities

$$D_\mu F_{\nu\lambda} + D_\lambda F_{\mu\nu} + D_\nu F_{\lambda\mu} = 0$$

for all indices λ, μ, ν. If there are external sources we have instead

$$D_\mu F^{\mu\nu} = j^\nu$$

for some Lie algebra valued current j^ν.

The Yang-Mills equations is a complicated nonlinear system of second order partial differential equations. Not much is known about the general solutions. However, there is a class of solutions which is well understood. These so-called (anti) instantons are characterized by the (anti) self-duality property $F = *F$ ($F = -*F$) in the case of a Riemannian 4-manifold M. Recall that

$$* : \Omega^k(M) \to \Omega^{n-k}(M)$$

is a linear map and $** = \pm 1$. When $n = 4$ and $k = 2$ the sign is + (exeercise) For this reason the eigenvalues of $*$ are ± 1, when restricted to 2-forms on a 4-manifold. In the case of Lorentzian metric $** = -1$ on 2-forms and therefore in this case there are no real eigenvalues (and no real (anti) instantons).

In the case of an instanton we have

$$Y(A) = -\frac{1}{2} \int_M \mathrm{tr}\, F \wedge F$$

and so the value of the YM functional is given by the second Chern class. In particular, when $M = S^4$ we get

$$Y(A) \sim \int_{S^3} \mathrm{tr}\, (g^{-1} dg)^3,$$

where $g : S^3 \to G$ is the transition function on the equator. Thus for self-dual solutions the YM functional is quantized in units $(2\pi)^2 k$ with $k \in \mathbb{Z}$.

5.2 Dirac equation For each positive integer n we construct a set of complex $2^{[n/2]} \times 2^{[n/2]}$ matrices γ_i, $i = 1, 2, \dots n$, with the relations

$$\gamma \gamma_j + \gamma_j \gamma_i = 2\eta_{ij},$$

where η is either the Minkowski or the Euclidean metric. Here $[x]$ is the integer part of a real number x. The matrices are constructed by induction on n. For $n = 1$ there is only one 1×1 matrix $\gamma_1 = 1$. The induction from odd to even n is as follows. The dimension of the matrices is increased by a factor 2. In the case of Euclidean metric we set

$$\gamma_i \mapsto \begin{pmatrix} 0 & \gamma_i \\ \gamma_i & 0 \end{pmatrix}$$

for $i = 1, 2, \ldots, n$ and we add

$$\gamma_{n+1} = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix},$$

where the blocks are $2^{(n-1)/2} \times 2^{(n-1)/2}$ matrices. The induction from even n to $n + 1$ is defined by adding the matrix

$$\gamma_{n+1} = (-1)^{n/4} \gamma_1 \gamma_2 \ldots \gamma_n$$

of same dimension.

In the case of the Minkowski metric there will be some sign changes. In the induction from odd to even $n = 2k$ we have

$$\gamma_i \mapsto \begin{pmatrix} 0 & \gamma_i \\ -\gamma_i & 0 \end{pmatrix}$$

where γ_i for $i = 1, 2, \ldots, 2k-1$ are the Euclidean $\gamma-$ matrices in $2k-1$ dimensions, and the 'time like' matrix is defined as

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In the induction from even $n = 2k$ to $2k + 1$ one can take γ_j for $j = 1, 2, \ldots, 2k$ as $i\times$ the Euclidean $\gamma-$ matrices in $2k$ dimensions and then set $\gamma_0 = e^{i\alpha}\gamma_1 \cdots \gamma_{2k}$ for an appropriate phase factor $e^{i\alpha}$.

The Dirac equation in the flat space $\mathbb{R}^n$ is then

$$(i\gamma^\mu \partial_\mu + m)\psi = 0,$$

where $\psi : \mathbb{R}^n \to \mathbb{C}^N$, with $N = 2^{[n/2]}$ and m is a constant which is interpreted as a mass of the field ψ. Because of the anticommutation relations of the $\gamma-$ matrices any solution of the Dirac equation satisfies also the Klein-Gordon equation

$$(\partial^\mu \partial_\mu + m^2)\psi = 0.$$

Using Fourier transform we can look for solutions of the form $u \cdot e^{-ip_\mu x^\mu}$, for some constant vector $u \in \mathbb{C}^N$ and a momentum vector $p \in \mathbb{R}^n$. The Dirac equation becomes now

$$(\gamma^\mu p_\mu + m)u = 0.$$

The Klein-Gordon equation requires that $-p^2 + m^2 = 0$.

Exercise Show that for $n = 4$, Minkowski metric, and for each momentum vector $p^2 = m^2$ there are exactly two linearly independent solutions of the matrix equation $(\gamma^\mu p_\mu + m)u = 0$.

In a curved space the Dirac equation needs a modification. First, we have to assume that the space M has a *spin structure.* This is defined as follows. We shall discuss the case of a Riemann metric. To begin with, the rotation group $SO(n)$, for $n > 2$, has a simply connected double covering $Spin(n)$. That is, there is a onto 2-1 group homomorphism $\phi : Spin(n) \to SO(n)$. We are looking for a principal bundle $P \to M$ with structure group $Spin(n)$ such that there is a 2-1 onto map $\theta : P \to LM$, where LM is the bundle of orthonormal oriented frames in the tangent bundle TM. The structure group of the principal bundle $LM \to M$ is $SO(n)$. The map θ should take the fiber P_x to the fiber L_xM for each $x \in M$ and we require that

$$\theta(pg) = \theta(p)\phi(g)$$

for all $p \in P$ and $g \in Spin(n)$. Such a bundle P is called a spin structure on M. Not every manifold has a spin structure and if there is a spin structure it does not need to be unique.

We shall from now on assume that M is a Riemannian spin manifold with a fixed spin structure.

The group $Spin(n)$ has a (faithful) representation in $\mathbb{C}^N$ for $N = 2^{[n/2]}$. The Lie algebra of $Spin(n)$ is isomorphic with the Lie algebra of $SO(n)$ since the groups are locally isomorphic. The Lie algebra $Lie(SO(n))$ consists of real antisymmetric $n \times n$ matrices and it is spanned by matrices $L_{ij} = -L_{ji}$ with commutation relations

$$[L_{ij}, L_{kl}] = \delta_{jk}L_{il} + \delta_{il}L_{jk} - \delta_{ik}L_{jl} - \delta_{jl}L_{ik}.$$

The representation in $\mathbb{C}^N$ is obtained by the mapping $L_{ij} \mapsto M_{ij} = \frac{1}{4}[\gamma_i, \gamma_j]$. One can then check by direct computation, using the anticommutation relations of

γ-matrices, that the matrices M_{ij} satisfy the same commutation relations as the matrices L_{ij} and thus we indeed have a representation of the Lie algebra $Lie(Spin(n))$ in $\mathbb{C}^N$. The group $Spin(n)$ is simply connected and for this reason the representation of its Lie algebra can be exponentiated to give a representation of the group $Spin(n)$. Denote this representation by ρ.

We can now define the *Dirac spinor bundle* S as the associated bundle $S = Spin(M) \times_\rho \mathbb{C}^N$. Sections of this vector bundle are *Dirac spinor fields.* We define a covariant derivative ∇_μ acting on Dirac spinor field. First, let e_a with $a = 1, 2, \ldots, n$ be a local oriented orthonormal basis in the tangent bundle TM. Then

$$\nabla_\mu e_a = \Gamma^b_{\mu a} e_b$$

defines the Christoffel symbols in the basis e_a. The Christoffel symbols can be computed from the fact that $e_a = e^\mu_a \partial_\mu$, and so

$$\nabla_\mu e_a = (\partial_\mu e^\nu_a)\partial_\nu + e^\nu_a \nabla_\mu \partial_\nu = (\partial_\mu e^\nu_a) e^b_\nu e_b + e^\nu_a \Gamma^\lambda_{\mu\nu} e^b_\lambda e_b,$$

where (e^a_μ) is the inverse to the matrix (e^μ_a). Inserting

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\alpha}(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu})$$

we get an explicit expression for the symbols $\Gamma^b_{\mu a}$. We set

$$\omega_\mu = \frac{1}{2}\Gamma^b_{\mu a} M_{ab}.$$

The Dirac equation on a curved manifold is then

$$i\gamma^\mu(\partial_\mu + \omega_\mu)\psi + m\psi = 0.$$

Here we have defined $\gamma^\mu = e^\mu_a \gamma^a$. They satisfy the anticommutation relations

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu},$$

whereas

$$\gamma_a\gamma_b + \gamma_b\gamma_a = 2\delta_{ab}.$$

Let E be a complex vector bundle over M with a connection, described locally by a vector potential A_μ. Consider the tensor product bundle $S \otimes E$. The Dirac equation for sections of the extended bundle is

$$i\gamma^\mu(\partial_\mu + \omega_\mu + A_\mu)\psi + m\psi = 0.$$

Here ω_μ acts on the first factor and A_μ on the second factor in the tensor product. To be precise, we should write $A_\mu \mapsto 1 \otimes A_\mu$ and $\omega_\mu \to \omega_\mu \otimes 1$.

We also need to fix a hermitean inner product $< \cdot, \cdot >_x$ in the fibers $(S \otimes E)_x$. We can then define the Hilbert space $H = L_2(S \otimes E)$ of square integrable sections of the bundle $S \otimes E \to M$. The inner product for a pair of sections ψ, ϕ is

$$< \psi, \phi >= \int_M < \psi(x), \phi(x) >_x d(vol_M).$$

Assuming that the gauge group is unitarily represented (the Lie algebra elements $A_\mu(x)$ are antihermitean matrices), in the case of Riemann metric the Dirac operator is self-adjoint (in an appropriate dense domain) in the Hilbert space H.

5.3 The index of the Dirac operator

Let $T : H \to H$ be a linear operator in a Hilbert space H. We set $\ker T = \{x \in H | Tx = 0\}$ and $\operatorname{coker} T = (TH)^\perp$. When $\ker T$ and $\operatorname{coker} T$ are finite dimensional then T is a *Fredholm operator* and its Fredholm index is

$$\operatorname{ind} T = \dim \ker T - \dim \operatorname{coker} T.$$

Example Let $T : H \to H$ be defined as $Te_n = e_{n+1}$, where $\{e_n\}_{n=1,2,\dots}$ is an orthonormal basis. Now $\ker T = 0$ and $\operatorname{coker} T$ consists of $\mathbb{C} \cdot e_1$. Thus $\operatorname{ind} T = -1$.

If T is a Fredholm operator then $\operatorname{ind} T = -\operatorname{ind} T^*$. This follows from $< y, T^* x >=< Ty, x >$ and so $T^* x = 0$ if and only if $< Ty, x >= 0$ for all y, which is equivalent to $x \in (TH)^\perp$ and so $\ker T^* = \operatorname{coker} T$. In particular, when T is self-adjoint the Fredholm index is zero.

Let us study the case of the Dirac operator on an even dimensional manifold M. In this case we have an hermitean operator Γ with $\Gamma^2 = 1$ which anticommutes with D. The *chirality operator* Γ is defined as $\gamma_{n+1} = e^{i\alpha}\gamma_1 \dots \gamma_n$ for an appropriate phase factor $e^{i\alpha}$.

We observe that the nonzero eigenvalues of D come in pairs $\pm\lambda$ because of $D(\Gamma\psi) = -\Gamma D\psi = -\lambda(\Gamma\psi)$ if $D\psi = \lambda\psi$.

The case of the zero eigenvalue is different. We can split the kernel of D to a pair of subspaces $\ker D = V_- \oplus V_+$ where $V_\pm$ are defined by diagonalizing Γ; the eigenvalues of Γ are ± 1. In general, the dimensions of $V_\pm$ are different. However, in the case of a compact manifold M all eigenspaces of D are finite dimensional and we can define the index $\mathrm{dim} V_+ - \mathrm{dim} V_-$. This is in fact the Fredholm index of a certain operator. Diagonalizing Γ we can write

$$D = \begin{pmatrix} 0 & D_+ \\ D_- & 0 \end{pmatrix},$$

where D_+ maps to eigenspace of Γ corresponding to the eigenvalue -1 to the eigenspace $+1$ and vice-versa for D_-.

We have now

$$\mathrm{ind} D_- = \dim V_+ - \dim V_-.$$

We have $D_+ = D_-^*$ and so $\mathrm{ind}\, D_+ = -\mathrm{ind}\, D_-$.

In functional analysis one proves that for bounded operators the index of a Fredholm operator is a continuous function (in operator norm) of the operator. For this reason the index is a *topological invariant.* It remains constant under continuous deformations of the operator. Although a Dirac operator is always unbounded, one can still show that its index is a continuous function of the parameters: vector potentials, choice of Riemann metric etc. For this reason the index $\mathrm{ind}\, D_-$ depends only on the homotopy class of the vector bundles S, E and defines a topological invariant for the bundles under consideration. On the other hand, we have explicitly used a metric and a vector potential in the construction of D.

We already know that there are characteristic classes (Chern classes, Pontrjagin classes) which are topological invariants. For this reason it is not so big surprize that the index can be expressed in terms of these classes.

Theorem. *(Atiyah-Singer)*

$$ind\, D_+ = \int_M \hat{A}(TM) \wedge ch(E).$$

We cannot prove this theorem here but instead we illustrate the philosophy behind the index theorems by an explicit calculation.

Let $H = H_+ \oplus H_-$ be a sum of two infinite-dimensional complex Hilbert spaces and denote by P_+ the projection on H_+. Let $u : H \to H$ be an invertible operator such that $P_+ u P_+$ is a Fredholm operator. We also assume that the trace of $[P_+, u]$ is absolutely converging. We claim that

$$\mathrm{ind}(P_+ u P_+) = \mathrm{tr}\, u^{-1}[u, P_+].$$

To prove this index theorem we need to check that the formula holds for some selection of operators u_n with $\operatorname{ind} P_+ u_n P_+ = n$ for $n \in \mathbb{Z}$. This is sufficient by the continuity of the index! So we may choose $H = L_2(S^1)$ and H_+ is defined by nonnegative Fourier modes and H_- by negative Fourier modes. We select u_n as the multiplication operator by the Fourier mode e^{-inx}. Then it is easy to see that $\operatorname{ind} P_+ u_n P_+ = n$. On the other hand u_n^{-1} is the multiplication operator by the function e^{inx} and by a simple computation $\operatorname{tr} u^{-1}[u, P_+] = n$.