# Estudo Geométrico de Soluções da Teoria de Yang-Mills no Espaço de Minkowski

Rogerio de Queiroz Chaves

Orientador: Prof. Dr. Márcio Antônio de Faria Rosa

Campinas, 11 de junho de 1999

UNICAMP
BIBLIOTECA CENTRAL

# Estudo Geométrico de Soluções da Teoria de Yang-Mills no Espaço de Minkowski

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por **Rogerio de Queiroz Chaves** e aprovada pela comissão julgadora.

Campinas, 11 de junho de 1999.

Prof. Dr. Márcio Antônio de Faria Rosa
Orientador

Banca examinadora:

Prof. Dr. Márcio Antônio de Faria Rosa

Prof. Dr. Caio José Coletti Negreiros

Prof. Dr. Jaime Vaz Junior

Dissertação apresentada ao **Instituto de Matemática, Estatística e Computação Científica, UNICAMP**, como requisito parcial para obtenção do Título de MESTRE em Matemática.



99.15342

UNIDADE BC
N.º CHAMADA:
V. Ex.
TOMBO BC/ 38986
PROC. 229/99
C D X
PREÇO R$ 11,00
DATA 09/12/99
N.º CPD

CM-00126433-6

## FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Chaves, Rogerio de Queiroz

C398e Estudo geométrico de soluções da teoria de Yang-Mills no espaço de Minkowski / Rogerio de Queiroz Chaves -- Campinas, [S.P. :s.n.], 1999.

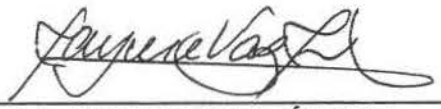
Orientador : Márcio Antônio de Faria Rosa

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

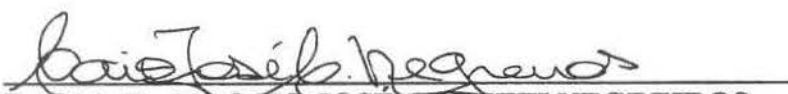
1. Geometria diferencial. 2. Espaços fibrados (Matemática). 3. Campos de calibre (Física). I. Rosa, Márcio Antônio de Faria. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

**Dissertação de Mestrado defendida em 29 de junho de 1999 e aprovada pela Banca Examinadora composta pelos Profs. Drs.**

**Prof (a). Dr (a). MÁRCIO ANTONIO DE FARIA ROSA**



**Prof (a). Dr (a). JAYME VAZ JÚNIOR**



**Prof (a). Dr (a). CAIO JOSÉ COLETTI NEGREIROS**

# Agradecimentos

A Deus, por permitir-me a imerecida honra de conhecê-lo, pouco a pouco, mesmo num meio tão árido e hostil a esse tipo de conhecimento.

Ao corpo docente da minha primeira e mais importante escola: meus pais Cirineu e Hilda Luiza, e as Roses.

À minha amada Michela, companheira na Escola do Amor Ágape.

Aos professores que aprendi a admirar e que deixaram um pouco deles mesmos em minha formação e, consequentemente, em tudo o que penso e fa,co. Da pré-escola até aqui: a Maria Anete, a Joana, a D. Iná, a Isabel, o Eurípedes, o Pedrão, o Édio, a Maria Ignez, o Evaldo, o Norton Moroczowiks, o Antônio Carlos (Macarrão), o Cristian, o Rafael, o Sávio, o Rigas, o Djairo, o Francesco Mercuri, e os outros que embora não tenham ficado em minha memória falha, podem alegrar-se vendo aqui o fruto de seu trabalho. Ao maior deles, o Márcio Rosa, que mais influenciou o presente trabalho, tanto pela via natural da orientação quanto pelo modelo de clareza de raciocínio que inspirou.

Pela hospitalidade, aos que permitiram que eu fizesse de seus lares meu lar, por algum ('as vezes longo) tempo: o professor Hans Kurt Edmund Liesenberg e sua família; o Marcos, a Marlene, o Adrian e a Paola; a família da Michela; o pessoal da Comunidade Cristã de Campinas; o pessoal da IMCG.

Pela amizade, idéias e comentários valiosos no nível acadêmico e para a vida, agradeo aos grandes amigos e irmãos da Moradia Estudantil: Edvaldo, Eduardo, Fernando, Albert, Robson, Anderson, Josias e Jean;
aos colegas do IMECC: Adilson, Alex, Mosna, Newton e por aíafora;
ao pessoal do Alfa e Ômega, Nilson e Angélica, Foster e Beth, Raul e Graça, Bill e Aida, e todos os outros;

Ao povo e ao CNPq pelo suporte financeiro;
aos poucos (dentre muitos) funcionários que fazem a UNICAMP funcionar a favor da produção de conhecimento.

# Resumo

Este trabalho está dividido em duas partes principais. Na primeira desenvolvemos a teoria básica das conexões em fibrados principais procurando fazer uma ponte entre a formulação local, típica dos físicos no estudo das teorias de calibre, e a formulação global, mais comum aos matemáticos, a partir de formas diferenciais sobre o fibrado que assumem valores numa álgebra de Lie. Na segunda parte, lidamos com o problema de encontrar soluções para as equações de campo de Yang-Mills formuladas sobre o espaço de Minkowski. Elas são equações diferenciais parciais não lineares de difícil solução. Estudamos e formalizamos matematicamente uma possibilidade de simplificação desse problema, proposta em um artigo por B. M. Schechter. Esta alternativa consiste de procurar soluções que sejam invariantes por transformações conformes do espaço de Minkowski. Para encontrá-las, projetamos o espaço de Minkowski sobre um hipertoro $S^3 \times S^1$, onde as transformações conformes que nos interessam reduzem-se a transformações ortogonais. Nesse novo ambiente, formulamos a teoria de Yang-Mills, deduzimos as equações de campo por um princípio variacional e, impondo s soluções a restrição de invariância por um grupo de transformações ortogonais, reduzimos as equações de Yang-Mills a uma equacão diferencial ordinária de resolução bem mais simples.

# Abstract

This work is divided in two main parts. In the first one we develop the basic theory of connections on principal fiber bundles, trying to make clear the link between the local approach, usual to the physicists in the study of gauge theories, and the global approach, due to the mathematicians, and using Lie algebra-valued differential forms over the bundle. In the second part, we deal with the problem of finding solutions to the Yang-Mills field equations over the Minkowski space. Those are non-linear partial differential equations very hard to solve. We study and detail mathematicaly a way, proposed in a paper by B. M. Schechter, to simplify this problem. This is done by looking for solutions that have conformal invariance. In order to find them, we project the Minkowski space onto a hipertorus $S^3 \times S^1$ where conformal transformations turns into orthogonal ones. In this new environment, once we have formulated the Yang-Mills theory, we deduce the field equations by means of a variational principle and, constraining the solutions to have some orthogonal invariance, we reduce the Yang-Mills equations to a single ordinary differential equation much easier to solve.

# Conteúdo


**1 Fibrados Principais: Conexões e Curvatura** **2**

1.1 Introdução . . . . . 2

1.2 Fibrados Principais . . . . . 3

1.3 Conexões . . . . . 4

1.3.1 Forma Local das Conexões . . . . . 8

1.4 Derivada Covariante e Curvatura . . . . . 10

1.4.1 Levantamento Horizontal e Transporte Paralelo . . . . . 11

1.4.2 A Derivada Covariante e o Tensor de Curvatura . . . . . 14

1.4.3 A Derivada Exterior Covariante e a Curvatura . . . . . 16

1.4.4 Expressão Local da Curvatura . . . . . 20

**2 A Teoria de Yang-Mills Clássica** **22**

2.1 Introdução e Histórico . . . . . 22

2.2 A Formulação da Teoria de Yang-Mills . . . . . 24

2.3 A Teoria de Yang-Mills no Hipertoro . . . . . 27

2.3.1 O Hipertoro e o Espaço de Minkowski . . . . . 28

2.3.2 As Equações de Yang-Mills no Hipertoro . . . . . 35

2.3.3 As Soluções . . . . . 38


-

# Capítulo 1

# Fibrados Principais: Conexões e Curvatura

"[O universo] nao pode ser lido até que tenhamos aprendido a linguagem e nos familiarizado com os caracteres em que ele é escrito. Ele é escrito em linguagem matemática e as letras são triangulos, círculos e outras figuras geométricas, sem as quais é humanamente impossível compreender uma única palavra."
Galileo Galilei (1564 - 1642),

## 1.1 Introdução

Neste capítulo, descreveremos a teoria básica das conexões em fibrados principais. Definiremos o conceito de curvatura de uma conexão partindo de dois pontos de vista diferentes e mostraremos sua equivalência.

Na literatura especializada há pelo menos duas abordagens para o tratamento das conexões que tiveram origens distintas. Da parte dos físicos, no estudo das teorias de calibre, surgiu a necessidade de se acrescentar aos sistemas em estudo, graus de liberdade extras, devidos às simetrias dos objetos considerados. Relacionados a essas simetrias sempre há grupos contínuos de transformações que chamamos de grupos de Lie. Os matemáticos por sua vez, no estudo da geometria diferencial, desenvolveram a estrutura matemática dos fibrados principais, que acopla a cada ponto de uma variedade diferenciável uma cópia de um grupo de Lie, denominada fibra, de forma que o fibrado principal localmente se parece com o produto cartesiano da variedade pelo grupo. Essa estrutura mostrou-se totalmente apropriada à forma-

lização das teorias que os físicos vinham desenvolvendo e permitiu uma compreensão mais aprofundada dos aspectos geométricos envolvidos em tais teorias.

## 1.2 Fibrados Principais

Para definir um fibrado principal, precisamos de uma variedade diferenciável $M$ que chamamos de **espaço base** e que, no caso da nossa aplicação às teorias de calibre, corresponde ao espaço-tempo, onde os fenômenos físicos ocorrem e são observados. A cada ponto $x \in M$, associamos uma fibra $G_x$ difeomorfa a um grupo de Lie $G$, denominado **grupo estrutural** e que, no nosso caso, é escolhido de acordo com as simetrias (propriedades de invariância em relação a certas transformações) dos objetos físicos que desejamos estudar (e.g. campos de interações e de partículas). Unindo todas as fibras sobre $M$, construimos o **espaço total** $P$ de forma que $\dim P = \dim M + \dim G$ com uma **projeção** $\pi : P \to M$, sobrejetiva e diferenciável, que identifica os pontos de $P$ pertencentes a uma mesma fibra, i.e., $\pi^{-1}(\{x\}) = G_x, \forall x \in M$. Precisamos ainda de uma ação à direita, livre e diferenciável do grupo $G$ sobre o espaço $P$, $\alpha : P \times G \to P$, com $\alpha(p, g)$ denotada por $pg$. Por ação livre, queremos dizer que $pg = p \Rightarrow g = e$ (o elemento neutro). Assim, para cada $g \in G$ existe um difeomorfismo $R_g : P \to P$, $R_g(p) = pg$ e para cada $p \in P$, a fibra que passa por $p$ é exatamente a órbita de $p$ pela ação do grupo, i. e., $G_{\pi(p)} = \{pg \mid g \in G\}$.

Agora podemos construir o fibrado principal, escolhendo como identificar cada fibra $G_x \subset P$ com o grupo $G$ [1]. Isso pode ser feito definindo funções chamadas **trivializações locais** de maneira que, para cada $x \in M$ exista um aberto $U \subset M$ com $x \in U$ e um difeomorfismo $T_U : \pi^{-1}(U) \to U \times G$, $T_U(p) = (\pi(p), s_U(p))$, onde $s_U : \pi^{-1}(U) \to G$ satisfaz $s_U(pg) = s_U(p)g$, $\forall g \in G$, $p \in \pi^{-1}(U)$. Cada aplicação $T_U$ é uma trivialização local e na linguagem dos físicos corresponde a uma escolha local de calibre. Decorre daíque o fibrado principal é localmente difeomorfo a $U \times G$ e, no caso em que $U = M$, dizemos que $T_U$ é uma **trivialização global** e o fibrado

[1] Há várias escolhas possíveis, o que equivale a dizer que, em geral, com um mesmo par de espaço base e grupo estrutural podemos construir fibrados principais topologicamente diferentes (não homeomorfos) entre si.

principal que a admite é dito **trivial**.

Suponhamos que dois abertos $U, V \subset M$ de trivializações locais $T_U$ e $T_V$ possuam intersecção não vazia. Então, para um $p \in P$ tal que $\pi(p) \in U \cap V$, é possivel que $T_U(p) \neq T_V(p)$, ou seja, $s_U(p) \neq s_V(p)$. Para esses casos, definimos uma **função de transição** de $T_V$ para $T_U$ por $g_{UV} : U \cap V \to G$, $g_{UV}(x) = s_U(p)s_V(p)^{-1}$, para $p \in \pi^{-1}(\{x\})$, o que não é ambíguo, uma vez que $s_U(pg)s_V(pg)^{-1} = s_U(p)g(s_V(p)g)^{-1} = s_U(p)gg^{-1}s_V(p)^{-1} = s_U(p)s_V(p)^{-1}$. Assim, $s_U(p) = g_{UV}(x)s_V(p)$ com $g_{UU}(x) = e$ $\forall x \in U$, $g_{VU}(x) = g_{UV}(x)^{-1}$ $\forall x \in U \cap V$ e $g_{UV}(x)g_{VW}(x) = g_{UW}(x)$ $\forall x \in U \cap V \cap W$.

Uma maneira equivalente de identificar cada fibra com o grupo estrutural é através da definição de **seções locais** do fibrado principal. Dado um aberto $U \subset M$, uma seção local é uma função diferenciável $\sigma_U : U \to P$ tal que $\pi \circ \sigma_U = 1_U$ (a função identidade em $U$). Assim, dada uma seção local, podemos construir a partir dela a trivialização local $T_U : \pi^{-1}(U) \to U \times G$ tal que $T_U(\sigma_U(x)g) = (x, g)$ (com isso levamos o ponto $\sigma_U(x)$ da fibra sobre x em $e \in G$). Por outro lado, a partir de uma Trivialização local $T_U : \pi^{-1}(U) \to U \times G$ obtemos uma seção local $\sigma_U : U \to P$ por $\sigma_U(x) = T_U^{-1}(x, e)$. Isso estabelece uma correspondência natural entre trivializações locais e seções locais (correspondendo às trivializações globais temos as **seções globais**, quando $U = M$, i.e., $\sigma : M \to P$). É natural agora (e será de utilidade mais adiante) investigarmos o efeito das funções de transição sobre as seções locais. Portanto, sejam $\sigma_U$ e $\sigma_V$ duas seções locais, correspondendo às trivializações locais $T_U$ e $T_V$, tais que $U \cap V \neq \emptyset$. Já vimos que $T_U(p) = (\pi(p), s_U(p))$ e que $T_U(\sigma_U(x)g) = (x, g)$. Para um ponto $p$ com $\pi(p) = x \in U \cap V$, temos $T_U(\sigma_U(x)s_U(p)) = (x, s_U(p)) = T_U(p)$. Então, como $T_U$ é injetiva, $p = \sigma_U(x)s_U(p)$, e pelas mesmas razões $p = \sigma_V(x)s_V(p)$. Logo $\sigma_V(x) = \sigma_U(x)s_U(p)s_V(p)^{-1} = \sigma_U(x)g_{UV}(x)$.

## 1.3 Conexões

Para nós, portanto, os fibrados principais vão servir para englobar numa só variedade os graus de liberdade "externos", devidos ao espaço-tempo (correspondendo

ao espaço base $M$), e os graus de liberdade "internos", devidos às propriedades de simetria dos objetos estudados (correspondendo ao grupo de Lie $G$). Estas idéias ficarão mais claras quando apresentarmos a aplicação delas a situações específicas, mas uma ilustração sugerida pelo professor Márcio pode ser muito esclarecedora também: imagine que estejamos descrevendo o movimento de um lápis no espaço. Então precisamos de três coordenadas para descrever a posição do centro de massa do lápis no espaço e mais duas coordenadas para descrever a atitude do lapis em relação ao centro de massa, pois fixado o centro de massa, a ponta do lápis tem liberdade para mover-se sobre uma superfície esférica. Se usássemos um fibrado principal para descrever esse movimento, as três primeiras coordenadas mencionadas estariam associadas ao espaço base, enquanto que as outras duas estariam ligadas ao grupo de rotações do $\mathbb{R}^3$, $SO(3)$, que como variedade diferenciável é difeomorfo a $S^2$.

Como estamos interessados em definir campos de vetores e de formas diferenciáveis sobre nossa variedade P, temos que falar de espaços tangentes. Chamaremos de $T_xM$ ao espaço vetorial tangente à variedade $M$ no ponto $x$ e analogamente temos $T_pP$, $p \in P$ e $T_gG$, $g \in G$. Será, então muito apropriado decompormos, para cada ponto $p \in P$, o espaço vetorial $T_pP$ como soma direta de um subespaço tangente em p à fibra $G_{\pi(p)}$ (que chamaremos de **subespaço vertical**, $V_p$, de $T_pP$) e um subespaço isomorfo a $T_{\pi(p)}M$ (que chamaremos de **subespaço horizontal**, $H_p$, de $T_pP$).

Utilizando a diferencial da projeção $\pi_* : T_pP \to T_{\pi(p)}M$, o subespaço vertical fica definido por $V_p = \ker(\pi_*)$. Entretanto, o subespaço horizontal não é unicamente determinado e justamente para determinarmos $H_p$ de maneira conveniente é que precisaremos das conexões.

Uma **conexão** em um fibrado Principal é uma decomposição do espaço tangente $T_pP$, em cada ponto $p \in P$, tal que

*(i)* $T_pP = H_p \oplus V_p$,

*(ii)* O subespaço $H_p$ depende diferenciavelmente[2] de $p$ e

[2]Em outras palavras, todo campo vetorial diferenciável $X$ sobre $P$ decompõe-se em campos diferenciáveis $X^H \in \{H_p, p \in P\}$ e $X^V \in \{V_p, p \in P\}$ com $X = X^H + X^V$.

*(iii)* $R_{g*}(H_p) = H_{pg}$.

É interessante observar que a condição *(iii)* vincula os subespaços horizontais em pontos de uma mesma fibra, relacionando-os pelo diferencial da ação à direita do grupo. Assim, estabelecendo o subespaço horizontal de $T_pP$ para um ponto $p \in P$, automaticamente fixamos o subespaço horizontal para todos os outros pontos da fibra, $\{pg, g \in G\}$.

Esta definição de conexão é conveniente para apreendermos seu significado geométrico e sua natureza intrínseca, mas é de pouca utilidade prática. Para separarmos $T_pP$ em $V_p$ e $H_p$ de uma maneira sistemática, vamos utilizar uma 1-forma sobre $P$ que assume valores na álgebra de Lie $\mathcal{G}$. Por esse motivo, será util fazer uma pequena pausa aqui para revisarmos alguns conceitos relativos a grupos de Lie e álgebras de Lie (para um estudo mais detalhado veja [17], [16] ou mesmo [13], §5.6).

Seja $G$ um grupo de Lie. Definimos ações (que são também difeomorfismos) de $G$ em si mesmo, à esquerda $L_g(h) = gh$ e à direita $R_g(h) = hg$, para $g, h \in G$. $L_g$ induz uma transformação linear $L_{g*} : T_hG \to T_{gh}G$ [3].

Dizemos que um campo vetorial $X$ sobre $G$ é **invariante à esquerda** se $L_{g*}X_h = X_{gh}, \forall g, h \in G$. O espaço dos campos vetoriais invariantes à esquerda sobre $G$, munido da operação colchete $[\ ,\ ]$ [4], forma a álgebra de Lie $\mathcal{G}$, que identificamos com $T_eG$, identificando cada campo vetorial com o vetor desse campo que é tangente a G no elemento neutro ($X \to X_e \in T_eG$) e tomando $[X_e, Y_e] = [X, Y]_e$ [5]. Definimos ainda a ação adjunta ad : $G \to G$ por $\mathrm{ad}_g h = ghg^{-1}$. Chamamos a diferencial da ação adjunta de transformação adjunta $\mathrm{Ad}_g : T_hG \to T_{ghg^{-1}}G$ com $\mathrm{Ad}_g(X) =$

[3] Geralmente, o grupo de Lie $G$ é um grupo de matrizes, ou pode ser representado como tal. Nesse caso, a ação de um elemento do grupo não passa da multiplicação por uma matriz, ou seja, é uma transformação linear. Como a derivada de uma transformação linear é ela própria, é muito comum não se fazer distinção entre as ações e suas derivadas, escrevendo por exemplo $Xg$ no lugar de $R_{g*}X$ ou $gXg^{-1}$ para $\mathrm{Ad}_g(X)$. No presente trabalho, preferimos não identificar a derivada da ação de um elemento do grupo com o próprio elemento, que, apesar de ser o mesmo objeto algébrico, possui um significado geométrico diferente.

[4] Dados dois campos vetoriais $X$ e $Y$ sobre uma variedade diferenciável $N$, o campo vetorial $[X, Y]$ é definido por $[X, Y]_p[f] = X_p[Y[f]] - Y_p[X[f]]$, para toda função $f \in C^\infty(N)$, onde $X[f]$ denota a derivada direcional de $f$ na direção do vetor $X$.

[5] De fato, um campo vetorial invariante à esquerda sobre um grupo de Lie $G$ fica completamente determinado por seu valor em $e \in G$, pois dado $X_e \in T_eG$ temos $X_g = X_{ge} = L_{g*}X_e$. Na verdade, nossa identificação é um isomorfismo.

$\mathrm{ad}_{g*}X$. Em particular, quando restrita a $T_eG \simeq \mathcal{G}$, Temos $\mathrm{Ad}_g : \mathcal{G} \to \mathcal{G}$, tal que $A \mapsto gAg^{-1}$, $A \in \mathcal{G}$. Além disso, nos será útil a idéia de curva contida no grupo $G$, para calcularmos vetores tangentes a $G$. Em particular, dado um vetor $A \in \mathcal{G}$ utilizaremos a transformação exponencial $\exp : T_eG \to G$ para construirmos a curva $g(t) = \exp(tA)$ em $G$, que tem $g(0) = e$ e $g'(0) = A$.

Um outro conceito que vamos utilizar mais de uma vez no que se segue, é o de **formas diferenciais a valores em um espaço vetorial** (não simplesmente em $\mathbb{R}$, como de costume). Dada uma variedade diferenciável $N$ e um espaço vetorial $V$, definimos o espaço das $k$-formas sobre $N$ (na verdade são campos de $k$-formas) assumindo valores em $V$ por $\Omega^k(N,V) = \Omega^k(N,\mathbb{R}) \otimes V$, ou seja, cada $k$-forma a valores em $V$ pode ser vista como um vetor de $k$-formas reais. Dada $\alpha \in \Omega^k(N,V)$ e uma base $\{e_1,\ldots,e_n\}$ de $V$, podemos escrever $\alpha(X_1,\ldots,X_k) = \alpha^i(X_1,\ldots,X_k)e_i$ [6], onde $\alpha^i \in \Omega^k(N,\mathbb{R})$, $i = 1,\ldots,n$. Definimos o operador de **derivação exterior** $\mathrm{d} : \Omega^k(N,V) \to \Omega^{k+1}(N,V)$ por $\mathrm{d}\alpha = (\mathrm{d}\alpha^i)e_i$, utilizando a derivada exterior usual para formas reais.

Agora estamos prontos para voltar ao assunto das conexões no nosso fibrado principal. Dado um vetor $A \in \mathcal{G}$, definimos o campo vetorial $A^\sharp$ sobre $P$, por

$$A^\sharp_p = \frac{\mathrm{d}}{\mathrm{dt}}(p\exp(tA))\mid_{t=0}.$$

Com isso, levamos a curva $g(t) = \exp(tA)$ de $G$ para a curva $\gamma(t) = pg(t)$ em $P$ sobre a fibra $G_{\pi(p)}$ e calculamos o vetor $A^\sharp_p = \gamma'(0)$, tangente a $\gamma(t)$ em $p = \gamma(0)$. O campo $A^\sharp$ é chamado de **campo vetorial fundamental** gerado por $A$ e a aplicação $\sharp : \mathcal{G} \to V_p$, $A \mapsto A^\sharp_p$, é um isomorfismo. Temos, portanto, uma maneira prática de construir o subespaço vertical $V_p$ [7].

Para estabelecer o subespaço horizontal $H_p$ definiremos as 1-formas de conexão, que farão a projeção dos vetores de $T_pP$ em $V_p \simeq \mathcal{G}$. Assim, uma **1-forma de conexão** é uma 1-forma $\omega \in \Omega^1(P,\mathcal{G})$ que satisfaz

[6] Estamos usando a convenção de Einstein para somatórios, i.e., se um índice aparece duas vezes em um produto, então fica subentendida uma soma com o índice mencionado percorrendo todos os valores que ele pode assumir (Ex.: $a^ib_i = \sum_{i=1}^n a^ib_i$).

[7] Convém observar que na construção do campo fundamental $A^\sharp$, as únicas características que utilizamos da curva $g(t) = \exp(tA)$ foram os fatos de que $g(0) = e$ e $g'(0) = A$. Portanto, qualquer curva em $G$ com estas características pode ser utilizada.

*(i)* $\omega(A^\sharp) = A$, $\forall A \in \mathcal{G}$ e

(1.1)

*(ii)* $R_g^*\omega = \mathrm{Ad}_{g^{-1}}\omega$ [8].

Dada uma 1-forma de conexão $\omega$, definimos o subespaço horizontal de $T_pP$ por $H_p = \ker(\omega) = \{X \in T_pP \mid \omega(X) = 0\}$. Para que este subespaço $H_p$ esteja de acordo com nossa definição de conexão, ele deve satisfazer, como já vimos, $R_{g*}(H_p) = H_{pg}$. E de fato, tomando $X \in H_p \subset T_pP$, temos $R_{g*}X \in T_{pg}P$ e $\omega(R_{g*}X) = g^{-1}\omega(X)g = 0$, pois $\omega(X) = 0$. Logo $R_{g*}X \in H_{pg}$ e então $R_{g*}(H_p) \subset H_{pg}$. Tomando $Y \in H_{pg}$, por raciocínio análogo vemos que $R_{g^{-1}*}(Y) \in H_p$ e como a transformada inversa de $R_{g*}$ é justamente $R_{g^{-1}*}$, temos $Y \in R_{g*}(H_p)$ e assim $H_{pg} \subset R_{g*}(H_p)$, o que nos leva a concluir que $R_{g*}(H_p) = H_{pg}$, exatamente como se requer de uma conexão. Por outro lado, convém observar que dada uma conexão, podemos obter de maneira única sua 1-forma de conexão, fazendo $\omega(X) = \omega(X^H + X^V) = \sharp^{-1}(X^V)$, onde $\sharp$ é o isomorfismo já definido entre $\mathcal{G}$ e $V_p$.

### 1.3.1 Forma Local das Conexões

Agora que definimos as 1-formas de conexão, vamos ver como elas se manifestam no espaço base $M$. Para isso, dada uma seção local $\sigma_U : U \to P$, $U \subset M$ aberto, consideremos uma 1-forma $\omega_U \in \Omega^1(U, \mathcal{G})$ definida por $\omega_U = \sigma_U^*\omega$. Assim, dadas seções locais vindo de uma coleção de abertos que cobre M, teremos 1-formas locais de conexão definidas sobre cada ponto de $M$. Seria muito interessante se pudéssemos fazer o caminho inverso, i.e., partindo de 1-formas de conexão definidas em abertos que formam uma cobertura da base, recuperar uma única 1-forma de conexão, definida globalmente no espaço total $P$ e cujo pull-back via seções locais dê exatamente as 1-formas nos abertos da base, de onde partimos. Veremos que isso é possível, desde que definamos as 1-formas nos abertos de $M$ satisfazendo certas condições, e então teremos uma completa equivalência entre as duas definições de 1-forma de conexão (global, em $P$ ou local, em $M$).

[8] Isto significa que $R_g^*\omega_{pg}(X) = \omega_{pg}(R_{g*}X) = L_{g*}^{-1}R_{g*}\omega_p(X), \forall X \in T_pP$.

Primeiro, observemos nossa situação anterior, em que temos uma 1-forma de conexão em $P$ e construímos as 1-formas sobre abertos de $M$. Suponhamos que duas seções locais $\sigma_U : U \to P$ e $\sigma_V : V \to P$ com $U \cap V \neq \emptyset$ nos levem às 1-formas locais $\omega_U = \sigma_U^* \omega$ e $\omega_V = \sigma_V^* \omega$. Vejamos que relação existe entre elas, relacionando $\sigma_U^*$ e $\sigma_V^*$ nos pontos de $U \cap V$.

Quando definimos seções locais, mostramos que $\sigma_V(x) = \sigma_U(x) g_{UV}(x)$ e para chegar a uma relação entre os pull-backs, precisamos das diferenciais dessas seções. Tomemos uma curva $\gamma : \mathbb{R} \to M$ com $\gamma(0) = x$ e $\gamma'(0) = Y \in T_x M$. Então,

$$\begin{aligned}
\sigma_{V*} Y &= \frac{\mathrm{d}}{\mathrm{dt}} \sigma_V(\gamma(t)) \mid_{t=0} = \frac{\mathrm{d}}{\mathrm{dt}} [\sigma_U(\gamma(t)) g_{UV}(\gamma(t))]_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{dt}} [\sigma_U(x) g_{UV}(\gamma(t))]_{t=0} + \frac{\mathrm{d}}{\mathrm{dt}} [\sigma_U(\gamma(t)) g_{UV}(x)]_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{dt}} [\sigma_V(x) g_{UV}(x)^{-1} g_{UV}(\gamma(t))]_{t=0} + \frac{\mathrm{d}}{\mathrm{dt}} [R_{g_{uv}(x)}(\sigma_U(\gamma(t)))]_{t=0} ,
\end{aligned}$$

onde trocamos $\sigma_U(x)$ por $\sigma_V(x) g_{UV}(x)^{-1}$ para obtermos um vetor fundamental tangente a $P$ em $\sigma_V(x)$ (considerando o ponto $\sigma_V(x) \in P$ e a curva em $G$ dada por $g_{UV}^{-1}(x) g_{UV}(\gamma(t)) = L^{-1}_{g_{uv}(x)} g_{UV}(\gamma(t))$ ). Segue-se que

$$\sigma_{V*} Y = [L^{-1}_{g_{uv}(x)*} g_{UV*} Y]^{\sharp}_{\sigma_V(x)} + R_{g_{uv}(x)*} \sigma_{U*} Y. \tag{1.2}$$

Portanto, usando as propriedades da conexão $\omega$, temos[9]

$$\omega_V(Y) = \sigma_V^* \omega = \omega(\sigma_V Y) = L^{-1}_{g_{uv}(x)*} g_{UV*} Y + \mathrm{Ad}_{g_{uv}(x)^{-1}} \omega_U(Y). \tag{1.3}$$

Resumindo, se temos uma 1-forma de conexão definida globalmente no espaço total $P$, podemos "puxá-la" para abertos da base via pull-back por seções locais e as 1-formas que obtemos nos abertos da base, se relacionam pela eq. (1.3), para pontos na intersecção de dois abertos.

Para fazer o caminho inverso (obter uma 1-forma $\omega$ em $P$ a partir de 1-formas dadas sobre abertos que cobrem a base $M$), o mínimo que devemos exigir das 1-formas locais, é que satisfaçam a eq. (1.3) nos pontos de intersecção entre entre

[9]Notemos que se $G$ é um grupo de matrizes,a lei de mudança de calibre (1.3) pode ser reescrita como $\omega_V = g_{UV}^{-1} \mathrm{d} g_{UV} + g_{UV}^{-1} \omega_U g_{UV}$ .

os abertos. Na verdade esta exigência é suficiente para construirmos uma 1-forma global $\omega$ sobre $P$. A construção pode ser feita da seguinte maneira: para cada 1-forma local de conexão $\omega_U$ e seção local $\sigma_U : U \to P$, definimos uma 1-forma $\omega^U \in \Omega^1(\pi^{-1}(U), \mathcal{G})$, tal que para $x \in U$, $Y \in T_xM$, $p = \sigma_U(x)$ e $A \in \mathcal{G}$ tenhamos $\omega^U(\sigma_{U*}Y + A_p^\sharp) = \omega^U(\sigma_{U*}Y) + \omega^U(A_p^\sharp) = \omega_U(Y) + A$. Isso define $\omega^U$ apenas em $\sigma_U(U)$, mas como desejamos fazer de $\omega^U$ uma 1-forma de conexão impomos as condições (1.1) de modo que para um vetor $X_{pg} \in T_{pg}P$, podemos escrever $\omega^U(X_{pg}) = \omega^U(R_{g*}R_{g*}^{-1}X_{pg}) = R_g^*\omega^U(R_{g*}^{-1}X_{pg}) = \mathrm{Ad}_{g^{-1}*}\omega^U(R_{g*}^{-1}X_{pg})$. Repetindo esta construção para os outros abertos da cobertura de $M$ (com suas respectivas 1-formas de conexão e seções locais), obteremos 1-formas de conexão em $T_pP$ para cada $p \in P$. O último passo para mostrar que assim construimos uma 1-forma global de conexão, é mostrar que se $U \cap V \neq \emptyset$ temos $\omega^U = \omega^V$ sobre $\pi^{-1}(U \cap V)$. Para $A \in \mathcal{G}$, temos naturalmente $\omega^U(A_p^\sharp) = \omega^V(A_p^\sharp) = A$. Para $Y \in T_xM$, tomando o vetor $\sigma_{V*}Y \in T_pP$, temos por um lado $\omega^V(\sigma_{V*}Y) = \omega_V(Y)$ e por outro lado $\omega^U(\sigma_{V*}Y) = \omega^U([L^{-1}_{g_{UV}(x)*}g_{UV*}Y]^\sharp_{\sigma_V(x)} + R_{g_{UV}(x)*}\sigma_{U*}Y) = L^{-1}_{g_{UV}(x)*}g_{UV*}Y + \mathrm{Ad}_{g_{UV}(x)^{-1}}\omega_U(Y) = \omega_V(Y)$, onde utilizamos (1.2) na primeira igualdade e (1.3) na última. Daí é direta a conclusão de que $\omega^U = \omega^V$ sobre $\pi^{-1}(U \cap V)$.

Resumindo novamente, vemos que tanto podemos definir a conexão via 1-forma globalmente definida em $P$ quanto por 1-formas locais definidas sobre abertos de $M$.

## 1.4 Derivada Covariante e Curvatura

Na seção anterior fizemos a decomposição do espaço tangente ao espaço total $P$ nos subespaços horizontal e vertical, de maneira que um vetor $X \in T_pP$, decompõe-se como $X = X^H + X^V$. Vamos agora definir um operador de derivação que nos permita medir a taxa de variação de campos de formas sobre o fibrado, em direções horizontais (em $H_p$). Isso desperta interesse porque, como já vimos o subespaço vertical está associado aos graus de liberdade internos (de transformações no grupo de simetria), enquanto que o subespaço horizontal está associado aos graus de liberdade

do espaço base.

Uma ilustração que pode clarear o significado geométrico e é inspirada na teoria da relatividade geral de Einstein, é a de que a presença de um potencial (representado pela conexão) deforma o espaço base, sendo que sua versão "deformada" apresenta-se mergulhada no espaço total $P$ e tem $H_p$ como espaço tangente em cada ponto $p$.

É necessário aqui um certo cuidado, pois apesar de esta ilustração ser útil, ela pode tornar-se perigosa se a levarmos a sério demais. De fato, por um ponto $p \in P$ podemos passar caminhos horizontais (curvas cujo vetor tangente em cada ponto está em $H_p$), entretanto, em geral não existe uma hipersuperfície horizontal contendo $p$ e difeomorfa a um aberto de $M$ que contém $\pi(p)$. Em outras palavras, o conjunto dos hiperplanos horizontais $\{H_p \mid p \in P\}$ em geral não é integrável, em qualquer vizinhança de um ponto $p$.

Assim, nossa "versão deformada de $M$" pode não corresponder a uma variedade diferenciável difeomorfa a $M$, mesmo localmente. Na verdade isso só será possível para conexões planas, que definiremos posteriormente.

Para derivar em direções horizontais no fibrado $P$ é que vamos definir uma regra de derivação covariante, que depende da conexão, uma vez que esta é quem determina quem é o subespaço horizontal.

A expressão "derivar em direções horizontais" leva-nos a duas abordagens para a definição da derivada covariante. Uma delas usa a idéia de derivada direcional (tomamos um caminho horizontal e derivamos ao longo do caminho) e o fato de que localmente temos um difeomorfismo $T_U : \pi^{-1}(U) \to U \times \mathcal{G}$. A outra abordagem, utilizando o cálculo diferencial exterior está associada à idéia de calcularmos o componente horizontal da derivada de uma $k$-forma. Desenvolveremos aqui as duas abordagens independentemente e depois mostraremos como estão relacionadas.

### 1.4.1 Levantamento Horizontal e Transporte Paralelo

A primeira abordagem lida principalmente com o fato de que a conexão impõe uma "alteração" nos sistemas locais de coordenadas e, uma vez que a derivada

covariante é tratada como uma derivada direcional em $P$, ela pode ser identificada com um campo vetorial, cuja direção em $T_pP$ depende da conexão. Após definirmos o que seja o levantamento horizontal de vetores tangentes a $M$, poderemos definir a derivada covariante como o levantamento horizontal de uma derivada direcional em $M$.

Dada uma curva $\gamma$ no espaço base $M$, será interessante definirmos o que seja o transporte paralelo de objetos do fibrado $P$ ao longo de $\gamma$. Quando transportarmos paralelamente uma fibra de $P$ ou um vetor, por exemplo, desejaremos que seu componente vertical permaneça constante. Para isso, será necessário percorrer uma curva de $P$ cuja projeção sobre $M$ seja $\gamma$ e cuja derivada tenha componente vertical nula. Fisicamente esta idéia equivale a fazer um movimento no campo de calibre sem gastar energia com mudança de posição (ou orientação) em relação à simetria interna.

Quando estabelecemos, para cada ponto $p \in P$, o subespaço horizontal $H_p \subset T_pP$, criamos a possibilidade de definir campos de vetores horizontais sobre $P$ ($Y$ é um campo horizontal se e somente se $Y_p \in H_p$, $\forall p \in P$). Além disso, a derivada da projeção $\pi : P \to M$, restrita a $H_p$, nos fornece um isomorfismo $\pi_* : H_p \to T_{\pi(p)}M$. Então, dado um campo vetorial $X$ sobre $M$, podemos definir o **levantamento horizontal** de X como sendo o campo $\tilde{X}$ sobre P tal que $\tilde{X}_p \in H_p$ e $\pi_*(\tilde{X}_p) = X_{\pi(p)}$, $\forall p \in P$. Do isomorfismo mencionado acima, entre $H_p$ e $T_{\pi(p)}M$, concluimos que $\tilde{X}$ existe e é único.

Dada uma curva diferenciável $\gamma : [0,1] \to M$ e um ponto $p \in \pi^{-1}(\{\gamma(0)\})$, dizemos que o **levantamento horizontal de $\gamma$ por $p$** é a curva $\tilde{\gamma} : [0,1] \to P$ tal que $\pi \circ \tilde{\gamma} = \gamma$, $\tilde{\gamma}(0) = p$ e $\tilde{\gamma}'(t) \in H_{\tilde{\gamma}(t)}$ $\forall t \in [0,1]$. Aqui é importante notar que dado um campo vetorial $X$ sobre $M$, os levantamentos horizontais das curvas integrais de X são as curvas integrais de $\tilde{X}$, uma vez que $\pi(\tilde{\gamma}(t)) = \gamma(t)$ implica em $\pi_*\tilde{\gamma}'(t) = \gamma'(t)$. Então a existência e unicidade do levantamento horizontal de campos de vetores garante a existencia e unicidade do levantamento horizontal de curvas. Isso decorre também do teorema de existência e unicidade para equações diferenciais ordinárias (EDO), pois dada a curva $\gamma$, seu levantamento $\tilde{\gamma}$ pode ser obtido de uma EDO, como mostramos a seguir.

Tomemos uma seção local $\sigma_U$ tal que $\gamma(0) \in U$ e $\sigma_U(\gamma(0)) = p$. Então podemos escrever

$$\tilde{\gamma}(t) \;=\; \sigma_U(\gamma(t))g(t) \tag{1.4}$$

para alguma curva diferenciável $g : [0,1] \to G$ com $g(0) = e$, pois $\pi(\tilde{\gamma}(t)) = \pi(\sigma_U(\gamma(t))) = \gamma(t)$. Fazendo $\alpha(t) = \sigma_U(\gamma(t))$ e derivando a expressão acima, calculamos

$$\begin{aligned}
\tilde{\gamma}'(t) \;&=\; \frac{\mathrm{d}}{\mathrm{d}s}[\alpha(s)g(s)]_{s=t} \;=\; \frac{\mathrm{d}}{\mathrm{d}s}[\alpha(s)g(t)]_{s=t} \;+\; \frac{\mathrm{d}}{\mathrm{d}s}[\alpha(t)g(s)]_{s=t} \\
&=\; \frac{\mathrm{d}}{\mathrm{d}s}[R_{g(t)}\alpha(s)]_{s=t} \;+\; \frac{\mathrm{d}}{\mathrm{d}s}[\alpha(t)g(t)g(t)^{-1}g(s)]_{s=t} \\
&=\; R_{g(t)*}\alpha'(t) \;+\; [L^{-1}_{g(t)*}g'(t)]^{\sharp}_{\alpha(t)g(t)} \;,
\end{aligned}$$

onde reescrevemos a segunda parcela de modo a obter um vetor fundamental em $T_{\tilde{\gamma}(t)}P$. Como queremos que $\tilde{\gamma}'(t)$ seja horizontal, devemos ter $\omega(\tilde{\gamma}'(t)) = 0$, o que significa, em termos da expressão acima,

$$R^{*}_{g(t)}\omega(\alpha'(t)) \;+\; L^{-1}_{g(t)*}g'(t) \;=\; 0,$$

e por (1.1) equivale a escrever

$$\mathrm{Ad}_{g^{-1}}\omega(\alpha'(t)) \;+\; L^{-1}_{g(t)*}g'(t) \;=\; L^{-1}_{g(t)*}R_{g(t)*}\omega(\alpha'(t)) \;+\; L^{-1}_{g(t)*}g'(t) \;=\; 0.$$

Aplicamos $L_{g(t)*}$ e isolamos $g'(t)$, chegando a

$$\begin{aligned}
g'(t) \;&=\; -R_{g(t)*}\omega(\alpha'(t)) \;=\; -R_{g(t)*}\omega[\sigma_{U*}(\gamma'(t))] \\
&=\; -R_{g(t)*}\sigma^{*}_{U}\omega(\gamma'(t)) \;=\; -R_{g(t)*}\omega_U(\gamma'(t)).
\end{aligned}$$

Sabemos que $\omega_U \in \Omega^1(U, \mathcal{G})$ e portanto podemos escrever $\omega_U = A_\mu \mathrm{d}x^\mu$, $A_\mu \in \mathcal{G}$. Se $\gamma(t) = (x_0(t), ..., x_4(t))$ e $\gamma' = (\dot{x}_0, ..., \dot{x}_4) = \dot{x}_\mu \frac{\partial}{\partial x_\mu}$, no caso em que $G$ é um grupo de matrizes, a equação acima torna-se

$$g'(t) \;=\; -A_\mu \dot{x}_\mu g(t),$$

que é uma equação diferencial linear para $g(t)$ (com condição inicial dada por $g(0) \;=\; e$), cuja solução substituída em (1.4) nos dá o levantamento horizontal de $\gamma(t)$.

Observemos que ao transportarmos um ponto $p \in P$ horizontalmente sobre uma curva $\gamma \subset M$, na verdade transportamos paralelamente toda a fibra que contém p, pois na definição de conexão vimos que $H_{pg} = R_{g*}H_p$. Ou seja, cada ponto $pg$ da fibra por $p$ é transportado ao longo de uma curva horizontal $\tilde{\gamma}(t)g = R_g(\tilde{\gamma}(t))$, com vetor tangente $R_{g*}\tilde{\gamma}'(t) \in H_{\tilde{\gamma}(t)g}$.

### 1.4.2 A Derivada Covariante e o Tensor de Curvatura

Seja $\sigma_U : U \to P$ uma seção local e $\omega$ a 1-forma de conexão. Como vimos, em cada ponto $x \in U$ temos $\omega_U = \sigma_U^* \omega = A_\mu \mathrm{d}x^\mu$, $A_\mu \in \mathcal{G}$ e como as 1-formas $\mathrm{d}x^\mu$ são duais aos vetores $\partial_\mu = \frac{\partial}{\partial x_\mu} \in T_xM$, i. e., $\mathrm{d}x^\mu(\frac{\partial}{\partial x_\nu}) = \delta^\mu_\nu$, então $\omega_U(\partial_\mu) = A_\mu$. Por outro lado, $\omega_U(\partial_\mu) = [\sigma_U^*\omega](\partial_\mu) = \omega(\sigma_{U*}\partial_\mu)$. Ou seja, $\omega(\sigma_{U*}\partial_\mu) = A_\mu = \omega(A^\sharp_\mu) \Rightarrow \omega(\sigma_{U*}\partial_\mu - A^\sharp_\mu) = 0$. Portanto $\sigma_{U*}\partial_\mu - A^\sharp_\mu \in H_{\sigma(x)}$ e é o levantamento horizontal de $\partial_\mu$ pelo ponto $p = \sigma(x)$. Por esse motivo escreveremos $\tilde{\partial}_\mu|_p = \sigma_{U*}\partial_\mu - A^\sharp_\mu$. Além disso, já vimos que a projeção induz um isomorfismo $\pi_* : H_p \to T_xM$ e como $\pi_*\tilde{\partial}_\mu = \partial_\mu \in T_xM$, o fato de os vetores $\partial_\mu$ formarem uma base para $T_xM$ implica que os $\tilde{\partial}_\mu$ formam uma base para $H_p$. Então qualquer vetor horizontal e consequentemente qualquer derivação em direção horizontal, pode ser escrita como combinação linear dos vetores $\tilde{\partial}_\mu|_p = \sigma_{U*}\partial_\mu - A^\sharp_\mu$.

Agora recordemo-nos de que à seção $\sigma_U$ está associada uma trivialização local $T_U : \pi^{-1}(U) \to U \times \mathcal{G}$ tal que $T_U(\sigma_U(x)) = (x, e), \forall x \in U$. Como $T_U$ é um difeomorfismo, para um ponto $p = \sigma_U(x), x \in U$, o diferencial $T_{U*}$ é um isomorfismo entre $T_pP$ e $T_xM \oplus \mathcal{G}$, o que nos permite identificar $\sigma_{U*}\partial_\mu - A^\sharp_\mu$ com $\partial_\mu - A_\mu$. Como $A_\mu \in \mathcal{G}$, podemos escrever $A_\mu = A^a_\mu \tau_a$, onde $\tau_a$ são geradores de $\mathcal{G}$ que formam uma base ortonormal. Assim, $T_{U*}(\sigma_{U*}\partial_\mu - A^\sharp_\mu) = \partial_\mu - A_\mu = \partial_\mu - A^a_\mu \tau_a \in T_xM \oplus \mathcal{G}$.

Portanto, dado um fibrado principal, uma seção local $\sigma_U : U \to P$ e uma conexão em $\pi^{-1}(U)$ representada por $A_\mu$, para cada direção em $M$ determinada pelos vetores $\partial_\mu$, definimos a **derivada covariante** por

$$\mathcal{D}_\mu = \partial_\mu - A_\mu = T_{U*}(\sigma_{U*}\partial_\mu - A^\sharp_\mu) \,. \tag{1.5}$$

Esta é a linguagem usual dos físicos para a definição de derivada covariante e surgiu antes da utilização do formalismo de fibrados principais. Com a introdução

desse formalismo, o símbolo $\partial_\mu - A_\mu$ tornou-se praticamente apenas uma notação diferente para $\sigma_{U*}\partial_\mu - A^\sharp_\mu$, que é quem realmente mora em $T_pP$.

Outra observação interessante é a de que localmente (em um aberto $U \subset M$), definir uma conexão no fibrado, i.e., escolher quem é o subespaço horizontal em cada ponto equivale a fornecer os coeficientes $A_\mu : M \to \mathcal{G}$ da 1-forma local de conexão $\omega_U$. É muito comum, principalmente entre os físicos, a identificação da conexão com esses $A_\mu$, denominados potenciais de calibre.

Não podemos perder de vista que esta formulação para a derivada covariante que desenvolvemos até aqui é dependente da seção local $\sigma_U$. Logo, se passarmos para outros pontos da fibra, dados por $\sigma_U(x)g(x) = \sigma_V(x)$, pela mesma construção que já fizemos, teremos $\mathcal{D}_\mu = \partial_\mu - B_\mu$, sendo $B_\mu$ os componentes da 1-forma $\omega_V$ e por (1.3), temos

$$B_\mu = \omega_V(\partial_\mu) = L^{-1}_{g*}\partial_\mu g + \mathrm{Ad}_{g^{-1}} A_\mu \tag{1.6}$$

e se $G$ for um grupo de matrizes significa o mesmo que

$$B_\mu = g^{-1}\partial_\mu g + g^{-1}A_\mu g. \tag{1.7}$$

Diante disso, é importante percebermos que se tivermos uma conexão e uma seção local $\sigma_U$ tais que os coeficientes $A_\mu$ sejam identicamente nulos em U, ao passarmos para outra seção $\sigma_V(x) = \sigma_U(x)g(x)$, teremos $\omega_V = B_\mu \mathrm{d}x^\mu$, com $B_\mu = L^{-1}_{g(x)*}\partial_\mu g(x)$. Além disso, o levantamento horizontal de $\partial_\mu$ fica reduzido a $\sigma_{U*}\partial_\mu$. Às conexões da forma $L^{-1}_{g*}\partial_\mu g$, onde $g : M \to G$ é uma função diferenciável, chamaremos "**conexões planas**", nome que ganhará mais significado ao definirmos o que é a curvatura de uma conexão. No caso da conexão nula temos $[\mathcal{D}_\mu, \mathcal{D}_\nu] = [\sigma_{U*}\partial_\mu, \sigma_{U*}\partial_\nu] = \sigma_{U*}[\partial_\mu, \partial_\nu] = 0$ e é fácil verificar que o comutador das derivadas covariantes é nulo para qualquer conexão plana. Entretanto, para outras conexões em geral temos

$$[\mathcal{D}_\mu, \mathcal{D}_\nu] = -\partial_\mu A_\nu + \partial_\nu A_\mu - [A_\mu, A_\nu]\ , \tag{1.8}$$

resultado que mostraremos posteriormente, quando tivermos apresentado outros elementos que possibilitarão uma demonstração mais direta. Esta última equação sugere a definição de um tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu], \tag{1.9}$$

denominado **tensor de curvatura** porque num certo sentido, avalia a curvatura da conexão, ou seja, a curvatura imposta ao subespaço horizontal pela conexão. Notemos que $F_{\mu\nu}$ é uma matriz de elementos de $\mathcal{G}$ e contém informação sobre o quanto as derivadas covariantes em diferentes direções (e consequentemente os transportes paralelos) falham em comutar.

### 1.4.3 A Derivada Exterior Covariante e a Curvatura

Nossa outra abordagem para a derivação covariante possui a vantagem de ser global e lidar com formas diferenciais globais sobre $P$ e não apenas com campos de vetores em sistemas locais de coordenadas.

Como as formas que vamos utilizar assumem valores na álgebra de Lie $\mathcal{G}$ precisamos rever rapidamente alguns conceitos que nos serão necessários. Omitiremos as demonstrações mais diretas e mais detalhes podem ser encontrados em [3].

Já definimos o operador de derivação exterior para formas a valores em um espaço vetorial, que em particular pode ser $\mathcal{G}$. Dada uma variedade diferenciável $N$ e duas formas $\varphi \in \Omega^p(N, \mathcal{G})$, e $\psi \in \Omega^q(N, \mathcal{G})$, definimos o **comutador**

$$[\varphi, \psi](X_1, ..., X_{p+q}) = \frac{1}{p!q!} \sum_{\sigma} (-1)^{\sigma} [\varphi(X_{\sigma(1)}, ..., X_{\sigma(p)}), \psi(X_{\sigma(p+1)}, ..., X_{\sigma(p+q)})], \tag{1.10}$$

onde $\sigma$ percorre todas as permutações de $\{1, 2, ..., p+q\}$[10]; $(-1)^{\sigma}$ é o sinal da permutação (se ela é par ou ímpar) e no segundo membro utilizamos o comutador de $\mathcal{G}$.

Se $\{\tau_1, \tau_2, ..., \tau_n\}$ é uma base de $\mathcal{G}$, já vimos que podemos escrever $\varphi = \varphi^i \tau_i$ e $\psi = \psi^j \tau_j$ $(i, j = 1, 2, ..., n)$, $\varphi^i \in \Omega^p(N, \mathbb{R})$, $\psi^j \in \Omega^q(N, \mathbb{R})$. Com isso a expressão (1.10) leva-nos a

$$[\varphi, \psi] = \sum_{i,j} (\varphi^i \wedge \psi^j)[\tau_i, \tau_j] \tag{1.11}$$

e com o auxílio das constantes de estrutura $c_{ij}^k$ (definidas por $[\tau_i, \tau_j] = c_{ij}^k \tau_k$) resulta

[10] Esta é a maneira natural de tornar o resultado antissimétrico, como devem ser as $(p+q)$-formas diferenciais.

em

$$[\varphi, \psi] = \sum_{i,j,k} c_{ij}^k (\varphi^i \wedge \psi^j) \tau_k. \tag{1.12}$$

Em particular, se $\mathcal{G}$ é uma álgebra de matrizes (e nesse caso $[A, B] = AB - BA$), temos

$$[\varphi, \psi] \;=\; \varphi \wedge \psi \;-\; (-1)^{ij} \psi \wedge \varphi. \tag{1.13}$$

E agora, fixada uma conexão, podemos definir a **derivada exterior covariante** por

$$\begin{array}{rccc} D: & \Omega^k(P, \mathcal{G}) & \to & \Omega^{k+1}(P, \mathcal{G}) \\ & \varphi & \mapsto & D\varphi = (\mathrm{d}\varphi)^H \end{array}$$

onde $\alpha^H(X_1, \ldots, X_k) = \alpha(X_1^H, \ldots, X_k^H), \forall \alpha \in \Omega^k(P, \mathcal{G}), k = 1, 2, ..., dim(P)$.

Definimos então a **curvatura** da conexão (ou a 2-forma de curvatura) como a derivada exterior covariante da 1-forma de conexão, denotada por

$$\Omega = D\omega \in \Omega^2(P, \mathcal{G}).$$

Para calcularmos curvatura diretamente a partir da 1-forma de conexão, mostraremos que vale a **equação estrutural de Cartan**

$$\Omega = \mathrm{d}\omega + \frac{1}{2}[\omega, \omega]. \tag{1.14}$$

Dados $X, Y \in T_pP$, temos por (1.10) que

$$[\omega, \omega](X, Y) = [\omega(X), \omega(Y)] - [\omega(Y), \omega(X)] = 2[\omega(X), \omega(Y)]$$

e então provar (1.14), equivale a provar que

$$\Omega(X, Y) = \mathrm{d}\omega(X, Y) + [\omega(X), \omega(Y)],$$

$\forall X, Y \in T_pP$. Já sabemos que podemos decompor $X = X^H + X^V$, $Y = Y^H + Y^V$ e que todos os termos da equação estrutural são bilineares e antissimétricos. Então basta mostrar que ela vale para três casos particulares: $X$ e $Y$ horizontais; $X$ horizontal e $Y$ vertical; $X$ e $Y$ verticais.

Para isso, precisaremos de dois resultados auxiliares. Um deles é o fato de que se $\alpha \in \Omega^k(P)$, $d\alpha$ pode ser definida como a $(k+1)$-forma tal que para quaisquer campos vetoriais $X_1, X_2, ..., X_{k+1} \in \Gamma(TP)$ [11], temos

$$d\alpha(X_1, ..., X_{k+1}) = \sum_{i=1}^{k+1} (-1)^{i+1} X_i[\alpha(X_1, ..., \hat{X}_i, ..., X_{k+1})]$$

$$+ \sum_{1 \leq i < j \leq k+1} (-1)^{i+j} \alpha([X_i, X_j], X_1, ..., \hat{X}_i, ..., \hat{X}_j, ..., X_{k+1}) \ , \qquad (1.15)$$

onde os termos sob acento circunflexo são omitidos na sequência. A demonstração desta expressão pode ser encontrada em [11] ou [20] e para [8] ela é a própria definição de derivação exterior. O outro resultado de que vamos precisar é que se $X$ é um campo vetorial horizontal sobre $P$ e $Y$ é um campo vetorial vertical, então, $[X, Y]$ é um vetor horizontal. De fato, se a curva integral de $Y$ pelo ponto $p$ é $pg(t)$, temos[12]

$$[X, Y]_p = -[Y, X]_p = \lim_{t \to 0} \left( \frac{1}{t}(X_p) - R^{-1}_{g(t)*} X_{pg(t)} \right)$$

e como já vimos, a transformação $R^{-1}_{g*} = R^{-1}_{(g^{-1})*}$ leva $H_{pg}$ em $H_p$. Portanto a diferença de vetores que aparece no limite acima está sempre em $H_p$ e, por isso, $[X, Y] \in H_p$.

Voltando à equação estrutural, se $X$ e $Y$ são horizontais, então $X = X^H, Y = Y^H$ e $\omega(X) = \omega(Y) = 0$. Logo a equação estrutural de Cartan fica reduzida à própria definição da curvatura, i. e., $\Omega(X, Y) = d\omega(X^H, Y^H) = D\omega$.

Se $X$ for horizontal e $Y$ vertical, $Y^H = 0 \Rightarrow \Omega(X, Y) = 0$ e $\omega(X) = 0 \Rightarrow [\omega(X), \omega(Y)] = 0$. Temos que mostrar, então, que $d\omega(X, Y) = 0$. Podemos supor que $Y = A^\sharp_p$, onde $A^\sharp$ é um campo fundamental e a equação (1.15) no caso da 1-forma $\omega$ reduz-se a

$$d\omega(X, A^\sharp) = X\omega(A^\sharp) - A^\sharp\omega(X) - \omega([X, A^\sharp]).$$

[11] $\Gamma(TP)$ aqui denota o conjunto de todos os campos vetoriais sobre $P$, i.e., todas as seções do fibrado tangente a $P$.

[12] Aqui estamos utilizando o fato de que $[X, Y] = \mathcal{L}_X Y$, a derivada de Lie do campo vetorial $Y$ ao longo do fluxo do campo $X$ (veja, por exemplo, [11] ou [13]).



UNICAMP
BIBLIOTECA CENTRAL

Já vimos que $[X, A^\sharp]$ é necessariamente horizontal e $X \in H_p$. Logo $\omega([X,Y]) = 0 = \omega(X)$. Como $\omega(A^\sharp) = A \in \mathcal{G}$ é uma função constante sobre $P$, $X\omega(A^\sharp) = 0$ e então $\mathrm{d}\omega(X, A^\sharp) = 0$, como queríamos mostrar.

Se $X$ e $Y$ forem verticais, podemos extender ambos por campos fundamentais de forma que $X = A^\sharp_p$ e $Y = B^\sharp_p$. Como ambos são verticais, $\Omega(X,Y) = 0$. Resta então mostrar que $\mathrm{d}\omega(X,Y) + [\omega(X), \omega(Y)] = 0$, com $\mathrm{d}\omega(A^\sharp, B^\sharp) = A^\sharp\omega(B^\sharp) - B^\sharp\omega(B^\sharp) - \omega([A^\sharp, B^\sharp]) = -\omega([A^\sharp, B^\sharp])$. Além disso, como a transformação $\sharp$ é um isomorfismo, $[A,B]^\sharp = [A^\sharp, B^\sharp]$ e $\omega([A^\sharp, B^\sharp]) = \omega([A,B]^\sharp) = [A,B] = [\omega(A^\sharp), \omega(B^\sharp)]$. Segue-se que o lado direito da equação estrutural de Cartan também se anula, e é o que queríamos.

Da definição do comutador para formas em (1.10), fica evidente que o pull-back do comutador (por qualquer transformação linear) é igual ao comutador dos pull-backs das formas ($F^*[\varphi, \psi] = [F^*\varphi, F^*\psi]$). Então podemos utilizar a equação estrutural e (1.1) para escrever

$$\begin{aligned} R_g^*\Omega \;&=\; R_g^*(\mathrm{d}\omega + \frac{1}{2}[\omega,\omega]) \;=\; \mathrm{d}(R_g^*\omega) + \frac{1}{2}[R_g^*\omega, R_g^*\omega] \\ &= \mathrm{d}(\mathrm{Ad}_{g^{-1}}\omega) + \frac{1}{2}[\mathrm{Ad}_{g^{-1}}\omega, \mathrm{Ad}_{g^{-1}}\omega] \;=\; \mathrm{Ad}_{g^{-1}}(\mathrm{d}\omega + \frac{1}{2}[\omega,\omega]) \\ &= \mathrm{Ad}_{g^{-1}}\Omega\ ,\ \forall g \in G. \end{aligned} \tag{1.16}$$

Outra característica relevante da 2-forma de curvatura $\Omega$ é o fato de sua derivada exterior covariante ser nula: pela equação estrutural de Cartan, $\mathrm{d}\Omega = \mathrm{d}^2\omega + \frac{1}{2}\mathrm{d}[\omega,\omega]$. Se $\{\tau_1, \tau_2, ..., \tau_n\}$ é uma base de $\mathcal{G}$, escrevemos $\omega = \omega^i\tau_i, \Omega = \Omega^i\tau_i$ e por (1.11) e (1.12), temos

$$[\omega,\omega] \;=\; \sum_{i,j}(\omega^i \wedge \omega^j)[\tau_i,\tau_j] \;=\; \sum_{i,j,k} c_{ij}^k(\omega^i \wedge \omega^j)\tau_k.$$

Logo

$$\mathrm{d}\Omega^k \;=\; \frac{1}{2}c_{ij}^k\mathrm{d}(\omega^i \wedge \omega^j) \;=\; \frac{1}{2}c_{ij}^k(\mathrm{d}\omega^i \wedge \omega^j + \omega^i \wedge \mathrm{d}\omega^j)$$

e como os componentes $\omega^i$ de $\omega$ anulam-se em vetores horizontais,

$$D\Omega^k(X,Y,Z) \;=\; \mathrm{d}\Omega^k(X^H, Y^H, Z^H) \;=\; 0 \quad \forall X,Y,Z \in T_pP\ ,$$

ou seja, $D\Omega = 0$, que é conhecida como **identidade de Bianchi**.

### 1.4.4 Expressão Local da Curvatura

Vimos na seção 1.3.1 que o pull-back da 1-forma de conexão $\omega$ por uma seção local $\sigma_U$ dá origem a uma 1-forma de conexão $\omega_U$, definida num aberto $U$ da base $M$. Analogamente, o pull-back da 2-forma de curvatura $\Omega$ pela mesma seção local origina uma 2-forma $\Omega_U$ em $U$, tal que $\Omega_U = F_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$ e $F_{\mu\nu}$ é exatamente o tensor de curvatura definido em (1.9).

De fato, usando a equação estrutural de Cartan e observando que $\sigma_U^*\mathrm{d}\omega = \mathrm{d}\sigma_U^*\omega$ e $\sigma_U^*[\omega,\omega] = [\sigma_U^*\omega, \sigma_U^*\omega]$ , podemos escrever

$$\Omega_U = \sigma_U^*\Omega = \sigma_U^*(\mathrm{d}\omega + \frac{1}{2}[\omega,\omega]) = \mathrm{d}\sigma_U^*\omega + \frac{1}{2}[\sigma_U^*\omega, \sigma_U^*\omega] = \mathrm{d}\omega_U + \frac{1}{2}[\omega_U, \omega_U].$$

Se $\omega_U = A_\mu \mathrm{d}x^\mu$, analogamente à equação (1.11) podemos escrever $[\omega_U, \omega_U] = 2[A_\mu, A_\nu]\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$ e como $\mathrm{d}\omega_U = \partial_\mu A_\nu - \partial_\nu A_\mu$, obtemos

$$\Omega_U = (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = F_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \ ,$$

conforme (1.9).

Como essas expressões locais são dependentes da seção local utilizada, precisamos determinar como elas comportam-se quando passamos a outra seção local, $\sigma_V$, por exemplo. Suponhamos que no aberto $U \cap V \in M$ tenhamos $\sigma_V(x) = \sigma_U(x)g(x)$. De (1.2) temos a relação

$$\sigma_{V*}Y = [L_{g*}^{-1}g_*Y]^\sharp_{\sigma_V(x)} + R_{g*}\sigma_{U*}Y$$

e pela definição da 2-forma $\Omega$ sabemos que $\Omega(Z,W) = 0$, se $Z$ ou $W$ é vertical. Portanto, para quaisquer $X, Y \in T_xM$

$$\Omega_V(X,Y) = \Omega(\sigma_{V*}X, \sigma_{V*}Y) = \Omega(R_{g*}\sigma_{U*}X, R_{g*}\sigma_{U*}Y) = \mathrm{Ad}_{g^{-1}}\Omega_U(X,Y)$$

Neste ponto já reunimos condições para a demonstração de (1.8). Primeiro, vamos mostrar que se $X$ e $Y$ são vetores horizontais $\Omega(X,Y) = -\omega([X,Y])$. Em seguida, aplicando isso aos vetores do tipo $\sigma_{U*}\partial_\mu - A_\mu^\sharp$, que dão origem às derivadas covariantes, obtemos a referida equação.

Dados $X, Y \in H_p$, como $\omega(X) = \omega(Y) = 0$ a equação estrutural de Cartan nos fornece $\Omega(X,Y) = \mathrm{d}\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]) = -\omega([X,Y])$. Consequentemente $\Omega(\sigma_{U*}\partial_\mu - A^\sharp_\mu, \sigma_{U*}\partial_\nu - A^\sharp_\nu) = \Omega(\sigma_{U*}\partial_\mu, \sigma_{U*}\partial_\nu) - \Omega(\sigma_{U*}\partial_\mu, A^\sharp_\nu) - \Omega(A^\sharp_\mu, \sigma_{U*}\partial_\nu) + \Omega(A^\sharp_\mu, A^\sharp_\nu) = \Omega(\sigma_{U*}\partial_\mu, \sigma_{U*}\partial_\nu)$ , pois $\Omega(X,Y) = 0$ se qualquer dos vetores for vertical. Mas $\Omega(\sigma_{U*}\partial_\mu, \sigma_{U*}\partial_\nu) = \sigma^*_U\Omega(\partial_\mu, \partial_\nu) = \Omega_U(\partial_\mu, \partial_\nu)$. Considerando que $\omega_U = A_\mu \mathrm{d}x^\mu$ e então $\Omega_U = (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = F_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$, temos

$$\Omega(\sigma_{U*}\partial_\mu - A^\sharp_\mu, \sigma_{U*}\partial_\nu - A^\sharp_\nu) = F_{\theta\eta}\mathrm{d}x^\theta \wedge \mathrm{d}x^\eta(\partial_\mu, \partial_\nu) = F_{\mu\nu}.$$

Logo,

$$\omega([\sigma_{U*}\partial_\mu - A^\sharp_\mu, \sigma_{U*}\partial_\nu - A^\sharp_\nu]) = -F_{\mu\nu} \ .$$

Por outro lado, $\pi_*[\sigma_{U*}\partial_\mu - A^\sharp_\mu, \sigma_{U*}\partial_\nu - A^\sharp_\nu] = [\pi_*(\sigma_{U*}\partial_\mu - A^\sharp_\mu), \pi_*(\sigma_{U*}\partial_\nu - A^\sharp_\nu)] = [\partial_\mu, \partial_\nu] = 0$ implica que $[\sigma_{U*}\partial_\mu - A^\sharp_\mu, \sigma_{U*}\partial_\nu - A^\sharp_\nu]$ é um vetor vertical. Portanto,

$$[\sigma_{U*}\partial_\mu - A^\sharp_\mu, \sigma_{U*}\partial_\nu - A^\sharp_\nu] = (-F_{\mu\nu})^\sharp$$

e, finalmente, lembrando que $T_U$ é a trivialização local associada à seção $\sigma_U$, podemos escrever para as derivadas covariantes

$$\begin{aligned}
[\mathcal{D}_\mu, \mathcal{D}_\nu] \ &= \ [\partial_\mu - A_\mu, \partial_\mu - A_\mu] \ = \ [T_{U*}(\sigma_{U*}\partial_\mu - A^\sharp_\mu), T_{U*}(\sigma_{U*}\partial_\nu - A^\sharp_\nu)] \\
&= \ T_{U*}[\sigma_{U*}\partial_\mu - A^\sharp_\mu, \sigma_{U*}\partial_\nu - A^\sharp_\nu] \ = \ T_{U*}(-F^\sharp_{\mu\nu}) \\
&= \ -F_{\mu\nu} \ = \ -\partial_\mu A_\nu + \partial_\nu A_\mu - [A_\mu, A_\nu] \ ,
\end{aligned}$$

como queríamos demonstrar.

# Capítulo 2

# A Teoria de Yang-Mills Clássica

> A Matemática não só é real, mas é a única realidade. De fato, o universo todo é feito de matéria, obviamente. E a matéria é feita de partículas. É feita de eletrons, neutrons e protons. Entao, o universo todo efeito de partículas. Agora, do que as partículas são constituídas? Elas não são constituídas de coisa alguma. A única coisa que se pode dizer sobre a realidade de um elétron é citar suas propriedades matemáticas. Logo, existe um sentido em que a matéria se dissolve completamente e o que resta é apenas uma estrutura matemática.
>
> Martin Gardner

## 2.1 Introdução e Histórico

Procederemos aqui um breve resumo histórico que nos ajudará a compreender o interesse e a motivação que há por trás de certos aspectos das teorias de Yang-Mills (em particular, a busca por soluções clássicas e a escolha da métrica do espaço-tempo entre a euclidiana e a de Minkowski). Uma ótima referência para mais detalhes e fonte de outras referências é [1].

Em 1954, C. N. Yang e R. L. Mills propuseram a idéia de teorias de calibre não abelianas (com grupo estrutural não comutativo), ainda sem saber se havia alguma interação na natureza que pudesse ser descrita por uma teoria desse tipo. Naquele tempo, essas teorias foram bastante exploradas, ao nível quântico, por sua elegância.

Por volta da década de 60, depois de terem explorado as características e técnicas de quantização e renormalização das teorias de Yang-Mills, os físicos fizeram várias tentativas de usá-las para descrever fenômenos físicos. Por fim, S. Weinberg e A. Salam (Prêmio Nobel em 1968) conseguiram chegar a uma teoria de campo

com grupo estrutural $SU(2) \times U(1)$, englobando a descrição do eletromagnetismo e da interação nuclear fraca. Posteriormente, o sucesso da teoria que decompôs os hadrons em quarks levou à necessidade de se descrever as interações entre quarks. Como uma possibilidade nesse sentido surgiu a cromodinâmica quântica (QCD), envolvendo apenas quarks e um campo de calibre com simetria SU(3).

Modelos como o de Weinberg-Salam ou o QCD podem ser formulados em termos das chamadas integrais de Feynman, que são integrais funcionais sobre todas as configurações de campo possíveis, ao nível clássico. Assim, quanto mais informações tivermos sobre os campos clássicos, poderemos alcançar mais compreensão da teoria quântica. Essa é uma das principais motivações para se investigar as teorias de Yang-Mills clássicas, além da riqueza de sua estrutura matemática, é claro. Hoje em dia um entendimento mais aprofundado das teorias de calibre tornou-se praticamente indispensável para quem deseja lidar com as idéias mais recentes da física matemática.

No que diz respeito à teoria de Yang-Mills com grupo estrutural $SU(2)$, alguns desafios ainda não foram superados completamente. As equações de Euler-Lagrange que resultam do princípio variacional aplicado à teoria de Yang-Mills, são as chamadas equações de Yang-Mills. Elas são equações diferenciais parciais não lineares de segunda ordem e, portanto, são difíceis de se resolver. A primeira dessas soluções foi encontrada em 1962 e não despertou tanto interesse, pois era ligada à teoria abeliana do eletromagnetismo (com grupo estrutural $U(1)$), mergulhada na teoria de Yang-Mills. No final da década de 60 e início da de 70 foram encontradas genuinas soluções não abelianas, algumas associadas a monopolos magnéticos e a partículas elementares. Talvez a descoberta que despertou mais interesse nessa área tenha sido a das soluções chamadas de "instantons" por A. A. Belavin, A. M. Polyakov, A. S. Schwartz e Yu. S. Tyupkin em 1975.

Polyakov [15] observou que certos tipos de campos, os auto-duais, satisfazem automaticamente as equações de Euler-Lagrange. A vantagem é que para encontrar potenciais que levem a campos auto-duais, precisamos apenas resolver equações de primeira ordem, o que simplifica consideravelmente o nosso problema quando trabalhamos com variedades com métrica euclidiana. Quando passamos ao espaço de

Minkowski, infelizmente surge outra inconveniência: as equações de auto-dualidade, para campos no espaço de Minkowski possuem apenas soluções complexas que nos forçam a admtir um grupo de calibre não compacto. Mas em geral, os grupos de simetria dos problemas físicos envolvidos são compactos e considerar grupos não compactos leva a uma ação que não é finita.

Nosso principal objetivo neste capítulo é apresentar e formalizar matematicamente uma alternativa, proposta por B. M. Schechter em [18], que permite obter soluções das equações de Yang-Mills no espaço de Minkowski, sem requerer que elas sejam auto-duais. Mais precisamente, forçaremos a propriedade de invarância das soluções por transformações conformes para simplificar as equações de Yang- Mills. Essa invariância por transformações conformes ser convertida em uma invarincia por transformaes ortogonais ao projetarmos o espao de Minkowski sobre um hipertoro $S^3 \times S^1$, de maneira semelhante projeo estereogrfica de um espao euclidiano sobre uma esfera.

## 2.2 A Formulação da Teoria de Yang-Mills

Baseados na estrutura matemática dos fibrados principais, desenvolvida no Capítulo 1, construiremos aqui a formulação típica das teorias de calibre, que consiste em determinar as conexões (ou escolhas de calibre) que tornem mínimo (ou simplesmente estacionário) um funcional denominado "ação". As contas serão feitas com mais detalhes quando lidarmos com um caso específico na seção 2.3.2.

Daqui em diante, sempre que nos referirmos à teoria de Yang-Mills, estaremos considerando uma teoria de calibre com grupo de simetria $SU(2)$. Neste trabalho nos restringimos ao estudo na teoria de Yang-Mills pura (sem fontes como cargas ou monopolos, o campo não interage com partículas) e é neste contexto que falamos de dualidade e auto-dualidade. O desenvolvimento desta seção servirá para esclarecer o significado desses termos.

Consideremos a teoria de calibre com grupo $SU(2)$ e tendo como espaço base uma variedade 4-dimensional $M$. Para formular essa teoria no contexto de fibrados principais, conforme o que desenvolvemos no Capítulo 1, representamos uma conexão

localmente por

$$\omega_U = A = A^a_\mu(x)\tau_a \mathrm{d}x^\mu,$$

onde $\tau_a$, $a \in \{1,2,3\}$, é uma base da álgebra de Lie $su(2)$. A curvatura, é dada por

$$\Omega_U = F = F^a_{\mu\nu}(x)\tau_a \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = \mathrm{d}A + \frac{1}{2}[A, A]. \tag{2.1}$$

Aqui $F^a_{\mu\nu} = \frac{\partial A^a_\nu}{\partial x^\mu} - \frac{\partial A^a_\mu}{\partial x^\nu} + C^a_{bc}A^b_\mu A^c_\nu$ , sendo $C^a_{bc}$ as constantes de estrutura da álgebra de Lie, referentes à base $\{\tau_1, \tau_2, \tau_3\}$. Daí vemos claramente que a curvatura depende da conexão ou, interpretando fisicamente, a intensidade de campo (dada pelos $F_{\mu\nu}$) nos pontos da variedade depende da configuração do potencial (dada pelos $A_\mu(x)$).

A medida de curvatura estabelecida no Capítulo 1 é dada por uma 2-forma ou equivalentemente, um 2-tensor e nos permite medir o quanto transportes paralelos em duas direções dadas falham em comutar. Entretanto, se calcularmos uma espécie de "norma" do tensor de curvatura, teremos um valor numérico que permite avaliar, num certo sentido, a curvatura da conexão num dado ponto. Integrando essa norma da conexão sobre toda a variedade, teremos uma idéia da curvatura "acumulada" em toda a variedade. Essas idéias ganham mais significado se as interpretarmos fisicamente, embora essa interpretação não seja única e haja controvérsias. Da maneira como alguns interpretam, essa nossa "norma da curvatura" corresponde a uma espécie de energia associada à deformação que o potencial impõe à variedade, ao integrarmos essa energia sobre todo o espaço, obtemos a energia "armazenada" no campo.

No caso da teoria de Yang-Mills que estamos considerando, essa norma da curvatura é calculada por[1]

$$\|F\|^2 = \frac{1}{2}\mathrm{tr}(F_{\mu\nu}F^{\mu\nu})\nu = \mathrm{tr}(F \wedge^* F)$$

e dá origem à lagrangiana

$$\mathcal{L} = \frac{-1}{2}\mathrm{tr}(F_{\mu\nu}F^{\mu\nu}).$$

[1]Aqui $^*F$ é a 2-forma dual de $F$ pelo operador estrela de Hodge e $\nu$ é a forma volume da variedade $M$.

Integrando essa lagrangiana sobre $M$, temos a ação

$$S = \int_M \frac{-1}{2}\mathrm{tr}(F_{\mu\nu}F^{\mu\nu})\nu = -\int_M \mathrm{tr}(F \wedge^* F),$$

que por (2.1) depende da conexão $A$. O objetivo central nas Teorias de calibre é determinar conexões que sejam "pontos críticos" da ação, ou seja, fazendo uma variação na conexão $A$, procuramos aquelas que tornem estacionário o funcional $S = S(A)$. Para isso, consideramos uma variação de $A$ em uma direção arbitrária $A + tB$ e procuramos conexões que satisfaçam

$$\frac{\mathrm{d}}{\mathrm{dt}}S(A + tB)|_{t=0} = 0.$$

Desse modo obtemos as equações de Euler-Lagrange

$$D^*F = 0,$$

nesse caso conhecidas como equações de Yang-Mills, a serem resolvidas para a conexão $A$. Entretanto, as equações de Yang-Mills acima constituem um sistema de equações diferenciais parciais não lineares acopladas e com termos de ordem 2 e 3 em $A$, o que torna difícil encontrar soluções. Em [15], Polyakov sugeriu a busca de soluções em que o campo e seu dual fossem múltiplos um do outro, ou seja

$$^*F = \lambda F. \tag{2.2}$$

Com isso, como $F = DA$ implica em $DF = 0$ (identidade de Bianchi), as equações de Yang-Mills são automaticamente satisfeitas ($D^*F = D\lambda F = \lambda DF = 0$) e ficamos com a tarefa bem mais simples de encontrar soluções para (2.2). Para isso, podemos aplicar novamente o operador $^*$ obtendo

$$^{**}F = \lambda^2 F.$$

Num espaço com métrica euclidiana o operador $^*$ composto com ele mesmo é igual à identidade e, portanto $\lambda^2 = 1 \Rightarrow \lambda = \pm 1$, o que significa que devemos procurar por campos que satisfaçam $^*F = F$ (autoduais) ou $^*F = -F$ (anti-autoduais).

As soluções desse tipo, baseadas em $\mathbb{R}^4$ são os chamados instantons e já foram quase que exaustivamente explorados (vide [19]).

Entretanto, se quisermos utilizar esse artifício para resolver as equações de Yang-Mills com base no espaço de Minkowski, devemos levar em conta que o quadrado do operador $^*$ na métrica de Minkowski é $-1$ o que nos leva a $\lambda = \pm i$ implicando em $^*F = \pm iF$. Assim, o operador $^*$ leva matrizes hermitianas em anti-hermitianas e vice-versa (se considerarmos seu efeito sobre os coeficientes de $F$) tornando impossível qualquer solução (exceto com conexões planas, é claro) com $G = SU(2)$ que satisfaça (2.2), pois para todo $F_{\mu\nu} \in su(2)$ não nulo, temos que $iF_{\mu\nu} \notin su(2)$. Para contornar esse problema, alguns artifícios têm sido tentados, como complexificar o grupo ou, equivalentemente, formular a teoria utilizando um grupo cuja álgebra de Lie possa ser descrita sobre um corpo complexo. Mas em suma podemos afirmar que soluções da teoria de Yang-Mills (com grupo de simetria $SU(2)$) com base no espaço de Minkowski e satisfazendo (2.2) não existem.

Na busca por soluções, outras maneiras de simplificar as equações de Yang-Mills foram propostas. Exploraremos aqui uma que impõe a condição de que a solução seja invariante por transformações conformes (fato que havia sido verificado para as soluções do caso euclidiano), além da invariância natural por mudança de calibre (mudança de escolha da seção local do fibrado principal).

## 2.3 A Teoria de Yang-Mills no Hipertoro

É fato bem conhecido que ao fazermos a projeção estereográfica da esfera de Riemann sobre o plano complexo, cada rotação da esfera (rotação em $\mathbb{R}^3$) induz uma transformação de Möbius do plano complexo (isso pode ser visto em [16]). Na verdade existe um isomorfismo entre o grupo $SO(3,1)$ e o grupo das transformações conformes do plano complexo (que passaremos a chamar simplesmente de grupo conforme do plano complexo), que quando restrito ao subgrupo $SO(3) \subset SO(3,1)$ leva rotações da esfera em transformações de Möbius do plano (um subconjunto delas). De modo semelhante, existe um isomorfismo entre o grupo conforme do espaço de Minkowski e o grupo $SO(4,2)$ das pseudo-rotações num espaço 6-dimensional com

métrica diag$(1,1,1,-1,1,-1)$, conforme [12].

Em [10] foram estudadas as propriedades de invariância dos instantons de Yang-Mills por transformações conformes. Aquele trabalho mostra que essa solução das equações de Yang-Mills (o instanton) é invariante por um subgrupo isomorfo a $O(5)$ do grupo de transformações conformes do espaço euclidiano $\mathbb{R}^4$ (que é isomorfo a $SO(5,1)$). Para tornar essa invariância explícita, foi desenvolvido um formalismo (que tornou-se padrão na literatura especializada) que consiste de projetar o espaço $\mathbb{R}^4$ sobre a esfera $S^4 \subset \mathbb{R}^5$ via projeção estereográfica (inversa). Desse modo, as equações de Yang-Mills, invariantes por transformações conformes no $\mathbb{R}^4$, quando transportadas para a esfera $S^4$ passam a ser invariantes pelo grupo $O(5)$ de transformações ortogonais da esfera (rotações e reflexões).

No caso da teoria de Yang-Mills com base no espaço de Minkowski, De Alfaro, Fubini e Furlan encontraram uma solução interessante [6], que eles mostraram ser invariante sob o grupo $O(4) \times O(2)$ que é o maior subgrupo compacto de $SO(4,2)$, o grupo conforme do espaço de Minkowski. Esse fato sugere fortemente uma construção análoga à do caso euclidiano. Do mesmo modo que a esfera $S^4$ era a variedade natural a ser usada para formularmos equações invariantes por $O(5)$, a invariância por $O(4) \times O(2)$ deve tornar-se explícita num hipertoro $S^3 \times S^1$.

### 2.3.1 O Hipertoro e o Espaço de Minkowski

Já mencionamos que o grupo conforme do espaço de Minkowski é isomorfo a $SO(4,2)$, que pode ser visto como um grupo de transformações pseudo-ortogonais num espaço 6-dimensional com métrica diag$(1,1,1,-1,1,-1)$. Essas transformações podem ser vistas como agindo sobre vetores do cone de luz

$$\xi_1^2 + \xi_2^2 + \xi_3^2 - \xi_4^2 + \xi_5^2 - \xi_6^2 = 0,$$

que projetamos sobre o espaço de Minkowski fazendo[2]

$$x_\mu = \frac{\xi_\mu}{\xi_5 + \xi_6} \ . \tag{2.3}$$

Como queremos nos restringir a um subgrupo $O(4) \times O(2)$ de $SO(4,2)$, trabalharemos com um subconjunto do cone de luz acima tal que

$$\xi_1^{\ 2} + \xi_2^{\ 2} + \xi_3^{\ 2} + \xi_5^{\ 2} = 1 \tag{2.4}$$

$$\xi_4^{\ 2} + \xi_6^{\ 2} = 1. \tag{2.5}$$

Isso equivale a restringir o vetor $\xi$ à superfície de um hipertoro $(S^3 \times S^1)$ contido no cone de luz. Será útil então definir dois vetores $\mathbf{r} = (r_0, r_1, r_2, r_3)$ e $\mathbf{R} = (R_0, R_1)$ como sendo coordenadas dos pontos em $S^3 \subset I\!R^4$ e $S^1 \subset I\!R^2$, respectivamente. Esses vetores ficam definidos em termos das coordenadas $\xi_A$ por

$$r_0 = \xi_5, \quad r_1 = \xi_1, \quad r_2 = \xi_2, \quad r_3 = \xi_3, \quad R_0 = \xi_6, \quad R_1 = \xi_4.$$

Assim, podemos usar o produto escalar em $I\!R^4$ e $I\!R^2$ para reescrever as equações (2.4) e (2.5) como

$$\mathbf{r}.\mathbf{r} = 1 \quad \text{e} \quad \mathbf{R}.\mathbf{R} = 1. \tag{2.6}$$

Além disso, as coordenadas no espaço de Minkowski podem ser reescritas a partir de (2.3) como

$$x_i = \frac{r_i}{r_0 + R_0} \quad \text{e} \quad x_0 = \frac{R_1}{r_0 + R_0}, \tag{2.7}$$

com o que obtemos uma função do hipertoro sobre o espaço de Minkowski. é interessante observar que essa função leva os pontos $(\mathbf{r}, \mathbf{R})$ e $(-\mathbf{r}, -\mathbf{R})$ no mesmo ponto do espaço de Minkowski, o que já era de se esperar devido à característica de espaço projetivo imposta pelas equações (2.3). Além disso, para $r_0 + R_0 > 0$ podemos achar uma função inversa do espaço de Minkowski no hipertoro da seguinte forma:

[2]Adotamos aqui a seguinte convenção para os conjuntos de índices: índices representados por letras gregas do final do alfabeto ($\mu, \nu$, etc.) assumem valores no conjunto {0, 1, 2, 3}; letras maiúsculas do alfabeto latino ($A, B, C$, etc.) assumem valores em {1, 2, ..., 6}; letras minúsculas do alfabeto latino ($i, j, k$, etc.) variam em {1, 2, 3}.

primeiro elevamos cada uma das equações (2.7) ao quadrado e usamos (2.6) para obtermos

$$(r_0 + R_0)^2 x_0^2 = 1 - R_0^2 \quad \text{e} \quad (r_0 + R_0)^2 |\vec{x}|^2 = 1 - r_0^2, \tag{2.8}$$

onde $\vec{x} = (x_1, x_2, x_3)$. Subtraindo a segunda equação da primeira chegamos a

$$(r_0 + R_0)^2 (x_0^2 - |\vec{x}|^2) = r_0^2 - R_0^2 = (r_0 + R_0)(r_0 - R_0). \tag{2.9}$$

Aqui é conveniente introduzirmos as variáveis

$$t_+ = x_0 + |\vec{x}| \quad \text{e} \quad t_- = x_0 - |\vec{x}|,$$

que os físicos costumam chamar de "tempo avançado" e "tempo retardado". Como $t_+ t_- = x_0^2 - |\vec{x}|^2$, $t_+ + t_- = 2x_0$ e lembrando que estamos supondo $r_0 + R_0 > 0$, (2.9) leva-nos a

$$(r_0 + R_0) t_+ t_- = r_0 - R_0 \quad \Rightarrow \quad r_0 (t_+ t_- - 1) + R_0 (t_+ t_- + 1) = 0 \quad \Rightarrow$$

$$R_0 = r_0 \left( \frac{1 - t_+ t_-}{1 + t_+ t_-} \right). \tag{2.10}$$

Substituindo $R_0$ na primeira equação de (2.8) obtemos

$$r_0^2 \left( 1 + \frac{1 - t_+ t_-}{1 + t_+ t_-} \right)^2 x_0^2 = 1 - r_0^2 \left( \frac{1 - t_+ t_-}{1 + t_+ t_-} \right)^2 \quad \Rightarrow$$

$$r_0^2 \left( \frac{2}{1 + t_+ t_-} \right)^2 x_0^2 + r_0^2 \left( \frac{1 - t_+ t_-}{1 + t_+ t_-} \right)^2 = 1 \quad \Rightarrow$$

$$r_0^2 \left[ \frac{4x_0^2 + 1 - 2t_+ t_- + t_+^2 t_-^2}{(1 + t_+ t_-)^2} \right] = 1 \quad \Rightarrow$$

$$r_0^2 = \frac{(1 + t_+ t_-)^2}{(t_+ + t_-)^2 + 1 - 2t_+ t_- + t_+^2 t_-^2} = \frac{(1 + t_+ t_-)^2}{t_+^2 + t_-^2 + 2t_+ t_- + 1 - 2t_+ t_- + t_+^2 t_-^2} \quad \Rightarrow$$

$$r_0 = \frac{1 + t_+ t_-}{[(1 + t_+^2)(1 + t_-^2)]^{1/2}}. \tag{2.11}$$

Seguindo um raciocínio análogo, invertemos (2.10) para

$$r_0 = R_0 \left( \frac{1 + t_+ t_-}{1 - t_+ t_-} \right). \tag{2.12}$$

e substituimos em (2.8), obtendo no final das contas

$$R_0 = \frac{1 - t_+ t_-}{[(1+t_+^2)(1+t_-^2)]^{1/2}} \,. \tag{2.13}$$

Agora, com (2.12) e (2.13) calculamos $r_0 + R_0$ e substituimos em (2.7), donde

$$r_i = x_i(r_0 + R_0) \quad \text{e} \quad R_1 = x_0(r_0 + R_0) \quad \Rightarrow$$

$$r_i = \frac{2x_i}{[(1+t_+^2)(1+t_-^2)]^{1/2}} \quad \text{e} \quad R_1 = \frac{2x_0}{[(1+t_+^2)(1+t_-^2)]^{1/2}} \,. \tag{2.14}$$

E com isso completamos a fórmula para obter as coordenadas no hipertoro a partir das coordenadas no espaço de Minkowski, através de (2.11), (2.13) e (2.14).

Obtida a transformação acima, nosso próximo passo é obter uma transformação entre os espaços tangentes ao hipertoro e ao espaço de Minkowski.

Antes de mais nada, vamos alistar alguns objetos que serão úteis na formalização dos nossos procedimentos. Chamemos de $\mathcal{N}^6$ o espaço 6-dimensional dos vetores $\xi$, de $\mathcal{H}$ o hipertoro definido por (2.4), contido em $\mathcal{N}^6$. Denotemos por $\mathcal{K} \subset \mathcal{H}$ o subconjunto do hipertoro tal que $\xi_5 + \xi_6 > 0$, i. e., $r_0 + R_0 > 0$. Seja $\mathcal{M}$ o espaço de Minkowski e T: $\mathcal{K} \to \mathcal{M}$ a transformação definida por (2.7), ou seja

$$\mathrm{T}(R_0, R_1, r_0, r_1, r_2, r_3) = \frac{1}{r_0 + R_0}(R_1, r_1, r_2, r_3) = (x_0, x_1, x_2, x_3). \tag{2.15}$$

Na verdade essa transformação estende-se a todo o espaço $\mathcal{N}^6$ por (2.3) e no que se segue, usaremos o símbolo T para denotar essa transformação tanto globalmente quanto sua restrição ao espaço $\mathcal{K}$, o que ficará claro em cada contexto. Entretanto, denotaremos por dT a diferencial de T agindo no espaço tangente a $\mathcal{N}^6$ e por $\mathrm{T}_*$ a restrição de dT ao espaço tangente a $\mathcal{K}$. O motivo de destacarmos essa restrição de T à subvariedade $\mathcal{K}$ é a possibilidade de obtermos uma transformação inversa, $\mathrm{T}^{-1} : \mathcal{M} \to \mathcal{K}$, definida por (2.11), (2.13), e (2.14).

Para trabalharmos com os espaços tangentes, vamos escolher uma base conveniente para os vetores em $\mathcal{N}^6$. Assim, em termos das coordenadas $\xi_A$ sejam os vetores

$$\mathbf{e}_0 = (0,0,0,0,1,0), \quad \text{i.e.,} \quad \xi_5 = 1 \quad \text{e} \quad \xi_A = 0 \quad \text{se} \quad A \neq 5$$

$$\mathbf{e}_i = (0,...,1,...,0) \text{ com } \xi_i = 1 \text{ e } \xi_A = 0 \text{ se } A \neq i \quad (\xi_A = \delta_{iA})$$

$$\mathbf{E}_0 = (0,0,0,0,0,1) \Rightarrow \xi_A = \delta_{A6}$$

$$\mathbf{E}_1 = (0,0,0,1,0,0) \Rightarrow \xi_A = \delta_{A4}.$$

Eles formam uma base para $\mathcal{N}^6$, pois apenas renomeamos os vetores da base canônica. Nesta base, os vetores que dão as coordenadas no hipertoro podem ser escritos como

$$\mathbf{r} = r_\mu \mathbf{e}_\mu \quad \text{e} \quad \mathbf{R} = R_a \mathbf{E}_a. \tag{2.16}$$

Dado um ponto $P$ sobre o hipertoro $\mathcal{H}$, os vetores $\mathbf{e}_i$ ou $\mathbf{E}_a$ acima em geral não pertencem a $T_P\mathcal{H}$, entretanto se subtrairmos seus componentes normais ao hipertoro, obteremos vetores tangentes a $\mathcal{H}$. Os vetores normais a $\mathcal{H}$ num ponto $P$ de coordenadas (**r**,**R**) são gerados pelos vetores **r** e **R**, uma vez que estes são múltiplos dos vetores gradiente para (2.4) e (2.5). E mais ainda, **r** e **R** formam uma base ortonormal para o espaço normal ao hipertoro em cada ponto. Sendo assim, podemos construir vetores tangentes ao hipertoro num ponto $P$, como por exemplo

$$\hat{\mathbf{e}}_0 = \mathbf{e}_0 - \langle \mathbf{e}_0, \mathbf{r} \rangle \mathbf{r} - \langle \mathbf{e}_0, \mathbf{R} \rangle \mathbf{R}$$

e de (2.16) vem

$$\hat{\mathbf{e}}_0 = \mathbf{e}_0 - \langle \mathbf{e}_0, r_\mu \mathbf{e}_\mu \rangle (r_\nu \mathbf{e}_\nu) - \langle \mathbf{e}_0, R_a \mathbf{E}_a \rangle (R_b \mathbf{E}_b),$$

ou seja,

$$\hat{\mathbf{e}}_0 = \mathbf{e}_0 - r_0 (r_\nu \mathbf{e}_\nu).$$

Analogamente, para cada ponto $P \in \mathcal{H}$ calculamos a projeção, em $T_P\mathcal{H}$, dos vetores $\mathbf{e}_\mu$ e $\mathbf{E}_a$, da base de $\mathcal{N}^6$, obtendo os vetores

$$\hat{\mathbf{e}}_\mu = \mathbf{e}_\mu - r_\mu (r_\nu \mathbf{e}_\nu) \quad \text{e}$$

$$\hat{\mathbf{E}}_a = \mathbf{E}_a - R_a (R_b \mathbf{E}_b).$$

Esses vetores obviamente geram o espaço $T_P\mathcal{H}$ e nos permitem definir derivações em direções tangentes ao hipertoro por

$$\hat{\partial}_\mu = \frac{\partial}{\partial r_\mu} - r_\mu \left( r_\nu \frac{\partial}{\partial r_\nu} \right) \quad \text{e}$$

$$\hat{\Delta}_a = \frac{\partial}{\partial R_a} - R_a \left( R_b \frac{\partial}{\partial R_b} \right). \tag{2.17}$$

Consideremos agora uma função diferenciável $g : \mathcal{M} \to \mathbb{R}$ e a transformação T: $\mathcal{N}^6 \to \mathcal{M}$ definida por (2.15). Pela regra da cadeia, podemos escrever

$$\frac{\partial (g \circ \mathrm{T})}{\partial r_\mu} = \frac{\partial g}{\partial x_\eta} \frac{\partial x_\eta}{\partial r_\mu},$$

de onde extraimos a seguinte relação entre as derivações em $\mathcal{N}^6$ e $\mathcal{M}$:

$$\mathrm{dT} \left( \frac{\partial}{\partial r_\mu} \right) = \frac{\partial x_\eta}{\partial r_\mu} \frac{\partial}{\partial x_\eta}. \tag{2.18}$$

Outra maneira de enxergar esse fato é através da matriz da transformação dT, derivada de (2.15) e dada por $t_{\eta\mu} = \frac{\partial x_\eta}{\partial r_\mu}$. Se $\{\bar{\mathbf{e}}_0, \bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \bar{\mathbf{e}}_3\}$ é a base canônica do espaço de Minkowski, $\mathcal{M}$, temos

$$\mathrm{dT}(\mathbf{e}_\mu) = \frac{\partial x_\eta}{\partial r_\mu} \bar{\mathbf{e}}_\eta,$$

que nos leva à mesma relação anterior para as derivadas parciais (é claro que vale o mesmo para os vetores $\mathbf{E}_a$ e as derivadas $\frac{\partial}{\partial R_a}$). Por isso, a partir de (2.15) calculamos

$$\frac{\partial x_0}{\partial r_0} = \frac{\partial x_0}{\partial R_0} = \frac{-R_1}{(r_0 + R_0)^2} \ ,$$

$$\frac{\partial x_i}{\partial r_0} = \frac{\partial x_i}{\partial R_0} = \frac{-r_i}{(r_0 + R_0)^2} \ ,$$

$$\frac{\partial x_0}{\partial r_i} = \frac{\partial x_i}{\partial R_1} = 0 \ ,$$

$$\frac{\partial x_0}{\partial R_1} = \frac{1}{r_0 + R_0} \ , \quad \frac{\partial x_j}{\partial r_i} = \frac{\delta_{ij}}{r_0 + R_0}$$

e então, de (2.18) temos, por exemplo

$$\mathrm{dT} \left( \frac{\partial}{\partial r_0} \right) = \frac{\partial x_\eta}{\partial r_0} \frac{\partial}{\partial x_\eta} = \frac{\partial x_0}{\partial r_0} \frac{\partial}{\partial x_0} + \frac{\partial x_i}{\partial r_0} \frac{\partial}{\partial x_i}$$

$$= \frac{-R_1}{(r_0 + R_0)^2} \frac{\partial}{\partial x_0} + \frac{-r_i}{(r_0 + R_0)^2} \frac{\partial}{\partial x_i} = \frac{-1}{r_0 + R_0} \left( x_0 \frac{\partial}{\partial x_0} + x_i \frac{\partial}{\partial x_i} \right) .$$

Do mesmo modo, podemos calcular a imagem por dT das outras derivações em $\mathcal{N}^6$ e aqui podemos abreviar a notação, fazendo $\nabla = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3})$, para obtermos

$$\mathrm{dT}\left(\frac{\partial}{\partial r_0}\right) = \mathrm{dT}\left(\frac{\partial}{\partial R_0}\right) = \frac{-1}{r_0 + R_0}\left(x_0 \frac{\partial}{\partial x_0} + \vec{x}\cdot\nabla\right),$$

$$\mathrm{dT}\left(\frac{\partial}{\partial r_i}\right) = \left(\frac{1}{r_0 + R_0}\right)\frac{\partial}{\partial x_i} \quad \text{e} \quad \mathrm{dT}\left(\frac{\partial}{\partial R_1}\right) = \frac{1}{r_0 + R_0}\frac{\partial}{\partial x_0}.$$

Agora, de (2.17) escrevemos

$$\mathrm{dT}(\hat{\partial}_\mu) = \mathrm{dT}\left(\frac{\partial}{\partial r_\mu}\right) - r_\mu r_\nu \mathrm{dT}\left(\frac{\partial}{\partial r_\nu}\right) \quad \text{e}$$

$$\mathrm{dT}(\hat{\Delta}_a) = \mathrm{dT}\left(\frac{\partial}{\partial R_a}\right) - R_a R_b \mathrm{dT}\left(\frac{\partial}{\partial R_b}\right),$$

De onde calculamos, por exemplo

$$\mathrm{dT}(\hat{\partial}_0) = \mathrm{dT}\left(\frac{\partial}{\partial r_0}\right) - r_0^2 \mathrm{dT}\left(\frac{\partial}{\partial r_0}\right) - r_0 r_i \mathrm{dT}\left(\frac{\partial}{\partial r_i}\right) =$$

$$- \frac{(1 - r_0^2)}{r_0 + R_0} x_0 \frac{\partial}{\partial x_0} - \frac{(1 - r_0^2)}{r_0 + R_0}\vec{x}\cdot\nabla - r_0 \vec{x}\cdot\nabla,$$

logo

$$\mathrm{dT}(\hat{\partial}_0) = \frac{-1}{r_0 + R_0}\left[(1 - r_0^2) x_0 \frac{\partial}{\partial x_0} + (1 + r_0 R_0)\vec{x}\cdot\nabla\right]. \tag{2.19}$$

Analogamente obtemos

$$\mathrm{dT}(\hat{\partial}_i) = \frac{1}{r_0 + R_0}\left[r_i\left(r_0 x_0 \frac{\partial}{\partial x_0} - R_0 \vec{x}\cdot\nabla\right) + \frac{\partial}{\partial x_i}\right], \tag{2.20}$$

$$\mathrm{dT}(\hat{\Delta}_0) = \frac{-1}{r_0 + R_0}\left[(1 + r_0 R_0) x_0 \frac{\partial}{\partial x_0} + (1 - R_0^2)\vec{x}\cdot\nabla\right] \quad \text{e} \tag{2.21}$$

$$\mathrm{dT}(\hat{\Delta}_1) = \frac{1}{r_0 + R_0}\left[R_1\left(r_0 x_0 \frac{\partial}{\partial x_0} - R_0 \vec{x}\cdot\nabla\right) + \frac{\partial}{\partial x_0}\right]. \tag{2.22}$$

Com isso podemos transportar qualquer vetor tangente a $\mathcal{H}$ para o espaço tangente a $\mathcal{M}$. Para fazer o transporte de $T_x\mathcal{M}$ para $T_p\mathcal{H}$, onde $p$=T$(x)$, basta invertermos as transformações acima, isolando cada derivada parcial de $\mathcal{M}$ e aplicando $T_*^{-1}$, o que resulta

$$T_*^{-1}\left(\frac{\partial}{\partial x_i}\right) = (r_0 + R_0)\hat{\partial}_i + r_i(\hat{\Delta}_0 - \hat{\partial}_0) \text{ e} \tag{2.23}$$

$$T_*^{-1}\left(\frac{\partial}{\partial x_0}\right) = (r_0 + R_0)\hat{\Delta}_1 + R_1(\hat{\Delta}_0 - \hat{\partial}_0)\,. \tag{2.24}$$

### 2.3.2 As Equações de Yang-Mills no Hipertoro

Desenvolvidas as ferramentas, voltamos nosso interesse agora para a formulação da teoria de Yang-Mills sobre o hipertoro ($\mathcal{H}$), de tal maneira que as soluções obtidas possam ser transformadas em soluções no espaço de Minkowski ($\mathcal{M}$), levando em conta a relação entre as coordenadas de $\mathcal{H}$ e de $\mathcal{M}$.

Como já vimos, resolver a teoria de Yang-Mills significa encontrar, a partir de um método variacional, conexões que tornem mínimo um funcional "ação", que pode ser visto como uma espécie de energia armazenada no campo definido pela conexão.

Essas conexões podem ser expressas localmente em termos dos coeficientes $A_\mu$ da 1-forma local de conexão, que permitem expressar as derivadas covariantes como $\partial_\mu - A_\mu$ (sec. 1.4.2). Quando transferidas ao hipertoro, onde decompomos os vetores tangentes em termos dos $\hat{\partial}_\mu$ e $\hat{\Delta}_a$, essas derivadas covariantes devem transformar-se em $\hat{\partial}_\mu - \hat{a}_\mu$ e $\hat{\Delta}_a - \hat{A}_a$, onde $\hat{a}_\mu$ e $\hat{A}_a$ são funções diferenciáveis de $\mathcal{H}$ em $su(2)$. Estendendo a transformação $T_*$ que aparece em (2.19) a (2.22), podemos definir um isomorfismo $\bar{T}_* : T_p\mathcal{H} \oplus su(2) \to T_x\mathcal{M} \oplus su(2)$, de forma que, por (2.23) e (2.24), obtenhamos

$$\bar{T}_*^{-1}\left(\frac{\partial}{\partial x_i} - A_i\right) = (r_0 + R_0)(\hat{\partial}_i - \hat{a}_i) + r_i[(\hat{\Delta}_0 - \hat{A}_0) - (\hat{\partial}_0 - \hat{a}_0)] \text{ e}$$

$$\bar{T}_*^{-1}\left(\frac{\partial}{\partial x_0} - A_0\right) = (r_0 + R_0)(\hat{\Delta}_1 - \hat{A}_1) + R_1[(\hat{\Delta}_0 - \hat{A}_0) - (\hat{\partial}_0 - \hat{a}_0)]\,.$$

E pela linearidade de $\bar{T}_*$, chegamos à relação entre as conexões nos dois espaços:

$$\bar{T}_*^{-1}(A_i) = (r_0 + R_0)\hat{a}_i + r_i(\hat{A}_0 - \hat{a}_0) \text{ e} \tag{2.25}$$

$$\bar{T}_*^{-1}(A_0) = (r_0 + R_0)\hat{A}_1 + R_1(\hat{A}_0 - \hat{a}_0)\,. \tag{2.26}$$

Precisamos agora transportar para o hipertoro a noção de curvatura da conexão. Já vimos na seção 1.4.2 que essa curvatura pode ser medida pelos comutadores das derivadas covariantes, com os quais construimos um tensor $F_{\mu\nu}$, e definimos o que era uma conexão plana (a que tinha curvatura nula). No caso do hipertoro, podemos proceder de modo análogo, mas devemos o cuidado de observar que para uma conexão nula o comutador de duas derivadas covariantes fica reduzido ao comutador de duas derivadas parciais o que nem sempre é nulo. Mais precisamente, a partir das equações (2.17) temos

$$[\hat{\Delta}_a, \hat{\Delta}_b] = 0 \ ,$$

$$[\hat{\Delta}_a, \hat{\partial}_\mu] = 0 \quad \text{e}$$

$$[\hat{\partial}_\mu, \hat{\partial}_\nu] = r_\mu \hat{\partial}_\nu - r_\nu \hat{\partial}_\mu \ . \tag{2.27}$$

Portanto, para definirmos um tensor de curvatura das conexões sobre o hipertoro que seja nulo para conexões planas, consideraremos os seguintes componentes:

$$\hat{E}_{a\mu} = -[\hat{\Delta}_a - \hat{A}_a, \hat{\partial}_\mu - \hat{a}_\mu] = \hat{\Delta}_a \hat{a}_\mu - \hat{\partial}_\mu \hat{A}_a + [\hat{A}_a, \hat{a}_\mu] \ , \tag{2.28}$$

$$\begin{aligned} \hat{H}_{\mu\nu} &= -\{[\hat{\partial}_\mu - \hat{a}_\mu, \hat{\partial}_\nu - \hat{a}_\nu] - r_\mu(\hat{\partial}_\nu - \hat{a}_\nu) + r_\nu(\hat{\partial}_\mu - \hat{a}_\mu)\} \\ &= \hat{\partial}_\mu \hat{a}_\nu - \hat{\partial}_\nu \hat{a}_\mu - r_\mu \hat{a}_\nu + r_\nu \hat{a}_\mu + [\hat{a}_\mu, \hat{a}_\nu] \ . \end{aligned} \tag{2.29}$$

Com isso, podemos construir um tensor de curvatura da conexão para o hipertoro, fazendo

$$\hat{F} = \begin{bmatrix} 0 & 0 & \hat{E}_{00} & \hat{E}_{01} & \hat{E}_{02} & \hat{E}_{03} \\ 0 & 0 & \hat{E}_{10} & \hat{E}_{11} & \hat{E}_{12} & \hat{E}_{13} \\ -\hat{E}_{00} & -\hat{E}_{10} & 0 & \hat{H}_{01} & \hat{H}_{02} & \hat{H}_{03} \\ -\hat{E}_{01} & -\hat{E}_{11} & \hat{H}_{10} & 0 & \hat{H}_{12} & \hat{H}_{13} \\ -\hat{E}_{02} & -\hat{E}_{12} & \hat{H}_{20} & \hat{H}_{21} & 0 & \hat{H}_{23} \\ -\hat{E}_{03} & -\hat{E}_{13} & \hat{H}_{30} & \hat{H}_{31} & \hat{H}_{32} & 0 \end{bmatrix}$$

Agora podemos construir a lagrangiana da teoria de Yang-Mills do modo usual, ou seja, tomando a norma de $\hat{F}$. Aqui é importante lembrar que da forma como estão organizadas as entradas na matriz acima, estamos supondo que as coordenadas no hipertoro estejam ordenadas como $(R_0, R_1, r_0, r_1, r_2, r_3)$, correspondendo respectivamente às coordenadas $(\xi_6, \xi_4, \xi_5, \xi_1, \xi_2, \xi_3)$ de $\mathcal{N}^6$. Por causa desta reordenação

das coordenadas, a métrica fica dada pela matriz $\eta = \mathrm{diag}(-1,-1,1,1,1,1)$. Além disso, cada entrada do tensor $\hat{F}$ é um elemento de $su(2)$, portanto a norma, que nos servirá de lagrangiana, é dada por

$$\|\hat{F}\| \;=\; \mathrm{Tr}(\hat{H}_{\mu\nu}\hat{H}_{\mu\nu} - 2\hat{E}_{a\mu}\hat{E}_{a\mu})\;.$$

Integrando a lagrangiana sobre o hipertoro, obtemos a ação

$$I = \int_{\mathcal{H}} \mathrm{Tr}(\hat{H}_{\mu\nu}\hat{H}_{\mu\nu} - 2\hat{E}_{a\mu}\hat{E}_{a\mu})\mathrm{d}S\;. \tag{2.30}$$

Variando este funcional em relação a $\hat{a}_\mu$ e $\hat{A}_a$ e procurando por pontos estacionários, obtemos as equações de Euler-Lagrange, que no nosso caso são as equações de Yang-Mills no hipertoro.

Mais precisamente, vamos considerar uma variação da conexão em uma direção arbitrária substituindo, na lagrangiana acima, as conexões por $\hat{a}_\mu + tb_\mu$ e $\hat{A}_a + tB_a$. Com isso, a partir de (2.28) e (2.29) temos os campos sujeitos à variação

$$\begin{aligned}
\tilde{H}_{\mu\nu} \;&=\; \hat{\partial}_\mu(\hat{a}_\nu + tb_\nu) - \hat{\partial}_\nu(\hat{a}_\mu + tb_\mu) - r_\mu(\hat{a}_\nu + tb_\nu) + r_\nu(\hat{a}_\mu + tb_\mu) + [\hat{a}_\mu + tb_\mu, \hat{a}_\nu + tb_\nu] \\
&=\; \hat{H}_{\mu\nu} \;+\; t(\hat{\partial}_\mu b_\nu - \hat{\partial}_\nu b_\mu - r_\mu b_\nu + r_\nu b_\mu + [\hat{a}_\mu, b_\nu] + [b_\mu, \hat{a}_\nu]) + t^2[b_\mu, b_\nu] \quad \text{e} \\
\tilde{E}_{a\mu} \;&=\; \hat{\Delta}_a(\hat{a}_\mu + tb_\mu) - \hat{\partial}_\mu(\hat{A}_a + tB_a) + [\hat{A}_a + tB_a, \hat{a}_\mu + tb_\mu] \\
&=\; \hat{E}_{a\mu} + t(\hat{\Delta}_a b_\mu - \hat{\partial}_\mu B_a + [\hat{A}_a, b_\mu] + [B_a, \hat{a}_\mu]) + t^2[B_a, b_\mu]\;.
\end{aligned}$$

Se $\tilde{S}$ é a ação sujeita à variação acima, teremos

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\tilde{S}|_{t=0} \;&=\; \frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{\mathcal{H}} \mathrm{Tr}(\tilde{H}_{\mu\nu}\tilde{H}_{\mu\nu} - 2\tilde{E}_{a\mu}\tilde{E}_{a\mu})\mathrm{d}S\right]_{t=0} \\
&=\; \int_{\mathcal{H}} \mathrm{Tr}\{2\hat{H}_{\mu\nu}(\hat{\partial}_\mu b_\nu - \hat{\partial}_\nu b_\mu - r_\mu b_\nu + r_\nu b_\mu + [\hat{a}_\mu, b_\nu] + [b_\mu, \hat{a}_\nu]) \\
&\qquad - 4\hat{E}_{a\mu}(\hat{\Delta}_a b_\mu - \hat{\partial}_\mu B_a + [\hat{A}_a, b_\mu] + [B_a, \hat{a}_\mu])\}\mathrm{d}S\;,
\end{aligned}$$

onde utilizamos o fato de que $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Agora, integrando por partes os termos do tipo $\hat{H}_{\mu\nu}\hat{\partial}_\mu b_\nu$ e lembrando que $\mathcal{H}$ é uma variedade compacta e sem bordo, transferimos as derivações para os campos $\hat{H}_{\mu\nu}$ e $\hat{E}_{a\mu}$ obtendo

$$\int_{\mathcal{H}} \mathrm{Tr}\{2(-\hat{\partial}_\mu\hat{H}_{\mu\nu}b_\nu + \hat{\partial}_\nu\hat{H}_{\mu\nu}b_\mu - r_\mu\hat{H}_{\mu\nu}b_\nu + r_\nu\hat{H}_{\mu\nu}b_\mu + \hat{H}_{\mu\nu}[\hat{a}_\mu, b_\nu] + \hat{H}_{\mu\nu}[b_\mu, \hat{a}_\nu])$$

$$- 4(-\hat{\Delta}_a \hat{E}_{a\mu} b_\mu + \hat{\partial}_\mu \hat{E}_{a\mu} B_a + \hat{E}_{a\mu}[\hat{A}_a, b_\mu] + \hat{E}_{a\mu}[B_a, \hat{a}_\mu])\}\mathrm{d}S.$$

Aqui há termos envolvendo $b_\nu$ e outros com $b_\mu$, por exemplo. Mas como os índices $\mu$ e $\nu$ são mudos na soma que estamos fazendo, podemos intercambiá-los, observando antissimetria dos tensores envolvidos ($\hat{H}_{\nu\mu} = -\hat{H}_{\mu\nu}$). Além disso, a propriedade já mencionada do traço transforma-se em uma propriedade cíclica quando temos o produto de mais que duas matrizes (Tr$(ABC)$ =Tr$(CAB)$). Tudo isso nos permite escrever a integral acima como

$$\int_{\mathcal{H}} \mathrm{Tr}\{2(-2\hat{\partial}_\mu \hat{H}_{\mu\nu} - 2r_\mu \hat{H}_{\mu\nu} - 2[\hat{a}_\mu, \hat{H}_{\mu\nu}])b_\nu \; - \; 4(-\hat{\Delta}_a \hat{E}_{a\nu} - [\hat{A}_a, \hat{E}_{a\nu}])b_\nu$$

$$- 4(\hat{\partial}_\mu \hat{E}_{a\mu} + [\hat{a}_\mu, \hat{E}_{a\mu}])B_a\}\mathrm{d}S \; .$$

Como as variações $b_\nu$ e $B_a$ são arbitrárias, a única maneira de a integral acima ser nula independentemente dessas variações é se

$$\hat{\partial}_\mu \hat{H}_{\mu\nu} + [\hat{a}_\mu, \hat{H}_{\mu\nu}] - (\hat{\Delta}_a \hat{E}_{a\nu} + [\hat{A}_a, \hat{E}_{a\nu}]) \; = \; 0 \quad \text{e} \tag{2.31}$$

$$\hat{\partial}_\mu \hat{E}_{a\mu} + [\hat{a}_\mu, \hat{E}_{a\mu}] \; = \; 0 \tag{2.32}$$

e estas são as esperadas equações de Yang-Mills no hipertoro.

### 2.3.3 As Soluções

Lembrando agora nossa discussão inicial sobre a correspondência entre transformações conformes no espaço de Minkowski e rotações no hipertoro, vamos procurar soluções para as equações (2.31) e (2.32) que sejam invariantes por $O(4)$, o que implicará em invariância por certas transformações conformes quando as transportarmos para o espaço de Minkowski.

Utilizaremos como base da álgebra de Lie $su(2)$, as matrizes anti-hermitianas

$$\tau_1 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \; , \quad \tau_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{e} \quad \tau_3 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \; ,$$

que satisfazem às relações

$$\tau_i \tau_j = \epsilon_{ijk} \tau_k = -\tau_j \tau_i \; , \quad \tau_i^2 = -I \quad \text{e} \quad \tau_i^* = -\tau_i = \tau_i^{-1} \; ,$$

onde $\tau_i^*$ denota o conjugado hermitiano de $\tau_i$. Acrescentando a estas a matriz identidade $\tau_0 = I$, podemos descrever os elementos do grupo $SU(2)$ por

$$g(x) \;=\; x_\mu \tau_\mu \;=\; \begin{bmatrix} x_0 + ix_3 & -x_2 + ix_1 \\ x_2 + ix_1 & x_0 - ix_3 \end{bmatrix} ,$$

para qualquer vetor $x \in S^3$. Para verificar isso, basta observar que $\det g(x) = \|x\| = 1$ e que transpondo e conjugando $g(x)$, obtemos $h(x) = x_0\tau_0 - x_k\tau_k$ que multiplicado por $g(x)$ resulta na matriz identidade. Ora, os elementos de $SU(2)$ são caracterizados justamente por ter o inverso igual ao conjugado hermitiano e terem determinante igual a 1.

Uma vez que as coordenadas de qualquer ponto do hipertoro satisfazem (2.6), podemos fazer $R_0 = \cos\theta$ e $R_1 = \sin\theta$. Baseado na forma das soluções discutidas em [10], B. M. Schechter (ver [18]) sugere a procura por soluções da forma[3]

$$\hat{a}_\mu \;=\; f(\theta) g(\mathbf{r})^{-1} \hat{\partial}_\mu g(\mathbf{r}) \quad \text{e} \quad \hat{A}_a \;=\; f(\theta) g(\mathbf{r})^{-1} \hat{\Delta}_a g(\mathbf{r}) \;=\; 0 \,, \tag{2.33}$$

onde $g(\mathbf{r}) = r_\mu \tau_\mu$ é o difeomorfismo natural de $S^3$ em $SU(3)$. De fato, isso nos leva a uma grande simplificação das equações de Yang-Mills sobre o hipertoro, como veremos a seguir. Chamemos de $\phi_\mu$ às funções $g(\mathbf{r})^{-1}\hat{\partial}_\mu g(\mathbf{r})$ do hipertoro na álgebra de Lie $su(2)$. Alternativamente, podemos escrever

$$\begin{aligned} \phi_\mu \;&=\; g^{-1}\hat{\partial}_\mu g \;=\; g^{-1}(\partial_\mu g - r_\mu r_\eta \partial_\eta g) \\ &=\; g^{-1}(\tau_\mu - r_\mu r_\eta \tau_\eta) \;=\; g^{-1}(\tau_\mu - r_\mu g) \end{aligned} \tag{2.34}$$

e então $\hat{a}_\mu = f(\theta)\phi_\mu = f(\theta)(g^{-1}\tau_\mu - r_\mu\tau_0)$. Substituindo em (2.28) obtemos

$$\hat{E}_{a\mu} = \hat{\Delta}_a \hat{a}_\mu = \hat{\Delta}_a f(\theta)\phi_\mu \,. \tag{2.35}$$

Para (2.29), temos

$$\hat{H}_{\mu\nu} \;=\; f(\theta)\left[\hat{\partial}_\mu\left(g^{-1}\hat{\partial}_\nu g\right) - \hat{\partial}_\nu\left(g^{-1}\hat{\partial}_\mu g\right) - r_\mu g^{-1}\hat{\partial}_\nu g + r_\nu g^{-1}\hat{\partial}_\mu g\right]$$

[3] O *ansatz* proposto em [18] parece diferente do aqui apresentado, pois utiliza matrizes hermitianas (as matrizes de Pauli) e é formulado de uma maneira diferente, entretanto é fácil verificar que as duas formulações são totalmente equivalentes, a menos de um fator $-1/2$.

$$+f^2(\theta)\left[g^{-1}\hat{\partial}_\mu g, g^{-1}\hat{\partial}_\nu g\right]$$

$$= f(\theta)\left(\hat{\partial}_\mu g^{-1}\hat{\partial}_\nu g + g^{-1}\hat{\partial}_\mu\hat{\partial}_\nu g - \hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g - g^{-1}\hat{\partial}_\nu\hat{\partial}_\mu g - r_\mu g^{-1}\hat{\partial}_\nu g + r_\nu g^{-1}\hat{\partial}_\mu g\right)$$

$$+f^2(\theta)\left(g^{-1}\hat{\partial}_\mu g g^{-1}\hat{\partial}_\nu g - g^{-1}\hat{\partial}_\nu g g^{-1}\hat{\partial}_\mu g\right) .$$

Lembrando que as derivações $\hat{\partial}_\mu$ são em direções tangentes ao hipertoro e que sobre o hipertoro temos $g^{-1}g = I$, podemos derivar esse produto obtendo $\hat{\partial}_\mu(g^{-1}g) = \hat{\partial}_\mu g^{-1} g + g^{-1}\hat{\partial}_\mu g = 0$. Logo $g^{-1}\hat{\partial}_\mu g = -\hat{\partial}_\mu g^{-1} g$. Além disso, podemos utilizar (2.27) para escrever

$$g^{-1}\hat{\partial}_\mu\hat{\partial}_\nu g - g^{-1}\hat{\partial}_\nu\hat{\partial}_\mu g = g^{-1}[\hat{\partial}_\mu, \hat{\partial}_\nu]g$$

$$= g^{-1}\left(r_\mu\hat{\partial}_\nu - r_\nu\hat{\partial}_\mu\right)g = r_\mu g^{-1}\hat{\partial}_\nu g - r_\nu g^{-1}\hat{\partial}_\mu g .$$

Portanto, voltando à expressão para $\hat{H}_{\mu\nu}$, temos

$$\hat{H}_{\mu\nu} = \left(f - f^2\right)\left(\hat{\partial}_\mu g^{-1}\hat{\partial}_\nu g - \hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g\right) . \tag{2.36}$$

Obtidas as expressões para $\hat{E}_{a\mu}$ e $\hat{H}_{\mu\nu}$, podemos levá-las às equações de Yang-Mills (2.31) e (2.32). Para a primeira parcela em (2.31) temos

$$\hat{\partial}_\mu\hat{H}_{\mu\nu} = \left(f - f^2\right)\left[\hat{\partial}_\mu\left(\hat{\partial}_\mu g^{-1}\hat{\partial}_\nu g\right) - \hat{\partial}_\mu\left(\hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g\right)\right]$$

$$= \left(f - f^2\right)\left(\hat{\partial}_\mu\hat{\partial}_\mu g^{-1}\hat{\partial}_\nu g + \hat{\partial}_\mu g^{-1}\hat{\partial}_\mu\hat{\partial}_\nu g - \hat{\partial}_\mu\hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g - \hat{\partial}_\nu g^{-1}\hat{\partial}_\mu\hat{\partial}_\mu g\right) .$$

Então calculamos

$$\hat{\partial}_\mu\hat{\partial}_\nu g = \hat{\partial}_\mu\left(\tau_\nu - r_\nu g\right) = -\partial_\mu\left(r_\nu g\right) + r_\mu r_\eta\partial_\eta\left(r_\nu g\right)$$

$$= -\delta_{\mu\nu}g - r_\nu\tau_\mu + 2r_\mu r_\nu g , \tag{2.37}$$

onde $-\delta_{\mu\nu}$ é o delta de Kronecker e de modo análogo, podemos obter

$$\hat{\partial}_\mu\hat{\partial}_\nu g^{-1} = -\delta_{\mu\nu}g^{-1} - r_\nu\partial_\mu g^{-1} + 2r_\mu r_\nu g^{-1} . \tag{2.38}$$

A partir daí, e lembrando que a repetição do índice $\mu$ num mesmo termo implica numa soma com $\mu$ indo de 0 a 3, podemos obter

$$\hat{\partial}_\mu\hat{\partial}_\mu g = -3g \quad \text{e} \quad \hat{\partial}_\mu\hat{\partial}_\mu g^{-1} = -3g^{-1} . \tag{2.39}$$

Além disso, temos

$$\hat{\partial}_\mu g^{-1}\hat{\partial}_\mu\hat{\partial}_\nu g = \left(\partial_\mu g^{-1} - r_\mu g^{-1}\right)\left(-\delta_{\mu\nu}g - r_\nu\tau_\mu + 2r_\mu r_\nu g\right)$$

$$= -\partial_\nu g^{-1} g - 4r_\nu\tau_0 + 2r_\nu r_\mu\partial_\mu g^{-1} g + r_\nu\tau_0 + r_\mu g^{-1} r_\nu\tau_\mu - 2r_\mu r_\mu r_\nu\tau_0$$

$$= -\partial_\nu g^{-1} g - 2r_\nu\tau_0 = -\hat{\partial}_\nu g^{-1} g - 3r_\nu\tau_0 \qquad (2.40)$$

e, analogamente,

$$\hat{\partial}_\mu\hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g = -\left(\delta_{\mu\nu}g^{-1} - r_\nu\partial_\mu g^{-1} + 2r_\mu r_\nu g^{-1}\right)\left(\tau_\mu - r_\mu g\right)$$

$$= -g^{-1}\partial_\nu g - 2r_\nu\tau_0 = -g^{-1}\hat{\partial}_\nu g - 3r_\nu\tau_0\ . \qquad (2.41)$$

Com isso obtemos

$$\hat{\partial}_\mu\hat{H}_{\mu\nu} = \left(f - f^2\right)\left(-3g^{-1}\hat{\partial}_\nu g - \hat{\partial}_\nu g^{-1} g - 3r_\nu\tau_0 + g^{-1}\hat{\partial}_\nu g + 3r_\nu\tau_0 + 3\hat{\partial}_\nu g^{-1} g\right)$$

$$= \left(f - f^2\right)\left(-4g^{-1}\hat{\partial}_\nu g\right) = 4\left(f^2 - f\right)\phi_\nu\ . \qquad (2.42)$$

Para a Segunda parcela de (2.31), temos

$$[\hat{a}_\mu, \hat{H}_{\mu\nu}] = (f)\left(f - f^2\right)[\phi_\mu, (\hat{\partial}_\mu g^{-1}\hat{\partial}_\nu g - \hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g)]$$

$$= \left(f^2 - f^3\right)\left(g^{-1}\hat{\partial}_\mu g\hat{\partial}_\mu g^{-1}\hat{\partial}_\nu g - g^{-1}\hat{\partial}_\mu g\hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g\right.$$

$$\left. - \hat{\partial}_\mu g^{-1}\hat{\partial}_\nu g g^{-1}\hat{\partial}_\mu g + \hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g g^{-1}\hat{\partial}_\mu g\right)$$

$$= \left(f^2 - f^3\right)\left(g^{-1}\hat{\partial}_\mu g\hat{\partial}_\mu g^{-1}\hat{\partial}_\nu g - 2g^{-1}\hat{\partial}_\mu g\hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g + \hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g g^{-1}\hat{\partial}_\mu g\right)\ . \qquad (2.43)$$

Não é difícil verificar que

$$\hat{\partial}_\mu g\hat{\partial}_\mu g^{-1} = (\partial_\mu g - r_\mu g)\left(\partial_\mu g^{-1} - r_\mu g^{-1}\right) = 3\tau_0 \qquad (2.44)$$

e que

$$\tau_\mu g^{-1}\tau_\mu = -2g\ , \qquad (2.45)$$

com o que calculamos

$$g^{-1}\hat{\partial}_\mu g\hat{\partial}_\nu g^{-1}\hat{\partial}_\mu g = g^{-1}(\partial_\mu g - r_\mu g)\left(\partial_\nu g^{-1} - r_\nu g^{-1}\right)(\partial_\mu g - r_\mu g)$$

$$= \left(g^{-1}\tau_\mu - r_\mu\tau_0\right)\left(\partial_\nu g^{-1}\tau_\mu - r_\nu g^{-1}\tau_\mu - \partial_\nu g^{-1} r_\mu g + r_\nu g^{-1} r_\mu g\right)$$

$$= g^{-1}\tau_\mu\partial_\nu g^{-1}\tau_\mu - r_\nu g^{-1}\tau_\mu g^{-1}\tau_\mu - g^{-1} r_\mu\tau_\mu\partial_\nu g^{-1} g + g^{-1} r_\mu\tau_\mu r_\nu g^{-1} g$$

$$- \partial_\nu g^{-1} r_\mu\tau_\mu + r_\nu g^{-1} r_\mu\tau_\mu + r_\mu r_\mu\partial_\nu g^{-1} g - r_\mu r_\mu r_\nu g^{-1} g$$

$$= -2g^{-1}\partial_\nu g + 2r_\nu\tau_0 - \partial_\nu g^{-1} g + r_\nu\tau_0$$

$$= -2g^{-1}\hat{\partial}_\nu g - \hat{\partial}_\nu g^{-1} g = -\phi_\nu .$$

Voltando a (2.43) podemos, então, escrever

$$[\hat{a}_\mu, \hat{H}_{\mu\nu}] = \left(f^2 - f^3\right)\left(3g^{-1}\hat{\partial}_\nu g + 2g^{-1}\hat{\partial}_\nu g - 3\hat{\partial}_\nu g^{-1} g\right)$$

$$= \left(f^2 - f^3\right)\left(8g^{-1}\hat{\partial}_\nu g\right) = 8\left(f^2 - f^3\right)\phi_\nu . \tag{2.46}$$

Para a terceira parcela em (2.31), a partir de (2.35) temos

$$\hat{\Delta}_a\hat{E}_{a\nu} = \hat{\Delta}_a[\hat{\Delta}_a f(\theta)]\phi_\nu = \{\hat{\Delta}_0[\hat{\Delta}_0 f(\theta)] + \hat{\Delta}_1[\hat{\Delta}_1 f(\theta)]\}\phi_\nu .$$

Escrevendo $\theta = \text{arctg}\left(\frac{R_1}{R_0}\right)$ temos $\frac{\partial\theta}{\partial R_0} = -R_1$ e $\frac{\partial\theta}{\partial R_1} = R_0$, com o que calculamos

$$\hat{\Delta}_0 f(\theta) = \frac{\partial f(\theta)}{\partial R_0} - R_0 R_b \frac{\partial f(\theta)}{\partial R_b}$$

$$= -R_1 f'(\theta) - R_0[-R_0 R_1 f'(\theta) + R_1 R_0 f'(\theta)] = -R_1 f'(\theta)$$

e analogamente

$$\hat{\Delta}_1 f(\theta) = R_0 f'(\theta) .$$

Logo

$$\hat{\Delta}_a\hat{E}_{a\nu} = \{\hat{\Delta}_0[-R_1 f'(\theta)] + \hat{\Delta}_1[R_0 f'(\theta)]\}\phi_\nu$$

$$= [R_0 R_1 + R_1^2 f''(\theta) - R_1 R_0 + R_0^2 f''(\theta)]\phi_\nu = f''(\theta)\phi_\nu . \tag{2.47}$$

A segunda equação de Yang-Mills (2.32) é automaticamente satisfeita para conexões da forma (2.33). De fato, para a primeira parcela temos

$$\hat{\partial}_\mu\hat{E}_{a\mu} = \hat{\Delta}_a f(\theta)\hat{\partial}_\mu\phi_\mu = \hat{\Delta}_a f(\theta)\left(\hat{\partial}_\mu g^{-1}\hat{\partial}_\mu g + g^{-1}\hat{\partial}_\mu\partial_\mu g\right) = 0 ,$$

por (2.39) e (2.44). Na Segunda parcela aparece um comutador $[\phi_\mu, \phi_\mu]$ que obviamente é nulo.

Reunindo as parcelas (2.42), (2.46) e (2.47), as equações de Yang-Mills ficam reduzidas a

$$(-4f + 12f^2 - 8f^3 - f'')\phi_\nu \;=\; 0\;,$$

o que significa que basta resolver a equação

$$f'' \;+\; 8f^3 \;-\; 12f^2 \;+\; 4f \;=\; 0\;, \tag{2.48}$$

ou seja,

$$f'' \;+\; 8f\left(f - \frac{1}{2}\right)(f-1) \;=\; 0\;.$$

Aqui é natural percebermos as soluções constantes, para $f = 0$, $\frac{1}{2}$ ou 1. Entretanto a primeira corresponde à conexão nula e a última a uma conexão plana $\hat{a}_\mu = g^{-1}\hat{\partial}_\mu g$, restando então $\hat{a}_\mu = \frac{1}{2}g^{-1}\hat{\partial}_\mu g$ como solução não trivial (i. e.,com curvatura não nula).

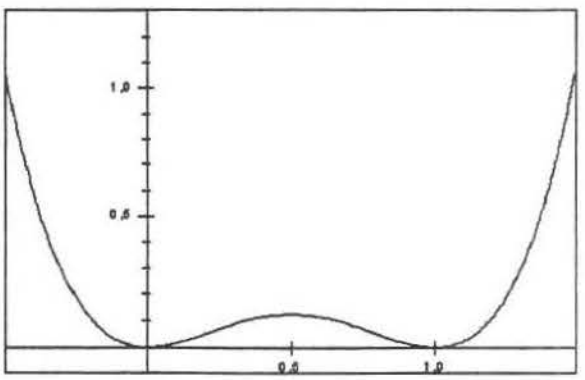
Para encontrar as soluções não constantes, podemos multiplicar a equação (2.48) por $f'(\theta)$, com o que obtemos

$$\begin{aligned} & f'f'' \;+\; 8f^3f' \;-\; 12f^2f' \;+\; 4ff' \;=\; 0 \\ \therefore\;\; & \frac{\mathrm{d}}{\mathrm{d}\theta}\left[\frac{1}{2}(f')^2 \;+\; 2f^4 \;-\; 4f^3 \;+\; 2f^2\right] \;=\; 0 \\ \therefore\;\; & \frac{1}{2}(f')^2 \;+\; 2f^2(f-1)^2 \;=\; C\;. \end{aligned}$$

Esta equação não linear de primeira ordem é equivalente à que surge do estudo do movimento de uma partícula de massa unitária num poço de potencial da forma $2f^2(f-1)^2$.Aqui a variável $\theta$ corresponde ao tempo, a função $f$ à posição da partícula, $\frac{1}{2}(f')^2$ é a energia cinética e a constante de integração $C$ corresponde à energia mecânica da partícula.

Para uma análise qualitativa das soluções, é suficiente observarmos o gráfico desse potencial, conforme aparece na figura 1. Ali podemos perceber, além das soluções constantes que já encontramos, que para $C < 1/8$ temos duas possibilidades de soluções periódicas com oscilações em torno de $f = 0$ ou de $f = 1$. Para $C = 1/8$, dependendo das condições iniciais, podemos ter a solução constante $f = 1/2$, já mencionada e duas soluções não periódicas com f aproximando-se assintoticamente

de $f = 1/2$. Para $C > 1/8$, voltamos a ter soluções periódicas, com oscilações mais amplas, percorrendo os dois "vales" em $f = 0$ e $f = 1$.




Para obter explicitamente as soluções dessa equação, basta observarmos que ela é do tipo elíptico, o que fica mais evidente se escrevermos

$$(f')^2 = C - 4f^2(f-1)^2$$

e substituindo a constante $C$ por $1/C^2$

$$f' = \frac{1}{C}\sqrt{1 - [2Cf(f-1)]^2}\,.$$

As soluções de equações desse tipo podem ser dadas em termos de funções elípticas como as de Jacobi, por exemplo.

# Bibliografia


[1] A. Actor, *Classical Solutions of SU(2) Yang-Mills Theories*, Rev. Mod. Phys., **51**, 461-525 (1979)

[2] A. O. Barut, **Electrodynamics and Classical Theory of Fields and Particles** (Dover, 1980)

[3] D. Bleecker, **Gauge Theory and Variational Principles** (Addison-Wesley, 1981)

[4] Y. Choquet-Bruhat, C. DeWitt-Morette, M. Dillard-Bleick, **Analysis, Manifolds and Physics,** Rev. Ed. (North-Holland,1982)

[5] M. Daniel, C. M. Viallet, *Geometrical Setting of Yang-Mills gauge theories*, Rev. Mod. Phys., **52** (1), 175-197 (1980)

[6] V. De Alfaro, S. Fubini, G. Furlan, *A New Classical Solution of the Yang-Mills Field Equations*, Phys. Lett. **65B** (2), 163-166 (1976)

[7] V. De Alfaro, S. Fubini, G. Furlan, *Properties of* $O(4) \times O(2)$ *Symmetric Solutions of the Yang-Mills Field Equations*, Phys. Lett. **72B** (2), 203-207 (1977)

[8] N. J. Hicks, **Notes on Differential Geometry**(Van Nostrand Reinhold Co., 1971)

[9] Chan Hong-Mo, J. Faridani, Tsou Sheung Tsun, *Non-Abelian Yang-Mills Analogue of Classical Electromagnetic Duality*, Phis. Rev. D, **52**, 6134 (1995)

[10] R. Jackiw, C. Rebbi, *Conformal Properties of a Yang-Mills Pseudoparticle*, Phys. Rev. D, **14**, 517-523 (1976)

[11] S. Kobayashi, K. Nomizu, **Foundations of Differential Geometry**, vol. 1 (Wiley, NY, 1963)

[12] G. Mack, A. Salam, *Finite Component Field Representations of the Conformal Group*, Ann. Phys. (N.Y.), **53**, 174-202 (1969)

[13] M. Nakahara, **Geometry, Topology and Physics** (IOP, 1990)

[14] C. Nash, S. Sen, **Topology and Geometry for Physicists** (Academic Press, 1983)

[15] A. M. Polyakov, *Compact gauge Fields and the Infrared Catastrophe*, Phys. Lett. **59B** (1), 82 (1975).

[16] D. H. Sattinger, O. L. Weaver, **Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics** (Springer-Verlag, NY, 1986)

[17] H. Samelson, **Notes on Lie Algebras** (Springer-Verlag, 1969)

[18] B. M. Schechter, *Yang-Mills Theory on the Hipertorus*, Phys. Rev. D, **16**, 3015-3020 (1976)

[19] M. O. Terra-Cunha, **A Geometria e os Instantons da Teoria de Yang & Mills** $SU(2)$, Tese de Mestrado do IFGW - UNICAMP (1997)

[20] F. W. Warner, **Foundations of Differentiable Manifolds and Lie Groups** (Scott, Foresman and Company, 1971)