Lecture 0: Differential Geometry, General Relativity, Classical Yang-Mills Theory

Let's start with a very condensed review of some of the elements of Differential Geometry, General Relativity, and Yang-Mills Theory. This should not be confused with an introduction to these subjects. If you plan to be a theoretical physicist, it is a good idea to eventually learn about these topics from a somewhat more mathematically sophisticated perspective than the one presented here: I recommend the relevant sections of Wald's "General Relativity" and Arnold's "Mathematical Methods of Classical Mechanics."

Let me start with a few notational remarks. If you find any of my notation confusing, please let me know – I have probably made a mistake, forgotten to explain something, or in any case will be happy to explain it again! I will use the notation $\partial v^\alpha/\partial x_\beta = \partial_\beta v^\alpha = v^\alpha_{,\beta}$ for ordinary partial derivatives, and $\nabla_\beta v^\alpha = v^\alpha_{;\beta}$ for covariant derivatives. I will follow the **Einstein summation convention**: whenever a term contains a repeated index, with one up and one down, a sum over that index is implied. I will take the "timelike" and "spacelike" components of the metric to have negative and positive signs, respectively: *e.g.* $ds^2 = -dt^2 + d\vec{x}^2$. More generally, I will use GR sign conventions that match those in the textbooks by Misner-Thorne-Wheeler (MTW), Wald, and Weinberg; these conventions are nicely summarized on the inside cover of MTW.

1 Differential Geometry

1.1 Manifolds and tensors.

Informally, an n-dimensional **manifold** $\mathcal{M}$ is a space that "looks like" $\mathbf{R}^n$ in the neighborhood of every point. For example, the plane, the surface of a ball, and the surface of a donut are all examples of 2-dimensional manifolds: a sufficiently small patch surrounding any point looks like a patch of $\mathbf{R}^2$.

We are free to choose many different coordinate systems to describe a given part of a manifold; let x^α and $\tilde{x}^\alpha$ be two such coordinate systems. A **tensor** of rank (m, n) has m upper ("contravariant") indices and n lower ("covariant") indices; if the components in the x-coordinate system are $T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n}$, then the components in the $\tilde{x}$-coordinate system are

$$\tilde{T}^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} = \frac{\partial\tilde{x}^{\alpha_1}}{\partial x^{\gamma_1}}\cdots\frac{\partial\tilde{x}^{\alpha_m}}{\partial x^{\gamma_m}}\frac{\partial x^{\delta_1}}{\partial\tilde{x}^{\beta_1}}\cdots\frac{\partial x^{\delta_n}}{\partial\tilde{x}^{\beta_n}}T^{\gamma_1\ldots\gamma_m}_{\delta_1\ldots\delta_n}. \tag{1}$$

Rank $(1,0)$ and $(0,1)$ tensors are called "contravariant vectors" and "covariant vectors," respectively.

1.2 Covariant derivatives and connections.

If we start with a rank (m,n) tensor field $T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n}$ and take its ordinary partial derivative ∂_μ, the resulting object $\partial_\mu T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n}$ no longer transforms like a tensor. (Check this.) We would like to define a new **covariant derivative** ∇_μ such that $\nabla_\mu T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n}$ *does* transform like a tensor of rank $(m,n+1)$. To achieve this, we introduce the **connection** whose components $\Gamma^\alpha_{\beta\gamma}$ do *not* transform like a tensor, but instead transform as

$$\tilde{\Gamma}^\alpha_{\beta\gamma} = \frac{\partial\tilde{x}^\alpha}{\partial x^\rho}\frac{\partial x^\sigma}{\partial\tilde{x}^\beta}\frac{\partial x^\tau}{\partial\tilde{x}^\gamma}\Gamma^\rho_{\sigma\tau} + \frac{\partial\tilde{x}^\alpha}{\partial x^\rho}\frac{\partial^2 x^\rho}{\partial\tilde{x}^\beta\partial\tilde{x}^\gamma} \qquad (2)$$

and then define the covariant derivative of a tensor as:

$$\begin{aligned}\nabla_\mu T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} \;=\;& \partial_\mu T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} \\ &+\Gamma^{\alpha_1}_{\mu\nu}\;T^{\nu\ldots\alpha_m}_{\beta_1\ldots\beta_n} + \ldots + \Gamma^{\alpha_m}_{\mu\nu}\;T^{\alpha_1\ldots\nu}_{\beta_1\ldots\beta_n} \\ &-\Gamma^\nu_{\mu\beta_1}T^{\alpha_1\ldots\alpha_m}_{\nu\ldots\beta_n} - \ldots - \Gamma^\nu_{\mu\beta_n}T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\nu}. \end{aligned} \qquad (3)$$

This does transform like a tensor. (Check this.) We will assume that the connection has **vanishing torsion**: $\Gamma^\alpha_{\beta\gamma} = \Gamma^\alpha_{\gamma\beta}$.

Now consider a curve $x^\mu(\tau)$ with tangent vector $u^\mu(\tau) = dx^\mu/d\tau$. A tensor $T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n}(\tau)$ is **parallel transported** if it evolves along the curve as

$$\begin{aligned}\frac{dT^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n}}{d\tau} + u^\mu\Big[&\Gamma^{\alpha_1}_{\mu\nu}T^{\nu\ldots\alpha_m}_{\beta_1\ldots\beta_n} + \ldots + \Gamma^{\alpha_m}_{\mu\nu}T^{\alpha_1\ldots\nu}_{\beta_1\ldots\beta_n} \\ &-\Gamma^\nu_{\mu\beta_1}T^{\alpha_1\ldots\alpha_m}_{\nu\ldots\beta_n} - \ldots - \Gamma^\nu_{\mu\beta_n}T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\nu}\Big] = 0\end{aligned} \qquad (4)$$

or, in other words, if it is covariantly constant along the direction of the curve

$$u^\mu\nabla_\mu T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} = 0. \qquad (5)$$

A **geodesic** is a curve $x^\mu(\tau)$ which parallel transports its own tangent vector:

$$\frac{du^\alpha}{d\tau} + \Gamma^\alpha_{\mu\nu}u^\mu u^\nu = 0 \quad\Leftrightarrow\quad u^\mu\nabla_\mu u^\alpha = 0. \qquad (6)$$

This is called the **geodesic equation**.

Unlike ordinary partial derivatives, covariant derivatives do *not* commute. The **Riemann curvature tensor** $R^\alpha{}_{\beta\gamma\delta}$ may be defined as a measure of their failure to commute

$$(\nabla_\gamma\nabla_\delta - \nabla_\delta\nabla_\gamma)v^\alpha = R^\alpha{}_{\beta\gamma\delta}v^\beta, \qquad (7)$$

from which we find (check this)

$$R^\alpha{}_{\beta\gamma\delta} = \Gamma^\alpha_{\beta\delta,\gamma} - \Gamma^\alpha_{\beta\gamma,\delta} + \Gamma^\alpha_{\gamma\sigma}\Gamma^\sigma_{\beta\delta} - \Gamma^\alpha_{\delta\sigma}\Gamma^\sigma_{\beta\gamma}. \qquad (8)$$

Eq. (7) is called the Ricci identity. More generally, the commutator of two covariant derivatives acting on a tensor is given by (check this):

$$\begin{aligned}(\nabla_\gamma\nabla_\delta - \nabla_\delta\nabla_\gamma)T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} &= R^{\alpha_1}{}_{\mu\gamma\delta}T^{\mu\ldots\alpha_m}_{\beta_1\ldots\beta_n} + \ldots + R^{\alpha_m}{}_{\mu\gamma\delta}T^{\alpha_1\ldots\mu}_{\beta_1\ldots\beta_n} \\ &\quad - R^{\mu}{}_{\beta_1\gamma\delta}T^{\alpha_1\ldots\alpha_m}_{\mu\ldots\beta_n} - \ldots - R^{\mu}{}_{\beta_n\gamma\delta}T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\mu}. \end{aligned} \tag{9}$$

The Riemann tensor may be contracted to give the **Ricci curvature tensor**:

$$R_{\alpha\beta} \equiv R^{\gamma}{}_{\alpha\gamma\beta}. \tag{10}$$

1.3 The metric.

Note that everything we have done so far does *not* make use of a metric – so we see that ideas like tensors, differentiation, parallel transport, geodesics, and curvature are all well-defined, even *before* we have introduced any notion of *distance* on our manifold. What new features arise from adding a metric tensor $g_{\alpha\beta}$, and its inverse $g^{\alpha\beta}$ (defined such that $g^{\alpha\beta}g_{\beta\gamma} = \delta^\alpha_\gamma$)?

We can raise and lower indices ($v^\alpha = g^{\alpha\beta}v_\beta$, $v_\alpha = g_{\alpha\beta}v^\beta$) and define an inner product ($a\cdot b = g_{\alpha\beta}a^\alpha b^\beta = g^{\alpha\beta}a_\alpha b_\beta$). We can define the **Ricci scalar**

$$R \equiv g^{\alpha\beta}R_{\alpha\beta}. \tag{11}$$

The metric $g_{\mu\nu}$ picks out a preferred connection on $\mathcal{M}$ (the **Levi-Civita connection**): the unique connection such that $\nabla_\mu g_{\alpha\beta} = 0$:

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2}g^{\alpha\delta}[g_{\beta\delta,\gamma} + g_{\gamma\delta,\beta} - g_{\beta\gamma,\delta}]. \tag{12}$$

(Derive this. Hint: start from $g_{\beta\delta;\gamma} + g_{\gamma\delta;\beta} - g_{\beta\gamma;\delta} = 0$.)

We can define the length of a curve $x^\mu(\tau)$, connecting two points A and B; it is $\int_A^B ds$ where the line element is

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu. \tag{13}$$

Now geodesics have another special property: they are paths from A to B whose length is stationary under small variations of the path. This is reminiscent of classical mechanics, in which classical trajectories are paths from A to B whose *action* is stationary under small variations of the path. Indeed, this perspective gives a different (and often more convenient) route to the geodesic equation, as the Euler-Lagrange equations

$$\frac{\partial L}{\partial x^\mu} = \frac{d}{d\tau}\left[\frac{\partial L}{\partial \dot{x}^\mu}\right] \qquad L = \frac{1}{2}g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta \tag{14}$$

where, in this equation, $\dot{x}^\mu = dx^\mu/d\tau$, where τ is an affine parameter along the curve (which, for a massive particle, is proportional to its proper time). Check that Eqs. (6) and (14) are equivalent.

Before moving on to General Relativity, note that the Riemann tensor satisfies the **Bianchi identity**:

$$R^{\alpha}{}_{\beta\gamma\delta;\epsilon} + R^{\alpha}{}_{\beta\delta\epsilon;\gamma} + R^{\alpha}{}_{\beta\epsilon\gamma;\delta} = 0. \tag{15}$$

and $R_{\alpha\beta\gamma\delta} = g_{\alpha\mu}R^{\mu}{}_{\beta\gamma\delta}$ has a number of symmetries under index permutation:

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta} = -R_{\beta\alpha\gamma\delta} = -R_{\alpha\beta\delta\gamma}, \quad R_{\alpha[\beta\gamma\delta]} = R_{[\alpha\beta\gamma\delta]} = 0, \tag{16}$$

where and $[\ldots]$ indicates anti-symmetrization of the enclosed indices.

2 General Relativity

General Relativity is defined by the action

$$S = \int d^4x\sqrt{-g}\left[\frac{R-2\Lambda}{16\pi G_N} + \mathcal{L}_m\right], \tag{17}$$

where g is the determinant of $g_{\mu\nu}$, G_N is Newton's gravitational constant, and Λ is the cosmological constant. The action is a functional of the gravitational field $g_{\mu\nu}$, as well as the various matter fields (which are all contained in $\mathcal{L}_m$). We obtain the classical equations of motion for the matter fields Φ by demanding that the action must be stationary ($\delta S = 0$) under an arbitrary variation $\delta\Phi$ around a classical solution Φ.

We could obtain the Einstein equations (the classical equations of motion for the metric) in a similar way, by demanding that $\delta S = 0$ under an arbitrary variation $\delta g_{\mu\nu}$ around a classical solution $g_{\mu\nu}$; but this turns out to be rather cumbersome, and there is an easier and better way: the **Palatini method**. In the Palatini method, we start by *forgetting* the relationship (12) between the metric and the connection, and regarding $g^{\mu\nu}$ and $\Gamma^{\alpha}_{\beta\gamma}$ as independent fields to be varied. Thus we write the action in the form

$$S = \int d^4x\sqrt{-g}\left[\frac{g^{\alpha\beta}R_{\alpha\beta}-2\Lambda}{16\pi G_N} + \mathcal{L}_m\right], \tag{18}$$

and note that $R_{\alpha\beta}$ is independent of the metric (it only depends on $\Gamma^{\alpha}_{\beta\gamma}$), while everything else in the action (except for $R_{\alpha\beta}$) is independent of $\Gamma^{\alpha}_{\beta\gamma}$.

We start by varying $\Gamma^{\alpha}_{\beta\gamma}$. Note that $\delta\Gamma^{\alpha}_{\beta\gamma}$ is a tensor, even though $\Gamma^{\alpha}_{\beta\gamma}$ is not; and the variation of the action is:

$$\begin{aligned}\delta S &= \int d^4x\sqrt{-g}\frac{g^{\alpha\beta}(\delta\Gamma^{\gamma}_{\alpha\beta;\gamma} - \delta\Gamma^{\gamma}_{\alpha\gamma;\beta})}{16\pi G_N} \\ &= \int d^4x\sqrt{-g}\frac{(g^{\alpha\mu}{}_{;\mu}\delta^{\beta}_{\gamma} - g^{\alpha\beta}{}_{;\gamma})\delta\Gamma^{\gamma}_{\alpha\beta}}{16\pi G_N} \end{aligned} \tag{19a}$$

where we have integrated by parts in the second line. Thus, $\delta S = 0$ implies $g^{\alpha\mu}{}_{;\mu}\delta^{\beta}_{\gamma} - g^{\alpha\beta}{}_{;\gamma} = 0$, which implies $g_{\alpha\beta;\gamma} = 0$. In other words, varying $\Gamma^{\alpha}_{\beta\gamma}$ leads us back to the Levi-Civita connection formula (12).

Before proceeding with the metric variation, we need the following two facts. First, if we make a variation $\delta g_{\mu\nu}$, what is the induced variation $\delta g^{\mu\nu}$ in the inverse metric? The answer is obtained by varying the relation $g^{\alpha\beta}g_{\beta\gamma} = \delta^{\alpha}_{\gamma}$ to obtain $g_{\gamma\beta}\delta g^{\beta\alpha} + g^{\alpha\beta}\delta g_{\beta\gamma} = 0$. Second, if we make a variation $\delta g_{\mu\nu}$ in the metric (or in any invertible matrix, for that matter), what is the induced variation δg in its determinant? The answer is given by Jacobi's formula from linear algebra: $\delta g = g g^{\alpha\beta}\delta g_{\alpha\beta}$; also note that this result can be derived from another famous fact in linear algebra: for a square matrix A, $\mathrm{Det}A = \exp[\mathrm{Tr}(\ln A)]$. With these facts, we can apply a variation in $\delta g^{\mu\nu}$ to Eq. (18)

$$\delta S = \int d^4x\sqrt{-g}\,\delta g^{\mu\nu}\left[\frac{R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu}}{16\pi G_N} - \frac{1}{2}g_{\mu\nu}\mathcal{L}_m + \frac{\partial\mathcal{L}_m}{\partial g^{\mu\nu}}\right] \tag{20}$$

from which we obtain the **Einstein equation**

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu} \tag{21}$$

where we have defined the **Einstein tensor**

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \tag{22}$$

and the **energy-momentum tensor** (a.k.a. the **stress-energy tensor**):

$$T_{\mu\nu} = g_{\mu\nu}\mathcal{L}_m - 2\frac{\partial\mathcal{L}_m}{\partial g^{\mu\nu}}. \tag{23}$$

(Check the preceding few steps.)

By contracting the Bianchi identities, we learn that $G^{\alpha\beta}{}_{;\beta} = 0$; together with the Einstein equations (21), this implies that $T^{\mu\nu}$ is locally conserved

$$T^{\alpha\beta}{}_{;\beta} = 0. \tag{24}$$

We will return, in a subsequent lecture, to a fuller discussion of $T_{\mu\nu}$ and its physical meaning. For now, consider the example in which the only matter field is a single real scalar field φ with Lagrangian

$$\mathcal{L}_m = -\frac{1}{2}g^{\mu\nu}\varphi_{,\mu}\varphi_{,\nu} - V(\varphi) \tag{25}$$

then $T_{\mu\nu}$ is given by

$$T_{\mu\nu} = \varphi_{,\mu}\varphi_{,\nu} - g_{\mu\nu}[\frac{1}{2}g^{\alpha\beta}\varphi_{,\alpha}\varphi_{,\beta} + V(\varphi)]. \tag{26}$$

A final point: in the absence of external (non-gravitational) forces, test particles move along geodesics of the metric $g_{\mu\nu}$.

3 Yang-Mill Theory

Let us briefly look at Yang-Mills theory. You can learn a lot by comparing its structure to that of General Relativity.

We first introduce the gauge covariant derivative

$$D_\mu \equiv \partial_\mu + igA_\mu. \tag{27}$$

Then we use the Ricci identity to define the Yang-Mills field strength $F_{\mu\nu}$ as the curvature of the gauge connection A_μ:

$$[D_\mu, D_\nu] = igF_{\mu\nu}. \tag{28}$$

This implies

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu]. \tag{29}$$

Under a Yang-Mills gauge transformation $U(x)$, we want the covariant derivative operator to transform "nicely":

$$D_\mu \to \tilde{D}_\mu = UD_\mu U^{-1} \tag{30}$$

which implies that A_μ transforms as

$$A_\mu \to \tilde{A}_\mu = UA_\mu U^{-1} - \frac{i}{g}U(\partial_\mu(U^{-1})) = UA_\mu U^{-1} + \frac{i}{g}(\partial_\mu U)U^{-1}. \tag{31}$$

From Eqs. (28, 30) we see that $F_{\mu\nu}$ also transforms nicely:

$$F_{\mu\nu} \to \tilde{F}_{\mu\nu} = UF_{\mu\nu}U^{-1}. \tag{32}$$

From here we see that the term

$$-\frac{1}{4\xi}\mathrm{Tr}[F_{\mu\nu}F^{\mu\nu}] \tag{33}$$

is both Lorentz- and gauge-invariant; this is the gauge-boson kinetic term, with ξ a normalization constant discussed below. For concreteness, let the Yang-Mills gauge group be a unitary (or special unitary) group, and let Φ and Ψ be multiplets of scalar and spinor fields, respectively, which transform as

$$\Phi \to \tilde{\Phi} = U\Phi \qquad\qquad \Psi \to \tilde{\Psi} = U\Psi; \tag{34}$$

then we can construct gauge invariant scalar and spinor kinetic terms as follows:

$$(D_\mu\Phi)^\dagger(D^\mu\Phi) \qquad\qquad i\bar{\Psi}\not{D}\Psi. \tag{35}$$

Now introduce a basis of generators

$$A_\mu = A^a_\mu T^a, \qquad F_{\mu\nu} = F^a_{\mu\nu}T^a, \qquad [T^a, T^b] = if^{abc}T^c \tag{36}$$

which are chosen to be orthonormal in the following sense

$$\mathrm{Tr}(T^a T^b) = \xi \delta^{ab} \tag{37}$$

where ξ is a constant which defines our generator normalization convention. From here it follows that

$$f^{abc} = \frac{-i}{\xi} Tr([T^a, T^b] T^c) \tag{38}$$

which implies that f^{abc} is totally anti-symmetric. When expanded in components, the field strength is

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f^{abc} A^b_\mu A^c_\nu \tag{39}$$

and the gauge kinetic term (33) becomes

$$-\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a . \tag{40}$$