Geometry & Physics

Prof. J.J. Binney

Oxford University

Michaelmas Term 2000

Books

The single most useful book is B. Schutz *Geometrical Methods of Mathematical Physics* (CUP). However, in preparing these lectures I have made extensive use of N.J. Hicks *Notes on Differential Geometry* (van Nostrand). I believe this book to be out of print, but it emphasizes completely coordinate free methods to a degree that I find useful. My treatment of classical mechanics leans on V.I. Arnold *Mathematical Methods of Classical Mechanics.* For gauge theories I have used L.H. Ryder *Quantum Field Theory* and C. Quigg *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions.*

1 Basic equipment

A **manifold** $\mathcal{M}$ is a topological space S together with a collection of continuous 1–1 maps $\phi_\alpha : U_\alpha \to \mathcal{R}^n$ between an open set $U_\alpha \subset S$ and an open set of $\mathcal{R}^n$ for some integer n. The collection $\{U_\alpha\}$ must cover S.

The pair (ϕ_α, U_α) is called a **chart** for S. U_α is called the **domain** of the chart. A complete set of charts is called an **atlas** for M.

Example 1.1

A circle is a manifold. One chart is provided by the angle θ and the open interval $0 < \theta < 2\pi$. Since the origin is excluded from the domain of this chart, we need another chart. We could use θ in the domain $-\pi < \theta < \pi$. Locally the circle looks like $\mathcal{R}^1$, but globally it is quite different. This difference gives rise to the need for more than one chart.

Since for $s \in S$, $\phi_\alpha(s) \in \mathcal{R}^n$ we may consider ϕ_α to provide n real-valued functions ϕ_α^i on S, the $i^{\rm th}$ such function returning the value of the $i^{\rm th}$ coordinate of the point in $\mathcal{R}^n$ to which s is mapped.

In $U_\alpha \cap U_\beta$ we obtain a map $\phi_\alpha \circ {\phi_\beta}^{-1} : \mathcal{R}^n \to \mathcal{R}^n$ of $\mathcal{R}^n$ into itself. The manifold is C^1 **differentiable** iff all of the n^2 derivatives

$$\frac{\partial}{\partial x^j}\left(\phi_\alpha \circ \phi_\beta^{-1}\right)^i\Big|_{(x^1,\ldots,x^n)}$$

exist. It is C^k differentiable if all higher-order derivatives up to $k^{\rm th}$ ones exist. I shall assume throughout that we shall are working with C^∞ manifolds.

Example 1.2

We require at least two charts for the 2-sphere S^2. These might be two systems of spherical polar coordinates $S \to (\vartheta, \varphi)$ with non-parallel polar axes. [So that the coordinate singularities at $(0,0)$ do not coincide.] Thus $\phi_1^1 = \vartheta$, $\phi_1^2 = \varphi$ for the first orientation of the polar axis. It is evident that the map $\phi_1 \circ {\phi_2}^{-1}$ is in this case a complex trigonometric function of two arguments. Alternative charts are provided by embedding the sphere in $\mathcal{R}^3$: then we can use $\phi_1^1 = x, \phi_1^2 = y$, with U_1 the hemisphere $z > 0$, and $\phi_2^1 = x, \phi_2^2 = y$, with U_2 the hemisphere $z < 0$.

Exercise (1):

Explain why additional charts are required for the sphere in the last example, and supply same.

Let f be a real-valued function on some open set $U \subset M$. We say that f is a C^k function iff all the maps $f \circ \phi_\alpha^{-1} : \mathcal{R}^n \to \mathcal{R}$ are k times differentiable at all points $\mathbf{x}$ that satisfy $\phi^{-1}(\mathbf{x}) \in U \cap U_\alpha$. I shall henceforth assume that functions are C^∞. Given a point $m \in M$, let $\mathcal{F}_m$ denote the set of all real-valued functions that are defined in a neighbourhod of m. The coordinate functions ϕ^i of any chart are elements of $\mathcal{F}_m$.

1.1 Tangent vectors

The big difference between curved and flat spaces lies in the nature of vectors. In a flat space we can think of a vector as joining two points in the space. This idea fails in a curved space and we proceed as follows. Let $f, g \in \mathcal{F}_m$ for $m \in M$. Then a **tangent vector** X_m at m is map $X_m : \mathcal{F}_m \to \mathcal{R}$ that has the following properties

$$\begin{aligned} X_m(af + bg) &= aX_m(f) + bX_m(g) \quad \text{with } a, b \in \mathcal{R} \\ X_m(fg) &= g(m)X_m(f) + f(m)X_m(g). \end{aligned} \tag{1.1}$$

The first condition states that X_m is a linear function. The second states that it obeys Leibnitz' rule for differentiating a product. We shall see that $X_m(f)$ returns the rate of change of f as one moves past m along a curve that is tangent to X_m at a rate that is determined by the magnitude of X_m.

Exercise (2):

Show from (1.1) that X_m annihilates any constant function.

It is clear that if X_m and Y_m are tangent vectors at m, then so is $Z_m \equiv aX_m + bY_m$ for any $a, b \in \mathcal{R}$. That is, the tangent vectors at m form a vector space in the sense of ordinary linear algebra. This space is called the **tangent space** at m, $\mathcal{T}_m$.

Given a chart (ϕ, U) with $m \in U$ it is easy to see that the following are elements of $\mathcal{T}_m$:

$$\left(\frac{\partial}{\partial\phi^i}\right)_m \quad \text{where} \quad \left(\frac{\partial}{\partial\phi^i}\right)_m f \equiv \frac{\partial}{\partial x^i} f \circ \phi^{-1}(\mathbf{x}) \quad \text{for } m = \phi^{-1}(\mathbf{x}). \tag{1.2}$$

On applying $(\partial/\partial\phi^i)_m$ to $\phi^j \in \mathcal{F}_m$ it becomes clear that these are linearly-independent elements of $\mathcal{T}_m$. It is not hard to show that they form a basis for the vector space $\mathcal{T}_m$. In fact, for any $X_m \in \mathcal{T}_m$ and $f \in \mathcal{F}_m$ we have[1]

$$X_m(f) = \sum_{i=1}^n X_m(\phi^i)\left(\frac{\partial}{\partial\phi^i}\right)_m f. \tag{1.3}$$

The numbers $X_m^i \equiv X_m(\phi^i)$ are called the **components** of X_m in the chart ϕ.

Equation (1.3) shows that $\mathcal{T}_m$ has (as a linear vector space) the same dimension n as M. If M is flat, and therefore itself a linear vector space, $\mathcal{T}_m$ becomes indistinguishable from M, and this is why we do not normally have to bother with $\mathcal{T}_m$.

Let C be a curve on M that passes through m and let $c \in \mathcal{R}$ be parameter which specifies a location on C in such a way that $c = 0$ corresponds to m. Then it is easy to check that $(\mathrm{d}/\mathrm{d}c)_0$ lies in $\mathcal{T}_m$.

1.2 Vector fields and Lie derivatives

A **vector field** X on $U \subset M$ is a rule that assigns to each $m \in U$ a tangent vector $X_m \in \mathcal{T}_m$. When we apply X to $f \in \mathcal{F}_m$ we obtain another function $g \in \mathcal{F}_m$: $g(m) \equiv X_m(f) \in \mathcal{R}$. So vector fields map $\mathcal{F}_m$ to itself.

If we apply a second vector field Y to the function Xf, we get a further function YXf. Does this imply that YX is itself a vector field? No, because we can show that $(YX)_m$ satisfies only the first of the defining conditions (1.1):

$$\begin{aligned}(YX)_m f_1 f_2 &= Y_m(Xf_1f_2) = Y_m\big(f_2X(f_1) + f_1X(f_2)\big)\\ &= Y_m(f_2)X_m(f_1) + f_2(m)Y_m\big(X(f_1)\big) + Y_m(f_1)X_m(f_2) + f_1(m)Y_m\big(X(f_2)\big).\end{aligned} \tag{1.4}$$

The second and final terms in the last line are all that is required for satisfaction of the second of conditions (1.1):

$$f_1(m)(YX)_m(f_2) + f_2(m)(YX)_m(f_1).$$

The other two terms in (1.4) are unwanted. Notice that we would get them also if we calculated $(XY)_m f_1 f_2$ rather than $(YX)_m f_1 f_2$. From this it follows that commutator

$$[X, Y] \equiv XY - YX \tag{1.5}$$

is a vector field, since the unwanted extra terms from the two products will cancel. $[X, Y]$ is called the **Lie bracket** of X and Y or the **Lie derivative** $\pounds_X Y$ of Y, for reasons which will emerge shortly.

Exercise (3):

Show that

$$\left[\frac{\partial}{\partial\phi^i}, \frac{\partial}{\partial\phi^j}\right] = 0 \quad \forall\ i, j. \tag{1.6}$$

[1] The proof involves writing $f(m') = f(m) + \sum_i x^i f_i(m')$, where $x^i \equiv \phi^i(m') - \phi^i(m)$ and the f_i equal $\partial f/\partial\phi^i$ at $m' = m$. This done, one exploits the linearity of X_m and the fact that it annihilates constant functions.

We showed above that for any curve C the differential operator $(\mathrm{d}/\mathrm{d}c)$ is an element of $\mathcal{T}_m$. In fact, given a vector field X and a point m we can construct a curve such that near m, $X = (\mathrm{d}/\mathrm{d}c)$ along the curve. To see this we consider this system of coupled o.d.e.s:

$$\frac{\mathrm{d}\phi^i(m_c)}{\mathrm{d}c} = (X\phi^i)(m_c) \qquad (i = 1, \ldots, n) \ , \tag{1.7}$$

where $m_c \in M$ is the point that corresponds to $c \in \mathcal{R}$. For sufficiently small U these 1st-order equations have a unique solution with $\phi^i(m_0) = \phi^i(m)$. $C(c)$ is called an **integral curve** of X.

Example 1.3

What is the integral curve of $(\partial/\partial\phi^j)$? For this choice of X equations (1.7) read

$$\frac{\mathrm{d}\phi^i(m_c)}{\mathrm{d}c} = \delta^i_j \qquad (i = 1, \ldots, n) \ . \tag{1.8}$$

so ϕ^i = const for $i \neq j$ and $\phi^j = c$. i.e., the integral curves of a basis field run parallel to the coordinate axes. Notice that $(\partial/\partial\phi^i)$ contains more information than just ϕ^i because its integral curve depends on all the coordinate functions, not just $\phi^i()$.

Note:

By analogy with the case of the pair $\big(\phi^i, (\partial/\partial\phi^i)\big)$ people often write $\mathrm{d}/\mathrm{d}c$ in place of X. The advantage of this notation is that it explicitly associates with a vector field its associated parameter c. The disadvantages are (i) that it uses the symbol 'd' for something other than exterior differentiation (see below), (ii) that it implies that X is defined by a function c, which it is not, and (iii) that it can obscure the fact that that c is a real parameter on a curve, and can be incremented without conceptual fuss. I usually prefer to denote a vector field by a capital letter and its parameter by the corresponding small letter.

Equation (1.7) defines exactly one integral curve of X through any $m \in M$. We use these curves to define the **exponential map** $\exp(X) : M \to M$ as follows. $\exp(X)(m)$ is the point one reaches by integrating equations (1.7) from $c = 0$ to $c = 1$ from initial condition $m_0 = m$. The notation follows from the evident fact that $\exp(X) \circ \exp(X) = \exp(2X)$, and more generally that $\exp(aX) \circ \exp(bX) = \exp\big((a + b)X\big)$. However, in general $\exp(X) \circ \exp(Y) \neq \exp(X + Y)$.

Taylor expanding the left side of (1.7) about $c = 0$ we have for sufficiently small c

$$\begin{aligned} \phi^i(m_c) &= \phi^i(m_0) + c\Big(\frac{\mathrm{d}\phi^i(m_c)}{\mathrm{d}c}\Big)_0 + \frac{c^2}{2!}\Big(\frac{\mathrm{d}^2\phi^i(m_c)}{\mathrm{d}c^2}\Big)_0 + \cdots \\ &= \phi^i(m) + cX_m(\phi^i) + \frac{c^2}{2!}X_m(X(\phi^i)) + \cdots \end{aligned} \tag{1.9}$$

Hence we have a Taylor expansion for the coordinates of $\exp(cX)$;

$$\phi^i\big(\exp(cX)(m)\big) = \bigg(\Big(1 + cX + \frac{(cX)^2}{2!} + \cdots\Big)\phi^i\bigg)(m). \tag{1.10}$$

Let us denote the power series in the operator X that appears on the right of this equation as e^{cX}. Then we have

$$\phi^i\big(\exp(cX)(m)\big) = \big(\mathrm{e}^{cX}\phi^i\big)(m) \ . \tag{1.11}$$

Now suppose we have two vector fields A, B and that we use the exponential map to push m first by a parallel to A and then by b parallel to B. The coordinates of the final point $\exp(bB) \circ \exp(aA)(m)$ are

$$\phi^i\big(\exp(bB) \circ \exp(aA)(m)\big) = \Big(\mathrm{e}^{bB}\mathrm{e}^{aA}\phi^i\Big)(m). \tag{1.12}$$

If, on the other hand we push m first along B and then along A our final point will have coordinates

$$\phi^i\big(\exp(aA)\circ\exp(bB)(m)\big) = \Big(\mathrm{e}^{aA}\mathrm{e}^{bB}\phi^i\Big)(m). \tag{1.13}$$

Hence recalling that e^{aA} is just a shorthand for the power series in (1.10) we have that the difference between the coordinates of the two final points is

$$\begin{aligned}\phi^i\big(\exp(aA)\circ\exp(bB)(m)\big)&\\ -\phi^i\big(\exp(bB)\circ\exp(aA)(m)\big) &= \Big(\Big[\mathrm{e}^{aA},\mathrm{e}^{bB}\Big]\phi^i\Big)(m)\\ &= \Big(\big(ab[A,B]+\mathrm{O}(a^2)\big)\phi^i\Big)(m).\end{aligned} \tag{1.14}$$

If the integral curves of A and B were to mesh together to form a coordinate system, both end points would have coordinates (a,b). Hence the integral curves of A and B form a coordinate system iff the difference (1.14), which to lowest order is proportional to the Lie bracket, vanishes.

1.3 Mapping manifolds

Let M_1 and M_2 be manifolds (not necessarily of the same dimension) and let $\alpha : M_1 \to M_2$ be a map between them. This map is said to be C^∞ if, in appropriate coordinate patches, $\phi_i\circ\alpha\circ\phi_j^{-1} : \mathcal{R}^{n_1}\to\mathcal{R}^{n_2}$ is analytic, where ϕ_j, ϕ_i are coordinates for M_1 and M_2, respectively.

α induces a map $\mathcal{T}_m \to \mathcal{T}_{\alpha(m)}$ as follows: given $f\in\mathcal{F}_{\alpha(m)}$ and $X\in\mathcal{T}_m$ define $(\alpha_* X)\in\mathcal{T}_{\alpha(m)}$ by

$$(\alpha_* X)(f) \equiv X(f\circ\alpha). \tag{1.15}$$

α_* is called the **Jacobian map** of α.

Exercise (4):

Given a third manifold M_3 and a C^∞ map $\beta : M_2\to M_3$, show that $(\beta\circ\alpha)_* = \beta_*\circ\alpha_*$.

For any vector field, the Jacobian of the exponential map maps tangent spaces into each other. The image $\exp(X)_* Y_m$ of a vector $Y_m\in\mathcal{T}_m$ under this map is said to the 'result of Lie-dragging Y_m along X'.

The following result explains the origin of the name of $\pounds_X Y$:

$$\pounds_X Y = \lim_{t\to 0}\frac{\exp(-tX)_* Y_{\exp(tX)(m)} - Y_m}{t}. \tag{1.16}$$

1.4 Tensors and forms

As in ordinary linear algebra, an important rôle is played by the linear real-valued functions on $\mathcal{T}_m$. We call these objects **covectors** or **1-forms**. They form an n-dimensional vector space $\mathcal{T}^*_m$, called the **cotangent space**. Thus if $\omega\in\mathcal{T}^*_m$ we have

$$\omega(X_m)\in\mathcal{R} \text{ with } \omega(aX_m+bY_m) = a\omega(X_m)+b\omega(Y_m)\ (X_m, Y_m\in\mathcal{T}_m;\quad a,b\in\mathcal{R}). \tag{1.17}$$

Given any $f\in\mathcal{F}_m$: we define a 1-form $\mathrm{d}f$ like this. For any $X_m\in\mathcal{T}_m$

$$\mathrm{d}f(X_m)\equiv X_m(f)\quad (f\in\mathcal{F}_m,\ X_m\in\mathcal{T}_m). \tag{1.18}$$

Notice that here 'd' has nothing to do with an infinitesimal. Soon we shall generalize d into an operator that makes a field of $(n+1)$-forms out of a field of n-forms, so it is handy to regard elements of $\mathcal{F}_m$ as fields of **0-forms**. The result of using d on a form ω is called the **exterior derivative** of ω.

Exercise (5):

Show that $\mathrm{d}(fg) = f(m)\mathrm{d}g + g(m)\mathrm{d}f$ for $f, g \in \mathcal{F}_m$.

Given a chart ϕ, we have n 1-forms $\mathrm{d}\phi^i$. With (1.18) we have that

$$\mathrm{d}\phi^i \Big(\frac{\partial}{\partial \phi^j}\Big)_m = \delta^i_j. \tag{1.19}$$

From this we can see (i) that the $\mathrm{d}\phi^i$ are n linearly independent elements of $\mathcal{T}^*_m$ and therefore constitute a basis for T^*_m, and (ii) that this basis is, in the usual sense of linear algebra, the one that is conjugate to the $(\partial/\partial\phi^i)_m$ basis for $\mathcal{T}_m$.

Exercise (6):

From (1.18) derive the **chain rule**

$$\mathrm{d}f(X_m) = \sum_i \frac{\partial f}{\partial \phi^i} \mathrm{d}\phi^i(X_m). \tag{1.20}$$

A **second-rank covariant tensor** or $\binom{0}{2}$ tensor S is a real-valued bilinear function on the tensor product $\mathcal{T}_m \otimes \mathcal{T}_m$. (The elements of a tensor product are simply ordered pairs of elements of the producted spaces.) Thus S is a real-valued linear function with two slots: $S(X, Y) \in \mathcal{R}$. A **second-rank contravariant tensor** or $\binom{2}{0}$ tensor T is a real-valued bilinear function on the tensor product $\mathcal{T}^*_m \otimes \mathcal{T}^*_m$, i.e., a function with two slots $T(\omega, \chi) \in \mathcal{R}$. It is obvious that one can go on to define all manner of tensors, including $\binom{0}{3}$ tensors, $\binom{1}{1}$ tensors (functions on $\mathcal{T}_m \otimes \mathcal{T}^*_m$) etc.

It is natural to classify homogeneous 2nd-rand tensors by symmetry: If T is a $\binom{0}{2}$ tensor, from it we extract antisymmetric and symmetric tensors:

$$\begin{aligned} A(X, Y) &\equiv T(X, Y) - T(Y, X); \quad A(Y, X) = -A(X, Y), \\ S(X, Y) &\equiv T(X, Y) + T(Y, X); \quad S(Y, X) = +S(X, Y). \end{aligned} \tag{1.21}$$

Antisymmetric $\binom{0}{k}$ tensors play an especially important rôle and are called **differential forms**. (0-forms and 1-forms are deemed to be antisymmetric.) A **2-form** is an antisymmetric $\binom{0}{2}$ tensor. A **3-form** is a totally antisymmetric $\binom{0}{3}$ tensor, etc. Thus if F and G are a 3-form and a 4-form, respectively,

$$\begin{aligned} F(X, Y, Z) &= -F(Y, X, Z) = F(Y, Z, X) = \cdots, \\ G(W, X, Y, Z) &= -G(X, W, Y, Z) = G(X, Y, W, Z) = -G(X, Y, Z, W) = \cdots \end{aligned} \tag{1.22}$$

Given two 1-forms ω, χ, we can form a 2-form as follows

$$(\omega \wedge \chi)(X, Y) \equiv \omega(X)\chi(Y) - \omega(Y)\chi(X). \tag{1.23}$$

Analogously we define the **wedge product** of three 1-forms as the totally antisymmetric part of the tensor product:[2]

$$\begin{aligned} (\omega \wedge \chi \wedge \psi)(X, Y, Z) =& \omega(X)\chi(Y)\psi(Z) + \omega(Z)\chi(X)\psi(Y) + \omega(Y)\chi(Z)\psi(X) \\ & - \omega(Y)\chi(X)\psi(Z) - \omega(Z)\chi(Y)\psi(X) - \omega(X)\chi(Z)\psi(Y). \end{aligned} \tag{1.24}$$

If we are now given a 1-form ω and a 2-form χ we must clearly define $\omega \wedge \chi$ such that, if χ were itself a wedge product $\chi = \psi \wedge \eta$ of 2 1-forms, we would have

$$\omega \wedge \chi = \omega \wedge (\psi \wedge \eta). \tag{1.25}$$

[2] There is *no* $1/n!$ in front of the definition of the wedge product of n 1-forms.

The required rule is

$$(\omega \wedge \chi)(X,Y,Z) = \omega(X)\chi(Y,Z) + \omega(Y)\chi(Z,X) + \omega(Z)\chi(X,Y). \tag{1.26}$$

Proceeding in this way one may show that if ω and χ are p- and q-forms, respectively, then

$$(\omega \wedge \chi)(X_1,\ldots,X_{p+q}) = \frac{1}{p!q!}\sum_\pi (-1)^\pi \omega(X_{\pi_1},\ldots,X_{\pi_p})\chi(X_{\pi_{p+1}},\ldots,X_{\pi_{p+q}}) \tag{1.27}$$

Here the sum is over all permutations π of the numbers $(1,2,\ldots,p+q)$ and $(-1)^\pi$ is ± 1 depending on whether the permutation is even or odd. With this rule we have

$$\omega \wedge \chi = (-1)^{pq}\chi \wedge \omega. \tag{1.28}$$

It is straightforward to check that with these definitions the product is associative:

$$\omega \wedge (\chi \wedge \psi) = (\omega \wedge \chi) \wedge \psi. \tag{1.29}$$

The wedge product of 4 or 5,..., 1-forms is similarly defined to be the totally antisymmetric part of the relevant tensor product. It is trivial to show that $\omega \wedge (a\chi) = a\omega \wedge \chi$ for $a \in \mathcal{R}$.

Since a 2-form is linear in each slot, we can evaluate it on any set of vectors if we know the values it takes for all possible pairs from a set of basis vectors:

$$\begin{aligned}\omega(X,Y) &= \omega\Big(\sum_i X^i \frac{\partial}{\partial\phi^i}, \sum_j Y^j \frac{\partial}{\partial\phi^i}\Big)\\ &= \sum_{ij} X^i Y^j \omega\Big(\frac{\partial}{\partial\phi^i}, \frac{\partial}{\partial\phi^j}\Big)\end{aligned} \tag{1.30}$$

Hence the number of linearly independent 2-forms is equal to the number of pairs of distinct basis vectors, namely $\frac{1}{2}n(n-1)$. Hence any 2-form can be expanded in terms of coordinate forms as

$$\omega = \sum_{\substack{j=1,n\\ i<j}} \omega_{ij}\mathrm{d}\phi^i \wedge \mathrm{d}\phi^j, \quad \text{where} \quad \omega_{ij} \equiv \omega\Big(\frac{\partial}{\partial\phi^i}, \frac{\partial}{\partial\phi^j}\Big). \tag{1.31}$$

Similarly, the number of linearly independent 3-forms is $\frac{1}{3!}n(n-1)(n-2)$ and so on. In particular there is only 1 linearly independent n-form, and there are exactly as many linearly independent $n-k$ forms as there are linearly independent k forms.

1.5 Exterior derivatives

Equation (1.18) defines the operator d that makes a field of one-forms from a field of 0-forms. It is a derivative operator of a sneaky kind: it uses the derivative power of the extra argument for which it makes space rather than doing any differentiation itself. Consider the following scheme for making a two-slot function out of a 1-form:

$$(\delta\omega)(X,Y) \equiv X\omega(Y) - Y\omega(X). \tag{1.32}$$

This object is antisymmetric and a linear function of both its arguments. Yet it in *not* a 2-form! The trouble is that its value at m does not depend only on X_m, but also on the values of X in an entire neighbourhood of m. (To see this, evaluate $\delta\omega$ on the field aX, where $a \in \mathcal{F}_m$ satisfies $a(m) = 1$.) An object that *does* depend only on X_m, Y_m is $\mathrm{d}\omega$, which is defined by

$$(\mathrm{d}\omega)(X,Y) \equiv X\omega(Y) - Y\omega(X) - \omega([X,Y]). \tag{1.33}$$

Exercise (7):
Prove the last statement by showing that $\mathrm{d}\omega(Z,Y) = 0 \ \forall\ Z$ s.t. $Z_m = 0$.

Exercise (8):
Show that

$$\mathrm{d}\mathrm{d}f = 0. \tag{1.34}$$

The action of d on a general p-form χ is defined by

$$\begin{aligned}(\mathrm{d}\chi)(X_1,\ldots,X_{p+1}) \equiv& \sum_{j=1}^{p+1}(-1)^{j+1}X_j\chi(X_1,\ldots,\hat{X}_j,\ldots,X_{p+1})\\ &+\sum_{i<j}(-1)^{i+j}\chi([X_i,X_j],X_1,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots,X_{p+1}),\end{aligned} \tag{1.35}$$

where a hat indicates that that item should be omitted from the list. In words this says, that first you use each argument in front with the others inside χ and then you put the commutator of each pair in the first slot and the remaining arguments following in order. Meanwhile you keep alternating the signs in front of your terms.

If we expand $\mathrm{d}\chi$ as a linear combination of coordinate forms as in (1.31), we get a reasonably simple expression for the expansion coefficients because all the basis fields $\partial/\partial\phi^i$ commute:

$$\begin{aligned}\text{if}\quad \chi &= \sum_{i<j<k\ldots}\chi_{ijk\ldots}\mathrm{d}\phi^i\wedge\mathrm{d}\phi^j\cdots\\ \text{then}\quad \mathrm{d}\chi &= \sum_{i<j<k\ldots}\mathrm{d}\chi_{ijk\ldots}\wedge\mathrm{d}\phi^i\wedge\mathrm{d}\phi^j\cdots\end{aligned} \tag{1.36}$$

(The proof of the second line simply involves evaluating the action of both sides on an arbitrary collection of basis fields $(\partial/\partial\phi^i)$.)

The following rules for exterior differentiation follow easily from the coordinate representation (1.36) and equation (1.34):

- $\mathrm{d}(\omega+\chi) = \mathrm{d}\omega + \mathrm{d}\chi$;
- $\mathrm{d}(\omega\wedge\chi) = (\mathrm{d}\omega)\wedge\chi + (-1)^p\omega\wedge(\mathrm{d}\chi)$ where ω,χ are p,q forms; (1.37)
- $\mathrm{d}\mathrm{d} = 0$.

When $\mathrm{d}\omega = 0$ we say that ω is **closed**. If $\omega = \mathrm{d}\chi$, we call ω **exact**. We have just seen that all exact forms are closed. On a non-trivial manifold the converse is untrue: there exist non-exact closed forms. The existence of such forms depends upon, and is an important diagnostic of, the global topology of the manifold. **Poincaré's lemma** states that on $\mathcal{R}^n$ all closed forms are exact. Since locally every M is like $\mathcal{R}^n$, given $\mathrm{d}\omega = 0$ we can always find χ such that $\omega = \mathrm{d}\chi$ in a (often large) neighbourhood of any point.

1.6 Integration of forms

If you have an p-form ω and a p-dimensional (sub)-manifold N, then you can integrate ω over N as follows. First set up a coordinate grid $(u^1,\ldots,u^p)$ on N. Then work out the Riemann integral

$$\int\omega \equiv \int \mathrm{d}u^1\cdots\mathrm{d}u^p\,\omega\Big(\frac{\partial}{\partial u^1},\ldots,\frac{\partial}{\partial u^p}\Big), \tag{1.38}$$

where $\mathrm{d}u$ is an infinitesimal, *not* a 1-form. That is, at the centre of each small cell on N we evaluate ω on the vectors conjugate to the coordinates, and then add the resulting numbers for each cell. For the definition (1.38) to make sense it is clearly important that $\int\omega$ should be independent of what coordinate system u^i that we place on N. Let's rewrite (1.38) in terms of new coords $(v^1,\ldots,v^p)$. By (1.3)

$$\frac{\partial}{\partial u^i} = \sum_j \frac{\partial v^j}{\partial u^i}\frac{\partial}{\partial v^j}.$$

We substitute this expansion into ω in (1.37) and exploit the linearity of ω to extract the coefficients of the $(\partial/\partial v^j)$. By the complete antisymmetry of ω we find

$$\omega\Big(\frac{\partial}{\partial u^1},\dots,\frac{\partial}{\partial u^p}\Big)=\frac{\partial(v^1,\dots,v^p)}{\partial(u^1,\dots,u^p)}\,\omega\Big(\frac{\partial}{\partial v^1},\dots,\frac{\partial}{\partial v^p}\Big), \tag{1.39}$$

where $\partial()/\partial()$ denotes the usual Jacobian determinant. On the other hand, from conventional calculus we have

$$\mathrm{d}u^1\cdots\mathrm{d}u^p=\frac{\partial(u^1,\dots,u^p)}{\partial(v^1,\dots,v^p)}\mathrm{d}v^1\cdots\mathrm{d}v^p. \tag{1.40}$$

When (1.39) and (1.40) are used in (1.37) the two Jacobian determinants cancel by the usual theorem, so that

$$\int\omega=\int\mathrm{d}v^1\cdots\mathrm{d}v^p\,\omega\Big(\frac{\partial}{\partial v^1},\dots,\frac{\partial}{\partial v^p}\Big). \tag{1.41}$$

Hence it does not matter what coordinate system you use to evaluate $\int\omega$.

The exterior derivative has been defined such that integration by parts yields **Stokes' theorem**

$$\int_N \mathrm{d}\omega=\int_{\partial N}\omega, \tag{1.42}$$

where ∂N denote the boundary of the $(p+1)$-dimensional region N.

Proof: We break N into small (but not infinitesimal) regions N' that can be covered by a single chart and first prove the result for such a region. We use coordinates designed for the job: the integral curves of $(\partial/\partial\phi^0),\dots,(\partial/\partial\phi^p)$ lie in N', and the boundaries of $\partial N'$ coincide with surfaces $\phi^i=\text{const}$. From the second of eqs (1.36) we have

$$\mathrm{d}\omega=\sum_{\alpha=0}^{n-1}\sum_{i<j<k\dots}\frac{\partial\omega_{ijk\dots}}{\partial\phi^\alpha}\mathrm{d}\phi^\alpha\wedge\mathrm{d}\phi^i\wedge\mathrm{d}\phi^j\cdots. \tag{1.43}$$

The value of this on $(\partial/\partial\phi^0),\dots,(\partial/\partial\phi^p)$ is

$$\mathrm{d}\omega\left(\frac{\partial}{\partial\phi^0},\dots,\frac{\partial}{\partial\phi^p}\right)=\sum_{\alpha=0}^{p}(-1)^\alpha\frac{\partial\omega_{01\dots\hat\alpha\dots p}}{\partial\phi^\alpha}, \tag{1.44}$$

where the hat again implies that that item should be omitted from the list. Hence

$$\begin{aligned}\int_{N'}\mathrm{d}\omega&=\sum_{\alpha=0}^{p}(-1)^\alpha\int\mathrm{d}\phi^0\cdots\mathrm{d}\phi^p\,\frac{\partial\omega_{01\dots\hat\alpha\dots p}}{\partial\phi^\alpha}\\&=\sum_{\alpha=0}^{p}(-1)^\alpha\int\mathrm{d}\phi^0\cdots\widehat{\mathrm{d}\phi^\alpha}\cdots\mathrm{d}\phi^p\,\big[\omega_{01\dots\hat\alpha\dots p}\big]_{\phi^\alpha_{\min}}^{\phi^\alpha_{\max}}.\end{aligned} \tag{1.45}$$

But the last expression is simply $\int_{\partial N'}\omega$. Indeed, by construction each of the $2(p+1)$ faces of N is a surface of constant ϕ^i and they come in pairs: left and right faces. Each term in the sum over α is the sum of integrals of ω over these faces. The alternating sign out front varies the order in which left and right are taken. This ensures that we integrate 'around' the boundary. For example the ordinary line integral

$$\oint_{\text{unit square}}\mathrm{d}\mathbf{l}\cdot\mathbf{A}=\int_0^1\mathrm{d}x\big(A_x(x,0)-A_x(x,1)\big)-\int_0^1\mathrm{d}y\big(A_y(0,y)-A(1,y)\big).$$

When we add the result we have proved for N' over all $N'\subset N$ (1.42) follows because the contributions from common faces of adjoining N' will cancel, leaving only the contribution from the various parts of ∂N.

2 Hamiltonian mechanics

The Hamiltonian mechanics of a time-independent system with N degrees of freedom takes place in phase space M of dimension $n = 2N$. This space is structured by a closed 2-form ω that has the property that if $\omega(X, Y) = 0\ \forall Y$, then $X = 0$. It is called the **symplectic form**.

The physics of the system is specified by a 0-form $H(m)$. Specifically, systems move with time t along a vector field $\mathcal{H}$, whose parameter t is. $\mathcal{H}$ is related to H by

$$\omega(,\mathcal{H}) = \mathrm{d}H. \tag{2.1}$$

Analogously with (2.1), ω enables us to associate a unique vector field $\mathcal{F}$ with *any* function f:

$$\omega(,\mathcal{F}) = \mathrm{d}f. \tag{2.2}$$

Note:

More generally, ω establishes a 1-1 correspondence between the elements of $\mathcal{T}_m^*$ and those of $\mathcal{T}_m$.

Exercise (9):

Prove that $\mathcal{F}$ is unique.

Now the rate of change of the local value of f as one moves with a system is

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \mathrm{d}f(\mathcal{H}) = \omega(\mathcal{H}, \mathcal{F}) = -\mathrm{d}H(\mathcal{F}). \tag{2.3}$$

The function $\{g, f\} \equiv \omega(\mathcal{F}, \mathcal{G})$ is called the **Poisson bracket** of f and g. (Notice the reversal of order!)

Corollary

$H =$ constant along any trajectory.

Let S be a surface in M and let S' be the surface obtained when every point in S has been evolved for time t. Then the **Poincaré invariant theorem** states that $\int_S \omega = \int_{S'} \omega$.

Proof: Consider the cylinder C that is bounded by S, S' and the trajectories that carry points on ∂S into points on $\partial S'$ – let R denote this last 'curved' part of the cylinder. Since $\mathrm{d}\omega = 0$, we have on applying Stokes' theorem to the interior of C that $\int_{\partial C} \omega = 0$. Since ∂C is made up of S, S' and R, we have

$$-\int_S \omega + \int_{S'} \omega + \int_R \omega = 0. \tag{2.4}$$

where the first term has a minus sign because we choose to orient both S and S' in the same sense while Stokes' theorem requires opposite ends of a volume to have opposite orientations. To evaluate $\int_R \omega$ we need to choose two coordinates for R. One obvious candidate is t. A second coordinate s can be generated by deciding on any scheme for denoting position around ∂S such that the same point corresponds to $s = 1$ as to $s = 0$. (i.e., after incrementing s by 1 you've gone right round ∂S.) Then the points $(s, t) \in R$ is the point $s \in \partial S$ reaches at time t. We have

$$\begin{aligned}\int_R \omega &= \int_0^t \mathrm{d}t \int_0^1 \mathrm{d}s\, \omega\Big(\frac{\partial}{\partial s}, \mathcal{H}\Big) \\ &= \int_0^t \mathrm{d}t \int_0^1 \mathrm{d}s\, \mathrm{d}H\Big(\frac{\partial}{\partial s}\Big) \\ &= \int_0^t \mathrm{d}t \int_0^1 \mathrm{d}s\, \Big(\frac{\partial H}{\partial s}\Big) \\ &= \int_0^t \mathrm{d}t\, [H(s(1)) - H(s(0))] = 0.\end{aligned} \tag{2.5}$$

Thus $\int_R \omega = 0$ and it follows that $\int_S \omega = \int_{S'} \omega$ as stated.

Immediate corollaries are that

$$\begin{aligned} \int_{S_4} \omega\wedge\omega &= \int_{S_4'} \omega\wedge\omega \\ \int_{S_6} \omega\wedge\omega\wedge\omega &= \int_{S_6'} \omega\wedge\omega\wedge\omega \\ &\cdots \end{aligned} \tag{2.6}$$

where successive lines relate integrals of dimension $4, 6, \ldots, n$. The n-dimensional version of this result is known as Liouville's theorem.

2.1 Canonical coordinates

Coordinates $(\mathbf{p}, \mathbf{q})$ in which ω takes the simple form

$$\omega = \sum_i^N \mathrm{d}p^i \wedge \mathrm{d}q^i \tag{2.7}$$

are called **canonical coordinates**. Notice that ω is closed because $\omega = \mathrm{d}\sum_i p^i \mathrm{d}q^i$. When N copies of ω are wedged together, the result is essentially the standard measure of phase-space volume:

$$\omega\wedge\omega\wedge\cdots\wedge\omega = N!\,\mathrm{d}p^1\wedge\mathrm{d}q^1\wedge\mathrm{d}p^2\wedge\cdots\wedge\mathrm{d}q^N.$$

Note:

Canonical coordinates may be constructed for *any* symplectic form such that any given field X coincides with $(\partial/\partial p^1)$. The construction is closely analogous to Gram-Schmidt orthogonalization.

We recover Hamilton's equations by determining the action of $\mathcal{H}$ on canonical coordinates:

$$\begin{aligned} \frac{\partial H}{\partial p^i} &= \mathrm{d}H\Big(\frac{\partial}{\partial p^i}\Big) = \Big(\sum_k \mathrm{d}p^k\wedge\mathrm{d}q^k\Big)\Big(\frac{\partial}{\partial p^i}, \mathcal{H}\Big) \\ &= \mathrm{d}q^i(\mathcal{H}) = \mathcal{H}(q^i) = \dot{q}^i \\ \frac{\partial H}{\partial q^i} &= \mathrm{d}H\Big(\frac{\partial}{\partial q^i}\Big) = \Big(\sum_k \mathrm{d}p^k\wedge\mathrm{d}q^k\Big)\Big(\frac{\partial}{\partial q^i}, \mathcal{H}\Big) \\ &= -\mathrm{d}p^i(\mathcal{H}) = -\mathcal{H}(p^i) = -\dot{p}^i. \end{aligned} \tag{2.8}$$

More generally this calculation shows that in canonical coordinates the relation between a function f and its vector field $\mathcal{F}$ is

$$\mathcal{F} = \sum_i^N \Big(\mathcal{F}(p^i)\frac{\partial}{\partial p^i} + \mathcal{F}(q^i)\frac{\partial}{\partial q^i}\Big) \text{ where } \begin{cases} \mathcal{F}(p^i) = -\dfrac{\partial f}{\partial q^i} \\ \mathcal{F}(q^i) = +\dfrac{\partial f}{\partial p^i}. \end{cases} \tag{2.9}$$

To find the canonical form of the Poisson bracket we calculate[3]

$$\begin{aligned} \{g, f\} \equiv \omega(\mathcal{F}, \mathcal{G}) &= \sum_k \mathrm{d}p^k\wedge\mathrm{d}q^k\Big(\sum_i \frac{\partial f}{\partial p^i}\frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial}{\partial p^i}, \sum_j \frac{\partial g}{\partial p^j}\frac{\partial}{\partial q^j} - \frac{\partial g}{\partial q^j}\frac{\partial}{\partial p^j}\Big) \\ &= \sum_k \Big(\frac{\partial g}{\partial q^k}\frac{\partial f}{\partial p^k} - \frac{\partial g}{\partial p^k}\frac{\partial f}{\partial q^k}\Big). \end{aligned} \tag{2.10}$$

For each function f the operator $\{, f\}$ is manifestly a vector field, and comparison of (2.10) with (2.9) shows that it is precisely the field $\mathcal{F}$ that the symplectic structure associates with f.

[3] Notice the reversal of the order of f, g in the definition of $\{g, f\}$.

Theorem 1

Poisson brackets satisfy the **Jacobi identity**

$$\{\{f,g\},h\}+\{\{g,h\},f\}+\{\{h,f\},g\}=0. \tag{2.11}$$

Proof: Every term in the sum on the l.h.s. of (2.11) contains a second derivative of one of f, g, h. Any 2nd derivatives of h must arise from the partial sum $\{\{g,h\},f\}+\{\{h,f\},g\}$. By the nature of the operator $\{,a\}$, this partial sum equals $[\mathcal{G},\mathcal{F}]h$. But since the commutator of any two fields is itself a field, $[\mathcal{G},\mathcal{F}]h$ can contain only first derivatives of h. By symmetry it follows that the l.h.s. of (2.11) can contain no 2nd derivatives of any function, and must vanish.

Exercise (10):

Show that $[\mathcal{F},\mathcal{G}]$ is the vector field associated with $\{f,g\}$.

Exercise (11):

An important tool in studies on integrability and chaos in systems with $N=2$ is a **surface of section** (**SoS**). This is a plot of (p^1,q^1) at points on an integral curve of $\mathcal{H}$ at which $q^2=0$ and $p^2>0$. Given a point in the SoS and a value E of H we can solve $H(\mathbf{p},\mathbf{q})=E$ for p^2, and then (numerically?) follow a trajectory until it next returns to the SoS. For a fixed value of E this procedure defines the **Poincaré return map** of the SoS onto itself. Show that this map is area-preserving. [Hint: first show that (P,Q,H,t) provide canonical coordinates for phase space, where P,Q are the coords at which the integral curve of $\mathcal{H}$ through a given point last hit the SoS and t is the time elapse between this hit and the given point.]

2.2 Canonical transformations

Suppose $\{P^i,Q^i\}$ is another set of canonical coordinates. Then $\omega=\mathrm{d}\sum_k P^k\mathrm{d}Q^k=\mathrm{d}\sum_k p^k\mathrm{d}q^k$ so (locally)

$$\sum_k(P^k\mathrm{d}Q^k-p^k\mathrm{d}q^k)=\mathrm{d}S \tag{2.12}$$

for some function S. If we take as coordinates of this function $S(\mathbf{Q},\mathbf{q})$, we have

$$\mathrm{d}S=\sum_k\left(\frac{\partial S}{\partial Q^k}\mathrm{d}Q^k+\frac{\partial S}{\partial q^k}\mathrm{d}q^k\right). \tag{2.13}$$

Equating coefficients of $\mathrm{d}Q^k$ and $\mathrm{d}q^k$ between (2.12) and (2.13) we conclude that

$$P^k=\frac{\partial S}{\partial Q^k},\qquad p^k=-\frac{\partial S}{\partial q^k}. \tag{2.14}$$

One says that S has **generated** the canonical transformation $\{\mathbf{p},\mathbf{q}\}\to\{\mathbf{P},\mathbf{Q}\}$ through (2.14)

Exercise (12):

By writing $\omega=-\mathrm{d}\sum_k Q^k\mathrm{d}P^k$ obtain

$$Q^k=\frac{\partial S}{\partial P^k}\qquad p^k=\frac{\partial S}{\partial q^k} \tag{2.15}$$

for a generating function of the form $S(\mathbf{P},\mathbf{q})$

Here there has been time to only sparsely sample the geometrical aspects of Hamiltonian mechanics. Much more on this topic can be found in V.I. Arnold's beautiful book *Mathematical Methods of Classical Mechanics* (Springer).

3 Fibre bundles

Much of physics takes place on a **fibre bundle**. To construct one of these, E, we attach a manifold F (the fibre) to each point on the **base space** M. So to specify $e \in E$ we specify $m \in M$ and $f \in F$ and say $e = (m, f)$. Usually the fibre carries more structure than M. For example it might be a vector space as in the **tangent bundle**, in which an element $t \in \mathcal{T}_m$ is associated with $m \in M$, and in the analogously defined **cotangent bundle**. Another example of a fibre bundle is provided by the quantum field of a spin-0 particle: at each point m the value of the field is a complex number $\psi \in \mathcal{C}$. In certain circumstances (a 'broken' symmetry) $|\psi|$ will everywhere take essentially the same value, and the fibre is for most purposes the circle S^1. Broken symmetries of fields that lie in spaces of higher dimension make F into still more non-trivial spaces.

A chart for E obviously contains charts for both M and F. Let $e = (m, f)$ lie in the intersection of the domains U_1, U_2 of two charts, ϕ_1 and ϕ_2. Then $\phi_1^{-1} \circ \phi_2$ is map $F \to \mathcal{R}^n \to F$ of F onto itself. By choosing charts with sensibly aligned coordinate systems we can ensure that this map is the identity and $\phi_1(f) = \phi_2(f)$. But can we choose coordinates for all fibres such that $\phi_i(f) = \phi_j(f)$ simultaneously for every pair of charts i, j at all points in $U_i \cap U_j$? If this can be done, the bundle is 'trivial' and is *globally* just the tensor product $M \otimes F$. The interesting case is when we cannot avoid mismatching coordinates somewhere. Then the smallest possible set of transformations $\phi_i^{-1} \circ \phi_j$ of $F \to F$ is a non-trivial group, the **structure group** of the bundle.

The classic example of a non-trivial bundle is for M a circle and F a line. This bundle can be represented by a strip of paper whose ends have been glued together. If the band forms a cylinder, the bundle is trivial, while if it forms a Möbius strip the bundle is non-trivial and the structure group is $\{1, -1\}$.

3.1 Connections

We assume that it makes sense to add points on a fibre F. For example, in the case of a spin-0 particle, a point on the fibre is a complex number $\psi(m)$, and these can be added. A 'connection' is a rule that specifies how we are to difference points that lie on *different* fibres, as is essential if we are to evaluate the gradient of ψ. We first define a connection for the tangent bundle.

A **connection** D for M assigns to each pair of vector fields X, Y on M a third vector field $D_X(Y)$ such that

- $D_X(Y + Z) = D_X(Y) + D_X(Z)$;
- $D_{X+Y}(Z) = D_X(Z) + D_Y(Z)$;
- $D_{fX}(Y) = f D_X(Y)$;
- $D_X(fY) = X(f)Y + f D_X(Y)$.

(3.1)

Here $f \in \mathcal{F}_m$. According to this definition D is fully linear in its subscript and is a derivation in its argument. Heuristically, $D_X Y$ is the rate of change of Y as you move along the integral curve of X.

Application of D to $f \in \mathcal{F}$ times a constant field shows that it is expedient to extend the operation of D to functions (rank $\binom{0}{0}$ tensors) by the rule

$$D_X f \equiv X(f). \tag{3.2}$$

We extend the action of D to $\binom{0}{k}$ tensors by the following rule:

$$\big(D_X T\big)(Y_1, \ldots, Y_k) \equiv X\big(T(Y_1, \ldots, Y_k)\big) - \sum_{i=1}^{k} T(Y_1, \ldots, D_X Y_i, \ldots, Y_k). \tag{3.3}$$

For example, for 1-forms ω and 2-forms ρ we have

$$\begin{aligned} \big(D_X \omega\big)(Y) &= X\big(\omega(Y)\big) - \omega(D_X Y) \\ \big(D_X \rho\big)(Y, Z) &= X\big(\rho(Y, Z)\big) - \rho(D_X Y, Z) - \rho(Y, D_X Z). \end{aligned} \tag{3.4}$$

Heuristically, $D_X T$ evaluated on the Y_i is the part of X acting on the scalar $T(Y_1, \ldots)$ that is due to the change in T rather than changes in the Y_i as one moves along an integral curve of X.

Now that we have defined how D acts on forms, it is straightforward to define the action of D on an arbitrary $\binom{j}{k}$ tensor T in precise analogy with (3.3).

If we are given a curve $C(c)$ on M through m and $Y_m \in \mathcal{T}_m$, then there is a unique field $Y(c)$ on C such that $Y(0) = Y_m$ and $D_{(\mathrm{d}/\mathrm{d}c)} Y = 0$. The values of Y for field are said to be the result of **parallel transporting** Y_m along C.

Proof: Let (ϕ, U) be a chart for $m \in U$. Then define functions $\Gamma^i_{jk} \in \mathcal{F}_m$ by

$$D_{(\partial/\partial\phi^k)}\left(\frac{\partial}{\partial\phi^j}\right) = \sum_i \Gamma^i_{jk} \frac{\partial}{\partial\phi^i}. \tag{3.5}$$

Then by the 4th rule above $Y = \sum_i Y^i(\partial/\partial\phi^i)$ has to satisfy

$$0 = D_{(\mathrm{d}/\mathrm{d}c)} Y = \sum_i \left(\frac{\mathrm{d}Y^i}{\mathrm{d}c} + \sum_j Y^j \frac{\mathrm{d}\phi^k}{\mathrm{d}c}\Gamma^i_{jk}\right)\frac{\partial}{\partial\phi^i}. \tag{3.6}$$

Thus we have to solve the linear first-order o.d.e.s

$$\frac{\mathrm{d}Y^i}{\mathrm{d}c} + \sum_{jk} Y^j \frac{\mathrm{d}\phi^k}{\mathrm{d}c}\Gamma^i_{jk} = 0. \tag{3.7}$$

These equations have a unique solution.

Γ^i_{jk} is called a **Christoffel symbol**. For fixed k the matrix $\delta^i_j + \delta\phi^k\Gamma^i_{jk}$ gives the amount by which the frame of the basis fields rotates as one moves $\delta\phi^k$ along the k-axis.

The definition of D above for the tangent bundle generalizes easily to other fibre bundles. We recognize that in general we have two sorts of vector fields, namely ones that lie in the tangent bundle of the base manifold and ones that lie in the fibre bundle under consideration. D's subscript field X specifies a direction, so it lies in the tangent bundle, while D's argument field Y lies in the fibre bundle. Suppose $e_1, \ldots, e_p$ are basis fields for the fibres. Then

$$X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial\phi^i} \qquad Y = \sum_{i=1}^{p} Y^i e_i \tag{3.8}$$

and the Christoffel symbols are now defined by

$$D_{(\partial/\partial\phi^k)} e_j = \sum_i \Gamma^i_{jk} e_i. \tag{3.9}$$

Hence the second lower index of Γ runs from 1 to k while the other two go from 1 to p.

3.2 Electromagnetism as a gauge field

There is one very important case in which we can suppress the first two indices of Γ entirely. This is when the fibre is a part of the complex plane; we define $e_1 \equiv 1$, $e_2 \equiv \mathrm{i}$ so that $Y = \Re(Y)e_1 + \Im(Y)e_2$ and we extend the operation of X to a complex function such as Y by the obvious rule $X(Y) = X\big(\Re(Y)\big) + X\big(\Im(Y)\big)\mathrm{i}$. Now for each k the sum $\sum_{ij} Y^j\Gamma^i_{jk}e_i$ is a complex number that is proportional to Y, so it is plausible that we can write it as a product of the complex product of Y and some complex number Γ_k. An explicit calculation reveals when this is possible. Obviously,

$$\sum_{ij} Y^j\Gamma^i_{jk}e_i = Y\Gamma_k = \big(\Re(Y)\Re(\Gamma_k) - \Im(Y)\Im(\Gamma_k)\big)e_1 + \big(\Re(Y)\Im(\Gamma_k) + \Im(Y)\Re(\Gamma_k)\big)e_2 \tag{3.10}$$

so comparing coefficients of, for example, $\Re(Y)$ and e_1 we discover that we require

$$\begin{aligned} \Gamma^1_{1k} &= \Re(\Gamma_k) \quad \Gamma^1_{2k} = -\Im(\Gamma_k) \\ \Gamma^2_{1k} &= \Im(\Gamma_k) \quad \Gamma^2_{2k} = \Re(\Gamma_k) \end{aligned} \tag{3.11}$$

These restrictions on the Γ^i_{jk} reflect the fact that multiplication by a complex number (Γ_k) can only effect a conformal transformation of the complex plane.

When the Christoffel symbols satisfy the symmetry of equations (3.11), we have for a 1-dimensional complex fibre

$$D_X Y = X(Y) + \sum_{k=1}^{n} X^k Y \Gamma_k. \tag{3.12}$$

Let's write this in conventional notation with $\phi^i = x^i$, $Y = \psi$, $\Gamma_k = -\mathrm{i}(q/\hbar)A_k$ and multiply through by $-\mathrm{i}\hbar$ for fun:

$$-\mathrm{i}\hbar D_{(\partial/\partial x^i)}\psi = -\mathrm{i}\hbar\left(\frac{\partial\psi}{\partial x^i} - \mathrm{i}(q/\hbar)A_i\psi\right). \tag{3.13}$$

We recognize this as just the i^{th} component of the quantity which appears squared in the Hamiltonian of a charged particle. Thus, in this interpretation the electromagnetic potential A is the Christoffel symbol of a connection. The idea behind all gauge field theories, such as QED and QCD, is that interactions between fields arise because some fields ('gauge fields') furnish connections for the fibre bundles that are associated with particle fields. We shall see below that the field equations of a gauge field are determined by the symmetry group of its fibre.

3.3 Torsion and curvature

$D_X Y$ is not a tensor because it does not depend only on the value that Y takes at m. Two tensors are associated with it though, the **torsion** T and the **curvature** R:

$$\begin{aligned} T(X,Y) &\equiv D_X Y - D_Y X - [X,Y] \\ R(X,Y)Z &\equiv D_X D_Y Z - D_Y D_X Z - D_{[X,Y]}Z \end{aligned} \tag{3.14}$$

T is manifestly antisymmetric in its arguments and R is manifestly antisymmetric in its first two arguments. Clearly T is defined only for the tangent bundle, whilst R is defined for any fibre bundle.

Exercise (13):
Show that T is a tensor by showing that $T(fX,Y) = fT(X,Y)$ for $f \in \mathcal{F}_m$.

Exercise (14):
Show that $R(fX,Y)Z = fR(X,Y)Z$ for $f \in \mathcal{F}_m$.

Given the result of the last exercise, to show that R is a tensor we have only to show that for $f \in \mathcal{F}_m$, $R(X,Y)fZ = fR(X,Y)Z$:

$$\begin{aligned} R(X,Y)fZ &= D_X\big((Yf)Z + fD_Y Z\big) - D_Y\big((Xf)Z + fD_X Z\big) - ([X,Y]f)Z - fD_{[X,Y]}Z \\ &= (XYf)Z + (Yf)D_X Z + (Xf)D_Y Z + fD_X D_Y Z - (YXf)Z \\ &\quad - (Xf)D_Y Z - (Yf)D_X Z - fD_Y D_X Z - ([X,Y]f)Z - fD_{[X,Y]}Z \\ &= fR(X,Y)Z. \end{aligned} \tag{3.15}$$

3.4 Cartan structural equations

Since D is linear in its subscript, from (3.9) we have

$$D_X e_j = \sum_k X^k D_{(\partial/\partial\phi^k)} e_j = \sum_{ik} X^k \Gamma^i_{jk} e_i \\ = \sum_i w^i_j(X) e_i. \tag{3.16a}$$

where

$$w^i_j(X) \equiv \sum_k X^k \Gamma^i_{jk} \quad \Leftrightarrow \quad \Gamma^i_{jk} \equiv w^i_j\Big(\frac{\partial}{\partial\phi^k}\Big). \tag{3.16b}$$

We can recast this definition in a form that does not mention the Christoffel symbols by introducing the 1-forms w^i dual to the basis vectors e^j:

$$w^i(e_j) = \delta^i_j \quad \Rightarrow \quad w^i_j(X) \equiv w^i\big(D_X e_j\big). \tag{3.17}$$

In terms of w^i_j, $D_X Y$ is

$$D_X Y = \sum_i \Big(X(Y^i) + \sum_j Y^j w^i_j(X)\Big) e_i. \tag{3.18}$$

We now calculate curvature tensor in terms of the **Cartan forms** w^i and w^i_j. We define a set of real-valued 2-forms by

$$R(X,Y)e_j = \sum_{i=1}^p R^i_j(X,Y) e_i \tag{3.19}$$

and substitute this equation into the definition of R along with $X = \sum X^i e_i$ etc. We find the **second Cartan structure equation**:

$$R^i_j = \mathrm{d}w^i_j + \sum_k w^i_k \wedge w^k_j. \tag{3.20}$$

This equation provides an alternative demonstration that R is a tensor by showing that it is soundly constructed from forms.

Exercise (15):

Define a set of 2-forms by $T(X,Y) = \sum_{i=1}^p T^i(X,Y)e_i$. By substituting this expression for T into (3.14), obtain the **first Cartan structure equation**:

$$T^i = \mathrm{d}w^i + \sum_j w^i_j \wedge w^j. \tag{3.21}$$

With (3.16b) it follows immediately from (3.20) that

$$R^i_{jab} \equiv R^i_j\Big(\frac{\partial}{\partial\phi^a}, \frac{\partial}{\partial\phi^b}\Big) \\ = \frac{\partial}{\partial\phi^a} w^i_j\Big(\frac{\partial}{\partial\phi^b}\Big) - \frac{\partial}{\partial\phi^b} w^i_j\Big(\frac{\partial}{\partial\phi^a}\Big) + \sum_k \Big\{ w^i_k\Big(\frac{\partial}{\partial\phi^a}\Big) w^k_j\Big(\frac{\partial}{\partial\phi^b}\Big) - w^i_k\Big(\frac{\partial}{\partial\phi^b}\Big) w^k_j\Big(\frac{\partial}{\partial\phi^a}\Big)\Big\} \\ = \frac{\partial\Gamma^i_{jb}}{\partial\phi^a} - \frac{\partial\Gamma^i_{ja}}{\partial\phi^b} + \sum_k \Big\{\Gamma^i_{ka}\Gamma^k_{jb} - \Gamma^i_{kb}\Gamma^k_{ja}\Big\}. \tag{3.22}$$

Taking the exterior derivative of (3.20) yields an important identity:

$$\mathrm{d}R^i_j = \sum_k \Big(\mathrm{d}w^i_k \wedge w^k_j - w^i_k \wedge \mathrm{d}w^k_j\Big) \\ = \sum_k \Big(\big(R^i_k - \sum_{k'} w^i_{k'} \wedge w^{k'}_k\big) \wedge w^k_j - w^i_k \wedge \big(R^k_j - \sum_{k'} w^k_{k'} \wedge w^{k'}_j\big)\Big) \\ = \sum_k \Big(R^i_k \wedge w^k_j - w^i_k \wedge R^k_j\Big) \tag{3.23}$$

This is the **Bianchi identity**.

Exercise (16):
Show from the Bianchi identity that

$$R^i_{jkl;m} + R^i_{jlm;k} + R^i_{jmk;l} = 0 \tag{3.24}$$

where $A_{;m} \equiv D_{(\partial/\partial\phi^m)}A$. [Hint: the calculation is horrendous unless carried out in a frame in which $w^i_j = 0$ at m. Argue that each term in (3.24) is a tensor and can be identified from its value in any given frame.]

Example 3.1
In electromagnetism $\Gamma_k = -\mathrm{i}(q/\hbar)A_k$, so equations (3.11) tell us that for real A_k the only non-zero Γ^i_{jk} are $\Gamma^1_{2k} = -\Gamma^2_{1k} = (q/\hbar)A_k$. Also $e_1 = 1$, $e_2 = \mathrm{i}$ so $w^1 = \Re$ and $w^2 = \Im$. Hence the only non-zero w^i_j are $w^1_2(\partial/\partial\phi^k) = -w^2_1(\partial/\partial\phi^k) = (q/\hbar)A_k$. It follows that $w^i_k \wedge w^k_j = 0$ so that $R^i_j = \mathrm{d}w^i_j$. There are two non-zero curvature forms, $R^1_2 = -R^2_1 = (q/\hbar)\mathrm{d}A$. Specifically,

$$R^1_2\Big(\frac{\partial}{\partial\phi^j}, \frac{\partial}{\partial\phi^k}\Big) = (q/\hbar)\Big(\frac{\partial A_k}{\partial\phi^j} - \frac{\partial A_j}{\partial\phi^k}\Big). \tag{3.25}$$

We recognize R^1_2 as $q/\hbar$ times the Maxwell field tensor F_{jk}. In this case the Bianchi identity reduces to $\mathrm{d}F = 0$, which represents four of Maxwell's equations.

The geometrical significance of R is revealed by calculating the change in a vector Y_m when it is parallel transported around a closed curve $C(c)$. By (3.7) we have at each point on the curve

$$\frac{\mathrm{d}Y^i}{\mathrm{d}c} = -\sum_{jb} Y^j \frac{\mathrm{d}\phi^b}{\mathrm{d}c}\Gamma^i_{jb}. \tag{3.26}$$

Consequently, the total change in each component Y^i on going around is

$$\Delta Y^i = -\oint \sum_{jb} \mathrm{d}\phi^b\, Y^j \Gamma^i_{jb}. \tag{3.27}$$

In this integral both Γ^i_{jb} and Y^j are functions of $m' \equiv C(c)$. However, if we consider only infinitesimal loops we may expand each component of Γ and Y in power series about m:

$$\begin{aligned} \Gamma^i_{jb}(m') &= \Gamma^i_{jb}(m) + \sum_a (\phi^a(m') - \phi^a(m)) \frac{\partial\Gamma^i_{jb}}{\partial\phi^a}\bigg|_m + \cdots \\ Y^j(m') &= Y^j(m) + \sum_a (\phi^a(m') - \phi^a(m)) \frac{\partial Y^j}{\partial\phi^a}\bigg|_m + \cdots \end{aligned} \tag{3.28}$$

Multiplying these two expansions together and substituting the result into (3.27), we get

$$\Delta Y^i = -\oint \sum_{jab} \mathrm{d}\phi^b(m') \Big\{ \big(\Gamma^i_{jb} Y^j\big)_m + \bigg(\Gamma^i_{jb}\frac{\partial Y^j}{\partial\phi^a} + Y^j \frac{\partial\Gamma^i_{jb}}{\partial\phi^a}\bigg)_m (\phi^a(m') - \phi^a(m)) + \cdots \Big\}. \tag{3.29}$$

Since the first term in Γ is constant, it can be taken outside the integral sign. Summing its coefficient $\mathrm{d}\phi^b(m')$ around our closed curve we then obtain zero. Now we note that (3.26) implies that

$$\sum_a \frac{\mathrm{d}\phi^a}{\mathrm{d}c}\bigg(\frac{\partial Y^j}{\partial\phi^a} + \sum_k Y^k \Gamma^j_{ka}\bigg) = 0. \tag{3.30}$$

Integrating this expression around C we find

$$\sum_a \frac{\partial Y^j}{\partial\phi^a}\bigg|_m (\phi^a(m') - \phi^a(m)) = -\sum_{ak}\Big(Y^k\Gamma^j_{ka}\Big)_m (\phi^a(m') - \phi^a(m)) + \mathrm{O}(c^2), \tag{3.31}$$

so we may eliminate $(\partial Y^j/\partial\phi^a)$ from (3.29). Then we have

$$\begin{aligned}\Delta Y^i &= -\sum_{ab}\left(\frac{\partial\Gamma^i_{jb}}{\partial\phi^a}Y^j - \sum_{jk}\Gamma^i_{jb}\Gamma^j_{ka}Y^k\right)_m \oint \mathrm{d}\phi^b(m')\left(\phi^a(m')-\phi^a(m)\right)+\cdots\\ &= -\sum_{ab}\left(\frac{\partial\Gamma^i_{kb}}{\partial\phi^a} - \sum_{jk}\Gamma^i_{jb}\Gamma^j_{ka}\right)_m Y^k \oint \mathrm{d}\phi^b(m')\,\phi^a(m')+\cdots\end{aligned}\tag{3.32}$$

A plot of the (ϕ^a,ϕ^b) plane shows that the integral in (3.32) equals an antisymmetric object S^{ab} which, for $a\neq b$, is numerically $\pm$ the area inside the projection of our path onto the (ϕ^a,ϕ^b) plane. Since only the antisymmetric part of an object contributes to the contraction of that object with an antisymmetric object, we may write

$$\begin{aligned}\Delta Y^i &= \tfrac{1}{2}\sum_{kab}\left(\frac{\partial\Gamma^i_{kb}}{\partial\phi^a} - \frac{\partial\Gamma^i_{ka}}{\partial\phi^b} + \sum_j \Gamma^i_{ja}\Gamma^j_{kb} - \Gamma^i_{jb}\Gamma^j_{ka}\right)Y^k S^{ab}\\ &= \tfrac{1}{2}\sum_{kab} R^i_{kab}Y^k S^{ab}.\end{aligned}\tag{3.33}$$

Here the (3.22) has been used to obtain the second equality. Since the numbers ΔY^i are the differences in the coordinates of two vectors in $\mathcal{T}_m$ (Y before and after parallel transport around the loop), they are themselves the coordinates of a vector in $\mathcal{T}_m$. In fact, we can write (3.33) in coordinate-free form as

$$\Delta Y = \tfrac{1}{2}\int_S R()Y,\tag{3.34}$$

where S is a 2-surface that is bounded by $C()$ and we exploit the fact that $R()$ is a (matrix-valued) 2-form.

The physical idea underlying a gauge theory is that it takes energy to bend a dynamical field, so that fields try to run straight, i.e., to be parallel at neighbouring points. In the presence of non-zero curvature it is *impossible* for a field to run straight. Hence curvature should be associated with a non-zero energy density. We shall ask below how this might be quantified.

3.5 General gauge fields and Lie groups

Before we try to quantify the energy density of curvature, let's look at connections when a bundle's fibres support a representation of a Lie group[4] – as they do, for example, in high-energy applications.

In these applications each fibre is a p-dimensional vector space which supports an irreducible representation ('irrep') of some Lie group $\mathcal{G}$. For example, the Dirac electron field ψ is a 4-dimensional object which supports the $(\frac{1}{2},0)$ irrep of the Lorentz group. Similarly, each quark field ψ is a 3-dimensional object that supports the $D^{(1,0)}$ irrep of $SU(3)$ (flavour). By 'supporting an irrep' we mean that each element of $\mathcal{G}$ is associated with a linear transformation of the fibre into itself. When these transformations are expressed as matrices by the choice of a coordinate basis $\{e_j\}$ for the fibre, we find that these matrices have the characters (traces) that are peculiar to the given irrep.

We are obliged to specify points on a fibre by a set of p numbers – the coordinates of that point in a given basis. But any two coordinate systems that can be made out of each other by one of $\mathcal{G}$'s representing transformations are in some respects physically equivalent. Hence on neighbouring fibres we are likely to be using coordinate systems that are related by a non-trivial transformation – a generalized rotation. The job of the connection is to tell us if there is such a rotation, and, by compensating for it, to recover the real difference between the values of ψ at neighbouring points of the base manifold M.

We first examine how the connection works when there is no curvature, so a universal standard of alignment can be set up by parallel-transporting a particular frame $\{e_j\}$ from m to all fibres. We call

[4] My lecture notes on group theory are on my webpage, www-thphys.physics.ox.ac.uk/users/JamesBinney/

this parallel frame $\{e_j^{(p)}\}$ and denote by $\{e_j\}$ the frame used at a general point m'. We assume that the generalized rotation between the $\{e_j\}$ and $\{e_j^{(p)}\}$ frames tends to the identity as $m' \to m$. Near the identity transformation, all representing matrices are of the form $\mathbf{T} = \exp(-\mathrm{i}\sum_k \alpha^{(k)}\boldsymbol{\tau}_{(k)})$, where the $\alpha^{(k)}$ are real numbers, the $\boldsymbol{\tau}_{(k)}$ are matrices that represent the **infinitesimal generators** of $\mathcal{G}$, and the operator exp may be defined by the usual power series.[5] So

$$\begin{aligned} e_j\big|_{m'} &= \exp\Big(-\mathrm{i}\sum_k \alpha^{(k)}\boldsymbol{\tau}_{(k)}\Big)\cdot e_j^{(p)} \\ &\simeq \Big(\mathbf{I} - \mathrm{i}\sum_k \alpha^{(k)}\boldsymbol{\tau}_{(k)}\Big)\cdot e_j^{(p)}. \end{aligned} \tag{3.35}$$

We now let $m' \to m$ bearing in mind our assumption that in this limit $\alpha^{(k)} \to 0$. We find

$$w^i\big(D_X e_j\big) = w^i_j(X) = -\mathrm{i}\sum_k \mathrm{d}\alpha^{(k)}(X)\tau^i_{(k)j} \tag{3.36}$$

where $\tau^i_{(k)j} \equiv w^i(\boldsymbol{\tau}_{(k)}\cdot e_j)$ is the k^{th} generator written in the $\{e_i\}$ basis. From (3.18) the action of the connection can now be written

$$D_X Y = \sum_i \Big(X(Y^i) - \mathrm{i}\sum_{jk} Y^j \mathrm{d}\alpha^{(k)}(X)\tau^i_{(k)j}\Big)e_i. \tag{3.37}$$

Equation (3.37) gives the form of the connection when there is no curvature. In the general case the 1-forms $\mathrm{d}\alpha^{(k)}$ in (3.37) are replaced by inexact 1-forms $A^{(k)}$, so

$$\begin{aligned} w^i_j(X) &= -\mathrm{i}\sum_k A^{(k)}(X)\tau^i_{(k)j} \\ D_X Y &= \sum_i \Big(X(Y^i) - \mathrm{i}\sum_{jk} Y^j A^{(k)}(X)\tau^i_{(k)j}\Big)e_i. \end{aligned} \tag{3.38}$$

The 1-forms $A^{(k)}$ are generally considered by be a set of n_g vector **gauge potentials** labelled by k. When X is a basis field, (3.38) reads

$$D_{(\partial/\partial\phi^a)}Y = \sum_i \Big(\frac{\partial Y^i}{\partial\phi^a} - \mathrm{i}\sum_{jk} Y^j A_a^{(k)}\tau^i_{(k)j}\Big)e_i, \tag{3.39a}$$

where

$$A_a^{(k)} \equiv A^{(k)}\Big(\frac{\partial}{\partial\phi^a}\Big) \qquad (a = 1,\ldots,n;\ k = 1,\ldots,n_g). \tag{3.39b}$$

Gauge transformations

Let $\{e_i\}$ and $\{e'_i\}$ be two sets of basis fields that are related by a position-dependent gauge transformation

$$e'_i = \sum_j T_i{}^j e_j, \tag{3.40}$$

where $\mathbf{T}(g)$ is one of the representing matrices of the fibre's group $\mathcal{G}$. We often need to be able to express the Cartan form w'^i_j of the $\{e'_i\}$ system in terms of w^i_j. Before we begin the calculation proper, we establish some basic relations. Let

$$Y = \sum_i Y^i e_i = \sum_j Y'^j e'_j. \tag{3.41}$$

[5] The number of infinitesimal generators is equal to the dimension n_g of $\mathcal{G}$: for $SU(2)$ it is 3; for $SU(3)$ it is 8 etc. The matrices, $\boldsymbol{\tau}$, used here are complex but Hermitian objects.

Then immediately

$$Y^k = \sum_j Y'^j T_j{}^k \quad \Rightarrow \quad Y'^l = \sum_k Y^k t_k{}^l, \quad \text{where} \quad \sum_i T_j{}^i t_i{}^k = \delta^k_j. \tag{3.42}$$

Since $\mathbf{t}$ is the right inverse of $\mathbf{T}$ and every right inverse is also a left inverse, we have $\sum_k t_i{}^k T_k{}^j = \delta^j_i$. Using this on (3.40) we find

$$e_j = \sum_i t_j{}^i e'_i. \tag{3.43}$$

Now we calculate

$$\begin{aligned}
\sum_m w'^m_j(X) e'_m &= D_X(e'_j) = \sum_{jk} D_X(T_j{}^k e_k)) \\
&= \sum_k \big(X(T_j{}^k) e_k + T_j{}^k D_X(e_k)\big) \\
&= \sum_k \big(X(T_j{}^k) e_k + T_j{}^k \sum_l w^l_k(X) e_l\big) \\
&= \sum_{km} \big(X(T_j{}^k) t_k{}^m + T_j{}^k \sum_l w^l_k(X) t_l{}^m\big) e'_m.
\end{aligned} \tag{3.44}$$

When we now use (3.38) to eliminate the Cartan forms we find that the gauge fields transform thus

$$\sum_n A'^{(n)}(X) \tau^m_{(n)j} = \sum_k \big(\mathrm{i}X(T_j{}^k) t_k{}^m + \sum_{ln} A^{(n)}(X) T_j{}^k \tau^l_{(n)k} t_l{}^m\big). \tag{3.45a}$$

Bearing in mind that the $\boldsymbol{\tau}$ are skew-symmetric, in matrix notation (3.45a) reads

$$\sum_n A'^{(n)}(X) \boldsymbol{\tau}_{(n)} = -\mathrm{i}X(\mathbf{T}) \cdot \mathbf{T}^{-1} + \sum_n A^{(n)}(X) \mathbf{T} \cdot \boldsymbol{\tau}_{(n)} \cdot \mathbf{T}^{-1}. \tag{3.45b}$$

We can clean up equation (3.45b) by defining a matrix-valued 1-form

$$\mathbf{A} \equiv \sum_{k=1}^{n_g} A^{(k)} \boldsymbol{\tau}_{(k)}. \tag{3.46}$$

Then

$$\mathbf{A}'(X) = -\mathrm{i}X(\mathbf{T}) \cdot \mathbf{T}^{-1} + \mathbf{T} \cdot \mathbf{A}(X) \cdot \mathbf{T}^{-1}. \tag{3.47}$$

Only the second term in this transformation law is proportional to A. When the curvature R vanishes, this fact can be exploited to ensure that $\mathbf{A}' = 0$.

Inserting $\mathbf{T} = \exp\big(-\mathrm{i}\sum_k \alpha^{(k)} \boldsymbol{\tau}_{(k)}\big)$ into (3.47) and evaluating for small $\alpha^{(k)}$, we find

$$\delta\mathbf{A}(X) \equiv (\mathbf{A}' - \mathbf{A})(X) = -\sum_k X(\alpha^{(k)}) \boldsymbol{\tau}_{(k)} + \mathrm{i}\sum_{jk} \mathbf{A}^{(k)}(X) \alpha^{(j)} [\boldsymbol{\tau}_{(k)}, \boldsymbol{\tau}_{(j)}]. \tag{3.48}$$

Curvature tensor of a gauge theory

Since for a gauge theory the Cartan form w^i_j is just $-\mathrm{i}\sum_k \tau^i_{(k)j} A^{(k)}$, (3.20) gives the curvature as

$$\begin{aligned}
R^i_j &= -\mathrm{i}\sum_k \tau^i_{(k)j} \mathrm{d}A^{(k)} - \sum_{k'k} \sum_l \tau^i_{(k)l} \tau^l_{(k')j} A^{(k)} \wedge A^{(k')} \\
&= -\mathrm{i}\sum_k \tau^i_{(k)j} \mathrm{d}A^{(k)} - \tfrac{1}{2} \sum_{k'k} \big[\boldsymbol{\tau}_{(k)}, \boldsymbol{\tau}_{(k')}\big]^i{}_j A^{(k)} \wedge A^{(k')}.
\end{aligned} \tag{3.49}$$

Here the square bracket in the last line denotes the commutator of two matrices. We have that $[\boldsymbol{\tau}_{(i)}, \boldsymbol{\tau}_{(j)}] = \sum_k c^k_{ij} \boldsymbol{\tau}_{(k)}$, where the numbers c^k_{ij} are the **structure constants** of $\mathcal{G}$.

By evaluating the r.h.s. of (3.49) on two fields X, Y, it is straightforward to show that in our cleaned-up notation, (3.49) can be written

$$\mathbf{R} = -\mathrm{i}\mathrm{d}\mathbf{A} - [\mathbf{A}, \mathbf{A}], \tag{3.50}$$

where every symbol is a matrix-valued form and $[\mathbf{A}, \mathbf{A}](X, Y) \equiv [\mathbf{A}(X), \mathbf{A}(Y)]$ does not vanish because the first and second occurrences of $\mathbf{A}$ evaluate to different matrices.

We know that R^i_j transforms like a tensor – that is, its transformation law involves only $\mathbf{T}$ and not its derivatives. This property suits it to a rôle in the Lagrangian of a field theory. In electromagnetism the Lagrangian density, $\mathcal{L} = \mathrm{Tr}\mathbf{F}\cdot\mathbf{F}/4\mu_0$ is essentially the square of the only independent, non-vanishing element of R^i_j. However, to form this square we need a metric, which do not yet have.

3.6 Yang–Mills theory

The prototype non-Abelian gauge theory is that worked out in 1954 by Yang & Mills (Phys. Rev., **96**, 191). In this theory $\mathcal{G} = SU(2)$, so $n_g = 3$ so there are 3 generators $\boldsymbol{\tau}_{(k)}$ and they satisfy the commutation relations

$$[\boldsymbol{\tau}_{(i)}, \boldsymbol{\tau}_{(j)}] = \mathrm{i}\sum_k \epsilon_{ijk} \tau_{(k)}. \tag{3.51}$$

There are two important representations of this theory: (i) the spinor one, when the fibres are 2-dim complex spaces and $\boldsymbol{\tau}_{(k)} = \frac{1}{2}\boldsymbol{\sigma}_{(k)}$ are proportional to the Pauli matrices; (ii) the vector one, when the fibres are real 3-dim spaces and $\tau^i_{(k)j} = \mathrm{i}\epsilon_{ijk}$.

In the vector rep. we have

$$\begin{array}{lll}
\text{Cartan form} & w^i_j(X) & = \sum_k \epsilon_{ijk} A^{(k)}(X) \\
\text{Connection} & (D_X\psi)^i & = X(\psi^i) + \sum_{jk} \epsilon_{ijk}\psi^j A^{(k)}(X) \\
\text{Gauge transf.} & \delta\mathbf{A}(X) & = -\sum_k X(\alpha^{(k)})\boldsymbol{\tau}_{(k)} - \sum_{kij} \epsilon_{kij}\alpha^{(i)} A^{(j)}(X)\boldsymbol{\tau}_{(k)} \\
\text{Curvature} & \mathbf{R}(X, Y) & = -\mathrm{i}\sum_k \mathrm{d}\mathbf{A}(X, Y)\boldsymbol{\tau}_{(k)} - \mathbf{A}(X)\times\mathbf{A}(Y) \quad ,
\end{array} \tag{3.52}$$

where $\mathbf{A}(X)\times\mathbf{A}(Y) \equiv \sum_{klm} \epsilon_{klm} A^{(k)}(X) A^{(l)}(Y)\boldsymbol{\tau}_{(m)}$.

4 Riemannian spaces

A **metric** is a symmetric $\binom{0}{2}$ tensor g that is non-degenerate in the sense that $g(X,Y)=0\ \forall Y \Rightarrow X=0$. A manifold that is equipped with a metric is called a **Riemannian manifold** if $g(X,Y)\geq 0\ \forall\ X,Y$, and is otherwise called a **pseudo-Riemannian manifold**. Clearly, space-time is a pseudo-Riemannian manifold. None the less, I shall concentrate on Riemannian manifolds below because they have everything that a pseudo-Riemannian manifold has and more.

Exercise (17):

Show that the metric tensor of a Riemannian manifold satisfies the **Schwartz inequality**, $g(X,Y)^2 \leq g(X,X)g(Y,Y)$, and the **triangle inequality**, $g(X+Y,X+Y)^{1/2} \leq g(X,X)^{1/2} + g(Y,Y)^{1/2}$.

On a Riemannian manifold any curve $C(c)$ has a length

$$l(C) \equiv \int_{c_1}^{c_2} \mathrm{d}c\, g\Big(\frac{\mathrm{d}}{\mathrm{d}c},\frac{\mathrm{d}}{\mathrm{d}c}\Big)^{1/2}. \tag{4.1}$$

Thus possession of a metric makes it possible for the first time to specify distances between points – hence g's name.

g associates the 1-form $g(,X)$ with every vector X. With every basis vector $(\partial/\partial\phi^i)$ we have already associated the 1-form $\mathrm{d}\phi^i$ by the rule $\mathrm{d}\phi^i(\partial/\partial\phi^j)=\delta^i_j$. A basis for which $\mathrm{d}\phi^i$ coincides with $g(,\partial/\partial\phi^i)$ is called an **orthonormal basis**:

$$g\Big(\frac{\partial}{\partial\phi^j},\frac{\partial}{\partial\phi^i}\Big)=\delta_{ij} \quad \text{(orthonormal basis)}. \tag{4.2}$$

The coordinate representation of this association is as follows. We define numbers X_i by

$$\sum_i X_i \mathrm{d}\phi^i = g(,X)$$

and compute

$$X_i = g\Big(\frac{\partial}{\partial\phi^i},X\Big)=\sum_j g_{ij}X^j \quad \text{where} \quad g_{ij}\equiv g\Big(\frac{\partial}{\partial\phi^i},\frac{\partial}{\partial\phi^j}\Big). \tag{4.3}$$

Thus g_{ij} is an index lowering operator.

The non-degeneracy of g implies that the matrix g_{ij} has an inverse. Let this be denoted g^{ij}. Then multiplying both sides of equation (4.3) by g^{ij} we have

$$X^i=\sum_j g^{ij}X_j \qquad \Big(\sum_k g^{ik}g_{kj}=\delta^i_j\Big). \tag{4.4}$$

We can use g_{ij} or g^{ij} to raise or lower any index in the coordinate expansion of a tensor of any type. For example

$$T_{ij\ldots k} \;\to\; T^{ij\ldots k} \equiv \sum_{pq\ldots r} g^{ip}g^{jq}\times\cdots\times g^{kr}T_{pq\ldots r}, \tag{4.5}$$

or, equivalently

$$T(\mathrm{d}\phi^i,\ldots,\mathrm{d}\phi^k)=\sum_{pq\ldots r} g^{ip}\times\cdots\times g^{kr}T\Big(\frac{\partial}{\partial\phi^p},\ldots,\frac{\partial}{\partial\phi^r}\Big). \tag{4.6}$$

4.1 The Hodge * operator

On an orientable manifold[6] g gives rise to a natural volume n-form Ω as follows. Let $\{(\partial/\partial\phi^i)\}$ be an orthonormal basis. Then

$$\Omega \equiv \mathrm{d}\phi^1 \wedge \cdots \wedge \mathrm{d}\phi^n. \tag{4.7}$$

One proves that Ω is unique up to a sign by showing that if $\{(\partial/\partial\psi^i)\}$ is a second orthonormal basis, then

$$\Omega \equiv \pm\mathrm{d}\psi^1 \wedge \cdots \wedge \mathrm{d}\psi^n. \tag{4.8}$$

Indeed, if

$$\frac{\partial}{\partial\phi^i} = \sum_j \Lambda_i^j \frac{\partial}{\partial\psi^j} \tag{4.9}$$

then

$$\begin{aligned}\delta_{ij} = g\Big(\frac{\partial}{\partial\phi^i}, \frac{\partial}{\partial\phi^j}\Big) &= \sum_{kl} \Lambda_i^k \Lambda_j^l g\Big(\frac{\partial}{\partial\psi^k}, \frac{\partial}{\partial\psi^l}\Big) \\ = \sum_{kl} \Lambda_i^k \Lambda_j^l \delta_{kl} &= \sum_k \Lambda_i^k \Lambda_j^k.\end{aligned} \tag{4.10}$$

So $\mathbf{\Lambda}$ is an orthogonal matrix and has determinant $\det\mathbf{\Lambda} = \pm 1$. On the other hand, if we set $\Omega = \lambda\mathrm{d}\psi^1 \ldots \mathrm{d}\psi^n$ and evaluate λ by requiring that $\Omega(\partial/\partial\phi^1, \ldots, \partial/\partial\phi^n) = 1$, we find with (4.9) that

$$\Omega = (\det\mathbf{\Lambda})^{-1}\,\mathrm{d}\psi^1 \wedge \cdots \wedge \mathrm{d}\psi^n \tag{4.11}$$

and the required result follows.

Exercise (18):

Adapt the proof just given to show that with respect to an arbitrary set of basis fields $(\partial/\partial\phi^i)$

$$\Omega = \big|\det(g_{ij})\big|^{1/2}\,\mathrm{d}\phi^1 \wedge \cdots \wedge \mathrm{d}\phi^n. \tag{4.12}$$

Exercise (19):

From the foregoing exercise show that $\Omega_{ij\ldots k} = \big|\det(g_{ab})\big|^{1/2}\epsilon_{ij\ldots k}$, where ϵ is the usual Levi-Civita symbol. [Note that $\epsilon_{ij\ldots k} = \epsilon^{ij\ldots k}$.]

In §1.4 we developed the theory of forms – totally antisymmetric $\binom{0}{p}$ tensors. By (4.6) g associates with every p-form a totally antisymmetric $\binom{p}{0}$ tensor. A volume form Ω establishes a 1-1 mapping between such tensors and $(n-p)$-forms. Indeed, let T be an antisymmetric $\binom{p}{0}$ tensor. Then with it we associate

$$\begin{aligned}\tau\Big(\frac{\partial}{\partial\phi^{p+1}}, \ldots, \frac{\partial}{\partial\phi^n}\Big) &= \Omega\Big(\frac{\partial}{\partial\phi^1}, \ldots, \frac{\partial}{\partial\phi^n}\Big)T(\mathrm{d}\phi^1, \ldots, \mathrm{d}\phi^p) \\ &\equiv \Omega\Big(T, \frac{\partial}{\partial\phi^{p+1}}, \ldots, \frac{\partial}{\partial\phi^n}\Big).\end{aligned} \tag{4.13}$$

That is, the $(n-p)$-form τ associated with T is what's left after you've stuffed the first p slots of Ω with T. The sequence of operations just described, namely the use of g to turn a p-form ω into a $\binom{p}{0}$ tensor and then the use of Ω to turn the latter into an $(n-p)$-form $*\omega$ is symbolized by the **Hodge * operator**. Substituting into (4.13) from (4.6) we have

$$\begin{aligned}*\omega\Big(\frac{\partial}{\partial\phi^{p+1}}, \ldots, \frac{\partial}{\partial\phi^n}\Big) =&\,\Omega\Big(\frac{\partial}{\partial\phi^1}, \ldots, \frac{\partial}{\partial\phi^n}\Big) \\ &\times \sum_{q\ldots r} g^{1q} \times \cdots \times g^{pr}\omega\Big(\frac{\partial}{\partial\phi^q}, \ldots, \frac{\partial}{\partial\phi^r}\Big)\end{aligned} \tag{4.14}$$

[6] An **orientable manifold** M is one on which one there exists an n-form that is everywhere non-zero.

The mapping accomplished by $*$ is possible because, as we saw in §1.4, there are exactly as many linearly independent $(n-p)$-forms as there are linearly independent p-forms.

In a general coordinate system the Hodge $*$ is a mess. It is best applied in an orthonormal coordinate system when $\Omega_{i,\dots,k} = \epsilon_{i,\dots,k}$ and $\mathbf{g}$ is the identity matrix. Then with

$$\omega = \sum_{i_1<\cdots<i_p} \omega_{i_1,\dots,i_p} \mathrm{d}\phi^{i_1} \wedge \cdots \wedge \mathrm{d}\phi^{i_p}, \tag{4.15}$$

$*\omega$ can be written down by replacing each ordered wedge product by an ordered wedge product of the remaining $\mathrm{d}\phi^i$. For example, in 4 dimensions $\mathrm{d}\phi^2 \wedge \mathrm{d}\phi^4$ is replaced by $-\mathrm{d}\phi^1 \wedge \mathrm{d}\phi^3$.

Theorem 2

For any two p-forms we have

$$*\phi \wedge \psi = *\psi \wedge \phi. \tag{4.16}$$

Proof: Since $*$ and $\wedge$ are linear operators, it suffices to show the result for ϕ and ψ monomials $\mathrm{d}\phi^1 \wedge \cdots \wedge \mathrm{d}\phi^p$. Both sides of (4.16) then vanish unless they are identical monomials and are trivially equal when they are the same monomial. □

Theorem 3

$*$ is, to within a sign, its own inverse:

$$**\omega = (-1)^{p(n-p)}\omega \quad p\text{-form } \omega \tag{4.17}$$

Proof: Since $*$ is linear, it again suffices to demonstrate the result for an arbitrary monomial $\omega = \mathrm{d}\phi^1 \wedge \cdots \wedge \mathrm{d}\phi^p$, where we have labelled our orthonormal coordinates to suit ω. Then

$$*\omega = \mathrm{d}\phi^{p+1} \wedge \cdots \wedge \mathrm{d}\phi^n \quad \Rightarrow \quad (*\omega)_{j_1\dots j_{n-p}} = \epsilon_{12\dots pj_1\dots j_{n-p}}. \tag{4.18}$$

Hence (4.14) yields,

$$**\omega_{i_1\dots i_p} = \epsilon_{i_{p+1}\dots i_n i_1\dots i_p}\epsilon_{12\dots p i_{p+1}\dots i_n}, \tag{4.19}$$

where $i_1 \dots i_n$ is an arbitrary permutation of the numbers $1\dots n$. Clearly the right side vanishes unless $i_1 \dots i_p$ between them take the values $1\dots p$, and is equal to ± 1 when they do take these values. To obtain the sign we move each of the $(n-p)$ first indices on the first ϵ to its proper place in the list, starting with i_n and picking up p minuses for each index. We finally conclude that $**\omega_1 = (-1)^{p(n-p)}\omega$ as required. □

Exercise (20):

Show that in 3d Euclidean space, with A and B 1-forms

$$\begin{aligned} A \times B &= *(A \wedge B) \\ \nabla \times A &= *\mathrm{d}A \\ \nabla \cdot A &= \mathrm{d} * A. \end{aligned} \tag{4.20}$$

Exercise (21):

Show that four of Maxwell's equations are contained in

$$*\mathrm{d} * F = \mu_0 j, \tag{4.21}$$

where j is the 1-form of the current density.

Each of the operators $*\mathrm{d}*\mathrm{d}$ and $\mathrm{d}*\mathrm{d}*$ carries a p-form into a p-form. It is straightforward to show that when applied to $f \in \mathcal{F}_m$, $*\mathrm{d}*\mathrm{d}$ yields $\nabla^2 f$, while $\mathrm{d}*\mathrm{d}*f$ yields zero. This motivates the definition of the **Laplace–Beltrami operator**

$$\triangle \equiv *\mathrm{d}*\mathrm{d} + (-1)^n \mathrm{d}*\mathrm{d}*. \tag{4.22}$$

$\triangle$ turns one p-form into another, and when used on 0-forms reduces to the Laplacian. A form that satisfies $\triangle A = 0$ is called a **harmonic form**. The theory of de Rahm cohomology – see below – shows that the existence of certain harmonic forms has remarkable implications for the global topology of M.

Exercise (22):

Show that in the Lorentz gauge ($\mathrm{d} * A = 0$) half of Maxwell's equations become $\triangle A = \mu_0 j$.

4.2 De Rham Cohomology

In Euclidean space closed forms are exact: $\mathrm{d}\omega = 0 \;\Rightarrow\; \omega = \mathrm{d}\psi$. On a topologically non-trivial manifold there are inexact closed forms and there proves to be a close relation between the number of such forms and the number of harmonic forms. This relation is surprising because d does not depend on the connection, whilst $\triangle$ does. To establish the relation we start by defining an inner product between any two p-forms ϕ, ψ:

$$(\phi, \psi) \equiv \int_M \phi \wedge *\psi \tag{4.23}$$

This equation makes sense because $*\psi$ is an $(n-p)$-form so $\phi \wedge *\psi$ is an n-form. We assume either that M is closed or that all forms vanish on its boundary.

By (4.16) the inner product (4.23) is symmetric: $(\phi, \psi) = (\psi, \phi)$.

By writing (4.23) in an orthonormal basis, we see that $(\phi, \phi) = 0$ iff $\phi = 0$.

Next we find the adjoint of d under the inner product. With ϕ a $(p-1)$-form, ψ a p-form and (1.37) we have:

$$\begin{aligned}(\mathrm{d}\phi, \psi) &= \int \mathrm{d}\phi \wedge *\psi \\ &= \int \mathrm{d}(\phi \wedge *\psi) - (-1)^{p-1} \int \phi \wedge \mathrm{d} * \psi \\ &= (-1)^p \int \phi \wedge (k * *)\mathrm{d} * \psi \\ &= (-1)^p k(\phi, *\mathrm{d} * \psi).\end{aligned} \tag{4.24}$$

From (4.17) $k = (-1)^{[n-(p-1)](p-1)} = (-1)^{n(p-1)-(p-1)}$, so the numerical prefactor in (4.24) is $(-1)^{np-n+1} = (-1)^{np+n+1}$, and

$$\delta \equiv (-1)^{np+n+1} * \mathrm{d}* \tag{4.25}$$

is the adjoint of d. With this definition (4.22) yields

$$\triangle = k(\delta\mathrm{d} + \mathrm{d}\delta), \qquad (k = \pm 1) \tag{4.26}$$

Theorem 4

A p-form ω is harmonic iff

$$\mathrm{d}\omega = \delta\omega = 0. \tag{4.27}$$

Proof: The if part is trivial given (4.26). Further

$$\triangle\omega = 0 \quad \Rightarrow \quad \begin{aligned} 0 &= (\omega, \triangle\omega) = (\omega, \delta\mathrm{d}\omega) + (\omega, \mathrm{d}\delta\omega) \\ &= (\mathrm{d}\omega, \mathrm{d}\omega) + (\delta\omega, \delta\omega), \end{aligned} \tag{4.28}$$

where we have used the mutual adjointness of d and δ. The result now follows because, as we have seen, each inner product is non-negative and a vanishing inner product implies that its contents vanish. □

Theorem 5

If ω is closed but not exact, then there exists ψ such that $\omega - \mathrm{d}\psi$ is harmonic.

Proof: Consider $I \equiv (\omega - \mathrm{d}\psi, \omega - \mathrm{d}\psi)$, which is always positive. Hence it must have a minimum as a functional of ψ. Let this minimum occur at ψ_0 and let η be any small variation around ψ_0. Then

$$\begin{aligned} 0 &= (\mathrm{d}\eta, \omega - \mathrm{d}\psi_0) + (\omega - \mathrm{d}\psi_0, \mathrm{d}\eta) \\ &= 2(\eta, \delta[\omega - \mathrm{d}\psi_0]) \\ &\Rightarrow \quad \delta[\omega - \mathrm{d}\psi_0] = 0 \end{aligned} \tag{4.29}$$

The result now follows from the last theorem because $\mathrm{d}[\omega - \mathrm{d}\psi_0]$ vanishes trivially. □

Theorem 6

The difference between two distinct harmonic p-forms is never exact.

Proof: Let ψ, ϕ be two harmonic p-forms. Then $\omega \equiv \psi - \phi$ is also harmonic, so $\delta\omega = 0$. Suppose $\omega = \mathrm{d}\gamma$. Then

$$0 = (\delta\omega, \gamma) = (\omega, \mathrm{d}\gamma) = (\omega, \omega)$$
$$\Rightarrow \quad \omega = 0$$

and $\psi = \phi$ contrary to conjecture. □

We now divide closed but inexact p-forms into **cohomology** classes such that the difference between any two forms in a given class is exact; this is a consistent proceeding because if $\omega_1 - \omega_2 = \mathrm{d}\phi_a$ and $\omega_2 - \omega_3 = \mathrm{d}\phi_b$, then $\omega_1 - \omega_3 = \mathrm{d}(\phi_a + \phi_b)$. The p^{th} **Betti number** b^p of M is the number of cohomology classes of p-forms.

Theorem 5 assures us that there is at least one harmonic form in each cohomology class by stating that for any ω we have that $\omega_0 \equiv \omega - \mathrm{d}\psi$ is harmonic. Theorem 6 states that there is only *one* harmonic form per cohomolgy class. Hence, there are precisely as many harmonic forms as cohomolgy classes.

The cohomology classes depend on d but not on the metric, so they are invariant if we change the metric whilst leaving the analytic structure alone. Physically changing the metric amounts to distorting the space in a continuous way without changing "the way it is joined up".

We can play a game with submanifolds and boundaries that is closely connected with the game we have been playing with closed forms and d. We focus on p-dimensional submanifolds C that have no boundary: $\partial C = 0$. Such objects are called **cycles**. For $p = 1$ we are speaking of closed curves; for $p = 2$ closed surfaces, etc. Sometimes the difference between two p-cycles that are not themselves boundaries of $p + 1$ manifolds will form such a boundary, sometimes not. (Consider closed curves on a torus.) If the difference is a boundary ($C_1 - C_2 = \partial S$), then we place C_1 and C_2 in the same **homology** class. The number of homology classes of p-cycles is denoted b_p. It turns out that $b_p = b^p$, so you can tell how many harmonic p-forms M has by counting the number of inequivalent p-cycles it admits. For example, there are no harmonic 1-forms on a 2-sphere because every closed curve on a 2-sphere is a boundary; there are no 'inexact' curves.

4.3 Riemannian connection

A metric induces a natural connection on the tangent bundle of M. Specifically, we seek a connection that has zero torsion and yields a vanishing derivative of g in the sense of the 2nd equation below:

$$\begin{aligned} 0 &= D_X Y - D_Y X - [X, Y], \\ Zg(X, Y) &= g(D_Z X, Y) + g(X, D_Z Y). \end{aligned} \tag{4.30}$$

To see that these equations suffice to define D, we write out the last equation twice more, cyclically permuting X, Y, Z as we go:

$$\begin{aligned} Xg(Y, Z) &= g(D_X Y, Z) + g(Y, D_X Z); \\ Yg(Z, X) &= g(D_Y Z, X) + g(Z, D_Y X). \end{aligned} \tag{4.31}$$

Now we assume that X, Y, Z are coordinate fields so that they commute (with the consequence that $D_X Y = D_Y X$ etc.) and subtract the 2nd of equations (4.30) from the sum of equations (4.31). In view of the symmetry of g various terms cancel and we have

$$Xg(Y, Z) + Yg(X, Z) - Zg(X, Y) = 2g(D_X Y, Z) \qquad \text{(commuting fields).} \tag{4.32}$$

Equation (4.32) clearly defines the action of D since it enables us to find an arbitrary component of $D_X Y$ for any two coordinate fields X, Y. In terms of Christoffel symbols [eq. (3.9)] we have

$$\begin{aligned} \tfrac{1}{2}\Big(\frac{\partial g_{jk}}{\partial \phi^i} + \frac{\partial g_{ik}}{\partial \phi^j} - \frac{\partial g_{ij}}{\partial \phi^k}\Big) &= g\Big(\sum_l \Gamma^l_{ij} \frac{\partial}{\partial \phi^l}, \frac{\partial}{\partial \phi^k}\Big) \\ &= \sum_l g_{kl} \Gamma^l_{ij}, \end{aligned} \tag{4.33}$$

which clearly implies that

$$\Gamma^l_{ij} = \tfrac{1}{2}\sum_k g^{lk}\Big(\frac{\partial g_{jk}}{\partial\phi^i} + \frac{\partial g_{ik}}{\partial\phi^j} - \frac{\partial g_{ij}}{\partial\phi^k}\Big). \tag{4.34}$$

Exercise (23):

Since $l(C) \geq 0$ there will be a shortest curve between two points. Show that this curve is a **geodesic**, i.e., $D_{(\mathrm{d}/\mathrm{d}c)}\dfrac{\mathrm{d}}{\mathrm{d}c} = 0$.

Suppose we have a bundle of nearby geodesics: in general relativity (GR) they could be the world-lines of a mass of balls that Gallileo has just dropped from the Leaning Tower of Pisa. Let V and v be the tangent vector and affine parameter for a typical geodesic.

Now imagine a curve that joins points associated with a given value of v. Let the local tangent vector to this curve be U, with u the corresponding affine parameter, so $U = \partial/\partial u$. By filling the whole region with curves joining events of equal v, we can extend U into a vector field and use u and v as 2 of the coordinates of M. Hence $[U, V] = 0$ and (4.30) yields $D_U V = D_V U$.

In the Gallileo example, the magnitude of U tells us how rapidly we move from ball to ball as we vary u, and thus is proportional to the distance between balls. If the directon of U changes, it's because the balls are moving around one another as they fall, so to get between two given balls, the direction you have to head off in is evolving in time (v). So there's interesting physical information in the "acceleration" $\partial^2 U/\partial v^2$, and we calculate it.

$$\begin{aligned}\frac{\partial^2 U}{\partial v^2} &= D_V D_V U = D_V D_U V = D_U D_V V + R(V,U)V \\ &= R(V,U)V.\end{aligned} \tag{4.35}$$

This is the **equation of geodesic deviation**.

When we change from the $\{\phi^a\}$ chart to the $\{\phi'^a\}$ chart, the natural basis vectors for $\mathcal{T}_m$ suffer a linear transformation

$$\frac{\partial}{\partial\phi'^a} = B^b_a\frac{\partial}{\partial\phi^b} \quad \text{where} \quad B^b_a \equiv \frac{\partial\phi^b}{\partial\phi'^a}. \tag{4.36}$$

We use the shorthand $D_a \equiv D_{\partial/\partial\phi^a}$ and $D'_a \equiv D_{\partial/\partial\phi'^a}$ and calculate

$$\begin{aligned}\Gamma'_{ab,c} \equiv g(D'_a\partial/\partial\phi'^b, \partial/\partial\phi'^c) &= g\left(B^f_a D_f(B^g_b\partial/\partial\phi^g), B^h_c\partial/\partial\phi^h\right) \\ &= g\left(B^f_a D_f(B^g_b)\partial/\partial\phi^g, B^h_c\partial/\partial\phi^h\right) + g\left(B^f_a B^g_b D_f\partial/\partial\phi^g, B^h_c\partial/\partial\phi^h\right) \\ &= B^f_a D_f(B^g_b)B^h_c g(\partial/\partial\phi^g, \partial/\partial\phi^h) + B^f_a B^g_b B^h_c g(D_f\partial/\partial\phi^g, \partial/\partial\phi^h) \\ &= B^f_a\frac{\partial B^g_b}{\partial\phi^f}B^h_c g(\partial/\partial\phi^g, \partial/\partial\phi^h) + B^f_a B^g_b B^h_c\Gamma_{fg,h}.\end{aligned} \tag{4.37}$$

This expression shows that in addition to the usual tensor transformation rule, Γ requires an additive term

$$B^f_a\frac{\partial B^g_b}{\partial\phi^f}B^h_c g(\partial/\partial\phi^g, \partial/\partial\phi^h) = \frac{\partial B^g_b}{\partial\phi'^a}B^h_c g(\partial/\partial\phi^g, \partial/\partial\phi^h) = \frac{\partial^2\phi^g}{\partial\phi'^a\partial\phi'^b}B^h_c g(\partial/\partial\phi^g, \partial/\partial\phi^h) \tag{4.38}$$

We can make $\Gamma' = 0$ if we can make this term cancel $B^f_a B^g_b B^h_c\Gamma_{fg,h}$ for any given Γ. To see that this is possible, define

$$A^a_b \equiv \frac{\partial\phi'^a}{\partial\phi^b} \quad \Rightarrow \quad A^a_b B^b_c = \delta^a_c \quad \text{and} \quad A^a_b B^c_a = \delta^c_b, \tag{4.39}$$

and multiply (4.37) through by $g^{lk}A^c_k$, where g^{lg} is the inverse of the metric tensor $[g^{lg}g(\partial/\partial\phi^g, \partial/\partial\phi^h) = \delta^l_h]$ and find that Γ' will vanish if

$$\frac{\partial^2\phi^l}{\partial\phi'^a\partial\phi'^b} = -B^f_a B^g_b g^{lk}\Gamma_{fg,k} = -B^f_a B^g_b\Gamma^l_{fg}. \tag{4.40}$$

From the Taylor series expansion of $\phi^l(\phi')$ we see that at any given point the double derivative matrix on the left can be set equal to any symmetric matrix we please, so we can set it equal to the rhs. Thus for any given point of M there are charts in which $\Gamma' = 0$. In GR such a chart is said to provide **local inertial coordinates**. The existence of these coordinates encapsulates the "principle of equivalence" that is the starting point for GR.

Since $0 = D'_a g'_{bc} = \partial'_a g'_{bc} - 2\Gamma'^d_{ab} g'_{dc}$, if $\Gamma' = 0$ all partial derivatives of g' also vanish.