Topics in Mathematical Physics

Ben Karsberg

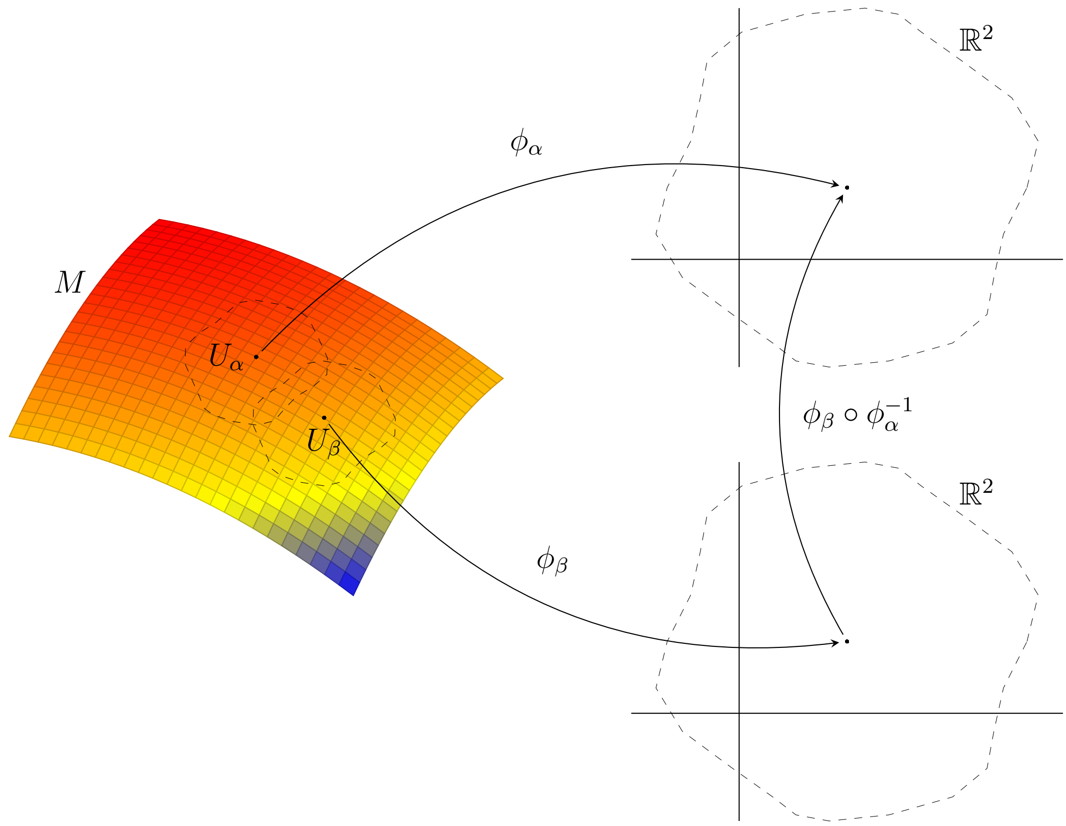
2021-22

1 Preface

- These notes were written up for the postgraduate masters course in mathematical physics at the University of Edinburgh
- The topic this year was 'Applications of Geometry in Physics'
- It was lectured by James Lucietti, who's notes are unfortunately only available on the Edinburgh intranet
- These notes were largely produced with those official notes as reference
- These notes were written primarily with myself in mind; it has some colloquialisms in it, and some topics are presented in the way I best understand them, and not necessarily the 'best' way
- The notes are also exclusively bullet pointed because I find prose difficult to digest
- Any errors are most certainly mine and not James's - let me know at benkarsberg@gmail.com if you note any bad ones
- Generally, text in *italic* is the definition of something the first time it shows up, and text in **bold** is something I think is important/want to remember/found tricky when writing these up
- Summation convention is implicit throughout this entire set of notes, unless explicitly stated otherwise
- Euclidean spatial vectors are denoted by e.g. $\mathbf{v}$, and 4-vectors are denoted by just $v = (v^0, \mathbf{v})$
- The metric signature is $\eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1)$ unless stated otherwise
- $\hbar = c = 1$
- I hope these notes are helpful to someone who isn't me - enjoy!

2 Differential Geometry Preliminaries

- This course is all about geometry, almost entirely differential
- Differentiable geometry is a listed prerequisite for this course, but I haven't done any properly, with the exception of pseudo-Riemannian geometry in coordinate bases in a GR course last year
- Therefore, we need to set the scene
- The principal arena for differential geometry is a **differential manifold**:
  - Informally, this is a set M that locally looks like $\mathbb{R}^n$
  - Formally, this is a topological space M equipped with a collection of charts $\{(U_\alpha, \phi_\alpha)\}$ called an **atlas**, where $\{U_\alpha\}$ is an open cover of M and $\phi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n$ are homeomorphisms
- The image we want in mind is something like this, generalised to $\mathbb{R}^n$ rather than just $\mathbb{R}^2$:



- Essentially, a chart (U, ϕ) assigns coordinates to a point $p \in U$ given by $\phi(p) = (x^1(p), \ldots, x^n(p))$, usually written $\phi = (x^i)$
- Importantly, the **transition functions** $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\beta(U_\alpha \cap U_\beta) \to \phi_\alpha(U_\alpha \cap U_\beta)$ should be $C^\infty(\mathbb{R}^n)$ bijections between open subsets of $\mathbb{R}^n$ (see diagram)
- Essentially, all changes of coordinates are given by a smooth map
- The atlas allows us to define calculus on a manifold, and we can classify a hierarchy of objects on M:

- $C^\infty(M)$ is the set of all smooth functions on M, with the function of a standard commutative algebra
- The **tangent space** T_pM at $p \in M$ is the n-dimensional vector space of directional derivatives at p; that is, the space of all $\mathbb{R}$-linear maps $X_p : C^\infty(M) \to \mathbb{R}$ obeying the Leibnitz/product rule

$$X_p(fg) = X_p(f)g(p) + f(p)X_p(g) \tag{2.1}$$

  for all $f, g \in C^\infty(M)$
- The elements $X_p \in T_pM$ are called **tangent vectors**; if we consider a smooth curve $\gamma : (a, a) \to M$ where $\gamma(0) = p$, we can define a tangent vector to the curve called the velocity $\dot{\gamma} \in T_pM$ by $\dot{\gamma}(f) = \frac{d}{dt} f(\gamma(t))\big|_{t=0}$, and all tangent vectors arise like this
- A **vector field** $X \in \mathfrak{X}(M)$ is a smooth assignment to each $p \in M$ of a tangent vector $X_p \in T_pM$ (or more formally, a section of the **tangent bundle** TM)
- A **differential** k**-form** α_p at $p \in M$ is a smooth alternating multilinear map $\alpha_p : T_pM \times \ldots \times T_pM \to \mathbb{R}$
- The set of all k-forms is denoted $\Omega^k(M)$, where $\alpha(X_1, \ldots, X_k) \in C^\infty(M)$ and $X_i \in \mathfrak{X}(M)$
- In particular, 1-forms at p are just dual vectors $\alpha_p \in T_p^*M$
- The **wedge produce** of $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^l(M)$ is $\alpha \wedge \beta \in \Omega^{k+l}(M)$ where $\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$; in particular, for 1-forms we have

$$(\alpha \wedge \beta)(X, Y) = \alpha(X)\beta(Y) - \beta(X)\alpha(Y) \tag{2.2}$$

- The **exterior derivative** is a $\mathbb{R}$-linear map $d : \Omega^k(M) \to \Omega^{k+1}(M)$ satisfying $d^2 = 0$ and $d(\alpha \wedge \beta = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge (d\beta)$
- The **interior derivative** $\iota_X : \Omega^k(M) \to \Omega^{k-1}(M)$ is defined by $\iota_X \alpha = \alpha(X, \ldots)$ for $X \in \mathfrak{X}(M)$
- In particular, if we regard a smooth function f on M as a 0-form, $(df)(X) = X(f)$
- **Tensors** and **tensor fields** of type (r, s) are very similar to vectors and vector fields, acting as multilinear maps $T : T_p^*M \times \ldots T_p^*M \times T_pM \times \ldots T_pM \to \mathbb{R}$
- The derivative of a smooth map $\phi : M \to N$ is called the **push-forward map** $\phi_* : T_pM \to T_{\phi(p)}N$ given by

$$(\phi_* X_p)(f) = X_p(f \circ \phi) \tag{2.3}$$

  where $f \in C^\infty(N)$
- Similarly, the **pull-back** is $\phi^* : T^*_{\phi(p)}N \to T_p^*M$ given by

$$(\phi^* \alpha_{\phi(p)})(X_p) = \alpha_{\phi(p)}(\phi_* X_p) \tag{2.4}$$

- A smooth bijection $\phi : M \to N$ is called a **diffeomorphism**, and for a diffeomorphism we have $(\phi^{-1})_* = \phi^*$
- The **flow** of a vector field X is a 1-parameter group of diffeomorphisms; i.e. a family $\phi_t : M \to M$ such that $\phi_{t+s} = \phi_t \circ \phi_s$ for all real t, s
- In particular, the integral curve of X through $p \in M$ is given by $\phi_t(p)$ so $X_p(f) = \frac{d}{dt} f(\phi_t(p))|_{t=0}$

- The **Lie derivative** of a tensor field T along a vector field X is

$$\mathcal{L}_X T = \frac{d}{dt}(\phi_t^* T)\bigg|_{t=0} \tag{2.5}$$

  where ϕ_t is the flow of $X \in \mathfrak{X}(M)$

- In particular, we have $\mathcal{L}_X f = X(f)$ for $f \in C^\infty(M)$; $\mathcal{L}_X Y = [X, Y]$ for $Y \in \mathfrak{X}(M)$; $\mathcal{L}_X \alpha = (d\iota_X + \iota_X d)\alpha$ for $\alpha \in \Omega^k(M)$

- Suppose we have a local basis $\{e_a \mid a = 1, \ldots, n\}$ of vectors; we can then expand a vector field as $X = X^a e_a$ where X^a are the components

- The dual basis is $\{f^a \mid a = 1, \ldots, n\}$ defined by $f^a(e_b) = \delta^a_b$; differential 1-forms can then be expanded as $\alpha = \alpha_a f^a$ where $\alpha_a = \alpha(e_a)$ are the components

- More generally, a k-form can be expanded as $\alpha = \frac{1}{k!}\alpha_{a_1 \ldots a_k} f^{a_1} \wedge \ldots \wedge f^{a_k}$ where $\alpha_{a_1 \ldots a_k} = \alpha(e_{a_1}, \ldots, e_{a_k})$

- This generalises to tensors in the expected way

- In particular, a chart (U, ϕ) where $\phi = x^i$ defines a basis of T_pM called the **coordinate basis**, $\left(\frac{\partial}{\partial x^i}\right)_p$

- In this basis, any vector field can be expanded as $X = X^i \frac{\partial}{\partial x^i}$ where $X^i = X(x^i)$ are smooth functions on U

- Similarly, the dual basis of T_p^*M is $(dx^i)_p$, where 1-forms can be expanded as $\alpha = \alpha_i dx^i$, and this can be expanded to a basis for k-forms by $dx^{i_1} \wedge \ldots \wedge dx^{i_k}$ where $i_1 < \ldots < i_k$

3 Symplectic Geometry and Classical Mechanics

- To motivate what follows, consider a classical Hamiltonian system with phase space $\mathbb{R}^{2n}$ and canonical coordinates $(\mathbf{p}, \mathbf{q})$
- Hamilton's equations for the time-evolution of the system are

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}, \quad \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} \tag{3.1}$$

where H is the Hamiltonian

- This can be rewritten in matrix form:

$$\begin{pmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{pmatrix} = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \begin{pmatrix} \dfrac{\partial H}{\partial \mathbf{q}} \\ \dfrac{\partial H}{\partial \mathbf{p}} \end{pmatrix} \iff \dot{\mathbf{x}} = \Omega^{-1} \nabla_{\mathbf{x}} H \tag{3.2}$$

where

$$\Omega = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} \tag{3.3}$$

is implicit

- As we shall see, this is the canonical example of a symplectic vector space

3.1 Symplectic Vector Spaces

- A **symplectic vector space** is a pair (V, ω), where V is a $2n$-dimensional real vector space, and $\omega : V \times V \to \mathbb{R}$ is a bilinear form satisfying:

  1. *Skew-symmetry*: $\omega(v, w) = -\omega(w, v)$, $\forall v, w \in V$
  2. *Non-degeneracy*: $\omega(v, w) = 0 \, \forall v \in V \implies w = 0$

- ω is called a **symplectic structure** on V
- Contrasting ω with the standard Euclidean inner product, the only real difference is skew vs non-skew symmetry
- In physics, we usually need a basis to work in - defining the arbitrary basis $\{e_a \,|\, a = 1, \ldots, 2n\}$, the components of ω are given by the usual $\omega_{ab} = \omega(e_a, e_b)$
- From this, it follows that ω_{ab} is an antisymmetric matrix
- Moreover, non-degeneracy of ω is equivalent to

$$\omega_{ab} w^b \implies w^a = 0 \tag{3.4}$$

which just means ω_{ab} is invertible

- **Example:** the canonical example of a symplectic vector space is $V = \mathbb{R}^{2n}$ and $\omega(\mathbf{v}, \mathbf{w}) = \mathbf{v}^T \Omega \mathbf{w}$, with Ω as defined in the introduction:

$$\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \tag{3.5}$$

- We now show that, in fact, all symplectic vector spaces are isomorphic to this example

- **Theorem:** any symplectic vector space (V, ω) has a **symplectic basis** $\{e_i, f_i \,|\, i = 1, \dots, n\}$ defined by

$$\omega(e_i, e_j) = \omega(f_i, f_j) = 0 \quad \text{and} \quad \omega(e_i, f_j) = \delta_{ij} \tag{3.6}$$

- **Proof:** by induction
  - For the $n = 1$ case, pick an arbitrary vector $0 \neq e_1 \in V$
  - By non-degeneracy, $\exists f_1 \in V$ such that $\omega(e_1, f_1) \neq 0$, and we can rescale so that $\omega(e_1, f_1) = 1$
  - Then, by skew-symmetry, we see $\omega(e_1, e_1) = \omega(f_1, f_1) = 0$ as required
  - Assume the hypothesis holds for $n = k \geq 1$, and consider the $n = k + 1$ case
  - By the same procedure that we used for the base case, we can construct e_1 and f_1
  - Then, define the *skew-orthogonal complement*:

$$W = \{w \in V \,|\, \omega(w, e_1) = \omega(w, f_1) = 0\} \tag{3.7}$$

  - W has dimension $2k$, and we now wish to show that it is a symplectic vector space
  - First note that any $v \in V$ can be uniquely decomposed as $v = u + w$, where $u \in \text{span}(e_1, f_1)$ and $w \in W$
  - Then consider $\omega|_W$, i.e. the symplectic structure restricted to W; pick $\tilde{w} \in W$, then:

$$\begin{aligned} \omega(\tilde{w}, w) = 0,\ \forall w \in W &\implies \omega(\tilde{w}, v) = 0,\ \forall v \in V \qquad &&\text{(by decomposition above)} \\ &\implies \tilde{w} = 0 &&\text{(by non-degeneracy)} \end{aligned} \tag{3.8}$$

  - $\omega|_W$ clearly inherits skew-symmetry from ω, so it is a symplectic form, and so W is indeed a symplectic vector space
  - By induction the conclusion follows ■

- Explicitly, this means in the symplectic basis $\omega_{ab} = \Omega_{ab}$, establishing the isomorphism

- The next result shows a radical departure from Euclidean geometry in these spaces

- **Theorem:** let (V, ω) be a $2n$-dimensional symplectic vector space, and $W \subset V$ a *null subspace* so $\omega|_W = 0$. Then $\dim W \leq n$

- **Proof:**
  - By the earlier theorem, we can assume $V = \mathbb{R}^{2n}$ and $\omega = \Omega$
  - Then, W is a null space iff $W \perp \Omega W$, where this is with respect to the Euclidean norm $\mathbf{w} \cdot \mathbf{v} = \mathbf{w}^T \mathbf{v}$
  - Therefore, we have

$$\dim W + \dim \Omega W \leq 2n \tag{3.9}$$

  - However, Ω is invertible so $\dim W = \dim \Omega W$, so we are done ■

- We now want to rewrite the symplectic structure in a symplectic basis, but in a more useful way

- Define the dual symplectic basis $\{e^i_*, f^i_* \,|\, i = 1, \dots, n\}$ of V^*, so specifically:

$$e^i_*(e_j) = f^i_*(f_j) = \delta^i_j, \quad e^i_*(f_j) = f^i_*(e_j) = 0 \tag{3.10}$$

- Then, the basis of 2-forms is given by the set

$$\{e_*^i \wedge e_*^j,\, e_*^i \wedge f_*^j,\, f_*^i \wedge f_*^j\} \tag{3.11}$$

as usual

- We therefore claim that

$$\omega = e_*^i \wedge f_*^i \tag{3.12}$$

with summation convention implied

- We can easily verify this by its action on basis elements:

$$\omega(e_a, e_b) = (e_*^i \wedge f_*^i)(e_a, e_b) = e_*^i(e_a) f_*^i(f_b) - e_*^i(f_b) f_*^i(e_a) = \delta_a^i \delta_b^i = \delta_{ab} \tag{3.13}$$

- More generally, consider the k-fold wedge product ω^k
- We have $\omega^k = 0$ for $k > n$ as ω^n is the highest ranked form on V, which has dimension $2n$
- **Lemma:** let (V, ω) be a $2n$-dimensional symplectic vector space. Then, in a symplectic basis we have

$$\omega^n = n! e_*^1 \wedge f_*^1 \wedge \ldots \wedge e_*^n \wedge f_*^n \tag{3.14}$$

- **Proof:** by induction (not given, but it's not hard - future Ben, just try and do it yourself for revision xx)

3.2 Symplectic Manifolds and the Cotangent Bundle

- A **symplectic manifold** is a pair (M, ω) where M is a 2-dimensional smooth manifold, and $\omega \in \Omega^2(M)$ is a 2-form satisfying:
  1. $d\omega = 0$ (closed)
  2. $\omega(X, Y) = 0 \; \forall X \in \mathfrak{X}(M) \implies Y = 0$ (non-degenerate)
- ω is called a **symplectic form**
- Note that if (M, ω) is a symplectic manifold, then $(T_x M, \omega_x)$ is a symplectic vector space at all points $x \in M$
- **Example:** let $M = \mathbb{R}^{2n}$ with coordinates (p_i, q_i), $i = 1, \ldots, n$, and $\omega = dp_i \wedge dq_i$ (summation implied)
  - $\omega = d(p_i dq_i)$, so it is exact and hence closed
  - Consider the components of ω in the coordinate basis $(\partial/\partial p_i, \partial/\partial q_i)$; we find

$$\begin{aligned} \omega\left(\frac{\partial}{\partial p_j}, \frac{\partial}{\partial p_k}\right) &= \omega\left(\frac{\partial}{\partial q_j}, \frac{\partial}{\partial q_k}\right) = 0 \\ \omega\left(\frac{\partial}{\partial p_j}, \frac{\partial}{\partial q_k}\right) &= dp_i\left(\frac{\partial}{\partial p_j}\right) dq_i\left(\frac{\partial}{\partial q_k}\right) = \delta_{ij}\delta_{ik} = \delta_{jk} \end{aligned} \tag{3.15}$$

    so in particular, $\omega_{ij} = \Omega_{ij}$ in this basis and is non-degenerate
  - In particular, this means ω is a symplectic form and that $(\partial/\partial p_i, \partial/\partial q_i)$ is a symplectic basis at every point
- **Theorem:** any symplectic manifold (M, ω) of dimension $2n$ is orientable with volume form ω^n

- **Proof:**
  - A manifold is **orientable** if there is a consistent definition of clockwise and anti-clockwise (e.g. a Mobius strip would be non-orientable), and a **volume form** is a highest-rank differentiable form on the manifold (of rank equal to the dimension of the manifold)
  - One characterisation of orientability is that a manifold is orientable if and only if there exists a nowhere vanishing top-ranked differential form/volume form
  - Therefore, if we can show $\omega^n = \omega \wedge \ldots \wedge \omega$ is nowhere vanishing, we are done
  - At each $x \in M$, $(T_x M, \omega_x)$ is a symplectic vector space
  - But we know that in symplectic vector spaces

    $$\omega_x^n = n! e_*^1 \wedge f_*^1 \wedge \ldots \wedge e_*^n \wedge f_*^n \tag{3.16}$$

    which is nowhere vanishing ■

- **Example:** consider the manifold given by the unit 2-sphere $M = S^2 = \{\mathbf{x} \in \mathbb{R}^3 \,|\, |\mathbf{x}| = 1\}$
  - The tangent space at $\mathbf{x}$ is $T_{\mathbf{x}} S^2 = span(\mathbf{x})^\perp$
  - The volume form at this point is given by

    $$\omega_{\mathbf{x}}(\mathbf{V}, \mathbf{w}) = \mathbf{x} \cdot (\mathbf{v} \times \mathbf{w}) \tag{3.17}$$

    for $\mathbf{v}, \mathbf{w} \in T_{\mathbf{x}} S^2$

3.2.1 The Cotangent Bundle

- The following discussion is why symplectic manifolds arise in classical mechanics
- Suppose Q is a $2n$-dimensional smooth manifold; its **cotangent bundle** is defined by

  $$T^*Q = \{(p_x, x) \,|\, p_x \in T_x^* Q,\ x \in Q\} \tag{3.18}$$

  which we think of as the set of all points x paired with all their covectors (one-forms)
  - The cotangent bundle is a $2n$-dimensional manifold
  - To show this, consider the chart on Q given by $\phi(x) = (q_i(x))$, containing $x \in Q$
  - We can see that $\Phi(p_x, x) = (p_i(x), q_i(x))$ is then a chart on T^*Q, where p_i are the components of $p_x = p_i(x)(dq_i)_x$
  - To check the transition functions are smooth, consider the projection $\pi : T^*Q \to Q,\ (p_x, x) \to x$; this is smooth and surjective, and the inverse projection has image $\pi^{-1}(x) = T_x^* Q$
  - Therefore, $\phi \circ \pi \circ \Phi^{-1}(p_i(x), q_i(x)) = (q_i(x))$ is smooth
- **Theorem:** T^*Q has a symplectic form ω, given in the above chart by

  $$\omega = dp_i \wedge dq_i \tag{3.19}$$

- **Proof:**
  - Recall that the derivative of a smooth map $\phi : M \to N$ is the push-forward map $\phi_* : T_p M \to T_{\phi(p)} N$, given by

    $$(\phi_* X_p)(f) = X_p(f\phi) \tag{3.20}$$

– Moreover, the pull-back of ϕ is $\phi^* : T^*_{\phi(p)}N \to T^*_p M$, given by

$$(\phi^* \alpha_{\phi(p)})(X_p) = \alpha_{\phi(p)}(\phi_* X_p) \tag{3.21}$$

– The projector above is a smooth map of this form, so its derivative is the push-forward $\pi_* : T_{(p_x,x)}(T^*Q) \to T_x Q$

– We now define $\theta_{(p_x,x)} \in T^*_{(p_x,x)}(T^*Q)$, given by

$$\theta_{(p_x,x)}(\xi) = p_x(\pi_* \xi), \quad \forall \xi \in T_{(p_x,x)}(T^*Q) \tag{3.22}$$

so $\theta_{(p_x,x)} = \pi^* p_x$ is the pull-back of p_x under π, meaning $\theta \in \Omega^1(T^*Q)$ is a one-form on T^*Q

– Define $\omega = d\theta \in \Omega^2(T^*Q)$, which is clearly closed since it is exact

– To show non-degeneracy, we check the components of θ in the chart $\Phi = (p_i, q_i)$ above of T^*Q, defined by chart $\phi = (q_i)$ of Q

– We find first that

$$\theta\left(\frac{\partial}{\partial q_i}\right) = p\left(\pi_* \frac{\partial}{\partial q_i}\right) \tag{3.23}$$

and note in particular that

$$\begin{aligned} \left(\pi_* \frac{\partial}{\partial q_i}\right)(f) &= \frac{\partial}{\partial q_i}(f \circ \pi) \\ &= \frac{d}{dt} f \circ \pi \circ \Phi^{-1}(p_1, \ldots, p_n, q_1, \ldots, q_i + t, \ldots, q_n)|_{t=0} \\ &= \frac{d}{dt} f \circ \phi^{-1}(q_1, \ldots, q_i + t, \ldots, q_n)|_{t=0} \\ &= \frac{\partial}{\partial q_i}(f) \end{aligned} \tag{3.24}$$

for all $f \in C^\infty(Q)$

– Therefore, $\theta(\partial/\partial q_i) = p(\partial/\partial q_i) = p_i$

– A similar calculation finds $\pi_* \partial/\partial p_i = 0$, so $\theta(\partial/\partial p_i) = 0$

– Therefore, $\theta = p_i dq_i$ in this chart ■

- This has a very clear link to classical mechanics: Q is the configuration space, and T^*Q is the phase space

- A symplectic manifold also induces a rather natural isomorphism between the tangent and cotangent space

- Specifically, this isomorphism is between the tangent space at each point and its dual

- **Theorem:** let (V, ω) be a symplectic vector space; then there is a 'natural' isomorphism called the musical isomorphism $V \equiv V^*$

- **Proof:**

  – Let $\xi \in V$; we define the linear map $\flat : V \to V^*$ by

$$\xi^\flat(\eta) = \omega(\eta, \xi), \quad \forall \eta \in V \tag{3.25}$$

  – $\flat$ is injective since ω is non-degenerate

- Instead, let $\lambda \in V^*$; we define the linear map $\sharp : V^* \to V$ by

$$\omega(\eta, \lambda^\sharp) = \lambda(\eta), \quad \forall \eta \in V \tag{3.26}$$

- This defines $\lambda^\sharp$ uniquely by non-degeneracy of ω, and it is again injective
- These two maps are mutual inverses:

$$\begin{aligned} \omega(\eta, (\xi^\flat)^\sharp) = \xi^\flat(\eta) = \omega(\eta, \xi), \quad \forall \eta \in V &\implies (\xi^\flat)^\sharp = \xi, \ \forall \xi \in V \\ (\lambda^\sharp)^\flat(\eta) = \omega(\eta, \lambda^\sharp) = \lambda(\eta), \quad \forall \eta \in V &\implies (\lambda^\sharp)^\flat = \lambda, \ \forall \lambda \in V^* \ \blacksquare \end{aligned} \tag{3.27}$$

- Moreover, the musical isomorphism defines a (2,0) tensor over V:

$$\omega^\sharp(\eta, \lambda) = -\omega(\eta^\sharp, \lambda^\sharp) \tag{3.28}$$

for all $\eta, \lambda \in V^*$

- This is called the **inverse symplectic form**
- Choosing a basis $\{e_a\}$ of V, we find:

$$(\xi^\flat)_a = \omega_{ab}\xi^b, \quad \text{and} \quad \omega_{ab}(\lambda^\sharp)^b = \lambda_a \implies (\lambda^\sharp)^a = \omega^{ab}\lambda_b \tag{3.29}$$

- This is, in a sense, raising and lowering indices using the symplectic form
- This makes sense as to why $\omega^\sharp$ is the inverse symplectic form:

$$(\omega^\sharp)^{ab}\eta_a\lambda_b = -\omega_{ab}(\eta^\sharp)^a(\lambda^\sharp)^b = -\omega_{ab}\omega^{ac}\omega^{bd}\eta_c\lambda_d = \delta^c_b\omega^{bd}\eta_c\lambda_d = \omega^{cd}\eta_c\lambda_d \tag{3.30}$$

- The analogous construction in pseudo-Riemannian geometry (c.f. GR) is defined by the metric tensor: an inner product on V
- This means that for any symplectic manifold (M, ω), there is a musical isomorphism $T_xM \equiv T_x^*M$ for every $x \in M$, which induces a musical isomorphism between vector fields and 1-forms on M

3.3 Hamiltonian Vector Fields and Flows

- Given arbitrary **Hamiltonian function** $H \in C^\infty(M)$, the vector field $X_H = (dH)^\sharp \in \mathfrak{X}(M)$ is the **Hamiltonian vector field**
- The set of all such vector fields is denoted $Ham(M)$
- Equivalently, by the musical isomorphism, we have

$$X_H^\flat = dH \implies X_H^\flat(\eta) = \omega(\eta, X_H) = -\omega(X_H, \eta) = -(\iota_{X_H}\omega)(\eta) \tag{3.31}$$

- Therefore, X_H is Hamiltonian if and only if

$$\iota_{X_H}\omega = -dH \tag{3.32}$$

- Just as a reminder, ι_X denotes the interior derivative of a k-form with respect to X, defined as contraction with X in the first argument
- In a coordinate basis $\{\partial_a : a = 1, \dots, 2n\}$, we find components

$$(\iota_{X_H}\omega)_a = \omega_{ab}X_H^a = -\omega_{ab}X_H^b = (-dH)_a = -\partial_a H \implies \omega_{ab}X_H^b = \partial_a H \tag{3.33}$$

- Equivalently:
$$X_H^a = \omega^{ab}\partial_b H \tag{3.34}$$

- **Example:** consider symplectic manifold $(\mathbb{R}^{2n}, \omega)$ with coordinates (p_i, q_i) and symplectic form $\omega = dp_i \wedge dq_i$
  - Pick some $H \in C^\infty(\mathbb{R}^{2n})$, and we compute the integral curves $\gamma(t)$ of X_H
  - Recall integral curves are defined by
$$\dot{\gamma}(t) = X_H(\gamma(t)) = X_H|_{\gamma(t)} \tag{3.35}$$
  - In our defining chart (p_i, q_i), we have
$$\dot{\gamma}(t) = \dot{p}_i \frac{\partial}{\partial p_i} + \dot{q}_i \frac{\partial}{\partial q_i} \tag{3.36}$$
    along the curve, where we slightly abuse notation to write $\dot{p}_i = \frac{d}{dt} p_i(\gamma(t))$ etc.
  - Therefore, along the curve we must have
$$\iota_{X_H}\omega|_{\gamma(t)} = (dp_i \wedge dq_i)(X_H(\gamma(t))) = dp_i(\dot{\gamma})dq_i - dq_i(\dot{\gamma})dp_i = \dot{p}_i dq_i - \dot{q}_i dp_i \tag{3.37}$$
  - We also have by definition:
$$dH = \frac{dH}{dp_i}dp_i + \frac{dH}{dq_i}dq_i \tag{3.38}$$
  - So by comparison, we deduce Hamilton's equations from purely geometric principles:
$$\dot{p}_i = -\frac{dH}{dq_i}, \quad \dot{q}_i = \frac{dH}{dp_i} \tag{3.39}$$

- This is a new way of looking at Hamilton's equations: solving them is equivalent to finding the integral curves of a certain vector field in phase space

- Now recall: integral curves of a complete vector field $X \in \mathfrak{X}(M)$ are in bijection with a one-parameter family of diffeomorphisms $\phi_t : M \to M$, $t \in \mathbb{R}$ satisfying composition law $\phi_{t+s} = \phi_t \circ \phi_s$ called **flows**

- Given flow ϕ_t, the corresponding vector field X can be reconstructed:
$$X_x(f) = \frac{d}{dt} f(\phi_t(x))\bigg|_{t=0} \quad \forall x \in M,\ f \in C^\infty(M) \tag{3.40}$$

- A **Hamiltonian flow** is the flow of a complete Hamiltonian vector field $X_H \in Ham(M)$

- A **Hamiltonian system** $(M, \omega; H)$ is a symplectic manifold (M, ω) with a chosen function $H \in C^\infty(M)$

- Time-evolution of a Hamiltonian system is defined by the Hamiltonian flow of X_H

- **Lemma:** for $H \in C^\infty(M)$ and $X_H \in Ham(M)$, $\mathcal{L}_{X_H}\omega = 0$ so Hamiltonian vector fields preserve the symplectic form

- **Proof:**
  - Recall Cartan's magic identity for the Lie derivative of k-forms:
$$\mathcal{L}_{X_H}\omega = d\iota_{X_H}\omega + \iota_{X_H}d\omega = -d^2H = 0 \quad \blacksquare \tag{3.41}$$

- **Theorem:** let ϕ_t be a Hamiltonian flow on symplectic manifold (M, ω); then the symplectic form is preserved by the flow, $\phi_t^* \omega = \omega$
  - At arbitrary $x \in M$, we have

$$\begin{aligned}(\phi_t^*\omega - \omega)_x &= \int_0^t \frac{d}{du}(\phi_u^*\omega)_x \, du \\ &= \int_0^t \frac{d}{ds}\left(\phi_{u+s}^*\omega\right)_x \bigg|_{s=0} du \quad \text{(chain rule)} \\ &= \int_0^t \left(\phi_u^* \frac{d}{ds}\phi_s^*\omega\bigg|_{s=0}\right)_x du \\ &= \int_0^t \phi_u^*(\mathcal{L}_{X_H}\omega)_{\phi_u(x)} \, du \\ &= 0\end{aligned} \tag{3.42}$$

  - The second last line comes from definition of the Lie derivative of a k-form:

$$\mathcal{L}_{X_H}\omega = \frac{d}{dt}\phi_t^*\omega\bigg|_{t=0} \quad \blacksquare \tag{3.43}$$

- **Corollary:** a Hamiltonian flow ϕ_t preserves any exterior power of the symplectic form; that is, $\phi_t^*\omega^k = \omega^k$
- **Corollary:** (Liouville's theorem) a Hamiltonian flow ϕ_t preserves the volume form $\phi_t^*\omega^n = \omega^n$
- This means that Hamiltonian flows preserve volumes in phase space
- **Theorem:** the Hamiltonian function $H \in C^\infty(M)$ is preserved by Hamiltonian flow associated to $X_H \in Ham(M)$
- **Proof:**
  - First, note

$$\mathcal{L}_{X_H}H = X_H(H) = dH(X_H) = -(\iota_{X_H}\omega)(X_H)) = -\omega(X_H, X_H) = 0 \tag{3.44}$$

  - Then, by a similar argument to the last theorem, we are done $\blacksquare$
- This result is essentially conservation of energy

3.4 The Poisson Bracket

- Given symplectic manifold (M, ω), we can define the **Poisson bracket** - a binary operation on $C^\infty(M)$
- Specifically, given $f, g \in C^\infty(M)$, the Poisson bracket is

$$\{f, g\} = X_g(f) \tag{3.45}$$

  where X_g is the Hamiltonian vector field associated with g
- Alternatively, we can think of this as the rate of change of f along the Hamiltonian flow ϕ_t of g:

$$\{f, g\}_x = \frac{d}{dt}f(\phi_t(x))\bigg|_{t=0}, \; \forall x \in M \tag{3.46}$$

- This definition is appealing geometrically, but in practice it is more useful to rewrite by noticing the following:

$$X_g(f) = df(X_g) = \iota_{X_g} df = -\iota_{X_g}\iota_{X_f}\omega = -\omega(X_f, X_g) \tag{3.47}$$

  and so

$$\{f, g\} = -\omega(X_f, X_g) \tag{3.48}$$

- This immediately implies that the Poisson bracket acts as $\{\cdot,\cdot\}: C^\infty(M) \to C^\infty(M) \to C^\infty(M)$, is $\mathbb{R}$-*bilinear*, and is *skew-symmetric*
- Moreover, using the musical isomorphism, we notice

$$\omega^\sharp(df, dg) = -\omega\left((df)^\sharp, (dg)^\sharp\right) = -\omega(X_f, X_g) = \{f, g\} \tag{3.49}$$

  which implies the following lemma

- **Lemma:** the Poisson bracket satisfies the *Leibnitz rule*

$$\{f_1 f_2, g\} = f_1\{f_2, g\} + \{f_1, g\}f_2 \tag{3.50}$$

  for all $f_1, f_2, g \in C^\infty(M)$

- As always, to do more useful things with the Poisson bracket, we need to express it in a coordinate basis
- Consider basis $e_a = \partial_a$ defined by chart (x^a), and the dual basis $f^a = dx^a$; then we find

$$\{f, g\} = \omega^\sharp\left(\partial_a f\, dx^a, \partial_b g\, dx^b\right) = \omega^{ab}\partial_a f \partial_b g \tag{3.51}$$

- **Example:** we look again at $(M, \omega) = (\mathbb{R}^{2n}, \omega)$ with coordinates (p_i, q_i) so $\omega = dp_i \wedge dq_i$
  - We know that in the $(\partial/\partial p_i, \partial/\partial q_i)$ basis

$$\omega_{ab} = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \implies \omega^{ab} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \tag{3.52}$$

  - Then, we find

$$\{f, g\} = \omega^{ab}\partial_a f \partial_b g = \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \tag{3.53}$$

    which is the usual definition of it from classical dynamics

- We also find that we can define the Poisson bracket for locally defined functions by just restricting the domain
- **Example:** consider the example from a while back, where Q is an n-dimensional smooth manifold with chart $(q_i(x))$ and the cotangent bundle $T^*Q = \{(p_x, x): p_x \in T^*_x Q,\ x \in Q\}$ is a $2n$-dimensional symplectic manifold with chart $(p_i(x), q_i(x))$ and $\omega = dp_i \wedge dq_i$
  - The coordinates are functions defined only on the chart $U \subset M$
  - Then by the earlier example, we find the *fundamental Poisson brackets*:

$$\{q_i, q_j\} = \{p_i, p_j\} = 0, \quad \{q_i, p_j\} = \delta_{ij} \tag{3.54}$$

- Now recall that the time-evolution of a Hamiltonian system $(M, \omega; H)$ is given by the flow ϕ_t of X_H

- Therefore for a smooth function $f \in C^\infty(M)$, along the flow we find

$$\frac{d}{dt}\phi_t^* f = X_H(f) = \{f, H\} \tag{3.55}$$

- We usually write this in the shorthand

$$\frac{d}{dt} f = \{f, H\} \tag{3.56}$$

- We can define a **first integral/constant of motion** of $(M, \omega; H)$ is any function $f \in C^\infty(M)$ which is invariant under the flow of X_H; that is, f is a constant of motion iff it Poisson commutes with H: $\{f, H\} = 0$

- **Theorem:** the Hamiltonian H is preserved by the flow of $X_f \in Ham(M)$ if and only if $f \in C^\infty(M)$ is a constant of motion

- **Proof:**

  – Note that
$$X_f(H) = \{H, f\} = -\{f, H\} \tag{3.57}$$
  and so $X_f(H) = 0$ iff $\{f, H\} = 0$ as required ■

- This is just Noether's theorem in phase space

- It turns out the Poisson bracket has a *Lie algebra* structure; before we get to this, we need to first relate the commutator of two Hamiltonian vector fields to their Poisson bracket

- Recall the definition of the commutator of two vector fields:

$$[X, Y](f) = X(Y(f)) - Y(X(f)), \quad \forall f \in C^\infty(M) \tag{3.58}$$

- **Theorem:** let X_f and X_g be two Hamiltonian vector fields; then

$$[X_f, X_g] = -X_{\{f,g\}} \tag{3.59}$$

  for $f, g \in C^\infty(M)$

- **Proof:**

  – Recall that X_H is Hamiltonian if and only if there exists $H \in C^\infty(M)$ such that $\iota_{X_H}\omega = -dH$

  – Therefore, we compute

$$\begin{aligned}
\iota_{[X_f, X_g]}\omega &= \left(\mathcal{L}_{X_f}\iota_{X_g} - \iota_{X_g}\mathcal{L}_{X_f}\right)\omega && \text{(Cartan identity)} \\
&= -\mathcal{L}_{X_f} dg && (\iota_{X_g}\omega = -dg \text{ and } \mathcal{L}_{X_f}\omega = 0) \\
&= -d\iota_{X_f} dg && \text{(Cartan's magic identity)} \\
&= -d\left(X_f(g)\right) && \text{(by definition)} \\
&= d\{f, g\} && \text{(by definition)}
\end{aligned} \tag{3.60}$$

  – The proof then follows from the recalled definition of Hamiltonian vector fields and non-degeneracy of ω ■

3.5 Lie Algebra of Hamiltonian and Symplectic Fields

- A **Lie algebra** $\mathfrak{g}$ is a vector space equipped with a bilinear map $[\cdot,\cdot]: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the **Lie bracket**, which satisfies:

  1. *Skew-symmetry:* $[x, y] = -[y, x]$
  2. *Jacobi identity:* $[x, [y, z]] + [y, [x, z]] + [z, [x, y]] = 0$

  for all $x, y, z \in \mathfrak{g}$

- The set of vector fields $\mathfrak{X}(M)$ equipped with the commutator (serving as the Lie bracket) is a Lie algebra

- A natural question is whether $Ham(M)$ is a Lie algebra - recall that $[X_f, X_g] = -X_{\{f,g\}}$, so the Lie bracket of Hamiltonian vector fields is Hamiltonian, and hence $Ham(M)$ is closed under $[\cdot,\cdot]$

- So all we need to do is check whether it is a vector space, which it is

- **Theorem:** the set of Hamiltonian vector fields $Ham(M)$ is a real Lie algebra under the Lie bracket of vector fields

- **Proof:**

  – If $X_f, X_g \in Ham(M)$, then $\alpha X_f + \beta X_g = X_{\alpha f + \beta g} \in Ham(M)$ for $\alpha, \beta \in \mathbb{R}$, so $Ham(M)$ is a real vector space

  – Moreover, the above discussion shows it is closed under the commutator of vector fields, which acts as a Lie bracket ■

- Note that we can define X_f for any $f \in C^\infty(M)$, so $Ham(M)$ is an infinite dimensional Lie algebra

- **Theorem:** the Poisson bracket satisfies the Jacobi identity

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0 \tag{3.61}$$

  for all $f, g, h \in C^\infty(M)$

- **Proof:**

  – By definition of the bracket, looking at the three terms in turn:

$$\begin{aligned}
\{\{f, g\}, h\} &= -\{h, \{f, g\}\} = -X_{\{f,g\}}(h) \\
\{\{g, h\}, f\} &= -\{\{h, g\}, f\} = -\{X_g(h), f\} = -X_f(X_g(h)) \\
\{\{h, f\}, g\} &= \{X_f(h), g\} = X_g(X_f(h))
\end{aligned} \tag{3.62}$$

  – Summing over the right hand sides, we find

$$-[X_f, X_g](h) - X_{\{f,g\}}(h) = 0 \tag{3.63}$$

  and so the left hand sides sum to zero too ■

- **Corollary:** $C^\infty(M)$ equipped with the Poisson bracket is a Lie algebra

- Note that $\phi: C^\infty(M) \to Ham(M)$, $f \mapsto -X_f$ is a Lie algebra homomorphism, given by $[\phi(f), \phi(g)] = \phi(\{f, g\})$

- The kernel of ϕ is given by constant functions on M, and so by virtue of $\ker\phi$ being non-empty, ϕ is not an isomorphism
- In the absence of additional structure, there is no canonical notion of a Lie bracket for $C^\infty(M)$ - a symplectic structure provides such a bracket
- **Theorem:** if f, g are constants of motion for a Hamiltonian system $(M, \omega; H)$, then $\{f, g\}$ is also a constant of motion
- **Proof:**
  - By definition, $\{f, H\} = \{g, H\} = 0$
  - Then, by Jacobi:

$$\{\{f, g\}, H\} = -\{\{g, H\}, f\} - \{\{H, f\}, g\} = 0 \quad \blacksquare \tag{3.64}$$

- This shows that constants of motion of a Hamiltonian system form a Lie algebra under the Poisson bracket
- From the theorem stating that H is invariant under the flow of X_f iff f is a constant of motion, this therefore corresponds to the Lie subalgebra of $Ham(M)$ preserving H
- Given (M, ω), a **symplectomorphism** is a diffeomorphism $\phi : M \to M$ which preserves the symplectic form $\phi^*\omega = \omega$
- A vector field $X \in \mathfrak{X}(M)$ is **symplectic** if $\mathcal{L}_X\omega = 0$; we denote this as $X \in Sym(M)$
- This means that if the flow $\phi_t : M \to M$ of a vector field X is a symplectomorphism, then X is symplectic
- This follows from $(\mathcal{L}_X\omega)_p = \frac{d}{dt}(\phi_t^*\omega)_p|_{t=0} = 0$ since $\phi_t^*\omega = 0$
- The converse is in fact true, following the same proof as earlier with all the integrals
- We know that Hamiltonian vector fields preserve the symplectic form: $\mathcal{L}_{X_H}\omega = 0$; Hamiltonian vector fields are therefore symplectic
- The converse is sometimes true: if X is symplectic, then

$$0 = \mathcal{L}_X\omega = d\iota_X\omega \quad \iff \quad \iota_X\omega \text{ is closed} \quad \iff \quad \iota_X\omega \text{ is locally exact} \tag{3.65}$$

  where the first implication follows from the definition of closure being $d\alpha = 0$, and the second from the *Poincare lemma* (every closed form is exact - the exterior derivative of another form)
- Therefore the converse is true if the de Rham cohomology $H^1(M, \mathbb{R}) = 0$ - every closed 1-form is exact - then there exists $f \in C^\infty(M)$ such that $\iota_X\omega = -df$, so X is Hamiltonian
- **Example:** consider the 2-torus $M = T^2$ with 2π-periodic coordinates (p, q), and symplectic form $\omega = dp \wedge dq$
  - The flow corresponding to $X = \partial/\partial p$ is $\phi_t(p, q) = (p + t, q)$; then $\mathcal{L}_X\omega = 0$ so X is symplectic
  - However, $\iota_X\omega = dq$ is not exact on $\mathrm{T}T^2$ since q is only defined locally (i.e. we have a discontinuous jump in q globally), and so X is not Hamiltonian
  - This is because $H^1(T^2, \mathbb{R}) \neq 0$

- **Theorem:** the set $Sym(M)$ forms a Lie algebra under the Lie bracket of vector fields
- **Proof:**
  - $Sym(M)$ is a real vector space since $\mathcal{L}_X$ is $\mathbb{R}$-linear in X
  - Let $X, Y \in Sym(M)$, and compute

$$\begin{aligned}
\iota_{[X,Y]}\omega &= \mathcal{L}_X \iota_Y \omega - \iota_Y \mathcal{L}_X \omega && \text{(Cartan identity)} \\
&= \mathcal{L}_X \iota_Y \omega && (X \in Sym(M) \implies \mathcal{L}_X \omega = 0) \\
&= (d\iota_X + \iota_X d)\iota_Y \omega && \text{(Cartan's magic identity)} \\
&= d(\iota_X \iota_Y \omega) && (Y \in Sym(M) \iff \iota_Y \omega \text{ closed})
\end{aligned} \tag{3.66}$$

  - This means $[X, Y] \in Ham(M)$ with Hamiltonian $\omega(X, Y)$; and therefore $[X, Y] \in Sym(M)$ meaning $Sym(M)$ is a Lie algebra ■
- We can write this conclusion as $[Sym(M), Sym(M)] \subseteq Ham(M)$
- But $Ham(M) \subseteq Sym(M)$ is then an ideal in $Sym(M)$:

$$[Sym(M), Ham(M)] \subseteq Ham(M) \tag{3.67}$$

3.6 Darboux's Theorem

- For the entirety of this section, (M, ω) is a $2n$-dimensional symplectic manifold
- A *symplectic chart* on (M, ω) is a coordinate chart (p_i, q_i), $i = 1, \ldots, n$ on M such that $\omega = dp_i \wedge dq_i$
- In physics, symplectic charts are called canonical coordinates
- **Lemma:** a chart (p_i, q_i) is symplectic iff $\{q_i, p_j\} = \delta_{ij}$, $\{q_i, q_j\} = \{p_i, p_j\} = 0$
- **Proof:**
  - The above Poisson brackets are equivalent to ω^{ab} taking the following form in the dp_i, dq_i basis:

$$\omega^{ab} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \tag{3.68}$$

  - This is in turn equivalent to ω_{ab} taking the usual canonical form in the $\partial/\partial p_i, \partial/\partial q_i$ basis ■
- **Theorem:** (Darboux) let (M, ω) be a symplectic manifold. Then any point $x \in M$ is contained in a symplectic chart
- **Proof:**
  - This is a long one, so strap in
  - Let $x \in M$, and pick some $p_1 \in C^\infty(M)$ such that $p_1(x) = 0$ but $dp_1|_x \neq 0$
  - This is always possible (e.g. just choose p_1 to be linear)
  - The Hamiltonian vector field X_{p_1} of p_1 (defined by $\iota_{X_{p_1}}\omega = -dp_1$) is unique and non-vanishing at x: $X_{p_1}|_x \neq 0$
  - Define N to be a $(2n-1)$-dimensional submanifold of M containing x, where $(X_{p_1})_x \notin T_x N$; N is *transverse* to X_{p_1} (i.e. we pick a chart adapted to X_{p_1})

- Now let ϕ_1^t be the Hamiltonian flow of X_{p_1}, and consider the map $F : N \times \mathbb{R} \to M$, defined by $F(y,t) = \phi_1^t(y)$ for $y \in N$, $t \in \mathbb{R}$
- Recall that the derivative of a smooth map $\phi : A \to B$ is the pull-back $\phi_* : T_pA \to T_{\phi(p)}B$, given by $(\phi_* X_p)(f) = X_p(f \circ \phi)$
- So the derivative of F at $(x,0)$ is the pull-back $F_* : T_xN \times T_0\mathbb{R} \to T_xM$, and $F_*(Y_x, (d/dt)_0) = (Y_x, (X_{p_1})_x)$
- This has full rank since $(X_{p_1})_x$ is linearly independent of all $Y_x \in T_xN$
- Therefore, F_* is invertible, and by the *inverse function theorem*, F is a local diffeomorphism
- What this means is that there is a neighbourhood (open set) of $(x,0) \in N \times \mathbb{R}$ that is diffeomorphic to a neighbourhood U of x in M
- For any $z \in U$, we can therefore assign a unique time t along the flow ϕ_1^t from $y \in N$, so $z = \phi_1^t(y)$ defines a smooth function $q_1(z) = t$ for $z \in U$
- Note that $q_1|_N = 0$, and we can calculate the derivative at z in the direction of X_{p_1}:
$$(X_{p_1})_z(q_1) = \frac{d}{ds} q_1(\phi_1^s(z))|_{s=0} = \frac{d}{ds} q_1(\phi_1^{s+t}(y))|_{s=0} = 1 \tag{3.69}$$
- Since this holds for any $z \in U$, we have that $X_{p_1}(q_1) = \{q_1, p_1\} = 1$
- For $n = 1$, this completes the construction of a symplectic chart
- For $n > 1$, we use induction
- Assume the theorem holds for $(2n-2)$-dimensional symplectic manifolds
- Note that the dp_1, dq_1 are linearly independent on U since $1 = \{q_1, p_1\} = \omega^\sharp(dq_1, dp_1)$
- So by the *implicit function theorem*, the level set $\bar{M} = \{w \in U \,|\, q_1(w) = p_1(w) = 0\}$ is a $(2n-2)$-dimensional submanifold of M containing x
- We now claim that $(\bar{M}, \omega|_{\bar{M}})$ is a $(2n-2)$-dimensional symplectic manifold
- $\omega|_{\bar{M}}$ inherits skew-symmetry from ω, so we need to check non-degeneracy
- To do this, take $\xi \in T_w\bar{M}$ to be a tangent vector on $\bar{M}$, so $\xi(p_1) = \xi(q_1) = 0$ since p_1 and q_1 are constant on $\bar{M}$
- Then:
$$\omega(X_{p_1}, \xi) = -dp_1(\xi) = 0 \quad \text{and} \quad \omega(X_{q_1}, \xi) = -dq_1(\xi) = 0 \tag{3.70}$$
- This means $T_w\bar{M}$ is skew-orthogonal (with respect to ω) to $span(X_{p_1}, X_{q_1})|_w$ for all $w \in \bar{M}$
- Recall the proof that every symplectic vector space has a symplectic basis: this same method of proof then shows that $\omega|_{\bar{M}}$ is non-degenerate
- Therefore by the induction hypothesis, there are symplectic coordinates (p_i, q_i), $i = 2, \ldots, n$ on a neighbourhood of x in $(\bar{M}, \omega|_{\bar{M}})$
- We can extend these to a neighbourhood of x in M (i.e. to p_1, $q_1 \neq 0$) as follows
- Any point z in a neighbourhood $\tilde{U} \subseteq U$ of x in M can be written uniquely as $z = \phi_1^t \psi_1^s w$ for some $w \in \bar{M}$, where ψ_1^s is the flow of X_{q_1}
- We therefore extend (p_i, q_i), $i = 2, \ldots, n$ to $\tilde{U}$ by keeping their values constant along flows of X_{p_1} and X_{q_1}:
$$X_{q_1}(q_i) = X_{q_1}(p_i) = X_{p_1}(q_i) = X_{p_1}(p_i) = 0 \tag{3.71}$$
so the coordinates of z are the same as w; that is, $p_i(z) = p_i(w)$, $q_i(z) = q_i(w)$

- Then $(p_1, \ldots, p_n, q_1, \ldots, q_n)$ are coordinates for $\tilde{U}$ of x in M; we now need to show these are symplectic
- We know already that $\{q_1, p_1\} = 1$, and note that $\{q_i, q_1\} = X_{q_1}(q_i) = 0$ and $\{p_i, q_1\} = \{q_i, p_1\} = \{p_i, p_1\} = 0$ similarly for $i = 2, \ldots, n$ on $\tilde{U}$
- By the induction hypothesis, $\{q_i, p_j\} = \delta_{ij}$ on $\bar{M} \cap \tilde{U}$
- For $i, j = 2, \ldots, n$ on $\tilde{U}$, we have:

$$X_{q_1}(\{q_i, p_j\}) = \{\{q_i, p_j\}, q_1\} = \{q_i, \{p_j, q_1\}\} + \{p_j, \{q_1, q_i\}\} = 0 \tag{3.72}$$

  and similarly, $X_{p_1}(\{q_i, p_j\}) = 0$
- Therefore, since the functions $\{q_i, p_j\}$ are unchanged on the flows of X_{p_1} and X_{q_1}, they still satisfy $\{q_i, p_j\} = \delta_{ij}$ on $\tilde{U}$
- Similar arguments hold for the remaining Poisson brackets, and so we are done ■

- Note that this shows that symplectic manifolds are locally isomorphic to symplectic vector spaces

3.7 Integrable Systems

- Loosely speaking, an integrable system is one which can be solved exactly; but what does this precisely mean?

- Liouville gave a definition: a $2n$-dimensional Hamiltonian system $(M, \omega; H)$ is *Liouville integrable* if:

  1. There exists $f_i \in C^\infty(M)$, $i = 1, \ldots, n$ in *involution*; that is,

  $$\{f_i, f_j\} = 0, \quad \forall i, j = 1, \ldots, n \tag{3.73}$$

  2. The functions f_i are functionally independent on the level sets

  $$M_{\mathbf{c}} = \{x \in M \,|\, f_i(x) = c_i,\, i = 1, \ldots, n\} \tag{3.74}$$

  (i.e. the one forms df_i are linearly independent on this set)

- Note that a $2n$-dimensional symplectic manifold can have at most n independent functions in involution

- If f_i, $i = 1, \ldots, k$ are such a set of functions, we find

$$\omega(X_{f_i}, X_{f_j}) = -\{f_i, f_j\} = 0 \tag{3.75}$$

  so $span(X_{f_1}, \ldots, X_{f_k})$ is a k-dim. null subspace of TM (i.e. all functions get mapped to zero), hence $k \leq n$

- **Theorem:** (Liouville-Arnold) let $(M, \omega; H)$ be a $2n$-dimensional Hamiltonian system which is Liouville integrable. Then:

  1. $M_{\mathbf{c}}$ is a smooth, n-dimensional submanifold of M, which is invariant under the flow of X_H; that is, if ϕ_t is the flow of X_H and $x \in M_{\mathbf{C}}$, then $\phi_t(x) \in M_{\mathbf{c}}$
  2. If $M_{\mathbf{c}}$ is compact and connected, it is diffeomorphic to an n-torus $T^n = \{(\theta_1, \ldots, \theta_n) \bmod 2\pi\}$
  3. The flow of X_H on a compact $M_{\mathbf{c}} \cong T^n$ takes the form $\dot{\theta}_i = \omega_i(\mathbf{c})$
  4. In a neighbourhood of $M_{\mathbf{c}} \cong T^n$, there exists a symplectic chart (I_i, θ_i) (*action-angle coordinates*) where $I_i = I_i(\mathbf{c})$ are constants of motion

- **Proof (1):**
  - By definition of Liouville integrable, the df_i are linearly independent on $M_{\mathbf{c}}$
  - So by the **implicit function theorem**, the level sets $M_{\mathbf{c}}$ are smooth, n-dimensional submanifolds of M
  - We know H is a constant of motion, so we can set $H = f_1$
  - Then, using the fact that the f_i are in involution, we have

$$X_H(f_i)) = -\{H, f_i\} = -\{f_1, f_i\} = 0 \tag{3.76}$$

    which shows that the $M_{\mathbf{c}}$ are tangent to X_H, and hence invariant submanifolds of its flow ■

- Before proving (2), we need a lemma
- **Lemma:** The Hamiltonian vector fields X_{f_i}, $i = 1, \ldots, n$ are tangent to $M_{\mathbf{c}}$, pairwise Lie commute, and are linearly independent at every point in $M_{\mathbf{c}}$
- **Proof:**
  - To prove pairwise commutation, consider the Hamiltonian vector fields X_{f_i} for $i = 1, \ldots, n$
  - We have:

$$[X_{f_i}, X_{f_j}] = -X_{\{f_i, f_j\}} = 0 \tag{3.77}$$

    by involution
  - To show linear independence, recall that ω is non-degenerate ($\omega(X, Y) = 0 \forall X \in \mathfrak{X}(M) \implies Y = 0$), and that $\iota_{X_{f_i}}\omega = -df_i$
  - Then since the df_i are linearly independent on $M_{\mathbf{c}}$ by definition of Liouville integrability, so too are the X_{f_i}
  - For tangent, note that $X_{f_i}(f_j) = -\{f_i, f_j\} = 0$, which means the X_{f_i} are tangent to $M_{\mathbf{c}}$
  - Therefore, X_{f_i} give a nowhere vanishing, pairwise commuting *frame* of vector fields on $M_{\mathbf{c}}$ (i.e. a basis for the tangent space at every point) ■
- We should note that there's actually a little more hidden under the surface of this lemma
- Since $X_{f_i}(f_j) = 0$, it follows that $M_{\mathbf{c}}$ is invariant under the flow of all the X_{f_i}
- Moreover, $\omega(X_{f_i}, X_{f_j}) = -\{f_i, f_j\} = 0$, which implies that $M_{\mathbf{c}}$ is an n-dimensional **null** submanifold; that is, $T_x M_{\mathbf{x}}$ is a null subspace of $T_x M$ for all $x \in M_{\mathbf{c}}$ (also known as a *Lagrangian submanifold*)
- We now use the following lemma for compact and connected manifolds
- **Lemma:** a compact, connected, n-dimensional smooth manifold N admitting a pairwise commuting frame of vector fields is diffeomorphic to T^n
- **Proof:**
  - Denote the commuting frame of vector fields by X_i, $i = 1, \ldots, n$, and its flows by ϕ_i^t
  - The frame is commuting by assumption, so the flows ϕ_i^t and ϕ_j^s commute too

- We can therefore define diffeomorphisms $\phi^{\mathbf{t}} : M \to M$ by $\phi^{\mathbf{t}} = \phi_1^{t_1} \circ \phi_2^{t_2} \circ \ldots \circ \phi_n^{t_n}$, where $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n$, and $\phi^{\mathbf{t}+\mathbf{s}} = \phi^{\mathbf{t}}\phi^{\mathbf{s}}$
- Now we fix a point $p \in N$, and define a smooth map $\phi : \mathbb{R}^n \to N$ by $\phi(\mathbf{t}) = \phi^{\mathbf{t}}(p)$ (i.e. p moves by t_i along the flow of X_i)
- The derivative of this map at the origin is the pull-back $\phi_* : T_0\mathbb{R}^n \to T_pN$, where $\phi_* \left(\frac{\partial}{\partial t_i}\right)_0 = X_i(p)$
- Since the $X_i(p)$ are linearly independent (as they form a frame), this derivative has full rank
- So by the inverse function theorem, ϕ defines a local diffeomorphism in a neighbourhood of the origin
- It can also be shown that ϕ is onto
- Now consider the set $\Gamma = \{\mathbf{t} \in \mathbb{R}^n \,|\, \phi(\mathbf{t}) = p\}$, which is a subgroup of $\mathbb{R}^n$ under addition and independent of p
- Moreover, $\mathbf{0} \in \Gamma$ and since ϕ is a diffeomorphism in a neighbourhood of $\mathbf{0}$, there can be no other elements of Γ in this neighbourhood
- Similarly, for any $\mathbf{t} \in \Gamma$ there is a neighbourhood containing it and no other elements of Γ, and so Γ is a discrete subgroup
- Any discrete subgroup of $\mathbb{R}^n$ has the form

$$\Gamma = \{n_1\mathbf{t}_1 + \ldots + n_k\mathbf{t}_k \,|\, (n_1, \ldots, n_k) \in \mathbb{Z}^k\} \tag{3.78}$$

  where $\mathbf{t}_i$ is a set of $0 \le k \le n$ linearly independent vectors on $\mathbb{R}^n$
- Now consider an isomorphism $A : \mathbb{R}^n \to \mathbb{R}^n$ such that $A(2\pi\mathbf{e}_i) = \mathbf{t}_i$ for $i = 1, \ldots, k$ and $\mathbf{e}_i$, $i = 1, \ldots, n$ is the standard orthonormal basis of $\mathbb{R}^n$
- Define the projection $\pi : \mathbb{R}^n \to T^k \times \mathbb{R}^{n-k}$ by

$$\pi(\theta_1, \ldots, \theta_k, x_{k+1}, \ldots, x_n) = (\theta_1 \bmod 2\pi, \ldots, \theta_k \bmod 2\pi, x_{k+1}, \ldots, x_n) \tag{3.79}$$

- Then, $\pi^{-1}(0) = \{2\pi(n_1\mathbf{e}_1 + \ldots + n_k\mathbf{e}_k) \,|\, (n_1, \ldots, n_k) \in \mathbb{Z}^k\}$, and hence $A \circ \pi^{-1}(0) = \Gamma$ and $\phi \circ A \circ \pi^{-1}(0) = p$
- Similarly, $\tilde{A} = \phi \circ A \circ \pi^{-1}$ maps any other point in $T^k \times \mathbb{R}^{n-k}$ to a unique point in N, and defines a diffeomorphism $\tilde{A} : T^k \times \mathbb{R}^{n-k} \to N$
- Since N is compact, we must have $k = n$, and hence $N \cong T^n$ ■

- **Proof (2):**
  - We know already that $M_{\mathbf{c}}$ admits a commuting frame of vector fields
  - So by the lemma, we are done ■
- **Proof (3):**
  - By the second lemma, $M_{\mathbf{c}} \cong T^n = \{(\theta_1, \ldots, \theta_n) \bmod 2\pi\}$, and $A\theta = \mathbf{t}$ where $\mathbf{t} = (t_1, \ldots, t_n)$ are the flow parameters of X_{f_i} and A is an invertible matrix
  - The diffeomorphism depends on $\mathbf{c} \in \mathbb{R}^n$, and so too must A then
  - To determine the dependence of the torus coordinates on the flow, we invert so $\theta = A^{-1}\mathbf{t}$
  - $H = f_1$, so the evolution under flow ϕ^t of X_H is obtained by setting $t = t_1$
  - Therefore, θ_i is linearly dependent on t, so $\dot{\theta}_i = \omega_i(\mathbf{c})$ ■

- We don't prove (4)
- In the angular coordinates on the torus, the Hamiltonian flow takes the simple form $\theta_i(t) = \theta_i(0) + \omega_i(\mathbf{c})$
- The functions f_i together with the torus coordinates θi define a chart for M in the neighbourhood of $M_{\mathbf{c}}$
- The flow in the coordinates (f_i, θ_i) is simply $\dot{f}_i = 0$ and $\dot{\theta}_i = \omega_i(f)$, which integrates as above
- In general, (f_i, θ_i) are not symplectic, but we can show there are functions $I_i(f)$ called actions which make (I_i, θ_i) symplectic
- The Hamiltonian flow defined by $H = f_1 = f_1(I)$ is therefore just given by Hamilton's equations

$$\dot{I}_i = 0, \quad \dot{\theta}_i = \frac{\partial H}{\partial I_i} \tag{3.80}$$

4 Lie Groups and Lie Algebras

- Motivation: Lie groups and their corresponding Lie algebras arise in studying symmetries in phsyics/geometry
- A *Lie group* is a smooth manifold G which has a group structure; that is, there are group operations at points $g, h \in G$:
  - *Multiplication:* $\mu : G \times G \to G$, with $\mu(g,h) = g \cdot h$
  - *Inversion:* $\iota : G \to G$, with $\iota(g) = g^{-1}$

  which are smooth
- We typically denote the identity element as $e \in G$
- The most important examples of these are matrix groups, the largest of which is the *general linear groups* $GL(n, \mathbb{R} \text{ or } \mathbb{C})$, consisting of the invertible $n \times n$ matrices over $\mathbb{R}$ or $\mathbb{C}$, with group multiplication just matrix multiplication
- The following theorem is stated without proof
- **Theorem:** any closed subgroup H of a Lie group G is itself a Lie group
- Given this theorem, we have some common Lie groups which are closed subgroups of the general linear groups:
  - *Special linear groups:* $SL(n, \mathbb{R} \text{ or } \mathbb{C}) = \{M \in GL(n, \mathbb{R} \text{ or } \mathbb{C}) \mid \det M = 1\}$
  - *Orthogonal groups:* $O(n) = \{M \in GL(n, \mathbb{R}) \mid M^T M = I_n\}$
  - *Special orthogonal groups:* $SO(n) = O(n) \cap SL(n, \mathbb{R})$
  - *Unitary groups:* $U(n) = \{M \in GL(n, \mathbb{C}) \mid M^\dagger M = I_n\}$
  - *Special unitary groups:* $SU(n) = U(n) \cap SL(n, \mathbb{C})$
  - *Symplectic groups:* $Sp(2n) = \{M \in GL(2n, \mathbb{R}) \mid M^T \Omega M = \Omega\}$ where Ω is the usual canonical symplectic form
  - *Lorentz groups:* $O(1,4) = \{M \in GL(4, \mathbb{R}) \mid M^T \eta M = \eta\}$ where $\eta = \text{diag}(-1,1,1,1)$
- Now, recall the definition of a *Lie algebra* as a vector space equipped with a Lie bracket, which is a skew-symmetric bilinear form satisfying the Jacobi identity
- An easy (yet important) example of a Lie algebra is the set of all $n \times n$ matrices $Mat(n)$, equipped with the standard matrix commutator as a Lie bracket
- We now show that given a Lie group G, there is a 'natural' associated Lie algebra
- First, we define *left and right multiplication*, which are maps defined for each $g, h \in G$:

$$\begin{aligned} L_g : G \to G \quad &, \quad L_g h \mapsto gh \\ R_g : G \to G \quad &, \quad R_g h \mapsto hg \end{aligned} \tag{4.1}$$

- We immediately note that these are both smooth and bijections, as $L_g^{-1} = L_{g^{-1}}$ and $R_g^{-1} = R_{g^{-1}}$, and so these are therefore diffeomorphisms for each $g \in G$
- We say a vector field $X \in \mathfrak{X}(G)$ is *left-invariant* if $(L_g)_* X_h = X_{gh}$ for all $g, h \in G$, where $(L_g)_* : T_h G \to T_{gh} G$ is the derivative of L_g
- We denote the set of all such left-invariant vector fields as $L(G)$

- There's also a nice equivalent definition: $X \in L(G)$ iff $X_g = (L_g)_* X_e$ (stated without proof)
- Thus left invariant vector fields are uniquely determined by their values at the identity e
- **Lemma:** $L(G)$ is a Lie algebra under the Lie bracket of vector fields
- **Proof:**
  - Recall the definition of the Lie bracket of vector fields $X, Y \in \mathfrak{X}(G)$ acting on $f \in C^\infty(G)$:
$$[X, Y](f) = X(Y(f)) - Y(X(f)) \tag{4.2}$$
  - We need to show two things: $L(G)$ is a vector space, and it is closed under the Lie bracket
  - It is clearly a vector space: if $X, Y \in L(G)$, then $X + Y \in L(G)$ and $\alpha X \in L(G)$ where $\alpha \in \mathbb{R}$, since $(L_g)_*$ is $\mathbb{R}$-linear
  - It is also closed: if $X, Y \in L(G)$, since $L_g : G \to G$ is a smooth bijection, $(L_g)_*[X, Y] = [(L_g)_* X, (L_g)_* Y] = [X, Y]$ ■
- In general, there is no obvious natural Lie bracket on a tangent space T_pM to some manifold M at p
- For Lie groups G, the tangent space at the identity e inherits a Lie bracket from the Lie bracket on $L(G)$; specifically, for $X_e, Y_e \in T_eG$, we have:
$$[X_e, Y_e] = [X, Y]_e \tag{4.3}$$
  where $X|_g = (L_g)_* X_e$ etc.
- **Theorem:** $L(G) \cong T_eG$ is a Lie algebra isomorphism
- **Proof:**
  - Let $X \in L(G)$, and define $\phi : L(G) \to T_eG$ by $\phi(X) = X_e$
  - This establishes a vector space homomorphism in one direction
  - Conversely, let $V \in T_eG$, and define
$$(X_V)_g = (L_g)_* V \tag{4.4}$$
  - Then:
$$(L_g)_*(X_V)_h = (L_g)_*(L_h)_* V = (L_{gh})_* V = (X_V)_{gh} \tag{4.5}$$
    and so $X_V \in L(G)$
  - We now claim that X_V is inverse to ϕ
  - Note that $\phi(X_V) = (X_V)_e = V$ since $(L_e)_* = id$
  - Therefore ϕ is a vector space isomorphism
  - Now note that since $[X_e, Y_e] = [X, Y]_e$ for all $X_e, Y_e \in T_e(G)$, we have
$$\phi([X, Y]) = [\phi(X), \phi(Y)] \tag{4.6}$$
    which is just the statement that ϕ is a Lie algebra isomorphism ■
- We refer to T_eG as **the** Lie algebra $\mathfrak{g}$ of the Lie group G

- Note that $\dim \mathfrak{g} = \dim(G)$
- As a simple example, consider the (additive) Lie group $G = (\mathbb{R}, +)$; any left invariant vector field looks like $X = c\frac{\partial}{\partial x}$ where $c \in \mathbb{R}$ and x is the defining chart. Therefore, $\left(\frac{\partial}{\partial x}\right)_a = (L_a)_* \left(\frac{\partial}{\partial x}\right)_0$, establishing the correspondence between $L(\mathbb{R})$ and $T_0\mathbb{R}$
- For matrix Lie groups, consider $G = GL(n, \mathbb{R})$
- The entries of a matrix $a = (a_{ij}) \in G$ define a global chart via $g_{ij} : G \to \mathbb{R}$ where $g_{ij}(a) = a_{ij}$
- Then any vector field on G can be written as

$$X = X_{ij} \frac{\partial}{\partial g_{ij}} \tag{4.7}$$

  where X_{ij} are smooth functions
- Now, consider $V \in T_eG$, which is $V = V_{ij} \left(\partial / \partial g_{ij}\right)_e$, where $V_{ij} \in Mat(n)$; we then have the following theorem
- **Theorem:** for Lie group $G = GL(n, \mathbb{R})$, for all $a \in G$ and $V, W \in T_eG$, we have:
  1. $X_{V_{ij}}(a) = (aV)_{ij}$
  2. $[X_V, X_W]_{ij}(a) = (a(VW - WV))_{ij}$
  3. $[V, W]_{ij} = (VW - WV)_{ij}$
- **Proof:** see tutorial sheet
- This theorem shows that for matrix Lie groups, left translation corresponds to left matrix multiplication
- We also state the following without proof
- **Theorem:** every finite dimensional Lie algebra is the Lie algebra of a Lie group

4.1 Frames and Structure Equation

- A **frame** of vector fields on an n-dimensional manifold M is a set of n linearly independent vector fields $\forall p \in M$
- Typically, a manifold does not admit a global frame, but usually does locally
- Lie groups however do admit a global frame
- **Theorem:** any lie group G is **parallelisable**; there exists a globally defined frame of vector fields
- **Proof:**
  - Let $V_1, \ldots, V_n$ be a basis of T_eG
  - Then the left-invariant vector fields $X_{V_1}, \ldots, X_{V_n}$ give a global frame ■
- We say a form $\omega \in \Omega^k(G)$ is **left-invariant** if $L_g^* \omega_{gh} = \omega_h$ for all $g, h \in G$, where L_g^* is the pull-back of L_g
- As for left-invariant vector fields, this has an equivalent characterisation in terms of the form at the identity only

- Specifically, $\omega_g = L^*_{g^{-1}}\omega_e$
- Now, consider an arbitrary basis $V_1, \ldots, V_n$ of $\mathfrak{g} \cong T_eG$; since it is a Lie algebra, we must have
$$[V_i, V_j] = c_{ij}{}^k V_k \tag{4.8}$$
where $c_{ij}{}^k \in \mathbb{R}$ are known as the **structure constants** of $\mathfrak{g}$
- Now recall the Lie algebra isomorphism $L(G) \cong T_eG$; this means that the left-invariant vector fields also satisfy
$$[X_i, X_j] = c_{ij}{}^k X_k \tag{4.9}$$
where we abbreviate $X_i = X_{V_i}$
- Consider the dual one-forms to the X_i given by $\theta^i \in \Omega^1(G)$, $i = 1, \ldots, n$, so $\theta^i(X_j) = \delta^i_j$
- These θ^i are clearly left-invariant:
$$(L^*_g\theta^i_g)((X_j)_e) = \theta^i_g((L_g)_*(X_j)_e) = \theta^i_g((X_i)_g) = \delta^i_j \implies L^*_g\theta^i_g = \theta^i_e \tag{4.10}$$
- Moreover, these forms satisfy the *Maurer-Cartan structure equation*
- **Theorem:** the left-invariant one-forms θ^i satisfy
$$d\theta^i = -\frac{1}{2}c_{jk}{}^i \theta^j \wedge \theta^k \tag{4.11}$$
- **Proof:**
  - Act on the dual frame of vector fields:
  $$d\theta^i(X_j, X_k) = X_j(\theta^i(X_k)) - X_k(\theta^i(X_j)) - \theta^i([X_j, X_k]) = -c_{jk}{}^i$$
  where we used a general identity for exterior derivatives of one-forms
  - Moreover, note that
  $$(\theta^l \wedge \theta^m)(X_j, X_k) = \delta^l_j\delta^m_k - \delta^l_k\delta^m_j$$
  and so we are done ■
- With this in mind, we define the *Maurer-Cartan one-form* on G, which is a **Lie algebra valued** one-form $\theta \in \Omega^1(G; \mathfrak{g})$ given by
$$\theta_g : T_gG \to \mathfrak{g}, \quad \theta_g(X_g) = (L_{g^{-1}})_*X_g \in \mathfrak{g} \tag{4.12}$$
for all $X_g \in T_gG$
- Note that $\theta_e(X_e) = X_e$, so $\theta_e = id_e$ on $\mathfrak{g}$
- Moreover:
$$(L^*_g\theta_g)V = \theta_g((L_g)_*V) = \theta_g((X_V)_g) = (L_{g^{-1}})_*(X_V)_g = V$$
for $V \in T_eG$, and so $L^*_g\theta_g = id_e$ on $\mathfrak{g}$, meaning θ is left-invariant
- **Lemma:** the Maurer-Cartan form θ on G can be written as $\theta = \theta^i \otimes V_i$ where θ^i is the dual frame of one-forms to $X_{V_i} \in L(G)$
- **Proof:**
  - We act on an arbitrary $Y_g \in T_gG$, and expand $Y = Y^iX_i$ in terms of left-invariant frame $\{X_i\}_{i=1}^n$

– The LHS gives us:

$$\theta_g(Y_g) = Y^i\theta_g(X_i) = Y^i(L_{g^{-1}})_*(L_g)_*V_i = Y^iV_i$$

and the RHS give us

$$(\theta_g^i \otimes V_i)(Y_g) = \theta_g^i(Y_g)V_i = Y^iV_i$$

and we are done ■

- **Theorem:** the Maurer-Cartan form on G satisfies the structure equation

$$d\theta + \frac{1}{2}[\theta, \theta] = 0 \tag{4.13}$$

where $[\cdot, \cdot]$ denotes **both** the Lie bracket on $\mathfrak{g}$ and the wedge product on $\Omega^1(G)$

- **Proof:**
  – By the previous lemma, we can write:

$$\begin{aligned} d\theta = d\theta^i \otimes V_i &= -\frac{1}{2}c_{jk}{}^i\theta^j \wedge \theta^k \otimes V_i && \text{(by (4.11))} \\ &= -\frac{1}{2}\theta^j \wedge \theta^k \otimes [V_k, V_l] && \text{(by (4.8))} \\ &= -\frac{1}{2}[\theta, \theta] && \text{(by def.)} \end{aligned}$$

and we are done ■

- For matrix groups, the Maurer-Cartan form has an explicit matrix representation
- **Theorem:** for $G = GL(n, \mathbb{R})$, the Maurer-Cartan form is

$$\theta_{ij} = (g^{-1}dg)_{ij} \tag{4.14}$$

where $g = (g_{ij})$ is the global chart such that for $A \in G$, $g_{ij}(A) = A_{ij}$, and θ_{ij} are the components of $\theta \in \Omega^1(G, \mathfrak{g})$ in the corresponding basis of $\mathfrak{g}$; specifically, we mean that $(\theta_a)_{ij} = (g(a)^{-1})_{ik}(dg_{kj})_a$

- **Proof:**
  – In the coordinate basis, the components of $\theta_a = \theta_a^{ij}(dg_{ij})_a$ are

$$\theta_a^{ij} = \theta_a\left(\frac{\partial}{\partial g_{ij}}\right)_a = (L_{a^{-1}})_*\left(\frac{\partial}{\partial g_{ij}}\right)_a$$

by definition of θ_a

  – Moreover, define a basis for T_eG by $E^{ij} = (\partial/\partial g_{ij})_e$; then

$$(L_a)_*E^{pq} = (aE^{pq})_{ij}\left(\frac{\partial}{\partial g_{ij}}\right)_a = a_{ip}\left(\frac{\partial}{\partial g_{iq}}\right)_a$$

  – Multiplying both sides by $a_{pk}^{-1}(L_{a^{-1}})$ gives

$$(:_{a^{-1}})_*\left(\frac{\partial}{\partial g_{kq}}\right)_a = a_{pk}^{-1}E^{pq} = g_{pk}^{-1}(a)E^{pq}$$

  – Relabelling indices as needed, we therefore obtain

$$\theta_a^{ij} = g_{pi}^{-1}(a)E^{pj} \implies \theta_a = g_{pi}^{-1}(a)(dg_{ij})_aE^{pj}$$

and so the matrix components relative to the basis E^{pj} are

$$(\theta_a)_{pj} = g_{pi}^{-1}(a)(dg_{ij})_a$$

for any $a \in G$, and we are done ■

- This proof works for any Lie subgroup of $GL(n, \mathbb{R})$ with the subtlety that g_{ij} is no longer a chart

4.2 The Exponential Map and Adjoint Representation

- A curve $c:\mathbb{R}\to G$ where G is a Lie group is a *one-dimensional subgroup of G* if

$$c(t)c(s) = c(t+s),\ \forall s,t\in\mathbb{R} \tag{4.15}$$

- Alternatively, c is a group homomorphism from $(\mathbb{R},+)$ to G
- Note the basic properties that $c(0) = e$ and $c(-t) = c(t)^{-1}$, and that the subgroup is Abelian in that $c(t)c(s) = c(s)c(t)$
- **Theorem:** the one-parameter subgroups $c_V(t)$ which have $c'(0) = V \in T_eG$ are in one-to-one correspondence with left-invariant vector fields $X_V \in L(G)$
- **Proof:**
  - A one-parameter subgroup $c(t)$ defines a vector field $X\in TG$ along $c(t)$ by $c'(t) = X(c(t))$
  - We now claim that X is left-invariant along $c(t)$
  - Consider the pushforward $c_*:T_t\mathbb{R}\to T_{c(t)}G$ defined by $c'(t) = c_*\left(\frac{d}{dt}\right)_t$
  - From tutorial sheet, we know that $\frac{d}{dt}\in L(\mathbb{R})$, and so $\left(\frac{d}{dt}\right)_t = L_{t*}\left(\frac{d}{dt}\right)_0$, which gives $c'(t) = (cL_t)_*\left(\frac{d}{dt}\right)_0$
  - Now, consider $(cL_t)_*$ acting on arbitrary $s\in\mathbb{R}$; we have

$$(cL_t)(s) = c(t+s) = c(t)c(s) = (L_{c(t)}c)(s) \implies cL_t = L_{c(t)}c \implies (cL_t)_* = (L_{c(t)})_*c_*$$

  - Putting this all together, we have

$$c'(t) = (L_{c(t)})_*c'(0)$$

    that is, $X_{c(t)} = (L_{c(t)})_*X_e$, which implies X is left-invariant along $c(t)$
  - For the converse, let $X\in L(G)$ be a left-invariant vector field
  - It has a unique integral curve $c:\mathbb{R}\to G$ through the identity, so $c(0) = e$ and $c'(t) = X(c(t))$
  - This curve in fact is a one-parameter subgroup
  - To show this, consider the integral curves $\gamma_1(t) = c(t+s)$ and $\gamma_2(t) = c(s)c(t) = L_{c(s)}c(t)$; we show these are equivalent
  - From the definition of the curves, we have immediately that $\gamma_1'(t) = X(\gamma_1(t))$
  - Moreover:

$$\begin{aligned}\gamma_2'(t) = \gamma_{2*}\left(\frac{d}{dt}\right)_t &= (L_{c(s)}c)_*\left(\frac{d}{dt}\right)_t = L_{c(s)*}c'(t)\\ &= L_{c(s)*}X(c(t)) = X(c(s)c(t)) = X(\gamma_2(t))\end{aligned}$$

    and so uniqueness of integral curves, $\gamma_1 = \gamma_2$ and we are done ■
- Note the expression $c'(t) = (L_{c(t)})_*c'(0)$ - for matrix groups, this is just the equation $c'(t) = c(t)c'(0)$, and so $c(t) = \exp(tc'(0))$ where exp is the matrix exponential
- This motivates the *exponential map*, given by $\exp: T_eG\to G$ where $\exp(V) = c_V(1)$, where c_V is the one-parameter subgroup generated by $V\in T_eG$
- **Theorem:** the one-parameter subgroup generated by $V\in T_eG$ is $C_V(t) = \exp(tV)$

- **Proof:**
  - Let $s \in \mathbb{R}$, $s \neq 0$; the curve $\tilde{c}(t) = c_V(st)$ is a one-parameter subgroup
  - It has tangent vector $\tilde{c}'(0) = sc'_V(0) = sV$
  - Therefore $\tilde{c}(t)$ is the one-parameter subgroup generated by sV, and by uniqueness $\tilde{c}(t) = c_{sV}(t)$
  - Then by definition, $\exp(sV) = c_{sV}(1) = \tilde{c}(1) = c_V(s)$ ■
- Note the notational fudging here: when we write the tangent vector $\gamma'(0) \in T_pM$ to a curve γ on manifold M, we often use shorthand $\gamma'(0) = \frac{d}{dt}\gamma(t)\big|_{t=0}$ to mean $\gamma'(0)(f) = \frac{d}{dt}f(\gamma(t))\big|_{t=0}$ for any $f \in C^\infty(M)$
- This means that if $\phi\colon M \to N$ is a smooth map, we can write $\phi_*\gamma'(0) = \frac{d}{dt}\phi \circ \gamma(t)\big|_{t=0}$
- The following corollaries (left unproven as they're quite easy lol) show why the map exp is called the exponential map
- **Corollary:** for $g \in G$, $V \in T_eG$, and $t, s \in \mathbb{R}$, we have:
  1. $\exp(tV)\exp(sV) = \exp((t+s)V)$
  2. $\frac{d}{dt}\exp(tV)\big|_{t=0} = V$
  3. $(X_V)_g = \frac{d}{dt}g\exp(tV)\big|_{t=0}$
  4. For matrix groups, exp coincides with the matrix exponential $\exp(V) = \sum_{k=0}^{\infty}\frac{1}{k!}V^k$
- We now discuss adjoint-adjacent things
- Given $g \in G$, we define the smooth map $Ad_g\colon G \to G$ by $Ad_g h = ghg^{-1}$ for all $h \in G$
- This clearly obeys $Ad_{gh} = Ad_g \circ Ad_h$, and $Ad_g(e) = e$ for any $g \in G$
- Therefore, the derivative of Ad_g at the identity is $ad_g = (Ad_g)_*\colon T_eG \to T_eG$
- Now for any $V \in T_eG$, we have:

$$(ad_g \circ ad_h)(V) = (Ad_g)_*(Ad_h)_*V = (Ad_g \circ Ad_h)_*V = (Ad_{gh})_*V = ad_{gh}V \tag{4.16}$$

- This means that $ad_g \circ ad_h = ad_{gh}$ for all $g, h, \in G$
- Now, recall that a *representation* of a group G on a vector space V is a homomorphism $\rho\colon G \to GL(V)$ where $GL(V)$ is the group of all invertible linear maps $V \to V$; in particular, $\rho(gh) = \rho(g) \circ \rho(h)$ for all $h, g, \in G$
- We therefore have that ad_g defines a representation of G on its Lie algebra $\mathfrak{g} = T_eG$
- This motivates the definition of the *adjoint representation* of G on its Lie algebra $\mathfrak{g}$, given by the representation $ad\colon G \to GL(\mathfrak{g})$ defined by $ad(g) = ad_g$
- Explicitly, for any $V \in T_eG$, the derivative ad_g is given by

$$ad_gV = \frac{d}{dt}(Ad_g c_V(t))\bigg|_{t=0} = \frac{d}{dt}(g\exp(tV)g^{-1})\bigg|_{t=0} \tag{4.17}$$

  where $c_V(t) = \exp(tV)$ is the one-parameter subgroup generated by V
- For matrix groups, the adjoint representation is very useful as it reduces to matrix multiplication

- Explicitly, we mean that

$$Ad_g c_V(t) = g\exp(tV)g^{-1} = \exp(tgVg^{-1}) \tag{4.18}$$

which implies that

$$ad_g V = \frac{d}{dt}\exp(tgVg^{-1})\bigg|_{t=0} = gVg^{-1} \tag{4.19}$$

where the RHS is just ordinary matrix multiplication

- Now consider taking a further derivative of the adjoint representation at the identity: $ad_* : \mathfrak{g} \to End(\mathfrak{g})$ (where we identify the tangent space of $GL(\mathfrak{g})$ at the identity with the endomorphisms of V)

- **Lemma:** for any $X, Y \in \mathfrak{g}$, we have $ad_{*X}Y = [X, Y]$ and $[ad_{*X}, ad_{*Y}] = ad_{*[X,Y]}$ where $ad_{*X} = ad_*(X)$

- **Proof:**
  - For any $X, Y \in \mathfrak{g}$, the derivative is given by

$$ad_{*X}(Y) = \frac{d}{dt}ad_{\exp(tX)}Y\bigg|_{t=0}$$

  - For a matrix group, this is easy to evaluate using the above corollary, whereupon we find

$$ad_{*X}(Y) = \frac{d}{dt}\exp(tX)Y\exp(-tX)\bigg|_{t=0} = [X, Y]$$

    where the RHS is the matrix commutator
  - This actually holds for all groups (not proved)
  - Finally, the Jacobi identity implies

$$[ad_{*X}, ad_{*Y}]Z = [X, [Y, Z]] - [Y, [X, Z]] = [[X, Y], Z] = ad_{*[X,Y]}Z$$

    and we are done ■

- Now, a *representation of a Lie algebra* $\mathfrak{g}$ on vector space V is a Lie algebra homomorphism $\phi : \mathfrak{g} \to End(V)$; that is, a linear map such that $\phi([X, Y]) = [\phi(X), \phi(Y)]$

- The above lemma shows that ad_* is a Lie algebra homomorphism on $V = \mathfrak{g}$ - a representation of $\mathfrak{g}$ on itself

- This defines the *adjoint representation of a Lie algebra* $\mathfrak{g}$ on itself, given by $ad_* : \mathfrak{g} \to End(\mathfrak{g})$

4.3 Group Actions on Manifolds

- For this section, G denotes a Lie group and M a smooth manifold

- We define a *left action* of G on M as a smooth map

$$\phi : G \times M \to M, \quad (g, p) \mapsto g \cdot p \tag{4.20}$$

obeying the properties that:

  1. $g_1 \cdot (g_2 \cdot p) = (g_1 g_2) \cdot p$ for all $g_1, g_2 \in G$ and $p \in M$
  2. $e \cdot p = p$ for all $p \in M$

- We can similarly define a right action if we like
- Inspired by this, given a left action and $g \in G$ we can define an associated diffeomorphism $\phi_g : M \to M$ by $\phi_g(p) = g \cdot p$
- In this notation, property (1) becomes $\phi_{g_1} \circ \phi_{g_2} = \phi_{g_1 g_2}$, so the map $g \mapsto \phi_g$ is a group homomorphism from G to the group of diffeomorphisms of G
- Simple examples of left actions of groups would be $GL(n, \mathbb{R})$ acting on $\mathbb{R}^n$ by matrix multiplication (i.e. $\phi(M, x) = Mx$ for $M \in GL(n, \mathbb{R})$ and $x \in \mathbb{R}^n$) or $O(n)$ acting on the unit n-sphere $S^{n-1} = \{x \in \mathbb{R}^n \,|\, |x| = 1\}$ similarly
- We can define three important properties of left actions; we say a left action ϕ of G on M is:
  - *Effective* if $\phi_g(p) = p$ for all $p \in M$ implies $g = e$. Essentially, the only element of G which acts trivially on M is the identity
  - *Transitive* if for all $p, q \in M$, there exists $g \in G$ such that $g \cdot p = q$. This says that any two points in M are related by the action of an element of G, and if this holds, we say M is a *homogenous space*
  - *Free* if every $g \neq e$ has no fixed points; that is, if $g \cdot p = p$ implies $g = e$
- There's a theorem (not proved) which says that given a non-effective group action, we can always define an effective one
- As some examples, the action of $GL(n, \mathbb{R})$ on $\mathbb{R}^n$ and of $O(n)$ on S^{n-1} are effective and transitive, but not free
- An example of an effective, transitive, and free action would be a Lie group G acting on itself by left-multiplication
- We now define the *orbit* of $p \in M$ under G as

$$G \cdot p = \{g \cdot p \in M \,|\, g \in G\} \tag{4.21}$$

- We think of this as the set of points in M related to p by an element of G
- Note that ϕ is transitive iff $M = G \cdot p$ for all $p \in M$
- The *stabiliser/isotropy subgroup* of $p \in M$ is defined as

$$G_p = \{g \in G \,|\, g \cdot p = p\} \tag{4.22}$$

- So ϕ is free iff $G_p = \{e\}$ for all $p \in M$
- **Theorem:** G_p is a Lie subgroup of G
- **Proof:**
  - Let $p \in M$, and define a smooth map by $\alpha_p : G \to M$ by $\alpha_p(g) = g \cdot p$
  - Then, $G_p = \{g \in G \,|\, \alpha_p(g) = p\}$, which says that G_p is the inverse image $\alpha_p^{-1}(p)$ of the point p
  - Therefore, G_p is a closed subgroup of G and hence a Lie group ■
- Now, consider a closed subgroup H of G; H acts on G by right-multiplication $R_h g = gh$

- The orbits of g under H are called the *left-cosets*, denoted gH
- The set of left-cosets G/H is in fact a smooth manifold (not proved), and G acts on G/H by a left action given by $g_1(g_2H) = (g_1g_2)H$
- In this language, we can reformulate the *orbit-stabiliser theorem*
- **Theorem:** given a left action of G on M, there is a diffeomorphism for each $p \in M$ given by

$$\psi : G \cdot p \to G/G_p, \quad \psi(g \cdot p) = gG_p \tag{4.23}$$

- Note that $\psi(g \cdot p) = g\psi(p)$

5 Gauge Theory

5.1 Integration on Manifolds

- For this section, M is an n-dimensional manifold
- Recall that M is *orientable* if there exists a nowhere vanishing, top-ranked form $\mu \in \Omega^n(M)$, called a *volume form/orientation*
- Note that since $\dim(\Omega^n(M)) = 1$, if we have two orientations μ and μ', we must have $\mu' = f\mu$ for $f \in C^\infty(M)$ and $f \neq 0$ on M
- This means we can classify orientations: $f > 0$ gives equivalent orientations, whereas $f < 0$ gives inequivalent ones
- So orientable M admits two inequivalent orientation classes
- Orientability is equivalent to existence of an *oriented atlas*; that is, an atlas where the Jacobian of all transition functions has positive determinant
- To see this equivalence, choose an *oriented chart* (U, ϕ), $\phi = (x^i)$ and take $\mu \in \Omega^n(U)$, so we can write
$$\mu = g\, dx^1 \wedge \ldots \wedge dx^n \tag{5.1}$$
with $g > 0$ on U (always possible due to continuity)
- Now, suppose we have an overlapping chart (V, ψ), $\psi = (y^i)$; the transition functions are $y(p) = \psi \circ \phi^{-1}(x(p))$ for all $p \in U \cap V$
- Therefore on this overlap, we have:
$$\mu = g'\, dy^1 \wedge \ldots \wedge dy^n = g' \det\left(\frac{\partial y}{\partial x}\right) dx^1 \wedge \ldots \wedge dx^n \tag{5.2}$$
since $dy^i = \frac{\partial y^i}{\partial x^j} dx^j$
- In particular, we can choose $\det(\partial y/\partial x) > 0$ so $g > 0$ implies $g' > 0$ on $U \cap V$, and so we can extend μ to all V such that the component $g' > 0$
- We can repeat this argument until we cover M with μ, and so we get a nowhere vanishing n-form on M
- Conversely, if $\mu \in \Omega^n(M)$ is nowhere vanishing, then we can always arrange the components of μ to be positive on overlapping charts, and since $g = \det(\partial y/\partial x)g'$, we can construct an oriented atlas
- We can now define the *integral* of $\alpha \in \Omega^n(U)$ on an oriented chart (U, ϕ); we can write $\alpha = a\, dx^1 \wedge \ldots \wedge dx^n$ for $a \in C^\infty(U)$, and then the integral is
$$\int_U \alpha = \int_{\phi(U)} (a \circ \phi^{-1})(x) dx \ldots dx^n \tag{5.3}$$
where the RHS is the usual integral on $\phi(u) \subset \mathbb{R}^n$
- Importantly, this definition is chart-independent

- If (U, ψ) is a different oriented chart, then $\alpha = a'\, dy^1 \wedge \ldots \wedge dy^n$, where $a = a' \det(\partial y/\partial x)$, and so:

$$\begin{aligned} \int_{\psi(U)} (a' \circ \psi^{-1})(y) dy^1 \ldots dy^n &= \int_{\phi(U)} (a' \circ \phi^{-1})(x) \det\left(\frac{\partial y}{\partial x}\right) dx^1 \ldots dx^n \\ &= \int_{\phi(U)} (a \circ \phi^{-1})(x) dx^1 \ldots dx^n \end{aligned} \tag{5.4}$$

  where the first line is just the usual change of variables formula

- We also generalise this definition to define integration of an n-form over all of M

- To do this, we take an open cover $\{U_\alpha\}$ of M and define a *partition of unity* of M: a set of functions $\rho_\alpha \in C^\infty(M)$ such that

  1. $\rho_\alpha|_{M/U_\alpha} = 0$
  2. $\sum_\alpha \rho_\alpha = 1$
  3. $0 \le \rho_\alpha \le 1$

- These always exist if M is *paracompact*, which we assume to be true in physics lol

- Then, the integral of $\omega \in \Omega^n(M)$ is

$$\int_M \alpha = \sum_\alpha \int_{U_\alpha} \rho_\alpha \omega \tag{5.5}$$

- It can be shown that this definition is independent of the partition of unity, and of the atlas

- **Theorem:** (Stokes) Let M be an n-dimensional oriented manifold, and $\omega \in \Omega^{n-1}(M)$ has compact support. Then:

$$\int_M d\omega = 0 \tag{5.6}$$

- Note that if M is compact and oriented, then $\int_M d\omega = 0$ for any $\omega \in \Omega^{n-1}(M)$

- Heuristically, this means that any volume form μ on M is not exact: if $\mu = d\omega$, then $\int_M \mu = 0$, but $\int_M \mu = \sum_\alpha \int_{U_\alpha} \rho_\alpha \mu > 0$ since $\rho_\alpha \ge 0$ and $g_\alpha > 0$

- Given a volume form μ, we can also define integration of functions $f \in C^\infty(M)$ by

$$\int_M f = \int_M f\mu \tag{5.7}$$

- But note: on an arbitrary oriented manifold, there is no canonical choice of a volume form

- We have seen so far that a $2n$-dimensional symplectic manifold (M, ω) admits a natural volume form ω^n, but another important class of manifolds admitting orientations are the *pseudo-Riemannian manifolds*

- These are n-dimensional manifold M equipped with a smooth family of non-degenerate symmetric bilinear forms $g_p : T_pM \times T_pM \to \mathbb{R}$ for all $p \in M$; this defines a $(0, 2)$ non-degenerate symmetric tensor field g on M called the *metric tensor*, and we write (M, g)

- The metric has a *signature* s, given by the number of negative eigenvalues

- In a coordinate basis, we write $g = g_{ij} dx^i dx^j$, where formally $dx^i dx^j = \frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i)$
- Simple examples would be the Euclidean metric $\delta = \delta_{ij} dx^i dx^j$ with $s = 0$ on $\mathbb{R}^n$, or the Minkowski metric $\eta = \eta_{\mu\nu} dx^\mu dx^\nu$ where $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, \ldots, 1)$ on $\mathbb{R}^n$
- Non-degeneracy guarantees the existence of a $(2, 0)$ tensor field called the *inverse metric*, which we denote g^{ab}, so $g^{ab} g_{bc} = \delta^a_c$
- Formally, this defines a musical isomorphism $T_p M \simeq T_p^* M$ just as for symplectic spaces
- The inverse metric also defines an *inner product of k-forms*: given $\alpha, \beta \in \Omega^k(M)$, the inner product is

$$\langle \alpha, \beta \rangle = \frac{1}{k!} \alpha_{a_1 \ldots a_k} \beta_{b_1 \ldots b_k} g^{a_1 b_1} \ldots g^{a_k b_k} \tag{5.8}$$

- Now, if M is orientable (which we assume), there is a natural nowhere vanishing n-form which we denote $\mathrm{dvol} \in \Omega^n(M)$, which is defined by the metric
- In a chart, this is explicitly:

$$\mathrm{dvol} = \sqrt{|\det g_{ij}|}\, dx^1 \wedge \ldots \wedge dx^n \tag{5.9}$$

- This is easily verified to be chart-independent, and can be extended to a volume form on M

5.2 Hodge Star Operator and Maxwell's Equations

- Maxwell's equations are most neatly expressed in terms of differential forms; this has the advantage that they can be formulated on any pseudo-Riemannian manifold
- We now define the *Hodge star operator*; let (M, g) be a pseudo-Riemannian manifold of dimension n, then $\star \colon \Omega^k(M) \to \Omega^{n-k}(M)$ defined by

$$\alpha \wedge \star\beta = \langle \alpha, \beta \rangle \,\mathrm{dvol} \tag{5.10}$$

  for all $\alpha, \beta \in \Omega^k(M)$ is the Hodge star
- While this is elegant, it is not immediately useful
- Let e_i, $i = 1, \ldots, n$ be a local orthonormal frame on (M, g), so $g_{ij} = g(e_i, e_j) = \eta_{ij}$ where $\eta_{ij} = \eta_i \delta_{ij}$ (no sum) and $\eta_1 = \ldots = \eta_s = -1$ and $\eta_{s+1} = \ldots = \eta_n = 1$; the dual vectors θ^i are also orthonormal $\langle \theta^i, \theta^j \rangle = \eta^{ij}$ where $\eta^{ij} = \eta_i \delta^{ij}$ (no sum)
- The volume form in an orthonormal frame is

$$\mathrm{dvol} = \theta^1 \wedge \ldots \wedge \theta^n \tag{5.11}$$

  since $\det g_{ij} = \pm 1$ in such a basis
- A basis of differential k-forms is given by

$$\theta^I = \theta^{i_1} \wedge \ldots \wedge \theta^{i_k} \tag{5.12}$$

  where $I = (i_1, \ldots, i_k)$ is a multi-index defined by $i_1 < \ldots < i_k$
- It can be checked that these are orthonormal with respect to the inner product on k-forms $\langle \theta^I, \theta^J \rangle = \eta^{IJ}$ where $\eta^{IJ} = \eta^{i_1 j_1} \ldots \eta^{i_k j_k}$

- Therefore, $\theta^I \wedge \star\theta^J = \eta^{IJ}\text{dvol}$, and so

$$\star\theta^I = \eta_I \sigma_I \theta^{\bar{I}} \tag{5.13}$$

where $\eta_I = \eta_{i_1} \dots \eta_{i_k}$, $\bar{I} = (\bar{i}_1, \dots, \bar{i}_{n-k})$ is the unique complementary multi-index to I and σ_I is the sign of the permutation $(I, \bar{I})$

- This gives us the action of $\star$ on an orthonormal basis of k-forms, and so we can compute its action on any k-form
- Note that $\star^2 : \Omega^k(M) \to \Omega^k(M)$
- **Lemma:** Let (M, g) be an n-dimensional pseudo-Riemannian manifold; on $\Omega^k(M)$, the operator $\star^2 = (-1)^{s+k(n-k)}$ where s is the signature of g
- **Proof:**
  - Working in an orthonormal basis as above:

$$\star^2\theta^{\bar{I}} = \eta_I \sigma_I \star \theta^{\bar{I}} = \eta_I \eta_{\bar{I}} \sigma_I \sigma_{\bar{I}} \theta^I \tag{5.14}$$

  - But, $\eta_I \eta_{\bar{I}} = (-1)^s$ is the product of diagonal elements of η_{ij}, and $\sigma_I \sigma_{\bar{I}} = (-1)^{k(n-k)}$ is the sign of the permutation taking $(I, \bar{I}) \to (\bar{I}, I)$, and we are done ■
- The *adjoint* of exterior derivative d is the operator $d^\dagger : \Omega^k(M) \to \Omega^{k-1}(M)$ defined by

$$\int_M \langle \alpha, d\beta \rangle \,\text{dvol} = \int_M \langle d^\dagger \alpha, \beta \rangle \,\text{dvol} \tag{5.15}$$

for all compactly supported $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^{k-1}(M)$

- **Lemma:** The adjoint of d is explicitly $d^\dagger = (-1)^{s+1+nk+n} \star d \star$ on $\Omega^k(M)$
- **Proof:**
  - For any $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^{k-1}(M)$ compactly supported, we have:

$$\begin{aligned}
\int_M \langle \alpha, d\beta \rangle \,\text{dvol} &= \int_M \langle d\beta, \alpha \rangle \,\text{dvol} = \int_M d\beta \wedge \star\alpha \\
&= \int_M d(\beta \wedge \star\alpha) - (-1)^{k-1} \beta \wedge d \star \alpha \\
&= (-1)^k \int_M \beta \wedge d \star \alpha \\
&= (-1)^k (-1)^{s+(n-k+1)(k-1)} \int_M \beta \wedge \star(\star d \star \alpha) \\
&= (-1)^{s+1+nk+n} \int_M \langle \beta, \star d \star \alpha \rangle \,\text{dvol}
\end{aligned} \tag{5.16}$$

  where we used Stokes', inserted $\star^2 = (-1)^{s+k'(n-k')}$ acting on the $k' = n - k + 1$ form $d \star \alpha$ and we are done ■
- In Euclidean space, the standard operations from vector calc. are $\nabla \cdot \mathbf{X} = \star d \star X$ and $\nabla \times \mathbf{X} = \star dX$, where the components of $\mathbf{X} \in \mathcal{X}(\mathbb{R}^3)$ are identified with those of one-form $X \in \Omega^1(\mathbb{R}^3)$ since $X_i = \delta_{ij} X^j$
- We can now write down Maxwell's equations on any pseudo-Riemannian manifold (M, g)

- The electric and magnetic fields can be packaged into the Maxwell field strength tensor $F \in \Omega^2(M)$, and the Maxwell equations in vacuum are then just

$$d \star F = 0 \quad \text{and} \quad dF = 0 \tag{5.17}$$

- Note the first is equivalent to $d^\dagger F = 0$, and the latter is the Bianchi identity

- Now consider the Maxwell equations on Minkowski spacetime $(\mathbb{R}^4, \eta)$, where in Cartesians (t, x^1, x^2, x^3), the metric is $\eta = -dt^2 + dx^i dx^i$

- In general, we can decompose any $F \in \Omega^2(\mathbb{R}^4)$ as

$$F = -E_j dt \wedge dx^j + \frac{1}{2} B_i \epsilon_{ijk} dx^j \wedge dx^k \tag{5.18}$$

- After some calculation, we find that the first of (5.17) is equivalent to

$$\partial_i E_i = 0, \quad \epsilon_{ijk} \partial_j B_k - \partial_t E_i = 0 \tag{5.19}$$

  and the second to

$$\partial_i B_i = 0, \quad \epsilon_{ijk} \partial_j E_k + \partial_t B_i = 0 \tag{5.20}$$

  which is just Maxwell's equations as usual

- Returning back to (M, g), note that F is a closed 2-form since $dF = 0$, so by the Poincare lemma we can always write it locally as $F = dA$ for some 1-form A, called a *gauge potential*

- Precisely, let $\{U_\alpha\}$ be an open cover of M, so we can write $F = dA_\alpha$ on each U_α for some $A_\alpha \in \Omega^1(U_\alpha)$

- A local gauge potential is only uniquely defined up to a local *gauge transformation*:

$$A'_\alpha = A_\alpha + d\lambda_\alpha \tag{5.21}$$

  for $\lambda_\alpha \in C^\infty(U_\alpha)$

- On a non-trivial overlap $U_{\alpha\beta} = U_\alpha \cap U_\beta$, we have $dA_\alpha = dA_\beta$, and so

$$A_\alpha = A_\beta + d\lambda_{\alpha\beta} \tag{5.22}$$

  where $\lambda_{\alpha\beta} \in C^\infty(U_{\alpha\beta})$

- This means that local gauge fields on overlapping charts are related by gauge transformations, which persists in more general gauge theories

5.3 Principal Bundles

- *Fibre bundles* are manifolds which locally look like a product space of a base manifold M and a fibre manifold F

- *Vector bundles* are fibre bundles where the fibre manifold F is a vector space

- *Principal bundles* are fibre bundles where F is a Lie group - these are the central object of gauge theory, so we study them in detail

- Formally, a principal (fibre) bundle is the following collection of data:
  - A (smooth) manifold P called the *total space*

- A Lie group G and a free right action on P:

$$P \times G \to P, \quad (p,g) \mapsto pg \tag{5.23}$$

- The orbit space $M = P/G$ is a manifold called the *base space*; there is a projection map

$$\pi : P \to M, \quad p \mapsto p \cdot G \tag{5.24}$$

  for all $p \in P$, so $\pi(p) = \pi(pg)$ for all $p \in P$, $g \in G$

- The *fibre* over $m \in M$ is $\pi^{-1}(m) = \{pg \,|\, g \in G, \pi(p) = m\}$. Note this means that $\pi^{-1}(m)$ is diffeomorphic to G for every $m \in M$, and the right action of G on $\pi^{-1}(m)$ is free and transitive
- A *local trivilisation* is an open cover $\{U_\alpha\}$ of M equipped with diffeomorphisms

$$\Psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times G, \quad \Psi_\alpha(p) = (\pi(p), g_\alpha(p)) \tag{5.25}$$

  where $p \in P$, the smooth maps $g_\alpha : \pi^{-1}(U_\alpha) \to G$ satisfy $g_\alpha(pg) = g_\alpha(p)g$, and $g_\alpha : \pi^{-1}(m) \to G$ are diffeomorphisms for all $m \in U_\alpha$

- Typical notations for a principal bundle are $P \xrightarrow{\pi} M$, $\pi : P \to M$, or $G \to P \xrightarrow{\pi} M$ etc.

- Consider the nonempty overlaps $U_{\alpha\beta} = U_\alpha \cap U_\beta$ of M; for $p \in \pi^{-1}(m)$ with $m \in U_{\alpha\beta}$, we have two different trivilisations:

$$\Psi_\alpha(p) = (m, g_\alpha(p)) \quad \text{and} \quad \Psi_\beta(p) = (m, g_\beta(p)) \tag{5.26}$$

- Therefore, $g_\alpha(p) = \tilde{g}_{\alpha\beta}(p) g_\beta(p)$ for some $\tilde{g}_{\alpha\beta}(p) \in G$, and so

$$\tilde{g}_{\alpha\beta}(p) = g_\alpha(p) g_\beta(p)^{-1} \tag{5.27}$$

  is a constant on each orbit: $\tilde{g}_{\alpha\beta}(pg) = \tilde{g}_{\alpha\beta}(p)$ for all $g \in G$

- This defines the *transition functions* of a principal G-bundle, which are the maps

$$g_{\alpha\beta} : U_{\alpha\beta} \to G, \quad g_{\alpha\beta}(m) = \tilde{g}_{\alpha\beta}(p) \tag{5.28}$$

  for any $m \in U_{\alpha\beta}$ and $p \in \pi^{-1}(m)$

- **Lemma:** The transition functions satisfy the *cocycle conditions*:

  1. $g_{\alpha\beta}(m) g_{\beta\alpha}(m) = e$ for all $m \in U_{\alpha\beta}$
  2. $g_{\alpha\gamma}(m) g_{\gamma\beta}(m) g_{\beta\alpha}(m) = e$ for all $m \in U_{\alpha\beta\gamma}$

  where $U_{\alpha\beta\gamma} = U_\alpha \cap U_\beta \cap U_\gamma$ are nonempty

- This follows immediately from the definitions

- Note that given a manifold M with open cover $\{U_\alpha\}$ and functions $\{g_{\alpha\beta} : U_{\alpha\beta} \to G\}$ satisfying the cocycle conditions, we can construct a principal bundle

- Define the equivalence relation $(m, g) \sim (m, g_{\alpha\beta}(m)g)$ for all $m \in U_{\alpha\beta}$, $g \in G$; then $P = \bigcap_\alpha (U_\alpha \times G)/\sim$ is a principal bundle with projection π onto the first factor, and right-action of G on the second factor

- We say a principal G-bundle is *trivial* is there is a diffeomorphism

$$\Psi : P \to M \times G, \quad p \mapsto (\pi(p), \gamma(p)) \tag{5.29}$$

  where $\gamma(pg) = \gamma(p)g$ (think of this as a global trivilisation)

- For example, $P = M \times G$ is a trivial principal bundle with π projecting onto the first factor and right action given by G on the second factor

- A *section* of P is a smooth map $s\colon M \to P$ with $\pi \circ s = id$; a smooth assignment to every $m \in M$ of a $p \in \pi^{-1}(m)$

- **Theorem:** A principal bundle is trivial if and only if it admits a section

- **Proof:**
  - First, suppose we have a section $s\colon M \to P$
  - Define the map $\Psi(p) = (\pi(p), \gamma(p))$ where $\gamma(p) \in G$ is uniquely defined by $p = s(\pi(p))\gamma(p)$
  - Note that:
  $$s(\pi(p))\gamma(p)g = pg = s(\pi(pg))\gamma(pg) = s(\pi(p))\gamma(pg) \tag{5.30}$$
  and so $\gamma(pg) = \gamma(p)g$
  - Therefore, Ψ is a global trivilisation
  - Conversely, suppose $\Psi\colon P \to M \times G$ is a global trivilisation, and define $s(m) = \Psi^{-1}(m, e) \in P$ for any $m \in M$
  - Then:
  $$\Psi(s(m)) = (\pi(s(m)), \gamma(s(m))) = (m, e) \tag{5.31}$$
  which implies that $\pi \circ s = id$ and $(\gamma \circ s)(m) = e$; the first of these means s is a section ■

- Note that this is specific to principal bundles: other fibre bundles can admit sections without being trivial

- By definition, P is locally trivial so local sections always exist

- That is, given a local trivilisation we can define smooth maps $s_\alpha\colon U_\alpha \to \pi^{-1}(U_\alpha)$ where $s_\alpha(m) = \Psi_\alpha^{-1}(m, e)$, and then the same calculation as in the proof shows that $\pi \circ s_\alpha = id$ and $(g_\alpha \circ s_\alpha)(m) = e$

- These are called the *canonical local sections*

- The converse also holds: given local sections $s_\alpha\colon U_\alpha \to \pi^{-1}(U_\alpha)$ with $\pi \circ s_\alpha = id$, we can define a local trivilisation

- For every $m \in U_\alpha$ and $p \in \pi^{-1}(m)$, there is a unique $g_\alpha(p) \in G$ such that $p = s_\alpha(m)g_\alpha(p)$ and $\Psi_\alpha(p) = (\pi(p), g_\alpha(p))$ is a local trivilisation

- So local trivilisations are equivalent to local sections

- **Lemma:** On any non-trivial overlap $U_{\alpha\beta}$, the canonical local sections are related by $s_\beta(m) = s_\alpha(m)g_{\alpha\beta}(m)$ for all $m \in U_{\alpha\beta}$

5.4 Connections on Principal Bundles

- For this section, we refer to principal bundle $P \xrightarrow{\pi} M$ with Lie group G, and $m \in M$ is a point on the base, and $p \in \pi^{-1}(m)$ is in the fibre over m

- The push-forward is (as usual) $\pi_*\colon T_pP \to T_mM$, given by $(\pi_* v_p)(f) = v_p(f \circ \pi)$ for $f \in C^\infty(M)$ and $v_p \in T_pP$

- We define the *vertical subspace* $V_p \subset T_pP$, given by $V_p = \ker \pi_*$
- We say a vector field $v \in \mathcal{X}(P)$ is *vertical* if $v_p \in V_p$ for all $p \in P$
- Note that if $\dot{\gamma} \in V_{\gamma(t)}$, then $0 = \pi_*\dot{\gamma} = \frac{d}{dt}\pi(\gamma(t))$, and so $\pi(\gamma(t))$ is constant: $\gamma(t)$ lies in a fixed fibre
- **Lemma:** If v_1, v_2 are vertical vector fields, then the Lie bracket $[v_1, v_2]$ is also a vertical vector field on P
- **Proof:**
  - By the standard property of the push-forward:
  $$\pi_*[v_1, v_2] = [\pi_* v_1, \pi_* v_2] = 0$$
  and we are done ■
- This motivates the definition of the *vertical distribution* $V = \cup_{p \in P} V_p \subset TP$, which is G-invariant as follows
- Note that $\pi(R_g p) = \pi(pg) = \pi(p)$, and so $\pi \circ R_g = \pi$; thus let $v_p \in V_p$ and consider $R_{g*}v_p \in T_{pg}P$
- We calculate $\pi_*(R_{g*}v_p) = (\pi \circ R_g)_* v_p = \pi_* v_p = 0$, and so $R_{g*}v_p \in V_{pg}$
- We write this G-invariance as $R_{g*}V_p = V_{pg}$
- The vertical subspace is defined naturally by the bundle structure, but there is no canonical complement of V_p in T_pP; a connection allows us to define such a complement
- A *connection* on P is a smooth choice of *horizontal subspace* $H_p \subset T_pP$ complementary to V_p; that is, $T_pP = H_p \oplus V_p$ such that $R_{g*}H_p = H_{pg}$ for all $p \in P$
- The latter property here states that horizontal subspaces define a G-invariant distribution $H \subset TP$
- Now, denote the Lie algebra of G by $\mathfrak{g} = T_eG$ as usual
- For any $X \in \mathfrak{g}$, we define the *fundamental vector field*
  $$\sigma_p(X) = \frac{d}{dt} p \exp(tX) \bigg|_{t=0} \tag{5.32}$$
  where $p\exp(tX)$ is a curve in P through p with tangent vector $\sigma_p(X) \in T_pP$ defined through the exponential map on G
- **Reminder:** a curve $c : \mathbb{R} \to G$ on a Lie group G is a one-parameter subgroup of G if $c(t)c(s) = c(t+s)$, and there is a bijection between one parameter subgroups $c_V(t)$ with $c'(0) = V$ where $X_V \in L(G)$ is a left-invariant vector field
- Then the exponential map $\exp : T_eG \to G$ is defined by $\exp(V) = c_V(1)$ where c_V is the one-parameter group generated by $V \in T_eG$ as above; in particular, $c_V(t) = \exp(tV)$
- In particular, $\exp(tX)$ is a curve in G through e with tangent $X \in T_eG$
- So, we have:
  $$\pi_*\sigma_p(X) = \frac{d}{dt}\pi(p\exp(tX))\bigg|_{t=0} = \frac{d}{dt}\pi(p)\bigg|_{t=0} = 0 \tag{5.33}$$
  where the second equality holds since $\exp(tX) \in G$ and so p and $p\exp(tX)$ are on the same orbit

- Therefore the fundamental vector field is always vertical: $\sigma_p(X) \in V_p$
- **Lemma:** Fundamental vector fields satisfy $R_{g*}\sigma_p(X) = \sigma_{pg}(ad_{g^{-1}}X)$ for $g \in G$, $X \in \mathfrak{g}$, $p \in P$
- **Proof:**
  - Recall the adjoint maps: $Ad_g : G \to G$ is $Ad_g h = ghg^{-1}$ for all $h \in G$, and the derivative of this at the identity is $ad_g = (Ad_g)_* : T_eG \to T_eG$
  - Therefore:

$$\begin{aligned} R_{g*}\sigma_p(X) &= \frac{d}{dt}R_g(p\exp(tX))\bigg|_{t=0} = \frac{d}{dt}p\exp(tX)g\bigg|_{t=0} \\ &= \frac{d}{dt}pgg^{-1}\exp(tX)g\bigg|_{t=0} = \frac{d}{dt}\,Ad_{g^{-1}}\exp(tX)\bigg|_{t=0} \\ &= \frac{d}{dt}pg\exp(t\,ad_{g^{-1}}X)\bigg|_{t=0} = \sigma_{pg}(ad_{g^{-1}}X) \end{aligned}$$

    where we use that $Ad_{g^{-1}}\exp(tX) = \exp(t\,ad_{g^{-1}}X)$ is the one-parameter subgroup generated by $ad_{g^{-1}}X$ and we are done ■

- Note: $\sigma_p : \mathfrak{g} \to V_p$ for each $p \in P$ defines an isomorphism (not proved)
- Thus $\dim V_p = \dim G$, and $\dim H_p = \dim P - \dim G$; H_p can be defined by $\dim G$ linear equations in T_pP (i.e. by $\dim G$ one-forms)
- This motivates the definition of the *connection 1-form* of a horizontal distribution $H \subset TP$: it is a $\mathfrak{g}$-valued 1-form $\omega \in \Omega^1(P;\mathfrak{g})$ defined by

$$\omega(v) = \begin{cases} 0 & \text{if } v \text{ is horizontal} \\ X & \text{if } v = \sigma(X) \end{cases} \tag{5.34}$$

- **Lemma:** The connection 1-form satisfies $R_g^*\omega = ad_{g^{-1}} \circ \omega$
- **Proof:**
  - If v is horizontal, then $\omega(v) = 0$, so since $R_{g*}H_p = H_{pg}$ it follows that $R_{g*}v$ is also horizontal
  - So: $0 = \omega(R_{g*}v) = (R_g^*)(v)$ as required
  - If v is vertical, then we can write $v = \sigma(X)$ for some $X \in \mathfrak{g}$ by the above isomorphism
  - Therefore:

$$\begin{aligned} (R_g^*\omega)(\sigma(X)) &= \omega(R_{g*}\sigma(X)) = \omega(\sigma(ad_{g^{-1}}X)) \\ &= ad_{g^{-1}}(X) = ad_{g^{-1}}\omega(\sigma(X)) \end{aligned}$$

    and we are done ■

- Conversely, it can be shown that any $\omega \in \Omega^1(P;\mathfrak{g})$ such that $\omega(\sigma(X)) = X$ and $R_g^*\omega = ad_{g^{-1}} \circ \omega$ defines a connection $H = \ker\omega \subset TP$
- Given such a connection 1-form $\omega \in \Omega^1(P;\mathfrak{g})$ and a canonical local section $s_\alpha : U_\alpha \to \pi^{-1}(U_\alpha)$, a *local gauge field* is defined as

$$A_\alpha = s_\alpha^*\omega \in \Omega^1(U_\alpha;\mathfrak{g}) \tag{5.35}$$

- **Lemma:** The restriction $\omega|_{\pi^{-1}(U_\alpha)}$ of a connection 1-form ω defined by a canonical local section is equal to
$$\omega_\alpha = ad_{g_\alpha^{-1}} \circ \pi^* A_\alpha + g_\alpha^* \theta \tag{5.36}$$
where $g_\alpha : \pi^{-1}(U_\alpha) \to G$ is defined by the corresponding local trivilisation, and $\theta \in \Omega^1(G;\mathfrak{g})$ is the Maurer-Cartan form on G (not proven)

- **Theorem:** On non-empty overlaps $U_{\alpha\beta}$, local gauge fields are related by
$$A_\alpha = ad_{g_{\alpha\beta}} \circ A_\beta + g_{\beta\alpha}^* \theta \tag{5.37}$$
where $g_{\alpha\beta} : U_{\alpha\beta}$ are the transition functions of P

- **Proof:**
  - $\omega \in \Omega^1(P;\mathfrak{g})$ is globally defined on P, so $\omega_\alpha = \omega_\beta$ on $\pi^{-1}(U_{\alpha\beta})$
  - Equation (5.36) allows us to relate the local gauge fields A_α and A_β on $U_{\alpha\beta}$ as
$$\begin{aligned} A_\alpha = s_\alpha^* \omega_\alpha = s_\alpha^* \omega_\beta &= s_\alpha^*(ad_{g_\beta^{-1}} \circ \pi^* A_\beta + g_\beta^* \theta) \\ &= ad_{(g_\beta \circ s_\alpha)^{-1}} \circ s_\alpha^* \pi^* A_\beta + s_\alpha^* g_\beta^* \theta \\ &= ad_{(g_\beta \circ s_\alpha)^{-1}} \circ A_\beta + (g_\beta \circ s_\alpha)^* \theta \end{aligned}$$
    since $s_\alpha^* \pi^* = (\pi \circ s_\alpha)^* = id$
  - We therefore need to consider map $g_\beta \circ s_\alpha : U_{\alpha\beta} \to G$, which we can write as $g_\beta \circ s_\alpha = (g_\beta g_\alpha^{-1} g_\alpha) \circ s_\alpha = (g_\beta g_\alpha^{-1}) \circ e = g_{\beta\alpha}$
  - Then finally, $g_{\beta\alpha}^{-1} = g_{\alpha\beta}$, and we are done ■

- **Corollary:** For a matrix group G:
$$A_\alpha = g_{\alpha\beta} A_\beta g_{\alpha\beta}^{-1} - dg_{\alpha\beta}\, g_{\alpha\beta}^{-1} \tag{5.38}$$

- **Proof:**
  - For matrix groups, $ad_g X = gXg^{-1}$ and $g_{\alpha\beta}^* \theta = g_{\alpha\beta}^{-1} dg_{\alpha\beta} = -g_{\alpha\beta} dg_{\alpha\beta}^{-1}$ is the pull-back of the Maurer-Cartan form ■

- So, given a collection of 1-forms $A_\alpha \in \Omega^1(U_\alpha;\mathfrak{g})$ that satisfy (5.37) on overlaps, we can reverse all the above logiv to construct a connection 1-form $\omega \in \Omega^1(P;\mathfrak{g})$ by defining $\omega = \omega_\alpha$ on $\pi^{-1}(U_\alpha)$ as in (5.36)

- We therefore get that the following give three equivalent descriptions of a connection on a principal bundle P:
  1. A horizontal distribution $H \subset TP$ that is G-invariant $R_{g*} H_p = H_{pg}$
  2. A 1-form $\omega \in \Omega^1(P;\mathfrak{g})$ satisfying $\omega(\sigma(X)) = X$, and $R_g^* \omega = ad_{g^{-1}} \circ \omega$
  3. A collection of 1-forms $A_\alpha \in \Omega^1(U_\alpha;\mathfrak{g})$ for an open cover $\{U_\alpha\}$ satisfying (5.36) on overlaps

5.5 Gauge Transformations

- We define a *gauge transformation* of a principal bundle P as a diffeomorphism $\Phi\colon P \to P$ such that both $\pi \circ \Phi = \pi$ and $\Phi(pg) = \Phi(p)g$
- The gauge transformations form a group under composition, denoted $\mathcal{G}$
- Note that Φ maps fibres to themselves: if $m \in M$ and $p \in \pi^{-1}(m)$ lies in the fibre above m, then $\Phi(p) \in \pi^{-1}(m)$ and vice versa since $\pi(\Phi(p)) = \pi(p)$
- Suppose we have a local trivilisation $\Psi_\alpha\colon \pi^{-1}(U_\alpha) \to U_\alpha \times G$ with $\Psi_\alpha(p) = (\pi(p), g_\alpha(p))$ with $g_\alpha(pg) = g_\alpha(p)g$; then

$$\Psi_\alpha(\Phi(p)) = (\pi(\Phi(p)), g_\alpha(\Phi(p))) = (\pi(p), \bar{\Phi}_\alpha(p)g_\alpha(p)) \tag{5.39}$$

  where we are implicit in defining $\bar{\Phi}_\alpha(pg) = g_\alpha(\Phi(p))g_\alpha(p)^{-1}$

- The $\bar{\Phi}_\alpha$s are constant on fibres:

$$\bar{\Phi}_\alpha(pg) = g_\alpha(\Phi(pg))g_\alpha(pg)^{-1} = g_\alpha(\Phi(p)g)(g_\alpha(p)g)^{-1} = \bar{\Phi}_\alpha(p) \tag{5.40}$$

  and so they define a smooth map $\Phi_\alpha\colon U_\alpha \to G$ by $\Phi_\alpha(m) = \bar{\Phi}_\alpha(p)$ where $p \in \pi^{-1}(m)$

- We call the Φ_α *local gauge transformations*
- Now, consider non-empty overlaps $U_{\alpha\beta}$; let $m \in U_{\alpha\beta}$ and $p \in \pi^{-1}(U_{\alpha\beta})$, and we can then compute (tediously) that $\Phi_\alpha(m) = g_{\alpha\beta}(m)\Phi_\beta(m)g_{\alpha\beta}(m)^{-1} = Ad_{g_{\alpha\beta}(m)}\Phi_\beta(m)$
- This means that local gauge transformations are related by conjugation under the transition functions
- We also have the converse: given a set $\{\Phi_\alpha \,|\, U_\alpha \to G\}$ satisfying this transformation law, we can reconstruct a global gauge transformation Φ by reversing the above procedure
- This is all a result of $\pi \circ \Phi = \pi$; we now consider the second property $\Phi(pg) = \Phi(p)g$
- Write $pg = R_g p$, and this property then becomes $(\Phi \circ R_g)(p) = (R_g \circ \Phi)(p)$
- Taking a derivative, we therefore deduce $R_{g*}\Phi_* = \Phi_* R_{g*}$
- Inspired by this, we define $H^\Phi = \Phi_* H$, where $H \subset TP$ is a connection; we calculate

$$R_{g*}H^\Phi_{\Phi(p)} = R_{g*}\Phi_* H_p = \Phi_* R_{g*} H_p = \Phi_* H_{pg} = H^\Phi_{\Phi(pg)} = H^\Phi_{\Phi(p)g} \tag{5.41}$$

  and so H^Φ is G-invariant

- In fact, it turns out H^Φ is complementary to vertical subspaces V and hence is a connection
- **Lemma:** $\Phi_*\sigma(X) = \sigma(X)$ for any $X \in \mathfrak{g}$; that is, $\Phi_* V_p = V_{\Phi(p)}$
- **Proof:**
  - We calculate:

$$\Phi_*\sigma_p(X) = \frac{d}{dt}\Phi(p\exp(tX))\bigg|_{t=0} = \frac{d}{dt}\Phi(p)\exp(tX)\bigg|_{t=0} = \sigma_{\Phi(p)}(X)$$

    and we are done ■

- This means $\Phi_* : T_pP \to T_{\Phi(p)}P$ is an isomorphism preserving vertical subspaces, so H^Φ is a horizontal distribution; combined with G-invariance, this establishes it is a connection

- **Lemma:** Let ω be the connection 1-form of a connection $H \subset TP$. Then $\omega^\Phi = (\Phi^*)^{-1}\omega$ is the connection 1-form for H^Φ

- **Proof:**

  - First, note that for all $X \in \mathfrak{g}$ we have

$$\omega^\Phi(\sigma(X)) = \omega(\Phi_*^{-1}\sigma(X)) = \omega(\sigma(X)) = X$$

    where the first equality is from the usual property of pull-backs/push-forwards, the second from the above lemma, and the third by definition of ω

  - On the other hand: $\ker \omega^\Phi$ os given by vector fields $v \in \mathfrak{X}(P)$ satisfying $0 = \omega^\Phi(v) = \omega(\Phi_*^{-1}v)$
  - This is equivalent to $(\Phi_*^{-1}v)_p \in H_p \implies v_{\Phi(p)} \in \Phi_* H_p = H^\Phi_{\Phi(p)}$
  - Thus $\ker \omega^\Phi$ is precisely the horizontal distribution H^Φ ■

- **Theorem:** The local gauge field $A_\alpha^\Phi \in \Omega^1(U_\alpha; \mathfrak{g})$ corresponding to connection 1-form ω^Φ is related to that of ω by

$$A_\alpha^\Phi = ad_{\Phi_\alpha} \circ (A_\alpha - \Phi_\alpha^*\theta) \tag{5.42}$$

  or, for matrix groups

$$A_\alpha^\Phi = \Phi_\alpha A_\alpha \Phi_\alpha^{-1} - d\Phi_\alpha \Phi_\alpha^{-1} \tag{5.43}$$

- **Proof:**

  - Let $m \in U_\alpha$, $p \in \pi^{-1}(m)$, and $q = \Phi^{-1}(p)$ lie in the same fibre (since $\pi \circ \Phi^{-1} = \pi$)
  - Recall the lemma in the last section:

$$\omega_q = ad_{g_\alpha(q)^{-1}} \circ \pi^* A_\alpha + g_\alpha^*\theta$$

  - We can therefore compute, using the lemma above as well:

$$\begin{aligned}\omega_p^\Phi = (\Phi^{-1})^*\omega_q &= ad_{g_\alpha(q)^{-1}} \circ \underbrace{(\Phi^{-1})^*\pi^*}_{=(\pi\circ\Phi^{-1})^*=\pi^*} A_\alpha + (\Phi^{-1})^* g_\alpha^*\theta \\ &= ad_{g_\alpha(q)^{-1}} \circ \pi^* A_\alpha + (g_\alpha \circ \Phi^{-1})^*\theta\end{aligned}$$

  - Consider the map $g_\alpha \circ \Phi^{-1} : \pi^{-1}(U_\alpha) \to G$; we compute

$$(g_\alpha \circ \Phi^{-1})(p) = g_\alpha(q) \quad \text{and} \quad \bar{\Phi}_\alpha(q) = g_\alpha(\Phi(q)) g_\alpha(q)^{-1}$$

    and so

$$g_\alpha(q) = \bar{\Phi}_\alpha(q)^{-1} g_\alpha(p) = \bar{\Phi}_\alpha(p)^{-1} g_\alpha(p)$$

    by constancy of $\bar{\Phi}$ on fibres

  - We can therefore write $g_\alpha \circ \Phi^{-1} = \bar{\Phi}_\alpha^{-1} g_\alpha$, and so

$$(g_\alpha \circ \Phi^{-1})^*\theta = g_\alpha^*\theta - ad_{g_\alpha(p)^{-1}\bar{\Phi}_\alpha(p)} \circ \bar{\Phi}_\alpha^*\theta$$

  - This is easily seen for matrix groups by using $g^*\theta = g^{-1}dg$ and the identity $dg^{-1} = -g^{-1}dgg^{-1}$:

$$(g_\alpha \circ \Phi^{-1})^*\theta = (\bar{\Phi}_\alpha^{-1} g_\alpha)^{-1} d(\bar{\Phi}_\alpha^{-1} g_\alpha) = g_\alpha^{-1} dg_\alpha - g_\alpha^{-1}\bar{\Phi}_\alpha(\bar{\Phi}_\alpha^{-1} d\bar{\Phi}_\alpha)(g_\alpha^{-1}\bar{\Phi}_\alpha)^{-1}$$

– Putting all this together:

$$\begin{aligned}\omega_p^\Phi &= ad_{g_\alpha(p)^{-1}\bar{\Phi}_\alpha(p)} \circ (\pi^* A_\alpha - \bar{\Phi}_\alpha^* \theta) + g_\alpha^* \\ &= ad_{g_\alpha(p)^{-1}} \circ \pi^*(ad_{\Phi_\alpha(m)} \circ A_\alpha - \Phi_\alpha^* \theta) + g_\alpha^* \theta\end{aligned}$$

where we use $\bar{\Phi}_\alpha(p) = \Phi_\alpha(\pi(p)) = \Phi_\alpha(m)$, that ad is a group homomorphism, and $\bar{\Phi}_\alpha^* = (\Phi_\alpha \circ \pi)^* = \pi^* \Phi_\alpha^*$

– On the other hand, the lemma from the last section applied to ω^Φ gives

$$\omega_p^\Phi = ad_{g_\alpha(p)^{-1}} \circ \pi^* A_\alpha^\Phi + g_\alpha^* \theta$$

so comparing the two gives the result

– For matrix groups, we use $\Phi_\alpha^* \theta = \Phi_\alpha^{-1} d\Phi_\alpha$ and $ad_g X = gXg^{-1}$, and we are done ■

- This theorem just states that local gauge fields transform under a gauge transformation

5.6 Curvature

- Given a connection $H \subset TP$ on principal G-bundle $P \xrightarrow{\pi} M$, we define the *horizontal projection map* $h : TP \to TP$ so that for every $p \in P$, we have

$$h_p(v) = \begin{cases} v & \text{if } v \in H_p \\ 0 & \text{if } v \in V_p \end{cases} \tag{5.44}$$

- We can extend this to be defined on forms: define $h_p^* : T_p^* P \to T_p^* P$ so for all $v \in T_p P$ and $\alpha \in T_p^* P$, we have

$$(h_p^* \alpha)(v) = \alpha(h_p(v)) \tag{5.45}$$

- For k-forms, we have $(h^*\alpha)(v_1, \ldots, v_k) = \alpha(hv_1, \ldots, hv_k)$, and note that $h^*(\alpha \wedge \beta) = (h^*\alpha) \wedge (h^*\beta)$

- The *curvature* of a connection is then $\Omega = h^* d\omega \in \Omega^2(P; \mathfrak{g})$, where ω is the connection 1-form

- Note the following for $u, v \in \mathfrak{X}(P)$:

$$\begin{aligned}\Omega(u,v) = (h^* d\omega)(u,v) &= d\omega(hu, hv) \\ &= (hu)(\omega(hv)) - (hv)(\omega(hu)) - \omega([hu, hv]) \\ &= -\omega([hu, hv])\end{aligned} \tag{5.46}$$

since hu, hv are horizontal and so are in $\ker \omega$

- This shows that $\Omega = 0 \iff [hu, hv]$ is horizontal for all $u, v \iff [u, v]$ is horizontal for all horizontal u, v

- **Theorem:** The curvature satisfies the structure equation

$$\Omega = d\omega + \frac{1}{2}[\omega, \omega] \tag{5.47}$$

where $[\cdot, \cdot]$ is both the Lie bracket on $\mathfrak{g}$ and the wedge product on $\Omega^k(P; \mathfrak{g})$

- **Proof:**

- First, note that for a basis $\{V_i\}$ of $\mathfrak{g}$, $\omega = \omega^i V_i$, and so

$$\begin{aligned}[]
[\omega, \omega](u, v) &= [V_i, V_j](\omega^i \wedge \omega^j)(u, v) = [V_i, V_j](\omega^i(u)\omega^j(v) - \omega^i(v)\omega^j(u)) \\
&= 2[V_i, V_j]\omega^i(u)\omega^j(v) = 2[\omega(u), \omega(v)]
\end{aligned}$$

- Therefore, the claim is equivalent to

$$\Omega(u, v) = d\omega(u, v) + [\omega(u), \omega(v)]$$

- We now have three cases to consider
- Case 1: u, v are both vertical, so $u = \sigma(X)$ and $v = \sigma(Y)$ for $X, Y \in \mathfrak{g}$
- Then: $\Omega(u, v) = d\omega(hu, hv) = 0$ since $hu = hv = 0$
- For the RHS, we have:

$$\begin{aligned}
d\omega(u, v) &= u(\omega(\sigma(Y))) - v(\omega(\sigma(X))) - \omega([\sigma(X), \sigma(Y)]) \\
&= u(Y) - v(X) - \omega(\sigma([X, Y])) \\
&= -[X, Y] \\
&= -[\omega(\sigma(X)), \omega(\sigma(Y))] \\
&= -[\omega(u), \omega(v)]
\end{aligned}$$

  where we use that $u(Y) = v(X) = 0$ since X, Y are constant Lie algebra-valued functions, and (not proven) $\sigma([X, Y]) = [\sigma(X), \sigma(Y)]$
- Case 2: u, v are both horizontal, so $hu = u$ and $hv = v$
- Then: $\Omega(u, v) = d\omega(hu, hv) = d\omega(u, v)$ and since $\omega(u) = \omega(v) = 0$, it holds in this case
- Case 3: WLOG say u is horizontal and v is vertical. so $hu = u$ and $v = \sigma(X)$
- Then:

$$\Omega(u, v) = d\omega(u, hv) = u(\omega(hv)) - (hv)(\omega(u)) - \omega([u, hv]) = 0$$

  since $\omega(u) = 0$ and $hv = 0$
- For the RHS:

$$\begin{aligned}
d\omega(u, v) &= u(\omega(v)) - v(\omega(u)) - \omega([u, v]) \\
&= u(X) - \omega([u, v]) = -\omega([u, v]) = 0
\end{aligned}$$

  using that $[u, v]$ is horizontal (try and prove this yourself) ■

- **Theorem:** The curvature satisfies the *Bianchi identity*

$$h^* d\Omega = 0 \tag{5.48}$$

- **Proof:**

  - From the structure equation, we have:

$$h^* d\omega = h^* d(d\omega + \tfrac{1}{2}[\omega, \omega]) = h^*[d\omega, \omega] = [h^* d\omega, h^* \omega] = 0$$

    since $h^* \omega = 0$ and $d[\omega, \omega] = 2[d\omega, \omega]$ ■

- We now define the *local field strengths*

$$F_\alpha = s_\alpha \Omega \in \Omega^2(P; \mathfrak{g}) \tag{5.49}$$

  where $s_\alpha : U_\alpha \to P$ are the canonical local sections

- Applying the curvature structure equation, we find

$$\begin{aligned} F_\alpha = s_\alpha^*\Omega &= s_\alpha^*(d\omega + \tfrac{1}{2}[\omega,\omega]) \\ &= ds_\alpha^*\omega + \tfrac{1}{2}[s_\alpha^*\omega, s_\alpha^*\omega] \\ &= dA_\alpha + \tfrac{1}{2}[A_\alpha, A_\alpha] \end{aligned} \tag{5.50}$$

  since pullbacks commute with exterior derivatives

- **Theorem:** On nonempty overlaps $U_{\alpha\beta}$, we have

$$F_\alpha = ad_{g_{\alpha\beta}} \circ F_\beta \tag{5.51}$$

  or, for matrix groups

$$F_\alpha = g_{\alpha\beta} F_\beta g_{\alpha\beta}^{-1} \tag{5.52}$$

- **Proof:**
  - Recall that

$$A_\alpha = ad_{g_{\alpha\beta}} \circ A_\beta + g_{\alpha\beta}^*\theta$$

    and plug this into the definition of local field strength
  - We also use the (not proved) identity

$$d(ad_{g_{\alpha\beta}} \circ A_\beta) = ad_{g_{\alpha\beta}} \circ dA_\beta - [g_{\beta\alpha}^*\theta, ad_{g_{\alpha\beta}} \circ A_\beta]$$

    which is easy to check for e.g. matrix groups
  - We find

$$\begin{aligned} F_\alpha &= ad_{g_{\alpha\beta}} \circ dA_\beta - [g_{\beta\alpha}^*\theta, ad_{g_{\alpha\beta}} \circ A_\beta] + d(g_{\beta\alpha}^*\theta) \\ &\quad + \tfrac{1}{2}[ad_{g_{\alpha\beta}} \circ A_\beta + g_{\beta\alpha}^*\theta, ad_{g_{\alpha\beta}} \circ A_\beta + g_{\beta\alpha}^*\theta] \\ &= ad_{g_{\alpha\beta}} \circ F_\beta + g_{\beta\alpha}^*(d\theta + \tfrac{1}{2}[\theta,\theta]) \\ &= ad_{g_{\alpha\beta}} \circ F_\beta \end{aligned}$$

    since $[ad_g X, ad_g Y] = ad_g[X,Y]$ ■

- This means that on overlaps, the local field strengths transform under the adjoint rep of G

- Now consider how curvature transforms under gauge transformation $\Phi : P \to P$

- Using the structure equation of $\omega^\Phi = (\Phi^*)^{-1}\omega$, we find that it transforms as

$$\Omega^\Phi = (\Phi^*)^{-1}\Omega \tag{5.53}$$

- In terms of local field strengths, we find that the theorem above holds similarly:

$$F_\alpha^\Phi = ad_{\Phi_\alpha} \circ F_\alpha \tag{5.54}$$

5.7 Associated Vector Bundles and Basic Forms

- Recall that a representation of a group G on a vector space V is a homomorphism $\rho : G \to GL(V)$, where $\rho(gh) = \rho(g) \circ \rho(h)$

- Suppose $P \to M$ is a principal G-bundle, and $\rho : G \to GL(V)$ is a representation on a vector space V

- We define a right G-action on $P \times V$ by $(p, v) \mapsto (pg, \rho(g^{-1})v)$

- This action is free, so the corresponding orbit space

$$P \times_\rho V = (P \times V)/G \tag{5.55}$$

  is a smooth manifold

- $\pi : P \to M$ defines a projection $\pi_V : P \times_\rho V \to M$ by $(p, v) \mapsto \pi(p)$, which is well-defined since $\pi(p) = \pi(pg)$

- $P \times_\rho V$ is the total space of the **associated vector bundle** of the principal G-bundle

- The associated bundle inherits a local trivilisation from P:

$$\Psi_\alpha^V : \pi_V^{-1}(U_\alpha) \to U_\alpha \times V \quad \Psi_\alpha^V(p, v) = (\pi(p), \rho(g_\alpha(p))v) \tag{5.56}$$

- This is well-defined since it is independent of the representative (p, v) under $(p, v) \sim (pg, \rho(g^{-1})v)$

- On the overlaps $U_{\alpha\beta}$, the transition functions $t_{\alpha\beta} : U_{\alpha\beta} \to GL(V)$ are defined by $\rho(g_\alpha) = t_{\alpha\beta} \circ \rho(g_\beta)$, giving

$$t_{\alpha\beta} = \rho(g_{\alpha\beta}) \tag{5.57}$$

  where $g_{\alpha\beta} : U_{\alpha\beta} \to G$ are the transition functions of P

- This means the transition functions of an associated bundle $P \times_\rho V$ are given by the values in a rep ρ on V of the transition functions of the original principal bundle P

- This works the other way around too - given the transition functions, we can construct the associated vector bundle

- We now establish a relation between a class of V-values forms on P and a class of forms on M

- First, a **basic form** $\bar{\zeta} \in \Omega^k(P)$ is both horizontal ($h^*\bar{\zeta} = \bar{\zeta}$) and G-invariant ($R_g^*\bar{\zeta} = \bar{\zeta}$)

- **Lemma:** $\bar{\zeta} \in \Omega^k(P)$ is basic iff $\bar{\zeta} = \pi^*\zeta$ for some $\zeta \in \Omega^k(M)$

- **Proof:**
  - If $\bar{\zeta} = \pi^*\zeta$, then for any $v \in \mathfrak{X}(P)$ we have

$$(h^*\bar{\zeta})(v) = \bar{\zeta}(hv) = (\pi^*\zeta)(hv) = \zeta(\pi_* hv) = \zeta(\pi_* v) = \bar{\zeta}(v)$$

    since $\pi_* h = \pi_*$
  - Verifying G-invariance:

$$R_g^*\bar{\zeta} = R_g^*\pi^*\zeta = (\pi \circ R_g)^*\zeta = \pi^*\zeta = \bar{\zeta}$$

    since $\pi \circ R_g = \pi$
  - Conversely, suppose $\bar{\zeta} \in \Omega^k(P)$ is basic
  - Define $\zeta_\alpha = s_\alpha^*\bar{\zeta} \in \Omega^k(U_\alpha)$, where $s_\alpha : U_\alpha \to P$ is a local section
  - On non-trivial overlaps $U_{\alpha\beta} \ni m$, we have $s_\beta(m) = s_\alpha(m)g_{\alpha\beta}(m) = R_{g_{\alpha\beta}(m)}s_\alpha(m)$ where $g_{\alpha\beta}$ are the transition functions
  - Therefore, for $m \in U_{\alpha\beta}$:

$$\begin{aligned}(\zeta_\beta)_m &= s_\beta^*\bar{\zeta}_{s_\beta(m)} = (R_{g_{\alpha\beta}(m)} \circ s_\alpha)^*\bar{\zeta}_{s_\alpha(m)g_{\alpha\beta}(m)} \\ &= s_\alpha^* R_{g_{\alpha\beta}(m)}^*\bar{\zeta}_{s_\alpha(m)g_{\alpha\beta}(m)} \\ &= s_\alpha^*\bar{\zeta}_{s_\alpha(m)} = (\zeta_\alpha)_m\end{aligned}$$

- Thus we can glue together the ζ_α to get a globally defined $\zeta \in \Omega^k(M)$
- Finally:
$$\pi^*\zeta = \pi^* s_\alpha^* \bar{\zeta} = (s_\alpha \circ \pi)^* \bar{\zeta} = (s_\alpha \circ \pi)^* h^* \bar{\zeta} = h^* \bar{\zeta} = \bar{\zeta}$$
using that $h \circ (s_\alpha \circ \pi)_* = h$ and we are done ■

- This shows that basic forms on P are in 1:1 correspondence with forms on M
- We can actually extend this to forms valued in vector space V
- Define
$$\Omega^k_G(P;V) = \left\{\bar{\zeta} \in \Omega^k(P;V) \,|\, h^*\bar{\zeta} = \bar{\zeta},\ R_g^*\bar{\zeta} = \rho(g^{-1}) \circ \bar{\zeta}\right\} \tag{5.58}$$
which is the set of V-values horizontal k-forms on P which are G-invariant
- Alternatively, consider $\Omega^k(M; P \times_\rho V)$, which is the set of k-forms on M with values in the associated bundle $P \times_\rho V$
- We can represent such k-forms locally on M by $\zeta_\alpha \in \Omega^k(U_\alpha; V)$ such that on overlaps $U_{\alpha\beta}$ they are related by the defining transition functions of the associated bundle:
$$\zeta_\alpha = \rho(g_{\alpha\beta}) \circ \zeta_\beta \tag{5.59}$$
- **Theorem:** There is an isomorphism
$$\Omega^k_G(P;V) \simeq \Omega^k(M; P \times_\rho V) \tag{5.60}$$
- **Proof:**
  - Proof idea: represent elements of $\Omega^k(M; P \times_\rho V)$ in terms of their local data
  - First, let $\bar{\zeta} \in \Omega^k_G(P;V)$, and again define $\zeta_\alpha = s_\alpha^* \bar{\zeta} \in \Omega^k(U_\alpha; V)$
  - Arguing analogously to the earlier lemma, we have
$$\zeta_\beta = s_\alpha^* \bar{\zeta} = s_\alpha^* R^*_{g_{\alpha\beta}} \bar{\zeta} = s_\alpha^* \rho(g_{\alpha\beta}^{-1}) \circ \bar{\zeta} = \rho(g_{\alpha\beta}^{-1}) \circ s_\alpha^* \bar{\zeta} = \rho(g_{\alpha\beta}^{-1}) \circ \zeta_\alpha$$
  - This is equivalent to (5.59), so defines an element $\zeta \in \Omega^k(M; P \times_\rho V)$
  - Conversely, let $\zeta_\alpha \in \Omega^k(U_\alpha; V)$ be a set of forms satisfying (5.59)
  - Define $\bar{\zeta}_\alpha = \rho(g_\alpha^{-1}) \circ \pi^* \zeta_\alpha \in \Omega^k(\pi^{-1}(U_\alpha; V)$; on an overlap, we have
$$\bar{\zeta}_\alpha = \rho(g_\alpha^{-1}) \circ \pi^* \rho(g_{\alpha\beta}) \circ \zeta_\beta = \rho(g_\alpha^{-1} g_{\alpha\beta}) \circ \pi^* \zeta_\beta = \rho(g_\beta^{-1}) \circ \pi^* \zeta_\beta = \bar{\zeta}_\beta$$
  - This means we have a globally defined $\bar{\zeta} \in \Omega^k(P;V)$, and we just need to check it is horizontal and G-invariant
  - For horizontal-ness:
$$(h^*\bar{\zeta})(v) = \bar{\zeta}(hv) = \rho(g_\alpha^{-1}) \circ \zeta_\alpha(\pi_* hv) = \rho(g_\alpha^{-1}) \circ \zeta_\alpha(\pi_* v) = \bar{\zeta}(v)$$
for any $v \in \mathfrak{X}(P)$ as required
  - For G-invariance:
$$\begin{aligned} R_g^*\bar{\zeta} &= \rho((R_g^* g_\alpha)^{-1}) \circ R_g^* \pi^* \zeta_\alpha = \rho(g^{-1} g_\alpha^{-1}) \circ (\pi \circ R_g)^* \zeta_\alpha \\ &= \rho(g^{-1} g_\alpha^{-1}) \circ \pi^* \zeta_\alpha = \rho(g^{-1}) \circ \rho(g_\alpha^{-1}) \circ \pi^* \zeta_\alpha = \rho(g^{-1}) \circ \bar{\zeta} \end{aligned}$$
and we are done ■

5.8 Yang-Mills Theory

- Yang-Mills (YM) theory is a nonlinear generalisation of Maxwell's equations
- Let $P \xrightarrow{\pi} M$ be a principal G-bundle, with a connection defined by local gauge potentials $A_\alpha \in \Omega^1(U_\alpha; \mathfrak{g})$
- We have local field strengths $F_\alpha \in \Omega^2(U_\alpha; \mathfrak{g})$, which are equivalent to globally defined $F_A \in \Omega^2(M; ad\, P)$
- Also recall the Bianchi identity:

$$d_A F_A = 0 \tag{5.61}$$

- The **Yang-Mills equation** on any pseudo-Riemannian (M, g) is

$$d_A \star F_A = 0 \tag{5.62}$$

where d_A is the exterior covariant derivative, and $\star$ is the Hodge star operator extended to $\Omega^k(M; P \times_\rho V)$

- We can write this locally as

$$d_A \star F_\alpha = d \star F_\alpha + [A_\alpha, \star F_\alpha] = 0 \tag{5.63}$$

- This is derived from a variational principle
- To do this, we first need an inner product on $\Omega^k(M; P \times_\rho V)$
- To define this, suppose $\langle ; \rangle_V : V \times V \to \mathbb{R}$ is an inner product on V
- To ensure this inner product glues together properly on overlaps, we insist it is G-invariant:

$$\langle \rho(g)v, \rho(g)w \rangle_V = \langle v, w \rangle_V \tag{5.64}$$

for all $g \in G$, $v, w \in V$

- Extending this to forms $\zeta \in \Omega^k(M; P \times_\rho V)$, we first represent such forms locally as $\zeta_\alpha \in \Omega^k(U_\alpha; V)$; then, we define $\langle \zeta_\alpha, \eta_\alpha \rangle \in C^\infty(U_\alpha)$ by

$$\langle \zeta_\alpha, \eta_\alpha \rangle = \frac{1}{k!} \langle (\zeta_\alpha)_{a_1 \dots a_k}, (\eta_\alpha)_{b_1 \dots b_k} \rangle_V \, g^{a_1 b_1} \dots g^{a_k b_k} \tag{5.65}$$

- On overlaps, we therefore have

$$\langle \zeta_\alpha, \eta_\alpha \rangle = \langle \rho(g_{\alpha\beta}) \circ \zeta_\beta, \rho(g_{\alpha\beta}) \circ \eta_\beta \rangle = \langle \zeta_\beta, \eta_\beta \rangle \tag{5.66}$$

and so we can define the global $\langle \zeta, \eta \rangle \in C^\infty(M)$, and hence an inner product on $\Omega^k(M; P \times_\rho V)$

- Note that a G-invariant inner product on V doesn't always exist
- For the adjoint bundle $ad\, P = P \times_{ad} \mathfrak{g}$, we need an ad-invariant inner product on $\mathfrak{g}$, so $\langle ad_g X, ad_g Y \rangle = \langle X, Y \rangle$ for all $X, Y \in \mathfrak{g}$ and $g \in G$
- To do this, we first note the useful identity: set $g = \exp(tZ)$, and then differentiating $\langle ad_{\exp(tZ)} X, ad_{\exp(tZ)} Y \rangle = \langle X, Y \rangle$ at $t = 0$ implies

$$\langle ad_{*Z} X, Y \rangle + \langle X, ad_{*Z} Y \rangle = 0 \tag{5.67}$$

- Therefore, by an earlier lemma we have

$$\langle [X,Z],Y\rangle = \langle X,[Y,Z]\rangle\,, \quad \forall X,Y,Z\in\mathfrak{g} \tag{5.68}$$

- For semi-simple Lie algebras (see SoPF notes), it turns out a negative multiple of the **Killing form** $\kappa(X,Y) = \operatorname{Tr} ad_X ad_Y$ is an ad-invariant inner product, and it is positive definite if G is compact

- Compact Lie algebras include important examples like $G = SU(n)$, so we just take all our groups of interest to be compact and hence the inner product on V to be positive definite

- We now define the notation $\operatorname{Tr}(\zeta_\alpha \wedge \eta_\alpha) = \zeta^i_\alpha \wedge \eta^i_\alpha \langle e_i, e_j\rangle$ for $\zeta_\alpha \in \Omega^k(U_\alpha;V)$ and $\eta_\alpha \in \Omega^{k'}(U_\alpha;V)$ where $\zeta_\alpha = \zeta^i_\alpha e_i$ for a basis $\{e_i\}$ of V

- On overlaps $U_{\alpha\beta}$, we get $\operatorname{Tr}(\zeta_\alpha\wedge\eta_\alpha) = \operatorname{Tr}(\zeta_\beta\wedge\eta_\beta)$

- We therefore have a globally defined form $\operatorname{Tr}(\zeta\wedge\eta)\in\Omega^{k+k'}(M)$ for the corresponding global $\zeta\in\Omega^k(M;P\times_\rho V)$ and $\eta\in\Omega^{k'}(M;P\times_\rho V)$

- Notably:

$$\operatorname{Tr}(\zeta\wedge\star\eta) = \langle\zeta,\eta\rangle\,\mathrm{dvol} \tag{5.69}$$

  for $\zeta,\eta\in\Omega^k(M;P\times_\rho V)$ by definition of $\star$

- For $ad\,P$, ad-invariance then implies

$$\operatorname{Tr}([\zeta,\zeta']\wedge\zeta'') = \operatorname{Tr}(\zeta\wedge[\zeta',\zeta'']), \quad \forall \zeta,\zeta',\zeta''\in\Omega^k(M;ad\,P) \tag{5.70}$$

- We now finally get to the definition of the **Yang-Mills functional**, which is

$$S_{YM} = \int_M |F_A|^2 \mathrm{dvol} \tag{5.71}$$

  where $F_A \in \Omega^2(M;ad\,P)$ is the curvature of a connection, and $|F_A|^2 = \langle F_A, F_A\rangle \in C^\infty(M)$

- Note that F_α depends on local sections $s_\alpha : U_\alpha \to \pi^{-1}(U_\alpha)$, but it can be shown that S_{YM} does *not*

- It follows that S_{YM} also does not depend on choice of local trivilisation, and so it only depends on the connection

- Defining the space of connections $\mathcal{A}$, we therefore have $S_{YM} : \mathcal{A}\to\mathbb{R}$

- **Theorem:** S_{YM} is gauge invariant

- **Proof:**
  - In a trivilisation, a gauge transformation is described locally by $\Phi_\alpha : U_\alpha \to G$
  - The field strengths transform as $F^\Phi_\alpha = ad_{\Phi_\alpha}\circ F_\alpha$, so by ad-invariance of the inner product, $\langle F^\Phi, F^\Phi\rangle = \langle F, F\rangle$ and we are done ■

- Note this really means $S_{YM} : \mathcal{A}/\mathcal{G} \to \mathbb{R}$, where the quotient is defined by equivalence under the group of gauge transformations $\mathcal{G}$

- Recall that $\mathcal{A}$ is affine, so if $A\in\mathcal{A}$ we have a curve $A + t\tau$ in the space of connections where $t\in\mathbb{R}$ and $\tau\in\Omega^1(M;ad\,P)$

- We say that A is a **critical point** of S_{YM} if for all $\tau \in \Omega^1(M; ad\, P)$ with compact support:

$$\frac{d}{dt} S_{YM}(A + t\tau)\bigg|_{t=0} = 0 \tag{5.72}$$

- **Theorem:** A connection is a critical point of S_{YM} iff it satisfies the YM equation

- **Proof:**

  – Working locally, we have $F_A = dA + \frac{1}{2}[A, A]$ (dropping explicit α labels)

  – This means:

$$F_{A+t\tau} = d(A + t\tau) + \tfrac{1}{2}[A + t\tau, A + t\tau] = F_A + t d_A \tau + \tfrac{1}{2}t^2[\tau, \tau]$$

  using $[\tau, A] = [A, \tau]$ and $d_A\tau = d\tau + [A, \tau]$

  – Therefore:

$$|F_{A+t\tau}|^2 = |F_A + t d_A\tau + \mathcal{O}(t^2)|^2 = |F_A|^2 + 2t \langle d_A\tau, F_A\rangle + \mathcal{O}(t^2)$$

  – Using (5.69), we then get

$$\begin{aligned}\frac{d}{dt} S_{YM}(A + t\tau)\bigg|_{t=0} &= 2\int_M \langle d_A\tau, F_A\rangle \,\mathrm{dvol} = 2\int_M \mathrm{Tr}(d_A\tau \wedge \star F_A) \\ &= 2\int_M \mathrm{Tr}(d\tau \wedge \star F_A) + \mathrm{Tr}([A, \tau] \wedge \star F_A) \\ &= 2\int_M \mathrm{Tr}(\tau \wedge d \star F_A) + \mathrm{Tr}(\tau \wedge [A, \star F_A]) \\ &= 2\int_M \mathrm{Tr}(\tau \wedge d_A \star F_A)\end{aligned}$$

  where on line 3 we used $d\mathrm{Tr}(\tau \wedge \star F_A) = \mathrm{Tr}(d\tau \wedge \star F_A) - \mathrm{Tr}(\tau \wedge d \star F_A)$ and Stokes' theorem to get rid of the boundary term, and $\mathrm{Tr}([\tau, A] \wedge \star F_A) = \mathrm{Tr}(\tau \wedge [A, \star F_A])$

  – The above variation has to vanish for all τ, so we get that the YM equation must be satisfied ■

5.9 Instantons

- An **instanton** is a localised solution to the YM equations that are 'existing at an instant of time'

- They are defined on *Riemannian* manifolds (M, g), so the signature $s = 0$

- We take (M, g) to have dimension $n = 4$, so the Hodge star acts as $\star\colon \Omega^2(M) \to \Omega^2(M)$ and satisfies $\star^2 = 1$

- This means we can decompose 2-forms pointwise as

$$\Omega^2(M) = \Omega^2_+(M) \oplus \Omega^2_-(M) \tag{5.73}$$

  where $\Omega_\pm(M)$ are the $\star = \pm 1$ eigenspaces respectively

- We say $\omega \in \Omega^2(M)$ is **self-dual** (SD) if $\omega \in \Omega^2_+(M)$, so $\star\omega = \omega$, and it is **anti self-dual** (ASD) if $\omega \in \Omega^2_-(M)$, so $\star\omega = -\omega$

- Any 2-form $\omega \in \Omega^2(M)$ can be uniquely decomposed into a sum of a SD and ASD 2-form $\omega = \omega_+ + \omega_-$ where $\omega_\pm \in \Omega^2_\pm(M)$

- Also:
$$\begin{aligned} \langle \omega_+, \omega_- \rangle \,\mathrm{dvol} &= \omega_+ \wedge \star\omega_- = -\omega_+ \wedge \omega_- = -\omega_- \wedge \omega_+ \\ &= -\omega_- \wedge \star\omega_+ = -\langle \omega_-, \omega_+ \rangle \,\mathrm{dvol} = -\langle \omega_+ \omega_- \rangle \,\mathrm{dvol} \end{aligned} \tag{5.74}$$
and so $\langle \omega_+, \omega_- \rangle = 0$, meaning the decomposition is orthogonal wrt the inner product on forms

- We can do the same decomposition of bundle-valued 2-forms on M so long as we have an inner product on V

- Given $F_A \in \Omega^2(M; ad\, P)$ associated to the curvature of a connection on a principal bundle, we can decompose $F_A = F_A^+ + F_A^-$, where $\star F_A^\pm = \pm F_A^\pm$

- This means
$$|F_A|^2 = \langle F_A^+ + F_A^-, F_A^+ + F_A^- \rangle = |F_A^+|^2 + |F_A^-|^2 \tag{5.75}$$
using orthogonality of the SD and ASD parts of F_A

- On the other hand:
$$\begin{aligned} \mathrm{Tr}(F_A \wedge F_A) &= \mathrm{Tr}(F_A^+ \wedge F_A^+) + 2\mathrm{Tr}(F_A^+ \wedge F_A^-) + \mathrm{Tr}(F_A^-) \\ &= \mathrm{Tr}(F_A^+ \wedge \star F_A^+) - 2\mathrm{Tr}(F_A^+ \wedge \star F_A^-) - \mathrm{Tr}(\star F_A^-) \\ &= (|F_A^+|^2 - |F_A^-|^2)\mathrm{dvol} \end{aligned} \tag{5.76}$$
using orthogonality again

- Therefore, defining
$$c = \int_M \mathrm{Tr}(F_A \wedge F_A) \tag{5.77}$$
we can write
$$S_{YM} = \mp c + 2\int_M |F_A^\pm|^2 \mathrm{dvol} \tag{5.78}$$

- We therefore deduce the following important result

- **Theorem:** The YM action on a 4-dimensional Riemannian manifold satisfies
$$S_{YM} \geq |c| \tag{5.79}$$
with equality iff $F_A^+ = 0$ or $F_A^- = 0$

- **Proof:**
  - We see instantly that $S_{YM} = -c$ iff $F_A^+ = 0$, which implies
  $$c = -\int_M |F_A^-|^2 < 0$$
  giving $S_{YM} = |c|$
  - Similarly, $S_{YM} = c$ iff $F_A^- = 0$ and the result follows similarly ■

- A connection which is (A)SD in the sense $F_A = F_A^\pm$ is called an **instanton**

- An instanton is clearly a solution to the YM equation by the Bianchi identity:
$$d_A \star F_A = \pm d_A F_A = 0 \tag{5.80}$$

- This means instantons are solutions to the YM equation which are global minimisers of the YM action

- The instanton equation $F_A^{\pm} = 0$ is a first order PDE for the connection A, which implies the YM equations (which are a second order PDE for A)
- It is therefore simpler to find instanton solutions
- Now, we assume M is compact in what follows
- Define $\Theta = \mathrm{Tr}(F_A \wedge F_A) \in \Omega^4(M)$; since $n = 4$ this is a top-ranked form, and so $d\Theta = 0$
- Hence this defines a cohomology class $[\Theta] \in H^4_{dR}(M)$ and $c = ([\Theta], [M]) = \int_M \Theta$ where $[M] \in H_4(M)$ is the fundamental homology class
- It turns out c is independent of the connection on P and is a topological invariant of the bundle called the **characteristic number**
- To see this, let $A_0, A_1 \in \mathcal{A}$ be connections, and consider the curve $A_t = A_0 + t\tau$ where $\tau = A_1 - A_0 \in \Omega^1(M; ad\, P)$ and $t \in [0, 1]$
- By the same logic as in the proof of connections being critical points of S_{YM} iff they satisfy the YM equation, we have $F_{A_t} = F_{A_0} + t d_{A_0}\tau + \frac{1}{2}t^2[\tau, \tau]$
- Therefore:
$$\frac{d}{dt}F_{A_t} = d_{A_0}\tau + t[\tau, \tau] = d_{A_t}\tau \tag{5.81}$$
- This gives:
$$\begin{aligned}\frac{d}{dt}\Theta_t &= 2\mathrm{Tr}\left(\frac{d}{dt}F_{A_t} \wedge F_{A_t}\right) = 2\mathrm{Tr}(d_{A_t}\tau \wedge F_{A_t}) \\ &= 2\mathrm{Tr}(d\tau \wedge F_{A_t} + [A_t, \tau] \wedge F_{A_t}) \\ &= 2\mathrm{Tr}(d\tau \wedge F_{A_t} + \tau \wedge [A_t, F_{A_t}]) \\ &= 2\mathrm{Tr}(d\tau \wedge F_{A_t} - \tau \wedge dF_{A_t}) \\ &= d\mathrm{Tr}(\tau \wedge F_{A_t})\end{aligned} \tag{5.82}$$
using the Bianchi identity in line 4
- Therefore:
$$\Theta_1 - \Theta_0 = d\left(2\int_0^1 \mathrm{Tr}(\tau \wedge F_{A_t})dt\right) \tag{5.83}$$
so Θ_0 and Θ_1 define the same cohomology class $[\Theta_0] = [\Theta_1]$
- By Stokes', $c = \int_M \Theta_0 = \int_M \Theta_1$ is independent of the connection
- For the adjoint bundle $ad\, P$, the **Pontrjagin class** is
$$p_1(ad\, P) = \frac{1}{4\pi^2}\mathrm{Tr}(F_A \wedge F_A) \in H^4(M, \mathbb{R}) \tag{5.84}$$
and the **Pontrjagin number** is
$$p_1(ad\, P)[M] = \frac{1}{4\pi^2}\int_M \mathrm{Tr}(F_A \wedge F_A) \tag{5.85}$$
- It can be shown that the Pontrjagin number is an integer
- This is a topological invariant of the bundle, in this context called the **instanton number** $k \in \mathbb{Z}$

- So $c = 4\pi^2 k$, and
$$S_{YM} \geq 4\pi^2 |k| \tag{5.86}$$
so the lower bound of the YM action is a topological invariant of the underlying bundle

- **Example:** The BPST instanton
  - The BPST instanton is the simplest non-trivial example of an instanton
  - It is a connection on a principal $SU(2)$-bundle over $\mathbb{R}^4$ with $k = 1$
  - It is a 5-parameter family of solutions where the parameters are the coordinates of its 'centre' in $\mathbb{R}^4$ and a scale
  - It has finite S_{YM} but not compact support, and by 'adding the point at infinity' it extends to an instanton on $S^4 = \mathbb{R}^4 \cup \{\infty\}$
  - The total space of this $SU(2)$-bundle over S^4 turns out to be S^7
  - We can study the **moduli space** (i.e. parameter space) of instantons as a geometric object in its own right - for $SU(2)$-bundles over S^4, the moduli space is $8k - 3$ dimensional, so for $k = 1$ the BPST instanton is the only instanton solution