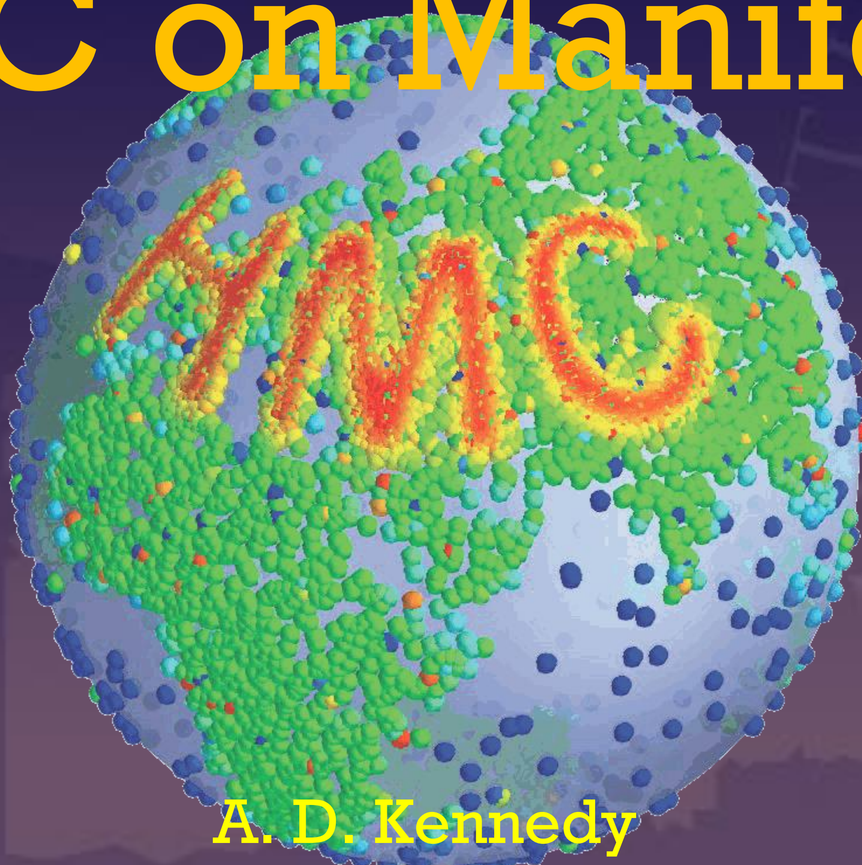
HMC on Manifolds



A. D. Kennedy
School *of* Physics *and* Astronomy
University of Edinburgh

Apology

« Je n'ai fait celle-ci plus longue que parce que je n'ai pas eu le loisir de la faire plus courte. »

Blaise Pascal

„ So lange kann ich auch nicht aufbleiben. "

Wolfgang Pauli

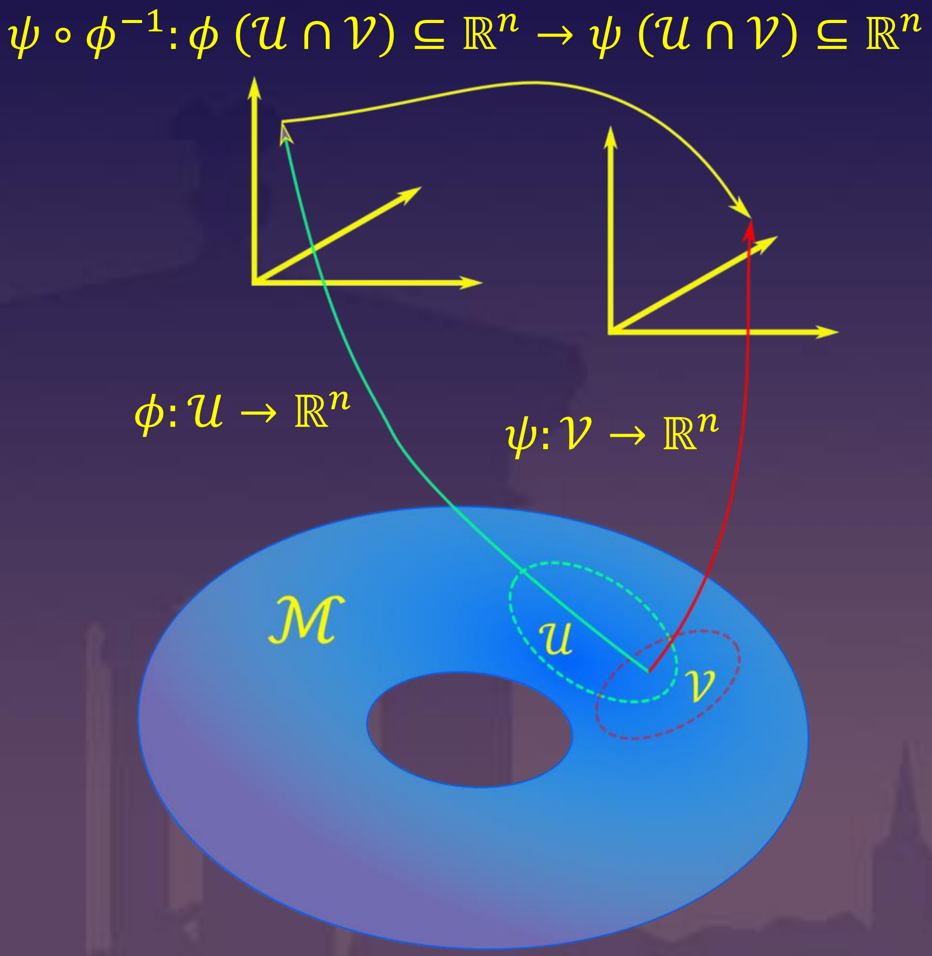
(When asked to schedule his lecture at 8 am)

Motivation

- Hybrid (or Hamiltonian) Monte Carlo is an efficient method for sampling from probability distributions in high-dimensional spaces
  - It is widely used in lattice field theory and Bayesian statistics
  - It is naturally defined on Riemannian manifolds
    - Gauge fields live on an $\mathrm{SU}(3,\mathbb{R})$ group manifold at each site
    - Non-linear σ models have fields in S_2 at each site
    - $\mathbb{CP}^n$ models have fields in $\mathbb{CP}^n$ at each site
    - Statisticians are interested in Stiefel manifolds $V_k(\mathbb{R}^n)$ (the space of k-frames in $\mathbb{R}^n$) and the space of covariance matrices (symmetric positive definite matrices)

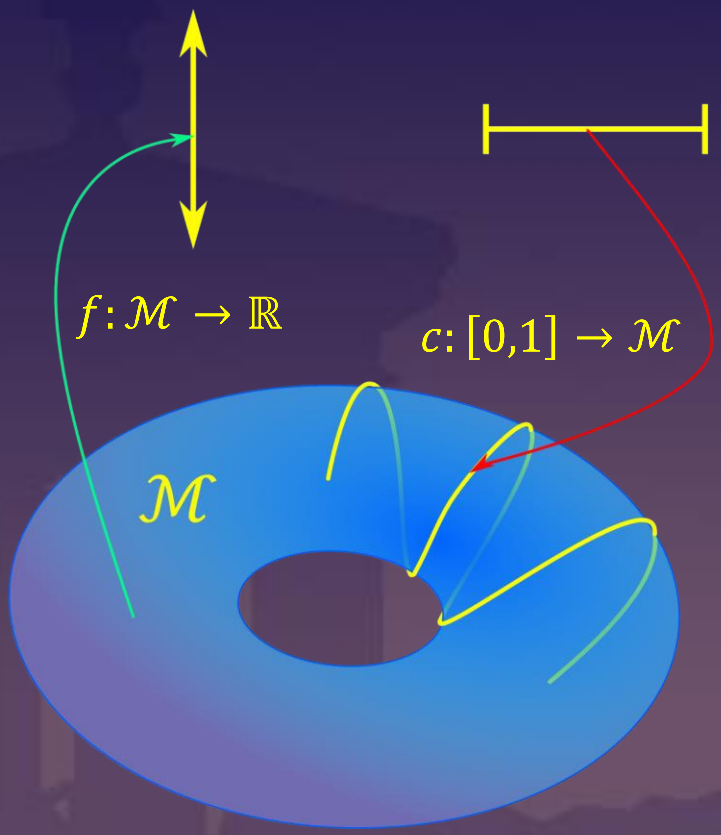
Differential Geometry

Differential Manifolds



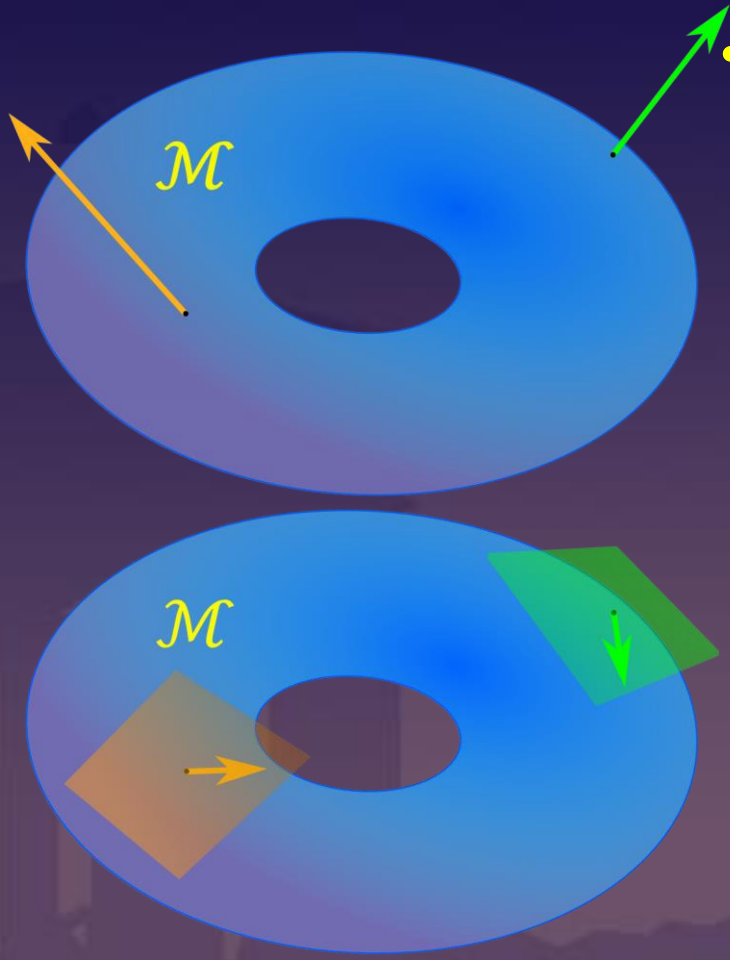
- Differential manifolds "look like" $\mathbb{R}^n$ locally
  - Have an atlas of coordinate charts
  - Smooth structure is endowed by C^∞ maps between charts
- Charts are local: we want to use global entities
  - Coordinate transformations are painful
  - Numerically unstable

Curves and Functions



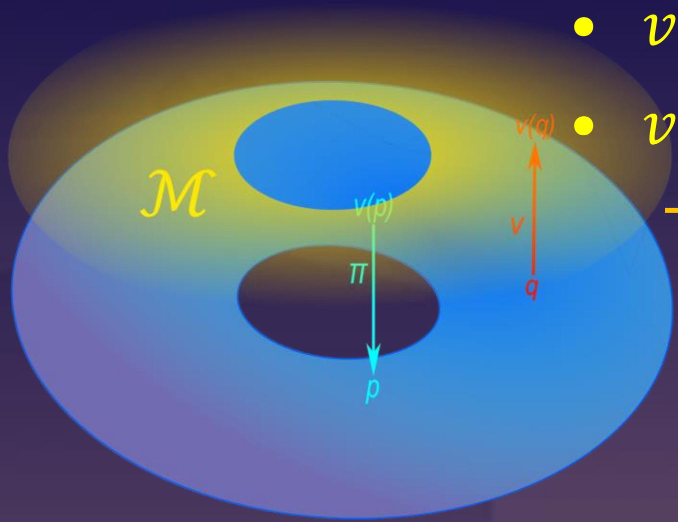
- Curves
  - A parametric curve is a smooth map from a closed interval of $\mathbb{R}$ into the manifold $\mathcal{M}$
- Functions $f \in \Lambda^0(\mathcal{M})$
  - 0-forms are smooth $\mathbb{R}$-valued functions over $\mathcal{M}$
- Smoothness will be implicit forthwith

Tangent Vector Fields

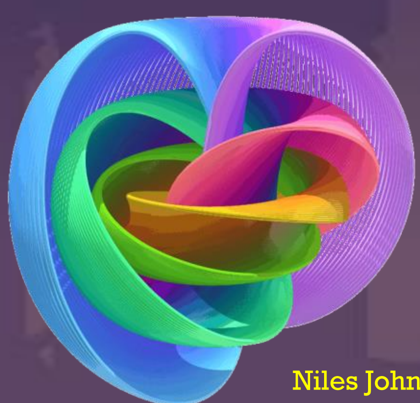


- A vector field $v \in \Gamma(\mathcal{M})$
$$v: \Lambda^0(\mathcal{M}) \to \Lambda^0(\mathcal{M})$$
is a linear differential operator
  - It satisfies the Leibniz rule $v(fg) = v(f)g + fv(g) \quad \forall f, g \in \Lambda_0$
    - In a chart $v = \sum_{j=1}^{n} v^i(x) \frac{\partial}{\partial x^i}$
  - The commutator $[u, v]$ of two vector fields is itself a vector field
    - In a chart
$$\left[u_j \frac{\partial}{\partial x^j}, v_k \frac{\partial}{\partial x^j}\right] = \left(u_j \frac{\partial v_\ell}{\partial x^j} - v_j \frac{\partial u_\ell}{\partial x^j}\right) \frac{\partial}{\partial x^\ell}$$

Fibre Bundles



- $v(p)$ lives in the tangent space $T_p\mathcal{M}$
- v is a section of the tangent bundle $T\mathcal{M}$
  - A fibre bundle is locally just a product $\mathcal{U} \times \mathcal{F}$ where $\mathcal{U} \subseteq \mathcal{M}$ is an open set and $\mathcal{F}$ is the vertical manifold
    - There is a "vertical" projection $\pi: \mathcal{U} \times \mathcal{F} \to \mathcal{U}$
    - For the tangent bundle the fibre is the tangent space $\mathcal{F} = T_x\mathcal{M}$
  - It is not necessarily a global product
    - Unlike a global product, there is no "horizontal" projection in general
    - E.g., S_3 is a fibre bundle with S_1 fibres over S_2. This is the Hopf fibration



Niles Johnson, OSU

k-Form Fields

- 1-form fields ω live in the dual space $\Lambda_1(\mathcal{M})$ to the space of vector fields $\Gamma(\mathcal{M})$
$$\omega: \Gamma(\mathcal{M}) \to \Lambda_0(\mathcal{M}): v \mapsto \omega(v)$$
  - They are sections of the cotangent bundle $T^*\mathcal{M}$
- A k-form field $\beta \in \Lambda_k(\mathcal{M})$ is a totally antisymmetric multilinear map
$$\beta: \Gamma(\mathcal{M})^{\otimes k} \to \Lambda_0(\mathcal{M}): (v_1, \dots, v_k) \mapsto \beta(v_1, \dots, v_k)$$
- There is an associative antisymmetric wedge product
$$\alpha \wedge \beta = (-1)^{jk} \beta \wedge \alpha \ \text{ for } \ \alpha \in \Lambda_j(\mathcal{M}), \beta \in \Lambda_k(\mathcal{M})$$
$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma) = \alpha \wedge \beta \wedge \gamma$$

Exterior Calculus

- Exterior derivatives map $d: \Lambda_k(\mathcal{M}) \to \Lambda_{k+1}(\mathcal{M})$
  - $d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^j \, \alpha \wedge d\beta$ (antiderivation)
  - $d^2 = 0$; central rôle in the de Rahm cohomology and the Chevalley—Eilenberg complex
    - $df(v) = vf \quad \forall f \in \Lambda_0(\mathcal{M}), v \in \Gamma(\mathcal{M})$
    - $d\alpha(u,v) = u\alpha(v) - v\alpha(u) - \alpha([u,v]) \;\; \forall \alpha \in \Lambda_1(\mathcal{M}), u, v \in \Gamma(\mathcal{M})$
      - $$\begin{aligned} d^2 f(u,v) &= u\, df(v) - v\, df(u) - df([u,v]) \\ &= (u \circ v - v \circ v)(f) - [u,v](f) = 0 \end{aligned}$$
    - $$\begin{aligned} d\omega(u,v,w) &= u\omega(v,w) + v\omega(w,u) + w\omega(u,v) \\ &\quad - \omega([u,v],w) - \omega([v,w],u) - \omega([w,u],v) \end{aligned}$$
      $\forall \omega \in \Lambda_2(\mathcal{M}), u, v, w \in \Gamma(\mathcal{M})$
      - $d^2\alpha(u,v,w) = \alpha\big([[u,v],w] + [[v,w],u] + [[w,u],v]\big) = 0$ using the Jacobi identity (for commutators)

Symplectic Geometry

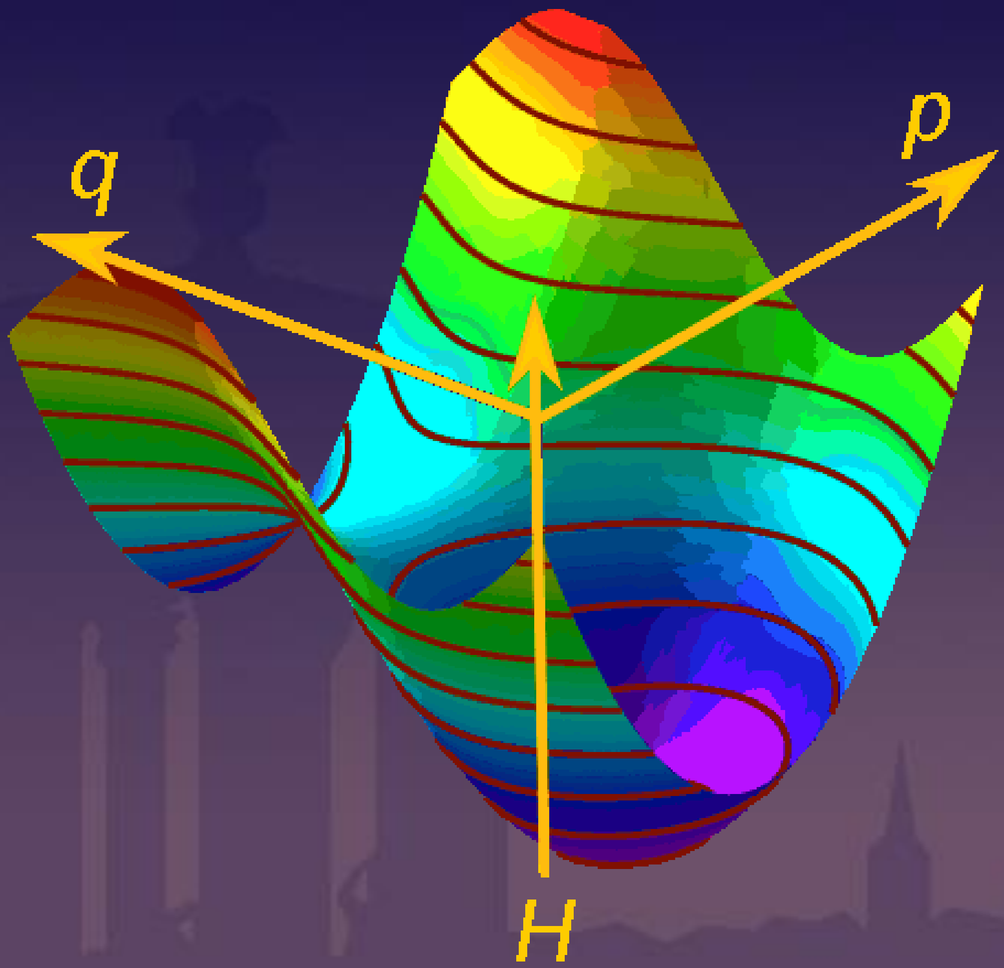
Symplectic Manifolds

- A <u>symplectic manifold</u> admits a fundamental non-degenerate closed 2-form field ω
  - Such manifolds must be even dimensional
  - <u>Darboux theorem</u>: Locally, there is always a chart with coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$ in which $\omega = \sum_{j=1}^{n} dp_j \wedge dq_j$
- The cotangent bundle $T^*\mathcal{M}$ is usually symplectic
  - Phase space over $\mathcal{M}$
  - <u>Liouville form</u> $\vartheta = \sum_{j=1}^{n} p_j \, dq_j$
  - Fundamental 2-form $\omega = d\vartheta$ is then automatically closed
  - This is well-defined if the cotangent bundle is a global product
    - We will see examples where this is not true later, time permitting

Hamiltonian Vector Fields

- For $A \in \Lambda_0(\mathcal{M})$ there is a corresponding Hamiltonian vector field $\hat{A} \in \mathrm{Ham}(\mathcal{M}) \subset \Gamma(\mathcal{M})$ such that for any vector field u we have $dA(u) = \omega(\hat{A}, u)$
  - This may be expressed as $dA = i_{\hat{A}}\omega$
- For HMC we define the Hamiltonian function $H = T + V$ on a cotangent bundle
  - V only depends on $\pi(x)$, the position in the base manifold
  - In simple cases T only depends on the "vertical" position in the fibre (the momentum)
    - This is not true in general, as we shall soon see
  - The integral curves of the Hamiltonian Hamiltonian vector field $\widehat{H}$ are the classical trajectories
  - We may build symplectic integrators using the Hamiltonian vector fields $\widehat{T}$ and $\widehat{V}$

Example — Pendulum



- Base manifold S_1
- Fibre $\mathbb{R}$
- Cotangent bundle $T^*\mathcal{M} = S_1 \times \mathbb{R}$
- Hamiltonian $$H = \frac{1}{2}p^2 + \sin(2\pi q)$$

Hamiltonian Flow

- Energy contours
- Hamiltonian Hamiltonian vector field

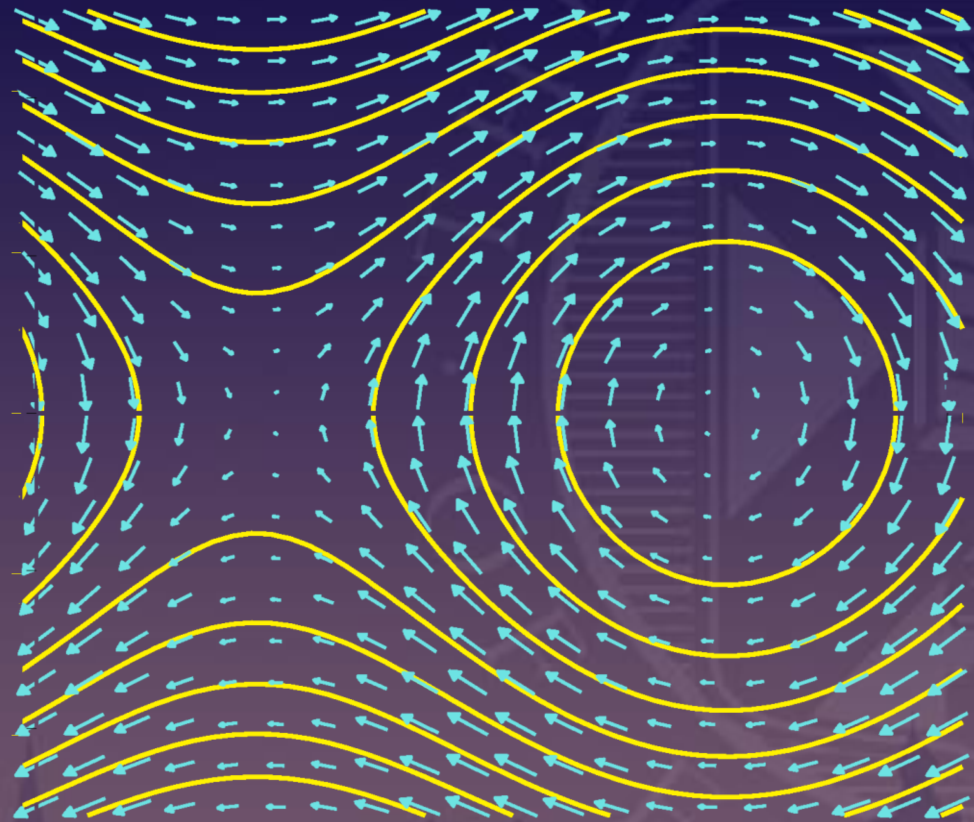
$$\hat{H} = p\frac{\partial}{\partial q} - 2\pi\cos(2\pi q)\frac{\partial}{\partial p}$$

- Integral curves

$$c: t \mapsto (q(t), p(t))$$

- Hamilton's equations

$$\dot{c} = \hat{H}c$$



HMC

- The HMC Markov chain has its fixed-point distribution
$$\propto e^{-H(q,p)} = e^{-T(q,p)} e^{-V(q)}$$
  - Momentum refreshment is a momentum Gibbs sampler (heatbath) that samples from the momentum distribution $e^{-T(q,p)}$
  - On the cotangent space of a Riemannian manifold with metric g
$$T(q,p) = \frac{1}{2} g^{-1}(p,p)$$
    - We require a Riemannian metric, $g > 0$, for this distribution to be normalizable
- The marginal distribution of q values is proportional to
$$\int d^n p \, e^{-\frac{1}{2} g^{-1}(p,p)} e^{-V(q)} \propto \sqrt{\det g(q)} \, e^{-V(q)}$$
  - If g is a constant, as it is on Lie groups and homogeneous spaces, then the factor of $\sqrt{\det g(q)}$ is immaterial, but on a general Riemannian manifold it is the natural measure
    - It is invariant under isometries

Poisson Geometry

Historical Remark

« … quelques remarques sur la plus profonde découverte de M. Poisson, mais qui, je crois, n’a pas été bien comprise ni par Lagrange, ni par les nombreux géomètres qui l’ont citée, ni par son auteur lui-même. »

Carl Gustav Jacob Jacobi

Poisson Brackets

- The Poisson bracket of the 0-forms (functions) A and B is
$$\{A,B\} = -\omega\left(\hat{A},\hat{B}\right)$$
  - By the following algebraic manipulations
    - $\hat{A}(B) = dB\left(\hat{A}\right) = \omega\left(\hat{B},\hat{A}\right) = \{A,B\}$
    - $\omega\left([\hat{A},\hat{B}],\hat{C}\right) = -\omega\left(\hat{C},[\hat{A},\hat{B}]\right) = -dC\left([\hat{A},\hat{B}]\right) = -[\hat{A},\hat{B}]C$
    - $[\hat{A},\hat{B}]C = \left(\hat{A}\hat{B}-\hat{B}\hat{A}\right)C = \hat{A}\{B,C\} - \hat{B}\{A,C\} = \{A,\{B,C\}\} - \{B,\{A,C\}\}$
    - $$\begin{aligned} d\omega\left(\hat{A},\hat{B},\hat{C}\right) = {} & -\hat{A}\{B,C\} - \hat{B}\{C,A\} - \hat{C}\{A,B\} \\ & - \omega\left([\hat{A},\hat{B}],\hat{C}\right) - \omega\left([\hat{B},\hat{C}],\hat{A}\right) - \omega\left([\hat{C},\hat{A}],\hat{B}\right) \\ = {} & 3\left(\{A,\{B,C\}\} + \{B,\{C,A\}\} + \{C,\{A,B\}\}\right) = 0 \end{aligned}$$
  - they form non-trivial Lie algebra
$$\{A,\{B,C\}\} + \{B,\{C,A\}\} + \{C,\{A,B\}\} = 0$$
  - Moreover, Hamiltonian vector fields are closed under commutation $[\hat{A},\hat{B}] = \widehat{\{A,B\}}$ since
    - $[\hat{A},\hat{B}]C = -\{C,\{A,B\}\} = \{\{A,B\},C\} = \widehat{\{A,B\}}C$

Structure-preserving maps

- Morphisms are called
  - Continuous for topological spaces
  - Smooth for differential manifolds
  - Symplectic or canonical transformations for Symplectic manifolds
- Isomorphisms of these structures are called
  - Homeomorphisms for topological spaces
  - Diffeomorphisms for differential manifolds
  - Symplectomorphisms for symplectic manifolds
  - Ichthyomorphisms for Poisson manifolds
    - For people who like multilingual puns

BCH and Shadow Hamiltonians

- The <u>Baker—Campbell —Hausdorff</u> formula is a formal expression for the product of matrix exponentials

$$e^X \cdot e^Y = \exp\left(X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}\left([X,[X,Y]] - [Y,[X,Y]]\right) + \cdots\right)$$

- Since $\left[\hat{X},\hat{Y}\right] = \widehat{\{X,Y\}}$ this gives $\widehat{e^X \cdot e^Y} = e^{\tilde{H}}$ with the <u>Shadow Hamiltonian</u>

$$\tilde{H} = \exp\left(\hat{X} + \hat{Y} + \frac{1}{2}\{\hat{X},\hat{Y}\} + \frac{1}{12}\left(\left\{\hat{X},\{\hat{X},\hat{Y}\}\right\} - \left\{\hat{Y},\{\hat{X},\hat{Y}\}\right\}\right) + \cdots\right)$$

  - For $\hat{X},\hat{Y} = \mathcal{O}(\delta\tau)$ this is an asymptotic expansion

- As (symmetric) symplectic integrators are products of such exponentials they are the geodesics of the <u>Shadow Hamiltonian</u> Hamiltonian vector field $\hat{\tilde{H}}$
  - Thus $\tilde{H}$ is a constant of motion, and as $H = \tilde{H} + \mathcal{O}(\delta\tau^n)$ the energy H is approximately conserved even for long trajectories

Historical Remark

"Metropolis was boss of the computer laboratory. We never had a single scientific discussion with him."

Marshall Rosenbluth

Symmetric Symplectic Integrators

- The key idea of HMC is that we may approximate Hamiltonian trajectories using symmetric symplectic integrators
  - The simplest example is the leapfrog (Störmer—Verlet) integrator
    $$\left[\exp\left(\frac{\delta\tau}{2}\,\hat{V}\right)\,\exp(\delta\tau\,\hat{T})\,\exp\left(\frac{\delta\tau}{2}\,\hat{V}\right)\right]^{\tau/\delta\tau} = \exp\left(\left(\hat{H}+\mathcal{O}(\delta\tau^2)\right)\tau\right)$$
    - This equation follows from the BCH formula
- This is used to suggest a candidate update for the Metropolis algorithm
  - The acceptance probability is $\min\left(1, e^{-\delta H}\right)$

Riemannian Geometry

Connections

- A <u>connection</u> on a manifold is bilinear map
$$\nabla\colon \Gamma(\mathcal{M}) \times \Gamma(\mathcal{M}) \to \Gamma(\mathcal{M})$$
- It is a derivation, so it satisfies
$$\left.\begin{aligned} \nabla_{fu} v &= f \nabla_u v \\ \nabla_u (fv) &= (uf)v + f\nabla_u v \end{aligned}\right\} \forall u, v \in \Gamma(\mathcal{M}), f \in \Lambda_0(\mathcal{M})$$
- $\nabla_u v$ is the <u>covariant derivative</u> of v with respect to u
  - In a chart $\nabla_{\partial_j} \partial_k = \Gamma_{jk}^{\ell} \partial_{\ell}$ with the notation $\partial_{\ell} = \frac{\partial}{\partial x_{\ell}}$
  - Γ_{jk}^{ℓ} are <u>Christoffel symbols</u>
- The <u>torsion</u> of the connection is
$$T(u, v) = \nabla_u v - \nabla_v u - [u, v]$$

Riemannian Manifolds

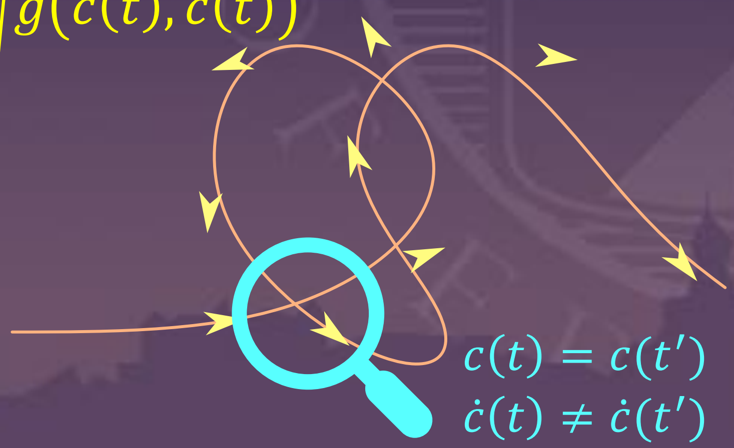
- A <u>Riemannian manifold</u> has the additional structure of a positive definite metric tensor

$$g\colon \Gamma(\mathcal{M}) \times \Gamma(\mathcal{M}) \to \Lambda_0(\mathcal{M})$$

  - At each point $q \in \mathcal{M}$ and vectors $u \neq v \in T_q\mathcal{M}$ we have $g(u,v) > 0$
- This defines the <u>norm</u> $\|v\| = \sqrt{g(v,v)}$ of a vector $v \in T_q\mathcal{M}$
- The <u>length</u> of a curve $c\colon [a,b] \to \mathcal{M}$ is defined to be

$$\ell(c) = \int_a^b dt\, \|\dot{c}(t)\| = \int_a^b dt\, \sqrt{g\big(\dot{c}(t), \dot{c}(t)\big)}$$

  - Where $\dot{c}\colon [a,b] \to \Gamma(\mathcal{M})$ is the tangent to the curve, $\dot{c} = \left.\frac{d}{dt}\right|_c$, and $\dot{c}(t)$ is thus the tangent vector at parameter t
    - We say it is the tangent at $c(t)$, but this may be ambiguous for a self-intersecting curve



Metric Connections

- On a Riemannian manifold a metric connection obeys

$$ug(v,w) = g(\nabla_u v, w) + g(v, \nabla_u w)$$

- There is a unique torsion-free Levi-Civita connection
  - It is given by the Koszul formula

$$2g(\nabla_u, v, w) = \begin{array}{r} g(v,w) + vg(w,u) - wg(u,v) \\ -g(u,[v,w]) + g(v,[w,u]) + g(w,[u,v]) \end{array}$$

  - In a chart Christoffel symbols are

$$\Gamma_{ij}^{k} = \frac{1}{2} g^{k\ell} \left(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij} \right)$$

Geodesics

- The distance $d(q, q')$ between two nearby points $q, q' \in \mathcal{M}$ is the length of the shortest curve connecting them
  - Such curves satisfy the geodesic equation $\nabla_{\dot{c}}\dot{c} = 0$ and are called minimal geodesics
  - In a chart the geodesic equation is $\ddot{c} + \Gamma^k_{ij}\dot{c}^i\dot{c}^j = 0$
- This provides $\mathcal{M}$ with a metric consistent with its topology
  - Not all points need be connected by a single geodesic

Riemannian Hamiltonian Systems

- The natural kinetic energy $T \in \Lambda_0(T^*\mathcal{M})$ is

$$T(q,p) = \tfrac{1}{2} g^{-1}(p,p)$$

  – The Hamiltonian vector field $\hat{T}$ in a Darboux chart is

$$\begin{pmatrix} \hat{T}^\mu \\ \hat{T}_\sigma \end{pmatrix} = \begin{pmatrix} \dfrac{\partial T}{\partial p_\mu} \\ -\dfrac{\partial T}{\partial q^\sigma} \end{pmatrix} = \begin{pmatrix} g^{\mu\nu} p_\nu \\ -\dfrac{1}{2}\dfrac{\partial g^{\mu\nu}}{\partial q^\sigma} p_\mu p_\nu \end{pmatrix}$$

    - For the integral curve $c = (Q,P)$ we have $\dot{c} = (\dot{Q}, \dot{P}) = \hat{T}\big|_c$

$$\dot{Q}^\mu = \hat{T}^\mu = g^{\mu\nu} P_\nu \quad \dot{P}_\sigma = \hat{T}_\sigma = -\frac{1}{2}\frac{\partial g^{\mu\nu}}{\partial q^\sigma} P_\mu P_\nu$$

Details…

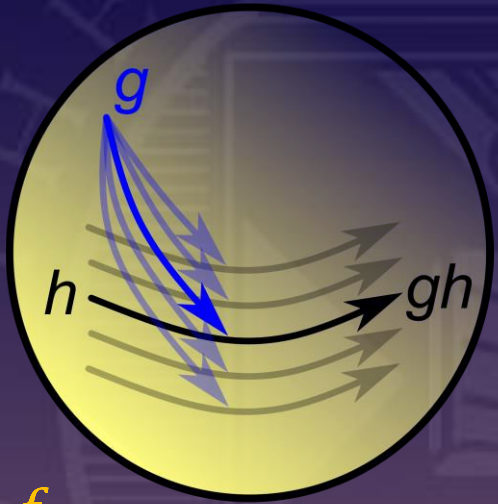
- Since
$$g_{\mu\nu}g^{\nu\rho} = \delta_\mu^\rho \Rightarrow \frac{\partial g_{\mu\nu}}{\partial q^\sigma} g^{\nu\rho} + g_{\mu\nu}\frac{\partial g^{\nu\rho}}{\partial q^\sigma} = 0 \Rightarrow \frac{\partial g^{\alpha\rho}}{\partial q^\sigma} = -g^{\alpha\mu}\frac{\partial g_{\mu\nu}}{\partial q^\sigma} g^{\nu\rho}$$
- and $P_\sigma = g_{\sigma\nu}\,\dot{Q}^\nu$ we may write Hamilton's equations in second-order form
$$\dot{P}_\sigma = \dot{g}_{\sigma\nu}\dot{Q}^\nu + g_{\sigma\nu}\ddot{Q}^\nu = \frac{1}{2}\left(\frac{\partial g_{\sigma\beta}}{\partial q^\alpha} + \frac{\partial g_{\sigma\alpha}}{\partial q^\beta}\right)\dot{Q}^\alpha\dot{Q}^\beta + g_{\sigma\nu}\ddot{Q}^\nu$$
$$= -\frac{1}{2}\frac{\partial g^{\mu\nu}}{\partial q^\sigma} g_{\mu\alpha}\dot{Q}^\alpha g_{\nu\beta}\dot{Q}^\beta = \frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial q^\sigma}\dot{Q}^\alpha\dot{Q}^\beta$$
$$\ddot{Q}^\mu = -\frac{1}{2}g^{\mu\sigma}\left(\frac{\partial g_{\sigma\beta}}{\partial q^\alpha} + \frac{\partial g_{\alpha\sigma}}{\partial q^\beta} - \frac{\partial g_{\alpha\beta}}{\partial q^\sigma}\right)\dot{Q}^\alpha\dot{Q}^\beta = -\Gamma^\mu_{\alpha\beta}\dot{Q}^\alpha\dot{Q}^\beta$$

Riemannian Manifolds

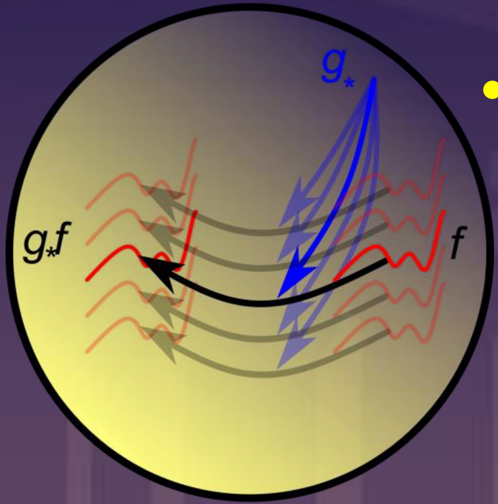
- The projection $\pi(c)$ of c onto the base manifold $\mathcal{M}$ satisfies the geodesic equation
$$\ddot{Q}^\mu + \Gamma^\mu_{\alpha\beta}\dot{Q}^\alpha\dot{Q}^\beta = 0$$
  - at least locally; the existence of global geodesics is a more subtle issue (c.f., the Hopf—Rinow theorem)
- The update step $\exp\colon \mathrm{Ham}(T^*\mathcal{M}) \to T^*\mathcal{M} \times T^*\mathcal{M} \colon \hat{T} \mapsto \mathrm{e}^{\hat{T}}$ on a Hamiltonian manifold corresponds to following a geodesic
  - This is called the exponential map
  - This is hard to do exactly in general
  - But in most applications g is "constant"

Lie Groups

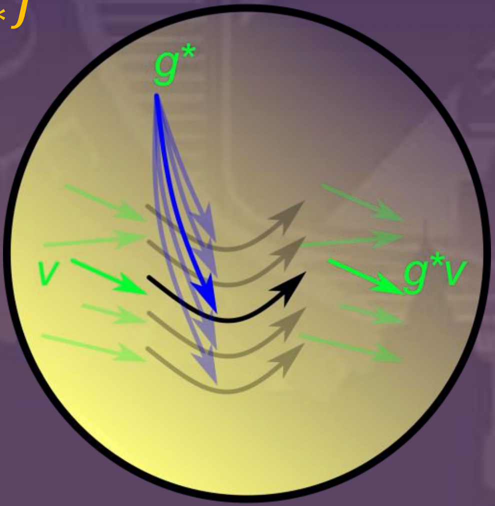
Lie Groups



- A <u>Lie group</u> $\mathcal{G}$ is a manifold with a group structure
  - Left action $g: \mathcal{G} \to \mathcal{G}: h \mapsto gh$



- Induced maps
  - $g_*: \Lambda_0(\mathcal{G}) \to \Lambda_0(\mathcal{G}): f \mapsto g_* f$
    $(g_* f)(h) = f(gh)$
  - $g^*: \Gamma(\mathcal{G}) \to \Gamma(\mathcal{G}): v \mapsto g^* v$
    $(g^* v) f = v(g_* f)$



Lie Algebras

- This allows us to define left-invariant vector fields
$$v = g^* v$$
  - These are "constant" in a natural way
  - They are not constant in any chart
- Every such field is determined by its value at the origin
$$v\Big|_g = (g^* v)\Big|_1$$
  - The space of left-invariant fields has dimension $\dim \mathcal{G}$
- The commutator of two left-invariant vector fields is itself left-invariant
$$[e_j, e_k] = c_{jk}^i e_i$$
  - c_{jk}^i are structure constants
  - The space of left-invariant vector fields is the Lie algebra $\mathfrak{g}$
  - Maurer—Cartan forms are duals of left-invariant 1-forms
$$\theta^i(e_j) = \delta_j^i$$
  - These satisfy the Maurer—Cartan equations
$$d\theta^i = -\frac{1}{2} c_{jk}^i \theta^j \wedge \theta^k$$

Symplectic Structure of Lie Groups

- There is a left-invariant Liouville form $\vartheta \in \Lambda_1(T^*\mathcal{G})$

$$\vartheta = \sum_{j=1}^{\dim \mathcal{G}} p_j \theta^j$$

- This gives the left-invariant fundamental form

$$\omega = d\vartheta = \sum_{j=1}^{\dim \mathcal{G}} \left(dp_j \wedge \theta^j + p_j d\theta^j\right) = \sum_{j=1}^{\dim \mathcal{G}} \left(dp_j \wedge \theta^j - c^j_{k\ell} p_j \theta^j \wedge \theta^\ell\right)$$

- Most Lie groups admit a pseudo-Riemannian metric
  - For matrix groups $g(u,v) = -\mathrm{tr}\big(U(u) \cdot U(v)\big)$ where $U: \mathfrak{g} \to \mathrm{GL}(n)$ is a representation of the Lie algebra $\mathfrak{g}$
  - This induces a metric for a connected compact Lie group in terms of minimal geodesics (the exponential map)

HMC on Lie Groups

Embedding

- Global quantities are nice, but how to represent them in a computer?
- For numerical computations we would like to express global quantities in global coordinates
- This may be done by embedding the manifold in a higher-dimensional Euclidean space
  - This is always possible by the Whitney embedding theorem which states that any manifold $\mathcal{M}$ may be embedded in $\mathbb{R}^n$ with $n \geq 2 \dim \mathcal{M}$
- For matrix Lie groups we may use the obvious embeddings

$$\mathrm{SO}(n, \mathbb{R}) \hookrightarrow \mathrm{GL}(n, \mathbb{R}) \hookrightarrow \mathbb{R}^{n^2}$$

$$\mathrm{SU}(n) \hookrightarrow \mathrm{GL}(n, \mathbb{C}) \hookrightarrow \mathbb{C}^{n^2}$$

  - This has the nice property that group multiplication is matrix multiplication
  - Moreover, the exponential map becomes a matrix exponential

More Embedding

- The embedding is just the defining representation
$$U: \mathrm{SU}(n) \hookrightarrow \mathbb{C}^{n^2}$$
  - We may consider this as a collection of component maps
  $$U_{ab}: \mathrm{SU}(n) \to \mathbb{R}: g \mapsto U_{ab}(g)$$
  - The matrix generators are $T_i = (e_i U)|_1$
  - Using left-invariance we have $(e_i U)_g = U(g) \cdot T_i$
  - We write $U: \mathfrak{su}(n) \hookrightarrow \mathbb{C}^{n^2}$ following the usual abuse of notation
- The computation of $\hat{V}$ is an application of the chain rule
$$\hat{V} = -(e_i V)\frac{\partial}{\partial p_i}; \quad U(\hat{V})\Big|_g = -\mathrm{tr}(\partial V_U \cdot U(g) \cdot T_i) g^{ij} T_j$$
  - with $V = V_U \circ U$ and $(\partial V_U)_{ab} = \frac{\partial V_U}{\partial U_{ba}}$
    - note the implicit transpose!

Summary

- The following table summaries the Hamiltonian system on a Lie group $\mathcal{G}$
  - The cotangent bundle (phase space) is $T^*\mathcal{G} = \mathcal{G} \times \mathfrak{g}$ in this case, where $\mathfrak{g} = T_1^*\mathcal{G}$ is the Lie algebra

Hamiltonian System	$\mathcal{G} \times \mathfrak{g}$	$\mathrm{GL}(n) \times \mathbb{C}^{n^2}$
Position	$g \in \mathcal{G}$	$U \in U(\mathcal{G}) \subseteq \mathrm{GL}(n) \subset \mathbb{C}^{n^2}$
Generators	$e_i(U(g))\Big\vert_{g=1} \in \mathfrak{g}$	$T_i = \partial U(e_i)\Big\vert_1 \in \mathbb{C}^{n^2}$
Momentum	$p = p^i e_i\Big\vert_1 \in \mathfrak{g}$	$P = p^i T_i \in \mathbb{C}^{n^2}$
Potential energy	$V: \mathcal{G} \to \mathbb{R}$	$V = V_U \circ U; \quad V_{\mathrm{U}}: \mathbb{C}^{n^2} \to \mathbb{R}$
Kinetic energy	$T(p) = \frac{1}{2} g(p, p)$	$T = T_U \circ U; \quad T_U(P) = -\frac{1}{2} \operatorname{tr} P^2$

HMC on Homogeneous Spaces

Homogeneous Spaces

- A <u>homogeneous space</u> is acted upon <u>transitively</u> by a Lie group
  - $\forall x, y \in \mathcal{M}$ there is a $g \in \mathcal{G}$ such that $x = gy$
- The action need not be <u>free</u>
  - Unlike for the group acting on itself
  - The <u>stabilizer</u> subgroup $\mathcal{S}_y$ of $y \in \mathcal{M}$ is
    $$S_y = \{g: x = gy\} \subseteq \mathcal{G}$$
    - The stabilizer subgroups of different points are isomorphic, $\mathcal{S}_y \cong \mathcal{S}$, but not equal in general
  - We find that $\mathcal{M} \cong \mathcal{G}/\mathcal{S}$

Hamiltonian Reduction

- Weinstein and Marsden introduced the Hamiltonian Reduction $\mathcal{M} // \mathcal{N}$
  - This defines "quotient phase space" over the quotient manifold $\mathcal{M}/\mathcal{N}$ induced from the cotangent bundle over $\mathcal{M}$
  - This allows us to construct a natural phase space over
    - Spheres $S_n = SO(n+1)/SO(n)$
    - Complex projective spaces $\mathbb{CP}^n = SU(n+1)/S(U(1) \times U(n))$
    - Oriented Grassmannians $\widetilde{Gr}(k,n) = SO(n)/(SO(k) \times SO(n-k))$
  - and many others
    - For details see Alessandro Barp, A D Kennedy, and Mark Girolami, "Hamiltonian Monte Carlo on Symmetric and Homogeneous Spaces via Symplectic Reduction", arXiv:1903.02699 (2019)

Hamiltonian Reduction

- The recipe is
  - Set the conserved momenta corresponding to $\mathcal{N}$ to zero
    - The conserved momenta correspond to the "momentum map"
    - If the momenta in $\mathcal{N}$ are non-zero then there are additional "magnetic forces"
      - "Centrifugal forces" may be a better name
  - Follow the Hamiltonian trajectories in $T^*\mathcal{M}$
  - Sample points in $\mathcal{M}/\mathcal{N}$ by "forgetting" the part in $\mathcal{N}$

Hamiltonian Reduction

- Examples
  - If q is a matrix in the defining representation of $\mathcal{M} = SO(n+1)$ and the representation of $\mathcal{N} = SO(n)$ is embedded in the lower right block, then a point in S_n is obtained by taking the first column of q
  - For $\mathbb{CP}^n$ the first column of a matrix defining representation of $SU(n+1)$ provides a suitable point <u>up to a phase</u>
    - This is typical of projective spaces
  - For Grassmannians we can may extract the points in the quotient using Plücker coordinates (ratios of determinants)

Questions?