# An Introduction to Lie Groups
# and Symplectic Geometry

A series of nine lectures on Lie groups and symplectic geometry delivered at the Regional Geometry Institute in Park City, Utah, 24 June–20 July 1991.

by

Robert L. Bryant
Duke University
Durham, NC
bryant@math.duke.edu

*This is an unofficial version of the notes and was last modified on 20 September 1993. The .dvi file for this preprint will be available by anonymous ftp from* `publications.math.duke.edu` *in the directory* `bryant` *until the manuscript is accepted for publication. You should get the ReadMe file first to see if the version there is more recent than this one.*

*Please send any comments, corrections or bug reports to the above e-mail address.*

# Introduction

These are the lecture notes for a short course entitled "Introduction to Lie groups and symplectic geometry" which I gave at the 1991 Regional Geometry Institute at Park City, Utah starting on 24 June and ending on 11 July.

The course really was designed to be an introduction, aimed at an audience of students who were familiar with basic constructions in differential topology and rudimentary differential geometry, who wanted to get a feel for Lie groups and symplectic geometry. My purpose was not to provide an exhaustive treatment of either Lie groups, which would have been impossible even if I had had an entire year, or of symplectic manifolds, which has lately undergone something of a revolution. Instead, I tried to provide an introduction to what I regard as the basic concepts of the two subjects, with an emphasis on examples which drove the development of the theory.

I deliberately tried to include a few topics which are not part of the mainstream subject, such as Lie's reduction of order for differential equations and its relation with the notion of a solvable group on the one hand and integration of ODE by quadrature on the other. I also tried, in the later lectures to introduce the reader to some of the global methods which are now becoming so important in symplectic geometry. However, a full treatment of these topics in the space of nine lectures beginning at the elementary level was beyond my abilities.

After the lectures were over, I contemplated reworking these notes into a comprehensive introduction to modern symplectic geometry and, after some soul-searching, finally decided against this. Thus, I have contented myself with making only minor modifications and corrections, with the hope that an interested person could read these notes in a few weeks and get some sense of what the subject was about.

An essential feature of the course was the exercise sets. Each set begins with elementary material and works up to more involved and delicate problems. My object was to provide a path to understanding of the material which could be entered at several different levels and so the exercises vary greatly in difficulty. Many of these exercise sets are obviously too long for any person to do them during the three weeks the course, so I provided extensive hints to aid the student in completing the exercises after the course was over.

I want to take this opportunity to thank the many people who made helpful suggestions for these notes both during and after the course. Particular thanks goes to Karen Uhlenbeck and Dan Freed, who invited me to give an introductory set of lectures at the RGI, and to my course assistant, Tom Ivey, who provided invaluable help and criticism in the early stages of the notes and tirelessly helped the students with the exercises. While the faults of the presentation are entirely my own, without the help, encouragement, and proofreading contributed by these folks and others, neither these notes nor the course would never have come to pass.

**Background Material and Basic Terminology.** In these lectures, I assume that the reader is familiar with the basic notions of manifolds, vector fields, and differential forms. All manifolds will be assumed to be both second countable and Hausdorff. Also, unless I say otherwise, I generally assume that all maps and manifolds are $C^\infty$.

Since it came up several times in the course of the course of the lectures, it is probably worth emphasizing the following point: A *submanifold* of a smooth manifold $X$ is, by definition, a pair $(S, f)$ where $S$ is a smooth manifold and $f: S \to X$ is a one-to-one immersion. In particular, $f$ need not be an embedding.

The notation I use for smooth manifolds and mappings is fairly standard, but with a few slight variations:

If $f: X \to Y$ is a smooth mapping, then $f': TX \to TY$ denotes the induced mapping on tangent bundles, with $f'(x)$ denoting its restriction to $T_x X$. (However, I follow tradition when $X = \mathbb{R}$ and let $f'(t)$ stand for $f'(t)(\partial/\partial t)$ for all $t \in \mathbb{R}$. I trust that this abuse of notation will not cause confusion.)

For any vector space $V$, I generally use $A^p(V)$ (instead of, say, $\Lambda^p(V^*)$) to denote the space of alternating (or exterior) p-forms on $V$. For a smooth manifold $M$, I denote the space of smooth, alternating $p$-forms on $M$ by $\mathcal{A}^p(M)$. The algebra of all (smooth) differential forms on $M$ is denoted by $\mathcal{A}^*(M)$.

I generally reserve the letter $d$ for the exterior derivative $d: \mathcal{A}^p(M) \to \mathcal{A}^{p+1}(M)$.

For any vector field $X$ on $M$, I will denote *left-hook with $X$* (often called *interior product with $X$*) by the symbol $X \lrcorner$. This is the graded derivation of degree $-1$ of $\mathcal{A}^*(M)$ which satisfies $X \lrcorner (df) = Xf$ for all smooth functions $f$ on $M$. For example, the Cartan formula for the Lie derivative of differential forms is written in the form

$$\mathcal{L}_X \phi = X \lrcorner d\phi + d(X \lrcorner \phi).$$

**Jets.** Occasionally, it will be convenient to use the language of jets in describing certain constructions. Jets provide a coordinate free way to talk about the Taylor expansion of some mapping up to a specified order. No detailed knowledge about these objects will be needed in these lectures, so the following comments should suffice:

If $f$ and $g$ are two smooth maps from a manifold $X^m$ to a manifold $Y^n$, we say that $f$ and $g$ *agree to order $k$ at $x \in X$* if, first, $f(x) = g(x) = y \in Y$ and, second, when $u: U \to \mathbb{R}^m$ and $v: V \to \mathbb{R}^n$ are local coordinate systems centered on $x$ and $y$ respectively, the functions $F = v \circ f \circ u^{-1}$ and $G = v \circ g \circ u^{-1}$ have the same Taylor series at $0 \in \mathbb{R}^m$ up to and including order $k$. Using the Chain Rule, it is not hard to show that this condition is independent of the choice of local coordinates $u$ and $v$ centered at $x$ and $y$ respectively.

The notation $f \equiv_{x,k} g$ will mean that $f$ and $g$ agree to order $k$ at $x$. This is easily seen to define an equivalence relation. Denote the $\equiv_{x,k}$-equivalence class of $f$ by $j^k(f)(x)$, and call it the *k-jet* of $f$ at $x$.

For example, knowing the 1-jet at $x$ of a map $f: X \to Y$ is equivalent to knowing both $f(x)$ and the linear map $f'(x): T_x \to T_{f(x)}Y$.

The set of $k$-jets of maps from $X$ to $Y$ is usually denoted by $J^k(X,Y)$. It is not hard to show that $J^k(X,Y)$ can be given a unique smooth manifold structure in such a way that, for any smooth $f\colon X \to Y$, the obvious map $j^k(f)\colon X \to J^k(X,Y)$ is also smooth.

These jet spaces have various functorial properties which we shall not need at all. The main reason for introducing this notion is to give meaning to concise statements like "The critical points of $f$ are determined by its 1-jet", "The curvature at $x$ of a Riemannian metric $g$ is determined by its 2-jet at $x$", or, from Lecture 8, "The integrability of an almost complex structure $J\colon TX \to TX$ is determined by its 1-jet". Should the reader wish to learn more about jets, I recommend the first two chapters of [GG].

**Basic and Semi-Basic.** Finally, I use the following terminology: If $\pi\colon V \to X$ is a smooth submersion, a $p$-form $\phi \in \mathcal{A}^p(V)$ is said to be $\pi$-*basic* if it can be written in the form $\phi = \pi^*(\varphi)$ for some $\varphi \in \mathcal{A}^p(X)$ and $\pi$-*semi-basic* if, for any $\pi$-vertical*vector field $X$, we have $X \lrcorner \phi = 0$. When the map $\pi$ is clear from context, the terms "basic" or "semi-basic" are used.

It is an elementary result that if the fibers of $\pi$ are connected and $\phi$ is a $p$-form on $V$ with the property that both $\phi$ and $d\phi$ are $\pi$-semi-basic, then $\phi$ is actually $\pi$-basic.

At least in the early lectures, we will need very little in the way of major theorems, but we will make extensive use of the following results:

• **The Implicit Function Theorem:** If $f\colon X \to Y$ is a smooth map of manifolds and $y \in Y$ is a regular value of $f$, then $f^{-1}(y) \subset X$ is a smooth embedded submanifold of $X$, with
$$T_x f^{-1}(y) = \ker(f'(x)\colon T_x X \to T_y Y)$$

• **Existence and Uniqueness of Solutions of ODE:** If $X$ is a vector field on a smooth manifold $M$, then there exists an open neighborhood $U$ of $\{0\} \times M$ in $\mathbb{R} \times M$ and a smooth mapping $F\colon U \to M$ with the following properties:

 i. $F(0,m) = m$ for all $m \in M$.

 ii. For each $m \in M$, the slice $U_m = \{t \in \mathbb{R} \,|\, (t,m) \in U\}$ is an open interval in $\mathbb{R}$ (containing 0) and the smooth mapping $\phi_m\colon U_m \to M$ defined by $\phi_m(t) = F(t,m)$ is an integral curve of $X$.

 iii. ( *Maximality* ) If $\phi\colon I \to M$ is any integral curve of $X$ where $I \subset \mathbb{R}$ is an interval containing 0, then $I \subset U_{\phi(0)}$ and $\phi(t) = \phi_{\phi(0)}(t)$ for all $t \in I$.

The mapping $F$ is called the (local) *flow* of $X$ and the open set $U$ is called the domain of the flow of $X$. If $U = \mathbb{R} \times M$, then we say that $X$ is *complete*.

Two useful properties of this flow are easy consequences of this existence and uniqueness theorem. First, the interval $U_{F(t,m)} \subset \mathbb{R}$ is simply the interval $U_m$ translated by $-t$. Second, $F(s+t,m) = F\left(s, F(t,m)\right)$ whenever $t$ and $s+t$ lie in $U_m$.

---

\* A vector field $X$ is $\pi$-vertical with respect to a map $\pi\colon V \to X$ if and only if $\pi'\left(X(v)\right) = 0$ for all $v \in V$

• **The Simultaneous Flow-Box Theorem:** If $X_1$, $X_2$, ..., $X_r$ are smooth vector fields on $M$ which satisfy the Lie bracket identities

$$[X_i, X_j] = 0$$

for all $i$ and $j$, and if $p \in M$ is a point where the $r$ vectors $X_1(p), X_2(p), \ldots, X_r(p)$ are linearly independent in $T_p M$, then there exists a local coordinate system $x^1, x^2, \ldots, x^n$ on an open neighborhood $U$ of $p$ so that, on $U$,

$$X_1 = \frac{\partial}{\partial x^1}, \qquad X_2 = \frac{\partial}{\partial x^2}, \qquad \ldots, \qquad X_r = \frac{\partial}{\partial x^r}.$$

The Simultaneous Flow-Box Theorem has two particularly useful consequences. Before describing them, we introduce an important concept.

Let $M$ be a smooth manifold and let $E \subset TM$ be a smooth subbundle of rank $p$. We say that $E$ is *integrable* if, for any two vector fields $X$ and $Y$ on $M$ which are sections of $E$, their Lie bracket $[X, Y]$ is also a section of $E$.

• **The Local Frobenius Theorem:** If $M^n$ is a smooth manifold and $E \subset TM$ is a smooth, integrable sub-bundle of rank $r$, then every $p$ in $M$ has a neighborhood $U$ on which there exist local coordinates $x^1, \ldots, x^r, y^1, \ldots, y^{n-r}$ so that the sections of $E$ over $U$ are spanned by the vector fields

$$\frac{\partial}{\partial x^1}, \qquad \frac{\partial}{\partial x^2}, \qquad \ldots, \qquad \frac{\partial}{\partial x^r}.$$

Associated to this local theorem is the following global version:

• **The Global Frobenius Theorem:** Let $M$ be a smooth manifold and let $E \subset TM$ be a smooth, integrable subbundle of rank $r$. Then for any $p \in M$, there exists a connected $r$-dimensional submanifold $L \subset M$ which contains $p$, which satisfies $T_q L = E_q$ for all $q \in S$, and which is maximal in the sense that any connected $r'$-dimensional submanifold $L' \subset M$ which contains $p$ and satisfies $T_q L' \subset E_q$ for all $q \in L'$ is a submanifold of $L$.

The submanifolds $L$ provided by this theorem are called the *leaves* of the sub-bundle $E$. (Some books call a sub-bundle $E \subset TM$ a *distribution* on $M$, but I avoid this since "distribution" already has a well-established meaning in analysis.)

# Contents

# Lecture 1:

# Introduction: Symmetry and Differential Equations

Consider the classical equations of motion for a particle in a conservative force field

$$\ddot{x} = -\mathrm{grad}\ V(x),$$

where $V \colon \mathbb{R}^n \to \mathbb{R}$ is some function on $\mathbb{R}^n$. If $V$ is proper (i.e. the inverse image under $V$ of a compact set is compact, as when $V(x) = |x|^2$), then, to a first approximation, $V$ is the potential for the motion of a ball of unit mass rolling around in a cup, moving only under the influence of gravity. For a general function $V$ we have only the grossest knowledge of how the solutions to this equation ought to behave.

Nevertheless, we can say a few things. The total energy ($=$ kinetic plus potential) is given by the formula $E = \frac{1}{2}|\dot{x}|^2 + V(x)$ and is easily shown to be constant on any solution (just differentiate $E\big(x(t)\big)$ and use the equation). Since, $V$ is proper, it follows that $x$ must stay inside a compact set $V^{-1}\big([0, E(x(0))]\big)$, and so the orbits are bounded. Without knowing any more about $V$, one can show (see Lecture 4 for a precise statement) that the motion has a certain "recurrent" behaviour: The trajectory resulting from "most" initial positions and velocities tends to return, infinitely often, to a small neighborhood of the initial position and velocity. Beyond this, very little is known is known about the behaviour of the trajectories for generic $V$.

Suppose now that the potential function $V$ is rotationally symmetric, i.e. that $V$ depends only on the distance from the origin and, for the sake of simplicity, let us take $n = 3$ as well. This is classically called the case of a central force field in space. If we let $V(x) = \frac{1}{2}v(|x|^2)$, then the equations of motion become

$$\ddot{x} = -v'\big(|x|^2\big)\ x.$$

As conserved quantities, i.e., functions of the position and velocity which stay constant on any solution of the equation, we still have the energy $E = \frac{1}{2}\big(|\dot{x}|^2 + v(|x|^2)\big)$, but is it also easy to see that the vector-valued function $x \times \dot{x}$ is conserved, since

$$\frac{d}{dt}(x \times \dot{x}) = \dot{x} \times \dot{x} - x \times v'(|x|^2)\ x.$$

Call this vector-valued function $\mu$. We can think of $E$ and $\mu$ as functions on the phase space $\mathbb{R}^6$. For generic values of $E_0$ and $\mu_0$, the simultaneous level set

$$\Sigma_{E_0,\mu_0} = \{\ (x, \dot{x}) \mid E(x, \dot{x}) = E_0,\ \mu(x, \dot{x}) = \mu_0\ \}$$

of these functions cut out a surface $\Sigma_{E_0,\mu_0} \subset \mathbb{R}^6$ and any integral of the equations of motion must lie in one of these surfaces. Since we know a great deal about integrals of ODEs on

surfaces, This problem is very tractable. (see Lecture 4 and its exercises for more details on this.)

The function $\mu$, known as the *angular momentum*, is called a *first integral* of the second-order ODE for $x(t)$, and somehow seems to correspond to the rotational symmetry of the original ODE. This vague relationship will be considerably sharpened and made precise in the upcoming lectures.

The relationship between symmetry and solvability in differential equations is profound and far reaching. The subjects which are now known as Lie groups and symplectic geometry got their beginnings from the study of symmetries of systems of ordinary differential equations and of integration techniques for them.

By the middle of the nineteenth century, Galois theory had clarified the relationship between the solvability of polynomial equations by radicals and the group of "symmetries" of the equations. Sophus Lie set out to do the same thing for differential equations and their symmetries.

Here is a "dictionary" showing the (rough) correspondence which Lie developed between these two achievements of nineteenth century mathematics.

| | |
|---|---|
| Galois theory | infinitesimal symmetries |
| finite groups | continuous groups |
| polynomial equations | differential equations |
| solvable by radicals | solvable by quadrature |

Although the full explanation of these correspondances must await the later lectures, we can at least begin the story in the simplest examples as motivation for developing the general theory. This is what I shall do for the rest of today's lecture.

**Classical Integration Techniques.** The very simplest ordinary differential equation that we ever encounter is the equation

$$(1) \qquad\qquad\qquad \dot{x}(t) = \alpha(t)$$

where $\alpha$ is a known function of $t$. The solution of this differential equation is simply

$$x(t) = x_0 + \int_0^x \alpha(\tau)\, d\tau.$$

The process of computing an integral was known as "quadrature" in the classical literature (a reference to the quadrangles appearing in what we now call Riemann sums), so it was said that (1) was "solvable by quadrature". Note that, once one finds a particular solution, all of the others are got by simply translating the particular solution by a constant, in this case, by $x_0$. Alternatively, one could say that the equation (1) itself was invariant under "translation in $x$".

The next most trivial case is the homogeneous linear equation

$$(2) \qquad\qquad\qquad \dot{x} = \beta(t)\, x.$$

This equation is invariant under scale transformations $x \mapsto rx$. Since the mapping $\log : \mathbb{R}^+ \to \mathbb{R}$ converts scaling to translation, it should not be surprising that the differential equation (2) is also solvable by a quadrature:

$$x(t) = x_0 e^{\int_0^t \beta(\tau)\, d\tau}.$$

Note that, again, the symmetries of the equation suffice to allow us to deduce the general solution from the particular.

Next, consider an equation where the right hand side is an affine function of $x$,

(3) $$\dot{x} = \alpha(t) + \beta(t)\, x.$$

This equation is still solvable in full generality, using two quadratures. For, if we set

$$x(t) = u(t) e^{\int_0^t \beta(\tau) d\tau},$$

then $u$ satisfies $\dot{u} = \alpha(t) e^{-\int_0^t \beta(\tau) d\tau}$, which can be solved for $u$ by another quadrature. It is not at all clear why one can somehow "combine" equations (1) and (2) and get an equation which is still solvable by quadrature, but this will become clear in Lecture 3.

Now consider an equation with a quadratic right-hand side, the so-called Riccati equation:

(4) $$\dot{x} = \alpha(t) + 2\beta(t)x + \gamma(t)x^2.$$

It can be shown that there is no method for solving this by quadratures and algebraic manipulations alone. However, there is a way of obtaining the general solution from a particular solution. If $s(t)$ is a particular solution of (4), try the *ansatz* $x(t) = s(t) + 1/u(t)$. The resulting differential equation for $u$ has the form (3) and hence is solvable by quadratures.

The equation (4), known as the Riccati equation, has an extensive history, and we will return to it often. Its remarkable property, that given one solution we can obtain the general solution, should be contrasted with the case of

(5) $$\dot{x} = \alpha(t) + \beta(t)x + \gamma(t)x^2 + \delta(t)x^3.$$

For equation (5), one solution does not give you the rest of the solutions. There is in fact a world of difference between this and the Riccati equation, although this is far from evident looking at them.

Before leaving these simple ODE, we note the following curious progression: If $x_1$ and $x_2$ are solutions of an equation of type (1), then clearly the difference $x_1 - x_2$ is constant. Similarly, if $x_1$ and $x_2 \neq 0$ are solutions of an equation of type (2), then the ratio $x_1/x_2$ is constant. Furthermore, if $x_1$, $x_2$, and $x_3 \neq x_1$ are solutions of an equation of type (3),

then the expression $(x_1 - x_2)/(x_1 - x_3)$ is constant. Finally, if $x_1$, $x_2$, $x_3 \neq x_1$, and $x_4 \neq x_2$ are solutions of an equation of type (4), then the *cross-ratio*

$$\frac{(x_1 - x_2)(x_4 - x_3)}{(x_1 - x_3)(x_4 - x_2)}$$

is constant. There is no such corresponding expression (for any number of particular solutions) for equations of type (5). The reason for this will be made clear in Lecture 3. For right now, we just want to remark on the fact that the linear fractional transformations of the real line, a group isomorphic to $\mathrm{SL}(2, \mathbb{R})$, are exactly the transformations which leave fixed the cross-ratio of any four points. As we shall see, the group $\mathrm{SL}(2, \mathbb{R})$ is closely connected with the Riccati equation and it is this connection which accounts for many of the special features of this equation.

We will conclude this lecture by discussing the group of rigid motions in Euclidean 3-space. These are transformations of the form

$$T(\mathbf{x}) = \mathbf{R}\,\mathbf{x} + \mathbf{t},$$

where $\mathbf{R}$ is a rotation in $\mathbb{E}^3$ and $\mathbf{t} \in \mathbb{E}^3$ is any vector. It is easy to check that the set of rigid motions form a group under composition which is, in fact, isomorphic to the group of 4-by-4 matrices

$$\left\{ \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \,\middle|\, {}^t\mathbf{R}\,\mathbf{R} = \mathbf{I}_3, \ \mathbf{t} \in \mathbb{R}^3 \right\}.$$

(Topologically, the group of rigid motions is just the product $\mathrm{O}(3) \times \mathbb{R}^3$.)

Now, suppose that we are asked to solve for a curve $\mathbf{x} \colon \mathbb{R} \to \mathbb{R}^3$ with a prescribed curvature $\kappa(t)$ and torsion $\tau(t)$. If $\mathbf{x}$ were such a curve, then we could calculate the curvature and torsion by defining an oriented orthonormal basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ along the curve, satisfying $\dot{x} = \mathbf{e}_1$, $\dot{\mathbf{e}}_1 = \kappa \mathbf{e}_2$, $\dot{\mathbf{e}}_2 = -\kappa \mathbf{e}_1 + \tau \mathbf{e}_3$. (Think of the torsion as measuring how $\mathbf{e}_2$ falls away from the $\mathbf{e}_1\mathbf{e}_2$-plane.) Form the 4-by-4 matrix

$$X = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{x} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

(where we always think of vectors in $\mathbb{R}^3$ as columns). Then we can express the ODE for prescribed curvature and torsion as

$$\dot{X} = X \begin{pmatrix} 0 & -\kappa & 0 & 1 \\ \kappa & 0 & -\tau & 0 \\ 0 & \tau & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We can think of this as a linear system of equations for a curve $X(t)$ in the group of rigid motions.

L.1.4

It is going to turn out that, just as in the case of the Riccati equation, the prescribed curvature and torsion equations cannot be solved by algebraic manipulations and quadrature alone. However, once we know one solution, all other solutions for that particular $(\kappa(t), \tau(t))$ can be obtained by rigid motions. In fact, though, we are going to see that one does not have to know a solution to the full set of equations before finding the rest of the solutions by quadrature, but only a solution to an equation connected to SO(3) just in the same way that the Riccati equation is connected to SL(2, $\mathbb{R}$), the group of transformations of the line which fix the cross-ratio of four points.

In fact, as we are going to see, $\mu$ "comes from" the group of rotations in three dimensions, which are symmetries of the ODE because they preserve $V$. That is, $V(R(x)) = V(x)$ whenever $R$ is a linear transformation satisfying $R^t R = I$. The equation $R^t R = I$ describes a locus in the space of $3 \times 3$ matrices. Later on we will see this locus is a smooth compact 3-manifold, which is also a group, called $O(3)$. The group of rotations, and generalizations thereof, will play a central role in subsequent lectures.

# Lecture 2:

# Lie Groups and Lie Algebras

**Lie Groups.** In this lecture, I define and develop some of the basic properties of the central objects of interest in these lectures: Lie groups and Lie algebras.

**Definition 1:** A *Lie group* is a pair $(G, \mu)$ where $G$ is a smooth manifold and $\mu \colon G \times G \to G$ is a smooth mapping which gives $G$ the structure of a group.

When the multiplication $\mu$ is clear from context, we usually just say "$G$ is a Lie group." Also, for the sake of notational sanity, I will follow the practice of writing $\mu(a, b)$ simply as $ab$ whenever this will not cause confusion. I will usually denote the multiplicative identity by $e \in G$ and the multiplicative inverse of $a \in G$ by $a^{-1} \in G$.

Most of the algebraic constructions in the theory of abstract groups have straightforward analogues for Lie groups:

**Definition 2:** A *Lie subgroup* of a Lie group $G$ is a subgroup $H \subset G$ which is also a submanifold of $G$. A *Lie group homomorphism* is a group homomorphism $\phi \colon H \to G$ which is also a smooth mapping of the underlying manifolds.

Here is the prototypical example of a Lie group:

**Example : The General Linear Group.** The (real) *general linear group* in dimension $n$, denoted $\mathrm{GL}(n, \mathbb{R})$, is the set of invertible $n$-by-$n$ real matrices regarded as an open submanifold of the $n^2$-dimensional vector space of all $n$-by-$n$ real matrices with multiplication map $\mu$ given by matrix multiplication: $\mu(a, b) = ab$. Since the matrix product $ab$ is defined by a formula which is polynomial in the matrix entries of $a$ and $b$, it is clear that $\mathrm{GL}(n, \mathbb{R})$ is a Lie group.

Actually, if $V$ is any finite dimensional real vector space, then $\mathrm{GL}(V)$, the set of bijective linear maps $\phi \colon V \to V$, is an open subset of the vector space $\mathrm{End}(V) = V \otimes V^*$ and becomes a Lie group when endowed with the multiplication $\mu \colon \mathrm{GL}(V) \times \mathrm{GL}(V) \to \mathrm{GL}(V)$ given by composition of maps: $\mu(\phi_1, \phi_2) = \phi_1 \circ \phi_2$. If $\dim(V) = n$, then $\mathrm{GL}(V)$ is isomorphic (as a Lie group) to $\mathrm{GL}(n, \mathbb{R})$, though not canonically.

The advantage of considering abstract vector spaces $V$ rather than just $\mathbb{R}^n$ is mainly conceptual, but, as we shall see, this conceptual advantage is great. In fact, Lie groups of linear transformations are so fundamental that a special terminology is reserved for them:

**Definition 3:** A *(linear) representation* of a Lie group $G$ is a Lie group homomorphism $\rho \colon G \to \mathrm{GL}(V)$ for some vector space $V$ called the *representation space*. Such a representation is said to be *faithful* (resp., *almost faithful* ) if $\rho$ is one-to-one (resp., has 0-dimensional kernel).

It is a consequence of a theorem of Ado and Iwasawa that every connected Lie group has an almost faithful, finite-dimensional representation. (In one of the later exercises, we will construct a connected Lie group which has no *faithful*, finite-dimensional representation, so *almost faithful* is the best we can hope for.)

**Example: Vector Spaces.** Any vector space over $\mathbb{R}$ becomes a Lie group when the group "multiplication" is taken to be addition.

**Example: Matrix Lie Groups.** The Lie subgroups of $\mathrm{GL}(n, \mathbb{R})$ are called *matrix Lie groups* and play an important role in the theory. Not only are they the most frequently encountered, but, because of the theorem of Ado and Iwasawa, practically anything which is true for matrix Lie groups has an analog for a general Lie group. In fact, for the first pass through, the reader can simply imagine that all of the Lie groups mentioned are matrix Lie groups. Here are a few simple examples:

1. Let $A_n$ be the set of diagonal $n$-by-$n$ matrices with positive entries on the diagonal.

2. Let $N_n$ be the set of upper triangular $n$-by-$n$ matrices with all diagonal entries all equal to 1.

3. ($n = 2$ only) Let $\mathbb{C}^\bullet = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a^2 + b^2 > 0 \right\}$. Then $\mathbb{C}^\bullet$ is a matrix Lie group diffeomorphic to $S^1 \times \mathbb{R}$. (You should check that this is actually a subgroup of $\mathrm{GL}(2, \mathbb{R})$!)

4. Let $\mathrm{GL}_+(n, \mathbb{R}) = \{ a \in \mathrm{GL}(n, \mathbb{R}) \mid \det(a) > 0 \}$

There are more interesting examples, of course. A few of these are

$$\mathrm{SL}(n, \mathbb{R}) = \{ a \in \mathrm{GL}(n, \mathbb{R}) \mid \det(a) = 1 \}$$
$$\mathrm{O}(n) = \{ a \in \mathrm{GL}(n, \mathbb{R}) \mid {}^t a\, a = I_n \}$$
$$\mathrm{SO}(n, \mathbb{R}) = \{ a \in \mathrm{O}(n) \mid \det(a) = 1 \}$$

which are known respectively as the *special linear group* , the *orthogonal group* , and the *special orthogonal group* in dimension $n$. In each case, one must check that the given subset is actually a subgroup and submanifold of $\mathrm{GL}(n, \mathbb{R})$. These are exercises for the reader. (See the problems at the end of this lecture for hints.)

A Lie group can have "wild" subgroups which cannot be given the structure of a Lie group. For example, $(\mathbb{R}, +)$ is a Lie group which contains totally disconnected, uncountable subgroups. Since all of our manifolds are second countable, such subgroups (by definition) cannot be given the structure of a (0-dimensional) Lie group.

However, it can be shown [Wa, pg. 110] that any *closed* subgroup of a Lie group $G$ is an embedded submanifold of $G$ and hence is a Lie subgroup. However, for reasons which will soon become apparent, it is disadvantageous to consider only closed subgroups.

**Example: A non-closed subgroup.** For example, even $\mathrm{GL}(n, \mathbb{R})$ can have Lie subgroups which are not closed. Here is a simple example: Let $\lambda$ be any irrational real number and define a homomorphism $\phi_\lambda: \mathbb{R} \to \mathrm{GL}(4, \mathbb{R})$ by the formula

$$
\phi_\lambda(t) = \begin{pmatrix} \cos t & -\sin t & 0 & 0 \\ \sin t & \cos t & 0 & 0 \\ 0 & 0 & \cos \lambda t & -\sin \lambda t \\ 0 & 0 & \sin \lambda t & \cos \lambda t \end{pmatrix}
$$

Then $\phi_\lambda$ is easily seen to be a one-to-one immersion so its image is a submanifold $G_\lambda \subset \mathrm{GL}(4, \mathbb{R})$ which is therefore a Lie subgroup. It is not hard to see that

$$
\overline{G_\lambda} = \left\{ \begin{pmatrix} \cos t & -\sin t & 0 & 0 \\ \sin t & \cos t & 0 & 0 \\ 0 & 0 & \cos s & -\sin s \\ 0 & 0 & \sin s & \cos s \end{pmatrix} \,\middle|\, s, t \in \mathbb{R} \right\}.
$$

Note that $G_\lambda$ is diffeomorphic to $\mathbb{R}$ while its closure in $\mathrm{GL}(4, \mathbb{R})$ is diffeomorphic to $S^1 \times S^1$!

It is also useful to consider matrix Lie groups with complex coefficients. However, complex matrix Lie groups are really no more general than real matrix Lie groups (though they may be more convenient to work with). To see why, note that we can write a complex $n$-by-$n$ matrix $A + Bi$ (where $A$ and $B$ are real $n$-by-$n$ matrices) as the $2n$-by-$2n$ matrix $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$. In this way, we can embed $\mathrm{GL}(n, \mathbb{C})$, the space of $n$-by-$n$ invertible complex matrices, as a closed submanifold of $\mathrm{GL}(2n, \mathbb{R})$. The reader should check that this mapping is actually a group homomorphism.

Among the more commonly encountered complex matrix Lie groups are the *complex special linear group*, denoted by $\mathrm{SL}(n, \mathbb{C})$, and the *unitary* and *special unitary groups*, denoted, respectively, as

$$
\mathrm{U}(n) = \{ a \in \mathrm{GL}(n, \mathbb{C}) \,|\, {}^*a\, a = I_n \}
$$
$$
\mathrm{SU}(n) = \{ a \in \mathrm{U}(n) \,|\, \det_{\mathbb{C}}(a) = 1 \}
$$

where ${}^*a = {}^t\bar{a}$ is the Hermitian adjoint of $a$. These groups will play an important role in what follows. The reader may want to familiarize himself with these groups by doing some of the exercises for this section.

**Basic General Properties.** If $G$ is a Lie group with $a \in G$, we let $L_a, R_a: G \to G$ denote the smooth mappings defined by

$$
L_a(b) = ab \qquad \text{and} \qquad R_a(b) = ba.
$$

**Proposition 1:** *For any Lie group $G$, the maps $L_a$ and $R_a$ are diffeomorphisms, the map $\mu: G \times G \to G$ is a submersion, and the inverse mapping $\iota: G \to G$ defined by $\iota(a) = a^{-1}$ is smooth.*

PROOF: By the axioms of group multiplication, $L_{a^{-1}}$ is both a left and right inverse to $L_a$. Since $(L_a)^{-1}$ exists and is smooth, $L_a$ is a diffeomorphism. The argument for $R_a$ is similar.

In particular, $L'_a\colon TG \to TG$ induces an isomorphism of tangent spaces $T_bG \overset{\sim}{\to} T_{ab}G$ for all $b \in G$ and $R'_a\colon TG \to TG$ induces an isomorphism of tangent spaces $T_bG \overset{\sim}{\to} T_{ba}G$ for all $b \in G$. Using the natural identification $T_{(a,b)}G \times G \simeq T_aG \oplus T_bG$, the formula for $\mu'(a,b)\colon T_{(a,b)}G \times G \to T_{ab}G$ is readily seen to be

$$\mu'(a,b)(v,w) = L'_a(w) + R'_b(v)$$

for all $v \in T_aG$ and $w \in T_bG$. In particular $\mu'(a,b)$ is surjective for all $(a,b) \in G \times G$, so $\mu\colon G \times G \to G$ is a submersion.

Then, by the Implicit Function Theorem, $\mu^{-1}(e)$ is a closed, embedded submanifold of $G \times G$ whose tangent space at $(a,b)$, by the above formula is

$$T_{(a,b)}\mu^{-1}(e) = \{(v,w) \in T_aG \times T_bG \,|\, L'_a(w) + R'_b(v) = 0\}.$$

Meanwhile, the group axioms imply that

$$\mu^{-1}(e) = \left\{(a, a^{-1}) \,|\, a \in G\right\},$$

which is precisely the graph of $\iota\colon G \to G$. Since $L'_a$ and $R'_a$ are isomorphisms at every point, it easily follows that the projection on the first factor $\pi_1\colon G \times G \to G$ restricts to $\mu^{-1}(e)$ to be a diffeomorphism of $\mu^{-1}(e)$ with $G$. Its inverse is therefore also smooth and is simply the graph of $\iota$. It follows that $\iota$ is smooth, as desired. $\qquad\square$

For any Lie group $G$, we let $G^\circ \subset G$ denote the connected component of $G$ which contains $e$. This is usually called the *identity component* of $G$.

**Proposition 2:** *For any Lie group $G$, the set $G^\circ$ is an open, normal subgroup of $G$. Moreover, if $U$ is any open neighborhood of $e$ in $G^\circ$, then $G^\circ$ is the union of the "powers" $U^n$ defined inductively by $U^1 = U$ and $U^{k+1} = \mu(U^k, U)$ for $k > 0$.*

PROOF: Since $G$ is a manifold, its connected components are open and path-connected, so $G^\circ$ is open and path-connected. If $\alpha, \beta\colon [0,1] \to G$ are two continuous maps with $\alpha(0) = \beta(0) = e$, then $\gamma\colon [0,1] \to G$ defined by $\gamma(t) = \alpha(t)\beta(t)^{-1}$ is a continuous path from $e$ to $\alpha(1)\beta(1)^{-1}$, so $G^\circ$ is closed under multiplication and inverse, and hence is a subgroup. It is a normal subgroup since, for any $a \in G$, the map

$$C_a = L_a \circ (R_a)^{-1}\colon G \to G$$

(conjugation by $a$) is a diffeomorphism which clearly fixes $e$ and hence fixes its connected component $G^\circ$ also.

Finally, let $U \subset G^\circ$ be any open neighborhood of $e$. For any $a \in G^\circ$, let $\gamma\colon [0,1] \to G$ be a path with $\gamma(0) = e$ and $\gamma(1) = a$. The open sets $\{L_{\gamma(t)}(U) \,|\, t \in [0,1]\}$ cover $\gamma([0,1])$,

so the compactness of $[0, 1]$ implies (via the Lebesgue Covering Lemma) that there is a finite subdivision $0 = t_0 < t_1 \cdots < t_n = 1$ so that $\gamma\big([t_k, t_{k+1}]\big) \subset L_{\gamma(t_k)}(U)$ for all $0 \le k < n$. But then each of the elements

$$a_k = \big(\gamma(t_k)^{-1}\big)\,\gamma(t_{k+1})$$

lies in $U$ and $a = \gamma(1) = a_0 a_1 \cdots a_{n-1} \in U^n$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

An immediate consequence of Proposition 2 is that, for a connected Lie group $H$, any Lie group homomorphism $\phi\colon H \to G$ is determined by its behavior on any open neighborhood of $e \in H$. We are soon going to show an even more striking fact, namely that, for connected $H$, any homomorphism $\phi\colon H \to G$ is determined by $\phi'(e)\colon T_e H \to T_e G$.

**The Adjoint Representation.** It is conventional to denote the tangent space at the identity of a Lie group by an appropriate lower case gothic letter. Thus, the vector space $T_e G$ is denoted $\mathfrak{g}$, the vector space $T_e \mathrm{GL}(n, \mathbb{R})$ is denoted $\mathfrak{gl}(n, \mathbb{R})$, etc.

For example, one can easily compute the tangent spaces at $e$ of the Lie groups defined so far. Here is a sample:

$$\mathfrak{sl}(n, \mathbb{R}) = \{a \in \mathfrak{gl}(n, \mathbb{R}) \,|\, \mathrm{tr}(a) = 0\}$$
$$\mathfrak{so}(n, \mathbb{R}) = \{a \in \mathfrak{gl}(n, \mathbb{R}) \,|\, a + {}^t a = 0\}$$
$$\mathfrak{u}(n, \mathbb{R}) = \{a \in \mathfrak{gl}(n, \mathbb{C}) \,|\, a + {}^t \bar{a} = 0\}$$

**Definition 4:** For any Lie group $G$, the *adjoint mapping* is the mapping $\mathrm{Ad}\colon G \to \mathrm{End}(\mathfrak{g})$ defined by

$$\mathrm{Ad}(a) = \Big(L_a \circ (R_a)^{-1}\Big)'(e)\colon T_e G \to T_e G.$$

As an example, for $G = \mathrm{GL}(n, \mathbb{R})$ it is easy to see that

$$\mathrm{Ad}(a)(x) = axa^{-1}$$

for all $a \in \mathrm{GL}(n, \mathbb{R})$ and $x \in \mathfrak{gl}(n, \mathbb{R})$. Of course, this formula is valid for any matrix Lie group.

The following proposition explains why the adjoint mapping is also called the adjoint representation.

**Proposition 3:** *The adjoint mapping is a linear representation* $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$.

PROOF: For any $a \in G$, let $C_a = L_a \circ R_{a^{-1}}$. Then $C_a\colon G \to G$ is a diffeomorphism which satisfies $C_a(e) = e$. In particular, $\mathrm{Ad}(a) = C_a'(e)\colon \mathfrak{g} \to \mathfrak{g}$ is an isomorphism and hence belongs to $\mathrm{GL}(\mathfrak{g})$.

L.2.5

The associative property of group multiplication implies $C_a \circ C_b = C_{ab}$, so the Chain Rule implies that $C_a'(e) \circ C_b'(e) = C_{ab}'(e)$. Hence, $\mathrm{Ad}(a)\mathrm{Ad}(b) = \mathrm{Ad}(ab)$, so Ad is a homomorphism.

It remains to show that Ad is smooth. However, if $C: G \times G \to G$ is defined by $C(a,b) = aba^{-1}$, then by Proposition 1, $C$ is a composition of smooth maps and hence is smooth. It follows easily that the map $c: G \times \mathfrak{g} \to \mathfrak{g}$ given by $c(a,v) = C_a'(e)(v) = \mathrm{Ad}(a)(v)$ is a composition of smooth maps. The smoothness of the map $c$ clearly implies the smoothness of $\mathrm{Ad}: G \to \mathfrak{g} \otimes \mathfrak{g}^*$. $\qquad\square$

**Left-invariant vector fields.** Because $L_a'$ induces an isomorphism from $\mathfrak{g}$ to $T_aG$ for all $a \in G$, it is easy to show that the map $\Psi: G \times \mathfrak{g} \to TG$ given by

$$\Psi(a,v) = L_a'(v)$$

is actually an isomorphism of vector bundles which makes the following diagram commute.

$$
\begin{array}{ccc}
G \times \mathfrak{g} & \xrightarrow{\Psi} & TG \\
\pi_1 \downarrow & & \downarrow \pi \\
G & \xrightarrow{id} & G
\end{array}
$$

Note that, in particular, $G$ is a parallelizable manifold. This implies, for example, that the only compact surface which can be given the structure of a Lie group is the torus $S^1 \times S^1$.

For each $v \in \mathfrak{g}$, we may use $\Psi$ to define a vector field $X_v$ on $G$ by the rule $X_v(a) = L_a'(v)$. Note that, by the Chain Rule and the definition of $X_v$, we have

$$L_a'(X_v(b)) = L_a'(L_b'(v)) = L_{ab}'(v) = X_v(ab).$$

Thus, the vector field $X_v$ is invariant under left translation by any element of $G$. Such vector fields turn out to be extremely useful in understanding the geometry of Lie groups, and are accorded a special name:

**Definition 5:** If $G$ is a Lie group, a *left-invariant vector field* on $G$ is a vector field $X$ on $G$ which satisfies $L_a'(X(b)) = X(ab)$.

For example, consider $\mathrm{GL}(n,\mathbb{R})$ as an open subset of the vector space of $n$-by-$n$ matrices with real entries. Here, $\mathfrak{gl}(n,\mathbb{R})$ is just the vector space of $n$-by-$n$ matrices with real entries itself and one easily sees that

$$X_v(a) = (a, av).$$

(Since $\mathrm{GL}(n,\mathbb{R})$ is an open subset of a vector space, namely, $\mathfrak{gl}(n,\mathbb{R})$, we are using the standard identification of the tangent bundle of $\mathrm{GL}(n,\mathbb{R})$ with $\mathrm{GL}(n,\mathbb{R}) \times \mathfrak{gl}(n,\mathbb{R})$.)

The following proposition determines all of the left-invariant vector fields on a Lie group.

**Proposition 4:** *Every left-invariant vector field $X$ on $G$ is of the form $X = X_v$ where $v = X(e)$ and hence is smooth. Moreover, such an $X$ is complete, i.e., the flow $\Phi$ associated to $X$ has domain $\mathbb{R} \times G$.*

PROOF: That every left-invariant vector field on $G$ has the stated form is an easy exercise for the reader. It remains to show that the flow of such an $X$ is complete, i.e., that for each $a \in G$, there exists a smooth curve $\gamma_a \colon \mathbb{R} \to G$ so that $\gamma_a(0) = a$ and $\gamma_a'(t) = X(\gamma_a(t))$ for all $t \in \mathbb{R}$.

It suffices to show that such a curve exists for $a = e$, since we may then define

$$\gamma_a(t) = a\,\gamma_e(t)$$

and see that $\gamma_a$ satisfies the necessary conditions: $\gamma_a(0) = a\,\gamma_e(0) = a$ and

$$\gamma_a'(t) = L_a'\left(\gamma_e'(t)\right) = L_a'\left(X\left(\gamma_e(t)\right)\right) = X\left(a\gamma_e(t)\right) = X\left(\gamma_a(t)\right).$$

Now, by the ODE existence theorem, there is an $\varepsilon > 0$ so that such a $\gamma_e$ can be defined on the interval $(-\varepsilon, \varepsilon) \subset \mathbb{R}$. If $\gamma_e$ could not be extended to all of $\mathbb{R}$, then there would be a maximum such $\varepsilon$. I will now show that there is no such maximum $\varepsilon$.

For each $s \in (-\varepsilon, \varepsilon)$, the curve $\alpha_s \colon (-\varepsilon + |s|, \varepsilon - |s|) \to G$ defined by

$$\alpha_s(t) = \gamma_e(s + t)$$

clearly satisfies $\alpha_s(0) = \gamma_e(s)$ and

$$\alpha_s'(t) = \gamma_e'(s + t) = X\left(\gamma_e(s + t)\right) = X\left(\alpha_s(t)\right),$$

so, by the ODE uniqueness theorem, $\alpha_s(t) = \gamma_e(s)\gamma_e(t)$. In particular, we have

$$\gamma_e(s + t) = \gamma_e(s)\gamma_e(t)$$

for all $s$ and $t$ satisfying $|s| + |t| < \varepsilon$.

Thus, I can extend the domain of $\gamma_e$ to $(-\frac{3}{2}\varepsilon, \frac{3}{2}\varepsilon)$ by the rule

$$\gamma_e(t) = \begin{cases} \gamma_e(-\frac{1}{2}\varepsilon)\gamma_e(t + \frac{1}{2}\varepsilon) & \text{if } t \in (-\frac{3}{2}\varepsilon, \frac{1}{2}\varepsilon); \\ \gamma_e(+\frac{1}{2}\varepsilon)\gamma_e(t - \frac{1}{2}\varepsilon) & \text{if } t \in (-\frac{1}{2}\varepsilon, \frac{3}{2}\varepsilon). \end{cases}$$

By our previous arguments, this extended $\gamma_e$ is still an integral curve of $X$, contradicting the assumption that $(-\varepsilon, \varepsilon)$ was maximal. $\square$

As an example, consider the flow of the left-invariant vector fields on $\mathrm{GL}(n, \mathbb{R})$ (or any matrix Lie group, for that matter): For any $v \in \mathfrak{gl}(n, \mathbb{R})$, the differential equation which $\gamma_e$ satisfies is simply

$$\gamma_e'(t) = \gamma_e(t)\,v.$$

This is a matrix differential equation and, in elementary ODE courses, we learn that the "fundamental solution" is

$$\gamma_e(t) = e^{tv} = I_n + \sum_{k=1}^{\infty} \frac{v^k}{k!} t^k$$

L.2.7

and that this series converges uniformly on compact sets in $\mathbb{R}$ to a smooth matrix-valued function of $t$.

Matrix Lie groups are by far the most commonly encountered and, for this reason, we often use the notation $\exp(tv)$ or even $e^{tv}$ for the integral curve $\gamma_e(t)$ associated to $X_v$ in a general Lie group $G$. (Actually, in order for this notation to be unambiguous, it has to be checked that if $tv = uw$ for $t, u \in \mathbb{R}$ and $v, w \in \mathfrak{g}$, then $\gamma_e(t) = \delta_e(u)$ where $\gamma_e$ is the integral curve of $X_v$ with initial condition $e$ and $\delta_e$ is the integral curve of $X_w$ initial condition $e$. However, this is an easy exercise in the use of the Chain Rule.)

It is worth remarking explicitly that for any $v \in \mathfrak{g}$ the formula for the flow of the left invariant vector field $X_v$ on $G$ is simply

$$\Phi(t, a) = a \exp(tv) = a\, e^{tv}.$$

(Warning: many beginners make the mistake of thinking that the formula for the flow of the left invariant vector field $X_v$ should be $\Phi(t, a) = \exp(tv)\, a$, instead. It is worth pausing for a moment to think why this is not so.)

It is now possible to describe all of the homomorphisms from the Lie group $(\mathbb{R}, +)$ into any given Lie group:

**Proposition 5:** *Every Lie group homomorphism* $\phi \colon \mathbb{R} \to G$ *is of the form* $\phi(t) = e^{tv}$ *where* $v = \phi'(0) \in \mathfrak{g}$.

PROOF: Let $v = \phi'(0) \in \mathfrak{g}$, and let $X_v$ be the associated left-invariant vector field on $G$. Since $\phi(0) = e$, by ODE uniqueness, it suffices to show that $\phi$ is an integral curve of $X_v$. However, $\phi(s + t) = \phi(s)\phi(t)$ implies $\phi'(s) = L'_{\phi(s)}\big(\phi'(0)\big) = X_v\big(\phi(s)\big)$, as desired. $\qquad\square$

**The Exponential Map.** We are now ready to introduce one of the principal tools in the study of Lie groups.

**Definition 6:** For any Lie group, the *exponential mapping* of $G$ is the mapping $\exp \colon \mathfrak{g} \to G$ defined by $\exp(v) = \gamma_e(1)$ where $\gamma_e$ is the integral curve of the vector field $X_v$ with initial condition $e$ .

It is an exercise for the reader to show that $\exp \colon \mathfrak{g} \to G$ is smooth and that

$$\exp'(0) \colon \mathfrak{g} \to T_e G = \mathfrak{g}$$

is just the identity mapping.

**Example:** As we have seen, for $\mathrm{GL}(n, \mathbb{R})$ (or $\mathrm{GL}(V)$ in general for that matter), the formula for the exponential mapping is just the usual power series:

$$e^x = I + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \cdots.$$

L.2.8

This formula works for all matrix Lie groups as well, and can simplify considerably in certain special cases. For example, for the group $N_3$ defined earlier (usually called the Heisenberg group), we have

$$
\mathfrak{n}_3 = \left\{ \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} \,\middle|\, x, y, z \in \mathbb{R} \right\},
$$

and $v^3 = 0$ for all $v \in \mathfrak{n}_3$. Thus

$$
\exp \left( \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} \right) = \begin{pmatrix} 1 & x & z + \frac{1}{2}xy \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.
$$

**The Lie Bracket.** Now, the mapping exp is not generally a homomorphism from $\mathfrak{g}$ (with its additive group structure) to $G$, although, in a certain sense, it comes as close as possible, since, by construction, it *is* a homomorphism when restricted to any one-dimensional linear subspace $\mathbb{R}v \subset \mathfrak{g}$. We now want to spend a few moments considering what the multiplication map on $G$ "looks like" when pulled back to $\mathfrak{g}$ via exp.

Since $\exp'(0): \mathfrak{g} \to T_eG = \mathfrak{g}$ is the identity mapping, it follows from the Implicit Function Theorem that there is a neighborhood $U$ of $0 \in \mathfrak{g}$ so that $\exp: U \to G$ is a diffeomorphism onto its image. Moreover, there must be a smaller open neighborhood $V \subset U$ of $0$ so that $\mu(\exp(V) \times \exp(V)) \subset \exp(U)$. It follows that there is a unique smooth mapping $\nu: V \times V \to U$ such that

$$
\mu(\exp(x), \exp(y)) = \exp(\nu(x, y)).
$$

Since exp is a homomorphism restricted to each line through $0$ in $\mathfrak{g}$, it follows that $\nu$ satisfies

$$
\nu(\alpha x, \beta x) = (\alpha + \beta)x
$$

for all $x \in V$ and $\alpha, \beta \in \mathbb{R}$ such that $\alpha x, \beta x \in V$.

Since $\nu(0, 0) = 0$, the Taylor expansion to second order of $\nu$ about $(0, 0)$ is of the form,

$$
\nu(x, y) = \nu_1(x, y) + \tfrac{1}{2}\nu_2(x, y) + R_3(x, y)
$$

where $\nu_i$ is a $\mathfrak{g}$-valued polynomial of degree $i$ on the vector space $\mathfrak{g} \oplus \mathfrak{g}$ and $R_3$ is a $\mathfrak{g}$-valued function on $V$ which vanishes to at least third order at $(0, 0)$.

Since $\nu(x, 0) = \nu(0, x) = x$, it easily follows that $\nu_1(x, y) = x + y$ and that $\nu_2(x, 0) = \nu_2(0, y) = 0$. Thus, the quadratic polynomial $\nu_2$ is linear in each $\mathfrak{g}$-variable separately.

Moreover, since $\nu(x, x) = 2x$ for all $x \in V$, substituting this into the above expansion and comparing terms of order 2 yields that $\nu_2(x, x) \equiv 0$. Of course, this implies that $\nu_2$ is actually skew-symmetric since

$$
0 = \nu_2(x + y, x + y) - \nu_2(x, x) - \nu_2(y, y) = \nu_2(x, y) + \nu_2(y, x).
$$

**Definition 7:** The skew-symmetric, bilinear multiplication $[,]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ defined by

$$[x, y] = \nu_2(x, y)$$

is called the *Lie bracket* in $\mathfrak{g}$. The pair $(\mathfrak{g}, [,])$ is called the *Lie algebra* of $G$.

With this notation, we have a formula

$$\exp(x)\exp(y) = \exp\big(x + y + \tfrac{1}{2}[x, y] + R_3(x, y)\big)$$

valid for all $x$ and $y$ in some fixed open neighborhood of $0$ in $\mathfrak{g}$.

One might think of the term involving $[,]$ as the first deviation of the Lie group multiplication from being just vector addition. In fact, it is clear from the above formula that, if the group $G$ is abelian, then $[x, y] = 0$ for all $x, y \in \mathfrak{g}$. For this reason, a Lie algebra in which all brackets vanish is called an *abelian* Lie algebra. (In fact, (see the Exercises) $\mathfrak{g}$ being abelian implies that $G^\circ$, the identity component of $G$, is abelian.)

**Example :** If $G = \mathrm{GL}(n, \mathbb{R})$, then it is easy to see that the induced bracket operation on $\mathfrak{gl}(n, \mathbb{R})$, the vector space of $n$-by-$n$ matrices, is just the matrix "commutator"

$$[x, y] = xy - yx.$$

In fact, the reader can verify this by examining the following second order expansion:

$$
\begin{aligned}
e^x e^y &= (I_n + x + \tfrac{1}{2}x^2 + \cdots)(I_n + y + \tfrac{1}{2}y^2 + \cdots)\\
&= (I_n + x + y + \tfrac{1}{2}(x^2 + 2xy + y^2) + \cdots)\\
&= (I_n + (x + y + \tfrac{1}{2}[x, y]) + \tfrac{1}{2}(x + y + \tfrac{1}{2}[x, y])^2 + \cdots)
\end{aligned}
$$

Moreover, this same formula is easily seen to hold for any $x$ and $y$ in $\mathfrak{gl}(V)$ where $V$ is any finite dimensional vector space.

**Theorem 1:** *If $\phi\colon H \to G$ is a Lie group homomorphism, then $\varphi = \phi'(e)\colon \mathfrak{h} \to \mathfrak{g}$ satisfies*

$$\exp_G(\varphi(x)) = \phi(\exp_H(x))$$

*for all $x \in \mathfrak{h}$. In other words, the diagram*

$$
\begin{array}{ccc}
\mathfrak{h} & \xrightarrow{\ \varphi\ } & \mathfrak{g}\\
{\scriptstyle \exp_H}\big\downarrow & & \big\downarrow{\scriptstyle \exp_G}\\
H & \xrightarrow{\ \phi\ } & G
\end{array}
$$

*commutes. Moreover, for all $x$ and $y$ in $\mathfrak{h}$,*

$$\varphi([x, y]_H) = [\varphi(x), \varphi(y)]_G.$$

PROOF: The first statement is an immediate consequence of Proposition 5 and the Chain Rule since, for every $x \in \mathfrak{h}$, the map $\gamma \colon \mathbb{R} \to G$ given by $\gamma(t) = \phi(e^{tx})$ is clearly a Lie group homomorphism with initial velocity $\gamma'(0) = \varphi(x)$ and hence must also satisfy $\gamma(t) = e^{t\varphi(x)}$.

To get the second statement, let $x$ and $y$ be elements of $\mathfrak{h}$ which are sufficiently close to zero. Then we have, using self-explanatory notation:

$$\phi(\exp_H(x)\exp_H(y)) = \phi(\exp_H(x))\phi(\exp_H(y)),$$

so

$$\phi(\exp_H(x + y + \tfrac{1}{2}[x,y]_H + R_3^H(x,y))) = \exp_G(\varphi(x))\exp_G(\varphi(y)),$$

and thus

$$\exp_G(\varphi(x + y + \tfrac{1}{2}[x,y]_H + R_3^H(x,y))) = \exp_G(\varphi(x) + \varphi(y) + \tfrac{1}{2}[\varphi(x),\varphi(y)]_G + R_3^G(\varphi(x),\varphi(y))),$$

finally giving

$$\varphi(x + y + \tfrac{1}{2}[x,y]_H + R_3^H(x,y)) = \varphi(x) + \varphi(y) + \tfrac{1}{2}[\varphi(x),\varphi(y)]_G + R_3^G(\varphi(x),\varphi(y)).$$

Now using the fact that $\varphi$ is linear and comparing second order terms gives the desired result. $\qquad\square$

On account of this theorem, it is usually not necessary to distinguish the map exp or the bracket $[\,,]$ according to the group in which it is being applied, so I will follow this practice also. Henceforth, these symbols will be used without group decorations whenever confusion seems unlikely.

Theorem 1 has many useful corollaries. Among them is

**Proposition 6:** *If $H$ is a connected Lie group and $\phi_1, \phi_2 \colon H \to G$ are two Lie group homomorphisms which satisfy $\phi_1'(e) = \phi_2'(e)$, then $\phi_1 = \phi_2$.*

PROOF: There is an open neighborhood $U$ of $e$ in $H$ so that $\exp_H$ is invertible on this neighborhood with inverse satisfying $\exp_H^{-1}(e) = 0$. Then for $a \in U$ we have, by Theorem 1,

$$\phi_i(a) = \exp_G(\varphi_i(\exp_H^{-1}(a))).$$

Since $\varphi_1 = \varphi_2$, we have $\phi_1 = \phi_2$ on $U$. By Proposition 2, every element of $H$ can be written as a finite product of elements of $U$, so we must have $\phi_1 = \phi_2$ everywhere. $\qquad\square$

We also have the following fundamental result:

**Proposition 7:** *If $\mathrm{Ad} \colon G \to \mathrm{GL}(\mathfrak{g})$ is the adjoint representation, then $\mathrm{ad} = \mathrm{Ad}'(e) \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is given by the formula $\mathrm{ad}(x)(y) = [x,y]$. In particular, we have the Jacobi identity*

$$\mathrm{ad}([x,y]) = [\mathrm{ad}(x), \mathrm{ad}(y)].$$

PROOF: This is simply a matter of unwinding the definitions. By definition, $\mathrm{Ad}(a) = C_a'(e)$ where $C_a \colon G \to G$ is defined by $C_a(b) = aba^{-1}$. In order to compute $C_a'(e)(y)$ for $y \in \mathfrak{g}$,

we may just compute $\gamma'(0)$ where $\gamma$ is the curve $\gamma(t) = a \exp(ty)a^{-1}$. Moreover, since $\exp'(0) \colon \mathfrak{g} \to \mathfrak{g}$ is the identity, we may as well compute $\beta'(0)$ where $\beta = \exp^{-1} \circ \gamma$. Now, assuming $a = \exp(x)$, we compute

$$
\begin{aligned}
\beta(t) &= \exp^{-1}\big(\exp(x)\exp(ty)\exp(-x)\big) \\
&= \exp^{-1}\big(\exp(x + ty + \tfrac{1}{2}[x, ty] + \cdots)\exp(-x)\big) \\
&= \exp^{-1}\big(\exp((x + ty + \tfrac{1}{2}[x, ty]) + (-x) + \tfrac{1}{2}[x + ty, -x] + \cdots\big) \\
&= ty + t[x, y] + E_3(x, ty)
\end{aligned}
$$

where the omitted terms and the function $E_3$ vanish to order at least 3 at $(x, y) = (0, 0)$. (Note that I used the identity $[y, x] = -[x, y]$.) It follows that

$$
\mathrm{Ad}(\exp(x))(y) = \beta'(0) = y + [x, y] + E_3'(x, 0)y
$$

where $E_3'(x, 0)$ denotes the derivative of $E_3$ with respect to $y$ evaluated at $(x, 0)$ and is hence a function of $x$ which vanishes to order at least 2 at $x = 0$. On the other hand, since, by the first part of Theorem 1, we have

$$
\mathrm{Ad}(\exp(x)) = \exp(\mathrm{ad}(x)) = I + \mathrm{ad}(x) + \tfrac{1}{2}(\mathrm{ad}(x))^2 + \cdots .
$$

Comparing the $x$-linear terms in the last two equations clearly gives the desired result. The validity of the Jacobi identity now follows by applying the second part of Theorem 1 to Proposition 3. $\qquad\square$

The Jacobi identity is often presented differently. The reader can verify that the equation $\mathrm{ad}\big([x, y]\big) = \big[\mathrm{ad}(x), \mathrm{ad}(y)\big]$ where $\mathrm{ad}(x)(y) = [x, y]$ is equivalent to the condition that

$$
\big[[x, y], z\big] + \big[[y, z], x\big] + \big[[z, x], y\big] = 0 \qquad \text{for all} \quad z \in \mathfrak{g}.
$$

This is a form in which the Jacobi identity is often stated. Unfortunately, although this is a very symmetric form of the identity, it somewhat obscures its importance and meaning.

The Jacobi identity is so important that the class of algebras in which it holds is given a name:

**Definition 8:** A *Lie algebra* is a pair $(\mathfrak{g}, [\,,\,])$ where $\mathfrak{g}$ is a vector space and $[\,,\,] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is a skew-symmetric bilinear multiplication which satisfies the Jacobi identity, i.e., $\mathrm{ad}([x, y]) = [\mathrm{ad}(x), \mathrm{ad}(y)]$, where $\mathrm{ad} \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is defined by $\mathrm{ad}(x)(y) = [x, y]$ A *Lie subalgebra* of $\mathfrak{g}$ is a linear subspace $\mathfrak{h} \subset \mathfrak{g}$ which is closed under bracket. A *homomorphism* of Lie algebras is a linear mapping of vector spaces $\varphi \colon \mathfrak{h} \to \mathfrak{g}$ which satisfies

$$
\varphi\big([x, y]\big) = \big[\varphi(x), \varphi(y)\big].
$$

At the moment, our only examples of Lie algebras are the ones provided by Proposition 6, namely, the Lie algebras of Lie groups. This is not accidental, for, as we shall see, every finite dimensional Lie algebra is the Lie algebra of some Lie group.

**Lie Brackets of Vector Fields.** There is another notion of Lie bracket, namely the Lie bracket of smooth vector fields on a smooth manifold. This bracket is also skew-symmetric and satisfies the Jacobi identity, so it is reasonable to ask how it might be related to the notion of Lie bracket that we have defined. Since Lie bracket of vector fields commutes with diffeomorphisms, it easily follows that the Lie bracket of two left-invariant vector fields on a Lie group $G$ is also a left-invariant vector field on $G$. The following result is, perhaps then, to be expected.

**Proposition 8:** *For any $x, y \in \mathfrak{g}$, we have $[X_x, X_y] = X_{[x,y]}$.*

PROOF: This is a direct calculation. For simplicity, we will use the following characterization of the Lie bracket for vector fields: If $\Phi_x$ and $\Phi_y$ are the flows associated to the vector fields $X_x$ and $X_y$, then for any function $f$ on $G$ we have the formula:

$$([X_x, X_y]f)(a) = \lim_{t \to 0^+} \frac{f(\Phi_y(-\sqrt{t}, \Phi_x(-\sqrt{t}, \Phi_y(\sqrt{t}, \Phi_x(\sqrt{t}, a))))) - f(a)}{t}.$$

Now, as we have seen, the formulas for the flows of $X_x$ and $X_y$ are given by $\Phi_x(t, a) = a \exp(tx)$ and $\Phi_y(t, a) = a \exp(ty)$. This implies that the general formula above simplifies to

$$([X_x, X_y]f)(a) = \lim_{t \to 0^+} \frac{f\big(a \exp(\sqrt{t}x) \exp(\sqrt{t}y) \exp(-\sqrt{t}x) \exp(-\sqrt{t}y)\big) - f(a)}{t}.$$

Now
$$\exp(\pm\sqrt{t}x) \exp(\pm\sqrt{t}y) = \exp(\pm\sqrt{t}(x + y) + \tfrac{t}{2}[x, y] + \cdots)$$

so $\exp(\sqrt{t}x) \exp(\sqrt{t}y) \exp(-\sqrt{t}x) \exp(-\sqrt{t}y)$ simplifies to $\exp(t[x, y] + \cdots)$ where the omitted terms vanish to higher $t$-order than $t$ itself. Thus, we have

$$([X_x, X_y]f)(a) = \lim_{t \to 0^+} \frac{f\big(a \exp(t[x, y] + \cdots)\big) - f(a)}{t}.$$

Since $[X_x, X_y]$ must be a left-invariant vector field and since

$$(X_{[x,y]}f)(a) = \lim_{t \to 0^+} \frac{f\big(a \exp(t[x, y])\big) - f(a)}{t},$$

the desired result follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We can now prove the following fundamental result.

**Theorem 2:** *For each Lie subgroup $H$ of a Lie group $G$, the subspace $\mathfrak{h} = T_e H$ is a Lie subalgebra of $\mathfrak{g}$. Moreover, every Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is $T_e H$ for a unique connected Lie subgroup $H$ of $G$.*

PROOF: Suppose that $H \subset G$ is a Lie subgroup. Then the inclusion map is a Lie group homomorphism and Theorem 1 thus implies that the inclusion map $\mathfrak{h} \hookrightarrow \mathfrak{g}$ is a Lie algebra homomorphism. In particular, $\mathfrak{h}$, when considered as a subspace of $\mathfrak{g}$, is closed under the Lie bracket in $G$ and hence is a subalgebra.

Suppose now that $\mathfrak{h} \subset \mathfrak{g}$ is a subalgebra.

First, let us show that there is at most one connected Lie subgroup of $G$ with Lie algebra $\mathfrak{h}$. Suppose that there were two, say $H_1$ and $H_2$. Then by Theorem 1, $\exp_G(\mathfrak{h})$ is a subset of both $H_1$ and $H_2$ and contains an open neighborhood of the identity element in each of them. However, since, by Proposition 2, each of $H_1$ and $H_2$ are generated by finite products of the elements in any open neighborhood of the identity, it follows that $H_1 \subset H_2$ and $H_2 \subset H_1$, so $H_1 = H_2$, as desired.

Second, to prove the existence of a subgroup $H$ with $T_e H = \mathfrak{h}$, we call on the Global Frobenius Theorem. Let $r = \dim(\mathfrak{h})$ and let $E \subset TG$ be the rank $r$ sub-bundle spanned by the vector fields $X_x$ where $x \in \mathfrak{h}$. Note that $E_a = L'_a(E_e) = L'_a(\mathfrak{h})$ for all $a \in G$, so $E$ is left-invariant. Since $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, Proposition 8 implies that $E$ is an integrable distribution on $G$. By the Global Frobenius Theorem, there is an $r$-dimensional leaf of $E$ through $e$. Call this submanifold $H$.

It remains is to show that $H$ is closed under multiplication and inverse. Inverse is easy: Let $a \in H$ be fixed. Then, since $H$ is path-connected, there exists a smooth curve $\alpha \colon [0,1] \to H$ so that $\alpha(0) = e$ and $\alpha(1) = a$. Now consider the curve $\bar{\alpha}$ defined on $[0,1]$ by $\bar{\alpha}(t) = a^{-1}\alpha(1-t)$. Because $E$ is left-invariant, $\bar{\alpha}$ is an integral curve of E and it joins $e$ to $a^{-1}$. Thus $a^{-1}$ must also lie in $H$. Multiplication is only slightly more difficult: Now suppose in addition that $b \in H$ and let $\beta \colon [0,1] \to H$ be a smooth curve so that $\beta(0) = e$ and $\beta(1) = b$. Then the piecewise smooth curve $\gamma \colon [0,2] \to G$ given by

$$\gamma(t) = \begin{cases} \alpha(t) & \text{if } 0 \le t \le 1; \\ a\beta(t-1) & \text{if } 1 \le t \le 2, \end{cases}$$

is an integral curve of $E$ joining $e$ to $ab$. Hence $ab$ belongs to $H$, as we wished to show. $\square$

**Theorem 3:** *If $H$ is a connected and simply connected Lie group, then, for any Lie group $G$, each Lie algebra homorphism $\varphi \colon \mathfrak{h} \to \mathfrak{g}$ is of the form $\varphi = \phi'(e)$ for some unique Lie group homorphism $\phi \colon H \to G$.*

PROOF: In light of Theorem 1 and Proposition 6, all that remains to be proved is that for each Lie algebra homorphism $\varphi \colon \mathfrak{h} \to \mathfrak{g}$ there exists a Lie group homomorphism $\phi$ satisfying $\phi'(e) = \varphi$.

We do this as follows: Suppose that $\varphi \colon \mathfrak{h} \to \mathfrak{g}$ is a Lie algebra homomorphism. Consider the product Lie group $H \times G$. Its Lie algebra is $\mathfrak{h} \oplus \mathfrak{g}$ with Lie bracket given by $[(h_1, g_1), (h_2, g_2)] = ([h_1, h_2], [g_1, g_2])$, as is easily verified. Now consider the subspace $\widehat{\mathfrak{h}} \subset \mathfrak{h} \oplus \mathfrak{g}$ spanned by elements of the form $(x, \varphi(x))$ where $x \in \mathfrak{h}$. Since $\varphi$ is a Lie algebra homomorphism, $\widehat{\mathfrak{h}}$ is a Lie subalgebra of $\mathfrak{h} \oplus \mathfrak{g}$ (and happens to be isomorphic to $\mathfrak{h}$). In

particular, by Theorem 2, it follows that there is a connected Lie subgroup $\widehat{H} \subset H \times G$, whose Lie algebra is $\widehat{\mathfrak{h}}$. We are now going to show that $\widehat{H}$ is the graph of the desired Lie group homomorphism $\phi \colon H \to G$.

Note that since $\widehat{H}$ is a Lie subgroup of $H \times G$, the projections $\pi_1 \colon \widehat{H} \to H$ and $\pi_2 \colon \widehat{H} \to G$ are Lie group homomorphisms. The associated Lie algebra homomorphisms $\varpi_1 \colon \widehat{\mathfrak{h}} \to \mathfrak{h}$ and $\varpi_2 \colon \widehat{\mathfrak{h}} \to \mathfrak{g}$ are clearly given by $\varpi_1(x, \varphi(x)) = x$ and $\varpi_2(x, \varphi(x)) = \varphi(x)$.

Now, I claim that $\pi_1$ is actually a surjective covering map: It is surjective since $\varpi_1 \colon \widehat{\mathfrak{h}} \to \mathfrak{h}$ is an isomorphism so $\pi_1(\widehat{H})$ contains a neighborhood of the identity in $H$ and hence, by Proposition 2 and the connectedness of $H$, must contain all of $H$. It remains to show that, under $\pi_1$, points of $H$ have evenly covered neighborhoods.

Let $\widehat{Z} = \ker(\pi_1)$. Then $\widehat{Z}$ is a closed discrete subgroup of $\widehat{H}$. Let $\widehat{U} \subset \widehat{H}$ be a neighborhood of the identity to which $\pi_1$ restricts to be a smooth diffeomorphism onto a neighborhood $U$ of $e$ in $H$. Then the reader can easily verify that for each $a \in \widehat{H}$ the map $\sigma_a \colon \widehat{Z} \times \widehat{U} \to \widehat{H}$ given by $\sigma_a(z, u) = azu$ is a diffeomorphism onto $(\pi_1)^{-1}(L_{\pi_1(a)}(U))$ which commutes with the appropriate projections and hence establishes the even covering property.

Finally, since $\widehat{H}$ is connected and, by hypothesis, $H$ is simply connected, it follows that $\pi_1$ must actually be a one-to-one and onto diffeomorphism. The map $\phi = \pi_2 \circ \pi_1^{-1}$ is then the desired homomorphism. $\qquad\square$

As our last general Theorem, we state, without proof, the following existence result.

**Theorem 4:** *For each finite dimensional Lie algebra $\mathfrak{g}$, there exists a Lie group $G$ whose Lie algebra is isomorphic to $\mathfrak{g}$.*

Unfortunately, this theorem is surprisingly difficult to prove. It would suffice, by Theorem 2, to show that every Lie algebra $\mathfrak{g}$ is isomorphic to a subalgebra of the Lie algebra of a Lie group. In fact, an even stronger statement is true. A theorem of Ado asserts that every finite dimensional Lie algebra is isomorphic to a subalgebra of $\mathfrak{gl}(n, \mathbb{R})$ for some $n$. Thus, to prove Theorem 4, it would be enough to prove Ado's theorem. Unfortunately, this theorem also turns out to be rather delicate (see [Po] for a proof). However, there are many interesting examples of $\mathfrak{g}$ for which a proof can be given by elementary means (see the Exercises).

On the other hand, this abstract existence theorem is not used very often anyway. It is rare that a (finite dimensional) Lie algebra arises in practice which is not readily representable as the Lie algebra of some Lie group.

The reader may be wondering about uniqueness: How many Lie groups are there whose Lie algebras are isomorphic to a given $\mathfrak{g}$? Since the Lie algebra of a Lie group $G$ only depends on the identity component $G$, it is reasonable to restrict to the case of connected Lie groups. Now, as you are asked to show in the Exercises, the universal cover $\widetilde{G}$ of a connected Lie group $G$ can be given a unique Lie group structure for which the covering map $\widetilde{G} \to G$ is a homomorphism. Thus, there always exists a connected and simply connected Lie group, say $G(\mathfrak{g})$, whose Lie algebra is isomorphic to $\mathfrak{g}$. A simple

application of Theorem 3 shows that if $G'$ is any other Lie group with Lie algebra $\mathfrak{g}$, then there is a homomorphism $\phi\colon G(\mathfrak{g}) \to G$ which induces an isomorphism on the Lie algebras. It follows easily that, up to isomorphism, there is only one simply connected and connected Lie group with Lie algebra $\mathfrak{g}$. Moreover, every other connected Lie group with Lie algebra $G$ is isomorphic to a quotient of $G(\mathfrak{g})$ by a discrete subgroup of $G$ which lies in the center of $G(\mathfrak{g})$ (see the Exercises).

**The Structure Constants.** Our work so far has shown that the problem of classifying the connected Lie groups up to isomorphism is very nearly the same thing as classifying the (finite dimensional) Lie algebras. (See the Exercises for a clarification of this point.) This is a remarkable state of affairs, since, *a priori*, Lie groups involve the topology of smooth manifolds and it is rather surprising that their classification can be reduced to what is essentially an algebra problem. It is worth taking a closer look at this algebra problem itself.

Let $\mathfrak{g}$ be a Lie algebra of dimension $n$, and let $x_1, x_2, \ldots, x_n$ be a basis for $\mathfrak{g}$. Then there exist constants $c_{ij}^k$ so that (using the summation convention)

$$[x_i, x_j] = c_{ij}^k \, x_k \, .$$

(These quantities $c$ are called the structure constants of $\mathfrak{g}$ relative to the given basis.) The skew-symmetry of the Lie bracket is is equivalent to the skew-symmetry of $c$ in its lower indices:

$$c_{ij}^k + c_{ji}^k = 0.$$

The Jacobi identity is equivalent to the quadratic equations:

$$c_{ij}^\ell \, c_{k\ell}^m + c_{jk}^\ell \, c_{i\ell}^m + c_{ki}^\ell \, c_{j\ell}^m = 0.$$

Conversely, any set of $n^3$ constants satisfying these relations defines an $n$-dimensional Lie algebra by the above bracket formula.

**Left-Invariant Forms and the Structure Equations.** Dual to the left-invariant vector fields on a Lie group $G$, there are the left-invariant 1-forms, which are indispensable as calculational tools.

**Definition 9:** For any Lie group $G$, the $\mathfrak{g}$-valued 1-form on $G$ defined by

$$\omega_G(v) = L'_{a^{-1}}(v) \qquad \text{for } v \in T_a G$$

is called the *canonical left-invariant 1-form* on $G$.

It is easy to see that $\omega_G$ is smooth. Moreover, $\omega_G$ is the unique left-invariant $\mathfrak{g}$-valued 1-form on $G$ which satisfies $\omega_G(v) = v$ for all $v \in \mathfrak{g} = T_e G$.

By a calculation which is left as an exercise for the reader,

$$\phi^*(\omega_G) = \varphi(\omega_H)$$

for any Lie group homomorphism $\phi\colon H \to G$ with $\varphi = \phi'(e)$. In particular, when $H$ is a subgroup of $G$, the pull back of $\omega_G$ to $H$ via the inclusion mapping is just $\omega_H$. For this reason, it is common to simply write $\omega$ for $\omega_G$ when there is no danger of confusion.

**Example:** If $G \subset \mathrm{GL}(n, \mathbb{R})$ is a matrix Lie group, then we may regard the inclusion $g\colon G \to \mathrm{GL}(n, \mathbb{R})$ as a matrix-valued function on $G$ and compute that $\omega$ is given by the simple formula

$$\omega = g^{-1}\, dg.$$

From this formula, the left-invariance of $\omega$ is obvious.

In the matrix Lie group case, it is also easy to compute the exterior derivative of $\omega$: Since $g\, g^{-1} = I_n$, we get

$$dg\, g^{-1} + g\, d(g^{-1}) = 0,$$

so

$$d(g^{-1}) = -g^{-1}\, dg\, g^{-1}.$$

This implies the formula

$$d\omega = -\omega \wedge \omega.$$

(Warning: Matrix multiplication is implicit in this formula!)

For a general Lie group, the formula for $d\omega$ is only slightly more complicated. To state the result, let me first define some notation. I will use $[\omega, \omega]$ to denote the $\mathfrak{g}$-valued 2-form on $G$ whose value on a pair of vectors $v, w \in T_a G$ is

$$[\omega, \omega](v, w) = [\omega(v), \omega(w)] - [\omega(w), \omega(v)] = 2[\omega(v), \omega(w)].$$

**Proposition 9:** *For any Lie group $G$, $\ d\omega = -\tfrac{1}{2}[\omega, \omega]$.*

PROOF: First, let $X_v$ and $X_w$ be the left-invariant vector fields on $G$ whose values at $e$ are $v$ and $w$ respectively. Then, by the usual formula for the exterior derivative

$$d\omega(X_v, X_w) = X_v\big(\omega(X_w)\big) - X_w\big(\omega(X_v)\big) - \omega\big([X_v, X_w]\big).$$

However, the $\mathfrak{g}$-valued functions $\omega(X_v)$ and $\omega(X_w)$ are clearly left-invariant and hence are constants and equal to $v$ and $w$ respectively. Moreover, by Proposition 8, $[X_v, X_w] = X_{[v,w]}$, so the formula simplifies to

$$d\omega(X_v, X_w) = -\omega\big(X_{[v,w]}\big).$$

The right hand side is, again, a left-invariant function, so it must equal its value at the identity, which is clearly $-[v, w]$, which equals $-[\omega(X_v), \omega(X_w)]$ Thus,

$$d\omega(X_v, X_w) = -\tfrac{1}{2}[\omega(X_v), \omega(X_w)]$$

for any pair of left-invariant vector fields on $G$. Since any pair of vectors in $T_a G$ can be written as $X_v(a)$ and $X_w(a)$ for some $v, w \in \mathfrak{g}$, the result follows. $\qquad\square$

L.2.17

The formula proved in Proposition 9 is often called the *structure equation* of Maurer and Cartan. It is also usually expressed slightly differently. If $x_1, x_2, \ldots, x_n$ is a basis for $\mathfrak{g}$ with structure constants $c^i_{jk}$, then $\omega$ can be written in the form

$$\omega = x_1 \omega^1 + \cdots + x_n \omega^n$$

where the $\omega^i$ are $\mathbb{R}$-valued left-invariant 1-forms and Proposition 9 can then be expanded to give

$$d\omega^i = -\tfrac{1}{2} c^i_{jk}\, \omega^j \wedge \omega^k,$$

which is the most common form in which the structure equations are given. Note that the identity $d(d(\omega^i)) = 0$ is equivalent to the Jacobi identity.

**An Extended Example: 2- and 3-dimensional Lie Algebras.** It is clear that up to isomorphism, there is only one (real) Lie algebra of dimension 1, namely $\mathfrak{g} = \mathbb{R}$ with the zero bracket. This is the Lie algebra of the connected Lie groups $\mathbb{R}$ and $S^1$. (You are asked to prove in an exercise that these are, in fact, the only connected one-dimensional Lie groups.)

The first interesting case, therefore, is dimension 2. If $\mathfrak{g}$ is a 2-dimensional Lie algebra with basis $x_1, x_2$, then the entire Lie algebra structure is determined by the bracket $[x_1, x_2] = a^1 x_1 + a^2 x_2$. If $a^1 = a^2 = 0$, then all brackets are zero, and the algebra is abelian. If one of $a^1$ or $a^2$ is non-zero, then, by switching $x_1$ and $x_2$ if necessary, we may assume that $a^1 \neq 0$. Then, considering the new basis $y_1 = a^1 x_1 + a^2 x_2$ and $y_2 = (1/a^1)x_2$, we get $[y_1, y_2] = y_1$. Since the Jacobi identity is easily verified for this Lie bracket, this does define a Lie algebra. Thus, up to isomorphism, there are only two distinct 2-dimensional Lie algebras.

The abelian example is, of course, the Lie algebra of the vector space $\mathbb{R}^2$ (as well as the Lie algebra of $S^1 \times \mathbb{R}$, and the Lie algebra of $S^1 \times S^1$).

An example of a Lie group of dimension 2 with a non-abelian Lie algebra is the matrix Lie group

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \,\middle|\, a \in \mathbb{R}^+,\ b \in \mathbb{R} \right\}.$$

In fact, it is not hard to show that, up to isomorphism, this is the only connected non-abelian Lie group (see the Exercises).

Now, let us pass on to the classification of the three dimensional Lie algebras. Here, the story becomes much more interesting. Let $\mathfrak{g}$ be a 3-dimensional Lie algebra, and let $x_1, x_2, x_3$ be a basis of $\mathfrak{g}$. Then, we may write the bracket relations in matrix form as

$$\begin{pmatrix} [x_2, x_3] & [x_3, x_1] & [x_1, x_2] \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} C$$

where $C$ is the 3-by-3 matrix of structure constants. How is this matrix affected by a change of basis? Well, let

$$\begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} A$$

where $A \in \mathrm{GL}(3, \mathbb{R})$. Then it is easy to compute that

$$( \ [y_2, y_3] \ \ [y_3, y_1] \ \ [y_1, y_2] \ ) = ( \ [x_2, x_3] \ \ [x_3, x_1] \ \ [x_1, x_2] \ ) \, \mathrm{Adj}(A)$$

where $\mathrm{Adj}(A)$ is the classical adjoint matrix of $A$, i.e., the matrix of 2-by-2 minors. Thus,

$$A^{-1} = (\det(A))^{-1} \, {}^t\mathrm{Adj}(A).$$

(Do not confuse this with the adjoint mapping defined earlier!) It then follows that

$$( \ [y_2, y_3] \ \ [y_3, y_1] \ \ [y_1, y_2] \ ) = ( \ y_1 \ \ y_2 \ \ y_3 \ ) \, C',$$

where

$$C' = A^{-1} \, C \, \mathrm{Adj}(A) = \det(A) \, A^{-1} \, C \, {}^t A^{-1}.$$

It follows without too much difficulty that, if we write $C = S + \hat{a}$, where $S$ is a symmetric 3-by-3 matrix and

$$\hat{a} = \begin{pmatrix} 0 & -a^3 & a^2 \\ a^3 & 0 & -a^1 \\ -a^2 & a^1 & 0 \end{pmatrix} \qquad \text{where} \qquad a = \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix},$$

then $C' = S' + \widehat{a'}$, where

$$S' = \det(A) \, A^{-1} \, S \, {}^t A^{-1} \qquad \text{and} \qquad a' = {}^t A a.$$

Now, I claim that the condition that the Jacobi identity hold for the bracket defined by the matrix $C$ is equivalent to the condition $Sa = 0$. To see this, note first that

$$\begin{aligned}
[[x_2, &x_3], x_1] + [[x_3, x_1], x_2] + [[x_1, x_2], x_3] \\
&= [C_1^1 x_1 + C_1^2 x_2 + C_1^3 x_3, x_1] + [C_2^1 x_1 + C_2^2 x_2 + C_2^3 x_3, x_2] + [C_3^1 x_1 + C_3^2 x_2 + C_3^3 x_3, x_3] \\
&= (C_3^2 - C_2^3)[x_2, x_3] + (C_1^3 - C_3^1)[x_3, x_1] + (C_2^1 - C_1^2)[x_1, x_2] \\
&= 2a^1[x_2, x_3] + 2a^2[x_3, x_1] + 2a^3[x_1, x_2] \\
&= 2 ( \ [x_2, x_3] \ \ [x_3, x_1] \ \ [x_1, x_2] \ ) \, a \\
&= 2 ( \ x_1 \ \ x_2 \ \ x_3 \ ) \, Ca,
\end{aligned}$$

and $Ca = (S + \hat{a})a = Sa$ since $\hat{a} \, a = 0$. Thus, the Jacobi identity applied to the basis $x_1, x_2, x_3$ implies that $Sa = 0$. However, if $y_1, y_2, y_3$ is any other triple of elements of $\mathfrak{g}$, then for some 3-by-3 matrix $B$, we have

$$( \ y_1 \ \ y_2 \ \ y_3 \ ) = ( \ x_1 \ \ x_2 \ \ x_3 \ ) \, B,$$

and I leave it to the reader to check that

$$[[y_2, y_3], y_1] + [[y_3, y_1], y_2] + [[y_1, y_2], y_3] = \det(B) \, ([[x_2, x_3], x_1] + [[x_3, x_1], x_2] + [[x_1, x_2], x_3])$$

L.2.19

in this case. Thus, $Sa = 0$ implies the full Jacobi identity.

There are now two essentially different cases to treat. In the first case, if $a = 0$, then the Jacobi identity is automatically satisfied, and $S$ can be any symmetric matrix. However, two such choices $S$ and $S'$ will clearly give rise to isomorphic Lie algebras if and only if there is an $A \in \mathrm{GL}(3, \mathbb{R})$ for which $S' = \det(A) A^{-1} S {}^t A^{-1}$. I leave as an exercise for the reader to show that every choice of $S$ yields an algebra (with $a = 0$) which is equivalent to exactly one of the algebras made by one of the following six choices:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

On the other hand, if $a \neq 0$, then by a suitable change of basis $A$, we see that we can assume that $a^1 = a^2 = 0$ and that $a^3 = 1$. Any change of basis $A$ which preserves this normalization is seen to be of the form

$$A = \begin{pmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since $Sa = 0$ and since $S$ is symmetric, it follows that $S$ must be of the form

$$S = \begin{pmatrix} s_{11} & s_{12} & 0 \\ s_{12} & s_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Moreover, a simple calculation shows that the result of applying a change of basis of the above form is to change the matrix $S$ into the matrix

$$S' = \begin{pmatrix} s'_{11} & s'_{12} & 0 \\ s'_{12} & s'_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where

$$\begin{pmatrix} s'_{11} & s'_{12} \\ s'_{12} & s'_{22} \end{pmatrix} = \frac{1}{A_1^1 A_2^2 - A_2^1 A_1^2} \begin{pmatrix} A_2^2 & -A_2^1 \\ -A_1^2 & A_1^1 \end{pmatrix} \begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} \begin{pmatrix} A_2^2 & -A_1^2 \\ -A_2^1 & A_1^1 \end{pmatrix}.$$

It follows that $s'_{11} s'_{22} - (s'_{12})^2 = s_{11} s_{22} - (s_{12})^2$, so there is an "invariant" to be dealt with. We leave it to the reader to show that the upper left-hand 2-by-2 block of $S$ can be brought by a change of basis of the above form into exactly one of the four forms

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix} \quad \begin{pmatrix} \sigma & 0 \\ 0 & -\sigma \end{pmatrix}$$

where $\sigma > 0$ is a real positive number.

L.2.20

To summarize, every 3-dimensional Lie algebra is isomorphic to exactly one of the following Lie algebras: Either

$$\mathfrak{so}(3): \begin{array}{l} [x_2, x_3] = x_1 \\ [x_3, x_1] = x_2 \\ [x_1, x_2] = x_3 \end{array} \qquad \text{or} \qquad \mathfrak{sl}(2, \mathbb{R}): \begin{array}{l} [x_2, x_3] = x_2 \\ [x_3, x_1] = x_1 \\ [x_1, x_2] = x_3 \end{array}$$

or an algebra of the form

$$[x_2, x_3] = b_{11}x_1 + b_{12}x_2$$
$$[x_3, x_1] = b_{21}x_1 + b_{22}x_2$$
$$[x_1, x_2] = 0$$

where the 2-by-2 matrix $B$ is one of the following

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} \sigma & 1 \\ -1 & \sigma \end{pmatrix} \qquad \begin{pmatrix} \sigma & 1 \\ -1 & -\sigma \end{pmatrix}$$

and, in the latter two cases, $\sigma$ is a positive real number. Each of these eight latter types can be represented as a subalgebra of $\mathfrak{gl}(3, \mathbb{R})$ in the form

$$\mathfrak{g} = \left\{ \begin{pmatrix} (1 + b_{21})z & -b_{11}z & x \\ b_{22}z & (1 - b_{12})z & y \\ 0 & 0 & z \end{pmatrix} \middle| \ x, y, z \in \mathbb{R} \right\}$$

I leave as an exercise for the reader to show that the corresponding subgroup of $\mathrm{GL}(3, \mathbb{R})$ is a closed, embedded, simply connected matrix Lie group whose underlying manifold is diffeomorphic to $\mathbb{R}^3$.

Actually, it is clear that, because of the skew-symmetry of the bracket, only $n\binom{n}{2}$ of these constants are independent. In fact, using the dual basis $x^1, \ldots, x^n$ of $\mathfrak{g}^*$, we can write the expression for the Lie bracket as an element $\beta \in \mathfrak{g} \otimes \Lambda^2(\mathfrak{g}^*)$, in the form

$$\beta = \tfrac{1}{2} c^i_{jk}\, x_i \otimes x^j \wedge x^k.$$

The Jacobi identity is then equivalent to the condition $J(\beta) = 0$, where

$$J: \mathfrak{g} \otimes \Lambda^2(\mathfrak{g}^*) \to \mathfrak{g} \otimes \Lambda^3(\mathfrak{g}^*)$$

is the quadratic polynomial map given in coordinates by

$$J(\beta) = \tfrac{1}{6} \left( c^\ell_{ij}\, c^m_{k\ell} + c^\ell_{jk}\, c^m_{i\ell} + c^\ell_{ki}\, c^m_{j\ell} \right) x_m \otimes x^i \wedge x^j \wedge x^k.$$

L.2.21

# Exercise Set 2:
# Lie Groups

**1**. Show that for any real vector space of dimension $n$, the Lie group $\mathrm{GL}(V)$ is isomorphic to $\mathrm{GL}(n, \mathbb{R})$. (Hint: Choose a basis $\mathbf{b}$ of $V$, use $\mathbf{b}$ to construct a mapping $\phi_{\mathbf{b}}\colon \mathrm{GL}(V) \to \mathrm{GL}(n, \mathbb{R})$, and then show that $\phi_{\mathbf{b}}$ is a smooth isomorphism.)

**2**. Let $G$ be a Lie group and let $H$ be an abstract subgroup. Show that if there is an open neighborhood $U$ of $e$ in $G$ so that $H \cap U$ is a smooth embedded submanifold of $G$, then $H$ is a Lie subgroup of $G$.

**3**. Show that $\mathrm{SL}(n, \mathbb{R})$ is an embedded Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$. (Hint: $\mathrm{SL}(n, \mathbb{R}) = \det^{-1}(1)$.)

**4**. Show that $\mathrm{O}(n)$ is an compact Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$. (Hint: $\mathrm{O}(n) = F^{-1}(I_n)$, where $F$ is the map from $\mathrm{GL}(n, \mathbb{R})$ to the vector space of $n$-by-$n$ symmetric matrices given by $F(A) = {}^{t}A\, A$. Taking note of Exercise 2, show that the Implicit Function Theorem applies. To show compactness, apply the Heine-Borel theorem.) Show also that $\mathrm{SO}(n)$ is an open-and-closed, index 2 subgroup of $\mathrm{O}(n)$.

**5**. Carry out the analysis in Exercise 3 for the complex matrix Lie group $\mathrm{SL}(n, \mathbb{C})$ and the analysis in Exercise 4 for the complex matrix Lie groups $\mathrm{U}(n)$ and $\mathrm{SU}(n)$. What are the (real) dimensions of all of these groups?

**6**. Show that the map $\mu\colon \mathrm{O}(n) \times A_n \times N_n \to \mathrm{GL}(n, \mathbb{R})$ defined by matrix multiplication is a diffeomorphism although it is *not* a group homomorphism. (Hint: The map is clearly smooth, you must only compute an inverse. To get the first factor $\nu_1\colon \mathrm{GL}(n, \mathbb{R}) \to \mathrm{O}(n)$ of the inverse map, think of an element $\mathbf{b} \in \mathrm{GL}(n, \mathbb{R})$ as a row of column vectors in $\mathbb{R}^n$ and let $\nu_1(\mathbf{b})$ be the row of column vectors which results from $\mathbf{b}$ by apply the Gram-Schmidt orthogonalization process. Why does this work and why is the resulting map $\nu_1$ smooth?) Show, similarly that the map

$$\mu\colon \mathrm{SO}(n) \times \bigl(A_n \cap \mathrm{SL}(n, \mathbb{R})\bigr) \times N_n \to \mathrm{SL}(n, \mathbb{R})$$

is a diffeomorphism. Are there similar factorizations for the groups $\mathrm{GL}(n, \mathbb{C})$ and $\mathrm{SL}(n, \mathbb{C})$? (Hint: Consider unitary bases rather than orthogonal ones.)

**7**. Show that

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \,\middle|\, a\overline{a} + b\overline{b} = 1 \right\}.$$

Conclude that $\mathrm{SU}(2)$ is diffeomorphic to the 3-sphere and, using the previous exercise, that, in particular, $\mathrm{SL}(2, \mathbb{C})$ is simply connected, while $\pi_1\bigl(\mathrm{SL}(2, \mathbb{R})\bigr) \simeq \mathbb{Z}$.

**8**. Show that, for any Lie group $G$, the mappings $L_a$ satisfy

$$L_a'(b) = L_{ab}'(e) \circ \left(L_b'(e)\right)^{-1}.$$

where $L_a'(b)\colon T_b G \to T_{ab} G$. (This shows that the effect of left translation is completely determined by what it does at $e$.) State and prove a similar formula for the mappings $R_a$.

**9**. Let $(G, \mu)$ be a Lie group. Using the canonical identification $T_{(a,b)}(G \times G) = T_a G \oplus T_b G$, prove the formula

$$\mu'(a,b)(v,w) = R_b'(a)(v) + L_a'(b)(w)$$

for all $v \in T_a G$ and $w \in T_b G$.

**10**. Complete the proof of Proposition 3 by explicitly exhibiting the map $c$ as a composition of known smooth maps. (Hint: if $f\colon X \to Y$ is smooth, then $f'\colon TX \to TY$ is also smooth.)

**11**. Show that, for any $v \in \mathfrak{g}$, the left-invariant vector field $X_v$ is indeed smooth. Also prove the first statement in Proposition 4. (Hint: Use $\Psi$ to write the mapping $X_v\colon G \to TG$ as a composition of smooth maps. Show that the assignment $v \mapsto X_v$ is linear. Finally, show that if a left-invariant vector field on $G$ vanishes anywhere, then it vanishes identically.)

**12**. Show that $\exp\colon \mathfrak{g} \to G$ is indeed smooth and that $\exp'\colon \mathfrak{g} \to \mathfrak{g}$ is the identity mapping. (Hint: Write down a smooth vector field $Y$ on $\mathfrak{g} \times G$ such that the integral curves of $Y$ are of the form $\gamma(t) = (v_0, a_0\, e^{tv_0})$. Now use the flow of $Y$,

$$\Psi\colon \mathbb{R} \times \mathfrak{g} \times G \to \mathfrak{g} \times G,$$

to write $\exp$ as the composition of smooth maps.)

**13**. Show that, for the homomorphism $\det\colon \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}^\bullet$, we have $\det'(I_n)(x) = \mathrm{tr}(x)$, where tr denotes the trace function. Conclude, using Theorem 1 that, for any matrix $a$,

$$\det(e^a) = e^{\mathrm{tr}(a)}.$$

**14**. Prove that, for any $g \in G$ and any $x \in \mathfrak{g}$, we have the identity

$$g \exp(x)\, g^{-1} = \exp\bigl(\mathrm{Ad}(g)(x)\bigr).$$

(Hint: Replace $x$ by $tx$ in the above formula and consider Proposition 5.) Use this to show that $\mathrm{tr}\bigl(\exp(x)\bigr) \geq -2$ for all $x \in \mathfrak{sl}(2, \mathbb{R})$. Conclude that $\exp\colon \mathfrak{sl}(2, \mathbb{R}) \to \mathrm{SL}(2, \mathbb{R})$ is *not* surjective. (Hint: show that every $x \in \mathfrak{sl}(2, \mathbb{R})$ is of the form $gyg^{-1}$ for some $g \in \mathrm{SL}(2, \mathbb{R})$ and some $y$ which is one of the matrices

$$\begin{pmatrix} 0 & \pm 1 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} 0 & -\lambda \\ \lambda & 0 \end{pmatrix}, \qquad (\lambda > 0).$$

Also, remember that $\mathrm{tr}\bigl(aba^{-1}\bigr) = \mathrm{tr}(b)$.)

E.2.2 $\hspace{10cm}$ 34

**15**. Using Theorem 1, show that if $H_1$ and $H_2$ are Lie subgroups of $G$, then $H_1 \cap H_2$ is also a Lie subgroup of $G$. (Hint: What should the Lie algebra of this intersection be? Be careful: $H_1 \cap H_2$ might have countably many distinct components even if $H_1$ and $H_2$ are connected!)

**16**. For any skew-commutative algebra $(\mathfrak{g}, [,])$, we define the map $\mathrm{ad}\colon \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ by $\mathrm{ad}(x)(y) = [x, y]$. Verify that the validity of the Jacobi identity $[\mathrm{ad}(x), \mathrm{ad}(y)] = \mathrm{ad}([x, y])$ (where, as usual, the bracket on $\mathrm{End}(\mathfrak{g})$ is the commutator) is equivalent to the validity of the identity
$$[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$$
for all $x, y, z \in \mathfrak{g}$.

**17**. Show that, as $\lambda \in \mathbb{R}$ varies, all of the groups
$$G_\lambda = \left\{ \begin{pmatrix} a & b \\ 0 & a^\lambda \end{pmatrix} \;\middle|\; a \in \mathbb{R}^+,\ b \in \mathbb{R} \right\}$$
with $\lambda \neq 1$ are isomorphic, but are not conjugate in $\mathrm{GL}(2, \mathbb{R})$. What happens when $\lambda = 1$?

**18**. Show that a connected Lie group $G$ is abelian if and only if its Lie algebra satisfies $[x, y] = 0$ for all $x, y \in \mathfrak{g}$. Conclude that a connected abelian Lie group of dimension n is isomorphic to $\mathbb{R}^n / \mathbb{Z}^d$ where $\mathbb{Z}^d$ is some discrete subgroup of rank $d \leq n$. (Hint: To show "$G$ abelian" implies "$\mathfrak{g}$ abelian", look at how $[,]$ was defined. To prove the converse, use Theorem 3 to construct a surjective homomorphism $\phi\colon \mathbb{R}^n \to G$ with discrete kernel.)

**19**. *(Covering Spaces of Lie groups.)* Let $G$ be a connected Lie group and let $\pi\colon \tilde{G} \to G$ be the universal covering space of $G$. (Recall that the points of $\tilde{G}$ can be regarded as the space of fixed-endpoint homotopy classes of continuous maps $\gamma\colon [0, 1] \to G$ with $\gamma(0) = e$.) Show that there is a unique Lie group structure $\tilde{\mu}\colon \tilde{G} \times \tilde{G} \to \tilde{G}$ for which the homotopy class of the constant map $\tilde{e} \in \tilde{G}$ is the identity and so that $\pi$ is a homomorphism. (Hints: Give $\tilde{G}$ the (unique) smooth structure for which $\pi$ is a local diffeomorphism. The multiplication $\tilde{\mu}$ can then be defined as follows: The map $\bar{\mu} = \mu \circ (\pi \times \pi)\colon \tilde{G} \times \tilde{G} \to G$ is a smooth map and satisfies $\bar{\mu}(\tilde{e}, \tilde{e}) = e$. Since $\tilde{G} \times \tilde{G}$ is simply connected, the universal lifting property of the covering map $\pi$ implies that there is a unique map $\tilde{\mu}\colon \tilde{G} \times \tilde{G} \to \tilde{G}$ which satisfies $\pi \circ \tilde{\mu} = \bar{\mu}$ and $\tilde{\mu}(\tilde{e}, \tilde{e}) = \tilde{e}$. Show that $\tilde{\mu}$ is smooth, that it satisfies the axioms for a group multiplication (associativity, existence of an identity, and existence of inverses), and that $\pi$ is a homomorphism. You will want to use the universal lifting property of covering spaces a few times.)

The kernel of $\pi$ is a discrete normal subgroup of $\tilde{G}$. Show that this kernel lies in the center of $G$. (Hint: For any $z \in \ker(\pi)$, the connected set $\{aza^{-1} \mid a \in G\}$ must also lie in $\ker(\pi)$.)

Show that the center of the simply connected Lie group
$$G = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \;\middle|\; a \in \mathbb{R}^+,\ b \in \mathbb{R} \right\}$$

is trivial, so any connected Lie group with the same Lie algebra is actually isomorphic to $G$.

(In the next Lecture, we will show that whenever $K$ is a closed normal subgroup of a Lie group $G$, the quotient group $G/K$ can be given the structure of a Lie group. Thus, in many cases, one can effectively list all of the connected Lie groups with a given Lie algebra.)

**20**. Show that $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ is not a matrix group! In fact, show that any homomorphism $\phi\colon \widetilde{\mathrm{SL}(2,\mathbb{R})} \to \mathrm{GL}(n,\mathbb{R})$ factors through the projections $\widetilde{\mathrm{SL}(2,\mathbb{R})} \to \mathrm{SL}(2,\mathbb{R})$. (Hint: Recall, from earlier exercises, that the inclusion map $\mathrm{SL}(2,\mathbb{R}) \to \mathrm{SL}(2,\mathbb{C})$ induces the zero map on $\pi_1$ since $\mathrm{SL}(2,\mathbb{C})$ is simply connected. Now, any homomorphism $\phi\colon \widetilde{\mathrm{SL}(2,\mathbb{R})} \to \mathrm{GL}(n,\mathbb{R})$ induces a Lie algebra homomorphism $\phi'(e)\colon \mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{gl}(n,\mathbb{R})$ and this may clearly be complexified to yield a Lie algebra homomorphism $\phi'(e)^{\mathbb{C}}\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(n,\mathbb{C})$. Since $\mathrm{SL}(2,\mathbb{C})$ is simply connected, there must be a corresponding Lie group homorphism $\phi^{\mathbb{C}}\colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{GL}(n,\mathbb{C})$. Now suppose that $\phi$ does not factor through $\mathrm{SL}(2,\mathbb{R})$, i.e., that $\phi$ is non-trivial on the kernel of $\widetilde{\mathrm{SL}(2,\mathbb{R})} \to \mathrm{SL}(2,\mathbb{R})$, and show that this leads to a contradiction.)

**21**. An *ideal* in a Lie algebra $\mathfrak{g}$ is a linear subspace $\mathfrak{h}$ which satisfies $[\mathfrak{h},\mathfrak{g}] \subset \mathfrak{h}$. Show that the kernel $\mathfrak{k}$ of a Lie algebra homomorphism $\varphi\colon \mathfrak{h} \to \mathfrak{g}$ is an ideal in $\mathfrak{h}$ and that the image $\varphi(\mathfrak{h})$ is a subalgebra of $\mathfrak{g}$. Conversely, show that if $\mathfrak{k} \subset \mathfrak{h}$ is an ideal, then the quotient vector space $\mathfrak{h}/\mathfrak{k}$ carries a unique Lie algebra structure for which the quotient mapping $\mathfrak{h} \to \mathfrak{h}/\mathfrak{k}$ is a homomorphism.

Show that the subspace $[\mathfrak{g},\mathfrak{g}]$ of $\mathfrak{g}$ which is generated by all brackets of the form $[x,y]$ is an ideal in $\mathfrak{g}$. What can you say about the quotient $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$?

**22**. Show that, for a connected Lie group $G$, a connected Lie subgroup $H$ is normal if and only if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. (Hint: Use Proposition 7 and the fact that $H \subset G$ is normal if and only if $e^x H e^{-x} = H$ for all $x \in \mathfrak{g}$.)

**23**. For any Lie algebra $\mathfrak{g}$, let $\mathfrak{z}(\mathfrak{g}) \subset \mathfrak{g}$ denote the kernel of the homomorphism $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. Use Theorem 2 and Exercise 16 to prove Theorem 4 for any Lie algebra $\mathfrak{g}$ for which $\mathfrak{z}(\mathfrak{g}) = 0$. (Hint: Look at the discussion after the statement of Theorem 4.)

Show also that if $\mathfrak{g}$ is the Lie algebra of the connected Lie group $G$, then the connected Lie subgroup $Z(\mathfrak{g}) \subset G$ which corresponds to $\mathfrak{z}(\mathfrak{g})$ lies in the center of $G$. (In the next lecture, we will be able to prove that the center of $G$ is a closed Lie subgroup of $G$ and that $Z(\mathfrak{g})$ is actually the identity component of the center of $G$.)

**24**. For any Lie algebra $\mathfrak{g}$, there is a canonical bilinear pairing $\kappa\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, called the *Killing form*, defined by the rule:

$$\kappa(x,y) = \mathrm{tr}\big(\mathrm{ad}(x)\mathrm{ad}(y)\big).$$

(i) Show that $\kappa$ is symmetric and, if $\mathfrak{g}$ is the Lie algebra of a Lie group $G$, then $\kappa$ is Ad-invariant:
$$\kappa\big(\mathrm{Ad}(g)x, \mathrm{Ad}(g)y\big) = \kappa(x, y) = \kappa(y, x).$$
Show also that
$$\kappa\big([z, x], y\big) = -\kappa\big(x, [z, y]\big).$$

A Lie algebra $\mathfrak{g}$ is said to be *semi-simple* if $\kappa$ is a non-degenerate bilinear form on $\mathfrak{g}$.

(ii) Show that, of all the 2- and 3-dimensional Lie algebras, only $\mathfrak{so}(3)$ and $\mathfrak{sl}(2, \mathbb{R})$ are semi-simple.

(iii) Show that if $\mathfrak{h} \subset \mathfrak{g}$ is an ideal in a semi-simple Lie algebra $\mathfrak{g}$, then the Killing form of $\mathfrak{h}$ as an algebra is equal to the restriction of the Killing form of $\mathfrak{g}$ to $\mathfrak{h}$. Show also that the subspace $\mathfrak{h}^{\perp} = \{x \in \mathfrak{g} \mid \kappa(x, y) = 0 \text{ for all } y \in \mathfrak{h}\}$ is also an ideal in $\mathfrak{g}$ and that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$ as Lie algebras. (Hint: For the first part, examine the effect of $\mathrm{ad}(x)$ on a basis of $\mathfrak{g}$ chosen so that the first $\dim \mathfrak{h}$ basis elements are a basis of $\mathfrak{h}$.)

(iv) Finally, show that a semi-simple Lie algebra can be written as a direct sum of ideals $\mathfrak{h}_i$, each of which has no proper ideals. (Hint: Apply (iii) as many times as you can find proper ideals of the summands found so far.)

A more general class of Lie algebras are the reductive ones. We say that a Lie algebra is *reductive* if there is a non-degenerate symmetric bilinear form $(\,,\,)\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ which satisifes the identity $\big([z, x], y\big) + \big(x, [z, y]\big) = 0$. Using the above arguments, it is easy to see that a reductive algebra can be written as the direct sum of an abelian algebra and some number of simple algebras in a unique way.

**25**. Show that, if $\omega$ is the canonical left-invariant 1-form on $G$ and $Y_v$ is the right-invariant vector field on $G$ satisfying $Y_v(e) = v$, then

$$\omega\big(Y_v(a)\big) = \mathrm{Ad}\big(a^{-1}\big)(v).$$

(Remark: For any skew-commutative algebra $(\mathfrak{a}, [,])$, the function $[\![\,,\,]\!]\colon \mathfrak{a} \times \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$ defined by
$$[\![x, y, z]\!] = [[x, y], z] + [[y, z], x] + [[z, x], y]$$
is tri-linear and skew-symmetric, and hence represents an element of $\mathfrak{a} \otimes \Lambda^3(\mathfrak{a}^*)$.)

<center>

## Lecture 3:

## Group Actions on Manifolds

</center>

In this lecture, I turn from the abstract study of Lie groups to their realizations as "transformation groups."

**Lie group actions.**

**Definition 1:** If $(G, \mu)$ is a Lie group and $M$ is a smooth manifold, then a *left action* of $G$ on $M$ is a smooth mapping $\lambda \colon G \times M \to M$ which satisfies $\lambda(e, m) = m$ for all $m \in M$ and

$$\lambda(\mu(a, b), m) = \lambda(a, \lambda(b, m)).$$

Similarly, a *right action* of $G$ on $M$ is a smooth mapping $\rho \colon M \times G \to M$, which satisfies $\rho(m, e) = m$ for all $m \in M$ and

$$\rho(m, \mu(a, b)) = \rho(\rho(m, a), b).$$

For notational sanity, whenever the action (left or right) can be easily inferred from context, we will usually write $a \cdot m$ instead of $\lambda(a, m)$ or $m \cdot a$ instead of $\rho(m, a)$. Thus, for example, the axioms for a left action in this abbreviated notation are simply $e \cdot m = m$ and $a \cdot (b \cdot m) = ab \cdot m$.

For a given a left action $\lambda \colon G \times M \to M$, it is easy to see that for each fixed $a \in G$ the map $\lambda_a \colon M \to M$ defined by $\lambda_a(m) = \lambda(a, m)$ is a smooth diffeomorphism of $M$ onto itself. Thus, $G$ gets represented as a group of diffeomorphisms, or "transformations" of a manifold $M$. This notion of "transformation group" was what motivated Lie to develop his theory in the first place. See the Appendix to this Lecture for a more complete discussion of this point.

**Equivalence of Left and Right Actions.** Note that every right action $\rho \colon M \times G \to M$ can be rewritten as a left action and vice versa. One merely defines

$$\tilde{\rho}(a, m) = \rho(m, a^{-1}).$$

(The reader should check that this $\tilde{\rho}$ is, in fact, a left action.) Thus, all theorems about left actions have analogues for right actions. The distinction between the two is mainly for notational and conceptual convenience. I will concentrate on left actions and only occasionally point out the places where right actions behave slightly differently (mainly changes of sign, etc.).

**Stabilizers and Orbits.** A left action is said to be *effective* if $g \cdot m = m$ for all $m \in M$ implies that $g = e$. (Sometimes, the word *faithful* is used instead.) A left action is said to be *free* if $g \neq e$ implies that $g \cdot m \neq m$ for all $m \in M$.

A left action is said to be *transitive* if, for any $x, y \in M$, there exists a $g \in G$ so that $g \cdot x = y$. In this case, $M$ is usually said to be *homogeneous* under the given action.

For any $m \in M$, the *G-orbit* of $m$ is defined to be the set

$$G \cdot m = \{g \cdot m \,|\, g \in G\}$$

and the *stabilizer* (or *isotropy group*) of $m$ is defined to be the subset

$$G_m = \{g \in G \,|\, g \cdot m = m\}.$$

Note that

$$G_{g \cdot m} = g \, G_m \, g^{-1}.$$

Thus, whenever $H \subset G$ is the stabilizer of a point of $M$, then all of the conjugate subgroups of $H$ are also stabilizers. These results imply that

$$G_M = \bigcap_{m \in M} G_m$$

is a closed normal subgroup of $G$ and consists of those $g \in G$ for which $g \cdot m = m$ for all $m \in M$. Often in practice, $G_M$ is a discrete (in fact, usually finite) subgroup of $G$. When this is so, we say that the action is *almost effective*.

The following theorem says that orbits and stabilizers are particularly nice objects. Though the proof is relatively straightforward, it is a little long, so we will consider a few examples before attempting it.

**Theorem 1:** *Let* $\lambda \colon G \times M \to M$ *be a left action of* $G$ *on* $M$. *Then, for all* $m \in M$, *the stabilizer* $G_m$ *is a closed Lie subgroup of* $G$. *Moreover, the orbit* $G \cdot m$ *can be given the structure of a smooth submanifold of* $M$ *in such a way that the map* $\phi \colon G \to G \cdot m$ *defined by* $\phi(g) = \lambda(g, m)$ *is a smooth submersion.*

**Example 1.** Any Lie group left-acts on itself by left multiplication. I.e., we set $M = G$ and define $\lambda \colon G \times M \to M$ to simply be $\mu$. This action is both free and transitive.

**Example 2.** Given a homomorphism of Lie groups $\phi \colon H \to G$, define a smooth left action $\lambda \colon H \times G \to G$ by the rule $\lambda(h, g) = \phi(h)g$. Then $H_e = \ker(\phi)$ and $H \cdot e = \phi(H) \subset G$.

▶      *In particular, Theorem 1 implies that the kernel of a Lie group homomorphism is a (closed, normal) Lie subgroup of the domain group and the image of a Lie group homomorphism is a Lie subgroup of the range group.*

**Example 3.** Any Lie group acts on itself by conjugation: $g \cdot g_0 = g g_0 g^{-1}$. This action is neither free nor transitive (unless $G = \{e\}$). Note that $G_e = G$ and, in general, $G_g$ is the centralizer of $g \in G$. This action is effective (respectively, almost effective) if and only if the center of $G$ is trivial (respectively, discrete). The orbits are the conjugacy classes of $G$.

**Example 4.** $GL(n, \mathbb{R})$ acts on $\mathbb{R}^n$ as usual by $A \cdot v = Av$. This action is effective but is neither free nor transitive since $GL(n, \mathbb{R})$ fixes $0 \in \mathbb{R}^n$ and acts transitively on $\mathbb{R}^n \backslash \{0\}$. Thus, there are exactly two orbits of this action, one closed and the other not.

**Example 5.** $\mathrm{SO}(n+1)$ acts on $S^n = \{x \in \mathbb{R}^{n+1} \,|\, x \cdot x = 1\}$ by the usual action $A \cdot x = Ax$. This action is transitive and effective, but not free (unless $n = 1$) since, for example, the stabilizer of $e_{n+1}$ is clearly isomorphic to $\mathrm{SO}(n)$.

**Example 6.** Let $\mathcal{S}_n$ be the $n(n+1)/2$-dimensional vector space of $n$-by-$n$ real symmetric matrices. Then $\mathrm{GL}(n, \mathbb{R})$ acts on $\mathcal{S}_n$ by $A \cdot S = A S \,{}^t\!A$. The orbit of the identity matrix $I_n$ is $\mathcal{S}_+(n)$, the set of all positive-definite $n$-by-$n$ real symmetric matrices (Why?). In fact, it is known that, if we define $I_{p,q} \in \mathcal{S}_n$ to be the matrix

$$
I_{p,q} = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix},
$$

(where the "0" entries have the appropriate dimensions) then $\mathcal{S}_n$ is the (disjoint) union of the orbits of the matrices $I_{p,q}$ where $0 \le p, q$ and $p + q \le n$ (see the Exercises).

The orbit of $I_{p,q}$ is open in $\mathcal{S}_n$ iff $p + q = n$. The stabilizer of $I_{p,q}$ in this case is defined to be $\mathrm{O}(p,q) \subset \mathrm{GL}(n,\mathbb{R})$.

Note that the action is merely almost effective since $\{\pm I_n\} \subset \mathrm{GL}(n, \mathbb{R})$ fixes every $S \in \mathcal{S}_n$.

**Example 7.** Let $\mathcal{J} = \{J \in \mathrm{GL}(2n, \mathbb{R}) \,|\, J^2 = -I_{2n}\}$. Then $\mathrm{GL}(2n, \mathbb{R})$ acts on $\mathcal{J}$ on the left by the formula $A \cdot J = AJA^{-1}$. I leave as exercises for the reader to prove that $\mathcal{J}$ is a smooth manifold and that this action of $\mathrm{GL}(2n, \mathbb{R})$ is transitive and almost effective. The stabilizer of $J_0 = $ multiplication by $i$ in $\mathbb{C}^n$ ($= \mathbb{R}^{2n}$) is simply $\mathrm{GL}(n, \mathbb{C}) \subset \mathrm{GL}(2n, \mathbb{R})$.

**Example 8.** Let $M = \mathbb{RP}^1$, denote the projective line, whose elements are the lines through the origin in $\mathbb{R}^2$. We will use the notation $\begin{bmatrix} x \\ y \end{bmatrix}$ to denote the line in $\mathbb{R}^2$ spanned by the non-zero vector $\begin{pmatrix} x \\ y \end{pmatrix}$.

Let $G = \mathrm{SL}(2, \mathbb{R})$ act on $\mathbb{RP}^1$ on the left by the formula

$$
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.
$$

This action is easily seen to be almost effective, with only $\pm I_2 \in \mathrm{SL}(2, \mathbb{R})$ acting trivially.

Actually, it is more common to write this action more informally by using the identification $\mathbb{RP}^1 = \mathbb{R} \cup \{\infty\}$ which identifies $\begin{bmatrix} x \\ y \end{bmatrix}$ when $y \ne 0$ with $x/y \in \mathbb{R}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with $\infty$. With this convention, the action takes on the more familiar "linear fractional" form

$$
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot x = \frac{ax + b}{cx + d}.
$$

Note that this form of the action makes it clear that the so-called "linear fractional" action or "Möbius" action on the real line is just the projectivization of the usual linear representation of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{R}^2$.

L.3.3

We now turn to the proof of Theorem 1.

PROOF OF THEOREM 1: Fix $m \in M$ and define $\phi: G \to M$ by $\phi(g) = \lambda(g, m)$ as in the theorem. Since $G_m = \phi^{-1}(m)$, it follows that $G_m$ is a closed subset of $G$. The axioms for a left action clearly imply that $G_m$ is closed under multiplication and inverse, so it is a subgroup.

I claim that $G_m$ is a submanifold of $G$. To see this, let $\mathfrak{g}_m \subset \mathfrak{g} = T_e G$ be the kernel of the mapping $\phi'(e): T_e G \to T_m M$. Since $\phi \circ L_g = \lambda_g \circ \phi$ for all $g \in G$, the Chain Rule yields a commutative diagram:

$$
\begin{array}{ccc}
\mathfrak{g} & \xrightarrow{L_g'(e)} & T_g G \\
{\scriptstyle \phi'(e)}\downarrow & & \downarrow{\scriptstyle \phi'(g)} \\
T_m M & \xrightarrow{\lambda_g'(m)} & T_{g \cdot m} M
\end{array}
$$

Since both $L_g'(e)$ and $\lambda_g'(m)$ are isomorphisms, it follows that $\ker(\phi'(g)) = L_g'(e)(\mathfrak{g}_m)$ for all $g \in G$. In particular, the rank of $\phi'(g)$ is independent of $g \in G$. By the Implicit Function Theorem (see Exercise 2), it follows that $\phi^{-1}(m) = G_m$ is a smooth submanifold of $G$.

It remains to show that the orbit $G \cdot m$ can be given the structure of a smooth submanifold of $M$ with the stated properties. That is, that $G \cdot m$ can be given a second countable, Hausdorff, locally Euclidean topology and a smooth structure for which the inclusion map $G \cdot m \hookrightarrow M$ is a smooth immersion and for which the map $\phi: G \to G \cdot m$ is a submersion.

Before embarking on this task, it is useful to remark on the nature of the fibers of the map $\phi$. By the axioms for left actions, $\phi(h) = h \cdot m = g \cdot m = \phi(g)$ if and only if $g^{-1}h \cdot m = m$, i.e., if and only if $g^{-1}h$ lies in $G_m$. This is equivalent to the condition that $h$ lie in the left $G_m$-coset $gG_m$. Thus, the fibers of the map $\phi$ are the left $G_m$-cosets in $G$. In particular, the map $\phi$ establishes a bijection $\bar{\phi}: G/G_m \to G \cdot m$.

First, I specify the topology on $G \cdot m$ to be quotient topology induced by the surjective map $\phi: G \to G \cdot m$. Thus, a set $U$ in $G \cdot m$ is open if and only if $\phi^{-1}(U)$ is open in $G$. Since $\phi: G \to M$ is continuous, the quotient topology on the image $G \cdot m$ is at least as fine as the subspace topology $G \cdot m$ inherits via inclusion into $M$. Since the subspace topology is Hausdorff, the quotient topology must be also. Moreover, the quotient topology on $G \cdot m$ is also second countable since the topology of $G$ is. For the rest of the proof, "the topology on $G \cdot m$" means the quotient topology.

I will both establish the locally Euclidean nature of this topology and construct a smooth structure on $G \cdot m$ at the same time by finding the required neighborhood charts and proving that they are smooth on overlaps. First, however, I need a lemma establishing the existence of a "tubular neighborhood" of the submanifold $G_m \subset G$. Let $d = \dim(G) - \dim(G_m)$. Then there exists a smooth mapping $\psi: B^d \to G$ (where $B^d$ is an open ball about 0 in $\mathbb{R}^d$) so that $\psi(0) = e$ and so that $\mathfrak{g}$ is the direct sum of the subspaces $\mathfrak{g}_m$ and $V = \psi'(0)(\mathbb{R}^d)$. By the Chain Rule and the definition of $\mathfrak{g}_m$, it follows that $(\phi \circ \psi)'(0): \mathbb{R}^d \to$

$T_m M$ is injective. Thus, by restricting to a smaller ball in $\mathbb{R}^d$ if necessary, I may assume henceforth that $\phi \circ \psi \colon B^d \to M$ is a smooth embedding.

Consider the mapping $\Psi \colon B^d \times G_m \to G$ defined by $\Psi(x, g) = \psi(x)g$. I claim that $\Psi$ is a diffeomorphism onto its image (which is an open set), say $U = \Psi(B^d \times G_m) \subset G$. (Thus, $U$ forms a sort of "tubular neighborhood" of the submanifold $G_m$ in $G$.)

To see this, first I show that $\Psi$ is one-to-one: If $\Psi(x_1, g_1) = \Psi(x_2, g_2)$, then

$$(\phi \circ \psi)(x_1) = \psi(x_1) \cdot m = (\psi(x_1)g_1) \cdot m = (\psi(x_2)g_2) \cdot m = \psi(x_2) \cdot m = (\phi \circ \psi)(x_2),$$

so the injectivity of $\phi \circ \psi$ implies $x_1 = x_2$. Since $\psi(x_1)g_1 = \psi(x_2)g_2$, this in turn implies that $g_1 = g_2$.

Second, I must show that the derivative

$$\Psi'(x, g) \colon T_x \mathbb{R}^d \oplus T_g G_m \to T_{\psi(x)g} G$$

is an isomorphism for all $(x, g) \in B^d \times G_m$. However, from the beginning of the proof, $\ker(\phi'(\psi(x)g)) = L'_{\psi(x)g}(e)(\mathfrak{g}_m)$ and this latter space is clearly $\Psi'(x, g)(0 \oplus T_g G_m)$. On the other hand, since $\phi(\Psi(x, g)) = \phi \circ \psi(x)$, it follows that

$$\phi'(\Psi(x, g))\big(\Psi'(x, g)(T_x \mathbb{R}^d \oplus 0)\big) = (\phi \circ \psi)'(x)(T_x \mathbb{R}^d)$$

and this latter space has dimension $d$ by construction. Hence, $\Psi'(x, g)(T_x \mathbb{R}^d \oplus 0)$ is a $d$-dimensional subspace of $T_{\psi(x)g} G$ which is transverse to $\Psi'(x, g)(0 \oplus T_g G_m)$. Thus, $\Psi'(x, g) \colon T_x \mathbb{R}^d \oplus T_g G_m \to T_{\psi(x)g} G$ is surjective and hence an isomorphism, as desired.

This completes the proof that $\Psi$ is a diffeomorphism onto U. It follows that the inverse of $\Psi$ is smooth and can be written in the form $\Psi^{-1} = \pi_1 \times \pi_2$ where $\pi_1 \colon U \to B^d$ and $\pi_2 \colon U \to G_m$ are smooth submersions.

Now, for each $g \in G$, define $\rho_g \colon B^d \to M$ by the formula $\rho_g(x) = \phi\big(g\psi(x)\big)$. Then $\rho_g = \lambda_g \circ \phi \circ \psi$, so $\rho_g$ is a smooth embedding of $B^d$ into $M$. By construction, $U = \phi^{-1}(\phi \circ \psi(B^d)) = \phi^{-1}(\rho_e(B^d))$ is an open set in $G$, so it follows that $\rho_e(B^d)$ is an open neighborhood of $e \cdot m = m$ in $G \cdot m$. By the axioms for left actions, it follows that $\phi^{-1}\big(\rho_g(B^d)\big) = L_g(U)$ (which is open in $G$) for all $g \in G$. Thus, $\rho_g(B^d)$ is an open neighborhood of $g \cdot m$ in $G \cdot m$ (in the quotient topology). Moreover, contemplating the commutative square

$$
\begin{array}{ccc}
U & \xrightarrow{L_g} & L_g(U) \\
{\scriptstyle \pi_1} \downarrow & & \downarrow {\scriptstyle \phi} \\
B^d & \xrightarrow{\rho_g} & \rho_g(B^d)
\end{array}
$$

whose upper horizontal arrow is a diffeomorphism which identifies the fibers of the vertical arrows (each of which is a topological identification map) implies that $\rho_g$ is, in fact, a homeomorphism onto its image. Thus, the quotient topology is locally Euclidean.

Finally, I show that the "patches" $\rho_g$ overlap smoothly. Suppose that

$$\rho_g(B^d) \cap \rho_h(B^d) \neq \emptyset.$$

Then, because the maps $\rho_g$ and $\rho_h$ are homeomorphisms,

$$\rho_g(B^d) \cap \rho_h(B^d) = \rho_g(W_1) = \rho_h(W_2)$$

where $W_i \neq \emptyset$ are open subsets of $B^d$. It follows that

$$L_g\big(\Psi(W_1 \times G_m)\big) = L_h\big(\Psi(W_2 \times G_m)\big).$$

Thus, if $\tau : W_1 \to W_2$ is defined by the rule $\tau = \pi_1 \circ L_{h^{-1}} \circ L_g \circ \psi$, then $\tau$ is a smooth map with smooth inverse $\tau^{-1} = \pi_1 \circ L_{g^{-1}} \circ L_h \circ \psi$ and hence is a diffeomorphism. Moreover, we have $\rho_g = \rho_h \circ \tau$, thus establishing that the patches $\rho_g$ overlap smoothly and hence that the patches define the structure of a smooth manifold on $G \cdot m$.

That the map $\phi : G \to G \cdot m$ is a smooth submersion and that the inclusion $G \cdot m \hookrightarrow M$ is a smooth one-to-one immersion are now clear. $\qquad\qquad\square$

It is worth remarking that the proof of Theorem 1 shows that the Lie algebra of $G_m$ is the subspace $\mathfrak{g}_m$. In particular, if $G_m = \{e\}$, then the map $\phi : G \to M$ is a one-to-one immersion.

The proof also brings out the fact that the orbit $G \cdot m$ can be identified with the left coset space $G/G_m$, which thereby inherits the structure of a smooth manifold. It is natural to wonder which subgroups $H$ of $G$ have the property that the coset space $G/H$ can be given the structure of a smooth manifold for which the coset projection $\pi : G \to G/H$ is a smooth map. This question is answered by the following result. The proof is quite similar to that of Theorem 1, so I will only provide an outline, leaving the details as exercises for the reader.

**Theorem 2:** *If $H$ is a closed subgroup of a Lie group $G$, then the left coset space $G/H$ can be given the structure of a smooth manifold in a unique way so that the coset mapping $\pi : G \to G/H$ is a smooth submersion. Moreover, with this smooth structure, the left action $\lambda : G \times G/H \to G/H$ defined by $\lambda(g, hH) = ghH$ is a transitive smooth left action.*

PROOF: (Outline.) If the coset mapping $\pi : G \to G/H$ is to be a smooth submersion, elementary linear algebra tells us that the dimension of $G/H$ will have to be $d = \dim(G) - \dim(H)$. Moreover, for every $g \in G$, there will have to exist a smooth mapping $\psi_g : B^d \to G$ with $\psi_g(0) = g$ which is transverse to the submanifold $gH$ at $g$ and so that the composition $\pi \circ \psi : B^d \to G/H$ is a diffeomorphism onto a neighborhood of $gH \in G/H$. It is not difficult to see that this is only possible if $G/H$ is endowed with the quotient topology. The hypothesis that $H$ be closed implies that the quotient topology is Hausdorff. It is automatic that the quotient topology is second countable. The proof that the quotient topology is locally Euclidean depends on being able to construct the "tubular neighborhood" $U$ of $H$ as constructed for the case of a stabilizer subgroup in the proof of Theorem 1. Once this is done, the rest of the construction of charts with smooth overlaps follows the end of the proof of Theorem 1 almost verbatim. $\qquad\square$

**Group Actions and Vector Fields.** A left action $\lambda\colon \mathbb{R} \times M \to M$ (where $\mathbb{R}$ has its usual additive Lie group structure) is, of course, the same thing as a flow. Associated to each flow on $M$ is a vector field which generates this flow. The generalization of this association to more general Lie group actions is the subject of this section.

Let $\lambda\colon G \times M \to M$ be a left action. Then, for each $v \in \mathfrak{g}$, there is a flow $\Psi_v^\lambda$ on $M$ defined by the formula

$$\Psi_v^\lambda(t, m) = e^{tv} \cdot m.$$

This flow is associated to a vector field on $M$ which we shall denote by $Y_v^\lambda$, or simply $Y_v$ if the action $\lambda$ is clear from context. This defines a mapping $\lambda_*\colon \mathfrak{g} \to \mathfrak{X}(M)$, where $\lambda_*(v) = Y_v^\lambda$.

**Proposition 1:** *For each left action $\lambda\colon G \times M \to M$, the mapping $\lambda_*$ is a linear anti-homomorphism from $\mathfrak{g}$ to $\mathfrak{X}(M)$. In other words, $\lambda_*$ is linear and*

$$\lambda_*([x, y]) = -[\lambda_*(x), \lambda_*(y)].$$

PROOF: For each $v \in \mathfrak{g}$, let $Y_v$ denote the *right* invariant vector field on $G$ whose value at $e$ is $v$. Then, according to Lecture 2, the flow of $Y_v$ on $G$ is given by the formula $\Psi_v(t, g) = \exp(tv)g$. As usual, let $\Phi_v$ denote the flow of the left invariant vector field $X_v$. Then the formula

$$\Psi_v(t, g) = \left( \Phi_{-v}(t, g^{-1}) \right)^{-1}$$

is immediate. If $\iota_*\colon \mathfrak{X}(G) \to \mathfrak{X}(G)$ is the map induced by the diffeomorphism $\iota(g) = g^{-1}$, then the above formula implies

$$\iota_*(X_{-v}) = Y_v.$$

In particular, since $\iota_*$ commutes with Lie bracket, it follows that

$$[Y_x, Y_y] = -Y_{[x,y]}$$

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

Now, regard $Y_v$ and $\Psi_v$ as being defined on $G \times M$ in the obvious way, i.e., $\Psi_v(g, m) = (e^{tv}g, m)$. Then $\lambda$ intertwines this flow with that of $\Psi_v^\lambda$:

$$\lambda \circ \Psi_v = \Psi_v^\lambda \circ \lambda.$$

It follows that the vector fields $Y_v$ and $Y_v^\lambda$ are $\lambda$-related. Thus, $[Y_x^\lambda, Y_y^\lambda]$ is $\lambda$-related to $[Y_x, Y_y] = -Y_{[x,y]}$ and hence must be equal to $-Y_{[x,y]}^\lambda$. Finally, since the map $v \mapsto Y_v$ is clearly linear, it follows that $\lambda_*$ is also linear. $\qquad\square$

▶     The appearance of the minus sign in the above formula is something of an annoyance and has led some authors (cf. [A]) to introduce a non-classical minus sign into either the definition of the Lie bracket of vector fields or the definition of the Lie bracket on $\mathfrak{g}$ in order to get rid of the minus sign in this theorem. Unfortunately, as logical as this revisionism is, it has not been particularly popular. However, let the reader of other sources beware when comparing formulas.

Even *with* a minus sign, however, Proposition 1 implies that the subspace $\lambda_*(\mathfrak{g}) \subset \mathfrak{X}(M)$ is a (finite dimensional) Lie subalgebra of the Lie algebra of all vector fields on $M$.

**Example: Linear Fractional Transformations.** Consider the Möbius action introduced earlier of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{RP}^1$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot s = \frac{as + b}{cs + d}.$$

A basis for the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$ is

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Thus, for example, the flow $\Psi_y^\lambda$ is given by

$$\Psi_y^\lambda(t, s) = \exp \begin{pmatrix} 0 & 0 \\ t & 0 \end{pmatrix} \cdot s = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \cdot s = \frac{s}{ts + 1} = s - s^2\, t + \cdots,$$

so $Y_y^\lambda = -s^2 \partial/\partial s$. In fact, it is easy to see that, in general,

$$\lambda_*(a_0 x + a_1 h + a_2 y) = (a_0 + 2a_1 s - a_2 s^2) \frac{\partial}{\partial s}.$$

The basic ODE existence theorem can be thought of as saying that every vector field $X \in \mathfrak{X}(M)$ arises as the "flow" of a "local" $\mathbb{R}$-action on $M$. There is a generalization of this to finite dimensional subalgebras of $\mathfrak{X}(M)$. To state it, we first define a *local left action* of a Lie group $G$ on a manifold $M$ to be an open neighborhood $U \subset G \times M$ of $\{e\} \times M$ together with a smooth map $\lambda: U \to M$ so that $\lambda(e, m) = m$ for all $m \in M$ and so that

$$\lambda\big(a, \lambda(b, m)\big) = \lambda(ab, m)$$

whenever this makes sense, i.e., whenever $(b, m)$, $(ab, m)$, and $\big(a, \lambda(b, m)\big)$ all lie in $U$.

It is easy to see that even a mere local Lie group action induces a map $\lambda_*: \mathfrak{g} \to \mathfrak{X}(M)$ as before. We can now state the following result, whose proof is left to the Exercises:

**Proposition 2:** *Let $G$ be a Lie group and let $\varphi: \mathfrak{g} \to \mathfrak{X}(M)$ be a Lie algebra homomorphism. Then there exists a local left action $(U, \lambda)$ of $G$ on $M$ so that $\lambda_* = -\varphi$.* $\qquad \square$

For example, the linear fractional transformations of the last example could just as easily been regarded as a local action of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{R}$, where the open set $U \subset \mathrm{SL}(2, \mathbb{R}) \times \mathbb{R}$ is just the set of pairs where $cs + d \neq 0$.

**Equations of Lie type.** Early in the theory of Lie groups, a special family of ordinary differential equations was singled out for study which generalized the theory of linear equations and the Riccati equation. These have come to be known as *equations of Lie type*. We are now going to describe this class.

Given a Lie algebra homomorphism $\lambda_* \colon \mathfrak{g} \to \mathfrak{X}(M)$ where $\mathfrak{g}$ is the Lie algebra of a Lie group $G$, and a curve $A \colon \mathbb{R} \to \mathfrak{g}$, the ordinary differential equation for a curve $\gamma \colon \mathbb{R} \to M$

$$\gamma'(t) = \lambda_* \big( A(t) \big) \big( \gamma(t) \big)$$

is known as an *equation of Lie type*.

**Example:** THE RICCATI EQUATION. By our previous example, the classical Riccati equation

$$s'(t) = a_0(t) + 2a_1(t)s(t) + a_2(t)\big(s(t)\big)^2$$

is an equation of Lie type for the (local) linear fractional action of $\mathrm{SL}(2,\mathbb{R})$ on $\mathbb{R}$. The curve $A$ is

$$A(t) = \begin{pmatrix} a_1(t) & a_0(t) \\ -a_2(t) & -a_1(t) \end{pmatrix}$$

**Example:** LINEAR EQUATIONS. Every linear equation is an equation of Lie type. Let $G$ be the matrix Lie subgroup of $\mathrm{GL}(n+1,\mathbb{R})$,

$$G = \left\{ \begin{pmatrix} A & B \\ 0 & 1 \end{pmatrix} \;\middle|\; A \in \mathrm{GL}(n,\mathbb{R}) \quad \text{and} \quad B \in \mathbb{R}^n \right\}.$$

Then $G$ acts on $\mathbb{R}^n$ by the standard affine action:

$$\begin{pmatrix} A & B \\ 0 & 1 \end{pmatrix} \cdot x = Ax + B.$$

It is easy to verify that the inhomogeneous linear differential equation

$$x'(t) = a(t)x(t) + b(t)$$

is then a Lie equation, with

$$A(t) = \begin{pmatrix} a(t) & b(t) \\ 0 & 0 \end{pmatrix}.$$

The following proposition follows from the fact that a left action $\lambda \colon G \times M \to M$ relates the right invariant vector field $Y_v$ to the vector field $\lambda_*(v)$ on $M$. Despite its simplicity, it has important consequences.

L.3.9

**Proposition 3:** *If $A\colon \mathbb{R} \to \mathfrak{g}$ is a curve in the Lie algebra of a Lie group $G$ and $S\colon \mathbb{R} \to G$ is the solution to the equation $S'(t) = Y_{A(t)}(S(t))$ with initial condition $S(0) = e$, then on any manifold $M$ endowed with a left $G$-action $\lambda$, the equation of Lie type*

$$\gamma'(t) = \lambda_*\big(A(t)\big)\big(\gamma(t)\big),$$

*with initial condition $\gamma(0) = m$ has, as its solution, $\gamma(t) = S(t) \cdot m$.* $\qquad\square$

The solution $S$ of Proposition 3 is often called the *fundamental solution* of the Lie equation associated to $A(t)$. The most classical example of this is the fundamental solution of a linear system of equations:

$$x'(t) = a(t)x(t)$$

where $a$ is an $n$-by-$n$ matrix of functions of $t$ and $x$ is to be a column of height $n$. In ODE classes, we learn that every solution of this equation is of the form $x(t) = X(t)x_0$ where $X$ is the $n$-by-$n$ matrix of functions of $t$ which solves the equation $X'(t) = a(t)X(t)$ with initial condition $X(0) = I_n$. Of course, this is a special case of Proposition 3 where $\mathrm{GL}(n, \mathbb{R})$ acts on $\mathbb{R}^n$ via the standard left action described in Example 4.

**Lie's Reduction Method.** I now want to explain Lie's method of analysing equations of Lie type. Suppose that $\lambda\colon G \times M \to M$ is a left action and that $A\colon \mathbb{R} \to \mathfrak{g}$ is a smooth curve. Suppose that we have found (by some method) a particular solution $\gamma\colon \mathbb{R} \to M$ of the equation of Lie type associated to $A$ with $\gamma(0) = m$. Select a curve $g\colon \mathbb{R} \to G$ so that $\gamma(t) = g(t) \cdot m$. Of course, this $g$ will not, in general be unique, but any other choice $\tilde{g}$ will be of the form $\tilde{g}(t) = g(t)h(t)$ where $h\colon \mathbb{R} \to G_m$.

I would like to choose $h$ so that $\tilde{g}$ is the fundamental solution of the Lie equation associated to $A$, i.e., so that

$$\tilde{g}'(t) = Y_{A(t)}\big(\tilde{g}(t)\big) = R'_{\tilde{g}(t)}\big(A(t)\big)$$

Unwinding the definitions, it follows that $h$ must satisfy

$$R'_{g(t)h(t)}\big(A(t)\big) = L'_{g(t)}\big(h'(t)\big) + R'_{h(t)}\big(g'(t)\big)$$

so

$$R'_{h(t)}\left(R'_{g(t)}\big(A(t)\big)\right) = L'_{g(t)}\big(h'(t)\big) + R'_{h(t)}\big(g'(t)\big)$$

Solving for $h'(t)$, we find that $h$ must satisfy the differential equation

$$h'(t) = R'_{h(t)}\left(L'_{g(t)^{-1}}\left(R'_{g(t)}\big(A(t)\big) - g'(t)\right)\right).$$

If we set

$$B(t) = L'_{g(t)^{-1}}\left(R'_{g(t)}\big(A(t)\big) - g'(t)\right),$$

L.3.10

then $B$ is clearly computable from $g$ and $A$ and hence may be regarded as known. Since $B = \left(R'_{h(t)}\right)^{-1}(h'(t))$ and since $h$ is a curve in $G_m$, it follows that $B$ must actually be a curve in $\mathfrak{g}_m$.

►      *It follows that the equation*

$$h'(t) = R'_{h(t)}\big(B(t)\big)$$

*is a Lie equation for $h$. In other words in order to find the fundamental solution of a Lie equation for $G$ when the particular solution with initial condition $g(0) = m \in M$ is known, it suffices to solve a Lie equation in $G_m$!*

This observation is known as *Lie's method of reduction*. It shows how knowledge of a particular solution to a Lie equation simplifies the search for the general solution. (Note that this is definitely *not* true of general differential equations.) Of course, Lie's method can be generalized. If one knows $k$ particular solutions with initial values $m_1, \ldots, m_k \in M$, then it is easy to see that one can reduce finding the fundamental solution to finding the fundamental solution of a Lie equation in

$$G_{m_1,\ldots,m_k} = G_{m_1} \cap G_{m_2} \cap \cdots \cap G_{m_k}.$$

If one can arrange that this intersection is discrete, then one can explicitly compute a fundamental solution which will then yield the general solution.

**Example:** THE RICCATI EQUATION AGAIN. Consider the Riccati equation

$$s'(t) = a_0(t) + 2a_1(t)s(t) + a_2(t)\big(s(t)\big)^2$$

and suppose that we know a particular solution $s_0(t)$. Then let

$$g(t) = \begin{pmatrix} 1 & s_0(t) \\ 0 & 1 \end{pmatrix},$$

so that $s_0(t) = g(t) \cdot 0$ (we are using the linear fractional action of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{R}$). The stabilizer of $0$ is the subgroup $G_0$ of matrices of the form:

$$\begin{pmatrix} u & 0 \\ v & u^{-1} \end{pmatrix}.$$

Thus, if we set, as usual,

$$A(t) = \begin{pmatrix} a_1(t) & a_0(t) \\ -a_2(t) & -a_1(t) \end{pmatrix},$$

then the fundamental solution of $S'(t) = A(t)S(t)$ can be written in the form

$$S(t) = g(t)h(t) = \begin{pmatrix} 1 & s_0(t) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u(t) & 0 \\ v(t) & \big(u(t)\big)^{-1} \end{pmatrix}$$

Solving for the matrix $B(t)$ (which we know will have values in the Lie algebra of $G_0$), we find

$$B(t) = \begin{pmatrix} b_1(t) & 0 \\ b_2(t) & -b_1(t) \end{pmatrix} = \begin{pmatrix} a_1(t) + a_2(t)s_0(t) & 0 \\ -a_2(t) & -a_1(t) - a_2(t)s_0(t) \end{pmatrix},$$

and the remaining equation to be solved is

$$h'(t) = B(t)h(t),$$

which is solvable by quadratures in the usual way:

$$u(t) = \exp\left( \int_0^t b_1(\tau)d\tau \right),$$

and, once $u(t)$ has been found,

$$v(t) = \left(u(t)\right)^{-1} \int_0^t b_2(\tau)\left(u(\tau)\right)^2 d\tau.$$

**Example:** LINEAR EQUATIONS AGAIN. Consider the general inhomogeneous $n$-by-$n$ system

$$x'(t) = a(t)x(t) + b(t).$$

Let $G$ be the matrix Lie subgroup of $\mathrm{GL}(n+1, \mathbb{R})$,

$$G = \left\{ \begin{pmatrix} A & B \\ 0 & 1 \end{pmatrix} \middle| A \in \mathrm{GL}(n, \mathbb{R}) \quad \text{and} \quad B \in \mathbb{R}^n \right\}.$$

acting on $\mathbb{R}^n$ by the standard affine action as before. If we embed $\mathbb{R}^n$ into $\mathbb{R}^{n+1}$ by the rule

$$x \mapsto \begin{pmatrix} x \\ 1 \end{pmatrix},$$

then the standard affine action of $G$ on $\mathbb{R}^n$ extends to the standard linear action of $G$ on $\mathbb{R}^{n+1}$. Note that $G$ leaves invariant the subspace $x^{n+1} = 0$, and solutions of the Lie equation corresponding to

$$A(t) = \begin{pmatrix} a(t) & b(t) \\ 0 & 0 \end{pmatrix}$$

which lie in this subspace are simply solutions to the homogeneous equation $x'(t) = a(t)x(t)$. Suppose that we knew a basis for the homogeneous solutions, i.e., the fundamental solution to $X'(t) = a(x)X(t)$ with $X(0) = I_n$. This corresponds to knowing the $n$ particular solutions to the Lie equation on $\mathbb{R}^{n+1}$ which have the initial conditions $e_1, \ldots, e_n$. The simultaneous stabilizer of all of these points in $\mathbb{R}^{n+1}$ is the subgroup $H \subset G$ of matrices of the form

$$\begin{pmatrix} I_n & y \\ 0 & 1 \end{pmatrix}$$

Thus, we choose

$$g(t) = \begin{pmatrix} X(t) & 0 \\ 0 & 1 \end{pmatrix}$$

as our initial guess and look for the fundamental solution in the form:

$$S(t) = g(t)h(t) = \begin{pmatrix} X(t) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I_n & y(t) \\ 0 & 1 \end{pmatrix}.$$

Expanding the condition $S'(t) = A(t)S(t)$ and using the equation $X'(t) = a(t)X(t)$ then reduces us to solving the equation

$$y'(t) = \left( X(t) \right)^{-1} b(t),$$

which is easily solved by integration. The reader will probably recognize that this is precisely the classical method of "variation of parameters".

**Solution by quadrature.** This brings us to an interesting point: Just how hard is it to compute the fundamental solution to a Lie equation of the form

$$\gamma'(t) = R'_{\gamma(t)} \left( A(t) \right)?$$

One case where it is easy is if the Lie group is abelian. We have already seen that if $T$ is a connected abelian Lie group with Lie algebra $\mathfrak{t}$, then the exponential map $\exp\colon \mathfrak{t} \to T$ is a surjective homomorphism. It follows that the fundamental solution of the Lie equation associated to $A\colon \mathbb{R} \to \mathfrak{t}$ is given in the form

$$S(t) = \exp \left( \int_0^t A(\tau)\, d\tau \right)$$

(Exercise: Why is this true?) Thus, the Lie equation for an abelian group is "solvable by quadrature" in the classical sense.

Another instance where one can at least reduce the problem somewhat is when one has a homomorphism $\phi\colon G \to H$ and knows the fundamental solution $S_H$ to the Lie equation for $\varphi \circ A\colon \mathbb{R} \to \mathfrak{h}$. In this case, $S_H$ is the particular solution (with initial condition $S_H(0) = e$) of the Lie equation on $H$ associated to $A$ by regarding $\phi$ as defining a left action on $H$. By Lie's method of reduction, therefore, we are reduced to solving a Lie equation for the group $\ker(\phi) \subset G$.

**Example.** Suppose that $G$ is connected and simply connected. Let $\mathfrak{g}$ be its Lie algebra and let $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$ be the linear subspace generated by all brackets of the form $[x, y]$ where $x$ and $y$ lie in $\mathfrak{g}$. Then, by the Exercises of Lecture 2, we know that $[\mathfrak{g}, \mathfrak{g}]$ is an ideal in $\mathfrak{g}$ (called the *commutator ideal* of $\mathfrak{g}$). Moreover, the quotient algebra $\mathfrak{t} = \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is abelian.

Since $G$ is connected and simply connected, Theorem 3 from Lecture 2 implies that there is a Lie group homomorphism $\phi_0\colon G \to T_0 = \mathfrak{t}$ whose induced Lie algebra homomorphism $\varphi_0\colon \mathfrak{g} \to \mathfrak{t} = \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is just the canonical quotient mapping. From our previous remarks, it follows that any Lie equation for $G$ can be reduced, by one quadrature, to a Lie equation for $G_1 = \ker \phi_0$. It is not difficult to check that the group $G_1$ constructed in this argument is also connected and simply connected.

The desire to iterate this process leads to the following construction: Define the sequence $\{\mathfrak{g}_k\}$ of commutator ideals of $\mathfrak{g}$ by the rules $\mathfrak{g}_0 = \mathfrak{g}$ and and $\mathfrak{g}_{k+1} = [\mathfrak{g}_k, \mathfrak{g}_k]$ for $k \geq 0$. Then we have the following result:

**Proposition 4:** *Let $G$ be a connected and simply connected Lie group for which the sequence $\{\mathfrak{g}_k\}$ of commutator ideals satisfies $\mathfrak{g}_N = (0)$ for some $N > 0$. Then any Lie equation for $G$ can be solved by a sequence of quadratures.*

A Lie algebra with the property described in Proposition 4 is called "solvable". For example, the subalgebra of upper triangular matrices in $\mathfrak{gl}(n, \mathbb{R})$ is solvable, as the reader is invited to check.

While it may seem that solvability is a lot to ask of a Lie algebra, it turns out that this property is surprisingly common. The reader can also check that, of all of the two and three dimensional Lie algebras found in Lecture 2, only $\mathfrak{sl}(2, \mathbb{R})$ and $\mathfrak{so}(3)$ fail to be solvable.

This (partly) explains why the Riccati equation holds such an important place in the theory of ODE. In some sense, it is the first Lie equation which cannot be solved by quadratures. (See the exercises for an interpretation and "proof" of this statement.)

In any case, the sequence of subalgebras $\{\mathfrak{g}_k\}$ eventually stabilizes at a subalgebra $\mathfrak{g}_N$ whose Lie algebra satisfies $[\mathfrak{g}_N, \mathfrak{g}_N] = \mathfrak{g}_N$. A Lie algebra $\mathfrak{g}$ for which $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$ is called "perfect". Our analysis of Lie equations shows that, by Lie's reduction method, we can, by quadrature alone, reduce the problem of solving Lie equations to the problem of solving Lie equations associated to Lie groups with perfect algebras. Further analysis of the relation between the structure of a Lie algebra and the solvability by quadratures of any associated Lie equation leads to the development of the so-called Jordan-Hölder decomposition theorems, see [?].

# Appendix: Lie's Transformation Groups, I

When Lie began his study of symmetry groups in the nineteenth century, the modern concepts of manifold theory were not available. Thus, the examples that he had to guide him were defined as "transformations in $n$ variables" which were often, like the Möbius transformations on the line or like conformal transformations in space, only defined "almost everywhere". Thus, at first glance, it might appear that Lie's concept of a "continuous transformation group" should correspond to what we have defined as a local Lie group action.

However, it turns out that Lie had in mind a much more general concept. For Lie, a set $\Gamma$ of local diffeomorphisms in $\mathbb{R}^n$ formed a "continuous transformation group" if it was closed under composition and inverse and moreover, the elements of $\Gamma$ were characterized as the solutions of some system of differential equations.

For example, the Möbius group on the line could be characterized as the set $\Gamma$ of (non-constant) solutions $f(x)$ of the differential equation

$$2f'''(x)f'(x) - 3\big(f''(x)\big)^2 = 0.$$

As another example, the "group" of area preserving transformations of the plane could be characterized as the set of solutions $\big(f(x,y), g(x,y)\big)$ to the equation

$$f_x\, g_y - g_x\, f_y \equiv 1,$$

while the "group" of holomorphic transformations of the plane

$$\mathbb{R}^2$$

(regarded as $\mathbb{C}$) was the set of solutions $\big(f(x,y), g(x,y)\big)$ to the equations

$$f_x - g_y = f_y + g_x = 0.$$

Notice a big difference between the first example and the other two. In the first example, there is only a 3-parameter family of local solutions and each of these solutions patches together on $\mathbb{RP}^1 = \mathbb{R} \cup \{\infty\}$ to become an element of the global Lie group action of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{RP}^1$. In the other two examples, there are many local solutions that cannot be extended to the entire plane, much less any "completion". Moreover in the volume preserving example, it is clear that no finite dimensional Lie group could ever contain all of the globally defined volume preserving transformations of the plane.

Lie regarded these latter two examples as "infinite continuous groups". Nowadays, we would call them "infinite dimensional pseudo-groups". I will say more about this point of view in an appendix to Lecture 6.

Since Lie did not have a group manifold to work with, he did not regard his "infinite groups" as pathological. Instead of trying to find a global description of the groups, he worked with what he called the "infinitesimal transformations" of $\Gamma$. We would say that,

for each of his groups $\Gamma$, he considered the space of vector fields $\gamma \subset \mathfrak{X}(\mathbb{R}^n)$ whose (local) flows were 1-parameter "subgroups" of $\Gamma$. For example, the infinitesimal transformations associated to the area preserving transformations are the vector fields

$$ X = f(x,y)\frac{\partial}{\partial x} + g(x,y)\frac{\partial}{\partial y} $$

which are divergence free, i.e., satisfy $f_x + g_y = 0$.

Lie "showed" that for any "continuous transformation group" $\Gamma$, the associated set of vector fields $\gamma$ was actually closed under addition, scalar multiplication (by constants), and, most significantly, the Lie bracket. (The reason for the quotes around "showed" is that Lie was not careful to specify the nature of the differential equations which he was using to define his groups. Without adding some sort of constant rank or non-degeneracy hypotheses, many of his proofs are incorrect.)

For Lie, every subalgebra $L$ of the algebra $\mathfrak{X}(\mathbb{R}^n)$ which could be characterized by some system of PDE was to be regarded the Lie algebra of some Lie group. Thus, rather than classify actual groups (which might not really be groups because of domain problems), Lie classified subalgebras of the algebra of vector fields.

In the case that $L$ was finite dimensional, Lie actually proved that there *was* a "germ" of a Lie group (in our sense) and a local Lie group action which generated this algebra of vector fields. This is Lie's so-called Third Fundamental Theorem.

The case where $L$ was infinite dimensional remained rather intractable. I will have more to say about this in Lecture 6. For now, though, I want to stress that there *is* a sort of analogue of actions for these "infinite dimensional Lie groups".

For example, if $M$ is a manifold and $\mathrm{Diff}(M)$ is the group of (global) diffeomorphisms, then we can regard the natural (evaluation) map $\lambda\colon \mathrm{Diff}(M) \times M \to M$ given by $\lambda(\phi, m) = \phi(m)$ as a faithful Lie group action. If $M$ is compact, then every vector field is complete, so, at least formally, the induced map $\lambda_*\colon T_{id}\mathrm{Diff}(M) \to \mathfrak{X}(M)$ ought to be an isomorphism of vector spaces. If our analogy with the finite dimensional case is to hold up, $\lambda_*$ must reverse the Lie bracket.

Of course, since we have not defined a smooth structure on $\mathrm{Diff}(M)$, it is not immediately clear how to make sense of $T_{id}\mathrm{Diff}(M)$. I will prefer to proceed formally and simply define the Lie algebra $\mathtt{diff}(M)$ of $\mathrm{Diff}(M)$ to be the vector space $\mathfrak{X}(M)$ with the Lie algebra bracket given by the negative of the vector field Lie bracket.

With this definition, it follows that a left action $\lambda\colon G \times M \to M$ where $G$ is finite dimensional can simply be regarded as a homomorphism $\Lambda\colon G \to \mathrm{Diff}(M)$ inducing a homomorphism of Lie algebras.

A modern treatment of this subject can be found in [SS].

# Appendix: Connections and Curvature

In this appendix, I want briefly to describe the notions of connections and curvature on principal bundles in the language that I will be using them in the examples in this Lecture.

Let $G$ be a Lie group with Lie algebra $\mathfrak{g}$ and let $\omega_G$ be the canonical $\mathfrak{g}$-valued, left-invariant 1-form on $G$.

**Principal Bundles.** Let $M$ be an $n$-manifold and let $P$ be a principal right $G$-bundle over $M$. Thus, $P$ comes equipped with a submersion $\pi \colon P \to M$ and a free right action $\rho \colon P \times G \to P$ so that the fibers of $\pi$ are the $G$-orbits of $\rho$.

**The Gauge Group.** The group $\mathrm{Aut}(P)$ of automorphisms of $P$ is, by definition, the set of diffeomorphisms $\phi \colon P \to P$ which are compatible with the two structure maps, i.e.,

$$\pi \circ \phi = \pi \qquad \text{and} \qquad \rho_g \circ \phi = \phi \circ \rho_g \ \text{ for all } g \in G.$$

For reasons having to do with Physics, this group is nowadays referred to as the *gauge group* of $P$. Of course, $\mathrm{Aut}(P)$ is not a finite dimensional Lie group, but it would have been considered by Lie himself as a perfectly reasonable "continuous transformation group" (although not a very interesting one for his purposes).

For any $\phi \in \mathrm{Aut}(P)$, there is a unique smooth map $\varphi \colon P \to G$ which satisfies $\phi(p) = p \cdot \varphi(p)$. The identity $\rho_g \circ \phi = \phi \circ \rho_g$ implies that $\varphi$ satisfies $\varphi(p \cdot g) = g^{-1}\varphi(p)g$ for all $g \in G$. Conversely, any smooth map $\varphi \colon P \to G$ satisfying this identity defines an element of $\mathrm{Aut}(P)$. It follows that $\mathrm{Aut}(P)$ is the space of sections of the bundle $C(P) = P \times_C G$ where $C \colon G \times G \to G$ is the conjugation action $C(a,b) = aba^{-1}$.

Moreover, it easily follows that the set of vector fields on $P$ whose flows generate 1-parameter subgroups of $\mathrm{Aut}(P)$ is identifiable with the space of sections of the vector bundle $\mathrm{Ad}(P) = P \times_{\mathrm{Ad}} \mathfrak{g}$.

**Connections.** Let $\mathfrak{A}(P)$ denote the space of *connections* on $P$. Thus, an element $A \in \mathfrak{A}(P)$ is, by definition, a $\mathfrak{g}$-valued 1-form $A$ on $P$ with the following two properties:

(1) For any $p \in P$, we have $\iota_p^*(A) = \omega_G$ where $\iota_p \colon G \to P$ is given by $\iota_p(g) = p \cdot g$.

(2) For all $g$ in $G$, we have $\rho_g^*(A) = \mathrm{Ad}(g^{-1})(A)$ where $\rho_g \colon P \to P$ is right action by $g$.

It follows from Property 1 that, for any connection $A$ on $P$, we have $A\big(\rho_*(x)\big) = x$ for all $x \in \mathfrak{g}$. It follows from Property 2 that $\mathfrak{L}_{\rho_*(x)}A = -[x, A]$ for all $x \in \mathfrak{g}$.

If $A_0$ and $A_1$ are connections on $P$, then it follows from Property 1 that the difference $\alpha = A_1 - A_0$ is a $\mathfrak{g}$-valued 1-form which is "semi-basic" in the sense that $\alpha(v) = 0$ for all $v \in \ker \pi'$. Moreover, Property 2 implies that $\alpha$ satisfies $\rho_g^*(\alpha) = \mathrm{Ad}(g^{-1})(\alpha)$. Conversely, if $\alpha$ is any $\mathfrak{g}$-valued 1-form on $P$ satisfying these latter two properties and $A \in \mathfrak{A}(P)$ is a connection, then $A + \alpha$ is also a connection. It is easy to see that a 1-form $\alpha$ with these two properties can be regarded as a 1-form on $M$ with values in $\mathrm{Ad}(P)$.

▶ Thus, $\mathfrak{A}(P)$ is an affine space modeled on the vector space $\mathcal{A}^1\big(\mathrm{Ad}(P)\big)$. In particular, if we regard $\mathfrak{A}(P)$ as an "infinite dimensional manifold", the tangent space $T_A\mathfrak{A}(P)$ at any point $A$ is naturally isomorphic to $\mathcal{A}^1\big(\mathrm{Ad}(P)\big)$.

L.3.17

**Curvature.** The *curvature* of a connection $A$ is the 2-form $F_A = dA + \frac{1}{2}[A, A]$. From our formulas above, it follows that

$$\rho_*(x) \lrcorner F_A = \rho_*(x) \lrcorner dA + [x, A] = \mathfrak{L}_{\rho_*(x)} A + [x, A] = 0.$$

Since the vector fields $\rho_*(x)$ span the vertical tangent spaces of $P$, it follows that $F_A$ is a "semi-basic" 2-form (with values in $\mathfrak{g}$). Moreover, the Ad-equivariance of $A$ implies that $\rho_g^*(F_A) = \mathrm{Ad}(g^{-1})(F_A)$. Thus, $F_A$ may be regarded as a section of the bundle of 2-forms on $M$ with values in the bundle $\mathrm{Ad}(P)$.

The group $\mathsf{Aut}(P)$ acts naturally on the *right* on $\mathfrak{A}(P)$ via pullback: $A \cdot \phi = \phi^*(A)$. In terms of the corresponding map $\varphi\colon P \to G$, we have

$$A \cdot \phi = \varphi^*(\omega_G) + \mathrm{Ad}(\varphi^{-1})(A).$$

It follows by direct computation that $F_{A \cdot \phi} = \phi^*(F_A) = \mathrm{Ad}(\varphi^{-1})(F_A)$.

We say that $A$ is *flat* if $F_A = 0$. It is an elementary ODE result that $A$ is flat if and only if, for every $m \in M$, there exists an open neighborhood $U$ of $m$ and a smooth map $\tau\colon \pi^{-1}(U) \to G$ which satisfies $\tau(p \cdot g) = \tau(p)g$ and $\tau^*(\omega_G) = A_{|U}$. In other words $A$ is flat if and only if the bundle-with-connection $(P, A)$ is locally diffeomorphic to the trivial bundle-with-connection $(M \times G, \omega_G)$.

**Covariant Differentiation.** The space $\mathcal{A}^p(\mathrm{Ad}(P))$ of $p$-forms on $M$ with values in $\mathrm{Ad}(P)$ can be identified with the space of $\mathfrak{g}$-valued, $p$-forms $\beta$ on $P$ which are both semi-basic and Ad-equivariant (i.e., $\rho_g^*(\beta) = \mathrm{Ad}(g^{-1})(\beta)$ for all $g \in G$). Given such a form $\beta$, the expression $d\beta + [A, \beta]$ is easily seen to be a $\mathfrak{g}$-valued $(p+1)$-form on $P$ which is also semi-basic and Ad-equivariant. It follows that this defines a first-order differential operator

$$d_A\colon \mathcal{A}^p(\mathrm{Ad}(P)) \to \mathcal{A}^{p+1}(\mathrm{Ad}(P))$$

called *covariant differentiation with respect to $A$*. It is elementary to check that

$$d_A(d_A\beta) = [F_A, \beta] = \mathrm{ad}(F_A)(\beta).$$

Thus, for a flat connection, $(\mathcal{A}^*(\mathrm{Ad}(P)), d_A)$ forms a complex over $M$.

We also have the *Bianchi identity $d_A F_A = 0$*.

For some, "covariant differentiation" means only $d_A\colon \mathcal{A}^0(\mathrm{Ad}(P)) \to \mathcal{A}^1(\mathrm{Ad}(P))$.

**Horizontal Lifts and Holonomy.** Let $A$ be a connection on $P$. If $\gamma\colon [0,1] \to M$ is a $C^1$ curve and $p \in \pi^{-1}(\gamma(0))$ is chosen, then there exists a unique $C^1$ curve $\tilde{\gamma}\colon [0,1] \to P$ which both "lifts" $\gamma$ in the sense that $\gamma = \pi \circ \tilde{\gamma}$ and also satisfies the differential equation $\tilde{\gamma}^*(A) = 0$.

(To see this, first choose any lift $\bar{\gamma}\colon [0,1] \to P$ which satisfies $\bar{\gamma}(0) = p$. Then the desired lifting will then be given by $\tilde{\gamma}(t) = \bar{\gamma}(t) \cdot g(t)$ where $g\colon [0,1] \to G$ is the solution of the Lie equation $g'(t) = -R_{g(t)}(A(\bar{\gamma}'(t)))$ satisfying the initial condition $g(0) = e$.)

The resulting curve $\tilde{\gamma}$ is called a *horizontal lift* of $\gamma$. If $\gamma$ is merely piecewise $C^1$, the horizontal lift can still be defined by piecing together horizontal lifts of the $C^1$-segments in the obvious way. Also, if $p' = p \cdot g_0$, then the horizontal lift of $\gamma$ with initial condition $p'$ is easily seen to be $\rho_{g_0} \circ \tilde{\gamma}$.

Let $p \in P$ be chosen and set $m = \pi(p)$. For every piecewise $C^1$-loop $\gamma \colon [0,1] \to M$ based at $m$, the horizontal lift $\tilde{\gamma}$ has the property that $\tilde{\gamma}(1) = p \cdot h(\gamma)$ for some unique $h(\gamma) \in G$. The *holonomy* of $A$ at $p$, denoted by $H_A(p)$ is, by definition, the set of all such elements $h(\gamma)$ of $G$ where $\gamma$ ranges over all of the piecewise $C^1$ closed loops based at $m$.

I leave it to the reader to show that $H_A(p \cdot g) = g^{-1} H_A(p) g$ and that, if $p$ and $p'$ can be joined by a horizontal curve in $P$, then $H_A(p) = H_A(p')$. Thus, the conjugacy class of $H_A(p)$ in $G$ is independent of $p$ if $M$ is connected.

A basic theorem due to Borel and Lichnerowitz (see [KN]) asserts that $H_A(p)$ is always a Lie subgroup of $G$.

# Exercise Set 3:

# Actions of Lie Groups

**1**. Verify the claim made in the lecture that every right (respectively, left) action of a Lie group on a manifold can be rewritten as a left (respectively, right) action. Is the assumption that a left action $\lambda \colon G \times M \to M$ satisfy $\lambda(e, m) = m$ for all $m \in M$ really necessary?

**2**. Show that if $f \colon X \to Y$ is a map of smooth manifolds for which the rank of $f'(x) \colon T_x X \to T_{f(x)} Y$ is independent of $x$, then $f^{-1}(y)$ is a (possibly empty) closed, smooth submanifold of $X$ for all $y \in Y$. Note that this properly generalizes the usual Implicit Function Theorem, which requires $f'(x)$ to be a surjection everywhere in order to conclude that $f^{-1}(y)$ is a smooth submanifold.

 (Hint: Suppose that the rank of $f'(x)$ is identically $k$. You want to show that $f^{-1}(y)$ (if non-empty) is a submanifold of $X$ of codimension $k$. To do this, let $x \in f^{-1}(y)$ be given and construct a map $\psi \colon V \to \mathbb{R}^k$ on a neighborhood $V$ of $y$ so that $\psi \circ f$ is a submersion near $x$. Then show that $(\psi \circ f)^{-1}\big(\psi(y)\big)$ (which, by the Implicit Function Theorem, is a closed codimension $k$ submanifold of the open set $f^{-1}(V) \subset X$) is actually equal to $f^{-1}(y)$ on some neighborhood of $x$. Where do you need the constant rank hypothesis?)

**3**. This exercise concerns the automorphism groups of Lie algebras and Lie groups.

 (i) Show that, for any Lie algebra $\mathfrak{g}$, the group of automorphisms $\mathrm{Aut}(\mathfrak{g})$ defined by

$$\mathrm{Aut}(\mathfrak{g}) = \{a \in \mathrm{End}(\mathfrak{g}) \,|\, \big[a(x), a(y)\big] = a\big([x, y]\big) \quad \text{for all } x, y \in \mathfrak{g}\}$$

 is a closed Lie subgroup of $\mathrm{GL}(\mathfrak{g})$. Show that its Lie algebra is

$$\mathfrak{der}(\mathfrak{g}) = \{a \in \mathrm{End}(\mathfrak{g}) \,|\, a\big([x, y]\big) = \big[a(x), y\big] + \big[x, a(y)\big] \quad \text{for all } x, y \in \mathfrak{g}\}.$$

 (Hint: Show that $\mathrm{Aut}(\mathfrak{g})$ is the stabilizer of some point in some representation of the Lie group $\mathrm{GL}(\mathfrak{g})$.)

 (ii) Show that if $G$ is a connected and simply connected Lie group with Lie algebra $\mathfrak{g}$, then the group of (Lie) automorphisms of $G$ is isomorphic to $\mathrm{Aut}(\mathfrak{g})$.

 (iii) Show that ad: $\mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ actually has its image in $\mathfrak{der}(\mathfrak{g})$, and that this image is an ideal in $\mathfrak{der}(\mathfrak{g})$. What is the interpretation of this fact in terms of "inner" and "outer" automorphisms of $G$? (Hint: Use the Jacobi identity.)

 (iv) Show that if the Killing form of $\mathfrak{g}$ is non-degenerate, then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. (Hint: Suppose that $[\mathfrak{g}, \mathfrak{g}]$ lies in a proper subspace of $\mathfrak{g}$. Then there exists an element $y \in \mathfrak{g}$ so that $\kappa\big([x, z], y\big) = 0$ for all $x, z \in \mathfrak{g}$. Show that this implies that $[x, y] = 0$ for all $x \in \mathfrak{g}$, and hence that $\mathrm{ad}(y) = 0$.)

 (v) Show that if the Killing form of $\mathfrak{g}$ is non-degenerate, then $\mathfrak{der}(\mathfrak{g}) = \mathrm{ad}(\mathfrak{g})$. This shows that all of the automorphisms of a simple Lie algebra are "inner". (Hint: Show that

the set $\mathfrak{p} = \{a \in \mathfrak{der}(\mathfrak{g}) \mid \operatorname{tr}(a \operatorname{ad}(x)) = 0 \text{ for all } x \in \mathfrak{g}\}$ is also an ideal in $\mathfrak{der}(\mathfrak{g})$ and hence that $\mathfrak{der}(\mathfrak{g}) = \mathfrak{p} \oplus \operatorname{ad}(\mathfrak{g})$ as algebras. Show that this forces $\mathfrak{p} = 0$ by considering what it means for elements of $\mathfrak{p}$ (which, after all, are derivations of $\mathfrak{g}$) to commute with elements in $\operatorname{ad}(\mathfrak{g})$.)

**4**. Consider the 1-parameter group which is generated by the flow of the vector field $X$ in the plane

$$X = \cos y \, \frac{\partial}{\partial x} + \sin^2 y \, \frac{\partial}{\partial y}.$$

Show that this vector field is complete and hence yields a free $\mathbb{R}$-action on the plane. Let $\mathbb{Z}$ also act on the plane by the action

$$m \cdot (x, y) = ((-1)^m x, y + m\pi).$$

Show that these two actions commute, and hence together define a free action of $G = \mathbb{R} \times \mathbb{Z}$ on the plane. Sketch the orbits and show that, even though the $G$-orbits of this action are closed, and the quotient space is Hausdorff, the quotient space is not a manifold. (The point of this problem is to warn the student not to make the common mistake of thinking that the quotient of a manifold by a free Lie group action is a manifold if it is Hausdorff.)

**5**. Show that if $\rho \colon M \times G \to M$ is a right action, then the induced map $\rho_* \colon \mathfrak{g} \to \mathfrak{X}(M)$ satisfies $\rho_*([x, y]) = [\rho_*(x), \rho_*(y)]$.

**6**. Prove Proposition 2. (Hint: you are trying to find an open neighborhood $U$ of $\{e\} \times M$ in $G \times M$ and a smooth map $\lambda \colon U \to M$ with the requisite properties. To do this, look for the graph of $\lambda$ as a submanifold $\Gamma \subset G \times M \times M$ which contains all the points $(e, m, m)$ and is tangent to a certain family of vector fields on $G \times M \times M$ constructed using the left invariant vector fields on $G$ and the corresponding vector fields on $M$ determined by the Lie algebra homomorphism $\phi \colon \mathfrak{g} \to \mathfrak{X}(M)$.)

**7**. Show that, if $A \colon \mathbb{R} \to \mathfrak{g}$ is a curve in the Lie algebra of a Lie group $G$, then there exists a unique solution to the ordinary differential equation $S'(t) = R_{S(t)}(A(t))$ with initial condition $S(0) = e$. (It is clear that a solution exists on some interval $(-\varepsilon, \varepsilon)$ in $\mathbb{R}$. The problem is to show that the solution exists on all of $\mathbb{R}$.)

**8**. Show that, under the action of $\operatorname{GL}(n, \mathbb{R})$ on the space of symmetric $n$-by-$n$ matrices defined in the Lecture, every symmetric $n$-by-$n$ matrix is in the orbit of an $I_{p,q}$.

**9**. This problem examines the geometry of the classical second order equation for one unknown.
  (i) Rewrite the second-order ODE

$$\frac{d^2 x}{dt^2} = F(t)x$$

as a system of first-order ODEs of Lie type for an action of $\operatorname{SL}(2, \mathbb{R})$ on $\mathbb{R}^2$.

E.3.2

(ii) Suppose in particular that $F(t)$ is of the form $\left(f(t)\right)^2 + f'(t)$, where $f(0) \neq 0$. Use the solution

$$x(t) = \exp\left(\int_0^t f(\tau)\,d\tau\right)$$

to write down the fundamental solution for this Lie equation up in $\mathrm{SL}(2, \mathbb{R})$.

(iii) Explain why the (more general) second order linear ODE

$$x'' = a(t)x' + b(t)x$$

is solvable by quadratures once we know a single solution with either $x(0) \neq 0$ or $x'(0) \neq 0$. (Hint: all two-dimensional Lie groups are solvable.)

**10\***. Show that the general equation of the form $y''(x) = f(x)\,y(x)$ is not integrable by quadratures. Specifically, show that there do not exist "universal" functions $F_0$ and $F_1$ of two and three variables respectively so that the function $y$ defined by taking the most general solution of

$$u'(x) = F_0\big(x, f(x)\big)$$
$$y'(x) = F_1\big(x, f(x), u(x)\big)$$

is the general solution of $y''(x) = f(x)\,y(x)$. Note that this shows that the general solution cannot be got by two quadratures, which one might expect to need since the general solution must involve two constants of integration. However, it can be shown that no matter how many quadratures one uses, one cannot get even a particular solution of $y''(x) = f(x)\,y(x)$ (other than the trivial solution $y \equiv 0$) by quadrature. (If one could get a (non-trivial) particular solution this way, then, by two more quadratures, one could get the general solution.)

**11**. The point of this exercise is to prove Lie's theorem (stated below) on (local) group actions on $\mathbb{R}$. This theorem "explains" the importance of the Riccati equation, and why there are so few actions of Lie groups on $\mathbb{R}$. Let $\mathfrak{g} \subset \mathfrak{X}(\mathbb{R})$ be a *finite dimensional* Lie algebra of vector fields on $\mathbb{R}$ with the property that, at every $x \in \mathbb{R}$, there is at least one $X \in \mathfrak{g}$ so that $X(x) \neq 0$. (Thus, the (local) flows of the vector fields in $\mathfrak{g}$ do not have any common fixed point.)

(i) For each $x \in \mathbb{R}$, let $\mathfrak{g}_x^k \subset \mathfrak{g}$ denote the subspace of vector fields which vanish to order at least $k+1$ at $x$. (Thus, $\mathfrak{g}_x^{-1} = \mathfrak{g}$ for all $x$.) Let $\mathfrak{g}_x^\infty \subset \mathfrak{g}$ denote the intersection of all the $\mathfrak{g}_x^k$. Show that $\mathfrak{g}_x^\infty = 0$ for all $x$. (Hint: Fix $a \in \mathbb{R}$ and choose an $X \in \mathfrak{g}$ so that $X(a) \neq 0$. Make a local change of coordinates near $a$ so that $X = \partial/\partial x$ on a neighborhood of $a$. Note that $[X, \mathfrak{g}_a^\infty] \subset \mathfrak{g}_a^\infty$. Now choose a basis $Y_1, \ldots, Y_N$ of $\mathfrak{g}_a^\infty$ and note that, near $a$, we have $Y_i = f_i\,\partial/\partial x$ for some functions $f_i$. Show that the $f_i$ must satisfy some differential equations and then apply ODE uniqueness. Now go on from there.)

---

\* This exercise is somewhat difficult, but you should enjoy seeing what is involved in trying to prove that an equation is not solvable by quadratures.

(ii) Show that the dimension of $\mathfrak{g}$ is at most 3. (Hint: First, show that $[\mathfrak{g}_x^j, \mathfrak{g}_x^k] = \subset \mathfrak{g}_x^{j+k}$. Now, by part (i), you know that there is a smallest integer $N$ (which may depend on $x$) so that $\mathfrak{g}_x^{N+1} = 0$. Show that if $X \in \mathfrak{g}$ does not vanish at $x$ and $Y_N \in \mathfrak{g}$ vanishes to exactly order $N$ at $x$, then $Y_{N-1} = [X, Y_N]$ vanishes to order exactly $N-1$. Conclude that the vectors $X, Y_0, \ldots, Y_N$ (where $Y_{i-1} = [X, Y_i]$ for $i > 0$) form a basis of $\mathfrak{g}$. Now, what do you know about $[Y_{N-1}, Y_N]$?).

(iii) (Lie's Theorem) Show that, if $\dim(\mathfrak{g}) = 2$, then $\mathfrak{g}$ is isomorphic to the (unique) non-abelian Lie algebra of that dimension and that there is a local change of coordinates so that

$$\mathfrak{g} = \{(a + bx)\partial/\partial x \,|\, a, b \in \mathbb{R}\}.$$

Show also that, if $\dim(\mathfrak{g}) = 3$, then $\mathfrak{g}$ is isomorphic to $\mathfrak{sl}(2, \mathbb{R})$ and that there exist local changes of coordinates so that

$$\mathfrak{g} = \{(a + bx + cx^2)\partial/\partial x \,|\, a, b, c \in \mathbb{R}\}.$$

(In the second case, after you have shown that the algebra is isomorphic to $\mathfrak{sl}(2, \mathbb{R})$, show that, at each point of $\mathbb{R}$, there exists a element $X \in \mathfrak{g}$ which does not vanish at the point and which satisfies $(\mathrm{ad}(X))^2 = 0$. Now put it in the form $X = \partial/\partial x$ for some local coordinate $x$ and ask what happens to the other elements of $\mathfrak{g}$.)

(iv) (This is somewhat harder.) Show that if $\dim(\mathfrak{g}) = 3$, then there is a diffeomorphism of $\mathbb{R}$ with an open interval $I \subset \mathbb{R}$ so that $\mathfrak{g}$ gets mapped to the algebra

$$\mathfrak{g} = \left\{ (a + b\cos x + c\sin x)\frac{\partial}{\partial x} \,\bigg|\, a, b, c \in \mathbb{R} \right\}.$$

In particular, this shows that every local action of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{R}$ is the restriction of the Möbius action on $\mathbb{RP}^1$ after "lifting" to its universal cover. Show that two intervals $I_1 = (0, a)$ and $I_2 = (0, b)$ are diffeomorphic in such a way as to preserve the Lie algebra $\mathfrak{g}$ if and only if either $a = b = 2n\pi$ for some positive integer $n$ or else $2n\pi < a, b < (2n + 2)\pi$ for some positive integer $n$. (Hint: Show, by a local analysis, that any vector field $X \in \mathfrak{g}$ which vanishes at any point of $\mathbb{R}$ must have $\kappa(X, X) \geq 0$. Now choose an $X$ so that $\kappa(X, X) = -2$ and choose a global coodinate $x: \mathbb{R} \to \mathbb{R}$ so that $X = \partial/\partial x$. You must still examine the effect of your choices on the image interval $x(\mathbb{R}) \subset \mathbb{R}$.)

Lie and his coworkers attempted to classify all of the finite dimensional Lie subalgebras of the vector fields on $\mathbb{R}^k$, for $k \leq 5$, since (they thought) this would give a classification of all of the equations of Lie type for at most 5 unknowns. The classification became extremely complex and lengthy by dimension 5 and it was abandoned. On the other hand, the project of classifying the *abstract* finite dimensional Lie algebras has enjoyed a great deal of success. In fact, one of the triumphs of nineteenth century mathematics was the classification, by Killing and Cartan, of all of the finite dimensional simple Lie algebras over $\mathbb{C}$ and $\mathbb{R}$.

# Lecture 4:

# Symmetries and Conservation Laws

**Variational Problems.** In this Lecture, I will introduce a particular set of variational problems, the so-called "first-order particle Lagrangian problems", which will serve as a link to the "symplectic" geometry to be developed in the next Lecture.

**Definition 1:** A *Lagrangian* on a manifold $M$ is a smooth function $L\colon TM \to \mathbb{R}$. For any smooth curve $\gamma\colon [a, b] \to M$, define

$$\mathcal{F}_L(\gamma) = \int_a^b L\big(\dot{\gamma}(t)\big)\, dt.$$

$\mathcal{F}_L$ is called the *functional* associated to $L$.

(The use of the word "functional" here is classical. The reader is supposed to think of the set of all smooth curves $\gamma\colon [a, b] \to M$ as a sort of infinite dimensional manifold and of $\mathcal{F}_L$ as a function on it.)

I have deliberately chosen to avoid the (mild) complications caused by allowing less smoothness for $L$ and $\gamma$, though for some purposes, it is essential to do so. The geometric points that I want to make, however will be clearest if we do not have to worry about determining the optimum regularity assumptions.

Also, some sources only require $L$ to be defined on some open set in $TM$. Others allow $L$ to "depend on $t$", i.e., take $L$ to be a function on $\mathbb{R} \times TM$. Though I will not go into any of these (slight) extensions, the reader should be aware that they exist. For example, see [A].

**Example:** Suppose that $L\colon TM \to \mathbb{R}$ restricts to each $T_xM$ to be a positive definite quadratic form. Then $L$ defines what is usually called a Riemannian metric on $M$. For a curve $\gamma$ in $M$, the functional $\mathcal{F}_L(\gamma)$ is then twice what is usually called the "action" of $\gamma$. This example is, by far, the most commonly occurring Lagrangian in differential geometry. We will have more to say about this below.

For a Lagrangian $L$, one is usually interested in finding the curves $\gamma\colon [a, b] \to M$ with given "endpoint conditions" $\gamma(a) = p$ and $\gamma(b) = q$ for which the functional $\mathcal{F}_L(\gamma)$ is a minimum. For example, in the case where $L$ defines a Riemannian metric on $M$, the curves with fixed endpoints of minimum "action" turn out also to be the *shortest* curves joining those endpoints. From calculus, we know that the way to find minima of a function on a manifold is to first find the "critical points" of the function and then look among those for the minima. As mentioned before, the set of curves in $M$ can be thought of as a sort of "infinite dimensional" manifold, but I won't go into details on this point. What I will do instead is describe what ought to be the set of "curves" in this space (classically called "variations") if it were a manifold.

L.4.1

Given a curve $\gamma\colon[a,b]\to M$, a *(smooth) variation of $\gamma$ with fixed endpoints* is, by definition, a smooth map

$$\Gamma\colon[a,b]\times(-\varepsilon,\varepsilon)\to M$$

for some $\varepsilon>0$ with the property that $\Gamma(t,0)=\gamma(t)$ for all $t\in[a,b]$ and that $\Gamma(a,s)=\gamma(a)$ and $\Gamma(b,s)=\gamma(b)$ for all $s\in(-\varepsilon,\varepsilon)$.

▶      *In this lecture, "variation" will always mean "smooth variation with fixed endpoints".*

If $L$ is a Lagrangian on $M$ and $\Gamma$ is a variation of $\gamma\colon[a,b]\to M$, then we can define a function $\mathcal{F}_{L,\Gamma}\colon(-\varepsilon,\varepsilon)\to\mathbb{R}$ by setting

$$\mathcal{F}_{L,\Gamma}(s)=\mathcal{F}_L(\gamma_s)$$

where $\gamma_s(t)=\Gamma(t,s)$.

**Definition 2:** A curve $\gamma\colon[a,b]\to M$ is *L-critical* if $\mathcal{F}'_{L,\Gamma}(0)=0$ for all variations of $\gamma$.

It is clear from calculus that a curve which minimizes $\mathcal{F}_L$ among all curves with the same endpoints will have to be $L$-critical, so the search for minimizers usually begins with the search for the critical curves.

**Canonical Coordinates.** I want to examine what the problem of finding $L$-critical curves "looks like" in local coordinates. If $U\subset M$ is an open set on which there exists a coordinate chart $x\colon U\to\mathbb{R}^n$, then there is a canonical extension of these coordinates to a coordinate chart $(x,p)\colon TU\to\mathbb{R}^n\times\mathbb{R}^n$ with the property that, for any curve $\gamma\colon[a,b]\to U$, with coordinates $y=x\circ\gamma$, the $p$-coordinates of the curve $\dot\gamma\colon[a,b]\to TU$ are given by $p\circ\dot\gamma=\dot y$. We shall call the coordinates $(x,p)$ on $TU$, the *canonical coordinates* associated to the coordinate system $x$ on $U$.

**The Euler-Lagrange Equations.** In a canonical coordinate system $(x,p)$ on $TU$ where $U$ is an open set in $M$, the function $L$ can be expressed as a function $L(x,p)$ of $x$ and $p$. For a curve $\gamma\colon[a,b]\to M$ which happens to lie in $U$, the functional $\mathcal{F}_L$ becomes simply

$$\mathcal{F}_L(\gamma)=\int_a^b L\big(y(t),\dot y(t)\big)\,dt.$$

I will now derive the classical conditions for such a $\gamma$ to be $L$-critical: Let $h\colon[a,b]\to\mathbb{R}^n$ be any smooth map which satisfies $h(a)=h(b)=0$. Then, for sufficiently small $\varepsilon$, there is a variation $\Gamma$ of $\gamma$ which is expressed in $(x,p)$-coordinates as

$$(x,p)\circ\Gamma=(y+sh,\dot y+s\dot h).$$

Then, by the classic integration-by-parts method,

$$\mathcal{F}'_{L,\Gamma}(0) = \frac{d}{ds}\Big|_{s=0} \left( \int_a^b L\big(y(t) + sh(t), \dot{y}(t) + s\dot{h}(t)\big)\, dt \right)$$

$$= \int_a^b \left( \frac{\partial L}{\partial x^k}\big(y(t), \dot{y}(t)\big)h^k(t) + \frac{\partial L}{\partial p^k}\big(y(t), \dot{y}(t)\big)\dot{h}^k(t) \right) dt$$

$$= \int_a^b \left( \frac{\partial L}{\partial x^k}\big(y(t), \dot{y}(t)\big) - \frac{d}{dt}\left( \frac{\partial L}{\partial p^k}\big(y(t), \dot{y}(t)\big) \right) \right) h^k(t)\, dt.$$

This formula is valid for any $h \colon [a, b] \to \mathbb{R}^n$ which vanishes at the endpoints. It follows without difficulty that the curve $\gamma$ is $L$-critical if and only if $y = x \circ \gamma$ satisfies the $n$ differential equations

$$\frac{\partial L}{\partial x^k}\big(y(t), \dot{y}(t)\big) - \frac{d}{dt}\left( \frac{\partial L}{\partial p^k}\big(y(t), \dot{y}(t)\big) \right) = 0, \qquad \text{for } 1 \le k \le n.$$

These are the famous *Euler-Lagrange equations*.

The main drawback of the Euler-Lagrange equations in this form is that they only give necessary and sufficient conditions for a curve to be $L$-critical if it lies in a coordinate neighborhood $U$. It is not hard to show that if $\gamma \colon [a, b] \to M$ is $L$-critical, then its restriction to any subinterval $[a', b'] \subset [a, b]$ is also $L$-critical. In particular, a necessary condition for $\gamma$ to be $L$-critical is that it satisfy the Euler-Lagrange equations on any subcurve which lies in a coordinate system. However, it is *not* clear that these "local conditions" are sufficient.

Another drawback is that, as derived, the equations depend on the choice of coordinates and it is not clear that one's success in solving them might not depend on a clever choice of coordinates.

In what follows, we want to remedy these defects. First, though, here are a couple of examples.

**Example:** RIEMANNIAN METRICS. Consider a Riemannian metric $L \colon TM \to \mathbb{R}$. Then, in local canonical coordinates,

$$L(x, p) = g_{ij}(x)p^i p^j.$$

where $g(x)$ is a positive definite symmetric matrix of functions. (Remember, the summation convention is in force.) In this case, the Euler-Lagrange equations are

$$\frac{\partial g_{ij}}{\partial x^k}\big(y(t)\big)\dot{y}^i(t)\dot{y}^j(t) = \frac{d}{dt}\left( 2g_{kj}\big(y(t)\big)\dot{y}^j(t) \right) = 2\frac{\partial g_{kj}}{\partial x^i}\big(y(t)\big)\dot{y}^i(t)\dot{y}^j(t) + 2g_{kj}\big(y(t)\big)\ddot{y}^j(t).$$

Since the matrix $g(x)$ is invertible for all $x$, these equations can be put in more familiar form by solving for the second derivatives to get

$$\ddot{y}^i = -\Gamma^i_{jk}(y)\dot{y}^j\dot{y}^k$$

L.4.3                                                                                      63

where the functions $\Gamma^i_{jk} = \Gamma^i_{kj}$ are given by the formula so familiar to geometers:

$$\Gamma^i_{jk} = \frac{1}{2} g^{i\ell} \left( \frac{\partial g_{\ell j}}{\partial x^k} + \frac{\partial g_{\ell k}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^\ell} \right)$$

where the matrix $(g^{ij})$ is the inverse of the matrix $(g_{ij})$.

**Example:** ONE-FORMS. Another interesting case is when $L$ is linear on each tangent space, i.e., $L = \omega$ where $\omega$ is a smooth 1-form on $M$. In local canonical coordinates,

$$L = a_i(x) p^i$$

for some functions $a_i$ and the Euler-Lagrange equations become:

$$\frac{\partial a_i}{\partial x^k}(y(t))\dot{y}^i(t) = \frac{d}{dt}\left( a_k(y(t)) \right) = \frac{\partial a_k}{\partial x^i}(y(t))\dot{y}^i(t)$$

or, simply,

$$\left( \frac{\partial a_i}{\partial x^k}(y) - \frac{\partial a_k}{\partial x^i}(y) \right) \dot{y}^i = 0.$$

This last equation should look familiar. Recall that the exterior derivative of $\omega$ has the coordinate expression

$$d\omega = \frac{1}{2} \left( \frac{\partial a_j}{\partial x^i} - \frac{\partial a_i}{\partial x^j} \right) dx^i \wedge dx^j.$$

If $\gamma:[a,b] \to U$ is $\mathcal{F}_\omega$-critical, then for every vector field $v$ along $\gamma$ the Euler-Lagrange equations imply that

$$d\omega\big(\dot{\gamma}(t), v(t)\big) = \frac{1}{2} \left( \frac{\partial a_j}{\partial x^i}(y(t)) - \frac{\partial a_i}{\partial x^j}(y(t)) \right) \dot{y}^i(t) v^j(t) = 0.$$

In other words, $\dot{\gamma}(t) \lrcorner d\omega = 0$. Conversely, if this identity holds, then $\gamma$ is clearly $\omega$-critical. This leads to the following global result:

**Proposition 1:** *A curve $\gamma:[a,b] \to M$ is $\omega$-critical for a 1-form $\omega$ on $M$ if and only if it satisfies the first order differential equation*

$$\dot{\gamma}(t) \lrcorner d\omega = 0.$$

PROOF: A straightforward integration-by-parts on $M$ yields the coordinate-free formula

$$\mathcal{F}'_{\omega,\Gamma}(0) = \int_a^b d\omega\big(\dot{\gamma}(t), \tfrac{\partial \Gamma}{\partial s}(t,0)\big) \, dt$$

where $\Gamma$ is any variation of $\gamma$ and $\frac{\partial \Gamma}{\partial s}$ is the "variation vector field" along $\gamma$. Since this vector field is arbitrary except for being required to vanish at the endpoints, we see that "$d\omega\big(\dot{\gamma}, v\big) = 0$ for all vector fields $v$ along $\gamma$" is the desired condition for $\omega$-criticality.   $\square$

The way is now paved for what will seem like a trivial observation, but, in fact, turns out to be of fundamental importance: It is the "seed" of Noether's Theorem.

**Proposition 2:**  *Suppose that $\omega$ is a 1-form on $M$ and that $X$ is a vector field on $M$ whose (local) flow leaves $\omega$ invariant. Then the function $\omega(X)$ is constant on all $\omega$-critical curves.*

PROOF:  The condition that the flow of $X$ leave $\omega$ invariant is just that $\mathfrak{L}_X(\omega) = 0$. However, by the Cartan formula,

$$0 = \mathfrak{L}_X(\omega) = d(X \lrcorner \omega) + X \lrcorner d\omega,$$

so for any curve $\gamma$ in $M$, we have

$$d\omega\big(\dot\gamma(t), X(\gamma(t))\big) = -d\omega\big(X(\gamma(t)), \dot\gamma(t)\big) = -(X \lrcorner d\omega)\big(\dot\gamma(t)\big) = d(X \lrcorner \omega)\big(\dot\gamma(t)\big)$$

and this last expression is clearly the derivative of the function $X \lrcorner \omega = \omega(X)$ along $\gamma$. Now apply Proposition 1. □

It is worth pausing a moment to think about what Proposition 2 means. The condition that the flow of $X$ leave $\omega$ invariant is essentially saying that the flow of $X$ is a "symmetry" of $\omega$ and hence of the functional $\mathcal{F}_\omega$. What Proposition 2 says is that a certain kind of symmetry of the functional gives rise to a "first integral" (sometimes called "conservation law") of the equation for $\omega$-critical curves. If the function $\omega(X)$ is not a constant function on $M$, then saying that the $\omega$-critical curves lie in its level sets is useful information about these critical curves.

Now, this idea can be applied to the general Lagrangian with symmetries. The only trick is to find the appropriate 1-form on which to evaluate "symmetry" vector fields.

**Proposition 3:**  *For any Lagrangian $L: TM \to \mathbb{R}$, there exist a unique function $E_L$ on $TM$ and a unique 1-form $\omega_L$ on $TM$ which, relative to any local coordinate system $x: U \to \mathbb{R}$, have the expressions*

$$E_L = p^i \frac{\partial L}{\partial p^i} - L \qquad \text{and} \qquad \omega_L = \frac{\partial L}{\partial p^i}\, dx^i.$$

*Moreover, if $\gamma: [a, b] \to M$ is any curve, then $\gamma$ satisfies the Euler-Lagrange equations for $L$ in every local coordinate system if and only if its canonical lift $\dot\gamma: [a, b] \to TM$ satisfies*

$$\ddot\gamma(t) \lrcorner d\omega_L = -dE_L(\dot\gamma(t)).$$

PROOF:  This will mainly be a sequence of applications of the Chain Rule.

There is an invariantly defined vector field $R$ on $TM$ which is simply the radial vector field on each subspace $T_m M$. It is expressed in canonical coordinates as $R = p^i \partial/\partial p^i$. Now, using this vector field, the quantity $E_L$ takes the form

$$E_L = -L + dL(R).$$

L.4.5

Thus, it is clear that $E_L$ is well-defined on $TM$.

Now we check the well-definition of $\omega_L$. If $z\colon U \to \mathbb{R}$ is any other local coordinate system, then $z = F(x)$ for some $F\colon \mathbb{R}^n \to \mathbb{R}^n$. The corresponding canonical coordinates on $TU$ are $(z, q)$ where $q = F'(x)p$. In particular,

$$\begin{pmatrix} dz \\ dq \end{pmatrix} = \begin{pmatrix} F'(x) & 0 \\ G(x, p) & F'(x) \end{pmatrix} \begin{pmatrix} dx \\ dp \end{pmatrix}.$$

where $G$ is some matrix function whose exact form is not relevant. Then writing $L_z$ for $\left( \frac{\partial L}{\partial z^1}, \ldots, \frac{\partial L}{\partial z^n} \right)$, etc., yields

$$\begin{aligned} dL &= L_z\, dz + L_q\, dq \\ &= \big( L_z F'(x) + L_q G(x, p) \big)\, dx + L_q F'(x)\, dp \\ &= L_x\, dx + L_p\, dp. \end{aligned}$$

Comparing $dp$-coefficients yields $L_p = L_q F'(x)$, so $L_p\, dx = L_q F'(x)\, dx = L_q\, dz$. In particular, as we wished to show, there exists a well-defined 1-form $\omega_L$ on $TM$ whose coordinate expression in local canonical coordinates $(x, p)$ is $L_p\, dx$.

The remainder of the proof is a coordinate calculation. The reader will want to note that I am using the expression $\ddot{\gamma}$ to denote the velocity of the curve $\dot{\gamma}$ in $TM$. The curve $\dot{\gamma}$ is described in $U$ as $(x, p) = (y, \dot{y})$ and its velocity vector $\ddot{\gamma}$ is simply $(\dot{x}, \dot{p}) = (\dot{y}, \ddot{y})$.

Now, the Euler-Lagrange equations are just

$$\frac{\partial L}{\partial x^i}(y, \dot{y}) = \frac{d}{dt}\left( \frac{\partial L}{\partial p^i}(y, \dot{y}) \right) = \frac{\partial^2 L}{\partial p^i \partial p^j}(y, \dot{y})\ddot{y}^j + \frac{\partial^2 L}{\partial p^i \partial x^j}(y, \dot{y})\dot{y}^j.$$

Meanwhile,

$$d\omega_L = \frac{\partial^2 L}{\partial p^i \partial p^j} dp^j \wedge dx^i + \frac{\partial^2 L}{\partial p^i \partial x^j} dx^j \wedge dx^i,$$

so

$$\ddot{\gamma} \lrcorner\, d\omega_L = \frac{\partial^2 L}{\partial p^i \partial p^j}(y, \dot{y}) \left( \ddot{y}^j\, dx^i - \dot{y}^i\, dp^j \right) + \frac{\partial^2 L}{\partial p^i \partial x^j}(y, \dot{y}) \left( \dot{y}^j\, dx^i - \dot{y}^i\, dx^j \right).$$

On the other hand, an easy computation yields

$$-dE_L(\dot{\gamma}) = \left( \frac{\partial L}{\partial x^i}(y, \dot{y}) - \frac{\partial^2 L}{\partial p^j \partial x^i}(y, \dot{y})\dot{y}^j \right) dx^i - \frac{\partial^2 L}{\partial p^i \partial p^j}(y, \dot{y})\dot{y}^i\, dp^j.$$

Comparing these last two equations, the condition $\ddot{\gamma} \lrcorner\, d\omega_L = -dE_L(\dot{\gamma})$ is seen to be the Euler-Lagrange equations, as desired. $\qquad\square$

**Conservation of Energy.** One important consequence of Proposition 3 is that the function $E_L$ is constant along the curve $\dot{\gamma}$ for any $L$-critical curve $\gamma\colon [a, b] \to M$. This follows since, for such a curve,

$$dE_L\big(\ddot{\gamma}(t)\big) = -d\omega_L\big(\ddot{\gamma}(t), \ddot{\gamma}(t)\big) = 0.$$

$E_L$ is generally interpreted as the "energy" of the Lagrangian $L$, and this constancy of $E_L$ on $L$-critical curves is often called the principle of Conservation of Energy.

▶    *Some sources define $E_L$ as $L - dL(R)$. My choice was to have $E_L$ agree with the classical energy in the classical problems.*

L.4.6                                                                             

**Definition 3:** If $L: TM \to \mathbb{R}$ is a Lagrangian on $M$, a diffeomorphism $f: M \to M$ is said to be a *symmetry* of $L$ if $L$ is invariant under the induced diffeomorphism $f': TM \to TM$, i.e., if $L \circ f' = L$. A vector field $X$ on $M$ is said to be an *infinitesimal symmetry* of $L$ if the (local) flow $\Phi_t$ of $X$ is a symmetry of $L$ for all $t$.

It is perhaps necessary to make a remark about the last part of this definition. For a vector field $X$ which is not necessarily complete, and for any $t \in \mathbb{R}$, the "time $t$" local flow of $X$ is well-defined on an open set $U_t \subset M$. The local flow of $X$ then gives a well-defined diffeomorphism $\Phi_t: U_t \to U_{-t}$. The requirement for $X$ is that, for each $t$ for which $U_t \neq \emptyset$, the induced map $\Phi'_t: TU_t \to TU_{-t}$ should satisfy $L \circ \Phi'_t = L$. (Of course, if $X$ is complete, then $U_t = M$ for all $t$, so symmetry has its usual meaning.)

Let $X$ be any vector field on $M$ with local flow $\Phi$. This induces a local flow on $TM$ which is associated to a vector field $X'$ on $TM$. If, in a local coordinate chart, $x: U \to \mathbb{R}^n$, the vector field $X$ has the expression

$$X = a^i(x) \, \frac{\partial}{\partial x^i},$$

then the reader may check that, in the associated canonical coordinates on $TU$,

$$X' = a^i \, \frac{\partial}{\partial x^i} + p^j \, \frac{\partial a^i}{\partial x^j} \, \frac{\partial}{\partial p^i}.$$

The condition that $X$ be an infinitesimal symmetry of $L$ is then that $L$ be invariant under the flow of $X'$, i.e., that

$$dL(X') = a^i \, \frac{\partial L}{\partial x^i} + p^j \, \frac{\partial a^i}{\partial x^j} \, \frac{\partial L}{\partial p^i} = 0.$$

The following theorem is now a simple calculation. Nevertheless, it is the foundation of a vast theory. It usually goes by the name "Noether's Theorem", though, in fact, Noether's Theorem is more general.

**Theorem 1:** *If $X$ is an infinitesimal symmetry of the Lagrangian $L$, then the function $\omega_L(X')$ is constant on $\dot{\gamma}: [a, b] \to TM$ for every $L$-critical path $\gamma: [a, b] \to M$.*

PROOF: Since the flow of $X'$ fixes $L$ it should not be too surprising that it also fixes $E_L$ and $\omega_L$. These facts are easily checked by the reader in local coordinates, so they are left as exercises. In particular,

$$\mathfrak{L}_{X'} \omega_L = d(X' \lrcorner \omega_L) + X' \lrcorner d\omega_L = 0 \qquad \text{and} \qquad \mathfrak{L}_{X'} E_L = dE_L(X') = 0.$$

Thus, for any $L$-critical curve $\gamma$ in $M$,

$$
\begin{aligned}
d\big(\omega_L(X')\big)\big(\ddot{\gamma}(t)\big) &= d(X' \lrcorner \omega_L)\big(\ddot{\gamma}(t)\big) = -(X' \lrcorner d\omega_L)\big(\ddot{\gamma}(t)\big) \\
&= d\omega_L\big(\ddot{\gamma}(t), X'(\dot{\gamma}(t))\big) = \big(\ddot{\gamma}(t) \lrcorner d\omega_L\big)\big(X'(\dot{\gamma}(t))\big) \\
&= -dE_L\big(X'(\dot{\gamma}(t))\big) = 0.
\end{aligned}
$$

L.4.7

Hence, the function $\omega_L(X')$ is constant on $\dot{\gamma}$, as desired. □

Of course, the formula for $\omega_L(X')$ in local canonical coordinates is simply

$$\omega_L(X') = a^i \frac{\partial L}{\partial p^i},$$

and the constancy of this function on the solution curves of the Euler-Lagrange equations is not difficult to check directly.

The principle

$$\textbf{Symmetry} \quad \Longrightarrow \quad \textbf{Conservation Law}$$

is so fundamental that whenever a new system of equations is encountered an enormous effort is expended to determine its symmetries. Moreover, the intuition is often expressed that "every conservation law ought to come from some symmetry", so whenever conserved quantities are observed in Nature (or, more accurately, our models of Nature) people nowadays look for a symmetry to explain it. Even when no symmetry is readily apparent, in many cases a sort of "hidden symmetry" can be found.

**Example:** MOTION IN A CENTRAL FORCE FIELD. Consider the Lagrangian of "kinetic minus potential energy" for an particle (of mass $m \neq 0$) moving in a "central force field". Here, we take $\mathbb{R}^n$ with its usual inner product and a function $V(|x|^2)$ (called the potential energy) which depends only on distance from the origin. The Lagrangian is

$$L(x,p) = \tfrac{m}{2}|p|^2 - V(|x|^2).$$

The function $E_L$ is given by

$$E_L(x,p) = \tfrac{m}{2}|p|^2 + V(|x|^2),$$

and $\omega_L = m\, p^i\, dx^i = m\, p \cdot dx$.

The Lagrangian $L$ is clearly symmetric with respect to rotations about the origin. For example, the rotation in the $ij$-plane is generated by the vector field

$$X_{ij} = x^j \frac{\partial}{\partial x^i} - x^i \frac{\partial}{\partial x^j}.$$

According to Noether's Theorem, then, the functions

$$\mu_{ij} = \omega_L(X'_{ij}) = m\big(x^j\, p^i - x^i\, p^j\big)$$

are constant on all solutions. These are usually called the "angular momenta". It follows from their constancy that the bivector $\xi = y(t) \wedge \dot{y}(t)$ is constant on any solution $x = y(t)$ of the Euler-Lagrange equations and hence that $y(t)$ moves in a fixed 2-plane. Thus, we are essentially reduced to the case $n = 2$. In this case, for constants $E_0$ and $\mu_0$, the equations

$$\tfrac{m}{2}|p|^2 + V(|x|^2) = E_0 \qquad \text{and} \qquad m(x^1 p^2 - x^2 p^1) = \mu_0$$

L.4.8

will generically define a surface in $T\mathbb{R}^2$. The solution curves to the Euler-Lagrange equations

$$\dot{x} = p \qquad \text{and} \qquad \dot{p} = -\frac{2}{m}V'(|x|^2)x$$

which lie on this surface can then be analysed by phase portrait methods. (In fact, they can be integrated by quadrature.)

**Example:** RIEMANNIAN METRICS WITH SYMMETRIES. As another example, consider the case of a Riemannian manifold with infinitesimal symmetries. If the flow of $X$ on $M$ preserves a Riemannian metric $L$, then, in local coordinates,

$$L = g_{ij}(x)p^i p^j$$

and

$$X = a^i(x)\frac{\partial}{\partial x^i}.$$

According to Conservation of Energy and Noether's Theorem, the functions

$$E_L = g_{ij}(x)p^i p^j \qquad \text{and} \qquad \omega_L(X') = 2g_{ij}(x)a^i(x)p^j$$

are first integrals of the geodesic equations.

For example, if a surface $S \subset \mathbb{R}^3$ is a surface of revolution, then the induced metric can locally be written in the form

$$I = E(r)\,dr^2 + 2\,F(r)dr\,d\theta + G(r)\,d\theta^2$$

where the rotational symmetry is generated by the vector field $X = \partial/\partial\theta$. The following functions are then constant on solutions of the geodesic equations:

$$E(r)\,\dot{r}^2 + 2\,F(r)\dot{r}\,\dot{\theta} + G(r)\,\dot{\theta}^2 \qquad \text{and} \qquad F(r)\dot{r} + G(r)\,\dot{\theta}.$$

This makes it possible to integrate by quadratures the geodesic equations on a surface of revolution, a classical accomplishment. (See the Exercises for details.)

**Subexample:** LEFT INVARIANT METRICS ON LIE GROUPS. Let $G$ be a Lie group and let $\omega^1, \omega^2, \ldots, \omega^n$ be any basis for the left-invariant 1-forms on $G$. Consider the Lagrangian

$$L = \left(\omega^1\right)^2 + \cdots + \left(\omega^n\right)^2,$$

which defines a left-invariant metric on $G$. Since left translations are symmetries of this metric and since the flows of the *right*-invariant vector fields $Y_i$ leave the left-invariant 1-forms fixed, we see that these generate symmetries of the Lagrangian $L$. In particular, the functions $E_L = L$ and

$$\mu_i = \omega^1(Y_i)\omega^1 + \cdots + \omega^n(Y_i)\omega^n$$

are functions on $TG$ which are constant on all of the geodesics of $G$ with the metric $L$. I will return to this example several times in future lectures.

Subsubexample: THE MOTION OF RIGID BODIES. A special case of the Lie group example is particularly noteworthy, namely the theory of the rigid body.

A *rigid body* (in $\mathbb{R}^n$) is a (finite) set of points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ with masses $m_1, \ldots, m_N$ such that the distances $d_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$ are fixed (hence the name "rigid"). The free motion of such a body is governed by the "kinetic energy" Lagrangian

$$L = \frac{m_1}{2}|\mathbf{p}_1|^2 + \cdots + \frac{m_N}{2}|\mathbf{p}_N|^2.$$

where $\mathbf{p}_i$ represents the velocity of the $i$'th point mass. Here is how this can be converted into a left-invariant Lagrangian variational problem on a Lie group:

Let $G$ be the matrix Lie group

$$G = \left\{ \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \ \middle|\ A \in \mathrm{O}(n), b \in \mathbb{R}^n \right\}.$$

Then $G$ acts as the space of isometries of $\mathbb{R}^n$ with its usual metric and thus also acts on the $N$-fold product

$$Y_N = \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$$

by the "diagonal" action. It is not difficult to show that $G$ acts transitively on the simultaneous level sets of the functions $f_{ij}(\mathbf{x}) = |\mathbf{x}_i - \mathbf{x}_j|$. Thus, for each symmetric matrix $\Delta = (d_{ij})$, the set

$$M_\Delta = \{\mathbf{x} \in Y_N \mid |\mathbf{x}_i - \mathbf{x}_j| = d_{ij} \}$$

is an orbit of $G$ (and hence a smooth manifold) when it is not empty. The set $M_\Delta$ is said to be the "configuration space" of the rigid body. (Question: Can you determine a necessary and sufficient condition on the matrix $\Delta$ so that $M_\Delta$ is not empty? In other words, which rigid bodies are possible?)

Let us suppose that $M_\Delta$ is not empty and let $\bar{\mathbf{x}} \in M_\Delta$ be a "reference configuration" which, for convenience, we shall suppose has its center of mass at the origin:

$$m_k\, \bar{\mathbf{x}}_k = 0.$$

(This can always be arranged by a simultaneous translation of all of the point masses.) Now let $\gamma\colon [a, b] \to M_\Delta$ be a curve in the configuration space. (Such curves are often called "trajectories".) Since $M_\Delta$ is a $G$-orbit, there is a curve $g\colon [a, b] \to G$ so that $\gamma(t) = g(t) \cdot \bar{\mathbf{x}}$. Let us write

$$\gamma(t) = (\mathbf{x}_1(t), \ldots, \mathbf{x}_N(t))$$

and let

$$g(t) = \begin{pmatrix} A(t) & b(t) \\ 0 & 1 \end{pmatrix}.$$

The value of the canonical left invariant form on $g$ is

$$g^{-1}\dot{g} = \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A^{-1}\dot{A} & A^{-1}\dot{b} \\ 0 & 0 \end{pmatrix}.$$

The kinetic energy along the trajectory $\gamma$ is then

$$\tfrac{1}{2}\sum_k m_k\,|\dot{\mathbf{x}}_k|^2 = \tfrac{1}{2}\sum_k m_k\,(\dot{\mathbf{x}}_k \cdot \dot{\mathbf{x}}_k) = \tfrac{1}{2}\sum_k m_k\,\left(\dot{A}\bar{\mathbf{x}}_k + \dot{b}\right) \cdot \left(\dot{A}\bar{\mathbf{x}}_k + \dot{b}\right).$$

Since $A$ is a curve in $\mathrm{O}(n)$, this becomes

$$= \tfrac{1}{2}\sum_k m_k\,\left((A^{-1}(\dot{A}\bar{\mathbf{x}}_k + \dot{b}))\right) \cdot \left((A^{-1}(\dot{A}\bar{\mathbf{x}}_k + \dot{b}))\right)$$

$$= \tfrac{1}{2}\sum_k m_k\,(\alpha\bar{\mathbf{x}}_k + \beta) \cdot (\alpha\bar{\mathbf{x}}_k + \beta).$$

Using the center-of-mass normalization, this simplifies to

$$= \tfrac{1}{2}\sum_k m_k\,\left(-{}^t\bar{\mathbf{x}}_k\,\alpha^2\,\bar{\mathbf{x}}_k + |\beta|^2\right).$$

With a slight rearrangement, this takes the simple form

$$L\big(\dot{\gamma}(t)\big) = -\operatorname{tr}\left(\big(\alpha(t)\big)^2 \mu\right) + \tfrac{1}{2}m\,|\beta(t)|^2$$

where $m = m_1 + \cdots + m_N$ is the total mass of the body and $\mu$ is the positive semi-definite symmetric $n$-by-$n$ matrix

$$\mu = \tfrac{1}{2}\sum_k m_k\,\bar{\mathbf{x}}_k\,{}^t\bar{\mathbf{x}}_k.$$

It is clear that we can interpret $L$ as a left-invariant Lagrangian on $G$. Actually, even the formula we have found so far can be simplified: If we write $\mu = R\,\delta\,{}^t R$ where $\delta$ is diagonal and $R$ is an orthogonal matrix (which we can always do), then right acting on $G$ by the element

$$\begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix}$$

will reduce the Lagrangian to the form

$$L\big(\dot{g}(t)\big) = -\operatorname{tr}\left(\big(\alpha(t)\big)^2 \delta\right) + \tfrac{1}{2}m\,|\beta(t)|^2.$$

Thus, only the eigenvalues of the matrix $\mu$ really matter in trying to solve the equations of motion of a rigid body. This observation is usually given an interpretation like "the motion of any rigid body is equivalent to the motion of its 'ellipsoid of inertia' ".

**Hamiltonian Form.** Let us return to the consideration of the Euler-Lagrange equations. As we have seen, in expanded form, the equations in local coordinates are

$$\frac{\partial^2 L}{\partial p^i \partial p^j}(y, \dot{y})\ddot{y}^j + \frac{\partial^2 L}{\partial p^i \partial x^j}(y, \dot{y})\dot{y}^j - \frac{\partial L}{\partial x^i}(y, \dot{y}) = 0.$$

L.4.11

In order for these equations to be solvable for the highest derivatives at every possible set of initial conditions, the symmetric matrix

$$H_L(x,p) = \left( \frac{\partial^2 L}{\partial p^i \partial p^j}(x,p) \right).$$

must be invertible at every point $(x,p)$.

**Definition 4:** A Lagrangian $L$ is said to be *non-degenerate* if, relative to every local coordinate system $x \colon U \to \mathbb{R}^n$, the matrix $H_L$ is invertible at every point of $TU$.

For example, if $L \colon TM \to \mathbb{R}$ restricts to each tangent space $T_m M$ to be a non-degenerate quadratic form, then $L$ is a non-degenerate Lagrangian. In particular, when $L$ is a Riemannian metric, $L$ is non-degenerate.

Although Definition 4 is fairly explicit, it is certainly not coordinate free. Here is a result which may clarify the meaning of non-degenerate.

**Proposition 4:** *The following are equivalent for a Lagrangian $L \colon TM \to \mathbb{R}$:*

(1) *$L$ is a non-degenerate Lagrangian.*
(2) *In local coordinates $(x,p)$, the functions $x^1, \ldots, x^n, \partial L/\partial p^1, \ldots, \partial L/\partial p^n$ have everywhere independent differentials.*
(3) *The 2-form $d\omega_L$ is non-degenerate at every point of $TM$, i.e., for any tangent vector $v \in T(TM)$, $v \lrcorner \, d\omega_L = 0$ implies that $v = 0$.*

PROOF: The equivalence of (1) and (2) follows directly from the Chain Rule and is left as an exercise. The equivalence of (2) and (3) can be seen as follows: Let $v \in T_a(TM)$ be a tangent vector based at $a \in T_m M$. Choose any local any canonical local coordinate system $(x,p)$ with $m \in U$ and write $q_i = \partial L/\partial p^i$ for $1 \le i \le n$. Then $\omega_L$ takes the form

$$d\omega_L = dq_i \wedge dx^i.$$

Thus,

$$v \lrcorner \, d\omega_L = dq_i(v)\, dx^i - dx^i(v)\, dq_i \, .$$

Now, suppose that the differentials $dx^1, \ldots, dx^n, dq_i, \ldots, dq_n$ are linearly independent at $a$ and hence span $T_a^*(TM)$. Then, if $v \lrcorner \, d\omega_L = 0$, we must have $dq_i(v) = dx^i(v) = 0$, which, because the given $2n$ differentials form a spanning set, implies that $v = 0$ Thus, $d\omega_L$ is non-degenerate at $a$.

On the other hand, suppose that that the differentials $dx^1, \ldots, dx^n, dq_1, \ldots, dq_n$ are linearly dependent at $a$. Then, by linear algebra, there exists a non-zero vector $v \in T_a^*(TM)$ so that $dq_i(v) = dx^i(v) = 0$. However, it is then clear that $v \lrcorner \, d\omega_L = 0$ for such a $v$, so that $d\omega_L$ will be degenerate at $a$. $\square$

For physical reasons, the function $q_i$ is usually called the *conjugate momentum* to the coordinate $x^i$.

L.4.12

Before exploring the geometric meaning of the coordinate system $(x, q)$, we want to give the following description of the $L$-critical curves of a non-degenerate Lagrangian.

**Proposition 5:** *If $L : TM \to \mathbb{R}$ is a non-degenerate Lagrangian, then there exists a unique vector field $Y$ on $TM$ so that, for every $L$-critical curve $\gamma : [a, b] \to M$, the associated curve $\dot\gamma : [a, b] \to TM$ is an integral curve of $Y$. Conversely, for any integral curve $\varphi : [a, b] \to TM$ of $Y$, the composition $\phi = \pi \circ \varphi : [a, b] \to M$ is an $L$-critical curve in $M$.*

PROOF: It is clear that we should take $Y$ to be the unique vector field on $TM$ which satisfies $Y \lrcorner\, d\omega_L = -dE_L$. (There is only one since, by Proposition 4, $d\omega_L$ is non-degenerate.) Proposition 3 then says that for every $L$-critical curve, its lift $\dot\gamma$ satisfies $\ddot\gamma(t) = Y(\dot\gamma(t))$ for all $t$, i.e., that $\dot\gamma$ is indeed an integral curve of $Y$.

The details of the converse will be left to the reader. First, one must check that, with $\phi$ defined as above, we have $\dot\phi = \varphi$. This is best done in local coordinates. Second, one must check that $\phi$ is indeed $L$-critical, even though it may not lie entirely within a coordinate neighborhood. This may be done by computing the variation of $\phi$ restricted to appropriate subintervals and taking account of the boundary terms introduced by integration by parts when the endpoints are not fixed. Details are in the Exercises. $\qquad\square$

The canonical vector field $Y$ on $TM$ is just the coordinate free way of expressing the fact that, for non-degenerate Lagrangians, the Euler-Lagrangian equations are simply a non-singular system of second order ODE for maps $\gamma : [a, b] \to M$

Unfortunately, the expression for $Y$ in canonical $(x, p)$-coordinates on $TM$ is not very nice; it involves the inverse of the matrix $H_L$. However, in the $(x, q)$-coordinates, it is a completely different story. In these coordinates, everything takes a remarkably simple form, a fact which is the cornerstone on symplectic geometry and the calculus of variations.

Before taking up the geometric interpretation of these new coordinates, let us do a few calculations. We have already seen that , in these coordinates, the canonical 1-form $\omega_L$ takes the simple form $\omega_L = q_i\, dx^i$.

We can also express $E_L$ as a function of $(x, q)$. It is traditional to denote this expression by $H(x, q)$ and call it the *Hamiltonian* of the variational problem (even though, in a certain sense, it is the same function as $E_L$). The equation determining the vector field $Y$ is expressed in these coordinates as

$$Y \lrcorner\, d\omega_L = Y \lrcorner\, (dq_i \wedge dx^i) = dq_i(Y)\, dx^i - dx^i(Y)\, dq_i$$

$$= -dH = -\frac{\partial H}{\partial x^i}\, dx^i - \frac{\partial H}{\partial q_i}\, dq_i \ ,$$

so the expression for $Y$ in these coordinates is

$$Y = \frac{\partial H}{\partial q_i}\frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i}\frac{\partial}{\partial q_i}.$$

In particular, the flow of $Y$ takes the form

$$\dot x^i = \frac{\partial H}{\partial q_i} \qquad \text{and} \qquad \dot q_i = -\frac{\partial H}{\partial x^i}.$$

L.4.13

These equations are known as *Hamilton's Equations* or, sometimes, as the *Hamiltonian form* of the Euler-Lagrange equations.

Part of the reason for the importance of the $(x, q)$ coordinates is the symmetric way they treat the positions and momenta. Another reason comes from the form the infinitesimal symmetries take in these coordinates: If $X$ is an infinitesimal symmetry of $L$ and $X'$ is the induced vector field on $TM$ with conserved quantity $G = \omega_L(X')$, then, since $\mathfrak{L}_{X'}\omega_L = 0$,

$$ X' \lrcorner \, d\omega_L = -d(\omega_L(X')) = -dG. $$

Thus, by the same analysis as above, the ODE represented by $X'$ in the $(x, q)$ coordinates becomes

$$ \dot{x}^i = \frac{\partial G}{\partial q_i} \qquad \text{and} \qquad \dot{q}_i = -\frac{\partial G}{\partial x^i}. $$

In other words, in the $(x, q)$ coordinates, the flow of a symmetry $X'$ has the same Hamiltonian form as the flow of the vector field $Y$ which gives the solutions of the Euler-Lagrange equations! This method of putting the symmetries of a Lagrangian and the solutions of the Lagrangian on a sort of equal footing will be seen to have powerful consequences.

**The Cotangent Bundle.** Early on in this lecture, we introduced, for each coordinate chart $x: U \to \mathbb{R}^n$, a canonical extension $(x, p): TU \to \mathbb{R}^n \times \mathbb{R}^n$ and characterized it by a geometric property. There is also a canonical extension $(x, \xi): T^*U \to \mathbb{R}^n \times \mathbb{R}^n$ where $\xi = (\xi_i): T^*U \to \mathbb{R}^n$ is characterized by the condition that, if $f: U \to \mathbb{R}$ is any smooth function on $U$, then, regarding its exterior derivative $df$ as a section $df: U \to T^*U$, we have

$$ \xi_i \circ df = \frac{\partial f}{\partial x^i}. $$

I will leave to the reader the task of showing that $(x, \xi)$ is indeed a coordinate system on $T^*U$.

It is a remarkable fact that the cotangent bundle $\pi: T^*M \to M$ of any smooth manifold carries a canonical 1-form $\omega$ defined by the following property: For each $\alpha \in T_x^*M$, we define the linear function $\omega_\alpha: T_\alpha(T^*M) \to \mathbb{R}$ by the rule $\omega_\alpha(v) = \alpha(\pi'(\alpha)(v))$. I leave to the reader the task of showing that, in canonical coordinates $(x, \xi): T^*U \to \mathbb{R}^n \times \mathbb{R}^n$, this canonical 1-form has the expression

$$ \omega = \xi_i \, dx^i. $$

**The Legendre transformation.** Now consider a smooth Lagrangian $L: TM \to \mathbb{R}$ as before. We can use $L$ to construct a smooth mapping $\tau_L: TM \to T^*M$ as follows: At each $v \in TM$, the 1-form $\omega_L(v)$ is semi-basic, i.e., there exists a (necessarily unique) 1-form $\tau_L(v) \in T^*_{\pi(v)}M$ so that $\omega_L(v) = \pi^*(\tau_L(v))$. This mapping is known as the *Legendre transformation* associated to the Lagrangian $L$.

L.4.14

74

This definition is rather abstract, but, in local coordinates, it takes a simple form. The reader can easily check that in canonical coordinates associated to a coordinate chart $x: U \to \mathbb{R}^n$, we have

$$(x, \xi) \circ \tau_L = (x, q) = \left( x^i, \frac{\partial L}{\partial p^i} \right).$$

In other words, the $(x, q)$ coordinates are just the canonical coordinates on the cotangent bundle composed with the Legendre transformation! It is now immediate that $\tau_L$ is a local diffeomorphism if and only if $L$ is a non-degenerate Lagrangian. Moreover, we clearly have $\omega_L = \tau_L^*(\omega)$, so the 1-form $\omega_L$ is also expressible in terms of the canonical 1-form $\omega$ and the Legendre transform.

What about the function $E_L$ on $TM$? Let us put the following condition on the Lagrangian $L$: Let us assume that $\tau_L: TM \to T^*M$ is a (one-to-one) diffeomorphism onto its image $\tau_L(TM) \subset T^*M$. (Note that this implies that $L$ is non-degenerate, but is stronger than this.) Then there clearly exists a function on $\tau_L(TM)$ which pulls back to $TM$ to be $E_L$. In fact, as the reader can easily verify, this is none other than the Hamiltonian function $H$ constructed above.

The fact that the Hamiltonian $H$ naturally "lives" on $T^*M$ (or at least an open subset thereof) rather than on $TM$ justifies it being regarded as distinct from the function $E_L$. There is another reason for moving over to the cotangent bundle when one can: The vector field $Y$ on $TM$ corresponds, under the Legendre transformation, to a vector field $Z$ on $\tau_L(TM)$ which is characterized by the simple rule $Z \lrcorner d\omega = -dH$. Thus, just knowing the Hamiltonian $H$ on an open set in $T^*M$ determines the vector field which sweeps out the solution curves! We will see that this is a very useful observation in what follows.

**Poincaré Recurrence.** To conclude this lecture, I want to give an application of the geometry of the form $\omega_L$ to understanding the global behavior of the $L$-critical curves when $L$ is a non-degenerate Lagrangian. First, I make the following observation:

**Proposition 6:** *Let $L: TM \to \mathbb{R}$ be a non-degenerate Lagrangian. Then $2n$-form $\mu_L = (d\omega_L)^n$ is a volume form on $TM$ (i.e., it is nowhere vanishing). Moreover the (local) flow of the vector field $Y$ preserves this volume form.*

PROOF: To see that $\mu_L$ is a volume form, just look in local $(x, q)$-coordinates:

$$\begin{aligned}
\mu_L = (d\omega_L)^n &= (dq_i \wedge dx^i)^n \\
&= n!\, dq_1 \wedge dx^2 \wedge dq_2 \wedge dx^2 \wedge \cdots \wedge dq_n \wedge dx^n.
\end{aligned}$$

By Proposition 4, this latter form is not zero.

Finally, since

$$\mathfrak{L}_Y(d\omega_L) = d(Y \lrcorner d\omega_L) = -d(dE_L) = 0,$$

it follows that the (local) flow of $Y$ preserves $d\omega_L$ and hence preserves $\mu_L$. $\qquad \square$

Now we shall give an application of Proposition 6. This is the famous Poincaré Recurrence Theorem.

**Theorem 2:** *Let $L\colon TM \to \mathbb{R}$ be a non-degenerate Lagrangian and suppose that $E_L$ is a proper function on $TM$. Then the vector field $Y$ is complete, with flow $\Phi\colon \mathbb{R} \times TM \to TM$. Moreover, this flow is recurrent in the following sense: For any point $v \in TM$, any open neighborhood $U$ of $v$, and any positive time interval $T > 0$, there exists an integer $N > 0$ so that $\Phi(TN, U) \cap U \neq \emptyset$.*

PROOF: The completeness of the flow of $Y$ follows immediately from the fact that the integral curve of $Y$ which passes through $v \in TM$ must stay in the compact set $E_L^{-1}\big(E_L(v)\big)$. (Recall that $E_L$ is constant on all of the integral curves of $Y$.) Details are left to the reader.

I now turn to the proof of the recurrence property. Let $E_0 = E_L(v)$. By hypothesis, the set $C = E_L^{-1}\big([E_0 - 1, E_0 + 1]\big)$ is compact, so the $\mu_L$-volume of the open set $W = E_L^{-1}\big((E_0 - 1, E_0 + 1)\big)$ (which lies inside C) is finite. It clearly suffices to prove the recurrence property for any open neighborhood $U$ of $v$ which lies inside $W$, so let us assume that $U \subset W$.

Let $\phi\colon W \to W$ be the diffeomorphism $\phi(w) = \Phi(T, w)$. This diffeomorphism is clearly invertible and preserves the $\mu_L$-volume of open sets in $W$. Consider the open sets $U^k = \phi^k(U)$ for $k > 0$ (integers). These open sets all have the same $\mu_L$-volume and hence cannot be all disjoint since then their union (which lies in $W$) would have infinite $\mu_L$-volume. Let $0 < j < k$ be two integers so that $U^j \cap U^k \neq \emptyset$. Then, since

$$U^j \cap U^k = \phi^j(U) \cap \phi^k(U) = \phi^j\big(U \cap \phi^{k-j}(U)\big),$$

it follows that $U \cap \phi^{k-j}(U) \neq \emptyset$, as we wished to show. $\qquad\square$

This theorem has the amazing consequence that, whenever one has a non-degenerate Lagrangian with a proper energy function, the corresponding mechanical system "recurs" in the sense that "arbitrarily near any given initial condition, there is another initial condition so that the evolution brings this initial condition back arbitrarily close to the first initial condition". I realize that this statement is somewhat vague and subject to misinterpretation, but the precise statement has already been given, so there seems not to be much harm in giving the paraphrase.

L.4.16

# Exercise Set 4:

# Symmetries and Conservation Laws

**1**. Show that two Lagrangians $L_1, L_2 \colon TM \to \mathbb{R}$ satisfy

$$E_{L_1} = E_{L_2} \qquad \text{and} \qquad d\omega_{L_1} = d\omega_{L_2}$$

if and only if there is a closed 1-form $\phi$ on $M$ so that $L_1 = L_2 + \phi$. (Note that, in this equation, we interpret $\phi$ as a function on $TM$.) Such Lagrangians are said to differ by a "divergence term." Show that such Lagrangians share the same critical curves and that one is non-degenerate if and only if the other is.

**2**. What does Conservation of Energy mean for the case where $L$ defines a Riemannian metric on $M$?

**3**. Show that the equations for geodesics of a rotationally invariant metric of the form

$$I = E(r)\, dr^2 + 2\, F(r) dr\, d\theta + G(r)\, d\theta^2$$

can be integrated by separation of variables and quadratures. (Hint: Start with the conservation laws we already know:

$$E(r)\, \dot{r}^2 + 2\, F(r)\dot{r}\, \dot{\theta} + G(r)\, \dot{\theta}^2 = v_0^2$$
$$F(r)\dot{r} + G(r)\, \dot{\theta} = u_0$$

where $v_0$ and $u_0$ are constants. Then eliminate $\dot{\theta}$ and go on from there.)

**4**. The definition of $\omega_L$ given in the text might be regarded as somewhat unsatisfactory since it is given in coordinates and not "invariantly". Show that the following invariant description of $\omega_L$ is valid: The manifold $TM$ inherits some extra structure by virtue of being the tangent bundle of another manifold $M$. Let $\pi \colon TM \to M$ be the basepoint projection. Then $\pi$ is a submersion: For every $a \in TM$,

$$\pi'(a) \colon T_a TM \to T_{\pi(a)} M$$

is a surjection and the fiber at $\pi(a)$ is equal to

$$\pi^{-1}\big(\pi(a)\big) = T_{\pi(a)} M.$$

It follows that the kernel of $\pi'(a)$ (i.e., the "vertical space" of the bundle $\pi \colon TM \to M$ at $a$) is naturally isomorphic to $T_{\pi(a)} M$. Call this isomorphism $\alpha \colon T_{\pi(a)} M \xrightarrow{\sim} \ker\big(\pi'(a)\big)$. Then the 1-form $\omega_L$ is defined by

$$\omega_L(v) = dL\big(\alpha \circ \pi'(v)\big) \qquad \text{for } v \in T(TM).$$

Hint: Show that, in local canonical coordinates, the map $\alpha \circ \pi'$ satisfies

$$\alpha \circ \pi' \left( a^i \frac{\partial}{\partial x^i} + b^i \frac{\partial}{\partial p^i} \right) = a^i \frac{\partial}{\partial p^i}.$$

**5**. For any vector field $X$ on $M$, let the associated vector field on $TM$ be denoted $X'$. Show that if $X$ has the form

$$X = a^i \frac{\partial}{\partial x^i}$$

in some local coordinate system, then, in the associated canonical $(x, p)$ coordinates, $X'$ has the form

$$X' = a^i \frac{\partial}{\partial x^i} + p^j \frac{\partial a^i}{\partial x^j} \frac{\partial}{\partial p^i}.$$

**6**. Show that conservation of angular momenta in the motion of a point mass in a central force field implies Kepler's Law that "equal areas are swept out over equal time intervals." Show also that, in the $n = 2$ case, employing the conservation of energy and angular momentum allows one to integrate the equations of motion by quadratures. (Hint: For the second part of the problem, introduce polar coordinates: $(x^1, x^2) = (r\cos\theta, r\sin\theta)$.)

**7**. In the example of the motion of a rigid body, show that the Lagrangian on $G$ is always non-negative and is non-degenerate (so that $L$ defines a left-invariant metric on $G$) if and only if the matrix $\mu$ has at most one zero eigenvalue. Show that $L$ is degenerate if and only if the rigid body lies in a subspace of dimension at most $n-2$.

**8**. Supply the details in the proof of Proposition 5. You will want to go back to the integration-by-parts derivation of the Euler-Lagrange equations and show that, even if the variation $\Gamma$ induced by $h$ does not have fixed endpoints, we still get a local coordinate formula of the form

$$\mathcal{F}'_{L,\Gamma}(0) = \frac{\partial L}{\partial p^k}\big(y(b), \dot{y}(b)\big) h^k(b) - \frac{\partial L}{\partial p^k}\big(y(a), \dot{y}(a)\big) h^k(a)$$

for any variation of a solution of the Euler-Lagrange equations. Give these "boundary terms" an invariant geometric meaning and show that they cancel out when we sum over a partition of a (fixed-endpoint) variation of an $L$-critical curve $\gamma$ into subcurves which lie in coordinate neighborhoods.)

**9**. (Alternate to Exercise 8.) Here is another approach to proving Proposition 5. Instead of dividing the curve up into sub-curves, show that for any variation $\Gamma$ of a curve $\gamma: [a, b] \to M$ (not necessarily with fixed endpoints), we have the formula

$$\mathcal{F}'_{L,\Gamma}(0) = \omega_L\big(V(b)\big) - \omega_L\big(V(a)\big) - \int_a^b d\omega_L\big(\ddot{\gamma}(t), V(t)\big) + dE_L\big(V(t)\big)\, dt$$

where $V(t) = \dot{\Gamma}'(t,0)(\partial/\partial s)$ is the "variation vector field" at $s = 0$ of the lifted variation $\dot{\Gamma}$ in $TM$. Conclude that, whether $L$ is non-degenerate or not, the condition $\ddot{\gamma} \lrcorner d\omega_L + dE_L\big(\dot{\gamma}(t)\big) = 0$ is the necessary and sufficient condition that $\gamma$ be $L$-critical.

**10**. **The Two Body Problem.** Consider a pair of point masses (with masses $m_1$ and $m_2$) which move freely subject to a force between them which depends only on the distance between the two bodies and is directed along the line joining the two bodies. This is what is classically known as the Two Body Problem. It is represented by a Lagrangian on the manifold $M = \mathbb{R}^n \times \mathbb{R}^n$ with position coordinates $x_1, x_2 \colon M \to \mathbb{R}^n$ of the form

$$L(x_1, x_2, p_1, p_2) = \frac{m_1}{2}|p_1|^2 + \frac{m_2}{2}|p_2|^2 - V(|x_1 - x_2|^2).$$

(Here, $(p_1, p_2)$ are the canonical fiber (velocity) coordinates on TM associated to the coordinate system $(x_1, x_2)$.) Notice that $L$ has the form "kinetic minus potential". Show that rotations and translations in $\mathbb{R}^n$ generate a group of symmetries of this Lagrangian and compute the conserved quantities. What is the interpretation of the conservation law associated to the translations?

**11**. **The Sliding Particle.** Suppose that a particle of unit weight and mass (remember: "geometric units" means never having to state your constants) slides without friction on a smooth hypersurface $x^{n+1} = F(x^1, \ldots, x^n)$ subject only to the force of gravity (which is directed downward along the $x^{n+1}$-axis). Show that the "kinetic-minus-potential" Lagrangian for this motion in the $x$-coordinates is

$$L = \tfrac{1}{2}\left((p^1)^2 + \cdots + (p^n)^2 + \left(\frac{\partial F}{\partial x^i}p^i\right)^2\right) - F(x^1, \ldots, x^n).$$

Show that this is a non-degenerate Lagrangian and that its energy $E_L$ is proper if and only if $F^{-1}\big((-\infty, a]\big)$ is compact for all $a \in \mathbb{R}$.

Suppose that $F$ is invariant under rotation, i.e., that

$$F(x^1, \ldots, x^n) = f\big((x^1)^2 + \cdots + (x^n)^2\big)$$

for some smooth function $f$. Show that the "shadow" of the particle in $\mathbb{R}^n$ stays in a fixed 2-plane. Show that the equations of motion can be integrated by quadrature.

Remark: This Lagrangian is also used to model a small ball of unit mass and weight "rolling without friction in a cup". Of course, in this formulation, the kinetic energy stored in the ball by its spinning is ignored. If you want to take this "spinning" energy into account, then you must study quite a different Lagrangian, especially if you assume that the ball rolls without slipping. This goes into the very interesting theory of "non-holonomic systems", which we (unfortunately) do not have time to go into.

**12**. Let $L$ be a Lagrangian which restricts to each fiber $T_xM$ to be a non-degenerate (though not necessarily positive definite) quadratic form. Show that $L$ is non-degnerate as a Lagrangian and that the Legendre mapping $\tau_L \colon TM \to T^*M$ is an isomorphism of vector bundles. Show that, if $L$ is, in addition a positive definite quadratic form on each fiber, then the new Lagrangian defined by

$$\tilde{L} = \big(L + 1\big)^{\frac{1}{2}}$$

is also a non-degenerate Lagrangian, but that the map $\tau_{\tilde{L}} \colon TM \to T^*M$, though one-to-one, is not onto.

# Lecture 5:

# Symplectic Manifolds, I

In Lecture 4, I associated a non-degenerate 2-form $d\omega_L$ on $TM$ to every non-degenerate Lagrangian $L: TM \to \mathbb{R}$. In this section, I want to begin a more systematic study of the geometry of manifolds on which there is specified a closed, non-degenerate 2-form.

## Symplectic Algebra.

First, I will develop the algebraic precursors of the manifold concepts which are to follow. For simplicity, all of these constructions will be carried out on vector spaces over the reals, but they could equally well have been carried out over any field of characteristic not equal to 2.

**Symplectic Vector Spaces.** A bilinear pairing $B: V \times V \to \mathbb{R}$ is said to be *skew-symmetric* (or *alternating*) if $B(x,y) = -B(y,x)$ for all $x, y$ in $V$. The space of skew-symmetric bilinear pairings on $V$ will be denoted by $A^2(V)$. The set $A^2(V)$ is a vector space under the obvious addition and scalar multiplication and is naturally identified with $\Lambda^2(V^*)$, the space of exterior 2-forms on $V$. The elements of $A^2(V)$ are often called skew-symmetric bilinear *forms* on V. A pairing $B \in A^2(V)$ is said to be *non-degenerate* if, for every non-zero $v \in V$, there is a $w \in V$ for which $B(v, w) \neq 0$.

**Definition 1:** A *symplectic space* is a pair $(V, B)$ where $V$ is a vector space and $B$ is a non-degenerate, skew-symmetric, bilinear pairing on $V$.

**Example.** Let $V = \mathbb{R}^{2n}$ and let $J_n$ be the $2n$-by-$2n$ matrix

$$ J_n = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}. $$

For vectors $v, w \in \mathbb{R}^{2n}$, define

$$ B_0(x, y) = {}^t x \, J_n \, y. $$

Then it is clear that $B_0$ is bilinear and skew-symmetric. Moreover, in components

$$ B_0(x, y) = x^1 y^{n+1} + \cdots + x^n y^{2n} - x^{n+1} y^1 - \cdots - x^{2n} y^n $$

so it is clear that if $B_0(x, y) = 0$ for all $y \in \mathbb{R}^{2n}$ then $x = 0$. Hence, $B_0$ is non-degenerate.

Generally, in order for $B(x, y) = {}^t x \, A \, y$ to define a skew-symmetric bilinear form on $\mathbb{R}^n$, it is only necessary that $A$ be a skew-symmetric $n$-by-$n$ matrix. Conversely, every skew-symmetric bilinear form $B$ on $\mathbb{R}^n$ can be written in this form for some unique skew-symmetric $n$-by-$n$ matrix $A$. In order that this $B$ be non-degenerate, it is necessary and sufficient that $A$ be invertible. (See the Exercises.)

**The Symplectic Group.** Now, a linear transformation $R: \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ preserves $B_0$, i.e., satisfies $B_0(Rx, Ry) = B_0(x, y)$ for all $x, y \in \mathbb{R}^{2n}$, if and only if ${}^t\!R\, J_n\, R = J_n$. This motivates the following definition:

**Definition 2:** The subgroup of $GL(2n, \mathbb{R})$ defined by

$$\mathrm{Sp}(n, \mathbb{R}) = \left\{ R \in GL(2n, \mathbb{R}) \mid {}^t\!R\, J_n\, R = J_n \right\}$$

is called the *symplectic group of rank n*.

It is clear that $\mathrm{Sp}(n, \mathbb{R})$ is a (closed) subgroup of $GL(2n, \mathbb{R})$. In the Exercises, you are asked to prove that $\mathrm{Sp}(n, \mathbb{R})$ is a Lie group of dimension $2n^2 + n$ and to derive other of its properties.

**Symplectic Normal Form.** The following proposition shows that there is a normal form for finite dimensional symplectic spaces.

**Proposition 1:** *If $(V, B)$ is a finite dimensional symplectic space, then there exists a basis $e_1, \ldots, e_n, f^1 \ldots, f^n$ of $V$ so that, for all $1 \le i, j \le n$,*

$$B(e_i, e_j) = 0, \qquad B(e_i, f^j) = \delta_i^j, \text{ and} \qquad B(f^i, f^j) = 0$$

PROOF: The desired basis will be constructed in two steps. Let $m = \dim(V)$.

Suppose that for some $n \ge 0$, we have found a sequence of linearly independent vectors $e_1, \ldots, e_n$ so that $B(e_i, e_j) = 0$ for all $1 \le i, j \le n$. Consider the vector space $W_n \subset V$ which consists of all vectors $w \in V$ so that $B(e_i, w) = 0$ for all $1 \le i \le n$. Since the $e_i$ are linearly independent and since $B$ is non-degenerate, it follows that $W_n$ has dimension $m - n$. We must have $n \le m - n$ since all of the vectors $e_1, \ldots, e_n$ clearly lie in $W_n$.

If $n < m - n$, then there exists a vector $e_{n+1} \in W_n$ which is linearly independent from $e_1, \ldots, e_n$. It follows that the sequence $e_1, \ldots, e_{n+1}$ satisfies $B(e_i, e_j) = 0$ for all $1 \le i, j \le n + 1$. (Since $B$ is skew-symmetric, $B(e_{n+1}, e_{n+1}) = 0$ is automatic.) This extension process can be repeated until we reach a stage where $n = m - n$, i.e., $m = 2n$. At that point, we will have a sequence $e_1, \ldots, e_n$ for which $B(e_i, e_j) = 0$ for all $1 \le i, j \le n$.

Next, we construct the sequence $f^1, \ldots, f^n$. For each $j$ in the range $1 \le j \le n$, consider the set of $n$ linear equations

$$B(e_i, w) = \delta_i^j, \qquad 1 \le i \le n.$$

We know that these $n$ equations are linearly independent, so there exists a solution $f_0^j$. Of course, once one particular solution is found, any other solution is of the form $f^j = f_0^j + a^{ji} e_i$ for some $n^2$ numbers $a^{ji}$. Thus, we have found the general solutions $f^j$ to the equations $B(e_i, f^j) = \delta_i^j$.

L.5.2

We now show that we can choose the $a^{ij}$ so as to satisfy the last remaining set of conditions, $B(f^i, f^j) = 0$. If we set $b^{ij} = B(f_0^i, f_0^j) = -b^{ji}$, then we can compute

$$\begin{aligned} B(f^i, f^j) &= B(f_0^i, f_0^j) + B(a^{ik}e_k, f_0^j) + B(f_0^i, a^{jl}e_l) + B(a^{ik}e_k, a^{jl}e_l) \\ &= b^{ij} + a^{ij} - a^{ji} + 0. \end{aligned}$$

Thus, it suffices to set $a^{ij} = -b^{ij}/2$. (This is where the hypothesis that the characteristic of $\mathbb{R}$ is not 2 is used.)

Finally, it remains to show that the vectors $e_1, \ldots, e_n, f^1 \ldots, f^n$ form a basis of $V$. Since we already know that $\dim(V) = 2n$, it is enough to show that these vectors are linearly independent. However, any linear relation of the form

$$a^i e_i + b_j f^j = 0,$$

implies $b_k = B(e_k, a^i e_i + b_j f^j) = 0$ and $a^k = -B(f^k, a^i e_i + b_j f^j) = 0$. $\qquad \square$

▶ We often say that a basis of the form found in Proposition 1 is a *symplectic* or *standard* basis of the symplectic space $(V, B)$.

**Symplectic Reduction of Vector Spaces.** If $B\colon V \times V \to \mathbb{R}$ is a skew-symmetric bilinear form which is not necessarily non-degenerate, then we define the *null space* of $B$ to be the subspace
$$N_B = \{v \in V \mid B(v, w) = 0 \text{ for all } w \in V\}.$$

On the quotient vector space $\overline{V} = V/N_B$, there is a well-defined skew-symmetric bilinear form $\overline{B}\colon \overline{V} \times \overline{V} \to \mathbb{R}$ given by
$$\overline{B}(\overline{x}, \overline{y}) = B(x, y)$$

where $\overline{x}$ and $\overline{y}$ are the cosets in $\overline{V}$ of $x$ and $y$ in $V$. It is easy to see that $(\overline{V}, \overline{B})$ is a symplectic space.

**Definition 2:** If $B$ is a skew-symmetric bilinear form on a vector space $V$, then the symplectic space $(\overline{V}, \overline{B})$ is called the *symplectic reduction* of $(V, B)$.

Here is an application of the symplectic reduction idea: Using the identification of $A^2(V)$ with $\Lambda^2(V^*)$ mentioned earlier, Proposition 1 allows us to write down a normal form for any alternating 2-form on any finite dimensional vector space.

**Proposition 2:** *For any non-zero $\beta \in \Lambda^2(V^*)$, there exist an integer $n \leq \frac{1}{2}\dim(V)$ and linearly independent 1-forms $\omega^1, \omega^2, \ldots, \omega^{2n} \in V^*$ for which*

$$\beta = \omega^1 \wedge \omega^2 + \omega^3 \wedge \omega^4 \ldots + \omega^{2n-1} \wedge \omega^{2n}.$$

*Thus, $n$ is the largest integer so that $\beta^n \neq 0$.*

L.5.3

PROOF: Regard $\beta$ as a skew-symmetric bilinear form $B$ on $V$ in the usual way. Let $(\overline{V}, \overline{B})$ be the symplectic reduction of $(V, B)$. Since $B \neq 0$, we known that $\overline{V} \neq \{0\}$. Let $\dim(\overline{V}) = 2n \geq 2$ and let $e_1, \ldots, e_n, f^1 \ldots, f^n$ be elements of $V$ so that $\overline{e}_1, \ldots, \overline{e}_n, \overline{f}^1 \ldots, \overline{f}^n$ forms a symplectic basis of $\overline{V}$ with respect to $\overline{B}$. Let $p = \dim(V) - 2n$, and let $b_1, \ldots, b_p$ be a basis of $N_B$.

It is easy to see that

$$\mathbf{b} = \begin{pmatrix} e_1 & f^1 & e_2 & f^2 & \cdots & e_n & f^n & b_1 & \cdots & b_p \end{pmatrix}$$

forms a basis of $V$. Let

$$\omega^1 \quad \cdots \quad \omega^{2n+p}$$

denote the dual basis of $V^*$. Then, as the reader can easily check, the 2-form

$$\Omega = \omega^1 \wedge \omega^2 + \omega^3 \wedge \omega^4 \ldots + \omega^{2n-1} \wedge \omega^{2n}$$

has the same values as $\beta$ does on all pairs of elements of $\mathbf{b}$. Of course this implies that $\beta = \Omega$. The rest of the Proposition also follows easily since, for example, we have

$$\beta^n = n! \, \omega^1 \wedge \cdots \wedge \omega^{2n} \neq 0,$$

although $\beta^{n+1}$ clearly vanishes. $\qquad \square$

▶      If we regard $\beta$ as an element of $A^2(V)$, then $n$ is one-half the dimension of $\overline{V}$. Some sources call the integer $n$ the half-rank of $\beta$ and others call $n$ the rank. I use "half-rank".

Note that, unlike the case of *symmetric* bilinear forms, there is no notion of signature type or "positive definiteness" for skew-symmetric forms.

▶      It follows from Proposition 2 that for $\beta$ in $A^2(V)$, where $V$ is finite dimensional, the pair $(V, \beta)$ is a symplectic space if and only if $V$ has dimension $2n$ for some $n$ and $\beta^n \neq 0$.

**Subspaces of Symplectic Vector Spaces.** Let $\Omega$ be a symplectic form on a vector space $V$. For any subspace $W \subset V$, we define the $\Omega$-complement to $W$ to be the subspace

$$W^\perp = \{ v \in V \,|\, \Omega(v, w) = 0 \text{ for all } w \in W \}.$$

The $\Omega$-complement of a subspace $W$ is sometimes called its *skew-complement*. It is an exercise for the reader to check that, because $\Omega$ is non-degenerate, $(W^\perp)^\perp = W$ and that, when $V$ is finite-dimensional,

$$\dim W + \dim W^\perp = \dim V.$$

However, unlike the case of an orthogonal with respect to a positive definite inner product, the intersection $W \cap W^\perp$ does not have to be the zero subspace. For example, in an $\Omega$-standard basis for $V$, the vectors $e_1, \ldots, e_n$ obviously span a subspace $L$ which satisfies $L^\perp = L$.

If $V$ is finite dimensional, it turns out (see the Exercises) that, up to symplectic linear transformations of $V$, a subspace $W \subset V$ is characterized by the numbers $d = \dim W$ and $\nu = \dim(W \cap W^\perp) \le d$. If $\nu = 0$ we say that $W$ is a *symplectic subspace* of $V$. This corresponds to the case that $\Omega$ restricts to $W$ to define a symplectic structure on $W$. At the other extreme is when $\nu = d$, for then we have $W \cap W^\perp = W$. Such a subspace is called *Lagrangian.*

## Symplectic Manifolds.

We are now ready to return to the study of manifolds.

**Definition 3:** A *symplectic structure* on a smooth manifold $M$ is a non-degenerate, closed 2-form $\Omega \in \mathcal{A}^2(M)$. The pair $(M, \Omega)$ is called a *symplectic manifold.* If $\Omega$ is a symplectic structure on $M$ and $\Upsilon$ is a symplectic structure on $N$, then a smooth map $\phi\colon M \to N$ satisfying $\phi^*(\Upsilon) = \Omega$ is called a *symplectic map.* If, in addition, $\phi$ is a diffeomorphism, we say that $\phi$ is a *symplectomorphism.*

Before developing any of the theory, it is helpful to see a few examples.

**Surfaces with Area Forms.** If $S$ is an orientable smooth surface, then there exists a volume form $\mu$ on $S$. By definition, $\mu$ is a non-degenerate closed 2-form on $S$ and hence defines a symplectic structure on $S$.

**Lagrangian Structures on $TM$.** From Lecture 4, any non-degenerate Lagrangian $L\colon TM \to \mathbb{R}$ defines the 2-form $d\omega_L$, which is a symplectic structure on $TM$.

**A "Standard" Structure on $\mathbb{R}^{2n}$.** Think of $\mathbb{R}^{2n}$ as a smooth manifold and let $\Omega$ be the 2-form with constant coefficients

$$\Omega = \tfrac{1}{2}{}^t dx \, J_n \, dx = dx^1 \wedge dx^{n+1} + \cdots + dx^n \wedge dx^{2n}.$$

**Symplectic Submanifolds.** Let $(M^{2m}, \Omega)$ be a symplectic manifold. Suppose that $P^{2p} \subset M^{2m}$ be any submanifold to which the form $\Omega$ pulls back to be a non-degenerate 2-form $\Omega_P$. Then $(P, \Omega_P)$ is a symplectic manifold. We say that $P$ is a *symplectic submanifold* of $M$.

It is not obvious just how to find symplectic submanifolds of $M$. Even though being a symplectic submanifold is an "open" condition on submanifolds of $M$, is is not "dense". One cannot hope to perturb an arbitrary even dimensional submanifold of $M$ slightly so as to make it symplectic. There are even restrictions on the topology of the submanifolds of $M$ on which a symplectic form restricts to be non-degenerate.

For example, no symplectic submanifold of $\mathbb{R}^{2n}$ (with any symplectic structure on $\mathbb{R}^{2n}$) could be compact for the following simple reason: Since $\mathbb{R}^{2n}$ is contractible, its second deRham cohomology group vanishes. In particular, for any symplectic form $\Omega$ on $\mathbb{R}^{2n}$, there must be a 1-form $\omega$ so that $\Omega = d\omega$ which implies that $\Omega^m = d\left(\omega \wedge \Omega^{m-1}\right)$. Thus, for all $m > 0$, the $2m$-form $\Omega^m$ is exact on $\mathbb{R}^{2n}$ (and every submanifold of $\mathbb{R}^{2n}$).

By Proposition 2, if $M^{2m}$ were a submanifold of $\mathbb{R}^{2n}$ on which $\Omega$ restricted to be non-degenerate, then $\Omega^m$ would be a volume form on $M$. However, on a compact manifold the volume form is never exact (just apply Stokes' Theorem).

**Example.** COMPLEX SUBMANIFOLDS. Nevertheless, there are many symplectic submanifolds of $\mathbb{R}^{2n}$. One way to construct them is to regard $\mathbb{R}^{2n}$ as $\mathbb{C}^n$ in such a way that the linear map $J{:}\mathbb{R}^{2n} \to \mathbb{R}^{2n}$ represented by $J_n$ becomes complex multiplication. (For example, just define the complex coordinates by $z^k = x^k + ix^{k+n}$.) Then, for any non-zero vector $v \in \mathbb{R}^{2n}$, we have $\Omega(v, Jv) = -|v|^2 \neq 0$. In particular, $\Omega$ is non-degenerate on every complex subspace $S \subset \mathbb{C}^n$. Thus, if $M^{2m} \subset \mathbb{C}^n$ is any complex submanifold (i.e., all of its tangent spaces are $m$-dimensional complex subspaces of $\mathbb{C}^m$), then $\Omega$ restricts to be non-degenerate on $M$.

**The Cotangent Bundle.** Let $M$ be any smooth manifold and let $T^*M$ be its cotangent bundle. As we saw in Lecture 4, there is a canonical 2-form on $T^*M$ which can be defined as follows: Let $\pi{:}T^*M \to M$ be the basepoint projection. Then, for every $v \in T_\alpha(T^*M)$, define

$$\omega(v) = \alpha\big(\pi'(v)\big).$$

I claim that $\omega$ is a smooth 1-form on $T^*M$ and that $\Omega = d\omega$ is a symplectic form on $T^*M$.

To see this, let us compute $\omega$ in local coordinates. Let $x{:}U \to \mathbb{R}^n$ be a local coordinate chart. Since the 1-forms $dx^1, \ldots, dx^n$ are linearly independent at every point of $U$, it follows that there are unique functions $\xi_i$ on $T^*U$ so that, for $\alpha \in T_a^*U$,

$$\alpha = \xi_1(\alpha)\, dx^1|_a + \cdots + \xi_n(\alpha)\, dx^n|_a\,.$$

The functions $x^1, \ldots, x^n, \xi_1, \ldots, \xi_n$ then form a smooth coordinate system on $T^*U$ in which the projection mapping $\pi$ is given by

$$\pi(x, p) = x.$$

It is then straightforward to compute that, in this coordinate system,

$$\omega = \xi_i\, dx^i.$$

Hence, $\Omega = d\xi_i \wedge dx^i$ and so is non-degenerate.

**Symplectic Products.** If $(M, \Omega)$ and $(N, \Upsilon)$ are symplectic manifolds, then $M \times N$ carries a natural symplectic structure, called the product symplectic structure $\Omega \oplus \Upsilon$, defined by

$$\Omega \oplus \Upsilon = \pi_1^*(\Omega) + \pi_2^*(\Upsilon).$$

Thus, for example, $n$-fold products of compact surfaces endowed with area forms give examples of compact symplectic $2n$-manifolds.

**Coadjoint Orbits.** Let $\mathrm{Ad}^*\colon G \to \mathrm{GL}(\mathfrak{g}^*)$ denote the coadjoint representation of $G$. This is the so-called "contragredient" representation to the adjoint representation. Thus, for any $a \in G$ and $\xi \in \mathfrak{g}^*$, the element $\mathrm{Ad}^*(a)(\xi) \in \mathfrak{g}^*$ is determined by the rule

$$\mathrm{Ad}^*(a)(\xi)(x) = \xi\big(\mathrm{Ad}(a^{-1})(x)\big) \qquad \text{for all } x \in \mathfrak{g}.$$

▶  *One must be careful not to confuse $Ad^*(a)$ with $\big(\mathrm{Ad}(a)\big)^*$. Instead, as our definition shows, $Ad^*(a) = \big(\mathrm{Ad}(a^{-1})\big)^*$.*

Note that the induced homomorphism of Lie algebras, $\mathrm{ad}^*\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}^*)$ is given by

$$\mathrm{ad}^*(x)(\xi)(y) = -\xi\big([x,y]\big)$$

The orbits $G \cdot \xi$ in $\mathfrak{g}^*$ are called the *coadjoint orbits*. Each of them carries a natural symplectic structure. To see how this is defined, let $\xi \in \mathfrak{g}^*$ be fixed, and let $\phi\colon G \to G \cdot \xi$ be the usual submersion induced by the $\mathrm{Ad}^*$-action, $\phi(a) = \mathrm{Ad}^*(a)(\xi) = a \cdot \xi$. Now let $\omega_\xi$ be the left-invariant 1-form on $G$ whose value at $e$ is $\xi$. Thus, $\omega_\xi = \xi(\omega)$ where $\omega$ is the canonical $\mathfrak{g}$-valued 1-form on $G$.

**Proposition 3:**  *There is a unique symplectic form $\Omega_\xi$ on the orbit $G \cdot \xi = G/G_\xi$ satisfying $\phi^*(\Omega_\xi) = d\omega_\xi$.*

PROOF:  If Proposition 3 is to be true, then $\Omega_\xi$ must satisfy the rule

$$\Omega_\xi\big(\phi'(v), \phi'(w)\big) = d\omega_\xi(v, w) \quad \text{for all } v, w \in T_a G.$$

What we must do is show that this rule actually does define a symplectic 2-form on $G \cdot \xi$.

First, note that, for $x, y \in \mathfrak{g} = T_e G$, we may compute via the structure equations that

$$d\omega_\xi(x, y) = \xi\big(d\omega(x, y)\big) = \xi\big(-[x,y]\big) = ad^*(x)(\xi)(y).$$

In particular, $ad^*(x)(\xi) = 0$, if and only if $x$ lies in the null space of the 2-form $d\omega_\xi(e)$. In other words, the null space of $d\omega_\xi(e)$ is $\mathfrak{g}_\xi$, the Lie algebra of $G_\xi$. Since $d\omega_\xi$ is left-invariant, it follows that the null space of $d\omega_\xi(a)$ is $L'_a(\mathfrak{g}_\xi) \subset T_a G$. Of course, this is precisely the tangent space at $a$ to the left coset $aG_\xi$. Thus, for each $a \in G$,

$$N_{d\omega_\xi(a)} = \ker \phi'(a),$$

It follows that, $T_{a \cdot \xi}(G \cdot \xi) = \phi'(a)(T_a G)$ is naturally isomorphic to the symplectic quotient space $(T_a G)/\big(L'_a(\mathfrak{g}_\xi)\big)$ for each $a \in G$. Thus, there is a unique, non-degenerate 2-form $\Omega_a$ on $T_{a \cdot \xi}(G \cdot \xi)$ so that $\big(\phi'(a)\big)^*(\Omega_a) = d\omega_\xi(a)$.

It remains to show that $\Omega_a = \Omega_b$ if $a \cdot \xi = b \cdot \xi$. However, this latter case occurs only if $a = bh$ where $h \in G_\xi$. Now, for any $h \in G_\xi$, we have

$$R_h^*(\omega_\xi) = \xi\big(R_h^*(\omega)\big) = \xi\big(\mathrm{Ad}(h^{-1})(\omega)\big) = \mathrm{Ad}^*(h)(\xi)(\omega) = \xi(\omega) = \omega_\xi.$$

Thus, $R_h^*(d\omega_\xi) = d\omega_\xi$. Since the following square commutes, it follows that $\Omega_a = \Omega_b$.

$$
\begin{array}{ccc}
T_aG & \xrightarrow{R_h'} & T_bG \\
\phi'(a) \downarrow & & \downarrow \phi'(b) \\
T_{a\cdot\xi}(G\cdot\xi) & \xrightarrow{id} & T_{b\cdot\xi}(G\cdot\xi)
\end{array}
$$

All this shows that there is a well-defined, non-degenerate 2-form $\Omega_\xi$ on $G\cdot\xi$ which satisfies $\phi^*(\Omega_\xi) = d\omega_\xi$. Since $\phi$ is a smooth submersion, the equation

$$
\phi^*(d\Omega_\xi) = d(d\omega_\xi) = 0
$$

implies that $d\Omega_\xi = 0$, as promised. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

▶       Note that a consequence of Proposition 3 is that all of the coadjoint orbits are actually even dimensional. As we shall see when we take up the subject of reduction, the coadjoint orbits are particularly interesting symplectic manifolds.

**Examples:** Let $G = \mathrm{O}(n)$, with Lie algebra $\mathfrak{so}(n)$, the space of skew-symmetric $n$-by-$n$ matrices. Now there is an $\mathrm{O}(n)$-equivariant positive definite pairing of $\mathfrak{so}(n)$ with itself $\langle,\rangle$ given by

$$
\langle x, y\rangle = -\mathrm{tr}(xy).
$$

Thus, we can identify $\mathfrak{so}(n)^*$ with $\mathfrak{so}(n)$ by this pairing. The reader can check that, in this case, the coadjoint action is isomorphic to the adjoint action

$$
\mathrm{Ad}(a)(x) = axa^{-1}.
$$

If $\xi$ is the rank 2 matrix

$$
\xi = \begin{pmatrix} \begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix} & 0 \\ 0 & 0 \end{pmatrix},
$$

then it is easy to check that the stabilizer $G_\xi$ is just the set of matrices of the form

$$
\begin{pmatrix} a & 0 \\ 0 & A \end{pmatrix}
$$

where $a \in \mathrm{SO}(2)$ and $A \in \mathrm{O}(n-2)$. The quotient $\mathrm{O}(n)/\big(\mathrm{SO}(2)\times\mathrm{O}(n-2)\big)$ thus has a symplectic structure. It is not difficult to see that this homogeneous space can be identified with the space of oriented 2-planes in $\mathbb{E}^n$.

As another example, if $n = 2m$, then $J_m$ lies in $\mathfrak{so}(2m)$, and its stabilizer is $\mathrm{U}(m) \subset \mathrm{SO}(2m)$. It follows that the quotient space $\mathrm{SO}(2m)/\mathrm{U}(m)$, which is identifiable as the set of orthogonal complex structures on $\mathbb{E}^{2m}$, is a symplectic space.

Finally, if $G = \mathrm{U}(n)$, then, again, we can identify $\mathfrak{u}(n)^*$ with $\mathfrak{u}(n)$ via the $\mathrm{U}(n)$-invariant, positive definite pairing

$$
\langle x, y\rangle = -\mathrm{Re}\big(\mathrm{tr}(xy)\big).
$$

Again, under this identification, the coadjoint action agrees with the adjoint action. For $0 < p < n$, the stabilizer of the element

$$\xi_p = \begin{pmatrix} iI_p & 0 \\ 0 & -iI_{n-p} \end{pmatrix}$$

is easily seen to be $\mathrm{U}(p) \times \mathrm{U}(n-p)$. The orbit of $\xi_p$ is identifiable with the space $\mathrm{Gr}_p(\mathbb{C}^n)$, i.e., the Grassmannian of (complex) $p$-planes in $\mathbb{C}^n$, and, by Proposition 3, carries a canonical, $\mathrm{U}(n)$-invariant symplectic structure.

**Darboux' Theorem.** There is a manifold analogue of Proposition 1 which says that symplectic manifolds of a given dimension are all locally "isomorphic". This fundamental result is known as *Darboux' Theorem*. I will give the classical proof (due to Darboux) here, deferring the more modern proof (due to Weinstein) to the next section.

**Theorem 1:** *(Darboux' Theorem) If $\Omega$ is a closed 2-form on a manifold $M^{2n}$ which satisfies the condition that $\Omega^n$ be nowhere vanishing, then for every $p \in M$, there is a neighborhood $U$ of $p$ and a coordinate system $x_1, x_2, \ldots, x_n, y^1, y^2, \ldots, y^n$ on $U$ so that*

$$\Omega_{|U} = dx_1 \wedge dy^1 + dx_2 \wedge dy^2 + \cdots + dx_n \wedge dy^n.$$

PROOF: We will proceed by induction on $n$. Assume that we know the theorem for $n-1 \geq 0$. We will prove it for $n$. Fix $p$, and let $y^1$ be a smooth function on $M$ for which $dy^1$ does not vanish at $p$. Now let $X$ be the unique (smooth) vector field which satisfies

$$X \lrcorner \Omega = dy^1.$$

This vector field does not vanish at $p$, so there is a function $x_1$ on a neighborhood $U$ of $p$ which satisfies $X(x_1) = 1$. Now let $Y$ be the vector field on $U$ which satisfies

$$Y \lrcorner \Omega = -dx_1.$$

Since $d\Omega = 0$, the Cartan formula, now gives

$$\mathcal{L}_X \Omega = \mathcal{L}_Y \Omega = 0.$$

We now compute

$$[X, Y] \lrcorner \Omega = \mathcal{L}_X Y \lrcorner \Omega = \mathcal{L}_X (Y \lrcorner \Omega) - Y \lrcorner (\mathcal{L}_X \Omega)$$
$$= \mathcal{L}_X (-dx_1) = -d\left(X(x^1)\right) = -d(1) = 0.$$

Since $\Omega$ has maximal rank, this implies $[X, Y] = 0$. By the simultaneous flow-box theorem, it follows that there exist local coordinates $x_1, y^1, z^1, z^2, \ldots, z^{2n-2}$ on some neighborhood $U_1 \subset U$ of $p$ so that

$$X = \frac{\partial}{\partial x_1} \qquad \text{and} \qquad Y = \frac{\partial}{\partial y^1}.$$

L.5.9

88

Now consider the form $\Omega' = \Omega - dx_1 \wedge dy^1$. Clearly $d\Omega' = 0$. Moreover,

$$X \lrcorner \Omega' = \mathfrak{L}_X \Omega' = Y \lrcorner \Omega' = \mathfrak{L}_Y \Omega' = 0.$$

It follows that $\Omega'$ can be expressed as a 2-form in the variables $z^1, z^2, \ldots, z^{2n-2}$ alone. Hence, in particular, $(\Omega')^{n+1} \equiv 0$. On the other hand, by the binomial theorem, then

$$0 \neq \Omega^n = n\, dx_1 \wedge dy^1 \wedge (\Omega')^{n-1}.$$

It follows that $\Omega'$ may be regarded as a closed 2-form of maximal half-rank $n-1$ on an open set in $\mathbb{R}^{2n-2}$. Now apply the inductive hypothesis to $\Omega'$. $\qquad\square$

Darboux' Theorem has a generalization which covers the case of closed 2-forms of constant (though not necessarily maximal) rank. It is the analogue for manifolds of the symplectic reduction of a vector space.

**Theorem 2:** *(Darboux' Reduction Theorem) Suppose that $\Omega$ is a closed 2-form of constant half-rank $n$ on a manifold $M^{2n+k}$. Then the "null bundle"*

$$N_\Omega = \big\{ v \in TM \mid \Omega(v, w) = 0 \text{ for all } w \in T_{\pi(v)}M \big\}$$

*is integrable and of constant rank $k$. Moreover, any point of $M$ has a neighborhood $U$ on which there exist local coordinates $x_1, \ldots, x_n, y^1, \ldots, y^n, z^1, \ldots z^k$ in which*

$$\Omega_{|U} = dx_1 \wedge dy^1 + dx_2 \wedge dy^2 + \cdots + dx_n \wedge dy^n.$$

PROOF: Note that a vector field $X$ on $M$ is a section of $N_\Omega$ if and only if $X \lrcorner \Omega = 0$. In particular, since $\Omega$ is closed, the Cartan formula implies that $\mathfrak{L}_X \Omega = 0$ for all such $X$.

If $X$ and $Y$ are two sections of $N_\Omega$, then

$$[X, Y] \lrcorner \Omega = \mathfrak{L}_X (Y \lrcorner \Omega) - Y(\mathfrak{L}_X \Omega) = 0 - 0 = 0,$$

so it follows that $[X, Y]$ is a section of $N_\Omega$ as well. Thus, $N_\Omega$ is integrable.

Now apply the Frobenius Theorem. For any point $p \in M$, there exists a neighborhood $U$ on which there exist local coordinates $z^1 \ldots, z^{2n+k}$ so that $N_\Omega$ restricted to $U$ is spanned by the vector fields $Z_i = \partial/\partial z^i$ for $1 \leq i \leq k$. Since $Z_i \lrcorner \Omega = \mathfrak{L}_{Z_i} \Omega = 0$ for $1 \leq i \leq k$, it follows that $\Omega$ can be expressed on $U$ in terms of the variables $z^{k+1}, \ldots, z^{2n+k}$ alone. In particular, $\Omega$ restricted to $U$ may be regarded as a non-degenerate closed 2-form on an open set in $\mathbb{R}^{2n}$. The stated result now follows from Darboux' Theorem. $\qquad\square$

**Symplectic and Hamiltonian vector fields.**

We now want to examine some of the special vector fields which are defined on symplectic manifolds. Let $M^{2n}$ be manifold and let $\Omega$ be a symplectic form on $M$. Let $\mathrm{Sp}(\Omega) \subset \mathrm{Diff}(M)$ denote the subgroup of symplectomorphisms of $(M, \Omega)$. We would like to follow Lie in regarding $\mathrm{Sp}(\Omega)$ as an "infinite dimensional Lie group". In that case, the Lie algebra of $\mathrm{Sp}(\Omega)$ should be the space of vector fields whose flows preserve $\Omega$. Of course, $\Omega$ will be invariant under the flow of a vector field $X$ if and only if $\mathfrak{L}_X \Omega = 0$. This motivates the following definition:

**Definition 4:** A vector field $X$ on $M$ is said to be *symplectic* if $\mathfrak{L}_X \Omega = 0$. The space of symplectic vector fields on $M$ will be denoted $\mathrm{sp}(\Omega)$.

It turns out that there is a very simple characterization of the symplectic vector fields on $M$: Since $d\Omega = 0$, it follows that for any vector field $X$ on $M$,
$$\mathfrak{L}_X \Omega = d(X \lrcorner \Omega).$$

▶   *Thus, $X$ is a symplectic vector field if and only if $X \lrcorner \Omega$ is closed.*

Now, since $\Omega$ is non-degenerate, for any vector field $X$ on $M$, the 1-form $\flat(X) = -X \lrcorner \Omega$ vanishes only where $X$ does. Since $TM$ and $T^*M$ have the same rank, it follows that the mapping $\flat \colon \mathfrak{X}(M) \to \mathcal{A}^1(M)$ is an isomorphism of $C^\infty(M)$-modules. In particular, $\flat$ has an inverse, $\sharp \colon \mathcal{A}^1(M) \to \mathfrak{X}(M)$.

With this notation, we can write $\mathrm{sp}(\Omega) = \sharp\big(\mathcal{Z}^1(M)\big)$ where $\mathcal{Z}^1(M)$ denotes the vector space of closed 1-forms on $M$. Now, $\mathcal{Z}^1(M)$ contains, as a subspace, $\mathcal{B}^1(M) = d\big(C^\infty(M)\big)$, the space of exact 1-forms on $M$. This subspace is of particular interest; we encountered it already in Lecture 4.

**Definition 5:** For each $f \in C^\infty(M)$, the vector field $X_f = \sharp(df)$ is called the *Hamiltonian vector field* associated to $f$. The set of all Hamiltonian vector fields on $M$ is denoted $\mathsf{h}(\Omega)$.

Thus, by definition, $\mathsf{h}(\Omega) = \sharp\big(\mathcal{B}^1(M)\big)$. For this reason, Hamiltonian vector fields are often called *exact*. Note that a Hamiltonian vector field is one whose equations, written in symplectic coordinates, represent an ODE in Hamiltonian form.

The following formula shows that, not only is $\mathrm{sp}(\Omega)$ a Lie algebra of vector fields, but that $\mathsf{h}(\Omega)$ is an ideal in $\mathrm{sp}(\Omega)$, i.e., that $[\mathrm{sp}(\Omega), \mathrm{sp}(\Omega)] \subset \mathsf{h}(\Omega)$.

**Proposition 4:**   *For $X, Y \in \mathrm{sp}(\Omega)$, we have*
$$[X, Y] = X_{\Omega(X,Y)}.$$
*In particular, $[X_f, X_g] = X_{\{f,g\}}$ where, by definition, $\{f, g\} = \Omega(X_f, X_g)$.*

PROOF: We use the fact that, for any vector field $X$, the operator $\mathfrak{L}_X$ is a derivation with respect to any natural pairing between tensors on $M$:
$$\begin{aligned}
[X, Y] \lrcorner \Omega &= \big(\mathfrak{L}_X Y\big) \lrcorner \Omega = \mathfrak{L}_X \big(Y \lrcorner \Omega\big) - Y \lrcorner \big(\mathfrak{L}_X \Omega\big) \\
&= d\big(X \lrcorner \big(Y \lrcorner \Omega\big)\big) + X \lrcorner d\big(Y \lrcorner \Omega\big) + 0 = d\big(\Omega(Y, X)\big) + 0 \\
&= -d\big(\Omega(X, Y)\big) = \big(X_{\Omega(X,Y)}\big) \lrcorner \Omega.
\end{aligned}$$
This proves our first equation. The remaining equation follows immediately.   □

The definition $\{f, g\} = \Omega(X_f, X_g)$ is an important one. The bracket $(f, g) \mapsto \{f, g\}$ is called the *Poisson bracket* of the functions $f$ and $g$. Proposition 4 implies that the Poisson bracket gives the functions on $M$ the structure of a Lie algebra. The Poisson bracket is slightly more subtle than the pairing $(X_f, X_g) \mapsto X_{\{f,g\}}$ since the mapping $f \mapsto X_f$ has a non-trivial kernel, namely, the locally constant functions.

Thus, if $M$ is connected, then we get an exact sequence of Lie algebras

$$0 \longrightarrow \mathbb{R} \longrightarrow C^\infty(M) \longrightarrow \mathfrak{h}(\Omega) \longrightarrow 0$$

which is not, in general, split (see the Exercises). Since $\{1, f\} = 0$ for all functions $f$ on $M$, it follows that the Poisson bracket on $C^\infty(M)$ makes it into a central extension of the algebra of Hamiltonian vector fields. The geometry of this central extension plays an important role in quantization theories on symplectic manifolds (see [GS 2] or [We]).

Also of great interest is the exact sequence

$$0 \longrightarrow \mathfrak{h}(\Omega) \longrightarrow \mathfrak{sp}(\Omega) \longrightarrow H^1_{dR}(M, \mathbb{R}) \longrightarrow 0,$$

where the right hand arrow is just the map described by $X \mapsto [X \lrcorner \Omega]$. Since the bracket of two elements in $\mathfrak{sp}(\Omega)$ lies in $\mathfrak{h}(\Omega)$, it follows that this linear map is actually a Lie algebra homomorphism when $H^1_{dR}(M, \mathbb{R})$ is given the abelian Lie algebra structure. This sequence also may or may not split (see the Exercises), and the properties of this extension have a great deal to do with the study of groups of symplectomorphisms of $M$. See the Exercises for further developments.

**Involution**

I now want to make some remarks about the meaning of the Poisson bracket and its applications.

**Definition 5:** Let $(M, \Omega)$ be a symplectic manifold. Two functions $f$ and $g$ are said to be *in involution* (with respect to $\Omega$) if they satisfy the condition $\{f, g\} = 0$.

▶ Note that, since $\{f, g\} = dg(X_f) = -df(X_g)$, it follows that two functions $f$ and $g$ are in involution if and only if each is constant on the integral curves of the other's Hamiltonian vector field.

Now, if one is trying to describe the integral curves of a Hamiltonian vector field, $X_f$, the more independent functions on $M$ that one can find which are constant on the integral curves of $X_f$, the more accurately one can describe those integral curves. If one were able find, in addition to $f$ itself, $2n-2$ additional independent functions on $M$ which are constant on the integral curves of $X_f$, then one could describe the integral curves of $X_f$ implicitly by setting those functions equal to a constant.

It turns out, however, that this is too much to hope for in general. It can happen that a Hamiltonian vector field $X_f$ has no functions in involution with it except for functions of the form $F(f)$.

Nevertheless, in many cases which arise in practice, we can find several functions in involution with a given function $f = f_1$ and, moreover, in involution with each other. In case one can find $n-1$ such independent functions, $f_2, \ldots, f_n$, we have the following theorem of Liouville which says that the remaining $n-1$ required functions can be found (at least locally) by quadrature alone. In the classical language, a vector field $X_f$ for which such functions are known is said to be "completely integrable by quadratures", or, more simply as "completely integrable".

**Theorem 3:** *Let $f^1, f^2, \ldots, f^n$ be $n$ functions in involution on a symplectic manifold $(M^{2n}, \Omega)$. Suppose that the functions $f^i$ are independent in the sense that the differentials $df^1, \ldots, df^n$ are linearly independent at every point of $M$. Then each point of $M$ has an open neighborhood $U$ on which there are functions $a_1, \ldots, a_n$ on $U$ so that*

$$\Omega = df^1 \wedge da_1 + \cdots + df^n \wedge da_n \,.$$

*Moreover, the functions $a_i$ can be found by "finite" operations and quadrature.*

PROOF: By hypothesis, the forms $df^1, \ldots, df^n$ are linearly independent at every point of $M$, so it follows that the Hamiltonian vector fields $X_{f^1}, \ldots, X_{f^n}$ are also linearly independent at every point of $M$. Also by hypothesis, the functions $f^i$ are in involution, so it follows that $df^i(X_{f^j}) = 0$ for all $i$ and $j$.

The vector fields $X_{f^i}$ are linearly independent on $M$, so by "finite" operations, we can construct 1-forms $\bar\beta_1, \ldots, \bar\beta_n$ which satisfy the conditions

$$\bar\beta_i(X_{f^j}) = \delta_{ij} \qquad \text{(Kronecker delta)}.$$

Any other set of forms $\beta_i$ which satisfy these conditions are given by expressions:

$$\beta_i = \bar\beta_i + g_{ij}\, df^j \,.$$

for some functions $g_{ij}$ on $M$. Let us regard the functions $g_{ij}$ as unknowns for a moment. Let $Y_1, \ldots, Y_n$ be the vector fields which satisfy

$$Y_i \lrcorner\, \Omega = \beta_i \,,$$

with $\bar Y_i$ denoting the corresponding quantities when the $g_{ij}$ are set to zero. Then it is easy to see that

$$Y_i = \bar Y_i - g_{ij}\, X_{f^j} \,.$$

Now, by construction,

$$\Omega(X_{f^i}, X_{f^j}) = 0 \qquad \text{and} \qquad \Omega(Y_i, X_{f^j}) = \delta_{ij} \,.$$

Moreover, as is easy to compute,

$$\Omega(Y_i, Y_j) = \Omega(\bar Y_i, \bar Y_j) - g_{ji} + g_{ij}.$$

L.5.13

Thus, choosing the functions $g_{ij}$ appropriately, say $g_{ij} = -\frac{1}{2}\Omega(\bar{Y}_i, \bar{Y}_j)$, we may assume that $\Omega(Y_i, Y_j) = 0$. It follows that the sequence of 1-forms $df^1, \ldots, df^n, \beta_1, \ldots, \beta_n$ is the dual basis to the sequence of vector fields $Y_1, \ldots, Y_n, X_{f^1}, \ldots, X_{f^n}$. In particular, we see that

$$\Omega = df^1 \wedge \beta_1 + \cdots + df^n \wedge \beta_n,$$

since the 2-forms on either side of this equation have the same values on all pairs of vector fields drawn from this basis.

Now, since $\Omega$ is closed, we have

$$d\Omega = df^1 \wedge d\beta_1 + \cdots + df^n \wedge d\beta_n = 0.$$

If, for example, we wedge both sides of this equation with $df^2, \ldots, df^n$, we see that

$$df^1 \wedge df^2 \wedge \ldots \wedge df^n \wedge d\beta_1 = 0.$$

Hence, it follows that $d\beta_1$ lies in the ideal generated by the forms $df^1, \ldots, df^n$. Of course, there was nothing special about the first term, so we clearly have

$$d\beta_i \equiv 0 \bmod df^1, \ldots, df^n \qquad \text{for all } 1 \le i \le n.$$

In particular, it follows that, if we pull back the 1-forms $\beta_i$ to any $n$-dimensional level set $M_c \subset M$ defined by equations $f^i = c^i$ where the $c^i$ are constants, then each $\beta_i$ becomes closed.

Let $m \in M$ be fixed and choose functions $g_1, \ldots, g_n$ on a neighborhood $U$ of $m$ in $M$ so that $g_i(m) = 0$ and so that the functions $g_1, \ldots, g_n, f^1, \ldots, f^n$ form a coordinate chart on $U$. By shrinking $U$ if necessary, we may assume that the image of this coordinate chart in $\mathbb{R}^n \times \mathbb{R}^n$ is an open set of the form $B_1 \times B_2$, where $B_1$ and $B_2$ are open balls in $\mathbb{R}^n$ (with $B_1$ centered on 0). In this coordinate chart, the $\beta_i$ can be expressed in the form

$$\beta_i = B_i^j(g, f)\, dg_j + C_{ij}(g, f)\, df^j.$$

Define new functions $a_i$ on $B^1 \times B^2$ by the rule

$$h_i(g, f) = \int_0^1 B_i^j(tg, f) g_j \, dt.$$

(This is just the Poincaré homotopy formula with the $f$'s held fixed. It is also the first place where we use "quadrature".) Since setting the $f$'s equals to constants makes $\beta_i$ a closed 1-form, it follows easily that

$$\beta_i = dh_i + A_{ij}(g, f)\, df^j$$

for some functions $A_{ij}$ on $B^1 \times B^2$. Thus, on $U$, the form $\Omega$ has the expression

$$\Omega = df^i \wedge dh_i + A_{ij}\, df^i \wedge df^j.$$

L.5.14

It follows that the 2-form $A = A_{ij}\, df^i \wedge df^j$ is closed on (the contractible open set) $B^1 \times B^2$. Thus, the functions $A_{ij}$ do not depend on the $g$-coordinates at all. Hence, by employing quadrature once more (i.e., the second time) in the Poincaré homotopy formula, we can write $A = -d(s_i\, df^i)$ for some functions $s_i$ of the $f$'s alone. Setting $a_i = h_i + s_i$, we have the desired local normal form $\Omega = df^i \wedge da_i$. $\qquad\qquad\qquad\qquad\Box$

In many useful situations, one does not need to restrict to a local neighborhood $U$ to define the functions $a_i$ (at least up to additive constants) and the 1-forms $da_i$ can be defined globally on $M$ (or, at least away from some small subset in $M$ where degeneracies occur). In this case, the construction above is often called the construction of "action-angle" coordinates. We will discuss this further in Lecture 6.

# Exercise Set 5:

# Symplectic Manifolds, I

**1.** Show that the bilinear form on $\mathbb{R}^n$ defined in the text by the rule $B(x, y) = {}^t x\, A\, y$ (where $A$ is a skew-symmetric $n$-by-$n$ matrix) is non-degenerate if and only if $A$ is invertible. Show directly (i.e., without using Proposition 1) that a skew-symmetric, $n$-by-$n$ matrix $A$ cannot be invertible if $n$ is odd. (Hint: For the last part, compute $\det(A)$ two ways.)

**2.** Let $(V, B)$ be a symplectic space and let $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ be a basis of $B$. Define the $m$-by-$m$ skew-symmetric matrix $A_{\mathbf{b}}$ whose $ij$-entry is $B(b_i, b_j)$. Show that if $\mathbf{b}' = \mathbf{b}R$ is any other basis of $V$ (where $R \in \mathrm{GL}(m, \mathbb{R})$), then

$$A_{\mathbf{b}'} = {}^t R\, A_{\mathbf{b}}\, R.$$

Use Proposition 1 and Exercise 1 to conclude that, if $A$ is an invertible, skew-symmetric $2n$-by-$2n$ matrix, then there exists a matrix $R \in \mathrm{GL}(2n, \mathbb{R})$ so that

$$A = {}^t R \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix} R.$$

In other words, the $\mathrm{GL}(2n, \mathbb{R})$-orbit of the matrix $J_n$ defined in the text (under the "standard" (right) action of $\mathrm{GL}(2n, \mathbb{R})$ on the skew-symmetric $2n$-by-$2n$ matrices) is the open set of all invertible skew-symmetric $2n$-by-$2n$ matrices.

**3.** Show that $\mathrm{Sp}(n, \mathbb{R})$, as defined in the text, is indeed a Lie subgroup of $\mathrm{GL}(2n, \mathbb{R})$ and has dimension $2n^2 + n$. Compute its Lie algebra $\mathfrak{sp}(n, \mathbb{R})$. Show that $\mathrm{Sp}(1, \mathbb{R}) = \mathrm{SL}(2, \mathbb{R})$.

**4.** In Lecture 2, we defined the groups $\mathrm{GL}(n, \mathbb{C}) = \{R \in \mathrm{GL}(2n, \mathbb{R}) \mid J_n\, R = R\, J_n\}$ and $\mathrm{O}(2n) = \{R \in \mathrm{GL}(2n, \mathbb{R}) \mid {}^t R\, R = I_{2n}\}$. Show that

$$\mathrm{GL}(n, \mathbb{C}) \cap \mathrm{Sp}(n, \mathbb{R}) = \mathrm{O}(2n) \cap \mathrm{Sp}(n, \mathbb{R}) = \mathrm{GL}(n, \mathbb{C}) \cap \mathrm{O}(2n) = \mathrm{U}(n).$$

**5.** Let $\Omega$ be a symplectic form on a vector space $V$ of dimension $2n$. Let $W \subset V$ be a subspace which satisfies $\dim W = d$ and $\dim(W \cap W^{\perp}) = \nu$. Show that there exists an $\Omega$-standard basis of $V$ so that $W$ is spanned by the vectors

$$e_1, \ \ldots, \ e_{\nu+m}, \ f_1, \ \ldots, \ f_m$$

where $d - \nu = 2m$. In this basis of $V$, what is a basis for $W^{\perp}$?

**6.** THE PFAFFIAN. Let $V$ be a vector space of dimension $2n$. Fix a basis $\mathbf{b} = (b_1, \ldots, b_{2n})$. For any skew-symmetric $2n$-by-$2n$ matrix $F = (f^{ij})$, define the 2-vector

$$\Phi_F = \tfrac{1}{2} f^{ij} b_i \wedge b_j = \tfrac{1}{2}\mathbf{b} \wedge F \wedge {}^t\mathbf{b}.$$

Then there is a unique polynomial function Pf, homogeneous of degree $n$, on the space of skew-symmetric $2n$-by-$2n$ matrices for which

$$(\Phi_F)^n = n! \operatorname{Pf}(F)\, b_1 \wedge \ldots \wedge b_{2n}.$$

Show that

$$\operatorname{Pf}(F) = f^{12} \qquad\qquad\qquad\qquad \text{when } n = 1,$$
$$\operatorname{Pf}(F) = f^{12} f^{34} + f^{13} f^{42} + f^{14} f^{23} \qquad \text{when } n = 2.$$

Show also that

$$\operatorname{Pf}(A\,F\,{}^tA) = \det(A)\operatorname{Pf}(F)$$

for all $A \in \operatorname{GL}(2n, \mathbb{R})$. (Hint: Examine the effect of a change of basis $\mathbf{b} = \mathbf{b}'A$. Compare Problem 2.) Use this to conclude that $\operatorname{Sp}(n, \mathbb{R})$ is a subgroup of $\operatorname{SL}(2n, \mathbb{R})$. Finally, show that $(\operatorname{Pf}(F))^2 = \det(F)$. (Hint: Show that the left and right hand sides are polynomial functions which agree on a certain open set in the space of skew-symmetric $2n$-by-$2n$ matrices.)

The polynomial function Pf is called the *Pfaffian*. It plays an important role in differential geometry.

**7.** Verify that, for any $B \in A^2(V)$, the symplectic reduction $(\overline{V}, \overline{B})$ is a well-defined symplectic space.

**8.** Show that if there is a $G$-invariant non-degnerate pairing $(\ ,\ )\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, then $\mathfrak{g}$ and $\mathfrak{g}^*$ are isomorphic as $G$-representations.

**9.** Compute the adjoint and coadjoint representations for

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \;\middle|\; a \in \mathbb{R}^+, b \in \mathbb{R} \right\}$$

Show that $\mathfrak{g}$ and $\mathfrak{g}^*$ are *not* isomorphic as $G$-spaces! (For a general $G$, the Ad-orbits of $G$ in $\mathfrak{g}$ are not even of even dimension in general, so they can't be symplectic manifolds.)

**10.** For any Lie group $G$ and any $\xi \in \mathfrak{g}^*$, show that the symplectic structures $\Omega_\xi$ and $\Omega_{a \cdot \xi}$ on $G \cdot \xi$ are the same for any $a \in G$.

**11.** This exercise concerns the splitting properties of the two Lie algebras sequences associated to any symplectic structure $\Omega$ on a connected manifold $M$:

$$0 \longrightarrow \mathbb{R} \longrightarrow C^\infty(M) \longrightarrow \mathfrak{h}(\Omega) \longrightarrow 0$$

and

$$0 \longrightarrow \mathfrak{h}(\Omega) \longrightarrow \mathfrak{sp}(\Omega) \longrightarrow H^1_{dR}(M, \mathbb{R}) \longrightarrow 0.$$

Define the "divided powers" of $\Omega$ by the rule $\Omega^{[k]} = (1/k!)\,\Omega^k$, for each $0 \le k \le n$.

(i) Show that, for any vector fields $X$ and $Y$ on $M$,

$$\Omega(X, Y)\,\Omega^{[n]} = -(X \lrcorner \Omega) \wedge (Y \lrcorner \Omega) \wedge \Omega^{[n-1]}.$$

Conclude that the first of the above two sequences splits if $M$ is compact. (Hint: For the latter statement, show that the set of functions $f$ on $M$ for which $\int_M f\,\Omega^{[n]} = 0$ forms a Poisson subalgebra of $C^\infty(M)$.)

(ii) On the other hand, show that for $\mathbb{R}^2$ with the symplectic structure $\Omega = dx \wedge dy$, the first sequence does *not* split. (Hint: Show that every smooth function on $\mathbb{R}^2$ is of the form $\{x, g\}$ for some $g \in C^\infty(\mathbb{R}^2)$. Why does this help?)

(iii) Suppose that $M$ is compact. Define a skew-symmetric pairing

$$\beta_\Omega \colon H^1_{dR}(M, \mathbb{R}) \times H^1_{dR}(M, \mathbb{R}) \to \mathbb{R}$$

by the formula

$$\beta_\Omega(a, b) = \int_M \tilde{a} \wedge \tilde{b} \wedge \Omega^{[n-1]},$$

where $\tilde{a}$ and $\tilde{b}$ are closed 1-forms representing the cohomology classes $a$ and $b$ respectively. Show that if there is a Lie algebra splitting $\sigma \colon H^1_{dR}(M, \mathbb{R}) \to \mathfrak{sp}(\Omega)$ then

$$\Omega\big(\sigma(a), \sigma(b)\big) = -\frac{\beta_\Omega(a, b)}{vol(M, \Omega^{[n]})}$$

for all $a, b \in H^1_{dR}(M, \mathbb{R})$. (Remember that the Lie algebra structure on $H^1_{dR}(M, \mathbb{R})$ is the abelian one.) Use this to conclude that the second sequence does split for a symplectic structure on the standard 1-holed torus, but does *not* split for any symplectic structure on the 2-holed torus. (Hint: To show the non-splitting result, use the fact that any tangent vector field on the 2-holed torus must have a zero.)

**12**. THE FLUX HOMOMORPHISM. The object of this exercise is to try to identify the subgroup of $\mathsf{Sp}(\Omega)$ whose Lie algebra is $\mathsf{h}(\Omega)$. Thus, let $(M, \Omega)$ be a symplectic manifold.

First, I remind you how the construction of the (smooth) universal cover of the identity component of $\mathsf{Sp}(\Omega)$ goes. Let $p \colon [0,1] \times M \to M$ be a smooth map with the property that the map $p_t \colon M \to M$ defined by $p_t(m) = p(t, m)$ is a symplectomorphism for all $0 \le t \le 1$. Such a $p$ is called a (smooth) path in $\mathsf{Sp}(M)$. We say that $p$ is *based at the identity map* $\mathsf{e} \colon M \to M$ if $p_0 = \mathsf{e}$. The set of smooth paths in $\mathsf{Sp}(M)$ which are based at $\mathsf{e}$ will be denoted by $\mathsf{P}_\mathsf{e}\big(\mathsf{Sp}(\Omega)\big)$.

Two paths $p$ and $p'$ in $\mathsf{P}_\mathsf{e}\big(\mathsf{Sp}(\Omega)\big)$ satisfying $p_1 = p_1'$ are said to be *homotopic* if there is a smooth map $P \colon [0,1] \times [0,1] \times M \to M$ which satisfies the following conditions: First, $P(s, 0, m) = m$ for all $s$ and $m$. Second, $P(s, 1, m) = p_1(m) = p_1'(m)$ for all $s$ and $m$. Third, $P(0, t, m) = p(t, m)$ and $P(1, t, m) = p'(t, m)$ for all $t$ and $m$. The set of homotopy classes of elements of $\mathsf{P}_\mathsf{e}\big(\mathsf{Sp}(\Omega)\big)$ is then denoted by $\widetilde{\mathsf{Sp}}^0(\Omega)$. In any reasonable topology on $\mathsf{Sp}(\Omega)$, this should to be the universal covering space of the identity component of $\mathsf{Sp}(\Omega)$. There is a natural group structure on $\widetilde{\mathsf{Sp}}^0(\Omega)$ in which $\tilde{\mathsf{e}}$, the homotopy class of the constant path at $\mathsf{e}$, is the identity element (cf., the covering spaces exercise in Exercise Set 2).

We are now going to construct a homomorphism $\Phi \colon \widetilde{\mathsf{Sp}}^0(\Omega) \to H^1(M, \mathbb{R})$, called the *flux homomorphism*. Let $p \in \mathsf{P}_\mathsf{e}\big(\mathsf{Sp}(\Omega)\big)$ be chosen, and let $\gamma \colon S^1 \to M$ be a closed curve representing an element of $H_1(M, \mathbb{Z})$. Then we can define

$$F(p, \gamma) = \int_{[0,1] \times S^1} (p \cdot \gamma)^*(\Omega)$$

where $(p \cdot \gamma) \colon [0,1] \times S^1 \to M$ is defined by $(p \cdot \gamma)(t, \theta) = p\big(t, \gamma(\theta)\big)$. The number $F(p, \gamma)$ is called the *flux* of $p$ through $\gamma$.

(i) Show that $F(p, \gamma) = F(p', \gamma')$ if $p$ is homotopic to $p'$ and $\gamma$ is homologous to $\gamma'$). (Hint: Use Stokes' Theorem several times.)

Thus, $F$ is actually well defined as a map $F \colon \widetilde{\mathsf{Sp}}^0(\Omega) \times H_1(M, \mathbb{R}) \to \mathbb{R}$.

(ii) Show that $F \colon \widetilde{\mathsf{Sp}}^0(\Omega) \times H_1(M, \mathbb{R}) \to \mathbb{R}$ is linear in its second variable and that, under the obvious multiplication, we have

$$F(p\,p', \gamma) = F(p, \gamma) + F(p', \gamma).$$

(Hint: Use Stokes' Theorem again.)

Thus, $F$ may be transposed to become a homomorphism

$$\Phi \colon \widetilde{\mathsf{Sp}}^0(\Omega) \to H^1(M, \mathbb{R}).$$

Show (by direct computation) that if $\zeta$ is a closed 1-form on $M$ for which the symplectic vector field $Z = \sharp\zeta$ is complete on $M$, then the path $p$ in $\mathsf{Sp}(M)$ defined by the flow of $Z$ from $t = 0$ to $t = 1$ satisfies $\Phi(p) = [\zeta] \in H^1(M, \mathbb{R})$. Conclude that the flux homomorphism $\Phi$ is always surjective and that its derivative $\Phi'(\tilde{\mathsf{e}}) \colon \mathsf{sp}(\Omega) \to H^1(M, \mathbb{R})$ is just the operation of taking cohomology classes. (Recall that we identify $\mathsf{sp}(\Omega)$ with $\mathcal{Z}^1(M)$.)

E.5.4                                                                                               98

(iii) Show that if $M$ is a compact surface of genus $g > 1$, then the flux homomorphism is actually well defined as a map from $\mathsf{Sp}(\Omega)$ to $H^1(M, \mathbb{R})$. Would the same result be true if $M$ were of genus 1? How could you modify the map so as to make it well-defined on $\mathsf{Sp}(\Omega)$ in the case of the torus? (Hint: Show that if you have two paths $p$ and $p'$ with the same endpoint, then you can express the difference of their fluxes across a circle $\gamma$ as an integral of the form

$$\int_{S^1 \times S^1} \Psi^*(\Omega)$$

where $\Psi \colon S^1 \times S^1 \to M$ is a certain piecewise smooth map from the torus into $M$. Now use the fact that, for any piecewise smooth map $\Psi \colon S^1 \times S^1 \to M$, the induced map $\Psi^* \colon H^2(M, \mathbb{R}) \to H^2(S^1 \times S^1, \mathbb{R})$ on cohomology is zero. (Why does this follow from the assumption that the genus of $M$ is greater than 1?) )

In any case, the subgroup $\ker \Phi$ (or its image under the natural projection from $\widetilde{\mathsf{Sp}}^0(\Omega)$ to $\mathsf{Sp}(\Omega)$) is known as the group $\mathsf{H}(\Omega)$ of *exact* or *Hamiltonian* symplectomorphisms. Note that, at least formally, its Lie algebra is $\mathsf{h}(\Omega)$.

**13**. In the case of the geodesic flow on a surface of revolution (see Lecture 4), show that the energy $f^1 = E_L$ and the conserved quantity $f^2 = F(r)\dot{r} + G(r)\dot{\theta}$ are in involution. Use the algorithm described in Theorem 3 to compute the functions $a_1$ and $a_2$, thus verifying that the geodesic equations on a surface of revolution are integrable by quadrature.

E.5.5

# Lecture 6:

# Symplectic Manifolds, II

**The Space of Symplectic Structures on $M$.**

I want to turn now to the problem of describing the symplectic structures a manifold $M$ can have. This is a surprisingly delicate problem and is currently a subject of research.

Of course, one fundamental question is whether a given manifold has any symplectic structures at all. I want to begin this lecture with a discussion of the two known obstructions for a manifold to have a symplectic structure.

**The cohomology ring condition.** If $\Omega \in \mathcal{A}^2(M^{2n})$ is a symplectic structure on a compact manifold $M$, then the cohomology class $[\Omega] \in H^2_{dR}(M, \mathbb{R})$ is non-zero. In fact, $[\Omega]^n = [\Omega^n]$, but the class $[\Omega^n]$ cannot vanish in $H^{2n}_{dR}(M)$ because the integral of $\Omega^n$ over $M$ is clearly non-zero. Thus, we have

**Proposition 1:** *If $M^{2n}$ is compact and has a symplectic structure, there must exist an element $u \in H^2(M, \mathbb{R})$ so that $u^n \neq 0 \in H^{2n}_{dR}(M)$.*

**Example.** This immediately rules out the existence of a symplectic structure on $S^{2n}$ for all $n > 1$. One consequence of this, as you are asked to show in the Exercises, is that there cannot be any simple notion of connected sum in the category of symplectic manifolds (except in dimension 2).

**The bundle obstruction.** If $M$ admits a symplectic structure $\Omega$, then, in particular, this defines a symplectic structure on each of the tangent spaces $T_m M$ which varies continuously with $m$. In other words, $TM$ must carry the structure of a symplectic vector bundle. There are topological obstructions to the existence of such a structure on the tangent bundle of a general manifold. As a simple example, if $M$ has a symplectic structure, then $TM$ must be orientable.

There are more subtle obstructions than orientation. Unfortunately, a description of these obstructions requires some acquaintance with the theory of characteristic classes. However, part of the following discussion will be useful even to those who aren't familiar with characteristic class theory, so I will give it now, even though the concepts will only reveal their importance in later Lectures.

**Definition 1:** An *almost symplectic structure* on a manifold $M^{2n}$ is a smooth 2-form $\Omega$ defined on $M$ which is non-degenerate but not necessarily closed. An *almost complex structure* on $M^{2n}$ is a smooth bundle map $J: TM \to TM$ which satisfies $J^2 v = -v$ for all $v$ in $TM$.

The reason that I have introduced both of these concepts at the same time is that they are intimately related. The really deep aspects of this relationship will only become apparent in the Lecture 9, but we can, at least, give the following result now.

L.6.1                                                                                                100

**Proposition 2:** *A manifold $M^{2n}$ has an almost symplectic structure if and only if it has an almost complex structure.*

PROOF: First, suppose that $M$ has an almost complex structure $J$. Let $g_0$ be any Riemannian metric on $M$. (Thus, $g_0 : TM \to \mathbb{R}$ is a smooth function which restricts to each $T_m M$ to be a positive definite quadratic form.) Now define a new Riemannian metric by the formula

$$g(v) = g_0(v) + g_0(Jv).$$

Then $g$ has the property that $g(Jv) = g(v)$ for all $v \in TM$ since

$$g(Jv) = g_0(Jv) + g_0(J^2 v) = g_0(Jv) + g_0(-v) = g(v).$$

Now let $\langle , \rangle$ denote the (symmetric) inner product associated with $g$. Thus, $\langle v, v \rangle = g(v)$, so we have $\langle Jx, Jy \rangle = \langle x, y \rangle$ when $x$ and $y$ are tangent vector with the same base point. For $x, y \in T_m M$ define $\Omega(x, y) = \langle Jx, y \rangle$. I claim that $\Omega$ is a non-degenerate 2-form on $M$. To see this, first note that

$$\Omega(x, y) = \langle Jx, y \rangle = -\langle Jx, J^2 y \rangle = -\langle J^2 y, Jx \rangle = -\langle Jy, x \rangle = -\Omega(y, x),$$

so $\Omega$ is a 2-form. Moreover, if $x$ is a non-zero tangent vector, then $\Omega(x, Jx) = \langle Jx, Jx \rangle = g(x) > 0$, so it follows that $x \lrcorner \Omega \neq 0$. Thus $\Omega$ is non-degenerate.

To go the other way is a little more delicate. Suppose that $\Omega$ is given and fix a Riemannian metric $g$ on $M$ with associated inner product $\langle , \rangle$. Then, by linear algebra there exists a unique bundle mapping $A : TM \to TM$ so that $\Omega(x, y) = \langle Ax, y \rangle$. Since $\Omega$ is skew-symmetric and non-degenerate, it follows that $A$ must be skew-symmetric relative to $\langle , \rangle$ and must be invertible. It follows that $-A^2$ must be symmetric and positive definite relative to $\langle , \rangle$ .

Now, standard results from linear algebra imply that there is a unique smooth bundle map $B : TM \to TM$ which positive definite and symmetric with respect to $\langle , \rangle$ and which satisfies $B^2 = -A^2$. Moreover, this linear mapping $B$ must commute with $A$. (See the Exercises if you are not familiar with this fact). Thus, the mapping $J = AB^{-1}$ satisfies $J^2 = -I$, as desired. $\qquad \square$

It is not hard to show that the mappings $(J, g_0) \mapsto \Omega$ and $(\Omega, g) \mapsto J$ constructed in the proof of Proposition 1 depend continuously (in fact, smoothly) on their arguments. Since the set of Riemannian metrics on $M$ is contractible, it follows that the set of homotopy classes of almost complex structures on $M$ is in natural one-to-one correspondence with the set of homotopy classes of almost symplectic structures.

(The reader who is familiar with the theory of principal bundles knows that at the heart of Proposition 1 is the fact that $\mathrm{Sp}(n, \mathbb{R})$ and $\mathrm{GL}(n, \mathbb{C})$ have the same maximal compact subgroup, namely $\mathrm{U}(n)$.)

L.6.2

Now I can describe some of the bundle obstructions. Suppose that $M$ has a symplectic structure $\Omega$ and let $J$ be any one of the almost complex structures on $M$ we constructed above. Then the tangent bundle of $M$ can be regarded as a complex bundle, which we will denote by $T^J$ and hence has a total Chern class

$$c(T^J) = \big(1 + c_1(J) + c_2(J) + \cdots + c_n(J)\big)$$

where $c_i(J) \in H^{2i}(M, \mathbb{Z})$. Now, by the properties of Chern classes, $c_n(J) = e(TM)$, where $e(TM)$ is the Euler class of the tangent bundle given the orientation determined by the volume form $\Omega^n$.

These classes are related to the Pontrijagin classes of $TM$ by the Whitney sum formula (see [MS]):

$$
\begin{aligned}
p(TM) &= 1 - p_1(TM) + p_2(TM) - \cdots + (-1)^{[n/2]} p_{[n/2]}(TM) \\
&= c(T^J \oplus T^{-J}) \\
&= \big(1 + c_1(J) + c_2(J) + \cdots + c_n(J)\big)\big(1 - c_1(J) + c_2(J) - \cdots + (-1)^n c_n(J)\big)
\end{aligned}
$$

Since $p(TM)$ depends only on the diffeomorphism class of $M$, this gives quadratic equations for the $c_i(J)$,

$$p_k(T) = \big(c_k(J)\big)^2 - 2c_{k-1}(J)c_{k+1}(J) + \cdots + (-1)^k 2c_0(J)c_{2k}(J),$$

to which any manifold with an almost complex structure must have solutions. Since not every $2n$-manifold has cohomology classes $c_i(J)$ satisfying these equations, it follows that some $2n$-manifolds have no almost complex structure and hence, by Proposition 2, no almost symplectic structure either.

**Examples.** Here are two examples in dimension 4 to show that the cohomology ring condition and the bundle obstruction are independent.

- $M = S^1 \times S^3$ does not have a symplectic structure because $H^2(M, \mathbb{R}) = 0$. However the bundle obstruction vanishes because $M$ is parallelizable (why?). Thus $M$ does have an almost symplectic structure.

- $M = \mathbb{CP}^2 \# \mathbb{CP}^2$. The cohomology ring of $M$ in this case is generated over $\mathbb{Z}$ by two generators $u_1$ and $u_2$ in $H^2(M, \mathbb{Z})$ which are subject to the relations $u_1 u_2 = 0$ and $u_1^2 = u_2^2 = v$ where $v$ generates $H^4(M, \mathbb{Z})$. For any non-zero class $u = n_1 u_1 + n_2 u_2$, we have $u^2 = (n_1^2 + n_2^2)v \neq 0$. Thus the cohomology ring condition is satisfied.

However, $M$ has no almost symplectic structure: If it did, then $T = TM$ would have a complex structure $J$, with total Chern class $c(J)$ and the equations above would give $p_1(T) = \big(c_1(J)\big)^2 - 2c_2(J)$. Moreover, we would have $e(T) = c_2(J)$. Thus, we would have to have

$$\big(c_1(J)\big)^2 = p_1(T) + 2e(T).$$

For any compact, simply-connected, oriented 4-manifold $M$ with orientation class $\mu \in H^4(M, \mathbb{Z})$, the Hirzebruch Signature Theorem (see [MS]) implies $p_1(T) = 3(b_2^+ - b_2^-)\mu$, where $b_2^\pm$ are the number of positive and negative eigenvalues respectively of the intersection pairing $H^2(M, \mathbb{Z}) \times H^2(M, \mathbb{Z}) \to \mathbb{Z}$. In addition, $e(T) = (2 + b_2^+ + b_2^-)\mu$. Substituting these into the above formula, we would have $\left(c_1(J)\right)^2 = (4 + 5b_2^+ - b_2^-)\mu$ for any complex structure $J$ on the tangent bundle of $M$.

In particular, if $M = \mathbb{CP}^2 \,\#\, \mathbb{CP}^2$ had an almost complex structure $J$, then $\left(c_1(J)\right)^2$ would be either $14v$ (if $\mu = v$, since then $b_2^+ = 2$ and $b_2^- = 0$) or $-2v$ (if $\mu = -v$, since then $b_2^+ = 0$ and $b_2^- = 2$). However, by our previous calculations, neither $14v$ nor $-2v$ is the square of a cohomology class in $H^2(\mathbb{CP}^2 \,\#\, \mathbb{CP}^2, \mathbb{Z})$.

▶     *This example shows that, in general, one cannot hope to have a connected sum operation for symplectic manifolds.*

The actual conditions for a manifold to have an almost symplectic structure can be expressed in terms of characteristic classes, so, in principle, this can always be determined once the manifold is given explicitly. In Lecture 9 we will describe more fully the following remarkable result of Gromov:

▶     *If $M^{2n}$ has no compact components and has an almost symplectic structure $\Upsilon$, then there exists a symplectic structure $\Omega$ on $M$ which is homotopic to $\Upsilon$ through almost symplectic structures.*

Thus, the problem of determining which manifolds have symplectic structures is now reduced to the compact case. In this case, no obstruction beyond what I have already described is known. Thus, I can state the following:

**Basic Open Problem:** *If a compact manifold $M^{2n}$ satisfies the cohomology ring condition and has an almost symplectic structure, does it have a symplectic structure?*

Even (perhaps especially) for 4-manifolds, this problem is extremely interesting and very poorly understood.

**Deformations of Symplectic Structures.** We will now turn to some of the features of the space of symplectic structures on a given manifold which does admit symplectic structures. First, we will examine the "deformation problem". The following theorem due to Moser (see [We]) shows that symplectic structures determining a fixed cohomology class in $H^2$ on a compact manifold are "rigid".

**Theorem 1:** *If $M^{2n}$ is a compact manifold and $\Omega_t$ for $t \in [0, 1]$ is a continuous 1-parameter family of smooth symplectic structures on $M$ which has the property that the cohomology classes $[\Omega_t]$ in $H^2_{dR}(M, \mathbb{R})$ are independent of $t$, then for each $t \in [0, 1]$, there exists a diffeomorphism $\phi_t$ so that $\phi_t^*(\Omega_t) = \Omega_0$.*

PROOF: We will start by proving a special case and then deduce the general case from it. Suppose that $\Omega_0$ is a symplectic structure on $M$ and that $\varphi \in \mathcal{A}^1(M)$ is a 1-form so that, for all $s \in (-1, 1)$, the 2-form

$$\Omega_s = \Omega_0 + s\,d\varphi$$

is a symplectic form on $M$ as well. (This is true for all sufficiently "small" 1-forms on $M$ since $M$ is compact.) Now consider the 2-form on $(-1, 1) \times M$ defined by the formula

$$\Omega = \Omega_0 + s\,d\varphi - \varphi \wedge ds.$$

(Here, we are using $s$ as the coordinate on the first factor $(-1, 1)$ and, as usual, we write $\Omega_0$ and $\varphi$ instead of $\pi_2^*(\Omega_0)$ and $\pi_2^*(\varphi)$ where $\pi_2\colon (-1, 1) \times M \to M$ is the projection on the second factor.)

The reader can check that $\Omega$ is closed on $(-1, 1) \times M$. Moreover, since $\Omega$ pulls back to each slice $\{s_0\} \times M$ to be the non-degenerate form $\Omega_{s_0}$ it follows that $\Omega$ has half-rank $n$ everywhere. Thus, the kernel $N_\Omega$ is 1-dimensional and is transverse to each of the slices $\{t\} \times M$. Hence there is a unique vector field $X$ which spans $N_\Omega$ and satisfies $ds(X) = 1$.

Now because $M$ is compact, it is not difficult to see that each integral curve of $X$ projects by $s = \pi_1$ diffeomorphically onto $(-1, 1)$. Moreover, it follows that there is a smooth map $\phi\colon (-1, 1) \times M \to M$ so that, for each $m$, the curve $t \mapsto \phi(t, m)$ is the integral curve of $X$ which passes through $(0, m)$.

It follows that the map $\Phi\colon (-1, 1) \times M \to (-1, 1) \times M$ defined by

$$\Phi(t, m) = \big(t, \phi(t, m)\big)$$

carries the vector field $\partial/\partial s$ to the vector field $X$. Moreover, since $\Omega_0$ and $\Omega$ have the same value when pulled back to the slice $\{0\} \times M$ and since

$$\begin{array}{ccc} \mathfrak{L}_{\partial/\partial s}\Omega_0 = 0 & & \mathfrak{L}_X\Omega = 0 \\ & \text{and} & \\ \partial/\partial s \lrcorner\, \Omega_0 = 0 & & X \lrcorner\, \Omega = 0, \end{array}$$

it follows easily that $\Phi^*(\Omega) = \Omega_0$. In particular, $\phi_t^*(\Omega_t) = \Omega_0$ where $\phi_t$ is the diffeomorphism of $M$ given by $\phi_t(m) = \phi(t, m)$.

Now let us turn to the general case. If $\Omega_t$ for $0 \le t \le 1$ is any continuous family of smooth closed 2-forms for which the cohomology classes $[\Omega_t]$ are all equal to $[\Omega_0]$, then for any two values $t_1$ and $t_2$ in the unit interval, consider the 1-parameter family of 2-forms

$$\Upsilon_s = (1-s)\Omega_{t_1} + s\Omega_{t_2}.$$

Using the compactness of $M$, it is not difficult to show that for $t_2$ sufficiently close to $t_1$, the family $\Upsilon_s$ is a 1-parameter family of symplectic forms on $M$ for $s$ in some open interval containing $[0, 1]$. Moreover, by hypothesis, $[\Omega_{t_2} - \Omega_{t_1}] = 0$, so there exists a 1-form $\varphi$ on $M$ so that $d\varphi = \Omega_{t_2} - \Omega_{t_1}$. Thus,

$$\Upsilon_s = \Omega_{t_1} + s\,d\varphi.$$

By the special case already treated, there exists a diffeomorphism $\phi_{t_2,t_1}$ of $M$ so that $\phi_{t_2,t_1}^*(\Omega_{t_2}) = \Omega_{t_1}$.

Finally, using the compactness of the interval $[0,t]$ for any $t \in [0,1]$, we can subdivide this interval into a finite number of intervals $[t_1, t_2]$ on which the above argument works. Then, by composing diffeomorphisms, we can construct a diffeomorphism $\phi_t$ of $M$ so that $\phi_t^*(\Omega_t) = \Omega_0$. $\hspace{2cm}$ $\square$

The reader may have wanted the family of diffeomorphisms $\phi_t$ to depend continuously on $t$ and smoothly on $t$ if the family $\Omega_t$ is smooth in $t$. This can, in fact, be arranged. However, it involves showing that there is a smooth family of 1-forms $\varphi_t$ on $M$ so that $\frac{d}{dt}\Omega_t = d\varphi_t$, i.e., smoothly solving the $d$-equation. This can be done, but requires some delicacy or use of elliptic machinery (e.g., Hodge-deRham theory).

▶ $\hspace{1cm}$ *Theorem 1 does not hold without the hypothesis of compactness.* For example, if $\Omega$ is the restriction of the standard structure on $\mathbb{R}^{2n}$ to the unit ball $B^{2n}$, then for the family $\Omega_t = e^t\Omega$ there cannot be any family of diffeomorphisms of the ball $\phi_t$ so that $\phi_t^*(\Omega_t) = \Omega$ since the integrals over $B$ of the volume forms $(\Omega_t)^n = e^{nt}\Omega^n$ are all different.

Intuitively, Theorem 1 says that the "connected components" of the space of symplectic structures on a manifold are orbits of the group $\text{Diff}^0(M)$ of diffeomorphisms isotopic to the identity. (The reason this is only intuitive is that we have not actually defined a topology on the space of symplectic structures on $M$.)

It is an interesting question as to how many "connected components" the space of symplectic structures on $M$ has. The work of Gromov has yielded methods to attack this problem and I will have more to say about this in Lecture 9.

### Submanifolds of Symplectic Manifolds

We will now pass on to the study of the geometry of submanifolds of a symplectic manifold. The following result describes the behaviour of symplectic structures near closed submanifolds. This theorem, due to Weinstein (see [Weinstein]), can be regarded as a generalization of Darboux' Theorem. The reader will note that the proof is quite similar to the proof of Theorem 1.

**Theorem 2:** *Let $P \subset M$ be a closed submanifold and let $\Omega_0$ and $\Omega_1$ be symplectic structures on $M$ which have the property that $\Omega_0(p) = \Omega_1(p)$ for all $p \in P$. Then there exist open neighborhoods $U_0$ and $U_1$ of $P$ and a diffeomorphism $\phi\colon U_0 \to U_1$ satisfying $\phi^*(\Omega_1) = \Omega_0$ and which moreover fixes $P$ pointwise and satisfies $\phi'(p) = id_p\colon T_pM \to T_pM$ for all $p \in P$.*

PROOF: Consider the linear family of 2-forms

$$\Omega_t = (1-t)\Omega_0 + t\Omega_1$$

which "interpolates" between the forms $\Omega_0$ and $\Omega_1$. Since $[0, 1]$ is compact and since, by hypothesis, $\Omega_0(p) = \Omega_1(p)$ for all $p \in P$, it easily follows that there is an open neighborhood $U$ of $P$ in $M$ so that $\Omega_t$ is a symplectic structure on $U$ for all $t$ in some open interval $I = (-\varepsilon, 1 + \varepsilon)$ containing $[0, 1]$.

We may even suppose that $U$ is a "tubular neighborhood" of $P$ which has a smooth retraction $R: [0, 1] \times U \to U$ into $P$. Since $\Phi = \Omega_1 - \Omega_0$ vanishes on $P$, it follows without too much difficulty (see the Exercises) that there is a 1-form $\varphi$ on $U$ which vanishes on $P$ and which satisfies $d\varphi = \Phi$.

Now, on $I \times U$, consider the 2-form

$$\Omega = \Omega_0 + s\, d\varphi - \varphi \wedge ds.$$

This is a closed 2-form of half-rank $n$ on $I \times U$. Just as in the previous theorem, it follows that there exists a unique vector field $X$ on $I \times U$ so that $ds(X) = 1$ and $X \lrcorner \Omega = 0$.

Since $\varphi$ and $d\varphi$ vanish on $P$, the vector field $X$ has the property that $X(s, p) = \partial/\partial s$ for all $p \in P$ and $s \in I$. In particular, the set $\{0\} \times P$ lies in the domain of the time 1 flow of $X$. Since this domain is an open set, it follows that there is an open neighborhood $U_0$ of $P$ in $U$ so that $\{0\} \times U_0$ lies in the domain of the time 1 flow of $X$. The image of $\{0\} \times U_0$ under the time 1 flow of $X$ is of the form $\{1\} \times U_1$ where $U_1$ is another open neighborhood of $P$ in $U$.

Thus, the time 1 flow of $X$ generates a diffeomorphism $\phi: U_0 \to U_1$. By the arguments of the previous theorem, it follows that $\phi^*(\Omega_1) = \Omega_0$. I leave it to the reader to check that $\phi$ fixes $P$ in the desired fashion. $\qquad\square$

Theorem 2 has a useful corollary:

**Corollary :** *Let $\Omega$ be a symplectic structure on $M$ and let $f_0$ and $f_1$ be smooth embeddings of a manifold $P$ into $M$ so that $f_0^*(\Omega) = f_1^*(\Omega)$ and so that there exists a smooth bundle isomorphism $\tau: f_0^*(TM) \to f_1^*(TM)$ which extends the identity map on the subbundle $TP \subset f_i^*(TM)$ and which identifies the symplectic structures on $f_i^*(TM)$. Then there exist open neighborhoods $U_i$ of $f_i(P)$ in $M$ and a diffeomorphism $\phi: U_0 \to U_1$ which satisfies $\phi^*(\Omega) = \Omega$ and, moreover, $\phi \circ f_0 = f_1$.*

PROOF: It is an elementary result in differential topology that, under the hypotheses of the Corollary, there exists an open neighborhood $W_0$ of $f_0(P)$ in $M$ and a smooth diffeomorphic embedding $\psi: W_0 \to M$ so that $\psi \circ f_0 = f_1$ and $\psi'\big(f_0(p)\big): T_{f_0(p)}(M) \to T_{f_1(p)}(M)$ is equal to $\tau(p)$. It follows that $\psi^*(\Omega)$ is a symplectic form on $W_0$ which agrees with $\Omega$ along $f_0(P)$. By Theorem 2, it follows that there is a neighborhood $U_0$ of $f_0(P)$ which lies in $W_0$ and a smooth map $\nu: U_0 \to W_0$ which is a diffeomorphism onto its image, fixes $f_0(P)$ pointwise, satisfies $\nu'\big(f_0(p)\big) = id_{f_0(p)}$ for all $p \in P$, and also satisfies $\nu^*\big(\psi^*(\Omega)\big) = \Omega$. Now just take $\phi = \psi \circ \nu$. $\qquad\square$

We will now give two particularly important applications of this result:

If $P \subset M$ is a symplectic submanifold, then by using $\Omega$, we can define a normal bundle for $P$ as follows:

$$\nu(P) = \{(p, v) \in P \times TM \mid v \in T_pM,\ \Omega(v, w) = 0 \text{ for all } w \in T_pP\}.$$

The bundle $\nu(P)$ has a natural symplectic structure on each of its fibers (see the Exercises), and hence is a symplectic vector bundle. The following proposition shows that, up to local diffeomorphism, this normal bundle determines the symplectic structure $\Omega$ on a neighborhood of $P$.

**Proposition 3:**  Let $(P, \Upsilon)$ be a symplectic manifold and let $f_0, f_1 : P \to M$ be two symplectic embeddings of $P$ as submanifolds of $M$ so that the normal bundles $\nu_0(P)$ and $\nu_1(P)$ are isomorphic as symplectic vector bundles. Then there are open neighborhoods $U_i$ of $f_i(P)$ in $M$ and a symplectic diffeomorphism $\phi : U_0 \to U_1$ which satisfies $f_1 = \phi \circ f_0$.

Proof:  It suffices to construct the map $\tau$ required by the hypotheses of Theorem 2. Now, we have a symplectic bundle decomposition $f_i^*(TM) = TP \oplus \nu_i(P)$ for $i = 1, 2$. If $\alpha : \nu_0(P) \to \nu_1(P)$ is a symplectic bundle isomorphism, we then define $\tau = id \oplus \alpha$ in the obvious way and we are done.                                                      □

At the other extreme, we want to consider submanifolds of $M$ to which the form $\Omega$ pulls back to be as degenerate as possible.

**Definition 2:**  If $\Omega$ is a symplectic structure on $M^{2n}$, an immersion $f : P \to M$ is said to be *isotropic* if $f^*(\Omega) = 0$. If the dimension of $P$ is $n$, we say that $f$ is a *Lagrangian* immersion. If in addition, $f$ is one-to-one, then we say that $f(P)$ is a *Lagrangian submanifold* of $M$.

Note that the dimension of an isotropic submanifold of $M^{2n}$ is at most $n$, so the Lagrangian submanifolds of $M$ have maximal dimension among all isotropic submanifolds.

**Example:** Graphs of Symplectic Mappings. If $f : M \to N$ is a symplectic mapping where $\Omega$ and $\Upsilon$ are the symplectic forms on $M$ and $N$ respectively, then the graph of $f$ in $M \times N$ is an isotropic submanifold of $M \times N$ endowed with the symplectic structure $(-\Omega) \oplus \Upsilon = \pi_1^*(-\Omega) + \pi_2^*(\Upsilon)$. If $M$ and $N$ have the same dimension, then the graph of $f$ in $M \times N$ is a Lagrangian submanifold.

**Example:** Closed 1-forms. If $\alpha$ is a 1-form on $M$, then the graph of $\alpha$ in $T^*M$ is a Lagrangian submanifold of $T^*M$ if and only if $d\alpha = 0$. This follows because $\Omega$ on $T^*M$ has the "reproducing property" that $\alpha^*(\Omega) = d\alpha$ for any 1-form on $M$.

**Proposition 4:**  Let $\Omega$ be a symplectic structure on $M$ and let $P$ be a closed Lagrangian submanifold of $M$. Then there exists an open neighborhood $U$ of the zero section in $T^*P$ and a smooth map $\phi : U \to M$ satisfying $\phi(0_p) = p$ which is a diffeomorphism onto an open neighborhood of $P$ in $M$, and which pulls back $\Omega$ to be the standard symplectic structure on $U$.

Proof:  From the earlier proofs, the reader probably can guess what we will do. Let $\iota : P \to M$ be the inclusion mapping and let $\zeta : P \to T^*P$ be the zero section of $T^*P$. I leave as an exercise for the reader to show that $\zeta^*\big(T(T^*P)\big) = TP \oplus T^*P$, and that the induced symplectic structure $\Upsilon$ on this sum is simply the natural one on the sum of a bundle and its dual:
$$\Upsilon\big((v_1, \xi_1), (v_2, \xi_2)\big) = \xi_1(v_2) - \xi_2(v_1)$$

I will show that there is a bundle isomorphism $\tau \colon TP \oplus T^*P \to \iota^*(TM)$ which restricts to the subbundle $TP$ to be $\iota' \colon TP \to \iota^*(TM)$.

First, select an $n$-dimensional subbundle $L \subset \iota^*(TM)$ which is complementary to $\iota'(TP) \subset \iota^*(TM)$. It is not difficult to show (and it is left as an exercise for the reader) that it is possible to choose $L$ so that it is a Lagrangian subbundle of $\iota^*(TM)$ so that there is an isomorphism $\alpha \colon T^*P \to L$ so that $\tau \colon TP \oplus T^*P \to \iota'(TP) \oplus L$ defined by $\tau = \iota' \oplus \alpha$ is a symplectic bundle isomorphism.

Now apply the Corollary to Theorem 2. $\qquad\square$

Proposition 4 shows that the symplectic structure on a manifold $M$ in a neighborhood of a closed Lagrangian submanifold $P$ is completely determined by the diffeomorphism type of $P$. This fact has several interesting applications. We will only give one of them here.

**Proposition 5:** *Let $(M, \Omega)$ be a compact symplectic manifold with $H^1_{dR}(M, \mathbb{R}) = 0$. Then in $\mathrm{Diff}(M)$ endowed with the $C^1$ topology, there exists an open neighborhood $\mathcal{U}$ of the identity map so that any symplectomorphism $\phi \colon M \to M$ which lies in $\mathcal{U}$ has at least two fixed points.*

PROOF: Consider the manifold $M \times M$ endowed with the symplectic structure $\Omega \oplus (-\Omega)$. The diagonal $\Delta \subset M \times M$ is a Lagrangian submanifold. Proposition 4 implies that there exists an open neighborhood $U$ of the zero section in $T^*M$ and a symplectic map $\psi \colon U \to M \times M$ which is a diffeomorphism onto its image so that $\psi(0_p) = (p, p)$.

Now, there is an open neighborhood $\mathcal{U}_0$ of the identity map on $M$ in $\mathrm{Diff}(M)$ endowed with the $C^0$ topology which is characterized by the condition that $\phi$ belongs to $\mathcal{U}_0$ if and only if the graph of $\phi$ in $M \times M$, namely $id \times \phi$ lies in the open set $\psi(U) \subset M \times M$. Moreover, there is an open neighborhood $\mathcal{U} \subset \mathcal{U}_0$ of the identity map on $M$ in $\mathrm{Diff}(M)$ endowed with the $C^1$ topology which is characterized by the condition that $\phi$ belongs to $\mathcal{U}$ if and only if $\psi^{-1} \circ (id \times \phi) \colon M \to T^*M$ is the graph of a 1-form $\alpha_\phi$.

Now suppose that $\phi \in \mathcal{U}$ is a symplectomorphism. By our previous discussion, it follows that the graph of $\phi$ in $M \times M$ is Lagrangian. This implies that the graph of $\alpha_\phi$ is Lagrangian in $T^*M$ which, by our second example, implies that $\alpha_\phi$ is closed. Since $H^1_{dR}(M, \mathbb{R}) = 0$, this, in turn, implies that $\alpha_\phi = df_\phi$ for some smooth function $f$ on $M$.

Since $M$ is compact, it follows that $f_\phi$ must have at least two critical points. However, these critical points are zeros of the 1-form $df_\phi = \alpha_\phi$. It is a consequence of our construction that these points must then be places where the graph of $\phi$ intersects the diagonal $\Delta$. In other words, they are fixed points of $\phi$. $\qquad\square$

This theorem can be generalized considerably. According to a theorem of Hamilton [Ha], if $M$ is compact, then there is an open neighborhood $\mathcal{U}$ of the identity map $id$ in $\mathrm{Sp}(\Omega)$ (with the $C^1$ topology) so that every $\phi \in \mathcal{U}$ is the time-one flow of a symplectic vector field $X_\phi \in \mathrm{sp}(\Omega)$. If $X_\phi$ is actually Hamiltonian (which would, of course, follow if $H^1_{dR}(M, \mathbb{R}) = 0$), then $-X_\phi \lrcorner \Omega = df_\phi$, so $X_\phi$ will vanish at the critical points of $f_\phi$ and these will be fixed points of $\phi$.

## Appendix: Lie's Transformation Groups, II

The reader who is learning symplectic geometry for the first time may be astonished by the richness of the subject and, at the same time, be wondering "Are there other geometries like symplectic geometry which remain to be explored?" The point of this appendix is to give one possible answer to this very vague question.

When Lie began his study of transformation groups in $n$ variables, he modeled his attack on the known study of the finite groups. Thus, his idea was that he would find all of the "simple groups" first and then assemble them (by solving the extension problem) to classify the general group. Thus, if one "group" $G$ had a homomorphism onto another "group" $H$

$$1 \longrightarrow K \longrightarrow G \longrightarrow H \longrightarrow 1$$

then one could regard $G$ as a semi-direct product of $H$ with the kernel subgroup $K$.

Guided by this idea, Lie decided that the first task was to classify the transitive transformation groups $G$, i.e., the ones which acted transitively on $\mathbb{R}^n$ (at least locally). The reason for this was that, if $G$ had an orbit $S$ of dimension $0 < k < n$, then the restriction of the action of $G$ to $S$ would give a non-trivial homomorphism of $G$ into a transformation group in fewer variables.

Second, Lie decided that he needed to classify first the "groups" which, in his language, "did not preserve any subset of the variables." The example he had in mind was the group of diffeomorphisms of $\mathbb{R}^2$ of the form

$$\phi(x, y) = \big(f(x), g(x, y)\big).$$

Clearly the assignment $\phi \mapsto f$ provides a homomorphism of this group into the group of diffeomorphisms in one variable. Lie called groups which "did not preserve any subset of the variables" *primitive*. In modern language, *primitive* is taken to mean that $G$ does not preserve any foliation on $\mathbb{R}^n$ (coordinates on the leaf space would furnish a "proper subset of the variables" which was preserved by $G$).

Thus, the fundamental problem was to classify the "primitive transitive continuous transformation groups".

When the algebra of infinitesimal generators of $G$ was finite dimensional, Lie and his coworkers made good progress. Their work culminated in the work of Cartan and Killing, classifying the finite dimensional simple Lie groups. (Interestingly enough, they did not then go on to solve the extension problem and so classify all Lie groups. Perhaps they regarded this as a problem of lesser order. Or, more likely, the classification turned out to be messy, uninteresting, and ultimately intractable.)

They found that the simple groups fell into two types. Besides the special linear groups, such as $\mathrm{SL}(n, \mathbb{R})$, $\mathrm{SL}(n, \mathbb{C})$ and other complex analogs; orthogonal groups, such as $\mathrm{SO}(p, q)$ and its complex analogs; and symplectic groups, such as $\mathrm{Sp}(n, \mathbb{R})$ and its complex analogs (which became known as the classical groups), there were five "exceptional" types. This story is quite long, but very interesting. The "finite dimensional Lie groups" went on

to become an essential part of the foundation of modern differential geometry. A complete account of this classification (along with very interesting historical notes) can be found in [He].

However, when the algebra of infinitesimal generators of $G$ was *infinite* dimensional, the story was not so complete. Lie himself identified four classes of these "infinite dimensional primitive transitive transformation groups". They were

- In every dimension $n$, the full diffeomorphism group, $\mathtt{Diff}(\mathbb{R}^n)$.

- In every dimension $n$, the group of diffeomorphisms which preserve a fixed volume form $\mu$, denoted by $\mathtt{SDiff}(\mu)$.

- In every even dimension $2n$, the group of diffeomorphisms which preserve the standard symplectic form
$$\Omega_n = dx_1 \wedge dy^1 + \cdots + dx_n \wedge dy^n,$$
denoted by $\mathtt{Sp}(\Omega_n)$.

- In every odd dimension $2n+1$, the group of diffeomorphisms which preserve, *up to a scalar function multiple*, the 1-form

$$\omega_n = dz + x_1\, dy^1 + \cdots + x_n\, dy^n.$$

This "group" was known as the *contact group* and I will denote it by $\mathtt{Ct}(\omega_n)$.

However, Lie and his coworkers were never able to discover any others, though they searched diligently. (By the way, Lie was aware that there were also holomorphic analogs acting in $\mathbb{C}^n$, but, at that time, the distinction between real and complex was not generally made explicit. Apparently, an educated reader was supposed to know or be able to guess what the generalizations to the complex category were.)

In a series of four papers spanning from 1902 to 1910, Élie Cartan reformulated Lie's problem in terms of systems of partial differential equations and, under the hypothesis of analyticity (real and complex were not carefully distinguished), he proved that Lie's classes were essentially all of the infinite dimensional primitive transitive transformation groups. The slight extension was that $\mathtt{SDiff}(\mu)$ had a companion extension to $\mathbb{R}\cdot\mathtt{SDiff}(\mu)$, the diffeomorphisms which preserve $\mu$ up to a constant multiple and that $\mathtt{Sp}(\Omega_n)$ had a companion extension to $\mathbb{R} \cdot \mathtt{Sp}(\Omega_n)$, the diffeomorphisms which preserve $\Omega_n$ up to a constant multiple. Of course, there were also the holomorphic analogues of these. Notice the remarkable fact that there are no "exceptional infinite dimensional primitive transitive transformation groups".

These papers are remarkable, not only for their results, but for the wealth of concepts which Cartan introduced in order to solve his problem. In these papers, Cartan introduces the notion of $G$-structures (of all orders), principal bundles and their connections, jet bundles, prolongation (both of group actions and exterior differential systems), and a host of other ideas which were only appreciated much later. Perhaps because of its originality, Cartan's work in this area was essentially ignored for many years.

L.6.11

In the 1950's, when algebraic varieties were being explored and developed as complex manifolds, it began to be understood that complex manifolds were to be thought of as manifolds with an atlas of coordinate charts whose "overlaps" were holomorphic. Generalizing this example, it became clear that, for any collection $\Gamma$ of local diffeomorphisms of $\mathbb{R}^n$ which satisfied the following definition, one could define a category of $\Gamma$-manifolds as manifolds endowed with an atlas $\mathcal{A}$ of coordinate charts whose overlaps lay in $\mathcal{A}$.

**Definition 3:** A *local diffeomorphism* of $\mathbb{R}^n$ is a pair $(U, \phi)$ where $U \subset \mathbb{R}^n$ is an open set and $\phi: U \to \mathbb{R}^n$ is a one-to-one diffeomorphism onto its image. A set $\Gamma$ of local diffeomorphisms of $\mathbb{R}^n$ is said to form a *pseudo-group* on $\mathbb{R}^n$ if it satisfies the following three properties:

(1) (Composition and Inverses) If $(U, \phi)$ and $(V, \psi)$ are in $\Gamma$, then $(\phi^{-1}(V), \psi \circ \phi)$ and $(\phi(U), \phi^{-1})$ also belong to $\Gamma$.

(2) (Localization and Globalization) If $(U, \phi)$ is in $\Gamma$, and $W \subset U$ is open, then $(W, \phi_{|W})$ is also in $\Gamma$. Moreover, if $(U, \phi)$ is a local diffeomorphism of $\mathbb{R}^n$ such that $U$ can be written as the union of open subsets $W_\alpha$ for which $(W_\alpha, \phi_{|W_\alpha})$ is in $\Gamma$ for all $\alpha$, then $(U, \phi)$ is in $\Gamma$.

(3) (Non-triviality) $(\mathbb{R}^n, id)$ is in $\Gamma$.

As it turned out, the pseudo-groups $\Gamma$ of interest in geometry were exactly the ones which could be characterized as the (local) solutions of a system of partial differential equations, i.e., they were Lie's transformation groups. This caused a revival of interest in Cartan's work. Consequently, much of Cartan's work has now been redone in modern language. In particular, Cartan's classification was redone according to modern standards of rigor and a very readable account of this theory can be found in [SS].

In any case, symplectic geometry, seen in this light, is one of a small handful of "natural" geometries that one can impose on manifolds.

# Exercise Set 6:

## Symplectic Manifolds, II

**1**. Assume $n > 1$. Show that if $A_{r,R} \subset \mathbb{R}^{2n}$ (with its standard symplectic structure) is the annulus described by the relations $r < |\mathbf{x}| < R$, then there cannot be a symplectic diffeomorphism $\phi \colon A_{r,R} \to A_{s,S}$ that "exchanges the boundaries". (Hint: Show that if $\phi$ existed one would be able to construct a symplectic structure on $S^{2n}$.) Conclude that one cannot naïvely define connected sum in the category of symplectic manifolds. (The "naïve" definition would be to try to take two symplectic manifolds $M_1$ and $M_2$ of the same dimension, choose an open ball in each one, cut out a sub-ball of each and identify the resulting annuli by an appropriate diffeomorphism that was chosen to be a symplectomorphism.)

**2**. This exercise completes the proof of Proposition 1.

(i) Let $\mathcal{S}_n^+$ denote the space of $n$-by-$n$ positive definite symmetric matrices. Show that the map $\sigma \colon \mathcal{S}_n^+ \to \mathcal{S}_n^+$ defined by $\sigma(s) = s^2$ is a one-to-one diffeomorphism of $\mathcal{S}_n^+$ onto itself. Conclude that every element of $\mathcal{S}_n^+$ has a unique positive definite square root and that the map $s \mapsto \sqrt{s}$ is a smooth mapping. Show also that, for any $r \in \mathrm{O}(n)$, we have $\sqrt{{}^t r a r} = {}^t r \sqrt{a}\, r$, so that the square root function is $\mathrm{O}(n)$-equivariant.

(ii) Let $\mathcal{A}_n^\bullet$ denote the space of $n$-by-$n$ invertible anti-symmetric matrices. Show that, for $a \in \mathcal{A}_n^\bullet$, the matrix $-a^2$ is symmetric and positive definite. Show that the matrix $b = \sqrt{-a^2}$ is the unique symmetric positive definite matrix that satisfies $b^2 = -a^2$ and moreover that $b$ commutes with $a$. Check also that the mapping $a \mapsto \sqrt{-a^2}$ is $\mathrm{O}(n)$-equivariant.

(iii) Now verify the claim made in the proof of Proposition 1 that, for any smooth vector bundle $E$ over a manifold $M$ endowed with a smooth inner product on the fibers and any smooth, invertible skew-symmetric bundle mapping $A \colon E \to E$, there exists a unique smooth positive definite symmetric bundle mapping $B \colon E \to E$ that satisfies $B^2 = -A^2$ and that commutes with $A$.

**3**. This exercise requires that you know something about characteristic classes.

(i) Show that $S^{4n}$ has no almost complex structure for any $n$. (Hint: What could the total Chern and Pontrijagin class of the tangent bundle be?)

   (Using the Bott Periodicity Theorem, it can be shown that the characteristic class $c_n(E)$ of any complex bundle $E$ over $S^{2n}$ is an integer multiple of $(n-1)!\, v$ where $v \in H^{2n}(S^{2n}, \mathbb{Z})$ is a generator. It follows that, among the spheres, only $S^2$ and $S^6$ could have almost complex structures and it turns out that they both do. It is a long standing problem whether or not $S^6$ has a complex structure.)

(ii) Using the formulas for 4-manifolds developed in the Lecture, determine how many possibilities there are for the first Chern class $c_1(J)$ of an almost complex structure J on $M$ where $M$ a connected sum of 3 or 4 copies of $\mathbb{CP}^2$.

**4.** Show that, if $\Omega_0$ is a symplectic structure on a compact manifold $M$, then there is an open neighborhood $U$ in $H^2(M, \mathbb{R})$ of $[\Omega_0]$, such that, for all $u \in U$, there is a symplectic structure $\Omega_u$ on $M$ with $[\Omega_u] = u$. (Hint: Since $M$ is compact, for any closed 2-form $\Upsilon$, the 2-form $\Omega + t\Upsilon$ is non-degenerate for all sufficiently small $t$.)

**5.** Mimic the proof of Theorem 1 to prove another theorem of Moser: For any compact, connected, oriented manifold $M$, two volume forms $\mu_0$ and $\mu_1$ differ by an oriented diffeomorphism (i.e., there exists an orientation preserving diffeomorphism $\phi: M \to M$ that satisfies $\phi^*(\mu_1) = \mu_0$) if and only if

$$\int_M \mu_0 = \int_M \mu_1.$$

(This theorem is also true without the hypothesis of compactness, but the proof is slightly more delicate.)

**6.** Let $M$ be a connected, smooth oriented 4-manifold and let $\mu \in \mathcal{A}^4(M)$ be a volume form that satisfies $\int_M \mu = 1$. (By the previous problem, any two such forms differ by an oriented diffeomorphism of $M$.) For any (smooth) $\Omega \in \mathcal{A}^2(M)$, define $*(\Omega^2) \in C^\infty(M)$ by the equation

$$\Omega^2 = *(\Omega^2)\, \mu.$$

Now, fix a cohomology class $u \in H^2_{dR}(M)$ satisfying $u^2 = r[\mu]$ where $r \neq 0$. Define the functional $\mathcal{F}: u \to \mathbb{R}$

$$\mathcal{F}(\Omega) = \int_M *(\Omega^2)\, \Omega^2 \qquad \text{for } \Omega \in u.$$

Show that any $\mathcal{F}$-critical 2-form $\Omega \in u$ is a symplectic form satisfying $*(\Omega^2) = r$ and that $\mathcal{F}$ has no critical values other than $r^2$. Show also that $\mathcal{F}(\Omega) \geq r^2$ for all $\Omega \in u$.

This motivates defining an invariant of the class $u$ by

$$\mathcal{I}(u) = \inf_{\Omega \in u} \mathcal{F}(\Omega).$$

Gromov has suggested (private communication) that perhaps $\mathcal{I}(u) = r^2$ for all $u$, even when the infimum is not attained.

**7.** Let $P \subset M$ be a closed submanifold and let $U \subset M$ be an open neighborhood of $P$ in $M$ that can be retracted onto $P$, i.e., there exists a smooth map $R: U \times [0,1] \to U$ so that $R(u, 1) = u$ for all $u \in U$, $R(p, t) = p$ for all $p \in P$ and $t \in [0, 1]$, and $R(u, 0)$ lies in $P$ for all $u \in U$. (Every closed submanifold of $M$ has such a neighborhood.)

Show that if $\Phi$ is a closed $k$-form on $U$ that vanishes at every point of $P$, then there exists a $(k{-}1)$-form $\phi$ on $U$ that vanishes on $P$ and satisfies $d\phi = \Phi$. (Hint: Mimic Poincaré's Homotopy Argument: Let $\Upsilon = R^*(\Phi)$ and set $v = \frac{\partial}{\partial t} \lrcorner \Upsilon$. Then, using the fact that $v(u, t)$ can be regarded as a $(k{-}1)$-form at $u$ for all $t$, define

$$\phi(u) = \int_0^1 v(u, t)\, dt.$$

Now verify that $\phi$ has the desired properties.)

**8**. Show that Theorem 2 implies Darboux' Theorem. (Hint: Take $P$ to be a point in a symplectic manifold $M$.)

**9**. This exercise assumes that you have done Exercise 5.10. Let $(M, \Omega)$ be a symplectic manifold. Show that the following description of the flux homomorphism is valid. Let $p$ be an $e$-based path in $\mathsf{Sp}(\Omega)$. Thus, $p \colon [0,1] \times M \to M$ satisfies $p_t^*(\Omega) = \Omega$ for all $0 \le t \le 1$. Show that $p^*(\Omega) = \Omega + \varphi \wedge dt$ for some 1-form $\varphi$ on $[0,1] \times M$. Let $\iota_t \colon M \to [0,1] \times M$ be the "$t$-slice inclusion": $\iota_t(m) = (t, m)$, and set $\varphi_t = \iota_t^*(\varphi)$.

Show that $\varphi_t$ is closed for all $0 \le t \le 1$. Show that if we set

$$\tilde{\Phi}(p) = \int_0^1 \varphi_t \, dt,$$

then the cohomology class $[\tilde{\Phi}(p)] \in H^1_{dR}(M, \mathbb{R})$ depends only on the homotopy class of $p$ and hence defines a map $\Phi \colon \widetilde{\mathsf{Sp}}^0(\Omega) \to H^1_{dR}(M, \mathbb{R})$. Verify that this map is the same as the flux homomorphism defined in Exercise 5.10.

Use this description to show that if $p$ is in the kernel of $\Phi$, then $p$ is homotopic to a path $p'$ for which the forms $\varphi_t'$ are all exact. This shows that the kernel of $\Phi$ is actually connected.

**10**. *The point of this exercise is to show that any symplectic vector bundle over a symplectic manifold $(M, \Omega)$ can occur as the symplectic normal bundle for some symplectic embedding $M$ into some other symplectic manifold.*

Let $(M, \Omega)$ be a symplectic manifold and let $\pi \colon E \to M$ be a symplectic vector bundle over $M$ of rank $2n$. (I.e., $E$ comes equipped with a section $B$ of $\Lambda^2(E^*)$ that restricts to each fiber $E_m$ to be a symplectic structure $B_m$.) Show that there exists a symplectic structure $\Psi$ on an open neighborhood in $E$ of the zero section of $E$ that satisfies the condition that $\Psi_{0_m} = \Omega_m + B_m$ under the natural identification $T_{0_m} E = T_m M \oplus E_m$.

(Hint: Choose a locally finite open cover $\mathcal{U} = \{U_\alpha \mid \alpha \in A\}$ of $M$ so that, if we define $E_\alpha = \pi^{-1}(U_\alpha)$, then there exists a symplectic trivialization $\tau_\alpha \colon E_\alpha \to \mathbb{R}^{2n}$ (where $\mathbb{R}^{2n}$ is given its standard symplectic structure $\Omega_0 = dx_i \wedge dy^i$). Now let $\{\lambda_\alpha \mid \alpha \in A\}$ be a partition of unity subordinate to the cover $\mathcal{U}$. Show that the form

$$\Psi = \pi^*(\Omega) + \sum_\alpha d\big(\lambda_\alpha \, \tau_\alpha^*(x_i \, dy^i)\big)$$

has the desired properties.)

**11**. Show that if $E$ is a symplectic vector bundle over $M$ and $L \subset E$ is a Lagrangian subbundle, then $E$ is isomorphic to $L \oplus L^*$ as a symplectic bundle. (The symplectic bundle structure $\Upsilon$ on $L \oplus L^*$ is the one that, on each fiber satisfies

$$\Upsilon\big((v, \alpha), (w, \beta)\big) = \alpha(w) - \beta(v). \,)$$

E.6.3

(Hint: First choose a complementary subbundle $F \subset E$ so that $E = L \oplus F$. Show that $F$ is naturally isomorphic to $L^*$ abstractly by using the fact that the symplectic structure on $E$ is non-degenerate. Then show that there exists a bundle map $A \colon F \to L$ so that

$$\tilde{F} = \{v + Av \mid v \in F\}$$

is also a Lagrangian subbundle of $E$ that is complementary to $L$ and isomorphic to $L^*$ via some bundle map $\alpha \colon L^* \to \tilde{F}$. Now show that $id \oplus \alpha \colon L \oplus L^* \to L \oplus \tilde{F} \simeq E$ is a symplectic bundle isomorphism.)

**12**. Action-Angle Coordinates. Proposition 4 can be used to show the existence of so-called action angle coordinates in the neighborhood of a compact level set of a completely integrable Hamiltonian system. (See Lecture 5). Here is how this goes: Let $(M^{2n}, \Omega)$ be a symplectic manifold and let $f = (f^1, \ldots, f^n) \colon M \to \mathbb{R}^n$ be a smooth submersion with the property that the coordinate functions $f^i$ are in involution, i.e., $\{f^i, f^j\} = 0$. Suppose that, for some $c \in \mathbb{R}^n$, the $f$-level set $M_c = f^{-1}(c)$ is compact. Replacing $f$ by $f - c$, we may assume that $c = 0$, which we do from now on.

Show that $M_0 \subset M$ is a closed Lagrangian submanifold of $M$.

Use Proposition 4 to show that there is an open neighborhood $B$ of $0 \in \mathbb{R}^n$ so that $\left(f^{-1}(B), \Omega\right)$ is symplectomorphic to a neighborhood $U$ of the zero section in $T^* M_0$ (endowed with its standard symplectic structure) in such a way that, for each $b \in B$, the submanifold $M_b = f^{-1}(b)$ is identified with the graph of a closed 1-form $\omega_b$ on $M_0$. Show that it is possible to choose $b_1, \ldots, b_n$ in $B$ so that the corresponding closed 1-forms $\omega_1, \ldots, \omega_n$ are linearly independent at every point of $M_0$.

Conclude that $M_0$ is diffeomorphic to a torus $T = \mathbb{R}^n / \Lambda$ where $\Lambda \subset \mathbb{R}^n$ is a lattice, in such a way that the forms $\omega_i$ become identified with $d\theta_i$ where $\theta_i$ are the corresponding linear coordinates on $\mathbb{R}^n$.

Now prove that for any $b \in B$, the 1-form $\omega_b$ must be a linear combination of the $\omega_i$ with constant coefficients. Thus, there are functions $a^i$ on $B$ so that $\omega_b = a^i(b)\omega_i$. (Hint: Show that the coefficients must be invariant under the flows of the vector fields dual to the $\omega_i$.)

Conclude that, under the symplectic map identifying $M_B$ with $U$, the form $\Omega$ gets identified with $da^i \wedge d\theta_i$. The functions $a^i$ and $\theta_i$ are the so-called "action-angle coordinates".

*Extra Credit:* Trace through the methods used to prove Proposition 4 and show that, in fact, the action-angle coordinates can be constructed using quadrature and "finite" operations.

# Lecture 7:

# Classical Reduction

In this section, we return to the study of group actions. This time, however, we will concentrate on group actions on symplectic manifolds that preserve the symplectic structure. Such actions happen to have quite interesting properties and moreover, turn out to have a wide variety of applications.

**Symplectic Group Actions.** First, the basic definition.

**Definition 1:** Let $(M, \Omega)$ be a symplectic manifold and let $G$ be a Lie group. A left action $\lambda: G \times M \to M$ of $G$ on $M$ is a *symplectic* action if $\lambda_a^*(\Omega) = \Omega$ for all $a \in G$.

We have already encountered several examples:

**Example:** LAGRANGIAN SYMMETRIES. If $G$ acts on a manifold $M$ is such a way that it preserves a non-degenerate Lagrangian $L: TM \to \mathbb{R}$, then, by construction, it preserves the symplectic 2-form $d\omega_L$.

**Example:** COTANGENT ACTIONS. A left $G$-action $\lambda: G \times M \to M$, induces an action $\tilde{\lambda}$ of $G$ on $T^*M$. Namely, for each $a \in G$, the diffeomorphism $\lambda_a: M \to M$ induces a diffeomorphism $\tilde{\lambda}_a: T^*M \to T^*M$. Since the natural symplectic structure on $T^*M$ is invariant under diffeomorphisms, it follows that $\tilde{\lambda}$ is a symplectic action.

**Example:** COADJOINT ORBITS. As we saw in Lecture 5, for every $\xi \in \mathfrak{g}^*$, the coadjoint orbit $G \cdot \xi$ carries a natural $G$-invariant symplectic structure $\Omega_\xi$. Thus, the left action of $G$ on $G \cdot \xi$ is symplectic.

**Example:** CIRCLE ACTIONS ON $\mathbb{C}^n$. Let $z^1, \ldots, z^n$ be linear complex coordinates on $\mathbb{C}^n$ and let this vector space be endowed with the symplectic structure

$$
\begin{aligned}
\Omega &= \tfrac{i}{2}\left(dz^1 \wedge d\bar{z}^1 + \cdots + dz^n \wedge d\bar{z}^n\right) \\
&= dx^1 \wedge dy^1 + \cdots + dx^n \wedge dy^n
\end{aligned}
$$

where $z^k = x^k + iy^k$. Then for any integers $(k_1, \ldots, k_n)$, we can define an action of $S^1$ on $\mathbb{C}^n$ by the formula

$$
e^{i\theta} \cdot \begin{pmatrix} z^1 \\ \vdots \\ z^n \end{pmatrix} = \begin{pmatrix} e^{ik_1\theta} z^1 \\ \vdots \\ e^{ik_n\theta} z^n \end{pmatrix}
$$

The reader can easily check that this defines a symplectic circle action on $\mathbb{C}^n$.

Generally what we will be interested in is the following: $Y$ will be a Hamiltonian vector field on a symplectic manifold $(M, \Omega)$ and $G$ will act symplectically on $M$ as a group of symmetries of the flow of $Y$. We want to understand how to use the action of $G$ to "reduce" the problem of integrating the flow of $Y$.

In Lecture 3, we saw that when $Y$ was the Euler-Lagrange vector field associated to a non-degenerate Lagrangian $L$, then the infinitesimal generators of symmetries of $L$ could be used to generate conserved quantities for the flow of $Y$. We want to extend this process (as far as is reasonable) to the general case.

▶      *For the rest of the lecture, I will assume that $G$ is a Lie group with a symplectic action $\lambda$ on a connected symplectic manifold $(M, \Omega)$.*

Since $\lambda$ is symplectic, it follows that the mapping $\lambda_*\colon \mathfrak{g} \to \mathfrak{X}(M)$ actually has image in $\mathfrak{sp}(\Omega)$, the algebra of symplectic vector fields on $M$. As we saw in Lecture 3, $\lambda_*$ is an anti-homomorphism, i.e., $\lambda_*\bigl([x,y]\bigr) = -\bigl[\lambda_*(x), \lambda_*(y)\bigr]$. Since, as we saw in Lecture 5, $[\mathfrak{sp}(\Omega), \mathfrak{sp}(\Omega)] \subset \mathfrak{h}(\Omega)$, it follows that $\lambda_*\bigl([\mathfrak{g}, \mathfrak{g}]\bigr) \subset \mathfrak{h}(\Omega)$. Thus, $H_\lambda\colon \mathfrak{g} \to H^1_{dR}(M, \mathbb{R})$ defined by $H_\lambda(x) = \bigl[\lambda_*(x) \lrcorner \Omega\bigr]$ is a homomorphism of Lie algebras with kernel containing the commutator subalgebra $[\mathfrak{g}, \mathfrak{g}]$.

The map $H_\lambda$ is the obstruction to finding a Hamiltonian function associated to each infinitesimal symmetry $\lambda_*(x)$ since $H_\lambda(x) = 0$ if and only if $\lambda_*(x) \lrcorner \Omega = -df$ for some $f \in C^\infty(M)$.

**Definition 2:** A symplectic action $\lambda\colon G \times M \to M$ is said to be *Hamiltonian* if $H_\lambda = 0$, i.e., if $\lambda_*(\mathfrak{g}) \subset \mathfrak{h}(\Omega)$.

There are a few particularly interesting cases where the obstruction $H_\lambda$ *must* vanish:

• If $H^1_{dR}(M, \mathbb{R}) = 0$. In particular, if $M$ is simply connected.

• If $\mathfrak{g}$ is perfect, i.e., $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. For example this happens whenever the Killing form on $\mathfrak{g}$ is non-degenerate (this is the first Whitehead Lemma, see Exercise 3). However, this is not the only case: For example, if $G$ is the group of rigid motions in $\mathbb{R}^n$ for $n \geq 3$, then $\mathfrak{g}$ has this property, even though its Killing form is degenerate.

• If there exists a 1-form $\omega$ on $M$ that is invariant under $G$ and satisfies $\Omega = d\omega$. (This is the case of symmetries of a Lagrangian.) To see this, note that if $X$ is a vector field on $M$ that preserves $\omega$, then

$$0 = \mathfrak{L}_X \omega = d(X \lrcorner \omega) + X \lrcorner \Omega,$$

so $X \lrcorner \Omega$ is exact.

For a Hamiltonian action $\lambda$, every infinitesimal symmetry $\lambda_*(x)$ has a Hamiltonian function $f_x \in C^\infty$. However, the choice of $f_x$ is not unique since we can add any constant to $f_x$ without changing its Hamiltonian vector field. This non-uniqueness causes some problems in the theory we wish to develop.

To see why, suppose that we choose a (linear) lifting $\rho\colon \mathfrak{g} \to C^\infty(M)$ of $-\lambda_*\colon \mathfrak{g} \to \mathfrak{h}(\Omega)$. (The choice of $-\lambda_*$ instead of $\lambda_*$ was made to get rid of the annoying sign in the formula for the bracket.)

$$
\begin{array}{ccccccc}
 & & & \mathfrak{g} & & & \\
 & & \rho \downarrow & & \searrow^{-\lambda_*} & & \\
0 & \to & \mathbb{R} & \to & C^\infty(M) & \to & \mathfrak{h}(\Omega) & \to & 0
\end{array}
$$

Thus, for every $x \in \mathfrak{g}$, we have $\lambda_*(x) \lrcorner \Omega = d(\rho(x))$. A short calculation (see the Exercises) now shows that $\{\rho(x), \rho(y)\}$ is a Hamiltonian function for $-\lambda_*([x,y])$, i.e., that

$$\lambda_*([x,y]) \lrcorner \Omega = d(\{\rho(x), \rho(y)\}).$$

In particular, it follows (since $M$ is connected) that there must be a skew-symmetric bilinear map $c_\rho \colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ so that

$$\{\rho(x), \rho(y)\} = \rho([x,y]) + c_\rho(x,y).$$

An application of the Jacobi identity implies that the map $c_\rho$ satisfies the condition

$$c_\rho([x,y], z) + c_\rho([y,z], x) + c_\rho([z,x], y) = 0 \quad \text{for all } x, y, z \in \mathfrak{g}.$$

This condition is known as the *2-cocycle condition* for $c_\rho$ regarded as an element of $A^2(\mathfrak{g}) = \Lambda^2(\mathfrak{g}^*)$. (See Exercise 3 for an explanation of this terminology.)

For purposes of simplicity, it would be nice if we could choose $\rho$ so that $c_\rho$ were identically zero. In order to see whether this is possible, let us choose another linear map $\tilde{\rho} \colon \mathfrak{g} \to C^\infty(M)$ that satisfies $\tilde{\rho}(x) = \rho(x) + \xi(x)$ where $\xi \colon \mathfrak{g} \to \mathbb{R}$ is any linear map. Every possible lifting of $-\lambda_*$ is clearly of this form for some $\xi$. Now we compute that

$$\begin{aligned}\{\tilde{\rho}(x), \tilde{\rho}(y)\} = \{\rho(x), \rho(y)\} &= \rho([x,y]) + c_\rho(x,y) \\ &= \tilde{\rho}([x,y]) + c_\rho(x,y) - \xi([x,y]).\end{aligned}$$

Thus, $c_{\tilde{\rho}}(x,y) = c_\rho(x,y) - \xi([x,y])$. Thus, in order to be able to choose $\tilde{\rho}$ so that $c_{\tilde{\rho}} = 0$, we see that there must exist a $\xi \in \mathfrak{g}^*$ so that $c_\rho = -\delta\xi$ where $\delta\xi$ is the skew-symmetric bilinear map on $\mathfrak{g}$ that satisfies $\delta\xi(x,y) = -\xi([x,y])$ (see the Exercises for an explanation of this notation). This is known as the *2-coboundary condition.*

There are several important cases where we can assure that $c_\rho$ can be written in the form $-\delta\xi$. Among them are:

• If $M$ is compact, then the sequence

$$0 \to \mathbb{R} \to C^\infty(M) \to \mathsf{H}(\Omega) \to 0$$

splits: If we let $C_0^\infty(M, \Omega) \subset C^\infty(M)$ denote the space of functions $f$ for which $\int_M f\, \Omega^n = 0$, then these functions are closed under Poisson bracket (see Exercise 5.6 for a hint as to why this is true) and we have a splitting of Lie algebras $C^\infty(M) = \mathbb{R} \oplus C_0^\infty(M, \Omega)$. Now just choose the unique $\rho$ so that it takes values in $C_0^\infty(M, \Omega)$. This will clearly have $c_\rho = 0$.

• If $\mathfrak{g}$ has the property that every 2-cocycle for $\mathfrak{g}$ is actually a 2-coboundary. This happens, for example, if the Killing form of $\mathfrak{g}$ is non-degenerate (this is the second Whitehead Lemma, see Exercise 3), though it can also happen for other Lie algebras. For example, for the non-abelian Lie algebra of dimension 2, it is easy to see that every 2-cocycle is a 2-coboundary.

• If there is a 1-form $\omega$ on $M$ that is preserved by the $G$ action and satisfies $d\omega = \Omega$. (This is true in the case of symmetries of a Lagrangian.) In this case, we can merely take $\rho(x) = -\omega(\lambda_*(x))$. I leave as an exercise for the reader to check that this works.

**Definition 3:** A Hamiltonian action $\lambda\colon G \times M \to M$ is said to be a *Poisson* action if there exists a lifting $\rho$ with $c_\rho = 0$.

Henceforth in this Lecture, I am only going to consider Poisson actions. By my previous remarks, this case includes all of the Lagrangians with symmetries, but it also includes many others.

I will assume that, in addition to having a Poisson action $\lambda\colon G \times M \to M$ specified, we have chosen a lifting $\rho\colon \mathfrak{g} \to C^\infty(M)$ of $-\lambda_*$ that satisfies $\{\rho(x), \rho(y)\} = \rho([x,y])$ for all $x, y \in \mathfrak{g}$. Note that such a $\rho$ is unique up to replacement by $\tilde\rho = \rho + \xi$ where $\xi : \mathfrak{g} \to \mathbb{R}$ satisfies $\delta\xi = 0$. Such $\xi$ (if any non-zero ones exist) are fixed under the co-adjoint action of the identity component of $G$.

**The Momentum Mapping.** We are now ready to make one of the most important constructions in the theory.

**Definition 4:** The *momentum mapping* associated to $\lambda$ and $\rho$ is the mapping $\mu\colon M \to \mathfrak{g}^*$ that satisfies

$$\mu(m)(y) = \rho(y)(m).$$

Note that, for fixed $m \in M$, the assignment $y \mapsto \rho(y)(m)$ is a linear map from $\mathfrak{g}$ to $\mathbb{R}$, so the definition makes sense.

It is worth pausing to consider why this mapping is called the momentum mapping. The reader should calculate this mapping in the case of a free particle or a rigid body moving in space. In either case, the Lagrangian is invariant under the action of the group $G$ of rigid motions of space. If $y \in \mathfrak{g}$ corresponds to a translation, then $\rho(y)$ gives the function on $T\mathbb{R}^3$ that evaluates at each point (i.e., each position-plus-velocity) to be the linear momentum in the direction of translation. If $y$ corresponds to rotation about a fixed axis, then $\rho(y)$ turns out to be the angular momentum of the body about that axis.

One important reason for studying the momentum mapping is the following formulation of the classical conservation of momentum theorems:

**Proposition 1:** *If $Y$ is a symplectic vector field on $M$ that is invariant under the action of $G$, then $\mu$ is constant on the integral curves of $Y$.*  □

In particular, $\mu$ provides conserved quantities for any $G$-invariant Hamiltonian.

The main result about the momentum mapping is the following one.

**Theorem 1:** *If $G$ is connected, then the momentum mapping $\mu\colon M \to \mathfrak{g}^*$ is $G$-equivariant.*

L.7.4

PROOF: Recall that the coadjoint action of $G$ on $\mathfrak{g}^*$ is defined by $\text{Ad}^*(g)(\xi)(x) = \xi\big(\text{Ad}(g^{-1})x\big)$. The condition that $\mu$ be $G$-equivariant, i.e., that $\mu(g \cdot m) = \text{Ad}^*(g)\big(\mu(m)\big)$ for all $m \in M$ and $g \in G$, is thus seen to be equivalent to the condition that

$$\rho\big(\text{Ad}(g^{-1})y\big)(m) = \rho(y)(g \cdot m)$$

for all $m \in M$, $g \in G$, and $y \in \mathfrak{g}$. This is the identity I shall prove.

Since $G$ is connected and since each side of the above equation represents a $G$-action, if we prove that the above formula holds for $g$ of the form $g = e^{tx}$ for any $x \in \mathfrak{g}$ and any $t \in \mathbb{R}$, the formula for general $g$ will follow. Thus, we want to prove that

$$\rho\big(\text{Ad}(e^{-tx})y\big)(m) = \rho(y)(e^{tx} \cdot m)$$

for all $t$. Since this latter equation holds at $t = 0$, it is enough to show that both sides have the same derivative with respect to $t$.

Now the derivative of the right hand side of the formula is

$$\begin{aligned}
d\big(\rho(y)\big)\big(\lambda_*(x)(e^{tx} \cdot m)\big) &= \Omega\big(\lambda_*(y)(e^{tx} \cdot m), \lambda_*(x)(e^{tx} \cdot m)\big) \\
&= \Omega\big(\lambda_*\big(\text{Ad}(e^{-tx})y\big)(m), \lambda_*\big(\text{Ad}(e^{-tx})x\big)(m)\big)\big) \\
&= \Omega\big(\lambda_*\big(\text{Ad}(e^{-tx})y\big)(m), \lambda_*(x)(m)\big)
\end{aligned}$$

where, to verify the second equality we have used the identity

$$\lambda'_a\big(\lambda_*(y)(m)\big) = \lambda_*\big(\text{Ad}(a)y\big)(a \cdot m)$$

and the fact that $\Omega$ is $G$-invariant.

On the other hand, the derivative of the left hand side of the formula is clearly

$$\begin{aligned}
\rho\big([-x, \text{Ad}(e^{-tx})y]\big)(m) &= -\big\{\rho(x), \rho\big(\text{Ad}(e^{-tx})y\big)\big\}(m) \\
&= \Omega\big(\lambda_*\big(\text{Ad}(e^{-tx})y\big)(m), \lambda_*(x)(m)\big)
\end{aligned}$$

so we are done. (Note that I have used my assumption that $c_\rho = 0$!)   $\square$

**Example:** LEFT-INVARIANT METRICS ON LIE GROUPS. Let $G$ be a Lie group and let $Q: \mathfrak{g} \to \mathbb{R}$ be a non-degenerate quadratic form with associated inner product $\langle,\rangle_Q$. Let $L: TG \to \mathbb{R}$ be the Lagrangian

$$L = \tfrac{1}{2}Q\big(\omega\big)$$

where $\omega: TG \to \mathfrak{g}$ is, as usual, the canonical left-invariant form on $G$. Then, using the basepoint map $\pi: TG \to G$, we compute that

$$\omega_L = \big\langle \omega, \pi^*(\omega)\big\rangle_Q.$$

As we saw in Lecture 3, the assumption that $Q$ is non-degenerate implies that $d\omega_L$ is a symplectic form on $TG$. Now, since the flow of a right-invariant vector field $Y_x$ is multiplication on the left by $e^{tx}$, it follows that, for this action, we may define

$$\rho(x) = -\omega_L(Y'_x) = -\big\langle \omega, \omega(Y_x)\big\rangle_Q = -\big\langle \omega, \text{Ad}(g^{-1})x\big\rangle_Q$$

(where $g\colon TG \to G$ is merely a more descriptive name for the base point map than $\pi$).

Now, there is an isomorphism $\tau_Q\colon \mathfrak{g} \to \mathfrak{g}^*$, called *transpose with respect to $Q$* that satisfies $\tau_Q(x)(y) = \langle x, y \rangle_Q$ for all $x, y \in \mathfrak{g}$. In terms of $\tau_Q$, we can express the momentum mapping as

$$\mu(v) = -\mathrm{Ad}^*(g)\big(\tau_Q(\omega(v))\big)$$

for all $v \in TG$. Note that $\mu$ is $G$-equivariant, as promised by the theorem.

According to the Proposition 1, the function $\mu$ is a conserved quantity for the solutions of the Euler-Lagrange equations. In one of the Exercises, you are asked to show how this information can be used to help solve the Euler-Lagrange equations for the $L$-critical curves.

**Example:** Coadjoint Orbits. Let $G$ be a Lie group and consider $\xi \in \mathfrak{g}^*$ with stabilizer subgroup $G_\xi \subset G$. The orbit $G \cdot \xi \subset \mathfrak{g}^*$ is canonically identified with $G/G_\xi$ ( identify $a \cdot \xi$ with $aG_\xi$) and we have seen that there is a canonical $G$-invariant symplectic form $\Omega_\xi$ on $G/G_\xi$ that satisfies $\pi_\xi^*(\Omega_\xi) = d\omega_\xi$ where $\pi_\xi : G \to G/G_\xi$ is the coset projection, $\omega$ is the tautological left-invariant 1-form on $G$, and $\omega_\xi = \xi(\omega)$.

Recall also that, for each $x \in \mathfrak{g}$, the right-invariant vector field $Y_x$ on $G$ is defined so that $Y_x(e) = x \in \mathfrak{g}$. Then the vector field $\lambda_*(x)$ on $G/G_\xi$ is $\pi_\xi$-related to $Y_x$, so

$$\pi_\xi^*\big(\lambda_*(x) \lrcorner \Omega_\xi\big) = Y_x \lrcorner d\omega_\xi = d\big(-\omega_\xi(Y_x)\big).$$

(This last equality follows because $\omega_\xi$, being left-invariant, is invariant under the flow of $Y_x$.) Now, the value of the function $\omega_\xi(Y_x)$ at $a \in G$ is

$$\omega_\xi(Y_x)(a) = \xi\big(\omega(Y_x(a))\big) = \xi\big(\mathrm{Ad}(a^{-1})(x)\big) = \mathrm{Ad}^*(a)(\xi)(x) = (a \cdot \xi)(x).$$

Thus, it follows that the natural left action of $G$ on $G{\cdot}\xi$ is Poisson, with momentum mapping $\mu : G{\cdot}\xi \to \mathfrak{g}^*$ given by

$$\mu(a \cdot \xi) = -a \cdot \xi.$$

(Note: some authors do not have a minus sign here, but that is because their $\Omega_\xi$ is the negative of ours.)

**Reduction.** I now want to discuss a method of taking quotients by group actions in the symplectic category. Now, when a Lie group $G$ acts symplectically on the left on a symplectic manifold $M$, it is not generally true that the space of orbits $G\backslash M$ can be given a symplectic structure, even when this orbit space can be given the structure of a smooth manifold (for example, the quotient need not be even dimensional).

However, when the action is Poisson, there is a natural method of breaking the orbit space $G\backslash M$ into a union of symplectic submanifolds provided that certain regularity criteria are met. The procedure I will describe is known as *symplectic reduction*. It is due, in its modern form, to Marsden and Weinstein (see [GS 2]).

The idea is simple: If $\mu\colon M \to \mathfrak{g}^*$ is the momentum mapping, then the $G$-equivariance of $\mu$ implies that there is a well-defined set map

$$\bar{\mu}\colon G\backslash M \to G\backslash \mathfrak{g}^*.$$

The theorem we are about to prove asserts that, provided certain regularity criteria are met, the subsets $M_\xi = \bar{\mu}^{-1}(\bar{\xi}) \subset G\backslash M$ are symplectic manifolds in a natural way.

**Definition 4:** Let $f\colon X \to Y$ be a smooth map. A point $y \in Y$ is a *clean* value of $f$ if the set $f^{-1}(y) \subset X$ is a smooth submanifold of $X$ and, moreover, if $T_x f^{-1}(y) = \ker f'(x)$ for each $x \in f^{-1}(y)$.

Note: While every regular value of $f$ is clean, not every clean value of $f$ need be regular. The concept of cleanliness is very frequently encountered in the reduction theory we are about to develop.

**Theorem 2:** Let $\lambda\colon G \times M \to M$ be a Poisson action on the symplectic manifold $M$. Let $\mu\colon M \to \mathfrak{g}^*$ be a momentum mapping for $\lambda$. Suppose that, $\xi \in \mathfrak{g}^*$ is a clean value of $\mu$. Then $G_\xi$ acts smoothly on $\mu^{-1}(\xi)$. Suppose further that the space of $G_\xi$-orbits in $\mu^{-1}(\xi)$, say, $M_\xi = G_\xi\backslash(\mu^{-1}(\xi))$, can be given the structure of a smooth manifold in such a way that the quotient mapping $\pi_\xi\colon \mu^{-1}(\xi) \to M_\xi$ is a smooth submersion. Then there exists a symplectic structure $\Omega_\xi$ on $M_\xi$ that is defined by the condition that $\pi_\xi^*(\Omega_\xi)$ be the pullback of $\Omega$ to $\mu^{-1}(\xi)$.

PROOF: Since $\xi$ is a clean value of $\mu$, we know that $\mu^{-1}(\xi)$ is a smooth submanifold of $M$. By the $G$-equivariance of the momentum mapping, the stabilizer subgroup $G_\xi \subset G$ acts on $M$ preserving the submanifold $\mu^{-1}(\xi)$. The restricted action of $G_\xi$ on $\mu^{-1}(\xi)$ is easily seen to be smooth.

Now, I claim that, for each $m \in \mu^{-1}(\xi)$, the $\Omega$-complementary subspace to $T_m(\mu^{-1}(\xi))$ is the space $T_m(G \cdot m)$, i.e., the tangent to the $G$-orbit through $m$. To see this, first note that the space $T_m(G \cdot m)$ is spanned by the values at $m$ assumed by the vector fields $\lambda_*(x)$ for $x \in \mathfrak{g}$. Thus, a vector $v \in T_m M$ lies in the $\Omega$-complementary space of $T_m(G \cdot m)$ if and only if $v$ satisfies $\Omega(\lambda_*(x)(m), v) = 0$ for all $x \in \mathfrak{g}$. Since, by definition, $\Omega(\lambda_*(x)(m), v) = d(\rho(x))(v)$, it follows that this condition on $v$ is equivalent to the condition that $v$ lie in $\ker \mu'(m)$. However, since $\xi$ is a clean value of $\mu$, we have $\ker \mu'(m) = T_m(\mu^{-1}(\xi))$, as claimed.

Now, the $G$-equivariance of $\mu$ implies that $\mu^{-1}(\xi) \cap (G \cdot m) = G_\xi \cdot m$ for all $m \in \mu^{-1}(\xi)$. In particular, $T_m(G_\xi \cdot m) \subseteq T_m(\mu^{-1}(\xi)) \cap T_m(G \cdot m)$. To demonstrate the reverse inclusion, suppose that $v$ lies in both $T_m(\mu^{-1}(\xi))$ and $T_m(G \cdot m)$. Then $v = \lambda_*(x)(m)$ for some $x \in \mathfrak{g}$, and, by the $G$-equivariance of the momentum mapping and the assumption that $\xi$ is clean (so that $T_m(\mu^{-1}(\xi)) = \ker \mu'(m)$) we have

$$0 = \mu'(m)(v) = \mu'(m)(\lambda_*(x)(m)) = (\text{Ad}^*)_*(x)(\mu(m)) = (\text{Ad}^*)_*(x)(\xi)$$

so that $x$ must lie in $\mathfrak{g}_\xi$. Consequently, $v = \lambda_*(x)(m)$ is tangent to the orbit $G_\xi \cdot m$ and thus,

$$T_m(\mu^{-1}(\xi)) \cap T_m(G \cdot m) = \ker \mu'(m) \cap T_m(G \cdot m) = T_m(G_\xi \cdot m).$$

As a result, since the $\Omega$-complementary spaces $T_m(\mu^{-1}(\xi))$ and $T_m(G \cdot m)$ intersect in the tangents to the $G_\xi$-orbits, it follows that if $\tilde{\Omega}_\xi$ denotes the pullback of $\Omega$ to $\mu^{-1}(\xi)$, then the null space of $\tilde{\Omega}_\xi$ at $m$ is precisely $T_m(G_\xi \cdot m)$.

Finally, let us assume, as in the theorem, that there is a smooth manifold structure on the orbit space $M_\xi = G_\xi \backslash \mu^{-1}(\xi)$ so that the orbit space projection $\pi_\xi \colon \mu^{-1}(\xi) \to M_\xi$ is a smooth submersion. Since $\tilde{\Omega}_\xi$ is clearly $G_\xi$ invariant and closed and moreover, since its null space at each point of $M$ is precisely the tangent space to the fibers of $\pi_\xi$, it follows that there exists a unique "push down" 2-form $\Omega_\xi$ on $M_\xi$ as described in the statement of the theorem. That $\Omega_\xi$ is closed and non-degenerate is now immediate. $\qquad\square$

The point of Theorem 2 is that, even though the quotient of a symplectic manifold by a symplectic group action is not, in general, a symplectic manifold, there *is* a way to produce a family of symplectic quotients parametrized by the elements of the space $\mathfrak{g}^*$. The quotients $M_\xi$ often turn out to be quite interesting, even when the original symplectic manifold $M$ is very simple.

Before I pass on to the examples, let me make a few comments about the hypotheses in Theorem 2.

First, there will always be clean values $\xi$ of $\mu$ for which $\mu^{-1}(\xi)$ is not empty (even when there are no such regular values). This follows because, if we look at the closed subset $D_\mu \subset M$ consisting of points $m$ where $\mu'(m)$ does not reach its maximum rank, then $\mu(D_\mu)$ can be shown (by a sort of Sard's Theorem argument) to be a proper subset of $\mu(M)$. Meanwhile, it is not hard to show that any element $\xi \in \mu(M)$ that does not lie in $\mu(D_\mu)$ is clean.

Second, it quite frequently does happen that the $G_\xi$-orbit space $M_\xi$ has a manifold structure for which $\pi_\xi$ is a submersion. This can be guaranteed by various hypotheses that are often met with in practice. For example, if $G_\xi$ is compact and acts freely on $\mu^{-1}(\xi)$, then $M_\xi$ will be a manifold. (More generally, if the orbits $G_\xi \cdot m$ are compact and all of the stabilizer subgroups $G_m \subset G_\xi$ are conjugate in $G_\xi$, then $M_\xi$ will have a manifold structure of the required kind.)

Weaker hypotheses also work. Basically, one needs to know that, at every point $m$ of $\mu^{-1}(\xi)$, there is a smooth *slice* to the action of $G_\xi$, i.e., a smoothly embedded disk $D$ in $\mu^{-1}(\xi)$ that passes through $m$ and intersects each $G_\xi$-orbit in $G_\xi \cdot D$ transversely and in exactly one point. (Compare the construction of a smooth structure on each $G$-orbit in Theorem 1 of Lecture 3.)

Even when there is not a slice around each point of $\mu^{-1}(\xi)$, there is very often a *near-slice*, i.e., a smoothly embedded disk $D$ in $\mu^{-1}(\xi)$ that passes through $m$ and intersects each $G_\xi$-orbit in $G_\xi \cdot D$ transversely and in a finite number of points. In this case, the quotient space $M_\xi$ inherits the structure of a symplectic *orbifold*, and these 'generalized manifolds' have turned out to be quite useful.

Finally, it is worth computing the dimension of $M_\xi$ when it does turn out to be a manifold. Let $G_m \subset G$ be the stabilizer of $m \in \mu^{-1}(\xi)$. I leave as an exercise for the reader to check that

$$\dim M_\xi = \dim M - \dim G - \dim G_\xi + 2 \dim G_m$$
$$= \dim M - 2 \dim G/G_m + \dim G/G_\xi.$$

Since we will see so many examples in the next Lecture, I will content myself with only mentioning two here:

• Let $M = T^*G$ and let $G$ act on $T^*G$ on the left in the obvious way. Then the reader can easily check that, for each $x \in \mathfrak{g}$, we have $\rho(x)(\alpha) = \alpha\big(Y_x(a)\big)$ for all $\alpha \in T_a^*G$ where, as usual, $Y_x$ denotes the right invariant vector field on $G$ whose value at $e$ is $x \in \mathfrak{g}$. Hence, $\mu \colon T^*G \to \mathfrak{g}^*$ is given by $\mu(\alpha) = R^*_{\pi(\alpha)}(\alpha)$.

Consequently, $\mu^{-1}(\xi) \subset T^*G$ is merely the graph in $T^*G$ of the left-invariant 1-form $\omega_\xi$ (i.e., the left-invariant 1-form whose value at $e$ is $\xi \in \mathfrak{g}^*$). Thus, we can use $\omega_\xi$ as a section of $T^*G$ to pull back $\Omega$ (the canonical symplectic form on $T^*G$, which is clearly $G$-invariant) to get the 2-form $d\omega_\xi$ on $G$. As we already saw in Lecture 5, and is now borne out by Theorem 2, the null space of $d\omega_\xi$ at any point $a \in G$ is $T_a aG_\xi$, the quotient by $G_\xi$ is merely the coadjoint orbit $G/G_\xi$, and the symplectic structure $\Omega_\xi$ is just the one we already constructed.

Note, by the way, that every value of $\mu$ is clean in this example (in fact, they are all regular), even though the dimensions of the quotients $G/G_\xi$ vary with $\xi$.

• Let $G = \mathrm{SO}(3)$ act on $\mathbb{R}^6 = T^*\mathbb{R}^3$ by the extension of rotation about the origin in $\mathbb{R}^3$. Then, in standard coordinates $(x, y)$ (where $x, y \in \mathbb{R}^3$), the action is simply $g \cdot (x, y) = (gx, gy)$, and the symplectic form is $\Omega = dx \cdot dy = {}^t dx \wedge dy$.

We can identify $\mathfrak{so}(3)^*$ with $\mathfrak{so}(3)$ itself by interpreting $a \in \mathfrak{so}(3)$ as the linear functional $b \mapsto -\mathrm{tr}(ab)$. It is easy to see that the co-adjoint action in this case gets identified with the adjoint action.

We compute that $\rho(a)(x, y) = -{}^t xay$, so it follows without too much difficulty that, with respect to our identification of $\mathfrak{so}(3)^*$ with $\mathfrak{so}(3)$, we have $\mu(x, y) = x{}^t y - y{}^t x$.

The reader can check that all of the values of $\mu$ are clean except for $0 \in \mathfrak{so}(3)$. Even this value would be clean if, instead of taking $M$ to be all of $\mathbb{R}^6$, we let $M$ be $\mathbb{R}^6$ minus the origin $(x, y) = (0, 0)$.

I leave it to the reader to check that the $G$-invariant map $P \colon \mathbb{R}^6 \to \mathbb{R}^3$ defined by

$$P(x, y) = (x \cdot x, \ x \cdot y, \ y \cdot y)$$

maps the set $\mu^{-1}(0)$ onto the "cone" consisting of those points $(a, b, c) \in \mathbb{R}^3$ with $a, c \geq 0$ and $b^2 = ac$ and the fibers of $P$ are the $G_0$-orbits of the points in $\mu^{-1}(0)$.

For $\xi \neq 0$, the $P$-image of the set $\mu^{-1}(\xi)$ is one nappe of the hyperboloid of two sheets described as $ac - b^2 = -\mathrm{tr}(\xi^2)$. The reader should compute the area forms $\Omega_\xi$ on these sheets.

# Exercise Set 7:

# Classical Reduction

**1**. Let $M$ be the torus $\mathbb{R}^2/\mathbb{Z}^2$ and let $dx$ and $dy$ be the standard 1-forms on $M$. Let $\Omega = dx \wedge dy$. Show that the "translation action" $(a,b) \cdot [x,y] = [x+a, y+b]$ of $\mathbb{R}^2$ on $M$ is symplectic but not Hamiltonian.

**2**. Let $(M, \Omega)$ be a connected symplectic manifold and let $\lambda: G \times M \to M$ be a Hamiltonian group action.
 (i) Prove that, if $\rho: \mathfrak{g} \to C^\infty(M)$ is a linear mapping that satisfies $\lambda_*(x) \lrcorner \Omega = d\big(\rho(x)\big)$, then $\lambda_*\big([x,y]\big) \lrcorner \Omega = d\big(\{\rho(x), \rho(y)\}\big)$.
 (ii) Show that the associated linear mapping $c_\rho: \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ defined in the text does indeed satisfy $c_\rho\big([x,y],z\big) + c_\rho\big([y,z],x\big) + c_\rho\big([z,x],y\big) = 0$ for all $x,y,z \in \mathfrak{g}$. (Hint: Use the fact the Poisson bracket satisfies the Jacobi identity and that the Poisson bracket of a constant function with any other function is zero.)

**3**. LIE ALGEBRA COHOMOLOGY. The purpose of this exercise is to acquaint the reader with the rudiments of Lie algebra cohomology.

 The Lie bracket of a Lie algebra $\mathfrak{g}$ can be regarded as a linear map $\partial: \Lambda^2(\mathfrak{g}) \to \mathfrak{g}$. The dual of this linear map is a map $-\delta: \mathfrak{g}^* \to \Lambda^2(\mathfrak{g}^*)$. (Thus, for $\xi \in \mathfrak{g}^*$, we have $\delta\xi(x,y) = -\xi\big([x,y]\big)$.) This map $\delta$ can be extended uniquely to a graded, degree-one derivation $\delta: \Lambda^*(\mathfrak{g}^*) \to \Lambda^*(\mathfrak{g}^*)$.

 (i) For any $c \in \Lambda^2(\mathfrak{g}^*)$, show that $\delta c(x,y,z) = -c\big([x,y],z\big) - c\big([y,z],x\big) - c\big([z,x],y\big)$. (Hint: Every $c \in \Lambda^2(\mathfrak{g}^*)$ is a sum of wedge products $\xi \wedge \eta$ where $\xi, \eta \in \mathfrak{g}^*$.) Conclude that the Jacobi identity in $\mathfrak{g}$ is equivalent to the condition that $\delta^2 = 0$ on all of $\Lambda^*(\mathfrak{g}^*)$.
Thus, for any Lie algebra $\mathfrak{g}$, we can define the *k'th cohomology group* of $\mathfrak{g}$, denoted $H^k(\mathfrak{g})$, as the kernel of $\delta$ in $\Lambda^k(\mathfrak{g}^*)$ modulo the subspace $\delta\big(\Lambda^{k-1}(\mathfrak{g}^*)\big)$.
 (ii) Let $G$ be a Lie group whose Lie algebra is $\mathfrak{g}$. For each $\Phi \in \Lambda^k(\mathfrak{g}^*)$, define $\omega_\Phi$ to be the left-invariant $k$-form on $G$ whose value at the identity is $\Phi$. Show that $d\omega_\Phi = \omega_{\delta\Phi}$. (Hint: the space of left-invariant forms on $G$ is clearly closed under exterior derivative and is generated over $\mathbb{R}$ by the left-invariant 1-forms. Thus, it suffices to prove this formula for $\Phi$ of degree 1. Why?)
Thus, the cohomology groups $H^k(\mathfrak{g})$ measure "closed-mod-exact" in the space of left-invariant forms on $G$. If $G$ is compact, then these cohomology groups are isomorphic to the corresponding deRham cohomology groups of the manifold $G$.
 (iii) (The Whitehead Lemmas) Show that if the Killing form of $\mathfrak{g}$ is non-degenerate, then $H^1(\mathfrak{g}) = H^2(\mathfrak{g}) = 0$. (Hint: You should have already shown that if $\kappa$ is non-degenerate, then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. Show that this implies that $H^1(\mathfrak{g})=0$. Next show that for $\Phi \in \Lambda^2(\mathfrak{g}^*)$, we can write $\Phi(x,y) = \kappa(Lx,y)$ where $L: \mathfrak{g} \to \mathfrak{g}$ is skew-symmetric. Then show that if $\delta\Phi = 0$, then $L$ is a derivation of $\mathfrak{g}$. Now see Exercise 3.3, part (iv).)

**4**. HOMOGENEOUS SYMPLECTIC MANIFOLDS. Suppose that $(M, \Omega)$ is a symplectic manifold and suppose that there exists a transitive symplectic action $\lambda: G \times M \to M$ where $G$ is a group whose Lie algebra satisfies $H^1(\mathfrak{g}) = H^2(\mathfrak{g}) = 0$. Show that there is a $G$-equivariant symplectic covering map $\pi: M \to G/G_\xi$ for some $\xi \in \mathfrak{g}^*$. Thus, up to passing to covers, the only symplectic homogeneous spaces of a Lie group satisfying $H^1(\mathfrak{g}) = H^2(\mathfrak{g}) = 0$ are the coadjoint orbits. This result is usually associated with the names Kostant, Souriau, and Symes.

(Hint: Since $G$ acts homogeneously on $M$, it follows that, as $G$-spaces, $M = G/H$ for some closed subgroup $H \subset G$ that is the stabilizer of a point $m$ of $M$. Let $\phi: G \to M$ be $\phi(g) = g \cdot m$. Now consider the left-invariant 2-form $\phi^*(\Omega)$ on $G$ in light of the previous Exercise. Why do we also need the hypothesis that $H^1(\mathfrak{g}) = 0$?)

Remark: This characterization of homogeneous symplectic spaces is sometimes misquoted. Either the covering ambiguity is overlooked or else, instead of hypotheses about the cohomology groups, sometimes compactness is assumed, either for $M$ or $G$. The example of $S^1 \times S^1$ acting on itself and preserving the bi-invariant area form shows that compactness is not generally helpful. Here is an example that shows that you must allow for the covering possibility: Let $H \subset \mathrm{SL}(2, \mathbb{R})$ be the subgroup of diagonal matrices with *positive* entries on the diagonal. Then $\mathrm{SL}(2, \mathbb{R})/H$ has an $\mathrm{SL}(2, \mathbb{R})$-invariant area form, but it double covers the associated coadjoint orbit.

**5**. Verify the claim made in the text that, if there exists a $G$-invariant 1-form $\omega$ on $M$ so that $d\omega = \Omega$, then the formula $\rho(x) = -\omega\big(\lambda_*(x)\big)$ yields a lifting $\rho$ for which $c_\rho = 0$.

**6**. Show that if $\mathbb{R}^2$ acts on itself by translation then, with respect to the standard area form $\Omega = dx \wedge dy$, this action is Hamiltonian but not Poisson.

**7**. Verify the claim made in the proof of Theorem 1 that the following identity holds for all $a \in G$, all $y \in \mathfrak{g}$, and all $m \in M$:

$$\lambda_a'\big(\lambda_*(y)(m)\big) = \lambda_*\big(\mathrm{Ad}(a)y\big)(a \cdot m).$$

**8**. Here are a few mechanical exercises that turn out to be useful in calculations:

(i) Show that if $\lambda_i : G \times M_i \to M_i$ for $i = 1, 2$ are Poisson actions on symplectic manifolds $(M_i, \Omega_i)$ with corresponding momentum mappings $\mu_i : M_i \to \mathfrak{g}^*$, then the induced product action of $G$ on $M = M_1 \times M_2$ (where $M$ is endowed with the product symplectic structure) is also Poisson, with momentum mapping $\mu : M \to \mathfrak{g}^*$ given by $\mu = \mu_1 \circ \pi_1 + \mu_2 \circ \pi_2$, where $\pi_i : M \to M_i$ is the projection onto the $i$-th factor.

(ii) Show that if $\lambda : G \times M \to M$ is a Poisson action of a connected group $G$ on a symplectic manifold $(M, \Omega)$ with equivariant momentum mapping $\mu : M \to \mathfrak{g}^*$ and $H \subset G$ is a (connected) Lie subgroup, then the restricted action of $H$ on $M$ is also Poisson and the associated momentum mapping is the composition of $\mu$ with the natural mapping $\mathfrak{g}^* \to \mathfrak{h}^*$ induced by the inclusion $\mathfrak{h} \hookrightarrow \mathfrak{g}$.

E.7.2

(iii) Let $(V, \Omega)$ be a symplectic vector space and let $G = \mathrm{Sp}(V, \Omega)$ ($\simeq \mathrm{Sp}(n, \mathbb{R})$) where the dimension of $V$ is $2n$. Show that the natural action of $G$ on $V$ is Poisson, with momentum mapping $\mu(x) = -\frac{1}{2}\big(x \otimes (x \lrcorner \Omega)\big)$. (Use the identification of $\mathfrak{g} = \mathfrak{sp}(V, \Omega)$ with $\mathfrak{g}^*$ defined by the nondegenerate quadratic form $\langle a, b \rangle = \mathrm{tr}(ab)$.) Show that the mapping $S^2(V) \to \mathfrak{sp}(V, \Omega)$ defined on decomposables by

$$x \circ y \mapsto -\frac{1}{2}\big(x \otimes (y \lrcorner \Omega) + y \otimes (x \lrcorner \Omega)\big)$$

is an isomorphism of $\mathrm{Sp}(V, \Omega)$-representations. Using this isomorphism, we can interpret the momentum mapping as the quadratic mapping $\tilde{\mu} : V \to S^2(V)$ defined by the rule $\tilde{\mu}(x) = \frac{1}{2}x^2$. What are the clean values of $\mu$? (There are no regular values.) Let $M_k$ be the product of $k$ copies of $V$ and let $G$ act 'diagonally' on $M_k$. Discuss the clean values and regular values (if any) of $\mu_k : M_k \to \mathfrak{g}^*$. What can you say about the corresponding symplectic quotients? (It may help to note that the group $\mathrm{O}(k)$ acts on $M_k$ in such a way that it commutes with the diagonal action and the corresponding momentum mapping.)

**9**. THE SHIFTING TRICK. It turns out that reduction at a general $\xi \in \mathfrak{g}$ can be reduced to reduction at $0 \in \mathfrak{g}$. Here is how this can be done: Suppose that $\lambda : G \times M \to M$ is a Poisson action on the symplectic manifold $(M, \Omega)$, that $\mu : M \to \mathfrak{g}^*$ is a corresponding momentum mapping, and that $\xi$ is an element of $\mathfrak{g}^*$. Let $\big(M^\xi, \Omega^\xi\big)$ be the symplectic product of $(M, \Omega)$ with $(G \cdot \xi, \Omega_\xi)$ and let $\mu^\xi : M^\xi \to \mathfrak{g}^*$ be the corresponding combined momentum mapping. (Thus, by the computation for coadjoint orbits done in the text, $\mu^\xi(m, a \cdot \xi) = \mu(m) - a \cdot \xi$.)

 (i) Show that $0 \in \mathfrak{g}^*$ is a clean value for $\mu^\xi$ if and only if $\xi$ is a clean value for $\mu$.

Assume for the rest of the problem that $\xi$ is a clean value of $\mu$.

(ii) There is a natural identification of the $G$-orbits in $(\mu^\xi)^{-1}(0) \subset M^\xi$ with the $G_\xi$-orbits in $\mu^{-1}(\xi)$ and that there is a smooth structure on $G \backslash (\mu^\xi)^{-1}(0)$ for which the map $\pi_0^\xi : (\mu^\xi)^{-1}(0) \to G \backslash (\mu^\xi)^{-1}(0)$ is a smooth submersion if and only if there is a smooth structure on $G_\xi \backslash \mu^{-1}(\xi)$ for which the map $\pi_\xi : \mu^{-1}(\xi) \to G_\xi \backslash \mu^{-1}(\xi)$ is a smooth submersion. In this case, the natural identification of the two quotient spaces is a diffeomorphism.

(iii) This natural identification is a symplectomorphism of $\big((M^\xi)_0, (\Omega^\xi)_0\big)$ with $(M_\xi, \Omega_\xi)$.

This shifting trick will be useful when we discuss Kähler reduction in the next Lecture.

**10**. MATRIX CALCULATIONS. The purpose of this exercise is to let you get some practice in a case where everything can be written out in coordinates.

Let $G = \mathrm{GL}(n, \mathbb{R})$ and let $Q: \mathfrak{gl}(n, \mathbb{R}) \to \mathbb{R}$ be a non-degenerate quadratic form. Show that if we use the inclusion mapping $x: \mathrm{GL}(n, \mathbb{R}) \to \mathcal{M}_{n \times n}$ as a coordinate chart, then, in the associated canonical coordinates $(x, p)$, the Lagrangian $L$ takes the form $L = \frac{1}{2}\langle x^{-1}p, x^{-1}p \rangle_Q$. Show also that $\omega_L = \langle x^{-1}p, x^{-1}dx \rangle_Q$.

Now compute the expression for the momentum mapping $\mu$ and the Euler-Lagrange equations for motion under the Lagrangian $L$. Show directly that $\mu$ is constant on the solutions of the Euler-Lagrange equations.

E.7.3

Suppose that $Q$ is Ad-invariant, i.e., $Q\big(\mathrm{Ad}(g)(x)\big) = Q(x)$ for all $g \in G$ and $x \in \mathfrak{g}$. Show that the constancy of $\mu$ is equivalent to the assertion that $p\, x^{-1}$ is constant on the solutions of the Euler-Lagrange equations. Show that, in this case, the $L$-critical curves in $G$ are just the curves $\gamma(t) = \gamma_0\, e^{tv}$ where $\gamma_0 \in G$ and $v \in \mathfrak{g}$ are arbitrary.

Finally, repeat all of these constructions for the general Lie group $G$, translating everything into invariant notation (as opposed to matrix notation).

**11**. EULER'S EQUATION. Look back over the example given in the Lecture of left-invariant metrics on Lie groups. Suppose that $\gamma: \mathbb{R} \to G$ is an $L$-critical curve. Define $\xi(t) = \tau_Q\big(\omega(\dot\gamma(t))\big)$. Thus, $\xi: \mathbb{R} \to \mathfrak{g}^*$. Show that the image of $\xi$ lies on a single coadjoint orbit. Moreover, show that $\xi$ satisfies *Euler's Equation:*

$$\dot\xi + \mathrm{ad}^*\big(\tau_Q^{-1}(\xi)\big)(\xi) = 0.$$

The reason Euler's Equation is so remarkable is that it only involves "half of the variables" of the curve $\dot\gamma$ in $TG$.

Once a solution to Euler's Equation is found, the equation for finding the original curve $\gamma$ is just $\dot\gamma = L'_\gamma\big(\tau_Q^{-1}(\xi)\big)$, which is a Lie equation for $\gamma$ and hence is amenable to Lie's method of reduction.

Actually more is true. Show that, if we set $\xi(0) = \xi_0$, then the equation $\mathrm{Ad}^*(\gamma)(\xi) = \xi_0$ determines the solution $\gamma$ of the Lie equation with initial condition $\gamma(0) = e$ up to right multiplication by a curve in the stabilizer subgroup $G_{\xi_0}$. Thus, we are reduced to solving a Lie equation for a curve in $G_{\xi_0}$. (It may be of some interest to recall that the stabilizer of the generic element $\eta \in \mathfrak{g}^*$ is an abelian group. Of course, for such $\eta$, the corresponding Lie equation can be solved by quadratures.)

**12**. PROJECT: ANALYSIS OF THE RIGID BODY IN $\mathbb{R}^3$. Go back to the example of the motion of a rigid body in $\mathbb{R}^3$ presented in Lecture 4. Use the information provided in the previous two Exercises to show that the equations of motion for a free rigid body are integrable by quadratures. You will want to first compute the coadjoint action and describe the coadjoint orbits and their stabilizers.

**13**. Verify that, under the hypotheses of Theorem 2, the dimension of the reduced space $M_\xi$ is given by the formula

$$\dim M_\xi = \dim M - \dim G - \dim G_\xi + 2 \dim G_m$$

where $G_m$ is the stabilizer of any $m \in \mu^{-1}(\xi)$.

(Hint: Show that for any $m \in \mu^{-1}(\xi)$, we have

$$\dim T_m \mu^{-1}(\xi) + \dim T_m\big(G \cdot m\big) = \dim M$$

and then do some arithmetic.)

**14**. In the reduction process, what is the relationship between $M_\xi$ and $M_{\mathrm{Ad}^*(g)(\xi)}$?

**15**. Suppose that $\lambda: G \times M \to M$ is a Poisson action and that $Y$ is a symplectic vector field on $M$ that is $G$-invariant. Then according to Proposition 1, $Y$ is tangent to each of the submanifolds $\mu^{-1}(\xi)$ (when $\xi$ is a clean value of $\mu$). Show that, when the symplectic quotient $M_\xi$ exists, then there exists a unique vector field $Y_\xi$ on $M_\xi$ that satisfies $Y_\xi\big(\pi_\xi(m)\big) = \pi'_\xi\big(Y(m)\big)$. Show also that $Y_\xi$ is symplectic. Finally show that, given an integral curve $\gamma: \mathbb{R} \to M_\xi$ of $Y_\xi$, then the problem of lifting this to an integral curve of $Y$ is reducible by "finite" operations to solving a Lie equation for $G_\xi$.

This procedure is extremely helpful for two reasons: First, since $M_\xi$ is generally quite a bit smaller than $M$, it should, in principle, be easier to find integral curves of $Y_\xi$ than integral curves of $Y$. For example, if $M_\xi$ is two dimensional, then $Y_\xi$ can be integrated by quadratures (Why?). Second, it very frequently happens that $G_\xi$ is a solvable group. As we have already seen, when this happens the "lifting problem" can be integrated by (a sequence of) quadratures.

# Lecture 8:

# Recent Applications of Reduction

In this Lecture, we will see some examples of symplectic reduction and its generalizations in somewhat non-classical settings.

In many cases, we will be concerned with extra structure on $M$ that can be carried along in the reduction process to produce extra structure on $M_\xi$. Often this extra structure takes the form of a Riemannian metric with special holonomy, so we begin with a short review of this topic.

**Riemannian Holonomy.** Let $M^n$ be a connected and simply connected $n$-manifold, and let $g$ be a Riemannian metric on $M$. Associated to $g$ is the notion of *parallel transport* along curves. Thus, for each (piecewise $C^1$) curve $\gamma \colon [0,1] \to M$, there is associated a linear mapping $P_\gamma \colon T_{\gamma(0)}M \to T_{\gamma(1)}M$, called parallel transport along $\gamma$, which is an isometry of vector spaces and which satisfies the conditions $P_{\bar\gamma} = P_\gamma^{-1}$ and $P_{\gamma_2\gamma_1} = P_{\gamma_2} \circ P_{\gamma_1}$ where $\bar\gamma$ is the path defined by $\bar\gamma(t) = \gamma(1-t)$ and $\gamma_2\gamma_1$ is defined only when $\gamma_1(1) = \gamma_2(0)$ and, in this case, is given by the formula

$$\gamma_2\gamma_1(t) = \begin{cases} \gamma_1(2t) & \text{for } 0 \le t \le \frac{1}{2}, \\ \gamma_2(2t-1) & \text{for } \frac{1}{2} \le t \le 1. \end{cases}$$

These properties imply that, for any $x \in M$, the set of linear transformations of the form $P_\gamma$ where $\gamma(0) = \gamma(1) = x$ is a subgroup $H_x \subset \mathrm{O}(T_xM)$ and that, for any other point $y \in M$, we have $H_y = P_\gamma H_x P_{\bar\gamma}$ where $\gamma \colon [0,1] \to M$ satisfies $\gamma(0) = x$ and $\gamma(1) = y$. Because we are assuming that $M$ is simply connected, it is easy to show that $H_x$ is actually connected and hence is a subgroup of $\mathrm{SO}(T_xM)$.

Élie Cartan was the first to define and study $H_x$. He called it the *holonomy* of $g$ at $x$. He assumed that $H_x$ was always a closed Lie subgroup of $\mathrm{SO}(T_xM)$, a result that was only later proved by Borel and Lichnerowitz (see [KN]).

Georges de Rham, a student of Cartan, proved that, if there is a splitting $T_xM = V_1 \oplus V_2$ that remains invariant under all the action of $H_x$, then, in fact, the metric $g$ is locally a product metric in the following sense: The metric $g$ can be written as a sum of the form $g = g_1 + g_2$ in such a way that, for every point $y \in M$ there exists a neighborhood $U$ of $y$, a coordinate chart $(x_1, x_2) \colon U \to \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, and metrics $\bar g_i$ on $\mathbb{R}^{d_i}$ so that $g_i = x_i^*(\bar g_i)$.

He also showed that in this reducible case the holonomy group $H_x$ is a direct product of the form $H_x^1 \times H_x^2$ where $H_x^i \subset \mathrm{SO}(V_i)$. Moreover, it turns out (although this is not obvious) that, for each of the factor groups $H_x^i$, there is a submanifold $M_i \subset M$ so that $T_xM_i = V_i$ and so that $H_x^i$ is the holonomy of the Riemannian metric $g_i$ on $M_i$.

From this discussion it follows that, in order to know which subgroups of $\mathrm{SO}(n)$ can occur as holonomy groups of simply connected Riemannian manifolds, it is enough to find the ones that, in addition, act irreducibly on $\mathbb{R}^n$. Using a great deal of machinery from the theory of representations of Lie groups, M. Berger [Ber] determined a relatively short list

of possibilities for irreducible Riemannian holonomy groups. This list was slightly reduced a few years later, independently by Alexseevski and by Brown and Gray. The result of their work can be stated as follows:

**Theorem 1:** *Suppose that $g$ is a Riemannian metric on a connected and simply connected $n$-manifold $M$ and that the holonomy $H_x$ acts irreducibly on $T_x M$ for some (and hence every) $x \in M$. Then either $(M, g)$ is locally isometric to an irreducible Riemannian symmetric space or else there is an isometry $\iota : T_x M \to \mathbb{R}^n$ so that $H = \iota \, H_x \, \iota^{-1}$ is one of the subgroups of $\mathrm{SO}(n)$ in the following table.*

### Irreducible Holonomies of Non-Symmetric Metrics

| Subgroup | Conditions | Geometrical Type |
|----------|------------|------------------|
| $\mathrm{SO}(n)$ | any $n$ | generic metric |
| $\mathrm{U}(m)$ | $n = 2m > 2$ | Kähler |
| $\mathrm{SU}(m)$ | $n = 2m > 2$ | Ricci-flat Kähler |
| $\mathrm{Sp}(m)\mathrm{Sp}(1)$ | $n = 4m > 4$ | Quaternionic Kähler |
| $\mathrm{Sp}(m)$ | $n = 4m > 4$ | hyperKähler |
| $\mathrm{G}_2$ | $n = 7$ | Associative |
| $\mathrm{Spin}(7)$ | $n = 8$ | Cayley |

A few words of explanation and comment about Theorem 1 are in order.

First, a Riemannian symmetric space is a Riemannian manifold diffeomorphic to a homogeneous space $G/H$ where $H \subset G$ is essentially the fixed subgroup of an involutory homomorphism $\sigma : G \to G$ that is endowed with a $G$-invariant metric $g$ that is also invariant under the involution $\iota : G/H \to G/H$ defined by $\iota(aH) = \sigma(a)H$. The classification of the Riemannian symmetric spaces reduces to a classification problem in the theory of Lie algebras and was solved by Cartan. Thus, the Riemannian symmetric spaces may be regarded as known.

Second, among the holonomies of non-symmetric metrics listed in the table, the ranges for $n$ have been restricted so as to avoid repetition or triviality. Thus, $\mathrm{U}(1) = \mathrm{SO}(2)$ and $\mathrm{SU}(1) = \{e\}$ while $\mathrm{Sp}(1) = \mathrm{SU}(2)$, and $\mathrm{Sp}(1)\mathrm{Sp}(1) = \mathrm{SO}(4)$.

Third, according to S. T. Yau's celebrated proof of the Calabi Conjecture, any compact complex manifold for which the canonical bundle is trivial and that has a Kähler metric also has a Ricci-flat Kähler metric (see [Bes]). For this reason, metrics with holonomy $\mathrm{SU}(m)$ are often referred to as Calabi-Yau metrics.

Finally, I will not attempt to discuss the proof of Theorem 1 in these notes. Even with modern methods, the proof of this result is non-trivial and, in any case, would take us far from our present interests. Instead, I will content myself with the remark that it is now known that every one of these groups does, in fact, occur as the holonomy of a Riemannian metric on a manifold of the appropriate dimension. I refer the reader to [Bes] for a complete discussion.

We will be particularly interested in the Kähler and hyperKähler cases since these cases can be characterized by the condition that the holonomy of $g$ leaves invariant certain closed non-degenerate 2-forms. Hence these cases represent symplectic manifolds with "extra structure", namely a compatible metric.

The basic result will be that, for a manifold $M$ that carries one of these two structures, there is a reduction process that can be applied to suitable group actions on $M$ that preserve the structure.

**Kähler Manifolds and Algebraic Geometry.**

In this section, we give a very brief introduction to Kähler manifolds. These are symplectic manifolds that are also complex manifolds in such a way that the complex structure is "maximally compatible" with the symplectic structure. These manifolds arise with great frequency in Algebraic Geometry, and it is beyond the scope of these Lectures to do more than make an introduction to their uses here.

**Hermitian Linear Algebra.** As usual, we begin with some linear algebra. Let $H \colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ be the hermitian inner product given by

$$H(z, w) = {}^t\bar{z}w = \bar{z}^1\, w^1 + \cdots + \bar{z}^n\, w^n.$$

Then $\mathrm{U}(n) \subset \mathrm{GL}(n, \mathbb{C})$ is the group of complex linear transformations of $\mathbb{C}^n$ that preserve $H$ since $H(Az, Aw) = H(z, w)$ for all $z, w \in \mathbb{C}^n$ if and only if ${}^t\bar{A}A = I_n$.

Now, $H$ can be split into real and imaginary parts as

$$H(z, w) = \langle z, w \rangle + \imath\, \Omega(z, w).$$

It is clear from the relation $H(z, w) = \overline{H(w, z)}$ that $\langle, \rangle$ is symmetric and $\Omega$ is skew-symmetric. I leave it to the reader to show that $\langle, \rangle$ is positive definite and that $\Omega$ is non-degenerate.

Moreover, since $H(z, \imath w) = \imath H(z, w)$, it also follows that $\Omega(z, w) = \langle \imath z, w \rangle$ and $\langle z, w \rangle = \Omega(z, \imath w)$. It easily follows from these equations that, if we let $J \colon \mathbb{C}^n \to \mathbb{C}^n$ denote multiplication by $\imath$, then knowing any two of the three objects $\langle, \rangle$, $\Omega$, or $J$ on $\mathbb{R}^{2n}$ determines the third.

**Definition 1:** Let $V$ be a vector space over $\mathbb{R}$. A non-degenerate 2-form $\Omega$ on $V$ and a complex structure $J \colon V \to V$ are said to be *compatible* if $\Omega(x, Jy) = \Omega(y, Jx)$ for all $x, y \in V$. If the pair $(\Omega, J)$ is compatible, then we say that the pair forms an *Hermitian structure* on $V$ if, in addition, $\Omega(x, Jx) > 0$ for all non-zero $x \in V$. The positive definite quadratic form $g(x, x) = \Omega(x, Jx)$ is called the *associated metric* on $V$.

I leave as an exercise for the reader the task of showing that any two Hermitian structures on $V$ are isomorphic via some invertible endomorphism of $V$.

It is easy to show that, if $g$ is the quadratic form associated to a compatible pair $(\Omega, J)$, then $\Omega(v, w) = g(Jv, w)$. It follows that any two elements of the triple $(\Omega, J, g)$ determine the third.

In an extension of the notion of compatibility, we define a quadratic form $g$ on $V$ to be compatible with a non-degenerate 2-form $\Omega$ on $V$ if the linear map $J\colon V \to V$ defined by the relation $\Omega(v, w) = g(Jv, w)$ satisfies $J^2 = -1$. Similarly, we define a quadratic form $g$ on $V$ to be compatible with a complex structure $J$ on $V$ if $g(Jv, w) = -g(Jw, v)$, so that $\Omega(v, w) = g(Jv, w)$ defines a 2-form on $V$.

**Almost Hermitian Manifolds.** Since our main interest is in symplectic and complex structures, I will introduce the notion of an almost Hermitian structure on a manifold in terms of its almost complex and almost symplectic structures:

**Definition 2:** Let $M^{2n}$ be a manifold. A 2-form $\Omega$ and an almost complex structure $J$ define an *almost Hermitian structure* on $M$ if, for each $m \in M$, the pair $(\Omega_m, J_m)$ defines a Hermitian structure on $T_m M$.

When $(\Omega, J)$ defines an almost Hermitian structure on $M$, the Riemannian metric $g$ on $M$ defined by $g(v) = \Omega(v, Jv)$ is called the *associated metric*.

Just as one must place conditions on an almost symplectic structure in order to get a symplectic structure, there are conditions that an almost complex structure must satisfy in order to be a complex structure.

**Definition 3:** An almost complex structure $J$ on $M^{2n}$ is *integrable* if each point of $M$ has a neighborhood $U$ on which there exists a coordinate chart $z\colon U \to \mathbb{C}^n$ so that $z'(Jv) = \imath\, z'(v)$ for all $v \in TU$. Such a coordinate chart is said to be *J-holomorphic*.

According to the Korn-Lichtenstein theorem, when $n = 1$ all almost complex structures are integrable. However, for $n \geq 2$, one can easily write down examples of almost complex structures $J$ that are not integrable. (See the Exercises.)

When $J$ is an integrable almost complex structure on $M$, the set

$$\mathcal{U}_J = \{(U, z) \mid z\colon U \to \mathbb{C}^n \text{ is } J\text{-holomorphic}\}$$

forms an atlas of charts that are holomorphic on overlaps. Thus, $\mathcal{U}_J$ defines a holomorphic structure on $M$.

The reader may be wondering just how one determines whether an almost complex structure is integrable or not. In the Exercises, you are asked to show that, for an integrable almost complex structure $J$, the identity $\mathcal{L}_{JX} J - J \circ \mathcal{L}_X J = 0$ must hold for all vector fields $X$ on $M$. It is a remarkable result, due to Newlander and Nirenberg, that this condition is sufficient for $J$ to be integrable.

▶ *The reason that I mention this condition is that it shows that integrability is determined by $J$ and its first derivatives in any local coordinate system. This condition can be rephrased as the condition that the vanishing of a certain tensor $N_J$, called the Nijnhuis tensor of $J$ and constructed out of the first-order jet of $J$ at each point, is necessary and sufficient for the integrability of $J$.*

We are now ready to name the various integrability conditions that can be defined for an almost Hermitian manifold.

**Definition 4:** We call an almost Hermitian pair $(\Omega, J)$ on a manifold $M$ *almost Kähler* if $\Omega$ is closed, *Hermitian* if $J$ is integrable, and *Kähler* if $\Omega$ is closed and $J$ is integrable.

We already saw in Lecture 6 that a manifold has an almost complex structure if and only if it has an almost symplectic structure. However, this relationship does not, in general, hold between complex structures and symplectic structures.

**Example:** Here is a complex manifold that has no symplectic structure. Let $\mathbb{Z}$ act on $M = \mathbb{C}^2 \backslash \{0\}$ by $n \cdot z = 2^n z$. This free action preserves the standard complex structure on $M$. Let $N = \mathbb{Z} \backslash \tilde{M}$, then, via the quotient mapping, $N$ inherits the structure of a complex manifold.

However, $N$ is diffeomorphic to $S^1 \times S^3$ as a smooth manifold. Thus $N$ is a compact manifold satisfying $H^2_{dR}(N, \mathbb{R}) = 0$. In particular, by the cohomology ring obstruction discussed in Lecture 6, we see that $M$ cannot be given a symplectic structure.

**Example:** Here is an example due to Thurston, of a compact 4-manifold that has a complex structure and has a symplectic structure, but has no Kähler structure.

Let $H_3 \subset \mathrm{GL}(3, \mathbb{R})$ be the Heisenberg group, defined in Lecture 2 as the set of matrices of the form
$$g = \begin{pmatrix} 1 & x & z + \frac{1}{2}xy \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

The left invariant forms and their structure equations on $H_3$ are easily computed in these coordinates as

$$\omega_1 = dx \qquad\qquad\qquad\qquad d\omega_1 = 0$$
$$\omega_2 = dy \qquad\qquad\qquad\qquad d\omega_2 = 0$$
$$\omega_3 = dz - \tfrac{1}{2}(x\,dy - y\,dx) \qquad d\omega_3 = -\omega_1 \wedge \omega_2$$

Now, let $\Gamma = H_3 \cap \mathrm{GL}(3, \mathbb{Z})$ be the subgroup of $H_3$ consisting of those elements of $H_3$ all of whose entries are integers. Let $X = \Gamma \backslash H_3$ be the space of right cosets of $\Gamma$. Since the forms $\omega_i$ are left-invariant, it follows that they are well-defined on $X$ and form a basis for the 1-forms on $X$.

Now let $M = X \times S^1$ and let $\omega_4 = d\theta$ be the standard 1-form on $S^1$. Then the forms $\omega_i$ for $1 \le i \le 4$ form a basis for the 1-forms on $M$. Since $d\omega_4 = 0$, it follows that the 2-form

$$\Omega = \omega_1 \wedge \omega_3 + \omega_2 \wedge \omega_4$$

is closed and non-degenerate on $M$. Thus, $M$ has a symplectic structure.

Next, I want to construct a complex structure on $M$. In order to do this, I will produce the appropriate local holomorphic coordinates on $M$. Let $\tilde{M} = H_3 \times \mathbb{R}$ be the simply connected cover of $M$ with coordinates $(x, y, z, \theta)$. We regard $\tilde{M}$ as a Lie group. Define the functions $w^1 = x + \imath y$ and $w^2 = z + \imath \left( \theta + \frac{1}{4}(x^2 + y^2) \right)$ on $\tilde{M}$. Then I leave to the reader to check that, if $g_0$ is the element of $\tilde{M}$ with coordinates $(x_0, y_0, z_0, \theta_0)$, then

$$L^*_{g_0}(w^1) = w^1 + w^1_0 \qquad \text{and} \qquad L^*_{g_0}(w^2) = w^2 + (\imath/2)\bar{w}^1_0\, w^1 + w^2_0.$$

Thus, the coordinates $w^1$ and $w^2$ define a left-invariant complex structure on $\tilde{M}$. Since $M$ is obtained from $\tilde{M}$ by dividing by the obvious left action of $\Gamma \times \mathbb{Z}$, it follows that there is a unique complex structure on $M$ for which the covering projection is holomorphic.

Finally, we show that $M$ cannot carry a Kähler structure. Since $\Gamma$ is a discrete subgroup of $H_3$, the projection $H_3 \to X$ is a covering map. Since $H_3 = \mathbb{R}^3$ as manifolds, it follows that $\pi_1(X) = \Gamma$. Moreover, $X$ is compact since it is the image under the projection of the cube in $H_3$ consisting of those elements whose entries lie in the closed interval $[0, 1]$. On the other hand, since $[\Gamma, \Gamma] \simeq \mathbb{Z}$, it follows that $\Gamma/[\Gamma, \Gamma] \simeq \mathbb{Z}^2$. Thus, $H^1(M, \mathbb{Z}) = H^1(X \times S^1, \mathbb{Z}) = \mathbb{Z}^2 \oplus \mathbb{Z}$. From this, we get that $H^1_{dR}(M, \mathbb{R}) = \mathbb{R}^3$. In particular, the first Betti number of $M$ is 3. Now, it is a standard result in Kähler geometry that the odd degree Betti numbers of a compact Kähler manifold must be even (for example, see [Ch]). Hence, $M$ cannot carry any Kähler metric.

**Example:** Because of the classification of compact complex surfaces due to Kodaira, we know exactly which compact 4-manifolds can carry complex structures. Fernandez, Gotay, and Gray [FGG] have constructed a compact, symplectic 4-manifold $M$ whose underlying manifold is not on Kodaira's list, thus, providing an example of a compact symplectic 4-manifold that carries no complex structure.

The fundamental theorem relating the two "integrability conditions" to the idea of holonomy is the following one. We only give the idea of the proof because a complete proof would require the development of considerable machinery.

**Theorem 2:** *An almost Hermitian structure $(\Omega, J)$ on a manifold $M$ is Kähler if and only if the form $\Omega$ is parallel with respect to the parallel transport of the associated metric $g$.*

PROOF: (IDEA) Once the formulas are developed, it is not difficult to see that the covariant derivatives of $\Omega$ with respect to the Levi-Civita connection of $g$ are expressible in terms of the exterior derivative of $\Omega$ and the Nijnhuis tensor of $J$. Conversely, the exterior derivative of $\Omega$ and the Nijnhuis tensor of $J$ can be expressed in terms of the covariant derivative of $\Omega$ with respect to the Levi-Civita connection of $g$. Thus, $\Omega$ is covariant constant (i.e., invariant under parallel translation with respect to $g$) if and only it is closed and $J$ is integrable. □

▶ *It is worth remarking that $J$ is invariant under parallel transport with respect to $g$ if and only if $\Omega$ is.*

The reason for this is that $J$ is determined from and determines $\Omega$ once $g$ is fixed. The observation now follows, since $g$ is invariant under parallel transport with respect to its own Levi-Civita connection.

**Kähler Reduction.** We are now ready to state the first of the reduction theorems we will discuss in this Lecture.

It turns out that it's a good idea to discuss a special case first.

**Theorem 3:** Kähler Reduction at 0. *Let $(\Omega, g)$ be a Kähler structure on $M^{2n}$. Let $\lambda: G \times M \to M$ be a left action that is Poisson with respect to $\Omega$ and preserves the metric $g$. Let $\mu: M \to \mathfrak{g}^*$ be the associated momentum mapping. Suppose that $0 \in \mathfrak{g}^*$ is a clean value of $\mu$ and that there is a smooth structure on the orbit space $M_0 = G \backslash \mu^{-1}(0)$ for which the natural projection $\pi_0 : \mu^{-1}(0) \to G \backslash \mu^{-1}(0)$ is a smooth submersion. Then there is a unique Kähler structure $(\Omega_0, g_0)$ on $M_0$ defined by the conditions that $\pi_0^*(\Omega_0)$ be equal to the pullback of $\Omega$ to $\mu^{-1}(0) \subset M$ and that $\pi_0: \mu^{-1}(0) \to M_0$ be a Riemannian submersion.*

Proof: Let $\tilde{g}_0$ and $\tilde{\Omega}_0$ be the pullbacks of $g$ and $\Omega$ respectively to $\mu^{-1}(0)$. By hypotheses, $\tilde{g}_0$ and $\tilde{\Omega}_0$ are invariant under the action of $G$.

From Theorem 2 of Lecture 7, we already know that there exists a unique symplectic structure $\Omega_0$ on $M_0$ for which $\pi_0^*(\Omega_0) = \tilde{\Omega}_0$.

Here is how we construct $g_0$. For any $m \in \mu^{-1}(0)$, there is a well defined $\tilde{g}_0$-orthogonal splitting

$$T_m \mu^{-1}(0) = T_m(G \cdot m) \oplus H_m$$

that is clearly $G$-invariant. Since, by hypothesis, $\pi_0: \mu^{-1}(0) \to M_0$ is a submersion, it easily follows that $\pi_0'(m): H_m \to T_{\pi_0(m)} M_0$ is an isomorphism of vector spaces. Moreover, the $G$-invariance of $\tilde{g}$ shows that there is a well-defined quadratic form $g_0(m)$ on $T_{\pi_0(m)} M_0$ that corresponds to the restriction of $\tilde{g}_0$ to $H_m$ under this isomorphism. By the very definition of Riemannian submersion, it follows that $g_0$ is a Riemannian metric on $M_0$ for which $\pi_0$ is a Riemannian submersion.

It remains to show that $(\Omega_0, g_0)$ defines a Kähler structure on $M_0$. First, we show that it is an almost Kähler structure, i.e., that $\Omega_0$ and $g_0$ are actually compatible. Since $\pi_0'(m): H_m \to T_{\pi_0(m)} M_0$ is an isomorphism of vector spaces that identifies $(\Omega_0, g_0)$ with the restriction of $(\Omega, g)$ to $H_m$, it suffices to show that $H_m$ is invariant under the action of $J$.

Here is how we do this. Tracing back through the definitions, we see that $x \in T_m M$ lies in the subspace $H_m$ if and only if $x$ satisfies both of the conditions $\Omega(x, y) = 0$ and $g(x, y) = 0$ for all $y \in T_m(G \cdot m)$. However, since $\Omega(x, y) = g(Jx, y)$ for all $y$, it follows that the necessary and sufficient conditions that $x$ lie in $H_m$ can also be expressed as the two conditions $g(Jx, y) = 0$ and $\Omega(Jx, y) = 0$ for all $y \in T_m(G \cdot m)$. Of course, these conditions are exactly the conditions that $Jx$ lie in $H_m$. Thus, $x \in H_m$ implies that $Jx \in H_m$, as desired.

Finally, in order to show that the almost Kähler structure on $M_0$ is actually Kähler, it must be shown that $\Omega_0$ is parallel with respect to the Levi-Civita connection of $g_0$. This is a straightforward calculation using the structure equations and will not be done here. (Alternatively, to prove that the structure is actually Kähler, one could instead show that the induced almost complex structure is integrable. This is somewhat easier and the interested reader can consult the Exercises, where a proof is outlined.)  $\square$

Now, it seems unreasonable to consider only reduction at $0 \in \mathfrak{g}^*$. However, some caution is in order because the naïve attempt to generalize Theorem 3 to reduction at a general $\xi \in \mathfrak{g}^*$ fails: Let $\lambda : G \times M \to M$ be a Poisson action on a Kähler manifold $(M, \Omega, g)$ that preserves $g$ and let $\mu : M \to \mathfrak{g}^*$ be a Poisson momentum mapping. Then for every clean value $\xi \in \mathfrak{g}^*$ for which the orbit space $M_\xi = G_\xi \backslash \mu^{-1}(\xi)$ has a smooth structure that makes $\pi_\xi : \mu^{-1}(\xi) \to M_\xi$ a smooth submersion, there is a symplectic structure $\Omega_\xi$ on $M_\xi$ that is induced by reduction in the usual way. Moreover, there is a unique metric $g_\xi$ on $M_\xi$ for which $\pi_\xi$ is a Riemannian submersion (when $\mu^{-1}(\xi) \subset M$ is given the induced submanifold metric). Unfortunately, it is *not*, in general, true that $g_\xi$ is compatible with $\Omega_\xi$. (See the Exercises for an example.)

If you examine the proof given above in the general case, you'll see that the main problem is that the 'horizontal space' $H_m$ need not be stable under $J$. In fact, what one knows in the general case is that $H_m$ is $g$-orthogonal to both $T_m G_\xi{\cdot}m$ and to $J(T_m G{\cdot}m)$. However, when $G_\xi$ is a proper subgroup of $G$ (i.e., when $\xi$ is not a fixed point of the coadjoint action), we won't have $T_m G_\xi{\cdot}m = T_m G{\cdot}m$, which is what we needed in the proof to show that $H_m$ is stable under $J$.

In fact, the proof does work when $G_\xi = G$, but this can be seen directly from the fact that, in this case, the 'shifted' momentum mapping $\mu^\xi = \mu - \xi$ still satisfies $G$-equivariance and we are simply performing reduction at 0 for the shifted momentum mapping $\mu^\xi$.

REDUCTION AT KÄHLER COADJOINT ORBITS. Generalizing the case where $G{\cdot}\xi = \{\xi\}$, there is a way to define Kähler reduction at certain values of $\xi \in \mathfrak{g}^*$, by relying on the 'shifting trick' described in the Exercises of Lecture 7:

In many cases, a coadjoint orbit $G{\cdot}\xi \subset \mathfrak{g}^*$ can be equipped with a $G$-invariant metric $h_\xi$ for which the pair $(\Omega_\xi, h_\xi)$ defines a Kähler structure on the orbit $G{\cdot}\xi$. (For example, this is always the case when $G$ is compact.) In such a case, the shifting trick allows us to define a Kähler metric on $M_\xi = G_\xi \backslash \mu^{-1}(0)$ by doing Kähler reduction at 0 on $M \times G{\cdot}\xi$ endowed with the product Kähler structure. In the cases in which there is only one $G$-invariant Kähler metric $h_\xi$ on $G{\cdot}\xi$ that is compatible with $\Omega_\xi$ (and, again, this always holds when $G$ is compact), this defines a canonical Kähler reduction procedure for $\xi \in \mathfrak{g}^*$.

**Example:** KÄHLER REDUCTION IN ALGEBRAIC GEOMETRY. By far the most common examples of Kähler manifolds arise in Algebraic Geometry. Here is a sample of what Kähler reduction yields:

Let $M = \mathbb{C}^{n+1}$ with complex coordinates $z^0, z^1, \ldots, z^n$. We let $z^k = x^k + \imath y^k$ define real coordinates on $M$. Let $G = S^1$ act on $M$ by the rule

$$e^{\imath\theta} \cdot z = e^{\imath\theta} z.$$

Then $G$ clearly preserves the Kähler structure defined by the natural complex structure on $M$ and the symplectic form

$$\Omega = \frac{\imath}{2}{}^t dz \wedge d\bar{z} = dx^1 \wedge dy^1 + \cdots + dx^n \wedge dy^n.$$

The associated metric is easily seen to be just

$$g = {}^t dz \circ d\bar{z} = (dx^1)^2 + (dy^1)^2 + \cdots + (dx^n)^2 + (dy^n)^2.$$

Now, setting $X = \frac{\partial}{\partial\theta}$, we can compute that

$$\lambda_*(X) = x^k \frac{\partial}{\partial y^k} - y^k \frac{\partial}{\partial x^k}.$$

Thus, it follows that

$$d\rho(X) = \lambda_*(X) \lrcorner \, \Omega = -x^k \, dx^k - y^k \, dy^k = d\left(-\tfrac{1}{2}|z|^2\right).$$

Thus, identifying $\mathfrak{g}^*$ with $\mathbb{R}$, we have that $\mu \colon \mathbb{C}^n \to \mathbb{R}$ is merely $\mu(z) = -\tfrac{1}{2}|z|^2$.

It follows that every negative number is a non-trivial clean value for $\mu$. For example, $S^{2n+1} = \mu^{-1}(-\tfrac{1}{2})$. Clearly $G = S^1$ itself is the stabilizer subgroup of all of the values of $\mu$. Thus, $M_{-\frac{1}{2}}$ is the quotient of the unit sphere by the action of $S^1$. Since each $G$-orbit is merely the intersection of $S^{2n+1}$ with a (unique) complex line through the origin, it is clear that $M_{-\frac{1}{2}}$ is diffeomorphic to $\mathbb{CP}^n$.

Since the coadjoint action is trivial, reduction at $\xi = -\tfrac{1}{2}$ will define a Kähler structure on $\mathbb{CP}^n$. It is instructive to compute what this Kähler structure looks like in local coordinates. Let $\mathbb{A}_0 \subset \mathbb{CP}^n$ be the subset consisting of those points $[z^0, \ldots, z^n]$ for which $z^0 \neq 0$. Then $\mathbb{A}_0$ can be parametrized by $\phi : \mathbb{C}^n \to \mathbb{A}_0$ where $\phi(w) = [1, w]$. Now, over $\mathbb{A}_0$, we can choose a section $\sigma \colon \mathbb{A}_0 \to S^{2n+1}$ by the rule

$$\sigma \circ \phi(w) = \frac{(1, w)}{W}$$

where $W^2 = 1 + |w^1|^2 + \cdots + |w^n|^2 > 0$. It follows that

$$
\begin{aligned}
\phi^*(\Omega_{-\frac{1}{2}}) = (\sigma \circ \phi)^*(\Omega) &= \frac{\imath}{2}\left(d\big(\frac{w^k}{W}\big) \wedge d\big(\frac{\bar{w}^k}{W}\big)\right) \\
&= \frac{\imath}{2}\left(\frac{dw^k \wedge d\bar{w}^k}{W^2} + (w^k \, d\bar{w}^k - \bar{w}^k \, dw^k) \wedge \frac{dW}{W^3}\right) \\
&= \frac{\imath}{2}\left(\frac{W^2 \delta_{jk} - \bar{w}^j \, w^k}{W^4}\right) dw^j \wedge d\bar{w}^k.
\end{aligned}
$$

I leave it to the reader to check that the quotient metric (i.e., the one for which the submersion $S^{2n+1} \to \mathbb{CP}^n$ is Riemannian) is given by the formula

$$g_{-\frac{1}{2}} = \left(\frac{W^2 \delta_{jk} - \bar{w}^j \, w^k}{W^4}\right) dw^j \circ d\bar{w}^k.$$

In particular, it follows that the functions $w^k$ are holomorphic functions with respect to the induced almost complex structure, verifying directly that the pair $(\Omega_{-\frac{1}{2}}, g_{-\frac{1}{2}})$ is indeed a Kähler structure on $\mathbb{CP}^n$. Up to a normalizing constant, this is the usual formula for the Fubini-Study Kähler structure on $\mathbb{CP}^n$ in an affine chart.

L.8.9

Of course, the Fubini-Study metric induces a Kähler structure on every complex sub-manifold of $\mathbb{CP}^n$. However, we can just as easily see how this arises from the reduction procedure: If $P(z^0, \ldots, z^n)$ is a non-zero homogeneous polynomial of degree $d$, then the set $\tilde{M}_P = P^{-1}(0) \subset \mathbb{C}^{n+1}$ is a complex subvariety of $\mathbb{C}^{n+1}$ that is invariant under the $S^1$ action since, by homogeneity, we have

$$P(e^{\imath\theta} \cdot z) = e^{\imath d\theta} P(z).$$

It is easy to show that if the variety $\tilde{M}_P$ has no singularity other than $0 \in \mathbb{C}^{n+1}$, then the Kähler reduction of the Kähler structure that it inherits from the standard structure on $\mathbb{C}^{n+1}$ is just the Kähler structure on the corresponding projectivized variety $M_P \subset \mathbb{CP}^n$ that is induced by restriction of the Fubini-Study structure.

**"Example":** FLAT BUNDLES OVER COMPACT RIEMANN SURFACES. The following is not really an example of the theory as we have developed it since it will deal with "infinite dimensional manifolds", however it is suggestive and the formal calculations yield an interesting result. (For a review of the terminology used in this and the next example, see the Appendix.)

Let $G$ be a Lie group with Lie algebra $\mathfrak{g}$, and let $\langle, \rangle$ be a positive definite, Ad-invariant inner product on $\mathfrak{g}$. (For example, if $G = \mathrm{SU}(n)$, we could take $\langle x, y \rangle = -\mathrm{tr}(xy)$.)

Let $\Sigma$ be a connected compact Riemann surface. Then there is a star operation $*\colon \mathcal{A}^1(\Sigma) \to \mathcal{A}^1(\Sigma)$ that satisfies $*^2 = -id$, and $\alpha \wedge *\alpha \geq 0$ for all 1-forms $\alpha$ on $\Sigma$.

Let $P$ be a principal right $G$-bundle over $\Sigma$, and let $\mathrm{Ad}(P) = P \times_{\mathrm{Ad}} \mathfrak{g}$ denote the vector bundle over $M$ associated to the adjoint representation $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathfrak{g})$. Let $\mathrm{Aut}(P)$ denote the group of automorphisms of $P$, also known as the *gauge group* of $P$.

Let $\mathfrak{A}(P)$ denote the space of connections on $P$. Then it is well known that $\mathfrak{A}(P)$ is an affine space modeled on the vector space $\mathcal{A}^1(\mathrm{Ad}(P))$, which consists of the 1-forms on $M$ with values in $\mathrm{Ad}(P)$. Thus, in particular, for every $A \in \mathfrak{A}(P)$, we have a natural isomorphism

$$T_A \mathfrak{A}(P) = \mathcal{A}^1(\mathrm{Ad}(P)).$$

I now want to define a "Kähler" structure on $\mathfrak{A}(P)$. In order to do this, I will define the metric $\mathbf{g}$ and the 2-form $\mathbf{\Omega}$.

First, for $\alpha \in T_A \mathfrak{A}(P)$, I define

$$\mathbf{g}(\alpha) = \int_\Sigma \langle \alpha, *\alpha \rangle.$$

(I extend the operator $*$ in the obvious way to $\mathcal{A}^1(\mathrm{Ad}(P))$.) It is clear that $\mathbf{g}(\alpha) \geq 0$ with equality if and only if $\alpha = 0$. Thus, $\mathbf{g}$ defines a "Riemannian metric" on $\mathfrak{A}(P)$. Since $\mathbf{g}$ is "translation invariant", it "follows" that $\mathbf{g}$ is "flat".

Second, I define $\mathbf{\Omega}$ by the rule:

$$\mathbf{\Omega}(\alpha, \beta) = \int_\Sigma \langle \alpha, \beta \rangle.$$

Since $\boldsymbol{\Omega}(\alpha, \beta) = \mathbf{g}(\alpha, *\beta)$, it follows that $\boldsymbol{\Omega}$ is actually non-degenerate. Moreover, because $\boldsymbol{\Omega}$ too is "translation invariant", it "must" be "parallel" with respect to $\mathbf{g}$.

Thus, $(\boldsymbol{\Omega}, \mathbf{g})$ is a "flat Kähler" structure on $\mathfrak{A}(P)$. Now, I claim that both $\boldsymbol{\Omega}$ and $\mathbf{g}$ are invariant under the natural right action of $\mathsf{Aut}(P)$ on $\mathfrak{A}(P)$. To see this, note that an element $\phi \in \mathsf{Aut}(P)$ determines a map $\varphi \colon P \to G$ by the rule $p \cdot \varphi(p) = \phi(p)$ and that this $\varphi$ satisfies the identity $\varphi(p \cdot g) = g^{-1}\varphi(p)g$. In terms of $\varphi$, the action of $\mathsf{Aut}(P)$ on $\mathfrak{A}(P)$ is given by the classical formula

$$A \cdot \phi = \phi^*(A) = \varphi^*(\omega_G) + \mathrm{Ad}\bigl(\varphi^{-1}\bigr)(A).$$

From this, it follows easily that $\boldsymbol{\Omega}$ and $\mathbf{g}$ are $\mathsf{Aut}(P)$-invariant.

Now, I want to compute the momentum mappping $\boldsymbol{\mu}$. The Lie algebra of $\mathsf{Aut}(P)$, namely $\mathfrak{aut}(P)$, can be naturally identified with $\mathcal{A}^0\bigl(\mathrm{Ad}(P)\bigr)$, the space of sections of the bundle $\mathrm{Ad}(P)$. I leave to the reader the task of showing that the induced map from $\mathfrak{aut}(P)$ to vector fields on $\mathfrak{A}(P)$ is given by $d_A \colon \mathcal{A}^0\bigl(\mathrm{Ad}(P)\bigr) \to \mathcal{A}^1\bigl(\mathrm{Ad}(P)\bigr)$. Thus, in order to construct the momentum mapping, we must find, for each $f \in \mathcal{A}^0\bigl(\mathrm{Ad}(P)\bigr)$, a function $\boldsymbol{\rho}(f)$ on $\mathfrak{A}$ so that the 1-form $d\boldsymbol{\rho}(f)$ is given by

$$d\boldsymbol{\rho}(f)(\alpha) = d_A f \lrcorner \boldsymbol{\Omega}(\alpha) = \int_\Sigma \langle d_A f, \alpha \rangle = -\int_\Sigma \langle f, d_A \alpha \rangle.$$

However, this is easy. We just set

$$\boldsymbol{\rho}(f)(A) = -\int_\Sigma \langle f, F_A \rangle$$

and the reader can easily check that

$$\frac{d}{dt}\Big|_{t=0} (\boldsymbol{\rho}(f)(A + t\alpha)) = -\int_\Sigma \langle f, d_A \alpha \rangle$$

as desired. Finally, using the natural isomorphism

$$\bigl(\mathfrak{aut}(P)\bigr)^* = \bigl(\mathcal{A}^0\bigl(\mathrm{Ad}(P)\bigr)\bigr)^* = \mathcal{A}^2\bigl(\mathrm{Ad}(P)\bigr),$$

we see that (up to sign) the formula for the momentum mapping simply becomes

$$\boldsymbol{\mu}(A) = F_A = dA + \tfrac{1}{2}[A, A].$$

Now, can we do reduction? What we need is a clean value of $\boldsymbol{\mu}$. As a reasonable first guess, let's try 0. Thus, $\boldsymbol{\mu}^{-1}(0)$ consists exactly of the flat connections on $P$ and the reduced space $\mathfrak{M}_0$ should be the flat connections modulo gauge equivalence, i.e., $\boldsymbol{\mu}^{-1}(0)/\mathsf{Aut}(P)$.

How can we tell whether 0 is a clean value? One way to know this would be to know that 0 is a regular value. We have already seen that $\boldsymbol{\mu}'(A)(\alpha) = d_A \alpha$, so we are asking

whether the map $d_A \colon \mathcal{A}^1\big(\mathrm{Ad}(P)\big) \to \mathcal{A}^2\big(\mathrm{Ad}(P)\big)$ is surjective for any flat connection $A$. Now, because $A$ is flat, the sequence

$$0 \longrightarrow \mathcal{A}^0\big(\mathrm{Ad}(P)\big) \xrightarrow{\ d_A\ } \mathcal{A}^1\big(\mathrm{Ad}(P)\big) \xrightarrow{\ d_A\ } \mathcal{A}^2\big(\mathrm{Ad}(P)\big) \longrightarrow 0$$

forms a complex and the usual Hodge theory pairing shows that $H^2(\Sigma, d_A)$ is the dual space of $H^0(\Sigma, d_A)$. Thus, $\boldsymbol{\mu}'(A)$ is surjective if and only if $H^0(\Sigma, d_A) = 0$. Now, an element $f \in \mathcal{A}^0\big(\mathrm{Ad}(P)\big)$ that satisfies $d_A f = 0$ exponentiates to a 1-parameter family of automorphisms of $P$ that commute with the parallel transport of $A$. I leave to the reader to show that $H^0(\Sigma, d_A) = 0$ is equivalent to the condition that the holonomy group $H_A(p) \subset G$ has a centralizer of positive dimension in $G$ for some (and hence every) point of $P$. For example, for $G = \mathrm{SU}(2)$, this would be equivalent to saying that the holonomy groups $H_A(p)$ were each contained in an $S^1 \subset G$.

Let us let $\widetilde{\mathfrak{M}}^* \subset \boldsymbol{\mu}^{-1}(0)$ denote the (open) subset consisting of those flat connections whose holonomy groups have at most discrete centralizers in $G$. If $G$ is compact, of course, this implies that these centralizers are finite. Then it "follows" that $\mathfrak{M}_0^* = \widetilde{\mathfrak{M}}^*/\mathrm{Aut}(P)$ is a Kähler manifold wherever it is a manifold. (In general, at the connections where the centralizer of the holonomy is trivial, one expects the quotient to be a manifold.)

Since the space of flat connections on $P$ modulo gauge equivalence is well-known to be identifiable as the space $R\big(\pi_1(\Sigma, s), G\big) = \mathrm{Hom}\big(\pi_1(\Sigma, s), G\big)/G$ of equivalence classes of representation of $\pi_1(\Sigma)$ into $G$, our discussion leads us to believe that this space (which is finite dimensional) should have a natural Kähler structure on it. This is indeed the case, and the geometry of this Kähler metric is the subject of current interest.

## HyperKähler Manifolds.

In this section, we will generalize the Kähler reduction procedure to the case of manifolds with holonomy $\mathrm{Sp}(m)$, the so-called hyperKähler case.

**Quaternion Hermitian Linear Algebra.** We begin with some linear algebra over the ring $\mathbb{H}$ of *quaternions*. For our purposes, $\mathbb{H}$ can be identified with the vector space of dimension 4 over $\mathbb{R}$ of matrices of the form

$$x = \begin{pmatrix} x^0 + \imath\, x^1 & x^2 + \imath\, x^3 \\ -x^2 + \imath\, x^3 & x^0 - \imath\, x^1 \end{pmatrix} \overset{\text{def}}{=} x^0\, 1 + x^1 i + x^2 j + x^3 k.$$

(We are identifying the 2-by-2 identity matrix with 1 in this representation.) It is easy to see that $\mathbb{H}$ is closed under matrix multiplication. If we define $\bar{x} = x^0 - x^1 i - x^2 j - x^3 k$, then we easily get $\overline{xy} = \bar{y}\bar{x}$ and

$$x\bar{x} = \big((x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2\big)\, 1 = \det(x)\, 1 \overset{\text{def}}{=} |x|^2\, 1.$$

It follows that every non-zero element of $\mathbb{H}$ has a multiplicative inverse. Note that the space of quaternions of unit norm, $S^3$ defined by $|x| = 1$, is simply $\mathrm{SU}(2)$.

Much of the linear algebra that works for the complex numbers can be generalized to the quaternions. However, some care must be taken since $\mathbb{H}$ is not commutative. In the following exposition, it turns out to be most convenient to define vector spaces over $\mathbb{H}$ as *right* vector spaces instead of left vector spaces. Thus, the standard $\mathbb{H}$-vector space of $\mathbb{H}$-dimension $n$ is $\mathbb{H}^n$ (thought of as columns of quaternions of height $n$) where the action of the scalars *on the right* is given by

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \cdot q = \begin{pmatrix} x^1 q \\ \vdots \\ x^n q \end{pmatrix}.$$

With this convention, a quaternion linear map $A \colon \mathbb{H}^n \to \mathbb{H}^m$, i.e., an additive map satisfying $A(v\,q) = A(v)\,q$, can be represented by an $m$-by-$n$ matrix of quaternions acting via matrix multiplication on the *left*.

Let $H \colon \mathbb{H}^n \times \mathbb{H}^n \to \mathbb{H}$ be the "quaternion Hermitian" inner product given by

$$H(z, w) = {}^t\bar{z}w = \bar{z}^1\,w^1 + \cdots + \bar{z}^n\,w^n.$$

Then by our conventions, we have

$$H(z\,q, w) = \bar{q}\,H(z, w) \qquad \text{and} \qquad H(z, w\,q) = H(z, w)\,q.$$

We also have $H(z, w) = \overline{H(w, z)}$, just as before.

We define $\mathrm{Sp}(n) \subset \mathrm{GL}(n, \mathbb{H})$ to be the group of $\mathbb{H}$-linear transformations of $\mathbb{H}^n$ that preserve $H$, i.e., $H(Az, Aw) = H(z, w)$ for all $z, w \in \mathbb{H}^n$. It is easy to see that

$$\mathrm{Sp}(n) = \left\{ A \in \mathrm{GL}(n, \mathbb{H}) \,\middle|\, {}^t\bar{A}A = I_n \right\}.$$

I leave as an exercise for the reader to show that $\mathrm{Sp}(n)$ is a compact Lie group of dimension $2n^2 + n$. Also, it is not difficult to show that $\mathrm{Sp}(n)$ is connected and acts irreducibly on $\mathbb{H}^n$. (see the Exercises)

Now $H$ can be split into one real and three imaginary parts as

$$H(z, w) = \langle z, w \rangle + \Omega_1(z, w)\,i + \Omega_2(z, w)\,j + \Omega_3(z, w)\,k.$$

It is clear from the relations above that $\langle , \rangle$ is symmetric and positive definite and that each of the $\Omega_a$ is skew-symmetric. Moreover, we have the following identities:

$$\langle z, w \rangle = \Omega_1(z, w\,i) = \Omega_2(z, w\,j) = \Omega_3(z, w\,k)$$

and

$$\Omega_2(z, w\,i) = \Omega_3(z, w)$$
$$\Omega_3(z, w\,j) = \Omega_1(z, w)$$
$$\Omega_1(z, w\,k) = \Omega_2(z, w).$$

**Proposition 1:** *The subgroup of* $\mathrm{GL}(4n, \mathbb{R})$ *that fixes the three 2-forms* $(\Omega_1, \Omega_2, \Omega_3)$ *is equal to* $\mathrm{Sp}(n)$.

PROOF: Let $G \subset \mathrm{GL}(4n, \mathbb{R})$ be the subgroup that fixes each of the $\Omega_a$. Clearly we have $\mathrm{Sp}(n) \subset G$.

Now, from the first of the identities above, it follows that each of the forms $\Omega_a$ is non-degenerate. Then, from the second set of these identities, it follows that the subgroup $G$ must also fix the linear transformations of $\mathbb{R}^{4n}$ that represent multiplication on the right by $i$, $j$, and $k$. Of course, this, by definition, implies that $G$ is a subgroup of $\mathrm{GL}(n, \mathbb{H})$. Returning to the first of the identities, it also follows that $G$ must preserve the inner product defined by $\langle, \rangle$. Finally, since we have now seen that $G$ must preserve all of the components of $H$, it follows that $G$ must preserve $H$ as well. However, this was the very definition of $\mathrm{Sp}(n)$. $\qquad \square$

Proposition 1 motivates the way we will want to define HyperKähler structures on manifolds: as triples of 2-forms that satisfy certain conditions. Here is the linear algebra definition on which the manifold definition will be based.

**Definition 5:** Let $V$ be a vector space over $\mathbb{R}$. A *hyperKähler structure* on $V$ is a choice of a triple of non-degenerate 2-forms $(\Omega_1, \Omega_2, \Omega_3)$ that satisfy the following properties: First, the linear maps $R_i$, $R_j$ that are defined by the equations

$$\Omega_2(v, R_i\, w) = \Omega_3(v, w) \qquad \qquad \Omega_1(v, R_j\, w) = -\Omega_3(v, w)$$

satisfy $\left(R_i\right)^2 = \left(R_j\right)^2 = -id$ and skew-commute, i.e., $R_i R_j = -R_j R_i$. Second, if we set $R_k = -R_i R_j$, then

$$\Omega_1(v, R_i\, w) = \Omega_2(v, R_j\, w) = \Omega_3(v, R_k\, w) = \langle v, w \rangle$$

where $\langle, \rangle$ (which is defined by these equations) is a positive definite symmetric bilinear form on $V$. The inner product $\langle, \rangle$ is called the *associated metric* on $V$.

This may seem to be a rather cumbersome definition (and I admit that it is), but it is sufficient to prove the following Proposition (which I leave as an exercise for the reader).

**Proposition 2:** *If* $(\Omega_1, \Omega_2, \Omega_3)$ *is a hyperKähler structure on a real vector space* $V$, *then* $\dim(V) = 4n$ *for some $n$ and, moreover, there is an $\mathbb{R}$-linear isomorphism of $V$ with $\mathbb{H}^n$ that identifies the hyperKähler structure on $V$ with the standard one on $\mathbb{H}^n$.* $\qquad \square$

We are now ready for the analogs of Definitions 3 and 4:

**Definition 6:** If $M$ is a manifold of dimension $4n$, an *almost hyperKähler structure* on $M$ is a triple $(\Omega_1, \Omega_2, \Omega_3)$ of 2-forms on $M$ that have the property that they induce a hyperKähler structure on each tangent space $T_m M$.

**Definition 7:** An almost hyperKähler structure $(\Omega_1, \Omega_2, \Omega_3)$ on a manifold $M^{4n}$ is a hyperKähler structure on $M$ if each of the forms $\Omega_a$ is closed.

At first glance, Definition 7 may seem surprising. After all, it appears to place no conditions on the almost complex structures $R_i$, $R_j$, and $R_k$ that are defined on $M$ by the almost hyperKähler structure on $M$ and one would surely want these to be integrable if the analogy with Kähler geometry is to be kept up. The nice result is that the integrability of these structures comes for free:

**Theorem 4:** *For an almost hyperKähler structure $(\Omega_1, \Omega_2, \Omega_3)$ on a manifold $M^{4n}$, the following are equivalent:*

(1) $d\Omega_1 = d\Omega_2 = d\Omega_3 = 0$.

(2) *Each of the 2-forms $\Omega_a$ is parallel with respect to the Levi-Civita connection of the associated metric.*

(3) *Each of the almost complex structures $R_i$, $R_j$, and $R_k$ are integrable.*

PROOF: (IDEA) The proof of Theorem 4 is much like the proof of Theorem 2. One shows by local calculations in Gauss normal coordinates at any point on $M$ that the covariant derivatives of the forms $\Omega_a$ with respect to the Levi-Civita connection of the associated metric can be expressed in terms of the coefficients of their exterior derivatives and vice-versa. Similarly, one shows that the formulas for the Nijnhuis tensors of the three almost complex structures on $M$ can be expressed in terms of the covariant derivatives of the three 2-forms and vice-versa. This is a rather formidable linear algebra problem, but it is nothing more. I will not do the computation here. $\qquad\square$

Note that Theorem 4 implies that the holonomy $H$ of the associated metric of a hyperKähler structure on $M^{4n}$ must be a subgroup of $\mathrm{Sp}(n)$. If $H$ is a proper subgroup of $\mathrm{Sp}(n)$, then by Theorem 1, the associated metric must be locally a product metric. Now, as is easy to verify, the only products from Berger's List that can appear as subgroups of $\mathrm{Sp}(n)$ are products of the form

$$\{e\}_{n_0} \times \mathrm{Sp}(n_1) \times \cdots \times \mathrm{Sp}(n_k)$$

where $\{e\}_{n_0} \subset \mathrm{Sp}(n_0)$ is just the identity subgroup and $n = n_0 + \cdots + n_k$. Thus, it follows that a hyperKähler structure can be decomposed locally into a product of the 'flat' example with hyperKähler structures whose holonomy is the full $\mathrm{Sp}(n_i)$. (If $M$ is simply connected and the associated metric is complete, then the de Rham Splitting Theorem asserts that $M$ can be globally written as a product of such metrics.) This motivates our calling a hyperKähler structure on $M^{4n}$ *irreducible* if its holonomy is equal to $\mathrm{Sp}(n)$.

The reader may be wondering just how common these hyperKähler structures are (aside from the flat ones of course). The answer is that they are not so easy to come by. The first known non-flat example was the Eguchi-Hanson metric (often called a "gravitational instanton") on $T^*\mathbb{CP}^1$. The first known irreducible example in dimensions greater than 4 was discovered by Eugenio Calabi, who, working independently from Eguchi and Hanson, constructed an irreducible hyperKähler structure on $T^*\mathbb{CP}^n$ for each $n$ that happened to

agree with the Eguchi-Hanson metric for $n = 1$. (We will see Calabi's examples a little further on.)

The first known compact example was furnished by Yau's solution of the Calabi Conjecture:

**Example:** K3 SURFACES. A $K3$ surface is a compact simply connected 2-dimensional complex manifold $S$ with trivial canonical bundle. What this latter condition means is that there is nowhere-vanishing holomorphic 2-form $\Upsilon$ on $S$. An example of such a surface is a smooth algebraic surface of degree 4 in $\mathbb{CP}^3$.

A fundamental result of Siu [Si] is that every $K3$ surface is Kähler, i.e., that there exists a 2-form $\Omega$ on $S$ so that the hermitian structure $(\Omega, J)$ on $S$ is actually Kähler. Moreover, Yau's solution of the Calabi Conjecture implies that $\Omega$ can be chosen so that $\Upsilon$ is parallel with respect to the Levi-Civita connection of the associated metric.

Multiplying $\Upsilon$ by an appropriate constant, we can arrange that $2\,\Omega^2 = \Upsilon \wedge \overline{\Upsilon}$. Since $\Omega \wedge \Upsilon = 0$ and $\Upsilon \wedge \Upsilon = 0$, it easily follows (see the Exercises) that if we write $\Omega = \Omega_1$ and $\Upsilon = \Omega_2 - \imath\,\Omega_3$, then the triple $(\Omega_1, \Omega_2, \Omega_3)$ defines a hyperKähler structure on $S$.

For a long time, the $K3$ surfaces were the only known compact manifolds with hyperKähler structures. In fact, a "proof" was published showing that there were no other compact ones. However, this turned out not to be correct.

**Example:** Let $M^{4n}$ be a simply connected, compact complex manifold (of complex dimension $2n$) with a *holomorphic* symplectic form $\Upsilon$. Then $\Upsilon^n$ is a non-vanishing holomorphic volume form, and hence the canonical bundle of $M$ is trivial. If $M$ has a Kähler structure that is compatible with its complex structure, then, by Yau's solution of the Calabi Conjecture, there is a Kähler metric $g$ on $M$ for which the volume form $\Upsilon^n$ is parallel. This implies that the holonomy of $g$ is a subgroup of $\mathrm{SU}(2n)$. However, this in turn implies that $g$ is Ricci-flat and hence, by a Bochner vanishing argument, that every holomorphic form on $M$ is parallel with respect to $g$. Thus, $\Upsilon$ is also parallel with respect to $g$ and hence the holonomy is a subgroup of $\mathrm{Sp}(n)$. If $M$ can be constructed in such a way that it cannot be written as a non-trivial product of complex submanifolds, then the holonomy of $g$ must act irreducibly on $\mathbb{C}^{2n}$ and hence must equal $\mathrm{Sp}(n)$.

Fujita was the first to construct a simply connected, compact complex 4-manifold that carried a holomorphic 2-form and that could not be written non-trivially as a product. This work is written up in detail in a survey article by [Bea].

**HyperKähler Reduction.** I am now ready to describe another method of constructing hyperKähler structures, known as hyperKähler reduction. This method first appeared in a famous paper by Hitchin, Karlhede, Lindström, and Roček, [HKLR].

**Theorem 5:** *Suppose that $(\Omega_1, \Omega_2, \Omega_3)$ is a hyperKähler structure on $M$ and that there is a left action $\lambda\colon G \times M \to M$ that is Poisson with respect to each of the three symplectic forms $\Omega_a$. Let*

$$\mu = (\mu_1, \mu_2, \mu_3)\colon M \to \mathfrak{g}^* \oplus \mathfrak{g}^* \oplus \mathfrak{g}^*$$

be a $G$-equivariant momentum mapping. Suppose that $0 \in \mathfrak{g}^* \oplus \mathfrak{g}^* \oplus \mathfrak{g}^*$ is a clean value for $\mu$ and that the quotient $M_0 = G \backslash \mu^{-1}(0)$ has a smooth structure for which the projection $\pi_0 : \mu^{-1}(0) \to M_0$ is a smooth submersion. Then there is a unique hyperKähler structure $(\Omega_1^0, \Omega_2^0, \Omega_3^0)$ on $M_0$ with the property that $\pi_0^*(\Omega_a^0)$ is the pull back of $\Omega_a$ to $\mu^{-1}(0) \subset M$ for each $a = 1$, 2, or 3.

PROOF: Assume the hypotheses of the Theorem. Let $\tilde{\Omega}_a^0$ be the pullback of $\Omega_a$ to $\mu^{-1}(0) \subset M$. It is clear that each of the forms $\tilde{\Omega}_a^0$ is a closed, $G$-invariant 2-form on $\mu^{-1}(0)$.

I first want to show that each of these can be written as a pullback of a 2-form on $M_0$, i.e., that each is semi-basic for $\pi_0$. To do this, I need to characterize $T_m \mu^{-1}(0)$ in an appropriate fashion. Now, the assumption that $0$ be a clean value for $\mu$ implies $\mu^{-1}(0)$ is a smooth submanifold of $M$ and that for $m \in \mu^{-1}(0)$ any $v \in T_m M$ lies in $T_m \mu^{-1}(0)$ if and only if $\Omega_a \big( v, \lambda_*(x)(m) \big) = 0$ for all $x \in \mathfrak{g}$ and all three values of $a$. Thus,

$$T_m \mu^{-1}(0) = \{ v \in T_m M \, \big| \, \langle v, w\, i \rangle = \langle v, w\, j \rangle = \langle v, w\, k \rangle = 0, \text{ for all } w \in T_m G{\cdot}m \}.$$

Also, by $G$-equivariance, $G{\cdot}m \subset \mu^{-1}(0)$ and hence $T_m G{\cdot}m \subset T_m \mu^{-1}(0)$. It follows that $v \in T_m G{\cdot}m$ implies that $v$ is in the null space of each of the forms $\tilde{\Omega}_a^0$. Thus, each of the forms $\tilde{\Omega}_a^0$ is semi-basic for $\pi_0$, as we wished to show. This, combined with $G$-invariance, implies that there exist unique forms $\Omega_a^0$ on $M_0$ that satisfy $\pi_0^*(\Omega_a^0) = \tilde{\Omega}_a^0$. Since $\pi_0$ is a submersion, the three 2-forms $\Omega_a^0$ are closed.

To complete the proof, it suffices to show that the triple $(\Omega_1^0, \Omega_2^0, \Omega_3^0)$ actually defines an almost hyperKähler structure on $M_0$, for then we can apply Theorem 4.

We do this as follows: Use the associated metric $\langle, \rangle$ to define an orthogonal splitting

$$T_m \mu^{-1}(0) = T_m G{\cdot}m \oplus H_m.$$

By the hypotheses of the theorem, the fibers of $\pi_0$ are the $G$-orbits in $\mu^{-1}(0)$ and, for each $m \in \mu^{-1}(0)$, the kernel of the differential $\pi_0'(m)$ is $T_m G{\cdot}m$. Thus, $\pi_0'(m)$ induces an isomorphism from $H_m$ to $T_{\pi_0(m)} M_0$ and, under this isomorphism, the restriction of the form $\tilde{\Omega}_a^0$ to $H_m$ is identified with $\Omega_a^0$.

Thus, it suffices to show that the forms $(\tilde{\Omega}_1^0, \tilde{\Omega}_2^0, \tilde{\Omega}_3^0)$ define a hyperKähler structure when restricted to $H_m$. By Proposition 2, to do this, it would suffice to show that $H_m$ is stable under the actions of $R_i$, $R_j$, and $R_k$. However, by definition, $H_m$ is the subspace of $T_m M$ that is orthogonal to the $\mathbb{H}$-linear subspace $\big( T_m G{\cdot}m \big) \cdot \mathbb{H} \subset T_m M$. Since the orthogonal complement of an $\mathbb{H}$-linear subspace of $T_m M$ is also an $\mathbb{H}$-linear subspace, we are done. □

Note that the proof also shows that the dimension of the reduced space $M_0$ is equal to $\dim M - 4 \dim(G/G_m)$, since, at each point $m \in \mu^{-1}(0)$, the space $T_m G{\cdot}m$ is perpendicular to $R_i \big( T_m G{\cdot}m \big) \oplus R_j \big( T_m G{\cdot}m \big) \oplus R_k \big( T_m G{\cdot}m \big)$ and this latter direct sum is orthogonal.

Unfortunately, it frequently happens that $0$ is not a clean value of $\mu$, in which case, Theorem 5 cannot be applied to the action. Moreover, there does not appear to be any simple way to perform hyperKähler reduction at the general clean value of $\mu$ in $\mathfrak{g}^* \oplus \mathfrak{g}^* \oplus \mathfrak{g}^*$

(in marked contrast to the Kähler case). In fact, for the general clean value $\xi \in \mathfrak{g}^* \oplus \mathfrak{g}^* \oplus \mathfrak{g}^*$ of $\mu$, the quotient space $M_\xi = G_\xi \backslash \mu^{-1}(\xi)$ need not even have its dimension be divisible by 4. (See the Exercises for a cautionary example.)

However, if $[\mathfrak{g}, \mathfrak{g}]^\perp \subset \mathfrak{g}^*$ denotes the annihilator of $[\mathfrak{g}, \mathfrak{g}]$ in $\mathfrak{g}$, then the points $\xi \in [\mathfrak{g}, \mathfrak{g}]^\perp$ are the fixed points of the coadjoint action of $G$. It is then possible to perform hyperKähler reduction at any clean value $\xi = (\xi_1, \xi_2, \xi_3) \in [\mathfrak{g}, \mathfrak{g}]^\perp \oplus [\mathfrak{g}, \mathfrak{g}]^\perp \oplus [\mathfrak{g}, \mathfrak{g}]^\perp$ since, in this case, we again have $G_\xi = G$, and so the argument in the proof above that $H_m \subset T_m \mu^{-1}(\xi)$ is a quaternionic subspace for each $m \in T_m \mu^{-1}(\xi)$ is still valid. Of course, this is not really much of a generalization, since reduction at such a $\xi$ is simply reduction at 0 for the modified (but still $G$-equivariant) momentum mapping $\mu^\xi = \mu - \xi$.

**Example:** One of the simplest things to do is take $M = \mathbb{H}^n$ and let $G \subset \mathrm{Sp}(n)$ be a closed subgroup. It is not difficult to show (see the Exercises) that the standard hyperKähler structure on $\mathbb{H}^n$ has its three 2-forms given by

$$ i\,\Omega_1 + j\,\Omega_2 + k\,\Omega_3 = \frac{1}{2}\,{}^t d\bar{q} \wedge dq $$

where $q : \mathbb{H}^n \to \mathbb{H}^n$ is the identity, thought of as a $\mathbb{H}^n$-valued function on $\mathbb{H}^n$. Using this formula, it is easy to show that the standard left action of $\mathrm{Sp}(n)$ on $\mathbb{H}^n$ is Poisson, with momentum mapping $\mu : \mathbb{H}^n \to \mathfrak{sp}(n) \oplus \mathfrak{sp}(n) \oplus \mathfrak{sp}(n)$ given by the formula*

$$ \mu(q) = \frac{1}{2}\big(q\,i\,{}^t\bar{q},\; q\,j\,{}^t\bar{q},\; q\,k\,{}^t\bar{q}\big). $$

Note that 0 is *not* a clean value of $\mu$ with respect to the full action of $\mathrm{Sp}(n)$. However, the situation can be very different for a closed subgroup $G \subset \mathrm{Sp}(n)$: Let $\pi_{\mathfrak{g}} : \mathfrak{sp}(n) \to \mathfrak{g}$ be the orthogonal projection relative to the $Ad$-invariant inner product on $\mathfrak{sp}(n)$. The momentum mapping for the action of $G$ on $\mathbb{H}^n$ is then given by

$$ \mu_G(q) = \frac{1}{2}\big(\pi_{\mathfrak{g}}(q\,i\,{}^t\bar{q}),\; \pi_{\mathfrak{g}}(q\,j\,{}^t\bar{q}),\; \pi_{\mathfrak{g}}(q\,k\,{}^t\bar{q})\big). $$

Since $G$ is compact, there is an orthogonal direct sum $\mathfrak{g} = \mathfrak{z} \oplus [\mathfrak{g}, \mathfrak{g}]$, where $\mathfrak{z}$ is the tangent algebra to the center of $G$. Thus, there will be a hyperKähler reduction for each clean value of $\mu_G$ that lies in $\mathfrak{z} \oplus \mathfrak{z} \oplus \mathfrak{z}$.

Let us now consider a very simple example: Let $S^1 \subset \mathrm{Sp}(n)$ act diagonally on $\mathbb{H}^n$ by the action

$$ e^{i\theta} \cdot \begin{pmatrix} q^1 \\ \vdots \\ q^n \end{pmatrix} = \begin{pmatrix} e^{i\theta}\,q^1 \\ \vdots \\ e^{i\theta}\,q^n \end{pmatrix}. $$

---

   \* Here, I am identifying $\mathfrak{sp}(n)$ with $\mathfrak{sp}(n)^*$ via the positive definite, $Ad$-invariant symmetric bilinear pairing defined by $\langle a, b \rangle = -\frac{1}{2}\,\mathrm{tr}(ab + ba)$. (Because of the noncommutativity of $\mathbb{H}$, we do not have $\mathrm{tr}(ab) = \mathrm{tr}(ba)$ for all $a, b \in \mathfrak{sp}(n)$.)

Then it is not difficult to see that the momentum mappping can be identified with the map

$$\mu(q) = {}^t\bar{q}\,i\,q.$$

The reduced space $M_p$ for any $p \neq 0$ is easily seen to be complex analytically equivalent to $T^*\mathbb{CP}^{n-1}$, and the induced hyperKähler structure is the one found by Calabi. In particular, for $n = 2$, we recover the Eguchi-Hansen metric.

In the Exercises, there are other examples for you to try.

The method of hyperKähler reduction has a wide variety of applications. Many of the interesting moduli spaces for Yang-Mills theory turn out to have hyperKähler structures because of this reduction procedure. For example, as Atiyah and Hitchin [AH] show, the space of magnetic monopoles of "charge" $k$ on $\mathbb{R}^3$ turns out to have a natural hyperKähler structure that is derived by methods extremely similar to the example presented earlier of a Kähler structure on the moduli space of flat connections over a Riemann surface.

Peter Kronheimer [Kr] has used the method of hyperKähler reduction to construct, for each quotient manifold $\Sigma$ of $S^3$, an asymptotically locally Euclidean (ALE) Ricci-flat self-dual Einstein metric on a 4-manifold $M_\Sigma$ whose boundary at infinity is $\Sigma$. He then went on to prove that all such metrics on 4-manifolds arise in this way.

Finally, it should also be mentioned that the case of metrics on manifolds $M^{4n}$ with holonomy $\mathrm{Sp}(n)\cdot\mathrm{Sp}(1)$ can also be treated by the method of reduction. I don't have time to go into this here, but the reader can find a complete account in [GL].

# Exercise Set 8:

# Recent Applications of Reduction

**1**. Show that the following two definitions of compatibility between an almost complex structure $J$ and a metric $g$ on $M^{2n}$ are equivalent

 (i) $(g, J)$ are compatible if $g(v) = g(Jv)$ for all $v \in TM$.

 (ii) $(g, J)$ are compatible if $\Omega(v, w) = \langle Jv, w \rangle$ defines a (skew-symmetric) 2-form on $M$.

**2**. A NON-INTEGRABLE ALMOST COMPLEX STRUCTURE. Let $J$ be an almost complex structure on $M$. Let $\mathcal{A}^{1,0} \subset \mathbb{C} \otimes \mathcal{A}^1(M)$ denote the space of $\mathbb{C}$-valued 1-forms on $M$ that satisfy $\alpha(Jv) = \imath \, \alpha(v)$ for all $v \in TM$.

 (i) Show that if we define $\mathcal{A}^{0,1}(M) \subset \mathbb{C} \otimes \mathcal{A}^1(M)$ to be the space of $\mathbb{C}$-valued 1-forms on $M$ that satisfy $\alpha(Jv) = -\imath \, \alpha(v)$ for all $v \in TM$, then $\mathcal{A}^{0,1}(M) = \overline{\mathcal{A}^{1,0}(M)}$ and that $\mathcal{A}^{1,0}(M) \cap \mathcal{A}^{0,1}(M) = \{0\}$.

 (ii) Show that if $J$ is an integrable almost complex structure, then, for any $\alpha \in \mathcal{A}^{1,0}(M)$, the 2-form $d\alpha$ is (at least locally) in the ideal generated by $\mathcal{A}^{1,0}(M)$. (Hint: Show that, if $z : U \to \mathbb{C}^n$ is a holomorphic coordinate chart, then, on $U$, the space $\mathcal{A}^{1,0}(U)$ is spanned by the forms $dz^1, \ldots, dz^n$. Now consider the exterior derivative of any linear combination of the $dz^i$.)

 It is a celebrated result of Newlander and Nirenberg that this condition is sufficient for $J$ to be integrable.

 (iii) Show that there is an almost complex structure on $\mathbb{C}^2$ for which $\mathcal{A}^{1,0}(\mathbb{C}^2)$ is spanned by the 1-forms
$$\omega^1 = dz^1 - \bar{z}^1 \, d\bar{z}^2$$
$$\omega^2 = dz^2$$

 and that this almost complex structure is not integrable.

**3**. Let U(2) act diagonally on $\mathbb{C}^{2n}$, thought of as $n > 2$ copies of $\mathbb{C}^2$ (the action on each factor is the standard one and the Kähler structure on each factor is the standard one). Regarding $\mathbb{C}^{2n}$ as the space of 2-by-$n$ matrices with complex entries, show that the momentum mapping in this case is (up to a constant factor) given by $\mu(z) = \imath \, z \, {}^t\bar{z}$. (As usual, identify $\mathfrak{u}(2)$ with $\mathfrak{u}(2)^*$ by using the nondegenerate bilinear pairing $\langle x, y \rangle = -\text{tr}(xy)$.) Note that $\xi_0 = \imath \, I_2 \in \mathfrak{u}(2)$ is a fixed point of the coadjoint action and describe Kähler reduction at $\xi_0$. On the other hand, if $\xi \in \mathfrak{u}(2)$ has eigenvalues $\imath\lambda_1$ and $\imath\lambda_2$ where $\lambda_1 > \lambda_2 > 0$, show that, even though $\xi$ is a clean value of $\mu$ and $G_\xi \subset \text{U}(2)$ acts freely on $\mu^{-1}(\xi)$, the metric $g_\xi$ defined on the quotient $M_\xi$ so that $\pi_\xi : \mu^{-1}(\xi) \to M_\xi$ is a Riemannian submersion is not compatible with $\Omega_\xi$.

**4.** Examine the coadjoint orbits of $G = \mathrm{SL}(2, \mathbb{R})$ and show that $G \cdot \xi$ carries a $G$-invariant Riemannian metric if and only if $G_\xi$ is compact. Classify the coadjoint orbits of $G = \mathrm{SL}(3, \mathbb{R})$ and show that none of them carry a $G$-invariant Riemannian metric. In particular, none of them carry a $G$-invariant Kähler metric.

**5.** The point of this problem is to examine the coadjoint orbits of the compact group $\mathrm{U}(n)$ and to construct, on each one, the Kähler metric compatible with the canonical symplectic structure. As usual, we identitfy $\mathfrak{u}(n)$ with $\mathfrak{u}(n)^*$ via the $Ad$-invariant positive definite symmetric bilinear form $\langle x, y \rangle = -\mathrm{tr}(xy)$. Thus, $\xi \in \mathfrak{u}(n)$ is to be regarded as the linear functional $x \mapsto \langle \xi, x \rangle$. This allows us to identify the adjoint and coadjoint representations. Of course, since $\mathrm{U}(n)$ is a matrix group $\mathrm{Ad}(a)(x) = axa^{-1} = ax\,^t\bar{a}$ for $a \in \mathrm{U}(n)$ and $x \in \mathfrak{u}(n)$. Recall that every skew-Hermitian matrix $\xi$ can be diagonalized by a unitary transformation. Consequently, each (co)adjoint orbit is the orbit of a unique matrix of the form

$$
\xi = \begin{pmatrix} i\xi_1 & 0 & \cdots & 0 \\ 0 & i\xi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & i\xi_n \end{pmatrix}, \qquad \xi_1 \geq \xi_2 \geq \cdots \geq \xi_n .
$$

Fix $\xi$ and let $n_1, \ldots, n_d \geq 1$ be the multiplicities of the eigenvalues (i.e., $n_1 + \cdots + n_d = n$ and $\xi_j = \xi_k$ if and only if, for some $r$, we have $n_1 + \cdots + n_r \leq j \leq k < n_1 + \cdots + n_{r+1}$). Then $\mathrm{U}(n)_\xi = \mathrm{U}(n_1) \times \mathrm{U}(n_2) \times \cdots \times \mathrm{U}(n_d)$ (i.e., the obvious block diagonal subgroup).

Let $\omega = g^{-1}\,dg = (\omega_{j\bar{k}})$ be the canonical left-invariant form on $\mathrm{U}(n)$ and let $\pi_\xi : \mathrm{U}(n) \to \mathrm{U}(n)/\mathrm{U}(n)_\xi = \mathrm{U}(n) \cdot \xi$ be the canonical projection. Show that the formulae

$$
\pi_\xi^*(\Omega_\xi) = \frac{i}{2} \sum_{k>j} 2(\xi_j - \xi_k)\, \omega_{k\bar{j}} \wedge \overline{\omega_{k\bar{j}}} \qquad \text{and} \qquad \pi_\xi^*(h_\xi) = \sum_{k>j} 2(\xi_j - \xi_k)\, \omega_{k\bar{j}} \circ \overline{\omega_{k\bar{j}}}
$$

define the symplectic form $\Omega_\xi$ and a compatible Kähler metric $h_\xi$ on the orbit $U(n) \cdot \xi$. (In fact, this $h_\xi$ is the *only* $\mathrm{U}(n)$-invariant, $\Omega_\xi$-compatible metric on $U(n) \cdot \xi$.)

REMARK: If you know about roots and weights, it is not hard to generalize this construction so that it works for the coadjoint orbits of any compact Lie group. The same uniqueness result is true as well: If $G$ is compact and $\xi \in \mathfrak{g}^*$ is any element, there is a unique $G$-invariant Kähler metric $h_\xi$ on $G \cdot \xi$ that is compatible with $\Omega_\xi$.

**6.** Let $M = \mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2} \setminus \{(0,0)\}$. Let $G = S^1$ act on $M$ by the action

$$
e^{i\theta} \cdot (z_1, z_2) = \left( e^{i d_1 \theta} z_1, e^{i d_2 \theta} z_2 \right)
$$

where $d_1$ and $d_2$ are relatively prime integers. Let $M$ have the standard flat Kähler structure. Compute the momentum mapping $\mu$ and the Kähler structures on the reduced spaces. How do the relative signs of $d_1$ and $d_2$ affect the answer? What interpretation can you give to these spaces?

**7**. Go back to the the example of the "Kähler structure" on the space $\mathfrak{A}(P)$ of connections on a principal right $G$-bundle $P$ over a connected compact Riemann surface $\Sigma$. Assume that $G = S^1$ and identify $\mathfrak{g}$ with $\mathbb{R}$ in the natural way. Thus, $F_A$ is a well-defined 2-form on $\Sigma$ and the cohomology class $[F_A] \in H^2_{dR}(\Sigma, \mathbb{R})$ is independent of the choice of $A$. Assume that $[F_A] \neq 0$. Then, in this case, $\boldsymbol{\mu}^{-1}(0)$ is empty so the construction we made in the example in the Lecture is vacuous. Here is how we can still get some information.

Fix any non-vanishing 2-form $\Psi$ on $\Sigma$ so that $[\Psi] = [F_A]$. Show that even though $\boldsymbol{\mu}$ has no regular values, $\Psi$ is a non-trivial clean value of $\boldsymbol{\mu}$. Show also that, for any $A \in \boldsymbol{\mu}^{-1}(\Psi)$, the stabilizer $\mathsf{G}_A \subset \mathsf{Aut}(P)$ is a discrete (and hence finite) subgroup of $S^1$. Describe, as fully as you can, the reduced space $\mathfrak{M}_\Psi$ and its Kähler structure. (Here, you will need to keep in mind that $\Psi$ is a fixed point of the coadjoint action of $\mathsf{Aut}(P)$, so that reduction away from 0 makes sense.)

**8**. Verify that $\mathrm{Sp}(n)$ is a connected Lie group of dimension $2n^2 + n$. (Hint: You will probably want to study the function $f(A) = {}^t\bar{A}A = I_n$.) Show that $\mathrm{Sp}(n)$ acts transitively on the unit sphere $S^{4n-1} \subset \mathbb{H}^n$ defined by the relation $H(x, x) = 1$. (Hint: First show that, by acting by diagonal matrices in $\mathrm{Sp}(n)$, you can move any element of $\mathbb{H}^n$ into the subspace $\mathbb{R}^n$. Then note that $\mathrm{Sp}(n)$ contains $\mathrm{SO}(n)$.) By analysing the stabilizer subgroup in $\mathrm{Sp}(n)$ of an element of $S^{4n-1}$, show that there is a fibration

$$
\begin{array}{ccc}
\mathrm{Sp}(n{-}1) & \to & \mathrm{Sp}(n) \\
& & \downarrow \\
& & S^{4n-1}
\end{array}
$$

and use this to conclude by induction that $\mathrm{Sp}(n)$ is connected and simply connected for all $n$.

**9**. Prove Proposition 2. (Hint: First show how the maps $R_i$, $R_j$, and $R_k$ define the structure of a right $\mathbb{H}$-module on $V$. Then show that $V$ has a basis $b_1, \ldots, b_n$ over $\mathbb{H}$ and use this to construct an $\mathbb{H}$-linear isomorphism of $V$ with $\mathbb{H}^n$. If you pick the basis $b_a$ carefully, you will be done at this point. Warning: You must use the positive definiteness of $\langle,\rangle$!)

**10**. Show that if $(\Omega_1, \Omega_2, \Omega_3)$ is a hyperKähler structure on a real vector space $V$, with associated defined maps $R_i$, $R_j$, and $R_k$ and metric $\langle,\rangle$, then $(\Omega_1, R_i)$ is a complex Hermitian structure on $V$ with associated metric $\langle,\rangle$. Moreover, the $\mathbb{C}$-valued 2-form $\Upsilon = \Omega_2 - \imath\,\Omega_3$ is $\mathbb{C}$-linear, i.e., $\Upsilon(R_i v, w) = \imath\,\Upsilon(v, w)$ for all $v, w \in V$. Show that $\Upsilon$ is non-degenerate on $V$ and hence that $\Upsilon$ defines a (complex) symplectic structure on $V$ (considered as a complex vector space).

**11**. Let $\mathrm{Sp}(1) \simeq \mathrm{SU}(2)$ act on $\mathbb{H}^n$ diagonally (i.e., as componentwise left multiplication $n$ copies of $\mathbb{H}$). Compute the momentum mapping $\mu : \mathbb{H}^n \to \mathfrak{sp}(1) \oplus \mathfrak{sp}(1) \oplus \mathfrak{sp}(1)$ and show that for all $n \geq 4$, the map $\mu$ is surjective and that, for generic $\xi \in \mathfrak{sp}(1) \oplus \mathfrak{sp}(1) \oplus \mathfrak{sp}(1)$, we have $G_\xi = \{\pm 1\}$. In particular, for nearly all nonzero regular values of $\mu$, the quotient space $M_\xi = G_\xi \backslash \mu^{-1}(\xi)$ has dimension $4n-9$. Consequently, this quotient space is not even Kähler. (Hint: You may find it useful to recall that $\mathfrak{sp}(1) = \mathrm{Im}\,\mathbb{H} \simeq \mathbb{R}^3$ and that the (co)adjoint action is identifiable with $\mathrm{SO}(3)$ acting by rotations on $\mathbb{R}^3$. In fact, you might want to note that it is possible to identify $\mathfrak{sp}(1) \oplus \mathfrak{sp}(1) \oplus \mathfrak{sp}(1)$ with $\mathbb{R}^9 \simeq M_{3,3}(\mathbb{R})$ in such a way that

$$\mu(pq^1\bar{u}, \cdots, pq^n\bar{u}) = R(p)\mu(q^1, \cdots, q^n)R(u)^{-1}$$

where $R : \mathrm{Sp}(1) \to \mathrm{SO}(3)$ is a covering homomorphism. Once this has been proved, you can use facts about matrix multiplication to simplify your computations.) Now, again, assuming that $n \geq 4$, compute $\mu^{-1}(0)$, show that, once the origin in $\mathbb{H}^n$ is removed, 0 is a regular value of $\mu$ and that $\mathrm{Sp}(1)$ acts freely on $\mu^{-1}(0)$. Can you describe the quotient space? (You may find it helpful to note that $\mathrm{SO}(n) \subset \mathrm{Sp}(n)$ is the commuting subgroup of $\mathrm{Sp}(1)$ embedded diagonally into $\mathrm{Sp}(n)$. What good is knowing this?)

**12**. Apply the hyperKähler reduction procedure to $\mathbb{H}^2$ with $G = S^1$ acting by the rule

$$e^{i\theta} \cdot \begin{pmatrix} q^1 \\ q^2 \end{pmatrix} = \begin{pmatrix} e^{im\theta} q^1 \\ e^{in\theta} q^2 \end{pmatrix},$$

where $m$ and $n$ are relatively prime integers. Determine which values of $\mu$ are clean and describe the resulting complex surfaces and their hyperKähler structures.

E.8.4

# Lecture 9:

# The Gromov School of Symplectic Geometry

In this lecture, I want to describe some of the remarkable new information we have about symplectic manifolds owing to the influence of the ideas of Mikhail Gromov. The basic reference for much of this material is Gromov's remarkable book *Partial Differential Relations*.

The fundamental idea of studying complex structures "tamed by" a given symplectic structure was developed by Gromov in a remarkable paper *Pseudo-holomorphic Curves on Almost Complex Manifolds* and has proved extraordinarily fruitful. In the latter part of this lecture, I will try to introduce the reader to this theory.

### Soft Techniques in Symplectic Manifolds

**Symplectic Immersions and Embeddings.** Before beginning on the topic of symplectic immersions, let me recall how the theory of immersions in the ordinary sense goes.

Recall that the Whitney Immersion Theorem (in the weak form) asserts that any smooth $n$-manifold $M$ has an immersion into $\mathbb{R}^{2n}$. This result is proved by first immersing $M$ into some $\mathbb{R}^N$ for $N \gg 0$ and then using Sard's Theorem to show that if $N > 2n$, one can find a vector $u \in \mathbb{R}^N$ so that $u$ is not tangent to $f(M)$ at any point. Then the projection of $f(M)$ onto a hyperplane orthogonal to $u$ is still an immersion, but now into $\mathbb{R}^{N-1}$.

This result is not the best possible. Whitney himself showed that one could always immerse $M^n$ into $\mathbb{R}^{2n-1}$, although "general position" arguments are not sufficient to do this. This raises the question of determining what the best possible immersion or embedding dimension is.

One topological obstruction to immersing $M^n$ into $\mathbb{R}^{n+k}$ can be described as follows: If $f \colon M \to \mathbb{R}^{n+k}$ is an immersion, then the trivial bundle $f^*(T\mathbb{R}^{n+k}) = M \times \mathbb{R}^{n+k}$ can be split into a direct sum $f^*(T\mathbb{R}^{n+k}) = TM \oplus \nu^f$ where $\nu^f$ is the normal bundle of the immersion $f$. Thus, if there is no bundle $\nu$ of rank $k$ over $M$ so that $TM \oplus \nu$ is trivial, then there can be no immersion of $M$ into $\mathbb{R}^{n+k}$.

The remarkable fact is that this topological necessary condition is almost sufficient. In fact, we have the following result of Hirsch and Smale for the general immersion problem.

**Theorem 1:** *Let $M$ and $N$ be connected smooth manifolds and suppose either that $M$ is non-compact or else that $\dim(M) < \dim(N)$. Let $f \colon M \to N$ be a continuous map, and suppose that there is a vector bundle $\nu$ over $M$ so that $f^*(TN) = TM \oplus \nu$. Then $f$ is homotopic to an immersion of $M$ into $N$*

Theorem 1 can be interpreted as an example of what Gromov calls the *h-principle*, which I now want to describe.

THE $h$-PRINCIPLE. Let $\pi\colon V \to X$ be a surjective submersion. A *section* of $\pi$ is, by definition, a map $\sigma\colon X \to V$ which satisfies $\pi \circ \sigma = id_X$. Let $J^k(X,V)$ denote the space of $k$-jets of sections of $V$, and let $\pi^k\colon J^k(X,V) \to X$ denote the basepoint or "source" projection. Given any section $s$ of $\pi$, there is an associated section $j^k(s)$ of $\pi^k$ which is defined by letting $j^k(s)(x)$ be the $k$-jet of $s$ at $x \in X$. A section $\sigma$ of $\pi^k$ is said to be *holonomic* if $\sigma = j^k(s)$ for some section $s$ of $\pi$.

A *partial differential relation* of order $k$ for $\pi$ is a subset $R \subset J^k(X,V)$. A section $s$ of $\pi$ is said to *satisfy* $R$ if $j^k(s)(X) \subset R$. We can now make the following definition:

**Definition 1:** A partial differential relation $R \subset J^k(X,V)$ satisfies the *h-principle* if, for every section $\sigma$ of $\pi^k$ which satisfies $\sigma(X) \subset R$, there is a one-parameter family of sections $\sigma_t$ $(0 \le t \le 1)$ of $\pi^k$ which satisfy the conditions that $\sigma_t(X) \subset R$ for all $t$, that $\sigma_0 = \sigma$, and that $\sigma_1$ is holonomic.

Very roughly speaking, a partial differential relation satisfies the $h$-principle if, whenever the "topological" conditions for a solution to exist are satisfied, then a solution exists.

For example, if $X = M$ and $V = M \times N$, where $\dim(N) \ge \dim(M)$, then there is an (open) subset $R = \mathrm{Imm}(M,N) \subset J^1(M, M \times N)$ which consists of the 1-jets of graphs of (local) immersions of $M$ into $N$. What the Hirsch-Smale immersion theory says is that $\mathrm{Imm}(M,N)$ satisfies the $h$-principle if either $\dim M = \dim N$ and $M$ has no compact component or else $\dim M < \dim N$.

Of course, the $h$-principle does not hold for every relation $R$. The real question is how to determine when the $h$-principle holds for a given $R$. Gromov has developed several extremely general methods for proving that the $h$-principle holds for various partial differential relations $R$ which arise in geometry. These methods include his theory of topological sheaves and techniques like his method of convex integration. They generally work in situations where the local solutions of a given partial differential relation $R$ are easy to come by and it is mainly a question of "patching together" local solutions which are fairly "flexible".

Gromov calls this collection of techniques "soft" to distinguish them from the "hard" techniques, such as elliptic theory, which come from analysis and deal with situations where the local solutions are somewhat "rigid".

Here is a sample of some of the results which Gromov obtains by these methods:

**Theorem 2:** *Let $X^{2n}$ be a smooth manifold and let $V \subset \Lambda^2\big(T^*(M)\big)$ denote the open subbundle consisting of non-degenerate 2-forms $\omega \in T_x X$. Let $Z^1(X,V) \subset J^1(X,V)$ denote the space of 1-jets of closed non-degenerate 2-forms on $X$. Then, if $X$ has no compact component, $Z^1(X,V)$ satisfies the h-principle.*

In particular, Theorem 2 implies that a non-compact, connected $X$ has a symplectic structure if and only if it has an almost symplectic structure.

Note that this result is definitely *not* true for compact manifolds. We have already seen several examples, e.g., $S^1 \times S^3$, which have almost symplectic structures but no symplectic structures because they do not satisfy the cohomology ring obstruction. Gromov has asked the following:

**Question :**  *If $X^{2n}$ is compact and connected and satisfies the condition that there exists an element $u \in H^2_{dR}(X, \mathbb{R})$ which satisfies $u^n \neq 0$, does $Z^1(X, V)$ satisfy the h-principle?*

The next result I want to describe is Gromov's theorem on symplectic immersions. This theorem is an example of a sort of "restricted $h$-principle" in that it is only required to apply to sections $\sigma$ which satisfy specified cohomological conditions.

First, let me make a few definitions: Let $(X, \Xi)$ and $(Y, \Psi)$ be two connected symplectic manifolds. Let $\mathcal{S}(X, Y) \subset J^1(X, X \times Y)$ denote the space of 1-jets of graphs of (local) symplectic maps $f: X \to Y$. i.e., (local) maps $f: X \to Y$ which satisfy $f^*(\Psi) = \Xi$. Let $\tau: \mathcal{S}(X, Y) \to Y$ be the obvious "target projection".

**Theorem 3:**  *If either $X$ is non-compact, or $\dim(X) < \dim(Y)$, then any section $\sigma$ of $\mathcal{S}(X, Y)$ for which the induced map $s = \tau \circ \sigma: X \to Y$ satisfies the cohomological condition $s^*([\Psi]) = [\Xi]$ is homotopic to a holonomic section of $\mathcal{S}(X, Y)$.*

This result can be also stated as follows: Suppose that either $X$ is non-compact or else that $\dim(X) < \dim(Y)$. Let $\phi: X \to Y$ be a smooth map which satisfies the cohomological condition $\phi^*([\Psi]) = [\Xi]$. Suppose that there exists a bundle map $f: TX \to \phi^*(TY)$ which is symplectic in the obvious sense. Then $\phi$ is homotopic to a symplectic immersion.

As an application of Theorem 3, we can now prove the following result of Narasimham and Ramanan.

**Corollary :**  *Any compact symplectic manifold $(M, \Omega)$ for which the cohomology class $[\Omega]$ is integral admits a symplectic immersion into $(\mathbb{CP}^N, \Omega_N)$ for some $N \gg n$.*

PROOF: Since the cohomology class $[\Omega]$ is integral, there exists a smooth map $\phi: M \to \mathbb{CP}^N$ for some $N$ sufficiently large so that $\phi^*([\Omega_N]) = [\Omega]$. Then, choosing $N \gg n$, we can arrange that there also exists a symplectic bundle map $f: TM \to f^*(T\mathbb{CP}^N)$ (see the Exercises). Now apply Theorem 3. $\qquad\square$

As a final example along these lines, let me state Gromov's embedding result. Here, the reader should be thinking of the difference between the Whitney Immersion Theorems and the Whitney Embedding Theorems: One needs slightly more room to embed than to immerse.

**Theorem 4:**  *Suppose that $(X, \Xi)$ and $(Y, \Psi)$ are connected symplectic manifolds and that either $X$ is non-compact and $\dim(X) < \dim(Y)$ or else that $\dim(X) < \dim(Y) - 2$. Suppose that there exists a smooth embedding $\phi: X \to Y$ and that the induced map on bundles $\phi': TX \to \phi^*(TY)$ is homotopic through a 1-parameter family of injective bundle maps $\varphi_t: TX \to \phi^*(TY)$ (with $\varphi_0 = \phi'$) to a symplectic bundle map $\varphi_1: TX \to \phi^*(TY)$. Then $\phi$ is isotopic to a symplectic embedding $\varphi: X \to Y$.*

This result is actually the best possible, for, as Gromov has shown using "hard" techniques (see below), there are counterexamples if one leaves out the dimensional restrictions. Note by the way that, because Theorem 4 deals with embeddings rather than immersions, it not straightforward to place it in the framework of the $h$-principle.

**Blowing Up in the Symplectic Category.** We have already seen in Lecture 6 that certain operations on smooth manifolds cannot be carried out in the symplectic category. For example, one cannot form connected sums in the symplectic category.

However, certain of the operations from the geometry of complex manifolds *can* be carried out. Gromov has shown how to define the operation of "blowing up" in the symplectic category.

Recall how one "blows up" the origin in $\mathbb{C}^n$. To avoid triviality, let me assume that $n > 1$. Consider the subvariety

$$ X = \{(v, [w]) \in \mathbb{C}^n \times \mathbb{CP}^{n-1} \,|\, v \in [w]\} \subset \mathbb{C}^n \times \mathbb{CP}^{n-1}. $$

It is easy to see that $X$ is a smooth embedded submanifold of the product and that the projection $\pi\colon X \to \mathbb{C}^n$ is a biholomorphism away from the "exceptional point" $0 \in \mathbb{C}^n$. Moreover, if $\Omega_0$ and $\Phi$ are the standard Kähler 2-forms on $\mathbb{C}^n$ and $\mathbb{CP}^{n-1}$ respectively, then, for each $\epsilon > 0$, the 2-form $\Omega_\epsilon = \Omega_0 + \epsilon\Phi$ is a Kähler 2-form on $X$.

Now, Gromov realized that this can be generalized to a "blow up" construction for any point $p$ on any symplectic manifold $(M^{2n}, \Omega)$. Here is how this goes:

First, choose a neighborhood $U$ of $p$ on which there exists a local chart $z\colon U \to \mathbb{C}^n$ which is symplectic, i.e., satisfies $z^*(\Omega_0) = \Omega$, and satisfies $z(p) = 0$. Suppose that the ball $B_{2\delta}(0)$ in $\mathbb{C}^n$ of radius $2\delta$ centered on $0$ lies inside $z(U)$. Since $\pi\colon \pi^{-1}\big(B_{2\delta}^*(0)\big) \to B_{2\delta}^*(0)$ is a diffeomorphism, there exists a closed 2-form $\tilde{\Phi}$ on $B_{2\delta}^*(0)$ so that $\pi^*(\tilde{\Phi}) = \Phi$. Since $H_{dR}^2\big(B_{2\delta}^*(0)\big) = 0$, there exists a 1-form $\varphi$ on $B_{2\delta}^*(0)$ so that $d\varphi = \Phi$.

Now consider the family of symplectic forms $\Omega_0 + \epsilon\,d\varphi$ on $B_{2\delta}^*(0)$. By using a homotopy argument exactly like the one used to Prove Theorem 1 in Lecture 6, it easily follows that for all $t > 0$ sufficiently small, there exists an open annulus $A(\delta - \varepsilon, \delta + \varepsilon)$ and a one-parameter family of diffeomorphisms $\phi_t\colon A(\delta - \varepsilon, \delta + \varepsilon) \to B_{2\delta}^*(0)$ so that

$$ \phi_t^*(\Omega_0) = \Omega_0 + t\,d\varphi. $$

It follows that we can set

$$ \hat{M} = \pi^{-1}\big(B_{\delta+\varepsilon}^*(0)\big) \cup_{\psi_t} M \setminus z^{-1}\big(\phi_t\big(B_{\delta-\varepsilon}^*(0)\big)\big) $$

where

$$ \psi_t\colon \pi^{-1}\big(A(\delta - \varepsilon, \delta + \varepsilon)\big) \to M \setminus z^{-1}\big(\phi_t\big(A(\delta - \varepsilon, \delta + \varepsilon)\big)\big) $$

is given by $\psi_t = z^{-1} \circ \phi_t \circ \pi$. Since $\psi_t$ identifies the symplectic structure $\Omega_t$ with $\Omega_0$ on the annulus "overlap", it follows that $\hat{M}$ is symplectic. This is Gromov's symplectic blow up procedure. Note that it can be effected in such a way that the symplectic structure on $M \setminus \{p\}$ is not disturbed outside of an arbitrarily small ball around $p$. Note also that there is a parameter involved, and that the symplectic structure is certainly not unique.

This is only describes a simple case. However, Gromov has shown how any compact symplectic submanifold $S^{2k}$ of $M^{2n}$ can be blown up to become a symplectic "hypersurface" $\hat{S}$ in a new symplectic manifold $\hat{M}$ which has the property that $M \setminus S$ is diffeomorphic to $\hat{M} \setminus \hat{S}$.

The basic idea is the same as what we have already done: First, one mimics the topological operations which would be performed if one were blowing up a complex sub-manifold of a complex manifold. Thus, the submanifold $S$ gets "replaced" by the complex projectivization $\hat{S} = \mathbb{P}N_S$ of a complex normal bundle. Second, one shows how to define a symplectic structure on the resulting smooth manifold which can be made to agree with the old structure outside of an arbitrarily small neighborhood of the blow up.

The details in the general case are somewhat more complicated than the case of blowing up a single point, and Dusa McDuff ([McDuff 1984]) has written out a careful construction. She has also used the method of blow ups to produce an example of a simply connected compact symplectic manifold which has no Kähler structure.

## Hard Techniques in Symplectic Manifolds.

**(Pseudo-) holomorphic curves.** I begin with a fundamental definition.

**Definition 2:** Let $M^{2n}$ be a smooth manifold and let $J\colon TM \to TM$ be an almost complex structure on $M$. For any Riemann surface $\Sigma$, we say that a map $f\colon \Sigma \to M$ is *J-holomorphic* if $f'(\imath v) = Jf'(v)$ for all $v \in T\Sigma$.

▶  Often, when $J$ is clear from context, I will simply say "$f$ is holomorphic". Several authors use the terminology "almost holomorphic" or "pseudo-holomorphic" for this concept, reserving the word "holomorphic" for use only when the almost complex structure on $M$ is integrable to an actual complex structure. This distinction does not seem to be particularly useful, so I will not maintain it.

It is instructive to see what this looks like in local coordinates. Let $z = x + \imath y$ be a local holomorphic coordinate on $\Sigma$ and let $w\colon U \to \mathbb{R}^{2n}$ be a local coordinate on $M$. Then there exists a matrix $\mathsf{J}(w)$ of functions on $w(U) \subset \mathbb{R}^{2n}$ which satisfies $w'(Jv) = \mathsf{J}\big(w(p)\big)w'(v)$ for all $v \in T_pU$. This matrix of functions satisfies the relation $\mathsf{J}^2 = -I_{2n}$. Now, if $f\colon \Sigma \to M$ is holomorphic and carries the domain of the $z$-coordinate into $U$, then $F = w \circ f$ is easily seen to satisfy the first order system of partial differential equations

$$\frac{\partial F}{\partial y} = \mathsf{J}(F)\frac{\partial F}{\partial x}.$$

Since $\mathsf{J}^2 = -I_{2n}$, it follows that this is a first-order, elliptic, determined system of partial differential equations for $F$. In fact, the "principal symbol" of these equations is the same as that for the Cauchy-Riemann equations. Assuming that $\mathsf{J}$ is sufficiently regular ($C^\infty$ is sufficient and we will always have this) there are plenty of local solutions. What is at issue is the nature of the global solutions.

A *parametrized holomorphic curve* in $M$ is a holomorphic map $f\colon \Sigma \to M$. Sometimes we will want to consider *unparametrized holomorphic curves* in $M$, namely equivalence classes $[\Sigma, f]$ of holomorphic curves in $M$ where $(\Sigma_1, f_1)$ is equivalent to $(\Sigma_2, f_2)$ if there exists a holomorphic map $\phi\colon \Sigma_1 \to \Sigma_2$ satisfying $f_1 = f_2 \circ \phi$.

We are going to be particularly interested in the space of holomorphic curves in $M$. Here are some properties that hold in the case of holomorphic curves in actual complex manifolds and it would be nice to know if they also hold for holomorphic curves in almost complex manifolds.

LOCAL FINITE DIMENSIONALITY. If $\Sigma$ is a compact Riemann surface, and $f\colon \Sigma \to M$ is a holomorphic curve, it is reasonable to ask what the space of "nearby" holomorphic curves looks like. Because the equations which determine these mappings are elliptic and because $\Sigma$ is compact, it follows without too much difficulty that the space of nearby holomorphic curves is finite dimensional. (We do *not*, in general know that it is a smooth manifold!)

INTERSECTIONS. A pair of distinct complex curves in a complex surface always intersect at isolated points and with positive "multiplicity." This follows from complex analytic geometry. This result is extremely useful because it allows us to derive information about actual numbers of intersection points of holomorphic curves by applying topological intersection formulas. (Usually, these topological intersection formulas only count the number of signed intersections, but if the surfaces can only intersect positively, then the topological intersection numbers (counted with multiplicity) are the actual intersection numbers.)

KÄHLER AREA BOUNDS. If $M$ happens to be a Kähler manifold, with Kähler form $\Omega$, then the area of the image of a holomorphic curve $f\colon \Sigma \to M$ is given by the formula

$$\mathrm{Area}\big(f(\Sigma)\big) = \int_\Sigma f^*(\Omega).$$

In particular, since $\Omega$ is closed, the right hand side of this equation depends only on the homotopy class of $f$ as a map into $M$. Thus, if $(\Sigma_t, f_t)$ is a continuous one-parameter family of closed holomorphic curves in a Kähler manifold, then they all have the same area. This is a powerful constraint on how the images can behave, as we shall see.

Now the first two of these properties go through without change in the case of almost complex manifolds.

In the case of local finiteness, this is purely an elliptic theory result. Studying the linearization of the equations at a solution will even allow one to predict, using the Atiyah-Singer Index Theorem, an upper bound for the local dimension of the moduli space and, in some cases, will allow us to conclude that the moduli space near a given closed curve is actually a smooth manifold (see below).

As for pairs of complex curves in an almost complex surface, Gromov has shown that they do indeed only intersect in isolated points and with positive multiplicity (unless they have a common component, of course). Both Gromov and Dusa McDuff have used this fact to study the geometry of symplectic 4-manifolds.

The third property is only valid for Kähler manifolds, but it is highly desirable. The behaviour of holomorphic curves in compact Kähler manifolds is well understood in a large part because of this area bound. This motivated Gromov to investigate ways of generalizing this property.

**Symplectic Tamings.** Following Gromov, we make the following definition.

**Definition 3:** A symplectic form $\Omega$ on $M$ *tames* an almost complex structure $J$ if it is $J$-positive, i.e., if it satisfies $\Omega(v, Jv) > 0$ for all non-zero tangent vectors $v \in TM$.

The reader should be thinking of Kähler geometry. In that case, the symplectic form $\Omega$ and the complex structure $J$ satisfy $\Omega(v, Jv) = \langle v, v \rangle > 0$. Of course, this generalizes to the case of an arbitrary almost Kähler structure.

Now, if $M$ is compact and $\Omega$ tames $J$, then for any Riemannian metric $g$ on $M$ (not necessarily compatible with either $J$ or $\Omega$) there is a constant $C > 0$ so that

$$|v \wedge Jv| \leq C\,\Omega(v, Jv)$$

where $|v \wedge Jv|$ represents the area in $T_pM$ of the parallelogram spanned by $v$ and $Jv$ in $T_pM$ (see the Exercises). In particular, it follows that, for any holomorphic curve $f \colon \Sigma \to M$, we have the inequality

$$\mathrm{Area}\big(f(\Sigma)\big) \leq C \int_\Sigma f^*(\Omega).$$

Just as in the Kähler case, the integral on the right hand side depends only on the homotopy class of $f$. Thus, if an almost complex structure can be tamed, it follows that, in any metric on $M$, there is a uniform upper bound on the areas of the curves in any continuous family of compact holomorphic curves in $M$.

**Example:** Let $N_3^{\mathbb{C}}$ denote the complex Heisenberg group. Thus, $N_3^{\mathbb{C}}$ is the complex Lie group of matrices of the form

$$g = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

Let $\Gamma \subset N_3^{\mathbb{C}}$ be the subgroup all of whose entries belong to the ring of Gaussian integers $\mathbb{Z}[\imath]$.

Let $M = N_3^{\mathbb{C}}/\Gamma$. Then $M$ is a compact complex 3-manifold. I claim that the complex structure on $M$ cannot be tamed by any symplectic form.

To see this, consider the right-invariant 1-form

$$dg\,g^{-1} = \begin{pmatrix} 0 & \omega_1 & \omega_3 \\ 0 & 0 & \omega_2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Since they are right-invariant, it follows that the complex 1-forms $\omega_1, \omega_2, \omega_3$ are also well-defined on $M$. Define the metric $G$ on $M$ to be the quadratic form

$$G = \omega_1 \circ \overline{\omega_1} + \omega_2 \circ \overline{\omega_2} + \omega_3 \circ \overline{\omega_3}.$$

Now consider the holomorphic curve $Y\colon\mathbb{C}\to N_3^{\mathbb{C}}$ defined by

$$Y(y) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

Let $\psi\colon\mathbb{C}\to M$ be the composition. It is clear that $\psi$ is doubly periodic and hence defines an embedding of a complex torus into $M$. It is clear that the $G$-area of this torus is 1.

Now $N_3^{\mathbb{C}}$ acts holomorphically on $M$ on the left (not by $G$-isometries, of course). We can consider what happens to the area of the torus $\psi(\mathbb{C})$ under the action of this group. Specifically, for $x\in\mathbb{C}$, let $\psi_x$ denote $\psi$ acted on by left multiplication by the matrix

$$\begin{pmatrix} 1 & x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then, as the reader can easily check, we have

$$\psi_x^*(G) = \bigl(1 + |x|^2\bigr)\bigl|dy\bigr|^2.$$

Thus, the $G$-area of the torus $\psi_x(\mathbb{C})$ goes off to infinity as $x$ tends to infinity. Obviously, there can be no taming of the complex manifold $M$. (In particular, $M$ cannot carry a Kähler structure compatible with its complex structure.)

**Gromov's Compactness Theorem.** In this section, I want to discuss Gromov's approach to compactifying the connected components of the space of unparametrized holomorphic curves in $M$.

**Example:** Before looking at the general case, let us look at what happens in a very familiar case: The case of algebraic curves in $\mathbb{CP}^2$ with its standard Fubini-Study metric and symplectic form $\Omega$ (normalized so as to give the lines in $\mathbb{CP}^2$ an area of 1).

Since this is a Kähler metric, we know that the area of a connected one-parameter family of holomorphic curves $(\Sigma_t, f_t)$ in $\mathbb{CP}^2$ is constant and is equal to an integer $d = \int_{\Sigma_t} f_t^*(\Omega)$, called the *degree*. To make matters as simple as possible, let me consider the curves degree by degree.

$d = 0$. In this case, the "curves" are just the constant maps and (in the unparametrized case) clearly constitute a copy of $\mathbb{CP}^2$ itself. Note that this is already compact.

$d = 1$. In this case, the only possibility is that each $\Sigma_t$ is just $\mathbb{CP}^1$ and the holomorphic map $f_t$ must be just a biholomorphism onto a line in $\mathbb{CP}^2$. Of course, the space of lines in $\mathbb{CP}^2$ is compact, just being a copy of the dual $\mathbb{CP}^2$. Thus, the space $\mathfrak{M}_1$ of unparametrized holomorphic curves in $\mathbb{CP}^2$ is compact. Note, however, that the space $\mathfrak{H}_1$ of holomorphic maps $f\colon\mathbb{CP}^1\to\mathbb{CP}^2$ of degree 1 is not compact. In fact, the fibers of the natural map $\mathfrak{H}_1\to\mathfrak{M}_1$ are copies of $\mathrm{Aut}(\mathbb{CP}^1) = \mathrm{PSL}(2,\mathbb{C})$.

$d = 2$. This is the first really interesting case. Here again, degree 2 (connected, parametrized) curves in $\mathbb{CP}^2$ consist of rational curves, and the images $f\colon \mathbb{CP}^1 \to \mathbb{CP}^2$ are of two kinds: the smooth conics and the double covers of lines. However, not only is this space not compact, the corresponding space of unparametrized curves is not compact either, for it is fairly clear that one can approach a pair of intersecting lines as closely as one wishes.

In fact, the reader may want to contemplate the one-parameter family of hyperbolas $xy = \lambda^2$ as $\lambda \to 0$. If we choose the parametrization

$$f_\lambda(t) = [t, \lambda t^2, \lambda] = [1, x, y],$$

then the pullback $\Phi_\lambda = f_\lambda^*(\Omega)$ is an area form on $\mathbb{CP}^1$ whose total integral is 2, but (and the reader should check this), as $\lambda \to 0$, the form $\Phi_\lambda$ accumulates equally at the points $t = 0$ and $t = \infty$ and goes to zero everywhere else. (See the Exercises for a further discussion.)

Now, if we go ahead and add in the pairs of lines, then this "completed" moduli space is indeed compact. It is just the space of non-zero quadratic forms in three variables (irreducible or not) up to constant multiples. It is well known that this forms a $\mathbb{CP}^5$.

In fact, a further analysis of low degree mappings indicates that the following phenomena are typical: If one takes a sequence $(\Sigma_k, f_k)$ of smooth holomorphic curves in $\mathbb{CP}^2$, then after reparametrizing and passing to a subsequence, one can arrange that the holomorphic curves have the property that, at a finite number of points $p_k^\alpha \in \Sigma$, the induced metric $f_k^*(\Omega)$ on the surface goes to infinity and the integral of the induced area form on a neighborhood of each of these points approaches an integer while along a finite number of loops $\gamma_i$, the induced metric goes to zero.

The first type of phenomenon is called "bubbling", for what is happening is that a small 2-sphere is inflating and "breaking off" from the surface and covering a line in $\mathbb{CP}^2$. The second type of phenomenon is called "vanishing cycles", a loop in the surface is literally contracting to a point. It turns out that the limiting object in $\mathbb{CP}^2$ is a union of algebraic curves whose total degree is the same as that of the members of the varying family.

Thus, for $\mathbb{CP}^2$, the moduli space $\mathfrak{M}_d$ of unparametrized curves of degree $d$ has a compactification $\overline{\mathfrak{M}_d}$ where the extra points represent decomposable or degenerate curves with "cusps".

Other instances of this "bubbling" phenomena have been discovered. Sacks and Uhlenbeck showed that when one wants to study the question of representing elements of $\pi_2(X)$ (where $X$ is a Riemannian manifold) by harmonic or minimal surfaces, one has to deal with the possibility of pieces of the surface "bubbling off" in exactly the fashion described above.

More recently, this sort of phenomenon has been used in "reverse" by Taubes to construct solutions to the (anti-)self dual Yang-Mills equations over compact 4-manifolds.

With all of this evidence of good compactifications of moduli spaces in other problems, Gromov had the idea of trying to compactify the connected components of the "moduli

space" $\mathfrak{M}$ of holomorphic curves in a general almost complex manifold $M$. Since one would certainly expect the area function to be continuous on each compactified component, it follows that there is not much hope of finding a good compactification of the components of $\mathfrak{M}$ in a case where the area function is not bounded on the components of $\mathfrak{M}$ (as in the case of the Heisenberg example above).

However, it is still possible that one might be able to produce such a compactification if one *can* get an area bound on the curves in each component. Gromov's insight was that having the area bound was enough to furnish *a priori* estimates on the derivatives of curves with an area bound, at least away from a finite number of points.

With all of this in mind, I can now very roughly state Gromov's Compactification Theorem:

**Theorem 5:** *Let $M$ be a compact almost complex manifold with almost complex structure $J$ and suppose that $\Omega$ tames $J$. Then every component $\mathfrak{M}_\alpha$ of the moduli space $\mathfrak{M}$ of connected unparametrized holomorphic curves in $M$ can be compactified to a space $\overline{\mathfrak{M}_\alpha}$ by adding a set of "cusp" curves, where a cusp curve is essentially a finite union of (possibly) singular holomorphic curves in $M$ which is obtained as a limit of a sequence of connected curves in $\mathfrak{M}_\alpha$ by "pinching loops" and "bubbling".*

For the precise definition of "cusp curve" consult [Gr 1], [Wo], or [Pa]. The method that Gromov uses to prove his compactness theorem is basically a generalization of the Schwarz Lemma. This allows him to get control of the sup-norm of the first derivatives of a holomorphic curve in $M$ terms of the $L_1^2$-norm (i.e., area norm) at least in regions where the area form stays bounded.

Unfortunately, although the ideas are intuitively compelling, the actual details are non-trivial. However, there are, by now, several good sources, from different viewpoints, for proofs of Gromov's Compactification Theorem. The articles [M 4] and [Wo] listed in the Bibliography are very readable accounts and are highly recommended. I hear that the (unpublished) [Pa] is also an excellent account which is closer in spirit to Gromov's original ideas of how the proof should go. Finally, there is the quite recent [Ye], which generalizes this compactification theorem to the case of curves with boundary.

Actually, the most fruitful applications of these ideas have been in the situation when, for various reasons, it turns out that there cannot be any cusp curves, so that, by Gromov's compactness theorem, the moduli space is already compact. Here is a case where this happens.

**Proposition 1:** *Suppose that $(M, J)$ is a compact, almost complex manifold and that $\Omega$ is a 2-form which tames $J$. Suppose that there exists a non-constant holomorphic curve $f: S^2 \to M$, and suppose that there is a number $A > 0$ so that $\int_{S^2} f^*(\Omega) \geq A$ for all non-constant holomorphic maps $f: S^2 \to M$. Then for any $B < 2A$, the set $\mathfrak{M}_B$ of unparametrized holomorphic curves $f(S^2) \subset M$ which satisfy $\int_{S^2} f^*(\Omega) = B$ is compact.*

PROOF: (IDEA) If the space $\mathfrak{M}_B$ were not compact, then a point of the compactification would would correspond a union of cusp curves which would contain at least two distinct non-constant holomorphic maps of $S^2$ into $M$. Of course, this would imply that the limiting value of the integral of $\Omega$ over this curve would be at least $2A > B$, a contradiction. $\square$

An example of this phenomenon is when the taming form $\Omega$ represents an integral class in cohomology. Then the presence of any holomorphic rational curves at all implies that there is a compact moduli space at some level.

**Applications.** It is reasonable to ask how the Compactness Theorem can be applied in symplectic geometry.

To do this, what one typically does is first fix a symplectic manifold $(M, \Omega)$ and then considers the space $\mathcal{J}(\Omega)$ of almost complex structures on $M$ which $\Omega$ tames. We already know from Lecture 5 that $\mathcal{J}(\Omega)$ is not empty. We even know that the space $\mathcal{K}(\Omega) \subset \mathcal{J}(\Omega)$ of $\Omega$-compatible almost complex structures is non-empty. Moreover, it is not hard to show that these spaces are contractible (see the Exercises).

▶      *Thus, any invariant of the almost complex structures $J \in \mathcal{J}(\Omega)$ or of the almost Kähler structures $J \in \mathcal{K}(\Omega)$ which is constant under homotopy through such structures is an invariant of the underlying symplectic manifold $(M, \Omega)$.*

This idea is extremely powerful. Gromov has used it to construct many new invariants of symplectic manifolds. He has then gone on to use these invariants to detect features of symplectic manifolds which are not presently accessible by any other means.

Here is a sample of some of the applications of Gromov's work on holomorphic curves. Unfortunately, I will not have time to discuss the proofs of any of these results.

**Theorem 6:** (GROMOV) *If there is a symplectic embedding of $B^{2n}(r) \subset \mathbb{R}^{2n}$ into $B^2(R) \times \mathbb{R}^{2n-2}$, then $r \leq R$.*

One corollary of Theorem 6 is that any diffeomorphism of a symplectic manifold which is a $C^0$-limit of symplectomorphisms is itself a symplectomorphism.

**Theorem 7:** (GROMOV) *If $\Omega$ is a symplectic structure on $\mathbb{CP}^2$ and there exists an embedded $\Omega$-symplectic sphere $S \subset \mathbb{CP}^2$, then $\Omega$ is equivalent to the standard symplectic structure.*

The next two theorems depend on the notion of asymptotic flatness: We say that a non-compact symplectic manifold $M^{2n}$ is *asymptotically flat* if there is a compact set $K_1 \subset M^{2n}$ and a compact set $K_2 \subset \mathbb{R}^{2n}$ so that $M \setminus K_1$ is symplectomorphic to $\mathbb{R}^{2n} \setminus K_2$ (with the standard symplectic structure on $\mathbb{R}^{2n}$).

**Theorem 8:** (MCDUFF) *Suppose that $M^4$ is a non-compact symplectic manifold which is asymptotically flat. Then $M^4$ is symplectomorphic to $\mathbb{R}^4$ with a finite number of points blown up.*

**Theorem 9:** (McDuff, Floer, Eliashberg) *Suppose that $M^{2n}$ is asymptotically flat and contains no symplectic 2-spheres. Then $M^{2n}$ is diffeomorphic to $\mathbb{R}^{2n}$.*

It is not known whether one might replace "diffeomorphic" with "symplectomorphic" in this theorem for $n > 2$.

## Epilogue

I hope that this Lecture has intrigued you as to the possibilities of applying the ideas of Gromov in modern geometry. Let me close by quoting from Gromov's survey paper on symplectic geometry in the Proceedings of the 1986 ICM:

> Differential forms (of any degree) taming partial differential equations provide a major (if not the only) source of integro-differential inequalities needed for a priori estimates and vanishing theorems. These forms are defined on spaces of jets (of solutions of equations) and they are often (e.g., in Bochner-Weitzenbock formulas) exact and invariant under pertinent (infinitesimal) symmetry groups. Similarly, convex (in an appropriate sense) functions on spaces of jets are responsible for the maximum principles. A great part of hard analysis of PDE will become redundant when the algebraic and geometric structure of taming forms and corresponding convex functions is clarified. (From the PDE point of view, symplectic geometry appears as a taming device on the space of 0-jets of solutions of the Cauchy-Riemann equation.)

# Exercise Set 9:

# The Gromov School of Symplectic Geometry

**1**. Use the fact that an orientable 3-manifold $M^3$ is parallelizable (i.e., its tangent bundle is trivial) and Theorem 1 to show that a compact 3-manifold can always be immersed in $\mathbb{R}^4$ and a 3-manifold with no compact component can always be immersed in $\mathbb{R}^3$.

**2**. Show that Theorem 2 does, in fact imply that any connected non-compact symplectic manifold which has an almost complex structure has a symplectic structure. (Hint: Show that the natural projection $Z^1(X, V) \to V$ has contractible fibers (in fact, $Z^1(X, V)$ is an affine bundle over $V$, and then use this to show that a non-degenerate 2-form on $X$ can be homotoped to a closed non-degenerate 2-form on $X$.)

**3**. Show that the hypothesis in Theorem 3 that $X$ either be non-compact or that $\dim(X) < \dim(Y)$ is essential.

**4**. Show that if $E$ is a symplectic bundle over a compact manifold $M^{2n}$ whose rank is $2n + 2k$ for some $k > 0$, then there exists a symplectic splitting $E = F \oplus T$ where $T$ is a trivial symplectic bundle over $M$ of rank $k$. (Hint: Use transversality to pick a non-vanishing section of $E$. Now what?)

    Show also that, if $E$ is a symplectic bundle over a compact manifold $M^{2n}$, then there exists another symplectic bundle $E'$ over $M$ so that $E \oplus E'$ is trivial. (Hint: Mimic the proof for complex bundles.)

    Finally, use these results to complete the proof of the Corollary to Theorem 3.

**5**. Show that if a symplectic manifold $M$ is simply connected, then the symplectic blow up $\hat{M}$ of $M$ along a symplectic submanifold $S$ of $M$ is also simply connected. (Hint: Any loop in $\hat{M}$ can be deformed into a loop which misses $\hat{S}$. Now what?)

**6**. Prove, as stated in the text, that if $M$ is compact and $\Omega$ tames $J$, then for any Riemannian metric $g$ on $M$ (not necessarily compatible with either $J$ or $\Omega$) there is a constant $C > 0$ so that

$$|v \wedge Jv| \le C\, \Omega(v, Jv)$$

where $|v \wedge Jv|$ represents the area in $T_p M$ of the parallelogram spanned by $v$ and $Jv$ in $T_p M$.

**7.** FIRST ORDER EQUATIONS AND HOLOMORPHIC CURVES. The point of this problem is to show how elliptic quasi-linear determined PDE for two functions of two unknowns can be reformulated as a problem in holomorphic curves in an almost complex manifold.

Suppose that $\pi \colon V^4 \to X^2$ is a smooth submersion from a 4-manifold onto a 2-manifold. Suppose also that $R \subset J^1(X,V)$ is smooth submanifold of dimension 6 which has the property that it locally represents an elliptic, quasi-linear pair of first order PDE for sections $s$ of $\pi$. Show that there exists a unique almost complex structure on $V$ so that a section $s$ of $\pi$ is a solution of $R$ if and only if its graph in $V$ is an (unparametrized) holomorphic curve in $V$.

(Hint: The hypotheses on the relation $R$ are equivalent to the following conditions. For every point $v \in V$, there are coordinates $x, y, f, g$ on a neighborhood of $v$ in $V$ with the property that $x$ and $y$ are local coordinates on a neighborhood of $\pi(v)$ and so that a local section $s$ of the form $f = F(x,y)$, $g = G(x,y)$ is a solution of $R$ if and only if they satisfy a pair of equations of the form

$$A_1\,f_x + B_1\,f_y + C_1\,g_x + D_1\,g_y + E_1 = 0$$
$$A_2\,f_x + B_2\,f_y + C_2\,g_x + D_2\,g_y + E_2 = 0$$

where the $A_1, \ldots, E_2$ are specific functions of $(x, y, f, g)$. The ellipticity condition is equivalent to the assumption that

$$\det \begin{pmatrix} A_1\xi + B_1\eta & C_1\xi + D_1\eta \\ A_2\xi + B_2\eta & C_2\xi + D_2\eta \end{pmatrix} > 0$$

for all $(\xi, \eta) \neq (0,0)$.)

Show that the problem of isometrically embedding a metric $g$ of positive Gauss curvature on a surface $\Sigma$ into $\mathbb{R}^3$ can be turned into a problem of finding a holomorphic section of an almost complex bundle $\pi \colon V \to \Sigma$. Do this by showing that the bundle $V$ whose sections are the quadratic forms which have positive $g$-trace and which satisfy the algebraic condition imposed by the Gauss equation on quadratic forms which are second fundamental forms for isometric embeddings of $g$ is a smooth rank 2 disk bundle over $\Sigma$ and that the Codazzi equations then reduce to a pair of elliptic first order quasi-linear PDE for sections of this bundle.

Show that, if $\Sigma$ is topologically $S^2$, then the topological self-intersection number of a global section of V is $-4$. Conclude, using the fact that distinct holomorphic curves in $V$ must have positive intersection number, that (up to sign) there cannot be more that one second fundamental form on $\Sigma$ which satisfies both the Gauss and Codazzi equations. Thus, conclude that a closed surface of positive Gauss curvature in $\mathbb{R}^3$ is rigid.

This approach to isometric embedding of surfaces has been extensively studied by Labourie [La].

**8**. Prove, as claimed in the text that, for the map $f_\lambda \colon \mathbb{CP}^1 \to \mathbb{CP}^2$ given by the rule

$$f_\lambda(t) = [t, \lambda t^2, \lambda] = [1, x, y],$$

the pull-back of the Fubini-Study metric accumulates at the points $t = 0$ and $t = \infty$ and goes to zero everywhere else. What would have happened if, instead we had used the map

$$g_\lambda(t) = [t, t^2, \lambda^2] = [1, x, y]?$$

Is there a contradiction here?

**9**. Verify the claim made in the text that, for a symplectic manifold $(M, \Omega)$, the spaces $\mathcal{K}(\Omega)$ and $\mathcal{J}(\Omega)$ of $\Omega$-compatible and $\Omega$-tame almost complex structures on $M$ are indeed contractible. (Hint: Fix an element $J_0 \in \mathcal{K}(\Omega)$, with associated inner product $\langle , \rangle_0$ and show that, for any $J \in \mathcal{J}(\Omega)$, we can write $J = J_0(S + A)$ where $S$ is symmetric and positive definite with respect to $\langle , \rangle_0$ and $A$ is anti-symmetric. Now what?)

**10**. The point of this exercise is to get a look at the pseudo-holomorphic curves of a non-integrable almost complex structure. Let $X^4 = \mathbb{C} \times \Delta = \{(w, z) \in \mathbb{C}^2 \,|\, |z| < 1\}$, and give $X^4$ the almost complex structure for which the complex valued 1-forms $\alpha = dw + \bar{z}\, d\bar{w}$ and $\beta = dz$ are a basis for the $(1, 0)$-forms. Verify that this does indeed define a non-integrable almost complex structure on the 4-manifold $X$. Show that the pseudo-holomorphic curves in $X$ can be described explicitly as follows: If $M$ is a Riemann surface and $\phi \colon M^2 \to X$ is a pseudo-holomorphic mapping, then one of the following is true: Either $\phi^*(\beta) = 0$ and there exists a holomorphic function $h$ on $M$ and a constant $z_0$ so that

$$\phi = \bigl(h - \bar{z}_0\, \bar{h},\ z_0 \bigr),$$

or else there exists a non-constant holomorphic function $g$ on $M$ which satisfies $|g| < 1$, a meromorphic function $f$ on $M$ so that $f\, dg$ and $fg\, dg$ are holomorphic 1-forms without periods on $M$, and a constant $w_0$ so that

$$\phi(p) = \left( w_0 + \int_{p_0}^{p} (\, f\, dg - \overline{fg\, dg}\,),\ \ g(p) \ \right)$$

where the integral is taken to be taken over any path from some basepoint $p_0$ to $p$ in $M$.

   (Hint: It is obvious that you must take $g = \phi^*(z)$, but it is not completely obvious where $f$ will be found. However, if $g$ is not a constant function, then it will clearly be holomorphic, now consider the "function"

$$f = \frac{\phi^*(\alpha)}{\bigl(1 - |g|^2\bigr)\, dg}$$

and show that it must be meromorphic, with poles at worst along the zeroes of $dg$.)

# Bibliography

[A]   V. I. ARNOL'D, *Mathematical Methods of Classical Mechanics, Second Edition,* Springer-Verlag, Berlin, Heidelberg, New York, 1989.

[AH]   M. F. ATIYAH and N. HITCHIN, *The Geometry and Dynamics of Magnetic Monopoles*, Princeton University Press, Princeton, 1988.

[Bea]   A. BEAUVILLE, *Variétés Kähleriennes dont la première classe de Chern est nulle,* J. Differential Geom. **18** (1983), 755–782.

[Ber]   M. BERGER, *Sur les groupes d'holonomie homogène des variétés a connexion affine et des variétés Riemanniennes,* Bull. Soc. Math. France **83** (1955), 279–330.

[Bes]   A. BESSE, *Einstein Manifolds*, Springer-Verlag, Berlin, Heidelberg, New York, 1987.

[Ch]   S.-S. CHERN, *On a generalization of Kähler geometry*, in *Algebraic Geometry and Topology: A Symposium in Honor of S. Lefshetz*, Princeton University Press, 1957, 103–121.

[FGG]   M. FERNANDEZ, M. GOTAY, and A. GRAY, *Four-dimensional parallelizable symplectic and complex manifolds,* Proc. Amer. Math. Soc. **103** (1988), 1209–1212.

[Fo]   A. T. FOMENKO, *Symplectic Geometry*, Gordon and Breach, New York, 1988.

[GL]   K. GALICKI and H.B. LAWSON, *Quaternionic Reduction and quaternionic orbifolds,* Math. Ann. **282** (1988), 1–21.

[GG]   M. GOLUBITSKY and V. GUILLEMIN , *Stable Mappings and their Singularities,* Graduate Texts in Mathematics **14**, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[Gr 1]   M. GROMOV, *Pseudo-holomorphic curves on almost complex manifolds,* Inventiones Mathematicae **82** (1985), 307–347.

[Gr 2]   M. GROMOV, *Partial Differential Relations,* Ergebnisse der Math., Springer-Verlag, Berlin, Heidelberg, New York, 1986.

[Gr 3]   M. GROMOV, *Soft and hard symplectic geometry,* Proceedings of the ICM at Berkeley, 1986, 1987, vol. 1, Amer. Math. Soc., Providence, R.I., 81–98.

[GS 1]   V. GUILLEMIN and S. STERNBERG, *Geometric Asymptotics,* Mathematical Surveys **14**, Amer. Math. Soc., Providence, R.I., 1977.

[GS 2]   V. GUILLEMIN and S. STERNBERG, *Symplectic Techniques in Physics,* Cambridge University, Cambridge and New York, 1984.

[GS 3]   V. GUILLEMIN and S. STERNBERG, *Variations on a Theme of Kepler,* Colloquium Publications **42**, Amer. Math. Soc., Providence, R.I., 1990.

[Ha]   R. HAMILTON, *The inverse function theorem of Nash and Moser,* Bulletin of the AMS **7** (1982), 65–222.

[He]  S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces,* Academic Press, Princeton, 1978.

[Hi]  N. Hitchin, *Metrics on moduli spaces,* in Contemporary Mathematics **58** (1986), Amer. Math. Soc., Providence, R.I., 157–178.

[HKLR]  N. Hitchin, A. Karlhede, U. Lindström, and M. Roček, *HyperKähler metrics and supersymmetry,* Comm. Math. Phys. **108** (1987), 535–589.

[Ki]  F. Kirwan, *Cohomology of Quotients in Symplectic and Algebraic Geometry*, Mathematical Notes **31**, Princeton University Press, Princeton, 1984.

[KN]  S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, vols. I and II*, John Wiley & Sons, New York, 1963.

[Kr]  P. Kronheimer, *The construction of ALE spaces as hyper-Kähler quotients*, J. Differential Geometry **29** (1989), 665-683.

[La]  F. Labourie, *Immersions isométriques elliptiques et courbes pseudo-holomorphes*, J. Differential Geometry **30** (1989), 395-424.

[M 1]  D. McDuff, *Symplectic diffeomorphisms and the flux homomorphism,* Inventiones Mathematicae **77** (1984), 353–366.

[M 2]  D. McDuff, *Examples of simply-connected symplectic non-Kählerian manifolds,* J. Differential Geom. **20** (1984), 267–277.

[M 3]  D. McDuff, *Examples of symplectic structures,* Inventiones Mathematicae **89** (1987), 13–36.

[M 4]  D. McDuff, *Symplectic 4-manifolds,* Proceedings of the ICM at Kyoto, 1990, to appear

[M 5]  D. McDuff, *Elliptic methods in symplectic geometry,* Bull. Amer. Math. Soc. **23** (1990), 311–358.

[MS]  J. Milnor and J. Stasheff, *Characteristic Classes,* Annals of Math. Studies **76**, Princeton University Press, Princeton, 1974.

[Mo]  J. Moser, *On the volume element on manifolds,* Trans. Amer. Math. Soc. **120** (1965), 280–296.

[Pa]  P. Pansu, *Pseudo-holomorphic curves in symplectic manifolds,* preprint, École Polytechnique, Palaiseau, 1986.

[Po]  I. Postnikov, *Groupes et algèbras de Lie,* (French translation of the Russian original), Éditions Mir, Moscow 1985.

[Sa]  S. Salamon, *Riemannian geometry and holonomy groups*, Pitman Research Notes in Math. no. 201, Longman scientific & Technical, Essex, 1989.

[Si]  Y. T. Siu, *Every K3 surface is Kähler*, Inventiones Math. **73** (1983), 139–150.

[SS] I. SINGER and S. STERNBERG, *The infinite groups of Lie and Cartan, I (The transitive groups),* Journal d'Analyse **15** (1965).

[Wa] F. WARNER, *Foundations of Differentiable Manifolds and Lie Groups,* Springer-Verlag, Berlin Heidelberg New York, 1983.

[We] A. WEINSTEIN, *Lectures on Symplectic Manifolds,* Regional Conference Series in Mathematics **29**, Amer. Math. Soc., Providence, R.I., 1976.

[Wo] J. WOLFSON, *Gromov's compactness of pseudo-holomorphic curves and symplectic geometry,* J. Differential Geom. **28** (1988), 383–405.

[Ye] R. YE, *Gromov's compactness theorem for pseudo-holomorphic curves,* Transactions of the AMS , to appear.