MATH 137 NOTES

ARUN DEBRAY
MARCH 28, 2013

Contents

1. Galilean Relativity and Newtonian Determinacy: 1/7/13

Mathematical methods of classical mechanics means mathematics — the course will deal with mathematical questions arising from physics rather than focusing on their physical meanings. Additionally, since this is classical mechanics, some of these claims will be slightly untrue due to special relativity and such, but this doesn't affect the scope of the class.

Classical mechanics is based on two principles: Galilean relativity and Newtonian determinacy.

Galilean relativity. Galilean relativity claims that, relative to some coordinate system called the inertial frame, spacetime is a 4-dimensional affine space A (i.e. a vector space, but without a fixed or preferred origin).

Definition. An affine function is a function on affine space that is linear up to a constant once the origin is identified. Note that there cannot be linear functions on affine spaces in general.

Additionally, there is a time function $A \xrightarrow{t} \mathbb{R}$ such that each "slice" of A with t constant has a Euclidean structure (i.e. a scalar product).

All laws of nature should be invariant under any transformation that preserves this structure. For example, there is no preferred transverse line to these slices, so shear transformations can go between these lines.

The Galileo group is this group of structure-preserving transformations of A. Different transverse lines represent different observers (in a fixed coordinate system): one might see the other traveling at a constant velocity, etc. There is no canonical projection onto space, unless such a transverse line is chosen. Then, the trajectory of a particle is a curve in three-dimensional Euclidean space. The projection onto time is canonical, though.

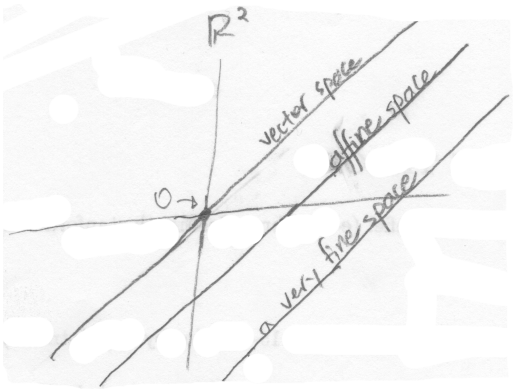

FIGURE 1. Vector spaces and affine spaces on $\mathbb{R}^2$.

Newtonian determinacy. For Newtonian determinacy, it is necessary to assume such an observer, so that velocity can be talked about. This principle states that if one knows the positions and velocities of all particles in a closed system, you can predict future states of that system.

The notion of degrees of freedom will be formalized more precisely later on, but one can consider systems with finitely many degrees of freedom (in which knowing a finite number of things completely determines the system), in which the positions and velocities lie in a configuration space $x \in \mathbb{R}^{3N}$, for some (large) N. Then, knowing $x(t)$ and $\dot{x}(t)$ at one $t \in \mathbb{R}$ determines the whole curve. In particular, it determines the second derivative, $\ddot{x} = F(x, \dot{x}, t)$, so the movement of a mechanical system is given by a 2^{nd}-order differential equation, and depends completely on F.

From a mathematical point of view, physics ends here; the solution of the differential equation is entirely mathematical.

Consider some closed system.[1] Then, Galilean relativity implies that the system cannot be dependent om time (since time-changing transformations are part of the Galileo group), giving $\ddot{x} = F(x, \dot{x})$.

If the closed system contains only one particle, both position and velocity can be changed by the group, so $\ddot{x} = c$. In addition, this constant must be invariant with respect to rotation, so $\ddot{x} = 0$, so $x = at + b$. In particular, there is some frame in which the particle is stationary.

Similar statements can be made for other systems: for example, a system with 2 particles for which $\dot{x}(0) = 0$ must oscillate along the line between them, since it would be otherwise variant under rotation. More generally, any system of two particles is confined to a single plane.

Systems with holonomic constraints. Describing a system as a trajectory in $\mathbb{R}^{3N}$ is sometimes a waste of space; for example, on a rigid body, it's possible to choose an orthogonal frame (an element of the orthogonal group, which is a 3-dimensional manifold) which accounts for the rotation of the object. Thus, a rigid 3-dimensional body can be described in a configuration space of $\mathbb{R}^3 \times \mathrm{O}(3)$, which is only 6-dimensional (3 fewer dimensions than before)!

Definition. A system with holonomic constraints is one in which the configuration space is given by a submanifold of $\mathbb{R}^{3N}$ as opposed to the entire manifold.

Skating is an example of a system with nonholonomic constraints: a skater lies on a plane and has a direction given by an angle φ, so the space is $\mathbb{R}^2 \times S^1$. Specifically, there is no constraint on $\dot{\varphi}$, but $\dot{x}_1 \cos\varphi + \dot{x}_2 \sin\varphi = 0$.

Consider the differential 1-form $\sin\varphi\,\mathrm{d}x_1 + \cos\varphi\,\mathrm{d}x_2$.[2] This yields a Pfaffian equation $) = \sin\varphi\,\mathrm{d}x_1 + \cos\varphi\,\mathrm{d}x_2$, which determines a family of planes at every point (i.e. one plane per point). Then, a *nonholonomic constraint* is a constraint that requires solutions to be tangent to the plane field at every point. In one dimension, this is always globally possible, and holonomic and nonholonomic constraints are equivalent, but in higher dimensions this isn't the case, and systems with nonholonomic constraints may not even have local solutions. However, this class will be more concerned with holonomic constraints.

[1] No system in the real world is entirely closed, though a suitable approximation may be obtained by isolating it from anything else, as in, say, Montana.

[2] A differential 1-form is a family of linear functions. Choose some $\mathbf{x} = (x_1, x_2, \varphi)$, and consider $\mathbb{R}^3$ with the origin at this point (which can be written $(x_1^0, x_2^0, \varphi^0)$) and write coordinates as h_1, h_2, ψ. Then, $\sin\varphi_0 h_1 + \cos\varphi_0 h_2$ is a linear function that depends on the choice of $\mathbf{x}$. This whole set of functions is called a differential 1-form, notated $\sin\varphi\,\mathrm{d}x_1 + \cos\varphi\,\mathrm{d}x_2$.

Thus, in a system with holonomic constraints and N particles, there is a submanifold[3] $X \subset \mathbb{R}^{3N}$. Often, X can be thought of abstractly rather than as a submanifold.[4]

Thus, motion of the system is given by a path $x : [a, b] \to X$, so that $x(t)$ describes the development of the system. Then, $\dot{x}(t)$ is the velocity, and $\ddot{x}$ the acceleration.

Newton's determinacy says that $\ddot{x} = F(x, \dot{x}, t)$. But where does F arise from? There are two formalisms of mechanics that answer this: the Lagrangian formalism and the Hamiltonian formalism.

The Lagrangian formalism. The Lagrangian formalism considers the (infinite-dimensional) space $P(X)$ or $\mathcal{P}$ of all paths on X. There exists a functional[5] called the action functional, and motion is determined by the extrema of this functional.

Consider the space TX of all tangent spaces; if $a \in X$, then the tangent space at a is T_aX, the set of all tangent vectors at a. TX is a $2n$-dimensional manifold if X is n-dimensional, and is called the tangent bundle of X.

Suppose X has local coordinates $x_1, \dots, x_n$ at some $a \in X$. These allow the definition of a basis at every tangent space, given by traveling in each direction $x_1, \dots, x_n$ with speed 1. This basis is denoted $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$, and can be thought of as taking the partial derivative in each direction.

Then, the overall coordinates are written $\dot{x}_1 \frac{\partial}{\partial x_1} + \cdots + \dot{x}_n \frac{\partial}{\partial x_n}$. This is confusing, but unfortunately is the standard notation; don't think of $\dot{x}_i$ as the derivative with respect to time, but just some other letter. The coordinates $(x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n)$ determine an element of TX with coordinates $\sum_{i=1}^n \dot{x}_i \frac{\partial}{\partial x_i}$. Changing the coordinates in X changes the coordinates $\dot{x}_1, \dots, \dot{x}_n$ by the Jacobi matrix (so it is a smooth transformation).

Then, consider a function $L : TX \times \mathbb{R} \to \mathbb{R}$ (i.e. to every tangent vector and time associate a number), $L(\mathbf{x}, \dot{\mathbf{x}}, t)$. Given such a function, one has the following functional on $\mathcal{P} = \{[0, 1] \to X\}$: if $\mathbf{x}(t) \in \mathcal{P}$, then

$$S(\mathbf{x}(t)) = \int_0^1 L(\mathbf{x}, \dot{\mathbf{x}}, t)\, \mathrm{d}t,$$

where the dot is a derivative with respect to time. This means that if $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))$ is a path on X, then it induces a path $\mathbf{X}(t) = (x_1(t), \dots, x_n(t), \frac{\partial x_1}{\partial t}, \dots, \frac{\partial x_n}{\partial t})$. However, the notational ambiguity returns: each of the $\frac{\partial x_i}{\partial t}$ can be written as $\dot{x}_i$, but then the dot also means coordinates in the tangent bundle.

Example 1.1. If $X = \mathbb{R}$, then $L(x, \dot{x}, t)$ is just a function of 3 real variables, which yields a number for any path x given by $S(x) = \int_0^1 L(x(t), \frac{\mathrm{d}x}{\mathrm{d}t}, t)\, \mathrm{d}t$.

Thus, mechanics should correspond to extrema of this functional. L is called the Lagrangian, and $S = \int_0^1 L\, \mathrm{d}t$ is called the action functional.

This formalism satisfies the least action principle, since the least actions are extrema. Thus, it also satisfies the Euler-Lagrange differential equations, so the question for a system boils down to writing down the Lagrangian. This is particularly nice because it is so much more general than elementary approaches to classical mechanics (which involve many changes of variables). However, some systems (e.g. the three-body problem) aren't completely integrable. Thus, questions in classical mechanics in the 20^{th} Century focused on qualitative statements about such systems (e.g. whether it is periodic, or stable; for example, the solar system is not believed to be stable over the long term) and numerical approximations.

2. The Euler-Lagrange Equation: 1/9/13

A submanifold is a structure that looks like a smooth function locally: if a point on a submanifold has coordinates $(x_1, \dots, x_k)$, then it can be rewritten as $(\mathbf{x}, \mathbf{x}')$, where $\mathbf{x} = (x_1, \dots, x_j)$ and $\mathbf{x}' = (x_{j+1}, \dots, x_n)$, such that $\mathbf{x}' = f(\mathbf{x})$ locally.

A configuration space is typically a submanifold, but sometimes may contain singular points (e.g. the union of two intersecting manifolds). In this case, it can be stratified, and the problem can be analyzed on each strata. For example, collisions in an n-body problem are singularities.

[3]Euclidean space has the property that everything can be determined by N coordinates. Some spaces satisfy this locally, even if not necessarily globally. For example, the angle on a circle gives a coordinate, but no single system works continuously over the whole space, and there must be coordinate changes. The space is a manifold if these coordinate changes are given by a smooth function.

[4]Not all systems will admit a manifold; one can obtain something such as $\{xy = 0\}$, which can be broken into different manifolds. This is also not part of the scope of this class.

[5]A function on a space of functions is traditionally called a functional, but this is just semantics; a functional is nothing more than a function.

Consider a configuration space $\mathbb{R}^{3N}$ and two points $\mathbf{x}_0, \mathbf{x}_1 \in \mathbb{R}^{3N}$. Then, the space of paths is $\mathcal{P} = \{\gamma : [0,1] \to \mathbb{R}^{3N}\}$, and the space of paths between $\mathbf{x}_0$ and $\mathbf{x}_1$ is $\mathcal{P}_{\mathbf{x}_0,\mathbf{x}_1} = \{\gamma \in \mathcal{P} \mid \gamma(0) = \mathbf{x}_0, \gamma(1) = \mathbf{x}_1\}$. $\mathcal{P}$ is a vector space, and has a norm $\|\gamma\| = \max_{t\in[0,1]} \|\gamma(t)\|$, called the C^0-norm. Thus, one can talk about continuous and differentiable functions on $\mathcal{P}$: a functional $S : \mathcal{P} \to \mathbb{R}$ is differentiable at a $\gamma_0 \in \mathcal{P}$ if $S(\gamma_0 + \mathbf{h}) = S(\gamma_0) + A_{\gamma_0}(\mathbf{h}) + o(\|\mathbf{h}\|)$, where $A_{\gamma_0}(\mathbf{h})$ is linear and the higher-order terms go to zero.

For example, consider $L : \mathbb{R}^{3N} \times \mathbb{R}^{3N} \times \mathbb{R} \to \mathbb{R}$, where the domain can be thought of as $\mathbb{R}^{6N+1} = \mathbb{R}^{3N} \times \mathbb{R}^{3N} \times \mathbb{R} = T(\mathbb{R}^{3N}) \times \mathbb{R}$ (since the tangent bundle of Euclidean space is this direct product). Denoting paths as $\mathbf{x}(t)$,[6] one can obtain a differentiable functional from L: let $S(\mathbf{x}(t)) = \int_0^1 L(\mathbf{x}, \dot{\mathbf{x}}, t)\,\mathrm{d}t$, where $\dot{\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$. Then, if L is differentiable with respect to $\mathbf{x}, \dot{\mathbf{x}}$, and t, then S is a differentiable functional: using a Taylor expansion,

$$\begin{aligned} S(\mathbf{x}+\mathbf{h}) &= \int_0^1 \left(L(\mathbf{x}, \dot{\mathbf{x}}, t) + \frac{\partial L}{\partial \mathbf{x}}\mathbf{h} + \frac{\partial L}{\partial \dot{\mathbf{x}}}\dot{\mathbf{h}} + o(\|\mathbf{h} + \dot{\mathbf{h}}\|) \right) \mathrm{d}t \\ &= \int_0^1 L(\mathbf{x}, \dot{\mathbf{x}}, t)\,\mathrm{d}t + A(\mathbf{x})\mathbf{h} + o(\|\mathbf{h} + \dot{\mathbf{h}}\|)\,\mathrm{d}t, \end{aligned}$$

where A is linear because differentiation is linear. Since $\mathbf{x}$ and $\mathbf{h}$ are vectors, then the notation above expands into $\frac{\partial L}{\partial \mathbf{x}} = \left(\frac{\partial L}{\partial x_1}, \ldots, \frac{\partial L}{\partial x_n}\right)$ and the vector products are the dot product.

Note that S is not necessarily differentiable if L is C^0-smooth; it must be C^1-smooth, which is given by the C^1 norm $\|\gamma\|_{C^1} = \max_{t\in[0,1]}(\|\gamma(t)\|, \|\gamma'(t)\|)$.

Just as in single-variable calculus, it is a necessary but not sufficient condition for (local) extrema of S that the differential vanishes. The points where these occur (sometimes critical points) are known as extremal points, though they might not be extrema.

In calculus of variations, the differential is called the first variation, and is written $A = \delta S(\mathbf{x})$ (with δ used because $\mathcal{P}$ is infinite-dimensional). Integrating by parts,

$$\delta S(\mathbf{x}) = \int_0^1 \left(\frac{\partial L}{\partial \mathbf{x}}\mathbf{h} + \frac{\partial L}{\partial \dot{\mathbf{x}}}\dot{\mathbf{h}} \right) \mathrm{d}t = \int_0^1 \left(\frac{\partial L}{\partial \mathbf{x}}\mathbf{h} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\mathbf{x}}}\mathbf{h} \right) \mathrm{d}t + \left.\frac{\partial L}{\partial \dot{\mathbf{x}}}\mathbf{h}\right|_0^1.$$

Now, restrict S to $\mathcal{P}_{\mathbf{x}_0,\mathbf{x}_1}$, so that the corresponding variations must also have their endpoints fixed. Thus, $\left.\frac{\partial L}{\partial \dot{\mathbf{x}}} \cdot \mathbf{h}\right|_0^1 = 0$, so $\mathbf{x}$ is a critical point if

$$\int_0^1 \left(\frac{\partial L}{\partial \mathbf{x}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \mathbf{x}} \right) \cdot \mathbf{h}\,\mathrm{d}t = 0$$

for any $\mathbf{h}$ which vanishes at the endpoints.

Suppose $\varphi : [0,1] \to \mathbb{R}$ is continuous. If $\int_0^1 \varphi(t)h(t)\,\mathrm{d}t = 0$ for all h such that $h(0) = h(1) = 0$, then $\varphi = 0$, because if $\varphi(x)$ is nonzero for some $x \in [0,1]$, then one can make a continuous function h that is zero everywhere except an arbitrarily small neighborhood of x but such that $h(x) = 1$, giving a positive integral.

This is also true in the vector-valued case, for the same reasons. Thus:

Theorem 2.1 (Euler-Lagrange equation)**.** $\mathbf{x}(t)$ *is extremal for* $S(\mathbf{x}) = \int_0^1 L(\mathbf{x}, \dot{\mathbf{x}}, t)\,\mathrm{d}t$ *iff*

$$\frac{\partial L}{\partial \mathbf{x}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\mathbf{x}}} = 0.$$

This makes "sense" mneomnically if one thinks of $\dot{\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$.

Example 2.1. Consider $\mathcal{P} = \{x : [0,1] \to \mathbb{R} \mid x(0) = a, x(1) = b\}$ and the functional $S(x) = \int_0^1 \sqrt{1 + \dot{x}^2}\,\mathrm{d}t$, which just computes arc length. Thus, $L = \sqrt{1 + \dot{x}^2}$, so

$$\frac{\partial L}{\partial x} = 0 \text{ and } \frac{\partial L}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}},$$

[6]... which is again going to cause confusing notation. It could be worse, I suppose; consider single-variable calculus students whose notation requires that $x' = 1$, but $c' = 0$.

and the Euler-Lagrange equation gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{x}}{\sqrt{1+\dot{x}^2}}\right)=0, \text{ which means that } \frac{\dot{x}}{\sqrt{1+\dot{x}^2}}=c$$

for some constant c. Thus, $\dot{x}^2=c_1$ for some other constant c_1 after rearranging, and $x(t)=c_1t+c_2$, with c_1 and c_2 determined by the boundary values a and b. Thus, the shortest path between two points is a straight line.

None of the above required Cartesian coordinates, so everything works just as well in another coordinate system. For example, if a curve is given in polar coordinates $r(t)$, $\varphi(t)$, take unit vectors $\mathbf{e}_r$ and $\mathbf{e}_\varphi$ in these directions, then $\mathbf{x}(t)=(r(t),\varphi(t))$ and $\dot{\mathbf{x}}(t)=\dot{r}\mathbf{e}_r+r\dot{\varphi}\mathbf{e}_\varphi$. Since $(\mathbf{e}_r,\mathbf{e}_\varphi)$ is orthnormal, then $|\dot{\mathbf{x}}|^2=\sqrt{\dot{r}^2+r^2\dot{\varphi}^2}$, so the arc length functional can be done in polar form and Example 2.1 still works.

Consider $X_i\in\mathbb{R}^3$ and $\mathbf{x}=(X_1,\dots,X_N)\in\mathbb{R}^{3N}$. According to Newton determinacy, $m_i\ddot{X}_i=F_i(\mathbf{x},\dot{\mathbf{x}},t)$, where the functions F_i are called forces. There is a class of mechanical systems in which F_i is dependent only on $\mathbf{x}$, which gives a function $U(X_1,\dots,X_n)$ such that

$$m_i\ddot{X}_i=F_i(\mathbf{x})=-\frac{\partial U}{\partial X_i}(\mathbf{x}). \tag{1}$$

This system is called a potential system, and corresponds to a set of points moving in a potential field. Many interesting problems take this form; for example, the n-body problem has (up to units)

$$U=\sum_{i,j=1}^{n}\frac{m_im_j}{|X_i-X_j|}.$$

However, this formalism is dependent on Euclidean coordinates, which is a problem. The Lagrangian formulation is more general and therefore more useful.

Considering the potential system above, define the kinetic energy to be $T=\sum\frac{m\|\dot{X}_i^2\|}{2}$ and define $U(\mathbf{X})$ to be the potential energy. Then, let $L:T(\mathbb{R}^{3N})\to\mathbb{R}$ (i.e. L is time-independent) be given by $L(\mathbf{x},\dot{\mathbf{x}})=T-U$, the Lagrange function of the system.

There is a nuance here: $\dot{X}_i$ isn't velocity or a derivative here, but should be thought of as coordinates in the tangent bundle, so that T computes some quadratic form on elements of the tangent bundle.

Theorem 2.2 (Least Action Principle of Hamilton)**.** *The Newton equations (1) are equivalent to the Euler-Lagrange equations for $L(\mathbf{x},\dot{\mathbf{x}})=T-U$.*[7]

Proof.

$$S(\mathbf{x})=\int_0^1 L(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t=\int_0^1\left(\sum_i\frac{m_i\|\dot{X}_i^2\|}{2}-U(X_1,\dots,X_n)\right)\mathrm{d}t.$$

The Euler-Lagrange equation for this is $\frac{\partial L}{\partial X_i}=\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{X}_i}$, but $\frac{\partial L}{\partial X_i}=-\frac{\mathrm{d}U}{\mathrm{d}X_i}$ and $\frac{\partial L}{\partial\dot{X}_i}=m_i\dot{X}_i$. Thus, this gives $-\frac{\partial U}{\partial X_i}=m_i\ddot{X}_i$, which is (1). ⊠

Thus, if one takes two points in the configuration space, understanding how the system moves from one to the other is in a way that "minimizes" (actually extremizes, though this is usually minimizing) this action functional. This is powerful because it is independent of coordinates.

A (Lagrangian) mechanical system where $L=T-U$, T is a quadratic form depending on velocity and U depends on $x_1,\dots,x_n$ is called a natural system. However, there are important systems for which L doesn't take this form.

If a system has holonomic constraints (such that the configuration space is a submanifold Σ of $\mathbb{R}^{3N}$), it can be thought of in two ways. The first is to think of it as a very powerful force pushing everything to Σ (so the configuration space is thought of as $\mathbb{R}^{3N}$). For example, if U is the potential function, one can make a new potential function $\tilde{U}(\mathbf{x})=U(\mathbf{x})+N\operatorname{dist}_\Sigma^2$ (i.e. distance to the surface), where N is very large, so that the gradient points into the surface and everything is pushed onto it. In the limit $N\to\infty$, the system is actually constrained to Σ, and cannot leave (e.g. in a rigid body, there is a force holding it together). Alternatively, one could just live on Σ, since there are always local coordinates, and the Euler-Lagrange equation is coordinate-independent.

For the Hamiltonian formalism, which will be discussed later, there is even more invariance under change (and thus power), since it doesn't require the coordinates in the tangent bundle to be induced from the coordinates on the manifold.

---

[7] Technically, $S(\mathbf{x})=\int_0^1 L(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t$, which is called the action function, is the equivalence.

Suppose $M \subset \mathbb{R}^{3N}$ is a configuration space (some submanifold of $\mathbb{R}^{3N}$). Then, TM is called the phase space, and is a manifold of twice the dimension of that of M. Then, $L : TM \to \mathbb{R}$ is the Lagrangian, and $S(\mathbf{x}) = \int_0^1 L(\mathbf{x}(t))\,\mathrm{d}t$ is the action functional. Any local system of coordinates on M are called generalized coordinates, typically denoted $q_1, \ldots, q_k$. The dual coordinates $\dot{q}_1, \ldots, \dot{q}_k$ are called generalized velocity. Then, the $\frac{\partial L}{\partial \dot{q}_i}$, typically written p_i, are called the generalized momenta and $\frac{\partial L}{\partial q_i}$ are called generalized forces. Then, the Lagrangian equation looks like $\dot{p}_i = \frac{\mathrm{d}L}{\mathrm{d}\dot{q}_i}$, which in specific cases implies $p_i = m_i \dot{q}_i$, which is the definition of momentum in elementary physics classes, so this formalism is fairly powerful. If L is independent of q_i, then the generalized momenta are preserved (in which case the coordinates are called cyclic). In general, if L is independent of some quantity, a conservation law emerges.

3. Suddenly, Riemannian Geometry: 1/14/13

In Lagrangian mechamics, one starts with an n-dimensional manifold M called the configuration space of the system, and the system is said to have n degrees of freedom. For example, one particle moving in 3-dimensional space has 3 degrees of freedom and the configuration space is $\mathbb{R}^3$. For a system of N particles moving in 3-dimensional space, there are $3N$ degrees of freedom and the configuration space is $\mathbb{R}^{3N}$.

In a constrained system, the system is restricted to a submanifold, which is still called the configuration space. For example, the motion of a rigid body is a constrained system.

The tangent bundle TM is a $2n$-dimensional manifold (and in particular is *not* a submanifold of whatever space M lies in). If $(q_1, \ldots, q_n)$ are local coordinates in an open neighborhood $U \subset M$, these coodrinates induce local coordinates on $TU \subset TM$ given by $(q_1, \ldots, q_m, \frac{\partial}{\partial q_1}, \ldots, \frac{\partial}{\partial q_n})$; the latter n coordinates represent velocity vectors. Then, every tangent vector is written $T = \sum_{i=1}^n \dot{q}_i \frac{\partial}{\partial q_i}$. Here, the dot is only notational; here it doesn't denote a derivative with respect to time.

Then, the Lagrangian formalism considers the Lagrangian function $L : TM \to \mathbb{R}$ (in an autonomous, or time-independent system) or sometimes $L : TM \times \mathbb{R} \to \mathbb{R}$ (non-autonomous, or time-dependent) $L(\mathbf{q}, \dot{\mathbf{q}})$ or $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ and a functional (i.e. a function on an infinite-dimensional space) on $\mathcal{P}_{a,b} = \{\gamma : [a, b] \to M \mid \gamma(0) = a, \gamma(1) = b\}$ (the space of paths on M from a to b, with $a, b \in M$) $S : \mathcal{P}_{a,b} \to \mathbb{R}$ and given by $S(\gamma) = \int_0^1 L(\gamma(t), \gamma'(t), t)\,\mathrm{d}t$.

This is usually written $S(\mathbf{q}) = \int_0^1 L(\mathbf{q}, \dot{\mathbf{q}}, t)\,\mathrm{d}t$, since paths and coordinates are conventionally written with the same letters. Here, $\dot{\mathbf{q}}$ means both coordinates of the velocity vector and $\frac{\mathrm{d}}{\mathrm{d}t}$, since a path p on M can be lifted into TM by taking the velocity, which is why this notation is used.

Formally, $\mathcal{P}_{a,b} \subset \mathcal{P}(M) = \{\gamma : [0, 1] \to M\}$. There is also $\mathcal{P}(TM) = \{\gamma : [0, 1] \to TM\}$. Then, there is a lifting map $\ell : \mathcal{P}(M) \to \mathcal{P}(TM)$ given by $\ell : \gamma(t) \to \frac{\mathrm{d}\gamma}{\mathrm{d}t}$ (i.e. $\gamma'(t)$).[8] Thus, if $\Gamma(T) = \ell(\gamma(t))$, then Γ is on TM, so one can take $\int_0^1 L(\Gamma(t))\,\mathrm{d}t$.

We also have the least action principle of Hamilton (i.e. Theorem 2.2): if a system travels from $a \in M$ to $b \in M$ over time, then the path it takes is an extremum for the action functional S (usually a minimum, but not always).

From the calculus of variations (and the previous lecture), the extrema of the action functional satisfy the Euler-Lagrange Equation (Theorem 2.1; called such in the calculus of variations, but just called the Lagrange equation in mechanics). This is a 2^{nd}-order differential equation.

In general systems L can be fairly arbitrary. but in many systems arising from Newtonian mechanics, L takes a special form. These systems are called natural (mechanical) systems, and will include most of the examples in this class. For example, if there are N particles moving in three-dimensional space, the configuration space is $\mathbb{R}^{3N}$ with natural coordinates $(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \mathbf{q}$, where each $\mathbf{r}_i = (x_i, y_i, z_i)$ is the path of one of the particles. Then, if particle i has mass m_i, there is a quadratic function $T : TM \to \mathbb{R}$ called the kinetic energy and another function $U(\mathbf{q})$ called the potential energy of the system, which depends only on the position. Since every function on M can be lifted to TM, then U can be thought of as a function on TM, so it makes sense to define the Lagrangian as $L = T - u$.

From the Euler-Lagrange equation,

$$-\frac{\partial U}{\partial \mathbf{r}_i} = \frac{\mathrm{d}}{\mathrm{d}t}(m_i \dot{\mathbf{r}}_i) \implies m_i \ddot{\mathbf{r}}_i = -\frac{\partial U}{\partial \mathbf{r}_i}.$$

Thus, for natural systems, the Lagrange equation is the same as the Newton equations.

[8] There is a canonical projection $\mathcal{P}(TM) \to \mathcal{P}(M)$, but lifting in the reverse direction requires a choice of coordinates.

This also holds true if $M \subsetneq \mathbb{R}^{3N}$ (i.e. the system is constrained), which can be seen by modelling the constraint as a very strong force onto M. Then, T can be thought of as T on $\mathbb{R}^{3N}$ restricted to M. In Riemannian geometry, this is an example of an induced Riemannian metric.

Definition. Let M be a manifold and $g : TM \to \mathbb{R}$ such that g is a positive definite quadratic form on (that is, when restricted to) any tangent space of TM. Then, g is a Riemannian metric.

A Riemannian metric induces a scalar product on each tangent space, so there is Euclidean geometry on every tangent space, which allows the calculation of (for example) arc length of curves on M. This leads to the subject known as Riemannian geometry.

The point of all this is that, while every manifold can be thought of as a submanifold of Euclidean space, this approach doesn't require that and is thus more convenient; in a sense, it is non-exploratory.

For example, on the surface of Earth (approximately S^2), one can introduce local coordinates at any point and measure things. Then, if $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$, then their scalar product will be some positive definite quadratic form $\sum_{i,j} a_{ij} u_i v_j$ — but the coefficients a_{ij} do not need to be constant, and their variation describes the geometry of the world.

If $q_1, \dots, q_n$ are coordinate functions, they induce the differential 1-forms $\mathrm{d}q_1, \dots, \mathrm{d}q_n$ ($\mathrm{d}q_i$ is the i^{th} coordinate function for the linearization of a function at a given point). Then, each quadratic form can be written as $\sum a_{ij}\, \mathrm{d}q_i \cdot \mathrm{d}q_j$ (here, the dot is just scalar product).[9] This is how the Riemannian metric is usually written in local coordinates.

Thus, for a system constrained to M, the kinetic energy function T can be some Riemannian metric; in particular, the coordinates for the quadratic form needn't be constant. Then, U is some (usually positive definite) function, and $L = T - U$.

If the system has no external forces (i.e. $U = 0$), then

$$\int_0^1 L(\mathbf{q}, \dot{\mathbf{q}})\, \mathrm{d}t = \int_0^1 T(\mathbf{q}, \dot{\mathbf{q}})\, \mathrm{d}t = \frac{1}{2} \int_0^1 \|\dot{\mathbf{q}}(t)\|^2\, \mathrm{d}t.$$

This latter quantity is called the energy or energy functional in Riemannian geometry.

It is a lemma in Riemannian geometry that minimizers of energy are also minimizers of length (intuitively, one of these minimizes $\|\dot{\mathbf{q}}\|$ and the other minimizes $\|\dot{\mathbf{q}}\|^2$). These minimizing curves are called geodesics.

Thus, Riemannian geometry is just a special case of Lagrangian mechanics for systems with holonomic constraints. Also, note that masses were largely absent from the above discussion, since they were absorbed into the Riemannian metric (and therefore they actually curve space, in some sense).

Now for some terminology: if some term exists in the Newton schema $\frac{\mathrm{d}}{\mathrm{d}t}(m_i \dot{\mathbf{r}}_i) = -\frac{\partial U}{\partial \mathbf{r}_i}$, then the generalized version of that term exists for the Lagrange schema $\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}$. Thus,

- If $q_1, \dots, q_n$ are local coordinates on M, they are called generalized coordinates (which doesn't add any more meaning).
- In the Newtonian schema, the $-\frac{\partial U}{\partial \mathbf{r}_i}$ are called forces, so the $\frac{\partial L}{\partial q_i} = F_i$ are called generalized forces.
- The $m_i \dot{\mathbf{r}}_i$ are called momenta, so the $p_i = \frac{\partial L}{\partial \dot{q}_i}$ are called generalized momenta.

The Newton form of these equations is $\frac{\mathrm{d}}{\mathrm{d}t}(p_i) = F_i$.

If $\frac{\partial L}{\partial q_i} = 0$ (in which case the potential and kinetic energies are independent of the coordinate q_i), the coordinate q_i is caled cyclic, and some quantity is preserved. In differential equations, this yields an integral of motion, which allows one to reduce the order of the equation, making it easier to solve.

In the case where the Lagrangian is time-independent (autonomous), there is also a conservation law:

Theorem 3.1 (Conservation of Energy)**.** *In an autonomous system,*[10] *define the energy to be*

$$H = \mathbf{p}\dot{\mathbf{q}} - L = \frac{\partial L}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}} = \sum \frac{\partial L}{\partial q_i} \dot{q}_i.$$

Then, H is preserved; it is an integral of motion, so it is constant for every trajectory.

Proof. Trivial; differentiate.

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(L - \sum \frac{\partial L}{\partial \dot{q}_i} \cdot \dot{q}_i \right) = \frac{\partial L}{\partial \mathbf{q}} \dot{\mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}} \frac{\mathrm{d}(\dot{\mathbf{q}})}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \mathbf{q}} \right) \dot{\mathbf{q}} - \frac{\partial L}{\partial \dot{\mathbf{q}}} \frac{\mathrm{d}(\dot{\mathbf{q}})}{\mathrm{d}t}.$$

Then, the first and third terms cancel out because of the Lagrange equation. ⊠

[9] Observe that this isn't a bilinear form; that would be written something like $\sum a_{ij}\, \mathrm{d}q_i \otimes \mathrm{d}q_j$.

[10] This is a system where L doesn't depend explicitly on t. There is an implicit dependence in the trajectories given by L.

H is also called the Hamiltonian, but not yet.

Theorem 3.2 (Euler equation for homogeneous systems)**.** *Suppose that $L = T - U$ is a natural system: $T = \sum a_{ij}\dot{q}_i\dot{q}_j$ and U is a function of $\mathbf{q}$. Then,*

$$\frac{\partial L}{\partial \dot{\mathbf{q}}}\dot{\mathbf{q}} = \sum a_{ij}\dot{q}_i\dot{q}_j = 2T.$$

The coefficient appears because in the calculation of the sum of squares, the symmetry of the equation ensures every term appears twice.

Thus, $L - \dot{\mathbf{q}}\frac{\partial L}{\partial \dot{\mathbf{q}}} = -T - U$.

Consider the motion of a particle in a field, so that $L = \frac{m|\dot{r}|^2}{2} - U(r)$, so that the equations of motion are the standard Newton equations: $m\ddot{r}_i = -\frac{\partial U}{\partial r}$. Suppose this field is symmetric to some axis; introduce cylindrical coordinates (r, φ, z), where the axis of symmetry is in the z-direction.

Thus, $\dot{\mathbf{r}}^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = \dot{r}^2 + r^2\dot{\varphi}^2 + \dot{z}^2$.[11] Then, φ is cyclic (this is actually the source of the word), and the momentum $p_\varphi - = \frac{\partial L}{\partial \dot{\varphi}} = r^2\dot{\varphi}$ is conserved. This is conservation of angular momentum about the z-axis; specifically, the quantity $r^2\dot{\varphi} = |\mathbf{r} \times \dot{\mathbf{r}}|$ is preserved, which is the area of the parallelogram spanned by $\mathbf{r}$ and $\dot{\mathbf{r}}$.

4. Examples of Natural Systems: 1/16/13

First consider a system with 1 degree of freedom, so that the configuration space is either a line or a circle. Then, $L = \dot{x}^2/2 - U(x)$ and $\ddot{x} = -U'(x)$. This system is always integrable and therefore easy to solve: the energy is $E = \dot{x}^2/2 + U(x)$, which is conserved, so $\dot{x} = \pm\sqrt{2(E - U(x))}$.

To solve this equation, think of it as a Pfaffian equation (which gives a line field, since this system has one degree of freedom), given by $0 = \mathrm{d}x - \sqrt{2(E - U(x))}\,\mathrm{d}t$. Then, a solution can be obtained by rearranging to

$$\mathrm{d}t = \frac{\mathrm{d}x}{\sqrt{2(E - U(x))}}$$

and integrating. The general goal is to reduce the order of the differential equation with an integral of motion.

Now consider a particle moving in a central field in 3-space. Thus, the configuration space is $\mathbb{R}^3$. if $\mathbf{x} \in \mathbb{R}^3$ is the path of the system (in this case, just the path of the particle), then the kinetic energy is $T = m|\dot{\mathbf{x}}|^2/2$ and the Lagrangian is

$$L = \frac{m|\dot{\mathbf{x}}|^2}{2} - U(|x|),$$

since the potential function of a central field depends only on distance.

This system is spherically symmetric, so it's cylindrically symmetric around any axis. Thus, angular momentum is conserved around any axis. If $(\mathbf{e}_r, \mathbf{e}_\varphi)$ is a basis for polar coordinates and $\mathbf{x} = r\mathbf{e}_r$, then $\dot{\mathbf{x}} = \dot{r}\mathbf{e}_r + r\dot{\varphi}\mathbf{e}_\varphi$, so $|\mathbf{x} \times \dot{\mathbf{x}}| = r^2\dot{\varphi}$. Thus, if the vector $\mathbf{x} \times \dot{\mathbf{x}}$ is projected to the plane orthogonal to any axis, it is conserved. This quantity is called the vector of angular momentum, $\mathbf{M} = \mathbf{x} \times \dot{\mathbf{x}}$. Since $\mathbf{M}$ is perpendicular to $\mathbf{x}$ and $\dot{\mathbf{x}}$, then it is perpendicular to any motion, so motion in a central field is planar.

Corollary 4.1. *If such a system is restricted to a sphere, then it must move in geodesics. But in a central field, it must also move in a plane. Thus, the geodesics on a sphere are great circles.*

It's possible to prove this more formally by introducing coordinates, but it's not nearly as clean.

Thus, everything can be done in the plane, treating $\mathbf{x}$ as a plane vector. Additionally, one can make the simplifying assumption that $m = 1$. Then, in polar coordinates, the Lagrangian is

$$L = \frac{\dot{r}^2 + r^2\dot{\varphi}^2}{2} - U(r),$$

since $|\dot{\mathbf{x}}|^2 = \dot{r}^2 + r^2\dot{\varphi}^2$, and $M = r^2\dot{\varphi}$ and $E = (\dot{r}^2 + r\dot{\varphi}^2)/2 + U(r)$ are integrals (and therefore constant). Thus, $\dot{\varphi} = M/r^2$ can be removed from the equation. Thus, $\dot{\varphi}^2 = M^2/r^4$ and $E = \dot{r}^2/4 + M^2/2r^2 + U(r)$, and L can be rewritten similarly as

$$L = \frac{\dot{r}^2}{2} - \left(U(r) - \frac{M^2}{2r^2}\right).$$

But this is just a system with one degree of freedom and a potential $V(r) = U(r) - M^2/2r^2$. This latter quantity is called effective potential energy.

[11] Here, $\mathbf{r}$ represents the path as a vector-valued function, and r represents its r-coordinate (cylindrical).

Solving this new system, one obtains $\dot{r} = \sqrt{2(E - V(r))}$, which can be rewritten in Pfaffian form as

$$\mathrm{d}t = \frac{\mathrm{d}r}{\sqrt{2(E - V(r))}}.$$

Thus, motion is given by

$$\int_{t_0}^{t_1} \mathrm{d}t = \int_{r_0}^{r_1} \frac{\mathrm{d}r}{\sqrt{2(E - V(r))}},$$

where $r_0 = r(t_0)$ and $r_1 = r(t_1)$. Additionally, it is possible to combine these equations and write

$$\frac{\mathrm{d}\varphi}{\mathrm{d}r} = \frac{M}{r^2\sqrt{2(E - V(r))}}. \tag{2}$$

After integrating, φ can be found as a function of r.

Since $r^2/2 + V(r) = E$, then $V(r) \leq E$, and at a minimum point (in which E is equal to a minimum of $V(r)$), r is constant, which means that the particle moves in a circular orbit. If the constraint is between two values $r_{\min}$ and $r_{\max}$, then the particle orbits between the two circles given by $r_{\min}$ and $r_{\max}$. The minimum value is called the pericenter, and the maximum is the apocenter.[12]

Computing the time needed to go from $r_{\min}$ to $r_{\max}$ is as easy as calculating

$$t = \int_{t_{\min}}^{t_{\max}} \mathrm{d}t = \int_{r_{\min}}^{r_{\max}} \frac{\mathrm{d}r}{\sqrt{2(E - V(r))}},$$

and computing the angle between the pericenter and apocenter just requires plugging this into (2). Note that this angle doesn't change as the trajectory evolves. In particular, this angle $\Delta\varphi$ is either:

- commensurate with 2π, in which case the orbit is periodic, or
- not, in case the orbit is everywhere dense between $r_{\max}$ and $r_{\min}$.

Suppose the original potential energy is $U(r) = r^\alpha$. Then, there are exactly 2 values of α for which all solutions are periodic: $\alpha = -1$, which is the Kepler solution (all orbits are ellipses), and $\alpha = 2$. This is pretty hard to show.

Consider the Kepler case, where $U(r) = -k/r$, so that $V = -k/r - M^2/2r^2$ and

$$\varphi = \int \frac{\frac{M}{r^2}\,\mathrm{d}r}{\sqrt{2(E - V(r))}} = \int \frac{\frac{M}{r^2}\,\mathrm{d}r}{\sqrt{2\left(E + \frac{k}{r} - \frac{M^2}{2r^2}\right)}}.$$

Let $u = 1/r$, so that

$$\begin{aligned} &= -\int \frac{M\,\mathrm{d}u}{\sqrt{2E + 2ku - M^2u^2}} \\ &= -\int \frac{\mathrm{d}u}{\sqrt{\frac{2E}{M^2} + \frac{2ku}{M^2} - u^2}}. \end{aligned}$$

Letting $t = u - k/M^2$, this equals

$$= -\int \frac{\mathrm{d}t}{\sqrt{\frac{2E}{M^2} + \frac{k^2}{M^4} - t^2}}.$$

Let $a = 2E/M^2 + k^2/M^4$.

$$\begin{aligned} &= \arccos\left(\frac{t}{a}\right) \\ &= \arccos\left(\frac{\frac{1}{r} - \frac{k}{M^2}}{\sqrt{\frac{2E}{M^2} + \frac{k^2}{M^4}}}\right), \end{aligned}$$

up to some constant, which is just choosing the initial point of the pericenter.

From this, one can express r through φ (or through $\cos\varphi$), obtaining $r = p/(1 + e\cos\varphi)$, where $p = M^2/k$ and $e = \sqrt{1 + 2EM^2/k^2}$ is the eccentricity. This is the polar form of a conic section; for example, if $2c$ is the distance between the two foci of an ellipse and a is the semimajor axis, then $e = c/a$ (so that $e < 1$ for an ellipse and $e = 0$

[12] In celestial mechanics, these are called the perigee and apogee, respectively.

for a circle, corresponding to negative energy; $e = 1$ for a parabola, corresponding to zero energy; and $e > 1$ for a hyperbola, corresponding to positive energy).

Consider an ellipse $x^2/a^2 + y^2/b^2 = 1$ and let $c = a^2 - b^2$. Then, $b = a\sqrt{1-e^2} = a(1 - e^2/2 + \dots)$, so the difference in the axes is in the second order of e, so an ellipse with small e is almost a circle.

Consider a satellite that is supposed to to orbit at a 300 km orbit, but was fired incorrectly, with a 1° deviation towards Earth. Then, what is its perigee (or rather, what is the order of magnitude of the mistake)? It's an ugly computation, but since the mistake is very small and the velocity doesn't change, then the orbit is still roughly circular, displaced by 1°. Thus, the displacement of this circle is equal to the distance between these two circles' centers, which is $6.7 \cdot 10^7 \cdot 2\pi/360 \approx 110\,\text{km}$, which is actually quite a lot! (This is because $r_\oplus \approx 6.4 \cdot 10^7\,\text{m}$.)

For a completely different example, consider a system whose configuration space is the upper-half plane $H = \mathbb{R}^2_+ = \{y > 0\}$ with Riemannian metric (kinetic energy)

$$\mathrm{d}s^2 = \frac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2}.$$

Here, $\mathrm{d}s$ measures the arc length, and the standard Euclidean metric is $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$. This looks similar, but this system's kinetic energy blows up as $y \to 0$. This object is called the Lobachevsky hyperbolic plane, and is very imortant in mathematics. The whole object can't be realized in $\mathbb{R}^3$ with the usual Euclidean metric, but parts of it can be.

If $T = (\dot{x}^2 + \dot{y}^2)/y^2$ and there are no forces, computing paths is possible (albeit slightly ugly). $L = T$ is independent of x, so $\frac{\partial L}{\partial \dot{x}} = \dot{x}/y^2$ is conserved. Thus, $\dot{x} = cy^2$. Additionally, since energy is conserved, then $Ey^2 = \dot{x}^2 + \dot{y}^2$, so $\dot{y}^2 = Ey^2 - c^2y^4$.

There are lots of isometries of this metric: the easiest example is translation along the x-axis, but another example is homothety ($(x, y) \mapsto (cx, cy)$). A more interesting isometry is called inversion, where $(x, y) \mapsto (x', y')$, such that $|(x, y)||(x', y')| = 1$. This is given by

$$(x, y) \mapsto \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2} \right),$$

and plugging into the axioms verifies that it preserves distances and lengths. Finally, reflection across any vertical line is an isometry, since the metric is independent of x.

Thus, vertical lines are geodesics, and since these isometries can map any point to any point and any direction to any direction, then all geodesics are the image of a vertical line under isometries. In particular, under inversion, one obtains semicircles around the x-axis. These vertical lines (which can be thought of as semicircles of infinite radius) and semicircles must be all possible isometries, since there is one for each point and direction.

Now consider $H \subset \mathbb{C}$, given by $\{\operatorname{Im} z > 0\}$. Then, these transformations are called M obius transformations, or conformal transformations, and are linear transformations of the form $\frac{az+b}{cz+d}$, with $a, b, c, d \in \mathbb{R}$ (up to coprime-ness). This set of transformations forms a group called $\mathrm{PSL}_2(\mathbb{R})$. The specific connections are: $1/z$ is an inversion composed with a reflection, az is scaling, and $z + b$ is translation. There is a nuance in that reflection is orientation-reversing, which doesn't play too well with hyperbolic geometry, but the important point is that all of the isometries of H preserve the complex structure. Thus, all surfaces can have a hyperbolic metric, and the equivalence between hyperbolic geometry and complex analysis is one of the most fundamental things in mathematics. Specifically, a theorem in the 19th century showed this could only hold for three metrics: the Euclidean metric, the spherical metric, and this hyperbolic metric (also sometimes called the metric of constant curvature -1).

5. More on the Kepler Problem: 1/23/13

Consider once again a particle moving in a central field; the sign errors that may lie in last week's lecture will be fixed. Its motion is planar, so one can consider motion in polar coordinates in this plane. Then, its velocity is $v^2 = \dot{r}^2 + r^2\dot{\varphi}^2$ and the Lagrange function is $L = \frac{1}{2}(\dot{r}^2 + r\dot{\varphi}^2) - U(r)$. Since this is independent of φ (i.e. φ is cyclic), then the angular momentum $\frac{\partial L}{\partial \dot{\varphi}} = r^2\dot{\varphi} = M$ is conserved. This allows the order of the equation to be reduced:

$$\begin{aligned}
\frac{\partial L}{\partial t} &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{r}} \\
\implies \ddot{r} &= -U'(r) + r\dot{\varphi}^2 \\
&= -U'(r) + \frac{M^2}{r^3} \\
&= -U'(r) + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{-M^2}{2r^2}\right).
\end{aligned}$$

Then, $V = M^2/2r^2 + U(r)$, so $\ddot{r} = -V'(r)$, which is a Newton equation with one degree of freedom. Then, the rest of the previous argument follows (this time with the right signs).[13]

If $X \subset M$ is a submanifold and a critical point on the space of paths of M lies in the space of paths of X, then it is also a point there. But the converse isn't necessarily true, so lifting from the plane to its tangent bundle doesn't quite work; if one takes a variation of r and adds some variation of φ or $\dot{\varphi}$, it isn't necessarily a lift.[14]

Returning to the Kepler problem, in which $U(r) = -k/r$ and $V(r) = -k/r + M^2/2r^2$, then $\ddot{r} = -V'(r)$ and there is an energy integral $E = V(r) + \dot{r}^2/2$. Thus,

$$\dot{r} = \sqrt{2(E - V(r))} \implies \mathrm{d}t = \frac{\mathrm{d}r}{\sqrt{2(E - V(r))}}.$$

Additionally, $r^2\dot{\varphi} = M$, so $\dot{\varphi} = M/r^2$. Thus,

$$\mathrm{d}\varphi = \frac{M}{r^2}\,\mathrm{d}t \implies \varphi = \int \frac{M\,\mathrm{d}r}{r^2\sqrt{2(E - V(r))}}.$$

In particular, the quadratic polynomial under the square root is a standard integral leading to an arcsin function. Solving it leads to

$$r = \frac{p}{1 + e\cos\varphi} \text{ where } p = \frac{M^2}{k} \text{ and } e = \sqrt{1 + \frac{2EM^2}{k^2}}. \tag{3}$$

These are equations for conic sections in polar coordinates.

Consider an ellipse given by the equation $x^2/a^2 + y^2/b^2 = 1$. Here, $c^2 = a^2 - b^2$ and $e = c/a$. if $e = 0$, it is a circle, and if $e > 1$, one obtains a hyperbola. This can be plugged ito (3) to determine how to obtain specific orbits (e.g. a parabola requires $E = 0$).

Suppose the orbit is periodic (so that $e < 1$). Then, $M = r^2\dot{\varphi}$ is twice the sectorial velocity (since the parallelogram has twice the area of the triangle) and the sectorial velocity is $M/2$. In partcular, it's constant (hello Kepler). If T is the period, then $T(M/2)$ is the area of the ellipse, so

$$\frac{TM}{2} = \pi ab \implies T = \frac{2\pi ab}{M}.$$

Thus,

$$a = \frac{p}{1 - e^2} = \frac{M^2}{2|E|M^2} = \frac{k}{2|E|} \text{since } E < 0, \text{ and similarly}$$

$$b = \sqrt{a^2 - c^2} = a\sqrt{1 - e^2} = \frac{k}{2|E|}\sqrt{\frac{2|E|M^2}{k^2}} = \frac{M}{\sqrt{2|E|}}$$

$$\implies T = \frac{2\pi ab}{M} = \frac{2\pi}{M}\left(\frac{k}{2|E|}\right)\left(\frac{M}{\sqrt{2|E|}}\right) = \frac{\pi k}{\sqrt{2}E^{3/2}}.$$

Since $2E = k/a$, then this becomes $T = \pi a^{3/2}/\sqrt{2k}$, which yields another of Kepler's laws. This is a very rich subject with lots to explore.

Consider a Russian cosmonaut on a spacewalk who takes a camera out to take a picture and finds he has nowhere to hold his lens cap.[15] He decides to throw his lens cap towards the Earth, but what happens? It turns out he will be hit by it from above.

Exercise 5.1. Prove this.

The best way to approach this is to avoid all of the ugly theory, but consider the linearization of this system: $\ddot{\mathbf{r}} = -\mathbf{r}/r^3$, up to units (so pick nice units which make this happen). Writing this in polar coordinates and writing variations, the system starts with $\dot{r}_0 = \dot{\varphi}_0 = 1$ and is perturbed to obtain $r = r_0 + r_1$ and $\varphi = \varphi_0 = \varphi_1$ (with $r_1 \ll r_0$). leaving only the linear term (not even constants) gives a mostly accurate description of the orbit and a system of linear equations that can be solved.

Suppose a system has some symmetry, such as a 1-parametric group of transformations that preserves the Lagrange function. For example, if φ is a cyclic coordinate, then $\frac{\partial}{\partial\varphi}$ is a vector field that preserves the Lagrangian. Such a vector field is called a symmetry. In this case, one can change coordinates so that the flow is in one of these directions and obtain an integral. But this is somewhat inconvenient, so a simpler solution was found by Noether.[16] First, though, a definition:

[13]The following assertions were wrong: that $L = 1/2(\dot{r}^2 + M^2/r^2) - U'(r) = \dot{r}^2/2 - V(r)$, where $V(r) = U(r) - M^2/2r^2$.

[14]To make life more confusing, this *does* work in the Hamiltonian case, since everything is set up differently.

[15]This is supposedly a true story.

[16]Noether was a great mathematician, but Theorem 5.1 is kind of trivial and not the entire reason.

Definition. A one-parametric group of transformations is a family of transformations $g^t : M \to M$ ($t \in \mathbb{R}$) such that $g^{t+t'} = g^t \circ g^{t'}$ (which implies a group structure).

This induces a vector field $\left.\frac{\mathrm{d}g^t(x)}{\mathrm{d}t}\right|_{g^t(x)} = \mathbf{v}$, leading to a local flow. It also induces another group $G^t : TM \to TM$, since the differential of g^t operates on tangent spaces.

If $g^t : M \to M$ is a symmetry of the Lagrangian system on M (i.e. $L(G^t(x)) = L(x)$ or $\frac{\mathrm{d}(L(G^t(x)))}{\mathrm{d}t} = 0$), then $\mathrm{d}L(\mathbf{v}) = 0$ as well. Thus, it becomes necessary to compute $\frac{\mathrm{d}G^t()}{\mathrm{d}t}$ as well (which is, oddly enough, tangent to the tangent bundle). The final result is

$$0 = \frac{\mathrm{d}(L(G^t(x)))}{\mathrm{d}t} = \mathrm{d}L(\mathbf{v}) = \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \mathbf{v} = \sum \frac{\partial L}{\partial \dot{q}_i} \cdot v_i,$$

assuming some coordinates $L(\mathbf{q}, \dot{\mathbf{q}}, t)$. The flow is $q_i \mapsto q_i + t$, which lifts upstairs (and the dot coordinates don't change at all).

Theorem 5.1 (Noether)**.** *Suppose the Lagrangian system $L(\mathbf{q}, \dot{\mathbf{q}})$ admits a 1-parametric group of symmetries generated by a vector field $\mathbf{v}$. Then, $I = \sum \frac{\partial L}{\partial \dot{q}_i} v_i$ is an integral of motion.*

Proof. First consider the special case $\mathbf{v} = \frac{\partial}{\partial q_1}$, so that $\mathbf{v}$ is a symmetry iff q_1 is cyclic. Then, $\sum \frac{\partial L}{\partial q_i} v_i = \frac{\partial L}{\partial \dot{q}_1}$ (i.e. the generalized momentum corresponding to q_1), so $\frac{\mathrm{d}I}{\mathrm{d}t} = 0$.

The generalized case is the same, since one can always choose coordinates in which $\mathbf{v} = \frac{\partial}{\partial q_1}$. ⊠

Consider a system of points with masses $m_1, \dots, m_k$ and coordinates $\mathbf{r}_1, \dots, \mathbf{r}_k \in \mathbb{R}^3$ moving in a potential field U that depends only on the distances $\|\mathbf{r}_i - \mathbf{r}_j\|$. Then, the momentum $\sum m_i \dot{\mathbf{r}}_i$ is an integral by Noether's Theorem: the Lagrangian is $L = \sum m_i \|\mathbf{r}_i\|^2/2 - U$, so L is independent by addition of any other vector (which is just translation).

Choose a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathbb{R}^3$; then, the vector field corresponding to $\mathbf{r}_i \mapsto \mathbf{r}_i + \mathbf{e}_1$ is $\mathbf{v} = (1, 0, 0, 1, 0, 0, \dots, 1, 0, 0)$. Thus,

$$L = \sum \frac{m_0(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2)}{2} - U,$$

so the Noether integral is $\sum \frac{\partial L}{\partial \dot{x}_i} = \sum m_i \dot{x}_i$. Thus, $\sum m_i \ddot{x}_i = 0$ since the integral is constant.

In some sense, preservation of momentum is a corollary of Noether's Theorem. Similarly, angular momentum, $\sum m_i(\mathbf{r}_i \times \dot{\mathbf{r}}_i)$ is preserved in the case of rotational symmetry.

6. Rigid Bodies I: 1/28/13

The sutdy of motion of rigid bodies was a major focus of mathematics in the 19th Century. Rigid bodies can be analyzed with the Lagrangian or Hamiltonian formalism; this lecture (as with the book) does the former.

The configuration space of a rigid body is $M = \mathbb{R}^3 \times \mathrm{SO}(3)$, corresponding to a rotational frame frozen in the body (whose orientation determines an orthogonal matrix in SO(3)) and a point that moves in $\mathbb{R}^3$.

Consider a rigid body in the absence of external forces, so that $L = T$, and approximate the rigid body as a set of i points with locations $\mathbf{q}_i$ and masses m_i, so that $T = \sum \frac{m_i \|\dot{\mathbf{q}}_i\|^2}{2}$. If instead the density function is given by a density distribution μ, then this becomes $T = 1/2 \int \mu \|\dot{\mathbf{q}}\|^2 \, \mathrm{d}t$. Since momentum is conserved, then $\sum m_i \dot{\mathbf{q}}_i$ (or the corresponding integral) is constant. Equivalently, the center of mass, $\mathbf{q}_C = \sum m_i \dot{\mathbf{q}}_i / \sum m_i$, moves with a constant velocity. Sometimes, the center of mass is fixed at the origin, so the configuration space is just SO(3). Thus, a rigid body in the absence of external forces moves as if fixed at a point.

Consider a rigid body fixed at a point (not necessarily the center of mass), so that the configuration space is SO(3). Pick some coordinates for $\mathbb{R}^3$, and denote this space k, and let K be some coordinate system fixed to the rigid body. Then, vectors in K will be denoted by uppercare letters, and those in k will be denoted by lowercase letters (the two spaces are the same, but with different coordinates, and a map relates them).

If $\mathbf{q} \in k$, then there is a corresponding $\mathbf{Q} \in K$ such that $\mathbf{q} = B(t)\mathbf{Q}$ for some orthogonal (but time-dependent) operator $B(t)$, the change-of-coordinates operator. Then, $\dot{\mathbf{q}} = \dot{B}\mathbf{Q} + B\dot{\mathbf{Q}}$ and $\dot{B}\mathbf{Q} = \dot{B}B^{-1}\mathbf{q}$. Then, $\dot{B}B^{-1}$ is an anti-self-adjoint (i.e. skew-symmetric) operator:

$$\dot{B}B^{-1} = (\dot{B}B^{\mathrm{T}})^{\mathrm{T}} = \dot{B}^{\mathrm{T}}B = -B^{\mathrm{T}}\dot{B} = -B^{-1}\dot{B},$$

since differentiation and taking a transpose commute.[17] And a skew-symmetric operator has a remarkable form in $\mathbb{R}^3$: if $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3) \in \mathbb{R}^3$, then the operator $\mathbf{q} \xrightarrow{C_{\boldsymbol{\omega}}} \boldsymbol{\omega} \times \mathbf{q}$ is skew-symmetric: $\langle C\mathbf{q}, \tilde{\mathbf{q}} \rangle = \langle \boldsymbol{\omega} \times \mathbf{q}, \tilde{\mathbf{q}} \rangle = \langle \mathbf{q}, \tilde{\mathbf{q}} \times \boldsymbol{\omega} \rangle = -\langle \mathbf{q}, C\tilde{\mathbf{q}} \rangle$, since things can be moved cyclically.

[17] This is a more general fact about this Lie algebra...

The dimension of the space of all $C_{\mathbf{q}}$ is 3, and the dimension of the space of skew-symmetric matrices is also 3 (since each such matrix is of the form

$$\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

which leaves only 3 degrees of freedom), and this map is linear, so it must be an isomorphism. In fact, this isomorphism can be explicitly written (up to sign):

$$(\omega_1, \omega_2, \omega_3) \mapsto \begin{pmatrix} 0 & \omega_1 & \omega_2 \\ -\omega_3 & 0 & \omega_3 \\ -\omega_2 & -\omega_1 & 0 \end{pmatrix}.$$

Thus, there exists an $\boldsymbol{\omega} \in \mathbb{R}^3$ such that $\dot{B}\mathbf{Q} = \boldsymbol{\omega} \times \mathbf{q}$. If $\boldsymbol{\omega}$ and $\mathbf{q}$ are fixed (i.e. frozen with respect to the body), then $\boldsymbol{\omega}$ is the angular velocity, representing the axis about which it rotates. In general, the axis of rotation can change, so $\boldsymbol{\omega}$ is called the instantaneous axis of rotation. Note that in the absence of external forces, the axis of rotation is periodic.

The Lagrangian of this system is independent of rotation, leading to three integrals, corresponding to angular momentum about any axis: $\sum \mu_i(\mathbf{q}_i \times \mathbf{v}_i) = \sum \mu_i(\mathbf{q}_i \times \dot{\mathbf{q}}_i)$ (or the corresponding integrals). This can be seen if the space is fixed or moving; let $\mathbf{m}$ be the angular momentum vector in absolute coordinates, and $\mathbf{M}$ be that relative to the body, so that $\mathbf{m} = B\mathbf{M}$. Given that $\dot{\mathbf{q}} = \boldsymbol{\omega}(t) \times \mathbf{q}$, one can pass $\boldsymbol{\omega}$ to the other set of coordinates as $\boldsymbol{\omega} = B\boldsymbol{\Omega}$ (whose length is also the angular velocity). Then, $\mathbf{m} = \sum \mu_i[\mathbf{q}_i, [\boldsymbol{\omega}, \mathbf{q}_i]]$ (where brackets denote the cross product), so $\mathbf{M} = B^{-1}\mathbf{m}$. Since the cross product is perserved by an orthogonal matrix, this also equals $\mathbf{M} = \sum \mu_i[\mathbf{Q}, [\boldsymbol{\Omega}, \mathbf{Q}]]$.

In particular, this entire thing depends on $\boldsymbol{\Omega}$ linearly: there is a linear operator A such that $\mathbf{M} = A(\boldsymbol{\Omega})$. A is called the operator of inertia.

Claim. A is symmetric.

Proof.

$$\langle A\mathbf{X}, \mathbf{Y}\rangle = \left\langle \sum \mu_i[\mathbf{Q}_i, [\mathbf{X}, \mathbf{Q}_i]], \mathbf{Y} \right\rangle = \sum \mu_i \left\langle [\mathbf{Q}_i, [\mathbf{X}, \mathbf{Q}_i]], \mathbf{Y} \right\rangle = \sum \mu_i \left\langle [\mathbf{X}, \mathbf{Q}_i], [\mathbf{Y}, \mathbf{Q}_i] \right\rangle,$$

which is clearly symmetric in $\mathbf{X}$ and $\mathbf{Y}$. ⊠

Thus,

$$\mathbf{M} = \langle A\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle = \sum \mu_i |\boldsymbol{\Omega} \times \mathbf{Q}_i|^2 = \sum \mu_i |\boldsymbol{\omega} \times \mathbf{q}_i|^2$$

since orthogonal operators preserve the cross product, and

$$= \sum \mu_i |\dot{\mathbf{q}}_i|^2 = 2T.$$

In particular, $T = 1/2 \langle A\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle$. Since A is symmetric, then it can be orthonormalized, so there is some et of coordinates in which this operator can be written in diagonal form: $\langle A\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle = I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2$. Consider the static body and some axis $\mathbf{e}$ (a unit vector in some direction) and let $I(\mathbf{e}) = \sum \mu_i r_i^2$, called the monent of inertia around the axis $\mathbf{e}$. Then, I_1, I_2, and I_3 are the moments of inertia with respect to the principal axes (i.e. those in which A is diagonal). This is because

$$\begin{aligned} \langle A\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle = 2T &= \sum \mu_i |\dot{\mathbf{q}}_i|^2 = \sum \mu_i |\boldsymbol{\omega} \times \mathbf{q}_i|^2 \\ &= \sum \mu_i |\boldsymbol{\omega}| r_i^2 = |\boldsymbol{\omega}| \sum \mu_i r_i^2, \end{aligned}$$

where r_i is the distance from $\mathbf{q}_i$ to the axis $\boldsymbol{\omega}$. Then, this last quantity is a moment of inertia. If one takes the moment of inertia in the first characteristic direction, one obtains the first eigenvalue, etc (obtained by plugging in an eigenvector instead of $\boldsymbol{\omega}$).

The equation $\langle A\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle = 1$ determines an ellipsoid (it also ends up being positive definite; the squares illustrate that it's nonnegative, at least) called the ellipsoid of inertia, which has semi-axes $1/\sqrt{I_i}$. If one of the $I_i \to 0$), the ellipse is very elongated.

Since $L = T = \langle A\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle$, then L depends only on this ellipsoid of inertia, so any teo rigid bodies with the same ellipsoid of inertia move in the same way.

Exercise 6.1. Show that an ellipsoid with uniform density is its own ellipsoid of inertia (so that one can replace any body with its ellipsoid).

The 3-dimensional system (a 2^{nd}-order equation) has 4 integrals: energy and the three angular momenta. But on the phase space, this reduces to 2 1^{st}-order equations (adding an extra coordinate), and such a differential equation is always a vector field in $\mathbb{R}^3$. These integrals correspond to hypersurfaces (oh which all the trajectories must lie), so the trajectories are on their intersection, which is a 2-dimensional manifold assuming the hyperplanes are all transverse.

In partiular, the vector field cannot have any zeros, so this manifold must be a torus (the only orientable manifold for which nice vector fields may be entirely nonzero). Motion on a torus must be quasiperiodic: $\dot{\varphi}_1 = c_1$ and $\dot{\varphi}_2 = c_2$ (though the rotation may not be commensurate), which illustrates that this problem can be integrated explicitly. However, the general problem of a rigid body in a central field can't be integrated, since two of the momentum integrals are lost.

Since $\mathbf{m} = B\mathbf{M}$ and $\mathbf{m}$ is constant, then $\mathbf{M}$ might be moving! Plugging it into the Euler equation yields $\dot{\mathbf{M}} = [\mathbf{M}, \mathbf{\Omega}]$ (which can be viewed in terms of $\mathbf{M}$ or $\mathbf{\Omega}$, since $\mathbf{M} = A\mathbf{\Omega}$), which can be solved to find $\mathbf{\Omega}$. This indicates how the axis of rotation changes; then, the motion of the system is given by rotation arouhd the changing axis.

Here is the reason that $\dot{\mathbf{M}} = [\mathbf{M}, \mathbf{\Omega}]$: $\dot{\mathbf{m}} = 0$, so

$$B\mathbf{\Omega} \times B\mathbf{M} = \boldsymbol{\omega} \times \mathbf{m} = \dot{B}\mathbf{M} + B\dot{\mathbf{M}} = B(\mathbf{\Omega} \times \mathbf{M}) = B(\mathbf{\Omega} \times \mathbf{M}) + B\dot{\mathbf{M}}.$$

To solve this equation, move to the principal axes: $\mathbf{M} = (M_1, M_2, M_3)$, $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$, and $A\mathbf{\Omega} = (I_1\Omega_1, I_2\Omega_2, I_3\Omega_3)$. Thus, $\mathbf{\Omega} = (M_1/I_1, M_2/I_2, M_3/I_3)$ and $\dot{M}_1 = M_2M_3(1/I_3 - 1/I_2) + \ldots$, so $M_1^2/I_1 + M_2^2/I - 2 + M_3^2/I_3 = 2E$ (the energy expressed through $\mathbf{m}$); this quantity is constant. The quantity $\|\mathbf{M}\| = \|\mathbf{m}\| = \sqrt{M_1^2 + M_2^2 + M_3^2}$ is also constant (thought $\mathbf{M}$ is not).

Thus, the trajectory is the intersection of this sphere and this ellipsoid, yielding three special cases:

(1) When the radius of the sphere is equal to the semiminor axis, the intersection is a point, so $\mathbf{M}$ is fixed, and the motion is rotation about the fixed axis.

(2 and 3) The same for the other two axes. The smallest axis is Lyapunov stable, but the other is unstable, and a displacement will push it away.

This is how motion looks to an observer on the rigid body.

Applying B, to step outside, the solution moves smoothly around the ellipsoid, causing rotation about the axis, which moves.

7. Rigid Bodies II: 1/30/13

Recall that a rigid body with one fixed point has configuration space SO(3). This is a system with 3 degrees of freedome, since SO(3) is a 3-dimensional manifold.

Each rotation is given by an axis and an angle; if this rotation is by an angle of less than π, then it can unambiguously be represented by a vector, where the magnitude represents the angle and the axis is given by the direction. But if the angle is π, there are two such possible points, causing an ambiguity. Topologically, this is fixed by "gluing" the opposite points together, and SO(3) is topologically equivalent to 3-dimensional real projective space. This is also the space of lines through the origin in $\mathbb{R}^4$ (the space of one-dimensional subspaces), or the 3-sphere with opposite points identified.

Studying the topological properties of a system tends to lead to qualitative physical properties. For example, SO(3) isn't simply connected, so there exist loops which cannot be shrunk to a point. Of these non-contractible loops, take the one with the smallest length; this loop is a geodesic. Thus, there is periodic motion!

Consider once again a system of coordinates k fixed to a body of motion, where K is the external system of coordinates, and use capital and lowercase letters as before: if $\mathbf{x} \in k$, let $\mathbf{X} = B(t)\mathbf{x} \in K$. Motion is instantaneously rotation about some axis, but that axis may change with respect to time; denote this axis by $\boldsymbol{\omega}$ or $\mathbf{\Omega}$, and denote the angular monentum as $\mathbf{m}$ or $\mathbf{M}$ (so that $\boldsymbol{\omega} = B\mathbf{\Omega}$, and $\mathbf{m} = B\mathbf{M}$). $\mathbf{m}$ is constant, and $\mathbf{M}$ moves with the body: there exists a symmetric A such that $\mathbf{M} = A\mathbf{\Omega}$, such that A depends only on the geometry of the body. Additionally, $\langle A\mathbf{\Omega}, \mathbf{\Omega}\rangle = 2T$, and if $\mathbf{e}$ is a unit vector, then the moment of inertia of the body about the $\mathbf{e}$-axis is $\langle A\mathbf{e}, \mathbf{e}\rangle$; if the system is made of discrete points with masses m_i and distances from the $\mathbf{e}$-axis r_i, then $\langle A\mathbf{e}, \mathbf{e}\rangle = \sum m_i r_i^2$ (this will be an integral in the continuous case).

Since $\langle A\mathbf{e}, \mathbf{e}\rangle$ is a positive-definite quadratic form, it is diagonalizable, so there are some axes in which $\langle A\mathbf{\Omega}, \mathbf{\Omega}\rangle = I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2$ (so that the eigenvalues are the moments of inertia about the principal axis).

Theorem 7.1 (Steiner)**.** *Writing $I(\mathbf{e}) = \langle A\mathbf{e}, \mathbf{e}\rangle$ (through the line $\mathbf{e}$), consider two lines $\mathbf{e}_0, \mathbf{e}_1$ separated by a distance ρ, such that $\mathbf{e}_0$ passes through the center of mass of the rigid body. Then, $I(\mathbf{e}_1) = I(\mathbf{e}_0) + 2m\rho^2$.*

Proof. In the discrete case (the continuous case is similar), $\sum m_i r_i = 0$, so

$$\begin{aligned}\sum m_i(r_i+\rho)^2 &= \sum m_i r_i^2 + 2\rho\sum m_i r_i + \sum p_i\rho^2\\ &= I(\mathbf{e}_0) + m\rho^2.\end{aligned}$$

⊠

In particular, the moment of inertia is lowest through the axis containing the center of mass.

The Euler equation can be written as $\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = [\mathbf{M}, \mathbf{\Omega}]$. Since $\mathbf{M} = A\mathbf{\Omega}$, then the equation can be written in terms of just one or the other; it is also $[\mathbf{M}, A^{-1}\mathbf{M}]$ (i.e. some quadratic equation in $\mathbf{M}$). Writing this in principal coordinates gives a quadratic system of 3 equations for $\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t}$ — which initially sounds complicated, but there are lots of integrals: $E = M_1^2/I_1 + M_2^2/I_2 + M_3^2/I_3$ is preserved, and $\|\mathbf{M}\|^2 = M_1^2 + M_2^2 + M_3^2$ is conserved as well. This means that the motion lies on the intersection of a sphere and ellipsoid, so an explicit but ugly formula can be written down, though it can't be integrated with elementary functions. In fact, Jacobi invented elliptic functions for this purpose.

Thus, all trajectories are periodic, except when the sphere is tangent to the middle axis of the ellipse; this is called a homoclinic orbit.

Definition. If O is a hyperbolic (i.e. unstable) point of a vector field and a trajectory starting from O returns to O eventually, then O is called a homoclinic point.

"Now, we don't want to be dizzy, so we view it from the outside, from the fixed coordinate system."

Theorem 7.2 (Poinsot)**.** *Represent a rigid body by its ellipsoid of inertia with semi-axes $1/\sqrt{I_1}$, $1/\sqrt{I_2}$, and $1/\sqrt{I_3}$ (such that $I_1 \le I_2 \le I_3$, so that the axes are in decreasing order). From stationary space, fix the center of the ellipsoid. Since the angular momentum $\mathbf{m}$ is fixed, then one can consider a plane tangent to the ellipsoid, orthogonal to $\mathbf{m}$.*

Then, the ellipsoid rolls around on the plane without slipping.[18]

This motion is actually somewhat complicated; if two of the axes are the same (in which case the ellipsoid is known as an ellipsoid of rotation), the motion is circular, but not necessarily periodic. The ellipsoid also rotates around its own axis, and if this isn't commensurate with 2π, then it never repeats. In the general case, the curve isn't even closed.

Proof of Theorem 7.2. Consider $\boldsymbol{\omega} = B\mathbf{\Omega}$, the axis of rotation of the ellipsoid. If one takes the normal vector to the ellipsoid where $\boldsymbol{\omega}$ intersects it, this vector must be parallel to $\mathbf{m}$. Fixing the ellipsoid, the normal is the gradient of the function $\nabla\langle A\mathbf{\Omega}, \mathbf{\Omega}\rangle = 2A\mathbf{\Omega} = 2\mathbf{M}$ (the derivative looks like this because the function is a quadratic form). Then, draw the plane paerpendicular to this normal at this point (so that it is perpendicular to $\mathbf{m}$). Then, all that needs to be checked is that the distance to this plane is constant.

A vector $\mathbf{X}$ on the elipsoid of inertia satisfies $\langle A\mathbf{X}, \mathbf{X}\rangle = 1$, and $\langle A\mathbf{\Omega}, \mathbf{\Omega}\rangle = 2T$, so scaling $\mathbf{\Omega}$ requires scaline by $\sqrt{2T}$, so the distance from the center (i.e. the fixed point) to the plane is constant. Specifically, the distance is some constant multiple of $\langle \mathbf{\Omega}, \mathbf{m}\rangle$, or the length of the projection, and this is equal to $\langle \mathbf{\Omega}, \mathbf{M}\rangle = \langle A\mathbf{\Omega}, \mathbf{\Omega}\rangle = 2T$, which is constant.

The ellipse does not slip because its instantaneous velocity is always zero. ⊠

Exercise 7.1. Find a formula for the motion's dependence on time in the case of an ellipsoid of rotation.

Now, suppose a rigid body moves in a constant field (i.e. a constant acceleration g, such as gravity). Consider a top: there is one fixed point, but it isn't necessarily the center of mass. If ℓ is the distance between these two points and θ is the angle between the line and the vertical axis, then the Lagrangian is $L = T - U = T - mg\cos\theta$.

Here everything is harder; only one of the integrals of angular momentum is still here, and the phase space is 6-dimensional, so 2 integrals aren't sufficient. In fact, an exact solution canot be computed (the system is non-integrable).

However, suppose that this ellipsoid of inertia is also an ellipsoid of rotation, so that one obtains an extra integral (corresponding to rotation about that axis), which is already sufficient to integrate the system. The Euler angles are two cyclic variables that allow the integrals to be obtains let $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ be a fixed frame and let $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ be coordinates in the moving system K. Let the angle between the vertical axes be θ, as before. Then, the planes perpendicular to $\mathbf{e}_z$ and $\mathbf{e}_3$ intersect at the angle θ. Let $\mathbf{e}_N$ be the unit normal on the line of nodes (referring to the intersection of the two planes — the name is due to Euler and has been religiously followed ever since). Let φ be the angle between $\mathbf{e}_x$ and $\mathbf{e}_N$, and ψ be the angle between $\mathbf{e}_1$ and $\mathbf{e}_N$. These three angles θ, φ, ψ are the Euler angles.

One can rotate $\mathbf{e}_x$ by φ, then around $\mathbf{e}_N$ by θ, then $\mathbf{e}_3$ by ψ, in order to change one basis into the other. But the system is invariant under φ, so it's cyclic. Additionally, ψ is cyclic (the axis of rotation), but θ isn't cyclic, unless the system is totally fixed.

[18] This is an excellent physical model for ROFLing, as it happens.

Computing kinetic energy in these coordinates isn't terribly simple. Note that the three rotations don't commute, but infinitesimally, they do: the error is quadratic, so on the small scale it doesn't make much of a difference. Then, for any point i in the body, its velocty is $\boldsymbol{\omega} \times \mathbf{r}_i$. Here, $\boldsymbol{\omega} = \boldsymbol{\omega}_\varphi + \boldsymbol{\omega}_\psi + \boldsymbol{\omega}_\theta = \dot{\varphi}\mathbf{e}_z + \dot{\psi}\mathbf{e}_3 + \dot{\theta}\mathbf{e}_N$: the three angles all change at some rate, so making this approximation has quadratic error. Since φ and ψ are fixed, the kinetic energy doesn't depend on them, so suppose they are both zero, so that $\mathbf{e}_N = \mathbf{e}_1 = \mathbf{e}_x$, and the only rotation is in the $\mathbf{e}_z\mathbf{e}_y$-plane. This implies that

$$\begin{aligned}\mathbf{e}_3 &= \cos\theta\mathbf{e}_z + \sin\theta\mathbf{e}_y \\ &= \dot{\varphi}\mathbf{e}_z + \dot{\psi}(\cos\theta\mathbf{e}_z + \sin\theta\mathbf{e}_y) + \dot{\theta}\mathbf{e}_x \\ &= \omega_x\mathbf{e}_x + \omega_y\mathbf{e}_y + \omega_z\mathbf{e}_z = \frac{1}{2}\left\langle A\boldsymbol{\Omega}, \boldsymbol{\Omega}\right\rangle.\end{aligned}$$

8. Rigid Bodies III: 2/4/13

The material on rigid bodies is hard to convey in lecture because of long computatiosn and such, so treat this lecture as a guide to reading the relevant chapter in the book.

Once again, denote the fixed coordinates as $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$, and the moving coordinates as $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, where the rigid body is symmetric about $\mathbf{e}_3$. Let θ be the angle between $\mathbf{e}_3$ and $\mathbf{e}_z$. Then, the $\mathbf{e}_x\mathbf{e}_y$-plane and the $\mathbf{e}_1\mathbf{e}_2$-plane intersect at a line; the coordinate in this line is $\mathbf{e}_N$. Then, φ is the angle between $\mathbf{e}_N$ and $\mathbf{e}_y$, and ψ is that between $\mathbf{e}_N$ and $\mathbf{e}_1$. These latter angles correspond to rotation of the body around an axis of symmetry. Additionally, assume a constant vertical field of gravity.

The Lagrangian is $L = T - U = T - mg\ell\cos\theta$, and $T = \frac{1}{2}\left\langle A\boldsymbol{\Omega}, \boldsymbol{\Omega}\right\rangle = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2)$, but the goal is to express it in the coordinates $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$.

As shown last week, in the infinitesimal case rotation can be thought of as changing only one of the angles for the purpose of calculating angular velocity. If only ψ changes, then the angular velocity is $\dot{\psi}\mathbf{e}_3$; similarly, if θ changes, it is $\dot{\theta}\mathbf{e}_N$, and for φ it is $\dot{\varphi}\mathbf{e}_z$.

In order to make everything nice, assume $\varphi = \psi = 0$. This is a dirty trick because we know the Lagrangian doesn't depend on them, so why even bother? This means that rotation will be in the $\mathbf{e}_2\mathbf{e}_3$-plane. This means that $\mathbf{e}_x = \mathbf{e}_N = \mathbf{e}_1$ and $\mathbf{e}_z = \mathbf{e}_3\cos\theta + \mathbf{e}_2\sin\theta$. Thus,

$$\begin{aligned}T &= \dot{\psi}\mathbf{e}_3 + \dot{\theta}\mathbf{e}_1 + \dot{\varphi}(\mathbf{e}_3\cos\theta + \mathbf{e}_2\sin\theta) \\ &= \dot{\theta}\mathbf{e}_1 + \dot{\varphi}\sin\theta\mathbf{e}_2 + (\dot{\psi} + \dot{\varphi}\cos\theta)\mathbf{e}_3 \\ &= \frac{I_1}{2}(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\varphi}\cos\theta)^2,\end{aligned}$$

since $I_1 = I_2$. Notice that this equation depends neither on φ nor upon ψ; what a surprise! Thus, $M_z = \frac{\partial L}{\partial\dot{\varphi}}$ and $M_3 = \frac{\partial L}{\partial\dot{\psi}}$ are constant, yielding the following formulas:

$$\begin{aligned}M_z &= \frac{\partial L}{\partial\dot{\varphi}} = \dot{\varphi}(I_1\sin^2\theta + I_3\cos^2\theta) + \dot{\psi}I_3\cos\theta \\ M_3 &= \frac{\partial L}{\partial\dot{\psi}} = \dot{\varphi}I_3\cos\theta + \dot{\psi}I_3.\end{aligned}$$

With respect to $\dot{\psi}$ and $\dot{\varphi}$, this is a system of linear equations, so they can be expressed in terms of M_z, M_3, and $\cos\theta$. Solving this yields

$$\dot{\varphi} = \frac{M_z - M_3\cos\theta}{I_1\sin^2\theta},$$

which can be plugged into the energy conservation law:

$$E = \frac{I_1}{2}\dot{\theta}^2 + \left(mg\ell\cos\theta + \frac{(M_z - M_3\cos\theta)^2}{2I_1\sin^2\theta}\right) + \frac{M_3^2}{2I_3}.$$

The right-hand term is constant, and the whols system looks like one with a single degree of freedom; the left-hand term is referred to as the effective potential energy $U_{\text{eff}}(\theta)$.

This can be solved, but it's very ugly and the resulting integral can't be written in terms of elementary functions. One can thus make numeric approximations or qualitative observations.

It's convenient to change variables: let $u = \cos\theta$, so that $\dot{u}^2 = \sin^2\theta\dot{\theta}^2$. Then, the equation becomes $\dot{u}^2 = f(u) = (\alpha - \beta u)(1 - u^2) - (a - bu)^2$ and $\dot{\varphi} = (a - bu)/(1 - \dot{u}^2)$. where α, β, a, and b are constants in terms of M_z, M_3, and so forth. Notice f is a degree-3 polynomial with a positive cubic coefficient. Then, $f(\pm1) < 0$ (since the square term is negative). If the maximum between -1 and 1 is positive, then there are two roots (and the third is meaningless,

since it would require $|u| = |\cos\theta| > 1$), and if not, there are none. If there are two roots u_1 and u_2, then motion has to happen between them, and the motion changes periodically between two angles θ_1 and θ_2 (such that $u_1 = \cos\theta_1$ and $u_2 = \cos\theta_2$). Then, $\dot{\varphi}$ is the angular velocity around the vertical axis, so if $\dot{\varphi}$ has constant sign, then the top traces a sinusoidal-like curve. But it may happen that things change sign, which yields a different picture, in which there may be loops. At a critical point, each curve has a cusp rather than a loop. Interestingly, this is an important case, and is physically observable, as in a spinning top starting at an angle.

The sign of $\dot{\varphi}$ depends entirely on $a - bu$, since $1 - u^2$ is always positive, so it is inportant whether $u = a/b$ is positive or negative.

Consider the special case of a slipping top: a small child tries to set up a top spinning perfectly vertically, so that $M_z = M_3 = I_3\omega_3$. Then,

$$\begin{aligned} U_{\text{eff}} &= \frac{M_3^1(1-\cos\theta)^2}{2I_1\sin^2\theta} + m_2\ell\cos\theta \\ &= \left(\frac{M_3^2}{8I_1} - \frac{mg\ell}{2}\right)\theta^2 + mg\ell = C + A\theta^2 \end{aligned}$$

after some questionable computation, where $C = mg\ell$ and $A = M_3^2/8I_1 - mg\ell/2$. Then, consider small θ. $\theta = 0$ is a minimum of this function and is thus a stationary point. Everything depends on the sign of C and A: if $A > 0$, the motion is stable.[19]

If the top begins rotating at some other non-vertical angle θ_0 and is rotating very fast, the main contributior to the energy is $\frac{1}{2}I_3\omega^2$. The Lagrange function sacles by a constant factor (so the qualitative behavior is the same) if the angular velocity is multiplied by C^2 and the gravity is multiplied by C for some constant C; thus, a top spinning very fast looks like a top spinning at a normal speed in very low gravity, so for now assume $g = 0$. Then, the effective potential energy is

$$U_{\text{eff}} = \frac{(M_z - M_3\cos\theta)^2}{2I_2\sin\theta}$$

and the minimum energy is $\cos\theta = M_z/M_3$ (which forces $M_z \le M_3$). Expanding at θ_0, this implies that $\theta = \theta_0 + x$ for some small variation x, and

$$U_{\text{eff}} = \frac{I_3^2\omega_3^2}{2I_1}x^2.$$

This implies that the equation of motion is $I_1\ddot{\theta}/2 = -U_{\text{eff}}$, or that $\ddot{\theta} = c\theta$ for a constant c.

This equation corresponds to a small oscillation with frequency $\sqrt{c} = \omega_{\text{nut}}$ called the frequency of nutation: $\omega_{\text{nut}} = I_3\omega - 3/I_1$. This is the frequency for how θ changes (though the rotation of the actual rigid body is more complicated than just a harmonic oscillator in the general case). Similarly, one can compute that $\dot{\varphi}$ is a harmonic oscillator with the same frequency.

Adding gravity back in shifts the minimum of the quadratic function U_{eff} very slightly, so the second derivative effectively doesn't change, and the formula for nutation still approximately holds.

9. The Legendre Transform and Projective Duality: 2/6/13

Suppose $y = f(x)$ is convex, or, more exactly, f' is a monotonically increasing map that is onto $\mathbb{R}$. Then, the goal is to construct a functional on the dual space to $\mathbb{R}$ (i.e. the set of linear functions on $\mathbb{R}$, each a function $\ell(x) = px$, so the space is isomorphic to $\mathbb{R}$). Then, define the Legendre transform of f to be a function $\hat{f}(p)$: if $p \in \mathbb{R}$, then there is exactly one x such that $f'(x) = p$; then, the tangent line to f at x looks like $y = px + b$. Then, $\hat{f}(p) = b$.

More explicitly, $\hat{f}(p) = \max_{x\in\mathbb{R}}(px - f(x)) = px - f(x)$ where $f'(x) = p$.

Example 9.1. If $f(x) = x^2/2$, then $\hat{f}(p) = px - x^2/2$ where $x - p$, so it simplifies to $p^2 - p^2/2 = p^2/2$, which is pretty nice.

If $f(x) = mx^2/2$, then $\hat{f}(p) = px - mx^2/2$ where $mx = p$, so it becomes $p^2/2m$, which is the expression for kinetic energy in terms of momentum.

Exercise 9.1. What is the Legendre transform of $f(x) = x^\alpha/\alpha$?

Solution. It will end up as $\hat{f}(p) = x^\beta/\beta$ such that $\alpha^{-1} + \beta^{-1} = 1$.

Here are some properties: $\hat{f}(p) \ge px - f(x)$, so $f(x) + \hat{f}(p) \ge px$. This is called the Young Inequality. It also has some non-obvious corollaries, such as $x^\alpha/\alpha + p^\beta/\beta \ge px$, which is hard to prove explicitly.

If a function has a monotone derivative that doesn't reach all values, the Legendre transform can still be computed: if $f(x) = e^x$, then $\hat{f}(p) = px - e^x$, with $p = e^x$, so $x = \ln p$. Thus, for positive p, $\hat{f}(p) = p\ln p - p$.

[19] Of course, if friction is accounted for, ω_3 decreases, so M_3 decreases, so the top falls.

Claim. The Legendre transformation of a convex f (such that f' is onto $\mathbb{R}$) has the same properties of convexity.

Proof. Exercise.

Theorem 9.1. *The Legendre transform is involutive: $\hat{\hat{f}}(x) = f(x)$ for any function f that meets the conditions.*

Proof. This is sort of a tautology: $\hat{f}(p) = \max(px - f(x))$, so $\hat{\hat{f}}(x) = \max(xp - \hat{f}(p))$, since the double-dual is naturally isomorphic to $\mathbb{R}$. Then, $px - \hat{f}(p)$ is just the tangent line as a function, and at x it is just $f(x)$.

The goal is to find the tangent line with the greatest intersection with the vertical line through x, and it is a property of convex functions that the maximum is $f(x)$. ⊠

One can generalize into more dimensions: let V be an n-dimensional vector space, and let $f : V \to R$ be convex. More precisely, there is an orientation-preserving map given by the differential that is one-to-one.[20] An example of a convex function is any nondegenerate quadratic form.

Then, $\hat{f}(\mathbf{x}) = \max_{\mathbf{x}\in V}(\mathbf{px} - f(\mathbf{x})) = \mathbf{px} - f(\mathbf{x})$ given $\mathbf{p} = \mathrm{d}f$ (so that $p_i = \frac{\partial f}{\partial x_i}$ for all i). Here, $\mathbf{px}$ can either mean the dot product (so $\mathbf{p} \in V$) or as the function $\mathbf{p}(\mathbf{x})$ (so that $\mathbf{p} \in V^*$).

This can be generalized yet further to the notion of projective duality. Projective space is an object of classical mathematics. It is an unfortunate fact of Euclidean geometry that some lines don't intersect (of course, the parallel ones). People have been annoyed by this, so they defined a point at infinity that is the class of all parallel lines (so that they "intersect" at infinity).

First consider the real projective space $\mathbb{R}P^2$. There are many different models of this, but one is the space of lines in $\mathbb{R}^3$ through 0. If $\mathbb{R}^3$ has coordinates x, y, z, consider the plane $xy = 1$. Then, all lines in $\mathbb{R}P^2$ intersect this plane at exactly one point, except for the one which is parallel (thus corresponding to adding one point). Similarly, one can take the real line and compacitfy it by adding one point (as the space of lines through the origin in $\mathbb{R}^2$) to get the projective line $\mathbb{R}P$, which is topologically equivalent to a circle. This can be generalized to $\mathbb{R}P^n$, which can also be thought of as affine space A^n with $\mathbb{R}P^{n-1}$ attached.

Life in projective space is like a utopia. Everybody's equal, and everyone intersects! But it's even better: in Euclidean space, there are lower classes (points) and higher classes (lines), so that two points determine a line, but not necessarily vice versa. However, in projective space, there is this kind of harmony: two points determine a line, and two lines determine a point.

Projective duality is the notion that there is a dual projective space where points go to lines and vice versa: send a line (as a hyperplane in $\mathbb{R}^{n+1}$) to the point (i.e. a line in $\mathbb{R}^{n+1}$) it is perpendicular to, and vice versa.[21]

Another nice fact is that in projective space, all quadratics are the same, only differing by how they intersect the point at infinity; all ellipses, parabolas, and hyperbolas are closed curves. An ellipse doesn't intersect the point at infinity; a parabola is tangent to it, and a hyperbola intersects it without being tangent.

The dual space has an affine part consisting of those lines which aren't vertical (since they're given by $\varphi = px - b$, so they can be seen as in $\mathbb{R}^2$). Take a graph of some function f on the projective plane, and take the space of all lines tangent to it. Then, in the orthogonal construction, these become points which trace out a curve. If the function is convex, this curve will lie in the affine part (since the tangent line is never vertical). This is just the Legendre transform!

Thus, the Legendre transform can be defined for nonconvex functions; if the function has an inflection point, its Legendre transform has a cusp (that looks like the semicubical parabola $x^{2/3}$), and vice versa. Similarly, if the tangent lines are ever parallel, the transform will self-intersect.

And projective duality is a special case of something much more general...

Returning to mechanics: in Lagrangian mechanics, one has a configuration space M and a phace space, or tangent bundle, $TM = \{T_{\mathbf{q}}M \mid \mathbf{q} \in M\}$, so that any local (generalized) coordinates $q_1, \dots, q_n$ extend to a basis of TM given by $\dot{q}_1 = \frac{\partial}{\partial q_1}, \dots, \dot{q}_n = \frac{\partial}{\partial q_n}$, so TM has the coordinates $\mathbf{q}, \dot{\mathbf{q}}$, as a $2n$-dimensional manifold.

Now, consider the cotangent bundle T^*M. If $\mathbf{q} \in M$, then $T^*_{\mathbf{q}}M$ is the dual space to $T_{\mathbf{q}}M$, the set of linear functionals on $T_{\mathbf{q}}M$, and $T^*M = \{T^*_{\mathbf{q}}M \mid \mathbf{q} \in M\}$. Then, the coordinates $q_1, \dots, q_n$ on M induce coordinates on T^*M $p_1 = \mathrm{d}q_1, \dots, p_n = \mathrm{d}q_n$.

It happens that $TM \cong T^*M$, but the isomorphism is not canonical. However, if the manifold is Riemannian (i.e. every tangent space has a scalar product, and therefore a metric and a Euclidean structure), then thee is a canonical choice: $V \to V^*$ given by $\mathbf{v} \mapsto \mathbf{v}(\mathbf{x}) = \langle \mathbf{v}, \mathbf{x} \rangle$.

In the Lagrangian case, the Lagrangian is a function $L : TM \to \mathbb{R}$. In some nice systems called natural systems, L is a positive-definite quadratic form on every tangent fiber and U depends only on $\mathbf{q}$. Then, one can take the Legedre transform of this quadratic form. The Legendre transform also can be thought of as a map $V \to V^*$ that sends

[20] This has the odd consequence that convexity and concavity are only distinguished by orientation.

[21] This also has the corollary that the spaces of lines and of hyperplanes are isomorphic, as with k-hyperplanes and $(n-k)$-hyperplanes.

$\mathbf{x} \mapsto \mathbf{p} = \mathrm{d}_{\mathbf{x}} f$ (given an f on V), or in coordinates, $x_i \mapsto \frac{\partial f(x_i)}{\partial x_i}$. This is a diffeomorphism. Thus, the Lagrangian, restricted to TM, induces a map $TM \mapsto T^*M$ given by its Legendre transform. If $L(\mathbf{q}, \dot{\mathbf{q}})$ is a quadratic in $\dot{\mathbf{q}}$, the Legendre transform is $H(\mathbf{q}, \mathbf{p}) = \hat{L}(\mathbf{q}, \mathbf{p})$, so $H : T^*M \to \mathbb{R}$. This function H is called the Hamilton function, or the Hamiltonian.

Working in coordinates, if $L = \sum m_i \dot{q}_i^2/2 - U(\mathbf{q})$, then $\hat{L} = \sum p_i \dot{q}_i - \sum m_i \dot{q}_i^2/2 + U(\mathbf{q})$ such that $p_i = \frac{\partial L(\dot{\mathbf{q}}, \mathbf{q})}{\partial \dot{q}_i}$. (so that $\dot{q}_i$ plays the role of x). This is a system of n equations in n variables, so it can be solved, giving a Legendre transform of $\hat{L} = \sum p_i/2m_i^2 - U(\mathbf{q})$. This is the kinetic energy again, but expressed in terms of the generalized momenta $\frac{\partial L}{\partial q_i}$ corresponding to the q_i. Thus, the Hamiltonian is just the full energy of the system expressed in terms of momentum rather than velocity.

Recall that if $\mathbf{q}(t)$ is a trajectory in M, then it lifts to a trajectory on TM. But the map also induces a trajectory $\mathbf{q}(t), \mathbf{p}(t)$ that is almost completely equivalent, and considerably more powerful, so one can solve question in T^*M and then transmit the answers back to TM.

In the one-dimensional case, $H(q, p) = \max(p\dot{q} - L(q, \dot{q}))$. The function being maximized is a function of three variables $G(q, \dot{q}, p)$ (or G such that $p = \frac{\partial L}{\partial q}$). Let Γ be the hypersurface given by $p = \frac{\partial L}{\partial q}$, so there is a map $(q, p) \overset{S}{\mapsto} (q, \dot{q})$ where $p = \frac{\partial L}{\partial q}$. Thus, Γ is the graph of S and also of something else, because it's one-to-one. Then, $\dot{q} = \Phi(q, p)$, so that $H(q, p) = G(q, \dot{q} = \Phi(q, p), p)$.

Then, the differential of H is a pullback: $\mathrm{d}H = S^* \,\mathrm{d}G$ (this is just restricting it to Γ and writing it in (q, p)-coordinates). Most books get this wrong, and even Arnold doesn't explain it (saying "I do not wish to break tradition"). Let G_p denote the partial derivative with respect to p, and so on; then,

$$\begin{aligned}
\mathrm{d}G &= G_p \,\mathrm{d}p + G_q \,\mathrm{d}q + G_{\dot{q}} \,\mathrm{d}\dot{q} \\
&= \dot{q} \,\mathrm{d}p + p \,\mathrm{d}q + \dot{L}_q \,\mathrm{d}q - L_{\dot{q}} \,\mathrm{d}\dot{q} \\
&= \Phi \,\mathrm{d}p + p \,\mathrm{d}\Phi - L_q \,\mathrm{d}q - L_{\dot{q}} \,\mathrm{d}\dot{q}
\end{aligned}$$

This requires plugging in $\mathrm{d}\Phi$ for q, which is what the pullback actually does. Then, $\mathrm{d}H$ falls out, giving a formula for H_p and H_q: $H_p = \dot{q}$, $H_q = -L_q$.

In general, nothing more can be said, but along the trajectory,

$$\begin{gathered}
\frac{\partial L}{\partial q} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right) \implies \frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\partial L}{\partial q} \\
\implies \dot{p} = \frac{\mathrm{d}p}{\mathrm{d}t} = -H_q \quad \text{and} \quad \dot{q} = \frac{\mathrm{d}q}{\mathrm{d}t} = H_p.
\end{gathered}$$

This is a system of two first-order differential equations, called the Hamilton canonical equations, and it's really nice and symmetric. A system of first-order differential equations is a vector field, so a trajectory is just an integral of this field. Diffeomorphisms of this vector field (called canonical transformations in mechanics or symplectomorphisms in mathematics) preserve the form of these equations, so there are a lot more options for cyclic coordinates. And if there are n cyclic coordinates, the system is explicitly integrable, so one calls in symplectic geometry to answer these sorts of questions.

10. The Least Action Principle in Hamiltonian Mechanics: 2/11/13

Recall some stuff from last week: if M is the configuration space of some system, then TM is the phase space considered in Lagrangian mechanics. Every coordinate system $\mathbf{q}$ on M induces coordinates $\dot{\mathbf{q}}$ on TM. The dymanics of this system are given by the Lagrangian $L : TM \to \mathbb{R}$. In some systems, called natural systems, $L = T - U$, where T is a fiberwise positive-definite quadratic form and U doesn't depend on $\mathbf{q}$.

The Hamiltonian formalism considers the cotangent bundle T^*M (the space of linear functions). A section of the tangent bundle is a vector field, but a section of the cotangent bundle is a linear function associating a linear function to each point, so it is a differential 1-form. The coordinates $q_1, \ldots, q_n$ on M induce coordinates $p_1 = \mathrm{d}q_1, \ldots, p_n = \mathrm{d}q_n$ on T^*M, often denoted $(\mathbf{q}, \mathbf{p})$.

The Legendre transform can be used to move between the tangent and cotangent bundles. On some vector space V, one obtains orientation-preserving diffeomorphisms $V \to V^*$ such that $\mathbf{x} \mapsto \mathrm{d}_{\mathbf{x}} f$. Then, $f(\mathbf{x})$ is transformed into some other function $\hat{f}(\mathbf{p}) = \mathbf{p}\mathbf{x} - f(\mathbf{x})$ at the maximum, so $\mathbf{p} = f'(\mathbf{x})$. This should be understood as a system of equations, so $\mathbf{x}$ can be expressed in terms of $\mathbf{p}$.

On each tangent space, the Lagrangian is a convex function (in the case of natural systems and some others), so the Legendre transform can be applied parametrically, with $\mathbf{q} \in M$ as the parameter. Then,

$$\hat{L}(\mathbf{p}, \mathbf{q}) = \mathbf{p}\dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}})\big|_{\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}}. \tag{4}$$

This Legendre transform is usually denoted $H(\mathbf{p},\mathbf{q})$, and is called the Hamiltonian function on the system.[22] The Legendre transform not only acts on functions, but can also act as a more general map on spaces.

Be careful of a nuance: usually, one can take pullbacks to compute the results of diffeomorphisms. However, suppose that X and Y are two sets and $h : X \to Y$ is one-to-one. Then, if $f : Y \to \mathbb{R}$, then $h^* f = f(h(x))$ is a function from $X \to \mathbb{R}$. if $g : X \to \mathbb{R}$, then $g \circ h^{-1}$ is a function on Y. This can always be done with this kind of map: it induces isomorphisms of functions. In the Hamiltonian case, this is exactly what is being done in (4), but putting the $\mathbf{pq}$ in front means it's not quite that simple.

Thus, systems can be mapped to paths in the cotangent bundle. These paths satisfy the following equations in local coordinates: $\dot{\mathbf{p}} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -H_{\mathbf{q}}(\mathbf{p},\mathbf{q})$ and $\dot{\mathbf{q}} = \frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = H_{\mathbf{p}}(\mathbf{p},\mathbf{q})$ (the Legendre transform of $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) = \frac{\partial L}{\partial \mathbf{q}}$).

Physically, this can be thought of as passing from velocities to momenta; H is the full energy of the system expressed in terms of the momentum; U is constant in the transform.

Write $\mathbf{u} = (\mathbf{p},\mathbf{q})$, so that $\dot{\mathbf{u}} = (-H_{\mathbf{q}}, H_{\mathbf{p}})$. This is a first-order system of differential equations, and it is therefore also a vector field $\mathbf{X}_H = -H_{\mathbf{q}}\frac{\partial}{\partial \mathbf{p}} + H_{\mathbf{p}}\frac{\partial}{\partial \mathbf{q}}$ (where the vector fields $\frac{\partial}{\partial p_i}, \frac{\partial}{\partial q_i}$ are the unit vector fields,[23] or the vectors with velocity 1 in the directions p_i and q_i), and the trajectory is $\dot{\mathbf{u}} = \mathbf{X}_H(\mathbf{u})$.

Consider the special case when $M = \mathbb{R}^n$. Then, $T^*\mathbb{R}^n = \mathbb{R}^{2n}$ with coordinates $(p_1,\dots,p_n,q_1,\dots,q_n)$, which we will want to think of as $\mathbb{R}^{2n} = \mathbb{C}^n$ with coordinates $(p_1 + iq_1,\dots,p_n + iq_n)$. If $H(\mathbf{p},\mathbf{q})$ is the Hamiltonian, then its gradient is $\nabla H = \left(\frac{\partial H}{\partial \mathbf{p}}, \frac{\partial H}{\partial \mathbf{q}}\right)$. Thus, $\mathbf{X}_H = i\nabla H$.

Geometrically, the level sets of ∇H are the level sets of the full energy. Then, the gradient is normal to these level sets $H(\mathbf{p},\mathbf{q}) = c$, so multiplying by i makes things tangent to these level sets. Then, conservation of energy becomes obvious: since $\mathbf{X}_H$ is always tangent to the level sets, then all trajectories must lie on these level sets.

Example 10.1. Consider $H = \frac{1}{2}(\|\mathbf{p}\|^2 + \|\mathbf{q}\|^2)$. If $H = E$ (i.e. some constant), then the level set is the sphere of radius $\sqrt{E}$. Thus, $\nabla H = \mathbf{p}\frac{\partial}{\partial \mathbf{p}} + \mathbf{q}\frac{\partial}{\partial \mathbf{q}}$. In a real vector space, every point belongs to a line, and in a complex vector space, every point intersects a complex line (a plane), so each line has a plane, which determines a circle that intersects with the sphere. The sphere is the union of such circles, which is referred to as a cofibration. If this is generalized to an ellipsoid, there are n different closed orbits, and everything else is not closed: if $H = \sum_{i=1}^n \frac{p_i^2+q_i^2}{2a_i^2}$, the closed sets happen only when intersecting with one of the complex axes.

Recall that in Lagrangian mechanics, the least action principle states that if γ is a path between $\mathbf{q}_1, \mathbf{q}_2 \in M$. Then, if γ is lifted to TM, then the principle claims that the paths are those which extremize $S(\gamma) = \int_0^1 L(\gamma,\dot{\gamma},t)\,\mathrm{d}t$.

In Hamiltonian mechanics, the least action principle is considerably more powerful. Take any path $\gamma(t) = (\mathbf{p}(t),\mathbf{q}(t))$, where $\gamma(0) = (\mathbf{p}_0,\mathbf{q}_0)$ and $\gamma(1) = (\mathbf{p}_1,\mathbf{q}_1)$. Other than the boundary conditions, there are no constraints on $\mathbf{p}$ and $\mathbf{q}$.

Theorem 10.1. *On this class of paths,[24] the system travels on a path that minimizes* $S(\gamma) = \int_\gamma (\mathbf{p}\,\mathrm{d}\mathbf{q} - H)\,\mathrm{d}t$.

In coordinates, there exists a differential form $\mathbf{p}\,\mathrm{d}\mathbf{q} = \sum_{i=1}^n p_i\,\mathrm{d}q_i$. There is a coordinate-invariant definition of this form, however; since a differential 1-form is a linear map on each cotangent fiber, vectors can be projected from the cotangent bundle to M via $\pi : T^*M \to M$. Then, $\mathrm{d}\pi : T(T^*M) \to TM$. If $\mathbf{v} \in T_{\mathbf{q},\mathbf{p}}(T^*M)$, then $\mathrm{d}\pi(\mathbf{v}) = T_{\mathbf{q}}(M)$, but since $\mathbf{p} \in T_{\mathbf{q}}M$, then $\mathbf{p}$ is a linear function, so $\mathbf{p}(\mathrm{d}\pi(\mathbf{v}))$ is a number, and $\mathbf{p}(\mathrm{d}\pi(\mathbf{v})) = \mathbf{p}\,\mathrm{d}\mathbf{q}(\mathbf{v})$, which is the same thing that was done in coordinates. Thus, the form $\mathbf{p}\,\mathrm{d}\mathbf{q}$ is universal, independent of coordinates. However, in coordinates induced from M, it always has the form $\sum p_i\,\mathrm{d}q_i$.

Evaluating the action functional, if $\gamma(t) = (\mathbf{p}(t),\mathbf{q}(t))$, then

$$S(\gamma) = \int_\gamma \mathbf{p}\,\mathrm{d}\mathbf{q} - H(\mathbf{p},\mathbf{q}) = \int_\gamma \sum p_i\,\mathrm{d}q_i - H(\mathbf{p},\mathbf{q}) = \int_0^1 \left(\sum p_i(t)\dot{q}_i(t) - H(\mathbf{p}(t),\mathbf{q}(t))\right)\,\mathrm{d}t.$$

For this function, the Euler-Lagrange equation is easy to write down, but since there is a different boundary condition, some proof is necessary. Consider a variation

$$\begin{aligned} S(\gamma+\delta\gamma) &= \int_0^1 (\mathbf{p}_0+\delta\mathbf{p})(\dot{\mathbf{q}}_0+\delta\dot{\mathbf{q}}) - H(\mathbf{p}_0+\delta\mathbf{p},\mathbf{q}_0+\delta\mathbf{q})\,\mathrm{d}t \\ &= \int_0^1 (\mathbf{p}_0\dot{\mathbf{q}}_0 - H(\mathbf{p},\mathbf{q}))\,\mathrm{d}t + \int_0^1 \left(\delta\mathbf{p}_0\dot{\mathbf{q}}_0 + \mathbf{q}_0\delta\dot{\mathbf{q}}_0 - \frac{\partial H}{\partial \mathbf{p}}\delta\mathbf{p} - \frac{\partial H}{\partial \mathbf{q}}\delta\mathbf{q}\right)\,\mathrm{d}t \end{aligned}$$

[22] In terms of the meanings of the coordinates, $H(\mathbf{q},\mathbf{p})$ would be a more natural ordering, but tradition and the alphabet have ensured that $H(\mathbf{p},\mathbf{q})$ is most commonly seen.

[23] This at first peculiar notation is used because when the vector field is thought of as an operator of differentiation, these coincide with partial differentiation in the appropriate direction.

[24] This class of paths is much larger, and is in fact $2n$-dimensional. This is the source of power of Hamiltonian mechanics.

plus some higher-order terms. At a critical point, this variation should be zero, so

$$= \int_0^1 \left(\delta \mathbf{p}_0 \dot{\mathbf{q}}_0 + \delta \dot{\mathbf{q}}_0 \mathbf{p}_0 - \frac{\partial H}{\partial \mathbf{p}} \delta \mathbf{p} - \frac{\partial H}{\partial \mathbf{q}} \delta \mathbf{q} \right) \, \mathrm{d}t$$

$$\implies \int_0^1 \delta \dot{\mathbf{q}} \mathbf{p}_0 \, \mathrm{d}t = \delta \mathbf{q} \mathbf{p}_0 |_0^1 - \int_0^1 \delta \mathbf{q} \dot{\mathbf{p}}_0 \, \mathrm{d}t$$

after integrating by parts. Then, if $\mathbf{q}$ is fixed, then the first term goes to zero (hooray for boundary conditions!), so when all of the terms are collected,

$$\int_0^1 \delta \mathbf{p} \left(\dot{\mathbf{q}}_0 - \frac{\partial H}{\partial \mathbf{p}} \right) + \delta \mathbf{q} \left(\dot{\mathbf{p}}_0 - \frac{\partial H}{\partial \mathbf{q}} \right) \, \mathrm{d}t = 0$$

for all $\delta \mathbf{p}$ and $\delta \mathbf{q}$. This implies that $\dot{\mathbf{q}} = -\frac{\partial H}{\partial \mathbf{p}}$ and $\dot{\mathbf{p}} = \frac{\partial H}{\partial \mathbf{q}}$, which was a variational proof of the Hamilton equations. There are many other boundary conditions that work; for example, if the space of paths is such that momenta (or equivalently, velocities) are zero at the endpoints, or at only one endpoint, things still work; this has many useful applications.

The form $\mathbf{p} \, \mathrm{d}\mathbf{q}$ is important for Hamiltonian mechanics, but there is a more important form called the symplectic form, $\mathrm{d}(\mathbf{p} \, \mathrm{d}\mathbf{q}) = \sum p_i \wedge q_i$. This is related to the quite important interpolation between real coordinates in $\mathbb{R}^{2n}$ and complex ones in $\mathbb{C}^n$.

A bilinear function is a function $f(\mathbf{x}, \mathbf{y})$ that is linear in each argument. A function can be complex bilinear (i.e. $f(\lambda \mathbf{x}, \mathbf{y}) = f(\mathbf{x}, \lambda \mathbf{y}) = \lambda f(\mathbf{x}, \mathbf{y})$ for any $\lambda \in \mathbb{C}$). If f is complex-valued, then it has real and imaginary parts, and it is possible to obtain from this a complex form whose real part is invariant under multiplication by i. However, if the function is bilinear, it's not possible to obtain anything interesting, so consider instead a Hermitian form, for which $f(\lambda \mathbf{x}, \mathbf{y}) = f(\mathbf{x}, \overline{\lambda} \mathbf{y})$, or $f(\mathbf{x}, \mathbf{y}) = \overline{f(\mathbf{y}, \mathbf{x})}$, where f is linear in the first argument.

There are some standard Hermitian forms, such as $f(\mathbf{x}, \mathbf{y}) = \sum x_i \overline{y}_i$. For example, if $x, y \in \mathbb{C}$ and $x = p + iq$, $y = u + iv$, then $x\overline{y} = (pu + qv) + i(pv - qu) = xy + i \det(x, y)$.[25] It turns out that the real part of any Hermitian form is symmetric, and the imaginary part is always skew-symmetric.

If $\mathbb{R}^2$ has some linear functions p and q, then there is a bilinear form called the exterior product: $p \wedge q(\mathbf{x}, \mathbf{y}) = \det(p(\mathbf{x}), q(\mathbf{y}))$. In higher dimensions one obtains $p_1 \wedge q_1 + \cdots + p_n \wedge q_n$. This is called the symplectic form.

Claim. Suppose V is a real vector space and F is a skew-symmetric bilinear function on V. Then,

(1) If F is nondegenerate, then $\dim V$ is even. (This is clear because $\det(A^{\mathrm{T}}) = \det(A) = \det(-A) = (-1)^n \det(A)$.)
(2) Given any nondegenerate, skew-symmetric, bilinear form, there exists a coordinate system $(p_1, , q_1, \ldots, p_n, q_n)$ such that $F = \sum_{j=1}^n p_j \wedge q_j$.

Thus, up to change of coordinates, this form is unique, which is an amazing property of skew-symmetric forms (which is completely untrue for, e.g., symmetric quadratic forms).

11. Some Linear Algebra: 2/20/13

Let V be a vector space and F be a bilinear function on V. Every bilinear function is determined, in a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$, by $F(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^{\mathrm{T}} A \mathbf{Y}$, where $a_{ij} = F(\mathbf{v}_i, \mathbf{v}_j)$. If F is skew-symmetric, so that $F(\mathbf{X}, \mathbf{Y}) = -F(\mathbf{Y}, \mathbf{X})$, then A is as well, so that $A^{\mathrm{T}} = -A$.

Exercise 11.1. Show that if A is the matrix of a skew-symmetric operator, then its rank must be even.

This implies in particular that if A has maximal rank, its determinant is nonzero, so $\dim V$ is even. If this is the case, then F is called nondegenerate. Such a form can only exist in an even-dimensional vector space.[26]

If F is a bilinear form and $\mathbf{X}$ is a vector, one can create a linear function by contraction: $\mathbf{X} \lrcorner F$ is given by $\mathbf{X} \lrcorner F(\mathbf{Y}) = F(\mathbf{X}, \mathbf{Y})$. Thus, F is nondegenerate if contraction with a nonzero vector gives a nonzero function.

[25] In the higher-dimensional case, this generalizes to $\sum_i \det(x_i, y_i)$, viewing each complex number as a 2-dimensional vector.
[26] An equivalent and more natural definition for nondegeneracy is that for all nonzero $\mathbf{X} \in V$, there exists a $\mathbf{Y}$ such that $F(\mathbf{X}, \mathbf{Y}) \neq 0$.

Theorem 11.1. *Let F be a nondegenerate, skew-symmetric bilinear function on V. Then, there exists a basis $\mathbf{e}_1, \mathbf{f}_1, \dots, \mathbf{e}_n, \mathbf{f}_n$ of V such that the matrix of F has the form*

$$\begin{pmatrix} 0 & 1 & & 0 & 0 \\ -1 & 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & \vdots & \\ 0 & 0 & & 0 & 1 \\ 0 & 0 & \cdots & -1 & 0 \end{pmatrix}$$

This is equivalent to the following:

- $F(\mathbf{e}_i, \mathbf{f}_i) = 1$ *for all i, and*
- $F(\mathbf{e}_i, \mathbf{f}_j) = F(\mathbf{e}_i, \mathbf{e}_j) = F(\mathbf{f}_i, \mathbf{f}_j)$ *if $i \neq j$.*

This means that if coordinates in this basis are denoted $x_1, y_1, \dots, x_n, y_n$, then $F = \sum_{i=1}^n x_i \wedge y_i$.

Proof. Proceed by induction on n (where $\dim(V) = 2n$).

If $n = 1$, then let $\mathbf{e}$ be any vector. Then, there exists an $\mathbf{f}$ such that $F(\mathbf{e}, \mathbf{f}) \neq 0$. Thus, $\mathbf{f}$ can be normalized such that $F(\mathbf{e}, \mathbf{f}) = 1$.

Suppose this holds true up to $n-1$. Then, let $\mathbf{e}_1$ be any vector, and find a corresponding $\mathbf{f}_1$, which will make up the first block. Let $L = \operatorname{span}(\mathbf{e}_1, \mathbf{f}_1)$, and consider $L^{\perp F} = \{\mathbf{Z} \in V \mid F(\mathbf{Z}, \mathbf{X}) = 0 \text{ for all } \mathbf{X} \in L\}$.[27] Then, $\dim(L) + \dim(L^{\perp F}) = \dim(V)$ by Exercise 11.2 below, so $\dim(L^{\perp F}) = 2n-2$, and $L^{\perp F} \pitchfork L$ (i.e. they intersect each other transversely, since they intersect only at the origin). By induction, there is a basis $\mathbf{e}_2, \mathbf{f}_2, \dots, \mathbf{e}_n, \mathbf{f}_n$ in which A has the right form, so adding on $\mathbf{e}_1$ and $\mathbf{f}_1$ leads to the desired matrix. ⊠

Exercise 11.2. Show that, in the above proof, if F is nondegenerate, then $\dim(L) + \dim(L^{\perp F}) = \dim(V)$. This can be done by reducing it to a system of equations.

This is a remarkable result — there is a whole zoo of symmetric forms, but there is only the one skew-symmetric form. This suggests that it is important.

Choose the standard basis $\{\mathbf{e}_1, \mathbf{f}_1, \dots, \mathbf{e}_n, \mathbf{f}_n\}$ of $\mathbb{R}^{2n}$ and let $\omega = \sum x_k \wedge y_k$. $\mathbb{R}^{2n}$ can also be viewed as a complex vector space $\mathbb{C}^n$, where multiplication by i is defined as $i\mathbf{e}_j = \mathbf{f}_j$ and $i\mathbf{f}_j = -\mathbf{e}_j$. This can be thought of as a linear operaotr whose square is $-\operatorname{Id}$. Then, one jas complex coordinates $z_k = x_k + iy_k$.

Here one can consider Hermitian forms. There is a unique positive-definite Hermitian form up to change of coordinates: if $\mathbf{z} = (z_1, \dots, z_n)$ and $\mathbf{w} = (w_1, \dots, w_n)$, then $H(\mathbf{z}, \mathbf{w}) = \sum z_i \overline{w}_i$.

Claim. The real part of H is $\langle \mathbf{z}, \mathbf{w} \rangle = \sum x_k u_k + y_k v_k$ (if $z_k = x_k + iy_k$ and $w_k = u_k + iv_k$), and the imaginary part is $-\omega(\mathbf{z}, \mathbf{w})$.

Proof. It's sufficient to check for $n = 1$, which is essentially the same thing: $z\overline{w} = (x+iy)(u-iv) = (xu+yv) - i(xv-yu)$. Then, the real part is $\langle w, z \rangle$ and the imaginary part is $\begin{vmatrix} x & u \\ y & v \end{vmatrix} = \omega(z, w)$. ⊠

$H(i\mathbf{z}, i\mathbf{w}) = H(\mathbf{z}, \mathbf{w})$ by the Hermitian property, so both the imaginary and real parts are invariant under multiplication by i. But if g denotes the standard scalar product, then $H(i\mathbf{z}, \mathbf{w}) = iH(\mathbf{z}, \mathbf{w}) = ig(\mathbf{z}, \mathbf{w}) + \omega(\mathbf{z}, \mathbf{w}) = g(i\mathbf{z}, \mathbf{w}) - \omega(i\mathbf{z}, \mathbf{w})$. Thus, equating real and imaginary parts, $g(i\mathbf{z}, \mathbf{w}) = \omega(\mathbf{z}, \mathbf{w})$ and $\omega(i\mathbf{z}, \mathbf{w}) = -g(\mathbf{z}, \mathbf{w})$ (which is the same thing, really).

Definition. A subspace $L \subset V$ (where V is a symplectic space, or a space with a symplectic form ω) is called isotropic if $\omega|_L = 0$ (i.e. $\omega(\mathbf{x}, \mathbf{y}) = 0$ for any $\mathbf{x}, \mathbf{y} \in L$).

Example 11.1. The trivial subspace is of course isotropic (and if ω were replaced with a symmetric form, this would be the only one). Another more interesting example is $\operatorname{span}\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ or $\operatorname{span}\{\mathbf{f}_1, \dots, \mathbf{f}_k\}$.

Exercise 11.3. Show that if L is an isotropic subspace, then $\dim L \leq n$. (As a hint, take the ω-orthogonal complement and then use the nondegeneracy of the form.)

Isotropic spaces of dimension n are called Lagrangian subspaces, and really play a central role in symplectic geometry and Hamiltonian systems. A quick check of the definition shows that Lagrangian spaces are precisely those for which $L^{\perp\omega} = L$. Since $\omega(\mathbf{x}, \mathbf{y}) = \langle i\mathbf{x}, \mathbf{y} \rangle$, then $L^{\perp\omega} = iL^{\perp}$, so if L is multiplied by i, then it becomes its own orthogonal complement!

[27] The usual orthogonal complement is the one for which F is the scalar product.

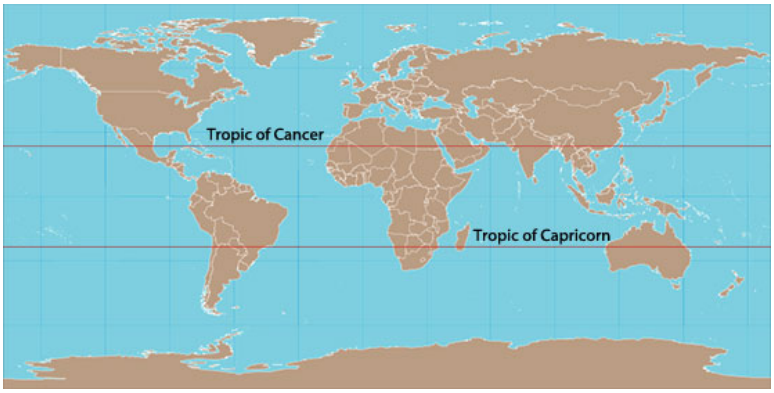

FIGURE 2. The Tropic of Cancer and the Tropic of Capricorn as isotropic subspaces of S^2, the Earth.

Suppose L is Lagrangian and graphical with respect to the coordinates $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$. Then, L is the graph of a symmetric linear function: $L = \{\mathbf{y} = A\mathbf{x}\}$, and the converse is also true: if $\mathbf{z} = (\mathbf{x}, A\mathbf{x}), \tilde{\mathbf{z}} = (\tilde{\mathbf{x}}, A\tilde{\mathbf{x}}) \in L$, then $\omega(\mathbf{z}, \tilde{\mathbf{z}}) = \langle i\mathbf{z}, \tilde{\mathbf{z}}\rangle = -\langle A\mathbf{x}, \tilde{\mathbf{x}}\rangle + \langle \mathbf{x}, A\tilde{\mathbf{x}}\rangle = 0$, so A is self-adjoint and can therefore be made symmetric.

If V is a symplectic space, the set of ω-preserving transformations ($F : V \to V$ such that $F^*\omega = \omega$, or $\omega(F(\mathbf{x}, \mathbf{y})) = \omega(\mathbf{x}, \mathbf{y})$, so that F is also called symplectic) form a group, called the symplectic group $\mathrm{Sp}(V, \omega)$. If $V = \mathbb{C}^n$, this is sometimes also written $\mathrm{Sp}(2n)$. The group of matrices that preserve the Hermitian operator is called the unitary group $\mathrm{U}(n)$, and those that preserve the Euclidean structure are called orthogonal, $\mathrm{O}(2n)$, so that $\mathrm{U}(n) = \mathrm{O}(2n) \cap \mathrm{Sp}(2n)$ by the above. Thus, acting on $\mathbb{R}^n$ by any unitary operator gives an isotropic subspace.

Now let's throw in some differential geometry; instead of thinking of ω as a bilinear form, think of it as a differential form (i.e. a bilinear form originating at any point). Then, $\omega = \sum_{i=1}^n \mathrm{d}x_i \wedge \mathrm{d}y_i$: this is the same form, but parallel-transported, so that it can be applied to vectors originating anywhere.

Definition. If $\Omega_1, \Omega_2 \subset \mathbb{R}^{2n}$ (as a symplectic space) and $f : \Omega_1 \to \Omega_2$ is a diffeomorphism[28] but not necessarily linear, then f is called a symplectic diffeomorphism or symplectomorphism if it preserves the symplectic form: $f^*\omega = \omega$, or $\omega(\mathrm{d}_{\mathbf{a}} f(\mathbf{x}), \mathrm{d}_{\mathbf{a}} f(\mathbf{y})) = \omega(\mathbf{x}, \mathbf{y})$ for all $\mathbf{a} \in \Omega_1$ and $\mathbf{x}, \mathbf{y} \in V_{\mathbf{a}}$.

Theorem 11.2 (Darboux)**.** *Let ω be a nondegenerate (i.e. nonzero everywhere), closed (i.e. $\mathrm{d}\omega = 0$) differential 2-form on some domain U. Then, for any $\mathbf{a} \in U$, there exists a neighborhood $\Omega \supset U$ of $\mathbf{a}$ and a symplectomorphism $F : (\Omega, \omega) \to (\tilde{\Omega} \subset \mathbb{R}^{2n}, \omega_{\mathrm{st}} = \sum \mathrm{d}x_i \wedge \mathrm{d}y_i)$.*

This theorem, which will be proven later, implies that $F^*\omega_{\mathrm{st}} = \omega$, so F defines coordinates. Thus, this theorem implies that in some neighborhood of $\mathbf{a}$, there are some coordinates in which ω has the canonical form. This is the complete analogue of the linear case: the symplectic form is still very unique locally.

This can be globalized:[29] let M be a $2n$-dimensional manifold and ω be a closed, nondegenerate differential 2-form. Then, (M, ω) is called a symplectic manifold, and Darboux's Theorem says there are local coordinates, called Darboux coordinates, in which M looks like $\mathbb{R}^{2n}$ with the standard symplectic form.

If $n = 1$, every differential 2-form is of the form $f(x, y)\,\mathrm{d}x \wedge \mathrm{d}y$, and is therefore trivially closed, and Darboux's Theorem says this can be reformulated as $\mathrm{d}x \wedge \mathrm{d}y$. This leads to the notion of area. Similarly, a symplectic structure in four-dimensional space allows integrating the symplectic form over any two-dimensional surface.

One important example of a symplectic manifold is $M = T^*N$, where N is some n-dimensional manifold. If $q_1, \dots, q_n$ are local coordinates on N and $p_1, \dots, p_n$ are dual coordinates on the cotangent fibers, then let $\lambda = \sum p_i\,\mathrm{d}q_i$, which is a 1-form. Then, $\mathrm{d}\lambda = \sum \mathrm{d}p_i \wedge \mathrm{d}q_i$ is the symplectic form, so on the phase space of any mechanical system there is a canonical symplectic structure (i.e. λ can be defined without any coordinates).

Recall that the Hamilton function is $H : T^*M \to \mathbb{R}$, with $\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}$, $\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}$, and $\omega = \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{q}$ (in coordinates). If $\mathbf{X}_H$ is the vector field associated with these coordinates, then $\mathbf{X}_H \lrcorner \omega = \mathrm{d}H$. As a sanity check, if $\mathbf{X}_H = a\frac{\partial}{\partial \mathbf{p}} + b\frac{\partial}{\partial \mathbf{q}}$, then $a\,\mathrm{d}\mathbf{q} - b\,\mathrm{d}\mathbf{p} = \mathbf{X}_H \lrcorner\, \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{q} = -\frac{\partial H}{\partial \mathbf{p}}\,\mathrm{d}\mathbf{p} + \frac{\partial H}{\partial \mathbf{q}}\,\mathrm{d}\mathbf{q}$, so $a = -H_{\mathbf{q}}$ and $b = H_{\mathbf{p}}$, which is exactly what was given for $\mathbf{X}_H$.

[28] That is, it is at least C^1-smooth, but one can assume C^∞ for now; it is invertible; and f^{-1} is also at least C^1-smooth.

[29] If you wish to globalize literally as well as metaphorically, let $M = S^{2n}$.

12. The Lie Derivative and the Poisson Bracket: 2/22/13

Let N be the configuration space of some mechanical system, and T^*N be the phase space. On T^*N, there is a canonical 1-form $\lambda = \mathbf{p}\,\mathrm{d}\mathbf{q}$ (with the notation chosen because in any set of coordinates $q_1, \dots, q_n$ on N and the induced coordinates $p_1, \dots, p_n$ on T^*N, $\lambda = \sum p_i\,\mathrm{d}q_i$). Then, mechanics is defined by a function $H : T^*N \to \mathbb{R}$. This defines a vector field

$$\mathbf{X}_H = \sum_{i=1}^{n} -\frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i} + \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i}, \tag{5}$$

and trajectories correspond to integral curves of the system (so that $\dot{p}_i = -\frac{\partial H}{\partial q_i}$ and $\dot{q}_i = \frac{\partial H}{\partial p_i}$). In order to get an equation in this form, it's not just unnecssary to choose specific $\mathbf{p}, \mathbf{q}$-coordinates, but in fact one can use much more general coordinates than the induced ones, since it's only necessary to know $\omega = \mathrm{d}\lambda = \sum \mathrm{d}p_i \wedge \mathrm{d}q_i$.[30]

Thus, if $\mathbb{R}^{2n}$ has coordinates $q_1, \dots, q_n, p_1, \dots, p_n$, then there is an invariant way to obtain the Hamilton canonical equations given $H(\mathbf{p}, \mathbf{q})$. Given a nondegenerate bilinear form Φ, one has an isomorphism $V \cong V^*$ given by $\mathbf{v} \mapsto \mathbf{v} \lrcorner \Phi$; since this is a nondegenerate linear function between spaces of the same dimension, it is an isomorphism. Thus, $\mathbf{v} \mapsto \mathbf{v} \lrcorner \omega$ induces an isomorphism $\mathrm{Vect}(U) \to \Omega^1(U)$ (the spaces of vector fields and differential forms, respectively). Denote by I the inverse of this map, which allows one to obtain vector fields from differential forms.

If $n = 1$, $\mathbf{v} = a\frac{\partial}{\partial p} + b\frac{\partial}{\partial q}$, so $\mathbf{v} \lrcorner \mathrm{d}p \wedge \mathrm{d}q = a\,\mathrm{d}q - b\,\mathrm{d}p$ (the minus sign occurs because the order is reversed). Thus, given an H, take $\mathrm{d}H = H_p\,\mathrm{d}p + H_q\,\mathrm{d}q$, so the corresponding vector field is $\mathbf{v} = H_q\frac{\partial}{\partial p} - H_p\frac{\partial}{\partial q}$. This is identical to the equations above, up to sign, so $\mathbf{X}_H = -I(\mathrm{d}H)$. Note that in higher dimensions this argument is essentially the same, with the addition of indices. Be warned, however: there is no way to choose signs in a reasonable, natural way. Different books tend to have different sign conventions; for example, Arnold defines contraction by placing the vector in the second argument, which means that $\mathbf{X}_H = I(\mathrm{d}H)$.

Thus, in order to define Hamiltonian mechanics, the only necessary ingredients are the phase space M with some closed nondegenerate differential 2-form ω. M doesn't need to be the cotangent bundle of anything, but can be any symplectic manifold, and indeed such examples appear in applications. For example, if one wants to find integrals of a system, a method called reduction sends a $2n$-dimensional manifold to a $(2n-2)$-dimensional one (which might not be a cotangent bundle).

Dynamics are determined by some function $H : M \to \mathbb{R}$, the full energy of the system.[31] Then, mechanics is determined by the vector field $\mathbf{X}_H = -I(\mathrm{d}H)$, or $\mathbf{X}_H \lrcorner \omega = -\mathrm{d}H$. Then, if $\mathbf{u}(t)$ is a trajectory, then $\dot{\mathbf{u}}(t) = \mathbf{X}_H(\mathbf{u}(t))$.

A mechanical system has degrees of freedom, usually defined as the dimension of the configuration space; without the configuration space, they can be defined as half the dimension of the phase space. If a system has 1 degree of freedom, then the phase space is a surface, so the symplectic form is an area form $f(p,q)\,\mathrm{d}p \wedge \mathrm{d}q$. Suppose A is a closed surfact or a compact region with boundary and ω_1, ω_2 are two area forms with the same total area (so that $\int_A \omega_1 = \int_A \omega_2$). Then, there exists a diffeomorphism $f : A \to A$ such that $f^*\omega_1 = \omega_2$. Thus, the area form is unique, though this is very untrue of the symplectic form in higher dimensions.

Consider a smooth vector field $\mathbf{X}$ on a manifold M (we'll be interested particularly in the Hamiltonian vector field). If the manifold is compact or has the appropriate boundary conditions, then the vector field integrates to a phase flow $\mathbf{X}^t : M \to M$, a 1-parametric group of diffeomorphisms, such that $\frac{\mathrm{d}\mathbf{X}^t(x)}{\mathrm{d}t} = \mathbf{X}(\mathbf{X}^t(x))$. Locally, this flow is always defined, and globally it depends on the conditions mentioned above (since otherwise, it could go to infinity). Then, finding the phase flow is equivalent to solving the differential equation $\dot{\mathbf{u}} = \mathbf{X}(\mathbf{u})$ for arbitrary initial data.

M can be thought of as a compelx manifold, in which case $\mathbf{X}_H = i\nabla H$. Thus, $\mathbf{X}_H$ is tangent to the level sets of H, so all dynamics lie inside the level sets. This implies that energy is conserved.

Returning to $\mathbb{R}^2$ with $\mathrm{d}p \wedge \mathrm{d}q$, consider $H(p,q) = (p^2+q^2)/2$, so that $\dot{p} = -q$ and $\dot{q} = p$, so that the phase flow is just rotations. If $H(p,q) = p^2 + 2q^2$, one has ellipses (with the major axis in the q-direction). Consider the transformation $(p,q) \mapsto (\lambda p, q/\lambda)$ for some $\lambda > 0$; this is a canonical transformation, because $\mathrm{d}p \wedge \mathrm{d}q$ is unchanged. Thus, the second system can be changed into $\lambda^2\tilde{p}^2 + 2\tilde{q}^2/\lambda^2$ (setting $p = \lambda\tilde{p}$ and $q = \tilde{q}/\lambda$). Then, if $\lambda = 4\sqrt{2}$, this is just the first system, $H(p,q) = p^2 + q^2$, up to a scaling factor. Thus, the system can be easily solved and passed back to the original coordinates. This is an example of the power of the Hamiltonian formalism; the second system can in fact be explicitly solved, but in general this might not be the case.

Suppose $\mathbf{X}$ is a vector field. If f is any function, one can take directional derivatives of the function as

$$\mathrm{d}_{\mathbf{a}}f(\mathbf{x}) = \lim_{t\to 0}\frac{f(\mathbf{a}+t\mathbf{x}) - f(\mathbf{a})}{t}.$$

[30] In order to clear up something potentially confusing, $\sum p_i \wedge q_i$ is a bilinear form on $\mathbb{R}^n$, which accepts vectors that originate at the origin, but $\sum \mathrm{d}p_i \wedge \mathrm{d}q_i$ is the same thing, except that since it's a differential form, the vectors can originate anywhere.

[31] Systems are assumed to be conservative; otherwise H would depend on time.

This moves points linearly, but one also can take the flow $\mathbf{X}^t$, which provides the alternate definition of $\frac{\partial}{\partial t} f(\mathbf{X}^t(\mathbf{u}))|_{t=0}$. This is the formal definition of the Lie derivative, and is a viewpoint used in some other situations (such as quantum mechanics). Though these formulas are equivalent, the latter one is a much better way to go about things, because it can differentiate any more objects.[32] The Lie derivative can also be given by $\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{X}^t)^* f$.

Similarly, if ω is a differential form, the Lie derivative is $L_{\mathbf{X}}\omega = \frac{\partial}{\partial t}((\mathbf{X}^t)^*\omega)|_{t=0}$, and the Lie derivative for a vector field can be defined in an analogous way. Differential forms in particular obey something called the Cartan formula, as seen in Math 53H: if $\mathbf{X}$ is a vector field and ω a differential form, then $L_{\mathbf{X}}\omega = \mathbf{X} \lrcorner\, \mathrm{d}\omega + \mathrm{d}(\mathbf{X} \lrcorner \omega)$.

If $\mathbf{Y}$ is another vector field, then $L_{\mathbf{X}}\mathbf{Y} = [\mathbf{X}, \mathbf{Y}]$, the Lie bracket of vector fields. If $\mathbf{X} = \sum f_i \frac{\partial}{\partial x_i}$ and $\mathbf{Y} = \sum g_i \frac{\partial}{\partial x_i}$, then

$$[\mathbf{X}, \mathbf{Y}] = \sum_{i,j} \left(f_i \frac{\partial g_j}{\partial x_i} - g_i \frac{\partial f_j}{\partial x_i} \right).$$

Thus, if f is any function, then $L_{\mathbf{X}} L_{\mathbf{Y}} f - L_{\mathbf{Y}} L_{\mathbf{X}} f = L_{[\mathbf{X},\mathbf{Y}]} f$, so this is a first-order differential operator, despite looking like a second-order one.

If ω is the symplectic form and $\mathbf{X}_H$ is the Hamiltonian vector field for some function H, then $L_{\mathbf{X}_H}\omega = \frac{\partial}{\partial t}(\mathbf{X}_H^t)^*\omega = \mathbf{X}_H \lrcorner\, \mathrm{d}\omega + \mathrm{d}(\mathbf{X}_H \lrcorner \omega)$. Since ω is closed, then $\mathrm{d}\omega = 0$, and $\mathbf{X}_H \lrcorner \omega = -\mathrm{d}H$, so $L_{\mathbf{X}_H}\omega = 0$, since $\mathrm{d}(\mathrm{d}H) = 0$. Thus, $(\mathbf{X}_H^t)^*\omega = \omega$, which is referred to as the Poincaré integral invariant, and implies that the symplectic form is preserved by the Hamiltonian flow. This is a useful fact, and on its own can be used to prove a lot of nontrivial things.

Corollary 12.1 (Liouville's Theorem)**.** *Define the phase volume to be* $\Omega = \mathrm{d}p_1 \wedge \mathrm{d}q_1 \wedge \cdots \wedge \mathrm{d}p_n \wedge \mathrm{d}q_n$*. Then,* $\omega^n = n!\Omega$*, which implies that* $(\mathbf{X}_H^t)^*\Omega = \Omega$*, so the Hamiltonian flow preserves the phase volume.*

This is a much cruder fact, and was known before Poincaré's.

Thus, if $\mathbf{X}$ is a volume-preserving vector field (i.e. it is divergence-free, so $\operatorname{div} \mathbf{X} = 0$) and the phase space is compact, then one has a very nontrivial fact called the Poincaré Recurrence Theorem: every trajectory returns arbitrarily closely to its starting point. In other words, if U is a neighborhood, then $(\mathbf{X}^{t_1})(U) \cap (\mathbf{X}^{t_2})(U)$ is nonempty for some t_1 and t_2, since volume is preserved. This has the curious consequence that if someone takes a box of gas molecules and releases them into a room, they will eventually return to the box, though after an extremely long time.

The Lie bracket has a lot of nice properties: it is linear in each argument, and it is skew-symmetric: $[\mathbf{X}, \mathbf{Y}] = -[\mathbf{Y}, \mathbf{X}]$.[33] Additionally, it satisfies something called the Jacobi identity: $[[\mathbf{X}, \mathbf{Y}], \mathbf{Z}] + [[\mathbf{Y}, \mathbf{Z}], \mathbf{X}] + [[\mathbf{Z}, \mathbf{X}], \mathbf{Y}] = 0$. In algebra, a vector space with a bilinear, skew-symmetric function that satisfies the Jacobi identity is called a Lie algebra, so these vector fields are a Lie algebra. These are deep connections to Lie groups, but that is beyond the scope of this class.

If (M, ω) is a $2n$-dimensional symplectic manifold, then there is a function on the space of functions $M \to \mathbb{R}$ called the Poisson bracket which gives the space of functions a Lie algebra structure: if $f, g : M \to \mathbb{R}$, then $[\mathbf{X}_f, \mathbf{X}_g] = \mathbf{X}_h$ for some other function h, since the Lie bracket of two Hamiltonian fields is also Hamiltonian. Then, the Poisson bracket is $h = \{f, g\}$. In coordinates, the Poisson bracket is given by

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i}.$$

There is yet another definition for the Poisson bracket: $\{f, g\} = \mathbf{X}_f \circ g = \mathrm{d}_g(\mathbf{X}_f)$, which can be thought of as the Lie derivative of the second function with respect to the first, at least up to sign. Since $\mathbf{X}_f \lrcorner\, \omega = -\mathrm{d}f$, then $\omega(\mathbf{X}_f(\mathbf{y})) = -\mathrm{d}f(\mathbf{y})$ for any vector $\mathbf{y}$; thus, $\{f, g\} = \omega(\mathbf{X}_f, \mathbf{X}_g)$.

Suppose $\{f, H\} = 0$. Then, $L_{\mathbf{X}_H} f = 0$, so f is constant on the trajectories of $\mathbf{X}_H$. Such functions are called integrals of motion. Thus, the integrals are precisely those functions which Poisson-commute with the Hamiltonian!

In the general case, if (M, ω) is symplectic and $f : M \to \mathbb{R}$ (called observable in quantum mechanics), one can look at how the function evolves on trajectories as $\{H, f\} = \frac{\partial}{\partial t}(f(\mathbf{X}_H(t))$, which corresponds to some measurement. Thus, one has the equation $\dot{f} = \{H, f\}$, so taking a coordinate function p_i for f corresponds to observing how the momentum changes: $\dot{p}_i = \{H, p_i\} = -H_{q_i}$.

13. Lagrangian Submanifolds: 2/25/13

First, a slight clarification from a previous lecture, the action functional is $S(\gamma) = \int_\gamma \mathbf{p}\,\mathrm{d}\mathbf{q} - H\,\mathrm{d}t$, and this is minimized over all paths such that $\mathbf{p}\,\mathrm{d}\mathbf{q}$ vanishes on the boundary fibers. This is very different from the Lagrangian because one doesn't really have to worry about the phase space.

[32] In French, the Lie derivative is also known as the fisherman's derivative, suggesting a fisherman who sits on his boat in one place and differentiates things as they flow by.

[33] Again, this leads to more weird sign conventions.

The Euler-Lagrange equations have the form $\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}$ and $\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}$, which induces the vector field $\mathbf{X}_H$ in (5). Then, if $\omega = \mathrm{d}(\mathbf{p}\,\mathrm{d}\mathbf{q})$ is the symplectic form, one has $\mathbf{X}_H \lrcorner \omega = -\mathrm{d}H$, which can actually be thought of as the definition of $\mathbf{X}_H$.

All of this structure seems like overkill, so it's possible to just take a $2n$-dimensional manifold M and a closed, nondegenerate differential 2-form ω. By Darboux's Theorem (i.e. Theorem 11.2), this form is locally unique up to change of coordinates (and the uniqueness of the skew-symmetric form is a corollary). Then, we can take any function $H : M \to \mathbb{R}$ and set up a vector field with $\mathbf{X}_H \lrcorner \omega = -\mathrm{d}H$.

Locally, this vector field integrates to a phase flow, and it integrates globally given appropriate boundary conditions (or if M is compact). Let $\mathbf{X}_H^t : M \to M$ be thbe phase flow; it allows differentiation of any object. If one wishes to differentiate something along a path, differentiating functions is nice, but (for example) differentiating vector fields depends on the embedding of the manifold into Euclidean space, which is extra information. Thus, one can instead use the Lie derivative, which avoids subtraction because it sits in one place.

Using the Cartan formula, $L_{\mathbf{X}_H}\omega = 0$, which means that $(\mathbf{X}_H^t)^*\omega$ is constant; since it is equal to ω at zero, then $(\mathbf{X}_H^t)^*\omega = \omega$. Thus, the Hamiltonian flow preserves the symplectic form (e.g. area when $\dim M = 2$).

Suppose $\mathbf{Y}$ is a vector field such that $L_{\mathbf{Y}}\omega = 0$. Then, by the Cartan formula, $\mathrm{d}(\mathbf{Y} \lrcorner \omega) = 0$, so this is equivalent to the 1-form $\mathbf{Y} \lrcorner \omega$ being closed, and therefore exact: locally, every closed form is exact, and on a star-shaped domain the converse is true. Thus, locally, $\mathbf{Y} \lrcorner \omega = -\mathrm{d}H$ for some H, so $\mathbf{Y}$ is locally a Hamiltonian flow.

Example 13.1. Let $M = T^2 = \mathbb{R}^2/\mathbb{Z}^2$ be the 2-dimensional torus, and let p and q be the Cartesian coordinates in $\mathbb{R}^2$ (and therefore also on M). Thus, $p\,\mathrm{d}q$ is invariant under translation. Let $\mathbf{Y} = \frac{\partial}{\partial p}$, which is obviously area-preserving, so $\frac{\partial}{\partial p} \lrcorner \omega = \mathrm{d}q$. Since q is a cyclic coordinate, but not a function, then $\mathrm{d}q$ isn't exact.

A vector field that preserves ω is called symplectic, and one for which $\mathbf{Y} \lrcorner \omega$ is exact is called Hamiltonian. Thus, all Hamiltonian vector fields are symplectic, but not vice versa. In simply-connected spaces, however, the two are equivalent.

Returning to the torus, choose a meridian and move it in the flow. The signed area should be 0 between the torus and its image (as the flow moves over the whole torus), and is called flux. (Again, in higher dimensions this becomes more complicated.)

Let (M, ω) be a $2n$-dimensional symplectic manifold and $L \subset M$ be a submanifold. L is called isotropic if $\omega|_L = 0$ (i.e. for all $\mathbf{a} \in L$ and $\mathbf{X}, \mathbf{Y} \in T_{\mathbf{a}}L$, $\omega_{\mathbf{a}}(\mathbf{X}, \mathbf{Y}) = 0$). This is similar to the discussion of isotropic subspaces previously, and is in fact just the generalization to submanifolds: a submanifold is isotropic if all of its tangent spaces are isotropic subspaces. Similarly, L is Lagrangian if L is isotropic and $\dim L = n$.

If $L = \{\mathbf{y} = A\mathbf{x}\}$, then L is Lagrangian iff A is symmetric in the linear case. In the more general case, consider $(\mathbb{R}^{2n}, \omega)$ and a Lagrangian submanifold L that is also graphical, so $L = \{\mathbf{p} = \mathbf{F}(\mathbf{q})\}$, where $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$. Then,

$$\omega|_L = \sum_{i=1}^{n} \mathrm{d}F_i \wedge \mathrm{d}q_i = \mathrm{d}\left(\sum_{i=1}^{n} F_i\,\mathrm{d}q_i\right) = 0,$$

where $\mathbf{F} = (F_1, \ldots, F_n)$. Thus, $\sum F_i\,\mathrm{d}q_i$ is closed, so it is exact (since $\mathbb{R}^{2n}$ is well-behaved), or $F_i = \frac{\partial f}{\partial q_i}$ for some f. Thus, Lagrangian submanifolds which are graphical are graphs of gradients of functions.

If $M = T^*N$ with $\omega = \mathrm{d}(\mathbf{p}\,\mathrm{d}\mathbf{q})$, then graphical also means that the projection onto N is one-to-one. Such manifolds are called sections. But this is just a 1-form — each point is associated with a covector (i.e. a linear functional). But this 1-form must be closed if it is graphical Lagrangian: $\alpha = \mathbf{F}(\mathbf{q})\,\mathrm{d}\mathbf{q}$, so $\mathrm{d}\alpha = \omega = 0$. Thus, the graphs of closed forms are Lagrangian, and in particular are the differentials of functions.

Globally, of course, a Lagrangian manifold need not be graphical. If M is 2-dimensional, so that if L is a Lagrangian submanifold, then L is 1-dimensional. Then, any curve is Lagrangian, because any 2-form on a 1-dimensional manifold is 0. Oops. If M is 4-dimensional, with coordinates $\mathbf{q} = (q_1, q_2)$, $\mathbf{p} = (p_1, p_2)$, then it happens that $(\mathbb{R}^4, \omega = \mathrm{d}p_1 \wedge \mathrm{d}q_1 + \mathrm{d}p_2 \wedge \mathrm{d}q_2) = (\mathbb{R}^2, \mathrm{d}p_1 \wedge \mathrm{d}q_1) \oplus (\mathbb{R}^2, \mathrm{d}p_2 \wedge \mathrm{d}q_2)$, so the product of Lagrangian submanifolds in $\mathbb{R}^2$ is a Lagrangian submanifold in $\mathbb{R}^4$. For example, one obtains a torus as the product of two circles. Considering $\mathbb{R}^4 = \mathbb{C}^2$, a submanifold is Lagrangian if $L^{\perp\omega} = iL = L$, with the former equality because it's complex, and the latter because it is Lagrangian.

On $(\mathbb{R}^2, \mathrm{d}p \wedge \mathrm{d}q)$, consider an area-preserving transformation $f : \mathbb{R}^2 \to \mathbb{R}^2$, so that $f^*(\mathrm{d}p \wedge \mathrm{d}q) = \mathrm{d}p \wedge \mathrm{d}q$, or, equivalently, $\det(Df) = 1$. On the space $\mathbb{R}^2 \times \mathbb{R}^2$, consider the symplectic form $\Omega = \mathrm{d}p_1 \wedge \mathrm{d}q_1 - \mathrm{d}p_2 \wedge \mathrm{d}q_2$. Of course, this is isomorphic to the conventional symplectic form, but consider the graph $\Gamma_f = \{(p_2, q_2) = f(p_1, q_1)\}$. Then, Γ_f is Lagrangian, because $\Omega|_{\Gamma_f} = \mathrm{d}p_1 \wedge \mathrm{d}q_1 - f^*(\mathrm{d}p_1 \wedge \mathrm{d}q_1) = 0$. This is in fact a necessary and sufficient condition: a function is area-preserving iff its graph is a Lagrangian submanifold.

Thus, one can split by Lagrangian subspaces, in which case H is the graph of a differential, and splitting as the direct product, it is the graph of an area-preserving transformation. However, this only works in $\mathbb{R}^{4n}$: in general, in

$(\mathbb{R}^{2n}, \omega)$, if $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is a symplectomorphism, then take $\Omega = \omega \oplus -\omega$ on $\mathbb{R}^{4n}$, and $\Gamma_f = \{(\mathbf{p}, \mathbf{q}), f(\mathbf{p}, \mathbf{q})\} \subset \mathbb{R}^{4n}$ is Lagrangian iff $f^*\omega = \omega$.

Poincaré was what we would call today an applied mathematician: when working towards what would be called his last geometric theorem, he was considering it in the more general context of the 3-body problem. He wanted to show the existence of certain periodic orbits, which relates to finding fixed points of symplectic, area-preserving transformations, specifically over an annulus. The condition he considered was that the edges of the annulus rotate in opposite directions.[34]

Under these conditions, the map has at least two fixed points. Poincaré proved this in a special case, but it was the last result he proved, and it arguably killed him. A solution was found shortly after he died, but it was a very specific trick that doesn't apply to the general field, and it set work back in that field by 50 years.

The basic ingredient of the proof is the following: associate to an area-preserving transformation its graph; then, fixed points correspond to intersections with the diagonal. But each of these is given by the differential of some functions $\mathrm{d}f$ and $\mathrm{d}g$, so the question boils down to finding the critical points of $f - g$. This leads to Morse theory, among other things.

14. The Last Geometric Theorem of Poincaré: 2/27/13

Consider again a symplectic manifold (M^{2n}, ω) with a Lagrangian submanifold $L \subset M$ (so that $\omega|_L = 0$ and $\dim L = n$), which is precisely the condition that $T_{\mathbf{a}}L$ is a Lagrangian subspace for all $\mathbf{a} \in L$. If L is graphical, given by $p_i = f_i(\mathbf{q})$, then the form $\sum f_i(\mathbf{q})\,\mathrm{d}\mathbf{q}$ is exact (and therefore closed in $\mathbb{R}^{2n}$), so it is equal to $\mathrm{d}F$ for some F and $f_i = \frac{\partial F}{\partial q_i}$.

Definition. Suppose $(V, \omega) = (R^{2n}, \sum \mathrm{d}p_i \wedge \mathrm{d}q_i)$ and L_1 and L_2 are two Lagrangian sub*spaces* with $L_1 \pitchfork L_2$. Then, by simple algebra, there is a symplectomorphism $S : V \to V$ such that $S(L_1) = \{\mathbf{p} = \mathbf{0}\}$ and $S(L_2) = \{\mathbf{q} = \mathbf{0}\}$. Such a pair L_1, L_2 is called a polarization of V.

The fact that a symplectic manifold tends to have lots of polarizations is sometimes useful.

Example 14.1. If $\mathbb{R}^4$ has coordinates p_1, p_2, q_1, q_2, then (p_1, p_2) and (q_1, q_2) is a polarization, as is (p_1, q_2) and (q_1, p_2).

There is another interpretation: consider the plane with coordinates p, q and an area form $\mathrm{d}p \wedge \mathrm{d}q$. Suppose $f : \mathbb{R}^2 \to \mathbb{R}^2$ is area-preserving (i.e. the Jacobian of f is 1, or f preserves measure). Then, extend the space to have two more coordinates P and Q. The graph of f is given as a surface $\Gamma_f = (P(p,q), Q(p,q))$, and will be Lagrangian with respect to the form $\Omega = \mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}P \wedge \mathrm{d}Q = \mathrm{d}p \wedge \mathrm{d}q + \mathrm{d}Q \wedge \mathrm{d}P$.

Then, $\{\Omega|_{\Gamma_f} = 0\}$ iff f is area-preserving: since Γ_f is graphical, then p, q-coordinates can be used on it. Then, if $\Omega|_{\Gamma_f} = 0$, then $\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}P \wedge \mathrm{d}Q = 0$, but $f^*(\mathrm{d}p \wedge \mathrm{d}q) = \mathrm{d}P \wedge \mathrm{d}Q$, so the form vanishes iff f preserves area. In higher dimensions, one can play the sam edoubling-dimension game. Then, the graph is Lagrangian iff f is a symplectomorphism.

Returning to the last geometric theorem of Poincaré:

Theorem 14.1 (Poincaré-Birkhoff, 1912)**.** *Consider $\mathbb{R}^2$ with coordinates q, p and take the strip $\{0 \le p \le 1\} = A$. Let $F : A \to A$ be area-preserving and satisfying the equivariancy condition $F(p, q+1) = (P(p,q), Q(p,q)+1)$, where $F = (P, Q)$ in coordinates.*[35] *Suppose $Q(1, q) > q$ and $Q(0, q) < q$ (called the twist condition). Then, F has a fixed point, which implies that on the annulus it has two fixed points.*

Poincaré's partial proof of Theorem 14.1. Assume F is C^1-small (which is the reason this is a partial proof). Then, extend F to the whole plane: in the upper half of the plane, let $Q(p, q) = q + C_1$ (where $p \ge 1$), and in the lower half, let $Q(p, q) = q - C_2$ (where $p \le 0$). Then, take the graph Γ_f with P and Q as above and the form $\mathrm{d}p \wedge \mathrm{d}q + \mathrm{d}Q \wedge \mathrm{d}P = \mathrm{d}(p\,\mathrm{d}q + Q\,\mathrm{d}P)$. The condition that F is C^1-small is that $P(p, q) = p + \varphi(p, q)$ and $Q(p, q) = q + \psi(p, q)$, with φ and ψ C^1-small (i.e. they don't deviate from the identity by very much). Then, let $\Delta = \{P = p, Q = q\}$ be the diagonal, so that Γ_f doesn't deviate from the diagonal very much, and neither do its tangent lines.

Take another presentation of $\mathbb{R}^4$ as $(q, P) \oplus (p, Q)$, and call it $(\widetilde{q}, \widetilde{Q}) \oplus (\widetilde{p}, \widetilde{P})$. In this polarization, $\omega = \mathrm{d}(p\,\mathrm{d}q + P\,\mathrm{d}Q)$. Since Δ is graphical with respect to this splitting, then Γ_f is also graphical, since it doesn't deviate very much from Δ. Thus, it is possible to write

$$\Gamma_f = \begin{cases} \widetilde{p} = A(\widetilde{q}, \widetilde{Q}) \\ \widetilde{P} = B(\widetilde{q}, \widetilde{Q}). \end{cases}$$

[34] Since rotation in one direction around a circle is equivalent to rotation in the opposite direction by a different angle, this is ambiguous, but can be made more precise, as will be shown.

[35] This condition arises so that F is well-defined on the annulus $A/\mathbb{Z}$.

Since Γ_f is Lagrangian, then $\widetilde{P}\,\mathrm{d}\widetilde{Q}+\widetilde{p}\,\mathrm{d}\widetilde{q}$ is exact, so there exists an $H:(\widetilde{Q},\widetilde{q})\to(\widetilde{P},\widetilde{p})$ such that $B\,\mathrm{d}\widetilde{Q}+A\,\mathrm{d}\widetilde{q}=\mathrm{d}H$.

We want to show that F has a fixed point, which is equivalent to an intersection with the graph of $\widetilde{p}=\widetilde{Q},\widetilde{q}=\widetilde{P}$, which is also Lagrangian. Thus, it is the graph of some function $G:(\widetilde{Q},\widetilde{q})\to(\widetilde{P},\widetilde{p})$ such that $G_{\widetilde{Q}}=\widetilde{P}$ and $G_{\widetilde{q}}=\widetilde{p}$.

Thus, $A=\widetilde{q}+s$, where s is something small, and similarly with B and $\widetilde{Q}$, and $G(\widetilde{Q},\widetilde{q})=\widetilde{q}\widetilde{Q}$.[36] Taking the two differential forms $A\,\mathrm{d}\widetilde{q}+B\,\mathrm{d}\widetilde{Q}$ and $\widetilde{Q}\,\mathrm{d}q+\widetilde{q}\,\mathrm{d}Q=\beta=\mathrm{d}G$, then a fixed point is equivaent to these coefficients coinciding. Thus, let $\alpha=\mathrm{d}H$, so that $\alpha-\beta=(A-\widetilde{Q})\,\mathrm{d}\widetilde{q}+(B-\widetilde{q})\,\mathrm{d}\widetilde{Q}=\mathrm{d}(H-G)=\mathrm{d}(H-\widetilde{q}\widetilde{Q})$. Since the goal is to get both partial derivatives to be zero, the goal is to find critical points of $H(\widetilde{q},\widetilde{Q})-\widetilde{q}\widetilde{Q}=H(q,P)-qP$, and the equations of motion give $\frac{\partial H}{\partial q}=p(q,P)$ and $\frac{\partial H}{\partial P}=Q(q,P)$.

If $P>1$, then $(p,q)\mapsto(p,q+c)=(P,Q)$, so $\frac{\partial H}{\partial q}=P$ and $\frac{\partial H}{\partial P}=q+C$, so $H=P(q+C)$, so $H-qP=Pc$ for some constant c. Similarly, downstairs $H-qP=-Pc$.

Claim. $T=H(q,P)=qp$ is periodic with respect to q: $T(q+1,P)=T(q,P)$.

Proof. The derivative is periodic, since $\frac{\partial H}{\partial q}=P$ and P is known to have period 1, so over a period, exactly the same constant is added to T. Since this is a constant that is known to be zero sufficiently far from the x-axis, then it must be zero everywhere. ⊠

Thus, quotient by q and define this function on the cylinder. T goes to $+\infty$ on both extremes, so it must have a minimum between 0 and 1. Thus, it has a critical point. ⊠

There is also a saddle point: tak any embedding of the circle, and take the maximum. Then, the minima of all such maxima can be shown to be a saddle point.

Poincaré didn't know what to do if F wasn't graphical or C^1-small. The general case is in fact extremely complicated. It's in fact related to a set of conjectures called the Arnold conjectures, some of which were proven quite recently.

15. Integrable Systems: 3/4/13

Consider a $2n$-dimensional symplectic manifold (M,ω) as before. Since ω determines the equations of motion, then the presence of lots of symplectomorphisms implies lots of possible coordinates. For example, if H only depends on p or on q in one system of coordinates, then the other coordinates are all cyclic, and the system can easily be solved. Thus, learning the structure of the group of canonical transformations is useful.

For simplicity, assume M is compact and let $\mathbf{X}_H$ denote the Hamiltonian vector field. Then, $(\mathbf{X}_H)^*\omega=\omega$, or $L_{\mathbf{X}_H}\omega=0$. This implies that a lot of symplectomorphisms exist. In some spaces (e.g. simply connected ones), any symplectomorphism that is isotopic to the identity is Hamiltonian. If a transformation f is generated by a Hamiltonian flow $\mathbf{X}_H$ and f has a fixed point at $\mathbf{a}$, then $\mathbf{X}_H$ has a critical point at $\mathbf{a}$.

Returning to the twist map from the previous lecture, if one assumes that f can be generated by a Hamiltonian function, then the Hamiltonian must be constant on the boundary, and based on the gradients, there must exist a minimum and a saddle point, so the two necessary fixed points exist. However, the more general proof involved generating functions: a Lagrangian manifold that is graphical with respect to a certain splitting can be given as the graph of the differential of a generating function.

The first known proof of Theorem 14.1 used a cheap trick that actually set the theory back for a long time, since it doesn't work anywhere else.

Proof of Theorem 14.1. Suppose the transformation f doesn't have a fixed point, and consider the vector field consisting of vectors from x to $f(x)$ (similar to the proof of Brower's Theorem), which nowhere vanishes. If one takes any curve from top to bottom, following the vector field, it rotates around the annulus by an angle of $(2k+1)\pi$; since it's homotopic, then k must vary continuously, but $k\in\mathbb{Z}$, so it's constant.

Consider the transformation $T_\varepsilon(p,q)=(p+\varepsilon,q)$ (i.e. shifting up by ε). If ε is small, take $f\circ T_\varepsilon=f_\varepsilon$ and, working in the plane, take the strip U of size ε below the bottom of the big strip corresponding to the annulus. Then, let $U_1=f_\varepsilon(U)$ and iterate, setting $U_2=f_\varepsilon(U_1)$, etc. None of these domains intersect, since the upper boundary of one becomes the lower boundary of the other. Since this transformation is periodic and area-preserving, there must be some point x where one of the U_i "sticks out" above the annulus. Then, take the sequence of points $x_0\in U$, $x_1=f_\varepsilon(x_0),\dots x_n=x$. Thus, this vector field is homotopic to a tangent field, and it is a theorem in topology that any such embedded curve is isotopic to an interval. Thus, the rotation was by an angle π, so $k=1$ (since the rotation was counterclockwise).

Doing the same thing with $f_{-\varepsilon}(U)$ shows that $k=-1$, which is a contradiction. ⊠

[36] QQ sounds about right... no wonder this theorem killed Poincaré.

Moving into a more classical subject, it will be worthwhile to study integrable systems. Here, the goal is to show that an integral exists, not necessarily that it's an elementary function. Integrable systems are among the most important aspects of classical mechanics. It's a slightly vague notion, but there's a nice definition for formality.

Definition. Let (M, ω) be the phase space of a mechanical system with Hamiltonian $H : M \to \mathbb{R}$. Then, $F : M \to \mathbb{R}$ is integrable iff $\mathrm{d}F(\mathbf{X}_H) = \mathbf{X}_H \circ F = \omega(\mathbf{X}_F, \mathbf{X}_H) = \{F, H\} = 0$. (Specifically, all of these quantities are identical, and the point is that it vanishes.)

Definition. If $\{F, G\} = 0$, where $\{\}$ is the Poisson bracket, F and G are said to Poisson-commute. Sometimes, they are also referred to as an involution.

This is nice, because when F is the Hamiltonian, the second definition implies the first. However, there is room for stupidity, too; you could claim n integrals, all equal to H. Thus, definitions also tend to require the linear independence of all of the ∇F_i.

Theorem 15.1. *Suppose (M, ω, H) is Liouville-integrable. Then, there exist canonical coordinates $(I_1, \dots, I_n, \varphi_1, \dots, \varphi_n)$ (i.e. such that $\omega = \sum \mathrm{d}I_j \wedge \mathrm{d}\varphi_j$) such that in these coordinates, $H = H(I)$ is independent of φ, so that all of the I_j are integrals. These coordinates are called action-angle coordinates.*

Hence, motion occurs on the intersections of the level sets of the I_j, and thus on an n-dimensional submanifold, since they are linearly independent. In other words, a level set $T_{\mathbf{c}} = \{I_1 = c_1, \dots, I_n = c_n\}$ is preserved by the Hamiltonian flow. If $T_{\mathbf{c}}$ is compact (e.g. when M is), then it is diffeomorphic to an n-dimensional torus, and on this torus the equations of motion have the form $\frac{\mathrm{d}\varphi_j}{\mathrm{d}t} = \omega_j(\mathbf{c})$ (i.e. constant when $\mathbf{c}$ is fixed). This type of motion is called quasi-periodic, since the torus is the quotient of $\mathbb{R}^n$ by a lattice, so everything can be lifted to $\mathbb{R}^n$, where the motion is linear. The motion will be periodic iff the slope is rational.

If the system has one degree of freedom, the Hamiltonian is an integral, so the system is integrable, and thus it lies on a circle. If there are two degrees of freedom, then one extra integral is needed. If the phase space is compact, one obtains tori (sometimes circles if they're not linearly independent), and the slope varies continuously.

This is interesting because systems in general aren't integrable; integrable systems can almost be thought of as degenerate.

This leads to a condition for non-integrability: if a trajectory is everywhere dense in some three-dimensional level set, it doesn't lie in the torus, so the system is non-integrable. This is how Poincaré solved the three-body problem.[37]

The real world isn't all that integrable, so why do we care? If something is not integrable, but close to some other integrable system (usually by adding some extra term), it can be viewed as a small perturbation of that integrable system, which is studied in Kolmogorov-Arnold-Moser (KAM) theory. It happens that if a system is almost integrable, then most of the tori are preserved. Specifically, in the space of tori, there is a nowhere-dense, positive-measure set of tori which aren't preserved, which imply chaotic motion. However, if the system (for given constraints) lies in between two tori, it can't escape, so the stability of the system can sometimes be deduced.

16. Liouville's Theorem: 3/8/13

Liouville's Theorem refers to several quite different results. In particular, the theorem that will be discussed in this lecture is distinct from Theorem 12.1.

Arnold's Principle. If a theorem is named after a mathematician M, then M was probably not the person responsible for the proof of the theorem.

Arnold's Principle applies to Liouville's Theorem, but also to itself, funnily enough.

Consider a symplectic manifold (M^{2n}, ω) with a Hamiltonian $H : M \to \mathbb{R}$. The Hamiltonian induces the phase flow $\mathbf{X}_H^t : M \to M$, which gives the dynamics of the system. Dynamics given by a vector field can be fairly complicated, but there are also lots of constraints; for example, the following theorem comes from the theory of ordinary differential equations.

Theorem 16.1 (Poincaré-Bendixson)**.** *If $\mathbf{X}$ is a nonvanishing vector field on a square S such that $\mathbf{X}|_{\partial S} = \frac{\partial}{\partial y}$, then there is a diffeomorphism $f : S \to S$ such that $f^*\mathbf{X} = \frac{\partial}{\partial y}$.*

Three-dimensional manifolds allow a lot of bad things to happen, though; the curve may be everywhere dense in the manifold. Generally, a poorly-behaved vector field can't be expected to have a nice, analytic solution.

Thus, it is nice to have integrals that reduce the dimension of the phase space. Conservation of energy (since H is constant on trajectories) is one already, but suppose that G is another integral, so that $\{G, H\} = 0$. Then, all trajectories lie on intersections of these level sets $\{G = a, H = b\}$, so if they intersect transversely, the Implicit

[37] . . . and got a big prise from the Swedish king. Later, he found an error in his proof, but was able to fix it.

Function Theorem implies that the intersection is 2-dimensional if the system has 2 degrees of freedom. Additionally, if they are compact, the intersection is a torus, and the vector field is given by some constant slope. This kind of motion is called quasi-periodic.

Theorem 16.2 (Liouville)**.** *In the general case (n degrees of freedom), suppose there are n integrals $F_1, \dots, F_n$,*[38] *and that they Poisson-commute: that $\{F_i, F_j\} = 0$ for $i \neq j$. If they are linearly independent, then they intersect transversely, so by the Implicit Function Theorem, their collective intersection is a manifold of dimension n, and if the level sets are compact, they form an n-dimensional torus.*

In the case of 2 degrees of freedom, choosing a level set $H = E$ gives a 3-manifold and another level set for the other integral gives the torus, so the space is foliated by tori. On the various tori, the slope changes, so consider the annulus between two tori at a cross-section somewhere, and follow it around the torus under the flow by one rotation.

Then, this is a twist map: each circular section of the torus is rotated by some angle. This map is area-preserving, since every parallelogram is sheared, and this is no accident: $L_{\mathbf{X}_H}\omega = 0$, and the restriction of ω to the torus is an area form.

Suppose this system is pertubed so that the boundaries do the same thing, but the middle does something new. Then, assuming the rotations are distinct, it is possible to pick a coordinate system on the torus such that the rotations of the outer and inner circles are in opposite directions. Specifically, choose two integer vectors with determinant 1, which can be used as a basis for the torus.

This is why Poincaré cared about his last geometric theorem; he arrived specifically at it through the 3-body problem. This problem has an 18-dimensional phase space and lots of integrals. including linear and angular momentum, but not enough to be solved. Thus, he considered specifically the restricted 3-body problem, in which one of the particles is small enough that its effect on the others is negligible, but it still moves under the influence of the other two. Poincaré wanted to find periodic orbits, which corresponded to fixed points of this map.

Proof of Theorem 16.2. Everything will be done in thw two-dimensional case, though the general case is very similar.

Suppose there are two integrals F_1, F_2 such that $\{F_1, F_2\} = 0$, and pick some constraints $F_1 = c_1$ and $F_2 = c_2$. Let $\mathbf{c} = (c_1, c_2)$, and define $T_{\mathbf{c}} = \{F_1 = c_1, F_2 = c_2\}$. Since the level sets intersect transversely and are compact, then $T_{\mathbf{c}}$ is a closed 2-dimensional manifold. Then, $T_{\mathbf{c}}$ is orientable, because at each tangent plane there is a canonical basis: let $\mathbf{X}_{F_1} = \mathbf{X}_1$ and $\mathbf{X}_{F_2} = \mathbf{X}_2$ be the Hamiltonian vector fields for F_1 and F_2, respectively. Then, $\mathbf{X}_1, \mathbf{X}_2 \in T(T_{\mathbf{c}})$ because they are tangent to the level sets of their respective functions, and are tangent to each others' level sets because they commute. This does also require that they don't vanish (which is known because they're linearly independent).

From topology, it is a fact that a closed, oriented 2-dimensional manifold is a sphere with n handles, where n is the genus (so that the sphere is a genus-0 surface, the torus is a genus-1 surface, etc.) However, the torus is the only such surface that admits a nonvanishing vector field (a consequence of the Hairy Ball Theorem). This is enough to imply the theorem in this case, but this doesn't generalize to higher dimensions, so another path will be sought.

Let $\mathbf{X}_1^t, \mathbf{X}_2^t$ be the flows due to $\mathbf{X}_1$ and $\mathbf{X}_2$. Then, they commute: $\mathbf{X}_1^t \circ \mathbf{X}_2^s = \mathbf{X}_2^s \circ \mathbf{X}_1^t$, because their Lie bracket at a vector $\mathbf{u}$ is

$$[\mathbf{X}_1, \mathbf{X}_2](\mathbf{u}) = \lim_{t,s\to 0} \frac{\mathbf{X}_1^s \mathbf{X}_2^t \mathbf{X}_1^{-s} \mathbf{X}_2^{-t}(\mathbf{u}) - \mathbf{u}}{ts},$$

and since the corresponding integrals Poisson-commute, their flows do as well.[39]

Consider the map $\varphi : \mathbb{R}^2 \to T_{\mathbf{c}}$ given by some $a \in T_{\mathbf{c}}$ and $\varphi(s,t) = \mathbf{X}_1^s \circ \mathbf{X}_2^t(a)$. This is a group action of $\mathbb{R}^2$ on $T_{\mathbf{c}}$, since $\mathbb{R}^2$ is an abelian group under addition, so $\varphi(u+v) = \varphi(u) + \varphi(v(a))$. Here are some properties of φ:

1. φ is an immersion (i.e. a map between two manifolds of maximal rank), so its Jacobian is nonvanishing. Specifically, its Jacobian is $(\mathbf{X}_1, \mathbf{X}_2)$, which we know to be nonvanihsing because of linear independence.
2. φ cannot be injective, since $T_{\mathbf{c}}$ is compact, but $\mathbb{R}^2$ is not.
3. φ doesn't have accumulation points. Topologically, φ is proper: the preimage of a compact set is compact.[40] It is a fact from elementary analysis that a proper map that maps to a connected space is onto, so φ is onto as well.

Consider $\Gamma = \{(s,t) \mid \varphi(s,t) = a\} \subseteq \mathbb{R}^2$. This is a subgroup of $\mathbb{R}^2$, called a stationary subgroup, but it is also a discrete subgroup: since φ is locally diffeomorphic, every $\mathbf{x} \in \Gamma$ has a neighborhood U such that $\Gamma \cap U = \mathbf{x}$.

This is pretty cool, so what are the discrete subgroups in $\mathbb{R}^2$ (or $\mathbb{R}^n$)? In $\mathbb{R}$, if Γ is a discrete subgroup, then Γ is trivial or $\Gamma = \{ne, n \in \mathbb{Z}\}$ for any $e \in \mathbb{R}$ (which can be determined by considering the smallest possible positive element in Γ). In $\mathbb{R}^2$, pick some $\mathbf{e}_1 \neq \mathbf{0}$ with $\mathbf{e}_1 \in \Gamma$. Then, there is some point $\mathbf{e}_2 \in \Gamma$ that is closest to the line spanned by $\mathbf{e}_1$, but not on it (if not, then $\mathbf{e}_1$ would be an accumulation point), so $\Gamma = \{n\mathbf{e}_1 + m\mathbf{e}_2 \mid m, n \in \mathbb{Z}\} = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle$. By applying induction on n, one obtains

[38] Usually, we have that H is one of these integrals, but it doesn't actually matter, and can be treated as a minor generalization.

[39] The actual argument requires a little more work; see Arnold's book for the details.

[40] This map is not in general proper: for example, if $\mathbb{R} \xrightarrow{\varphi} S^1$, then $\varphi^{-1}(S^1) = \mathbb{R}$, which is noncompact.

Theorem 16.3. *If $\Gamma \leq \mathbb{R}^n$ is a discrete subgroup, then there exist k linearly independent vectors $\mathbf{e}_1, \dots, \mathbf{e}_k$ (for some $k \leq n$) such that*

$$\Gamma = \left\{ \sum_{i=1}^{k} n_i \mathbf{e}_i \mid n_i \in \mathbb{Z} \right\} = \langle \mathbf{e}_1, \dots, \mathbf{e}_k \rangle .$$

These subgroups are called lattices.

Group-theoretically, this means that $T_{\mathbf{c}} = \mathbb{R}^2/\Gamma$ (or, in the general case, $T_{\mathbf{c}} = \mathbb{R}^n/\Gamma$, which will be used in what follows). If the rank of this lattice is n, this implies $T_{\mathbf{c}}$ is a torus; if not, the $T_{\mathbf{c}}$ is the product of a torus and Euclidean space. Since $T_{\mathbf{c}}$ is compact, then the rank of Γ must be n, so $T_{\mathbf{c}}$ is in fact a torus.

Finally, it is necessary to show the motion is quasi-periodic. The lattice Γ determines a coordinate system φ_1, φ_2, which are cyclic coordinates defined mod 1. If they're scaled mod 2π, something more angle-like is obtained. They are given by basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, so the flows of $\mathbf{X}_1^t$ and $\mathbf{X}_2^t$ are lienar combinations of $\mathbf{e}_1$ and $\mathbf{e}_2$: $\mathbf{X}^t(a) = a_1 t \mathbf{e}_1 + a_2 t \mathbf{e}_2 = (a_1 t, a_2 t)$, so the flow will be linear starting from any point. ⊠

Exercise 16.1. The above argument requires $F_1 = H$. If F_1 is instead a function of H, why does this also follow?

This proof was very theoretical, so how might one actually find the flow? φ_1 and φ_2 seem pretty mysterious. Tune in next time for some general methods for doing this. One example: if $\mathbf{p}$ and $\mathbf{q}$ are some local canonical coordinates, with $\mathbf{p} = (p_1, p_2)$ and $\mathbf{q} = (q_1, q_2)$, then the goal is to find some new canonical coordinates $(\mathbf{I}, \boldsymbol{\varphi})$ with $\mathbf{I} = (I_1, I_2)$ and $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$, such that $\omega = \mathrm{d}I_1 \wedge \mathrm{d}\varphi_1 + \mathrm{d}I_2 \wedge \mathrm{d}\varphi_2$, $H = h(\mathbf{I})$ for some h (in particular, H is independent of φ), and so thsat $\mathbf{I}$ is the vector of integrals and $\dot{\varphi}_1 = \frac{\partial H}{\partial I_1}$ and $\dot{\varphi}_2 = \frac{\partial H}{\partial I_2}$ are the equations of motion. Then, $\mathbf{I}$ parameterizes the torus, and the right-hand sides of these equations are constant on a given torus. Thus, one can obtain the angles for quasi-periodic motion. Then, $\mathbf{I}$ and $\boldsymbol{\varphi}$ are given by generating functions: they can be defined globally even if $\mathbf{p}$ and $\mathbf{q}$ are only local.

Fix a point on the torus, and observe that the torus is Lagrangian: $\omega(\mathbf{X}_1, \mathbf{X}_2) = \{F_1, F_2\} = 0$, since ω vanishes on the tangent space. Then, take any closed surve γ on the torus, and assume γ is exact (which isn't strictly necessary, but is helpful). Then, $\omega = d(\mathbf{p}\,\mathrm{d}\mathbf{q}) = 0$, so $\mathbf{p}\,\mathrm{d}\mathbf{q}|_{T_{\mathbf{c}}}$ is closed, and thus $\int_\gamma \mathbf{p}\,\mathrm{d}\mathbf{q}$ is homotopy-independent (in terms of changing γ). There are two classes of curves, so one obtains I_1 and I_2. Then, doing this on every torus yields the functions everywhere.

17. The Hamilton-Jacobi Method: 3/11/13

Thanks to Liouville's Theorem, a system on a $2n$-dimensional phase space with n integrals $F_1, \dots, F_n$ has quasi-periodic motion given by $\varphi_i \mapsto \varphi_i + \omega_i t$, or motion with veloity $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)$. On the torus, each coordinate looks like $e^{c_k + 2\pi i t \omega_k}$, winding around the torus. If the periods are incommensurate, the motion won't be periodic. The great question is, how should these φ_i be obtained?

Example 17.1. Consider a system with one degree of freedom and a Hamiltonian $H = p^2 + q^2$. This is a complete system of commuting integrals (since only one is necessary), so the goal is to find a change of coordinates so that one of them is cyclic. The symplectic form is $\omega = \mathrm{d}p \wedge \mathrm{d}q = r\,\mathrm{d}r \wedge \mathrm{d}\varphi - \mathrm{d}(r^2/2) \wedge \mathrm{d}\varphi$. Letting $I = r^2/2$, this is $\mathrm{d}I \wedge \mathrm{d}\varphi$, so φ is a cyclic coordinate with period 2π, and these coordinates are called action-angle coordinates.

Pick some level set $T_h = \{H = h\}$, so that $\oint_{T_h} p\,\mathrm{d}q = \pi h = 2\pi I$, where $I = h/2$. This is the area of a circle with radius $\sqrt{h}$. Then, fix some point $q_0 \in T_h$ such that $q = 0$ and take the coordinate chart, which will be $p = \sqrt{h - q^2}$. If q is any other point on T_h, let $T_h^{q_0 q}$ be the arc from q_0 to q with counter-clockwise orientation. Then, designate

$$S(I, q) = \int_{T_h^{q_0 q}} p\,\mathrm{d}q = \int_{q_0}^{q} -\sqrt{2I - q^2}\,\mathrm{d}q. \tag{6}$$

The goal is to consider this as the generating transformation $\mathrm{d}p \wedge \mathrm{d}q \overset{\oplus}{\mapsto} \mathrm{d}I \wedge \mathrm{d}\varphi$. The graph must be Lagrangian, and if it's graphical with respect to the Lagrangian subspaces $(I, q) \oplus (\varphi, p)$, then it's also the graph of a differential form, so $\varphi = \frac{\partial S(I,q)}{\partial I}$ and $p = \frac{\partial S(I,q)}{\partial q}$.

With S as in (6), $\varphi = \int_{q_0}^{q} \frac{\mathrm{d}q}{\sqrt{2I - q}}$ and $p = -\sqrt{2I - q^2}$ (which is just a restatement of the coordinate chart). Geometrically, $\sin\varphi = q/\sqrt{2I}$, and this is exactly what φ integrates to. Also, it is possible to check that the period of φ is 2π: $\Delta\varphi = \frac{\partial(\Delta S)}{\partial I}$. But S changes by $2\pi I$ over a period, so $\Delta\varphi = 2\pi$.

This example illustrates a much more general recipe. Take the torus T_h, and take loops $\gamma_1, \dots, \gamma_n$ such that $T_h = \gamma_1 \times \cdots \times \gamma_n$. Then, let $I_j = \frac{1}{2\pi} \int_{\gamma_i} \mathbf{p}\,\mathrm{d}\mathbf{q}_i$, and again define the generating function $S(\mathbf{q}, \mathbf{I})$ by fixing some

$\mathbf{q}_0 \in T_h$ and defining $S(\mathbf{q}, \mathbf{I}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{p}\,\mathrm{d}\mathbf{q}$, which is well-defined because of homotopy invariance,[41] at least within some coordinate neighborhood. Then, $\varphi_j = \frac{\partial S}{\partial I_j}$, so by the same logic as above, each φ_j is 2π-periodic.

It turns out that actually computing this is a nightmare, and it's mostly a theoretical result. The Hamilton-Jacobi method is kind of weird, but it's much easier to use. Historically, people have tried to invent methods that yield integrable systems and classified the things that could be solved, under the reasoning that integrable systems are rare and thus important.

Suppose a Hamiltonian system is given by $H(\mathbf{p}, \mathbf{q})$. If it were possible to change coordinates $(\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{P}, \mathbf{Q})$ such that $H(\mathbf{p}, \mathbf{q}) = K(\mathbf{Q})$, then the system is automatically integrable, and by definition is an involution (since these coordinates would be canonical). Thus, the question returns as always to finding these coordinates. In fact, the goal is to find a generating function $S(\mathbf{q}, \mathbf{Q})$ such that $\mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}$ and $\mathbf{P} = \frac{\partial S}{\partial \mathbf{Q}}$. Then, the problem reduces to solving $H(\frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \mathbf{Q}), \mathbf{q}) = K(\mathbf{Q})$. This is a 1^{st}-order partial differential equation — and it's being used to solve an ODE. Though this seems like it shouldn't make any sense, it works well in many cases. Alternatively, it can be run in reverse to solve some PDEs.

Viewing $\mathbf{Q}$ as a parameter, the right-hand side is constant for each $\mathbf{Q}$, so the goal is to find an n-parametric family of solutions. In some sense, $K(\mathbf{Q})$ can be thought of as one such coordinate, and $n-1$ more are needed. This family of solutions is called a full solution.

The general idea is as follows: if the right-hand side can just be written as

$$K(\mathbf{Q}) = h_1\left(\frac{\partial S}{\partial q_1}, q_1, c\right) + h_2\left(\frac{\partial S}{\partial q_2}, q_2, c_1, c_2\right) + h_3\left(\frac{\partial S}{\partial q_3}, q_3, c_1, c_2, c_3\right) + \dots,$$

then one can solve $h_1\left(\frac{\partial S}{\partial q_1}, q_1\right) = c_1$, since it's just an ordinary differential equation. Then, the next equation can be solved, since the variables separate.

Example 17.2. Consider the motion of a particle in a field with two centers. Each center produces a field $k/r + 1$, k/r_2 (e.g. a particle moving in the gravitational field of an ellipsoid). Then, the trick is to introduce good coordinates: if r_1 and r_2 are the distances of the particle to the respective centers, introduce $\xi = r_1 + r_2$ and $\eta = r_1 - r_2$, which are called the elliptic coordinates of Jacobi. The dual coordinates are also used, and are called p_ξ and p_η. One page of computation later, it is possible to show that

$$H = 2p_\xi^2 \frac{\xi^2 - 4c^2}{\xi^2 - \eta^2} + 2p_\eta^2 \frac{4c^2 - \eta^2}{\xi^2 - \eta^2} - \frac{4k\xi}{\xi^2 - \eta^2}, \tag{7}$$

where k is as above (some sort of gravitational constant). Then, the goal is to find $p_\xi = \frac{\partial S}{\partial \xi}$ and $p_\eta = \frac{\partial S}{\partial \eta}$:

$$\begin{aligned} K(\xi^2 - \eta^2) &= 2p_\xi^2(\xi^2 - 4c^2) + 2p_\eta^2(4c^2 - \eta^2) - 4k\xi \\ &= 2\left(\frac{\partial S}{\partial \xi}\right)^2 (\xi^2 - 4c^2) + 2\left(\frac{\partial S}{\partial \eta}\right)^2 (4c^2 - \eta^2) - 4k\xi. \end{aligned} \tag{8}$$

Then, group together the ξ and η terms and solve.

18. Optics and Contact Geometry: 3/13/13

The idea behind the Hamilton-Jacobi method is to consider a Hamiltonian function $H(\mathbf{p}, \mathbf{q})$ (it could even depend on time, though we won't consider that) and find a change of canonical coordinates $\mathbf{P}, \mathbf{Q}$ such that $H(\mathbf{p}, \mathbf{q}) = K(\mathbf{Q})$. Then, the coordinates automatically Poison-commute and there are n integrals.

The map $\mathbf{p}, \mathbf{q} \mapsto \mathbf{P}, \mathbf{Q}$ is a symplectomorphism, and its graph in $\mathbb{R}^{4n}$ is Lagrangian with respect to the form $\mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{q} - \mathrm{d}\mathbf{P} \wedge \mathrm{d}\mathbf{Q}$. If this space can be split into two Lagragian subspaces and the map is graphical with respect to this splitting, then it is the differential of a function S: $= \frac{\partial S(\mathbf{q}, \mathbf{Q})}{\partial \mathbf{q}}$, and $\mathbf{P} = -\frac{\partial S(\mathbf{q}, \mathbf{Q})}{\partial \mathbf{Q}}$ (in this case, the splitting is $(\mathbf{p}, \mathbf{P}) \oplus (\mathbf{q}, \mathbf{Q})$). This can be thought of as a system of equations, and if the map is graphical, then a solution exists, at least locally.

Thus, the goal isn't as much to find $\mathbf{p}, \mathbf{q} \overset{\Phi}{\to} \mathbf{P}, \mathbf{Q}$ as much as it is to find the generating function $S(\mathbf{q}, \mathbf{Q})$, so that a family of solutions to $H\left(\frac{\partial S(\mathbf{q})}{\partial \mathbf{q}}, \mathbf{q}\right) = K$ is obtained (with $\mathbf{Q}$ as a parameter). This is a first-order PDE, and the Hamiltonian is a first-order ODE, but they are connected: knowing the wolution to one implies a method for finding a solution for the other.

Continuing with Example 17.2 from the previous lecture, a particle moved in a gravitational field generated by two point masses and had distances r_1 and r_2 from these point masses. Then, Jacobi introduced coordinates $\xi = r_1 + r_2$

[41] Speaking more precisely, each I_j is homotopy-invariant because $\mathbf{p}\,\mathrm{d}\mathbf{q}$ is a closed form, which is because $\mathbf{p}\,\mathrm{d}\mathbf{q}|_{T_h} = 0$, and the integral of a closed form over homotopic paths is constant.

and $\eta = r_1 - r_2$, giving a Hamiltonian as in (7), where k is the gravitational constant. Then, the differential equation in (8) can be rewritten so the variables are separated:

$$2\left(\frac{\partial S}{\partial \xi}\right)^2 (\xi^2 - 4c^2) - (C\xi^2 + 4k\xi) + 2\left(\frac{\partial S}{\partial \eta}\right)^2 (4c^2 - \eta^2) + C\eta^2 = 0.$$

Thus, the PDE can be broken into two ODEs. Letting

$$c_1 = 2\left(\frac{\partial S}{\partial \xi}\right)^2 (\xi^2 - 4c^2) - (C\xi^2 + 4k\xi) \text{ and}$$
$$-c_1 = 2\left(\frac{\partial S}{\partial \eta}\right)^2 (4c^2 - \eta^2) + C\eta^2,$$

one obtains some solutions $S_1(\xi, c_1, C)$ and $S_2 = (\eta, c_1, C)$, and the overall function is $S(\xi, \eta, c_1, C) = S_1(\xi, c_1, C) + S_2 = (\eta, c_1, C)$. This can be explicitly solved in terms of elliptic integrals, which aren't elementary functions but are at least integrals of elementary functions.

Given a problem, it's difficult to determine which coordinates work for the system (in terms of performing a separation as above), or even if such a system exists at all. Landau and Lifshcitz list some standard coordinates transforms (which involves a scary amount of casework and computation). Many of these were discovered by taking a method and trying to find problems that can be solved with it rather than the other way around.

This method also admits a more geometric approach. The general approach to a first-order PDE is to take oordinates $\mathbf{q} = q_1, \dots, q_n$ and find some function $S(q_1, \dots, q_n)$ such that

$$F\left(q_1, \dots, q_n, \frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_n}, S\right) = 0.$$

For simpliclty, assume F doesn't depend on S; then, $F(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}})$ is exactly the Hamilton-Jacobi equation!

Consider the $\mathbf{pq}$-space (if $\mathbf{q}$ gives coordinates on a manifold M, this is T^*M). It's easier to solve algebraic equations than differential ones, so consider $F(\mathbf{q}, \mathbf{p}) = 0$, assuming $\mathbf{p}$ and $\mathbf{q}$ are independent. This defines a hypersurface in T^*M: for example, if

$$F\left(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}\right) = \sum_{i=1}^{n} \left(\frac{\partial S}{\partial q_i}\right)^2 - 1,$$

then one obtains a hypersurface $\sum p_i^2 = 1$, which is a cylinder $S^n \times \mathbb{R}^n$.

Suppose $p_i = \varphi_i(\mathbf{q})$, which is a necessary (yet not sufficient) condition for solving the Hamilton-Jacobi equation, and consider the map $\Phi = (\varphi_1, \dots, \varphi_n)$ such that $\Phi(q_1, \dots, q_n) = (\varphi_1(\mathbf{q}), \dots, \varphi_n(\mathbf{q})$ and let $\Sigma = \{F(\mathbf{q}, \mathbf{p}) = 0\}$ (i.e. the hypersurface in question). Then, we want the graph of Φ to sit within they hypersurface Σ, and specifically such that the graphs of the φ_i are partial derivatives. This is equivalent to the graph being a Lagrangian submanifold, so the problem can be framed as finding graphical Lagrangian submanifolds of this hypersurface.

Σ can be thought of as a level set of some Hamiltonian function F, so the Hamiltonian vector field is tangent to Σ. This is slightly dishonest because it depends on a choice of F, but if $\Sigma = \{\widetilde{F} = 0\}$ for some $\widetilde{F} \neq F$, the direction of the Hamiltonian vector field doesn't change, since it's still tangent, though the magnitudes can. Thus, the corresponding line field (also called a direction field) is invariant, and gives a 1-dimensional foliation of Σ.

Claim. Thus, every Lagrangian submanifold must be tangent to this line field.

Proof. Consider symplectic space $(\mathbb{R}^{2n}, \omega)$ and some surface $\Pi^{2n-1} \subset \mathbb{R}^{2n}$. Then, $\omega|_\Pi$ is a skew-symmetric form, so it must be degenerate (since $2n - 1$ is odd). Thus, it has a kernel. This kernel must be one-dimensional; if it weren't, it couldn't be killed by going back up to $\mathbb{R}^{2n}$. In particular, it must lie in the Hamiltonian direction, since it is perpendicular to the normal (viewing $\mathbb{R}^{2n}$ as $\mathbb{C}^n$ again). Thus, if a Lagrangian submanifold is transverse to this line field, it can be extended in that direction to an $(n+1)$-dimensional manifold on which ω is zero, which is impossible. ⊠

Note that the integral curves in the direction of the vector field are sometimes called characteristics.

In ODEs, one generally specifies initial conditions in order to obtain specific data. In PDEs, the corresponding starting conditions are called boundary conditions. One example is the Cauchy problem: if a function h is given on $\{q_n = 0\}$, the goal is to extend it in some neighborhood (spoiler: this isn't always possible): $S(q_1, \dots, q_{n-1}, 0) = h(q_1, \dots, q_{n-1})$. Suppose F is resolved in one partial derivative: $\frac{\partial S}{\partial \mathbf{q}_n} = G\left(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}, q_n\right)$ for some G, so the hypersurface is $p_n = G(\mathbf{q}', \mathbf{p}, q_n)$, where $\mathbf{q}' = (q_1, \dots, q_{n-1})$ and $\mathbf{p}' = (p_1, \dots, p_{n-1}) = \frac{\partial S}{\partial \mathbf{q}'}$. Locally, this can be done with some

coordinates if the method applies at all, so

$$G\left(\mathbf{q}, \frac{\partial h}{\partial \mathbf{q}}, 0\right) = p_n = \frac{\partial S}{\partial q_n}(0),$$

and some initial data is obtained.

The vector field in the directon of q_n is nonzero, so one can flow the initial data out to an $(n-1)$-dimensional surface, which also happens to be a Lagrangian submanifold. Thus, it's the graph of the gradient of something, which will be the solution to the equation. Since there is lots of possible choices of initial data, there are lots of possible solutions.

This discussion naturally leads to geometric optics. An optical system is a medium in which light propagates, such that the speed of light may be different at different points or in different directions (e.g. in an inhomogeneous material), but it doesn't change with respect to time. Thus, the phase space of an optical system is $(2n-1)$-dimensional, since the magnitude of the velocity is given, and the only necessary information is the direction.

There exists a Lagrangian formalism, but the Hamiltonian one is much nicer. Consider a configuration space M, which is also the space where light propagates.[42] Then, take the cotangent bundle T^*M. Since it's only necessary to know the momenta up to proportionality, take the projectivized cotangent bundle, PT^*M (all vectors in T^*M up to constant, so in some sense the space of directions), which is a $(2n-1)$-dimensional manifold.

A point in $\ell \in T^*M$ is a linear function, and any such linear function defines a hyperplane $\{\ell = 0\}$ that is tangent to M; if ℓ_1, ℓ_2 define the same hyperplane, then they differ by a constant. Thus, an $\ell \in PT^*M$ can be thought of as a hyperplane tangent to M., and PT^*M is called the space of contact elements.

Recall that projective space also came up in the discussion of a system of non-holonomic constraints, in which velocity must be tangent to a given direction, so the space is projective.

The cotangent bundle has a canonical form $\mathbf{p}\,\mathrm{d}\mathbf{q}$, but this doesn't descend into projective space. Its kernel does, though, and on the space of contact elements, there is a canonical hyperplane field defined by a skating condition (so named because of the example commonly given for systems of non-holonomic constraints): there is a component in the direction of the skate $\dot{q}$ and some angular velocity $\dot{\varphi}$, and the condition is that $\dot{q}$ must point in the direction of the skate.

In optics, there are 2 languages: one can describe light as rays emanating from a point, or one can consider a light source and consider where light can go in a fixed time. In the latter case, one obtains a hyperurface, called a wavefront. The light source might not be a point, however. The equidistant condition implies that at some point there is a cusp, where the wavefront intersects itself. The set of all of these singularities takes the form of an astroid. These points are specially lit, and called caustic.

Huygens' Principle. In order to describe a front at a time $t+s$, the following two descriptions are equivalent:

(1) The front can be considered as everywhere the light can reach in time $t+s$.
(2) One can take the front of light at time t and treat it as a light source. Then, the space that light from that source can reach in time s can be taken as the overall front at time $t+s$.

From the perspective of contact geometry, tangent lines can be added, so one obtains a submanifld that is Legendrian, or tangent to the hyperplane field in the space of contact elements. This is the contact analogue of Lagrangian. Thus, one has a vector field which moves Legendrian manifolds to Legendrian manifolds. These manifolds are smooth, though their projections onto the configuration space may contain singularities.

The dynamics are no different than in mechanics: one can take the trajectories of the Hamiltonian system to obtain rays.

19. Legendrian Knots: An Introduction to Contact Geometry: 3/14/13

This lecture wasn't technically part of the class, as it was the last lecture of the SUMO Speaker Series for this quarter. However, Dr. Eliashberg gave the talk and it was very relevant to Math 137, so it has been included in these notes.

Consider a function of one variable $z = f(x)$. Its graph exists in the xz-plane, but we're in 3-space, so we can do much better: let $y = f'(x)$, so that one gets a curve which is $f(x)$ when projected onto one axis and $f'(x)$ when projected onto the other.

In this space, there is a remarkable plane field given by the Pfaffian equation $\mathrm{d}z - y\,\mathrm{d}x = 0$. At every point there is a plane whose slope in the xz-direction is y. Thus, these planes twist in the y-direction, as in Figure 3.

[42] This can have more than 3 dimensions, such as if there are several sources of light.

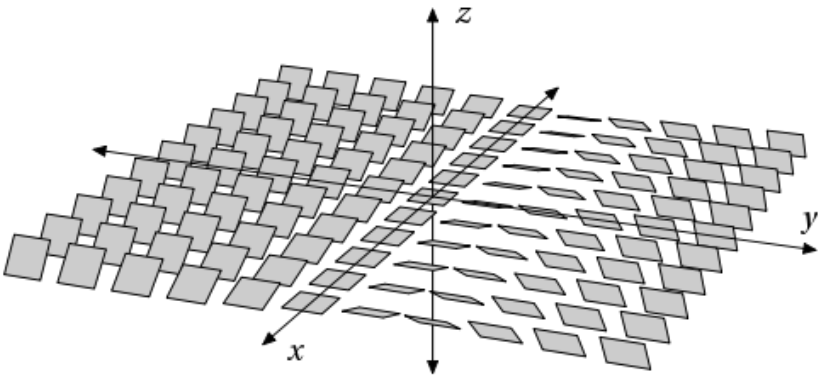

FIGURE 3. The contact structure $\mathrm{d}z - y\,\mathrm{d}x = 0$. Source: Wikipedia

This plane field is called a contact structure, and is completely nonintegrable: that is, there is no surface that is tangent to it, even locally. This can be proven in many ways. However, there are curves which are tangent to it, which are called Legendrian curves.[43] These curves appear a lot in everyday life.

Consider the skating problem: a skater skates on the plane. Its skate touches the plane at some point (x, y) with an angle φ. Thus, motion is given by 3 functions $x(t)$, $y(t)$, and $\varphi(t)$. The skater can rotate and move forward, but only in the direction of the skate: $(\dot{x}, \dot{y})$ and $(\cos\varphi, \sin\varphi)$ must point in the same direction, which implies that $\dot{y}\cos\varphi - \dot{x}\sin\varphi = 0$.

In the three-dimensional space given by (x, y, φ)-coordinates, consider the form given by $\cos\varphi\,\mathrm{d}y - \sin\varphi\,\mathrm{d}x = 0$. This is identical to the previous plane field up to a change of coordinates.

Parking a car is another example of the skating condition, since the car must travel in the direction of the wheel. Thus, a car traces a Legendrian curve in the (x, y, φ) phase space. There are plenty of other applications, such as Carnot cycles, but that isn't really the point of the talk.

Suppose that the plane fied were integrable; then, locally, the space will be foliated by leaves tangent to the curve. Then, motion between two distinct leaves would be impossible. Of course, this can't be correct, because parallel parking exists, and this is precisely motion between two parallel curves.

We have already seen that the simultaneous graph of a function and its derivative is a Legendrian curve. However, there is a nice converse: a graphical Legendrian curve such that $z = f(x)$ requires that $y = f'(x)$. There are non-graphical Legendrian curves, however, such as those places where the tangent line is vertical. Suppose this is nondegenerate (dominated by the quadratic term), so $x = y^2$ and $y = \pm\sqrt{x}$. Then, $z = \pm 2x^{3/2}/3$, which is a semicubical parabola. In general, the projection of this curve to the xz-plane is called the front of the curve. Geometrically speaking, there is a cusp. Thus, a generic Legendrian curve has a lot of smooth pieces and these cusps. Going from one point to another on this curve gives a piecewise smooth curve with these cusps. In fact, the curve can be made with a specified arbitrary slope, which means that every curve can be C^0-approximatd with a Legendrian curve. In the real world, this implies one can parallel park in a very small space.

Legendrian curves also naturally arise in optics. Consider a flat, homogeneous, isotropic medium, so that light propagates in straight lines and with constant speed in all directions. For some given light source, at a fixed time, light is in a circle around that source. This is also called a front. Optics can consider where the light rays go, but also how the fronts change. For example, if one takes an elliptical source and "moves backward," the fronts are no longer smooth, but they still look like Legendrian curves.

At every point (x, y), consider the co-oriented tangent line to the front, and let φ be the angle of the unit normal. Then, the phase space of the system is (x, y, φ)-space, or $\mathbb{R}^2 \times S^1$. Qualitatively, each point in $\mathbb{R}^2$ can be lifted to $\mathbb{R}^2 \times S^1$ by adding on the slope of the unit normal. Then, the contact structure $\cos\varphi\,\mathrm{d}x + \sin\varphi\,\mathrm{d}y$ (which is the same as in the skating problem) makes this a Legendrian curve. Then, the condition that light moves at constant speed perpendicular to the wavefront implies that the curve is always Legendrian. The singular points of the curve are the points where several light rays come together (such as the center of curvature), called caustic points.

Legendrian curves are also quite relevant in mathematics. For example, consider a continuous function φ on an interval (a, b) such that $\varphi'(x) > 0$. Then, Rolle's Theorem implies that $\varphi(b) > \varphi(a)$. This can be extended in some sense for Legendrian curves. If a Legendrian curve L has positive slope, then $L \subset \{y > 0\}$, but again, by parallel parking, it's possible to create a counterexample. However, there are restrictions for which the theorem holds.

[43] Of course, Legendre had never heard of these curves; the term is due to Arnold.

Theorem 19.1. *Let $\mathcal{L}$ be the space of all embedded*[44] *Legendrian curves which lie between 2 planes and with boundary on these planes. Then, let $\mathcal{L}_0 \subset \mathcal{L}$ denote the connected component of curves which can be deformed to a graphical curve without self-intersection. Then, Rolle's Theorem holds in $\mathcal{L}_0$.*

Thus, there exist Legendrian curves which aren't isotopic to a graphical curve. This is related to applications of contact geometry in the outside world.

So, what does Legendrian isotopy actually mean? Such a curve, when projected to the xz-plane, looks like a front, but on the xy-plane, each cusp corresponds to the tangent line parallel to the y-axis. Thus, to get the original curve from the xz-plane, one needs only to add the slope at each point. The curve is embedded if every loop given by a self-intersection has nowhere zero algebraic area; thus, self-intersections can only happen where two branches are tangent on the front in the z-plane (since that implies they have the same slope).

In some sense, this can be framed combinatorially, without any topology. One can prove the following, albeit not easily: that the curve resulting from such a deformation might not be graphical, but is close; there are a finite number of points such that splitting them in a certain way yields a graphical curve. Thus, the curve has a bunch of smooth sections and a finite number of "flying saucers," implying it is piecewise continuous. Thus, Rolle's Theorem still holds for piecewise constant functions. The proof of this would take another lecture (the first proof was 1500 pages and wrong; the next one was 100 pages but incomprehensible).

There's also a wonderful subject of Legendrian knots (closed Legendrian curves), which can be classified up to Legendrian isotopy. This is a very rich theory, and there are a lot of invariants:

- The most obvious is the topological class of the knots.
- More subtly, there are numerical invariants, such as the tangential degree of projection. This is the number of loops made by the tangent line over one traversal of the knot.
- The Thurston-Bennequin invariant (related to something called the Arnold principle): if the knot is moved slightly, what is the signed number of self-intersections?

If the knot is topologically trivial (i.e. the unknot), these are the only invariants. If otherwise, there are some more. For example, let the self-intersections be given by $a_1, \dots, a_n$ and take the associative algebra on $\mathbb{Z}_2$ of these letters: $A = \mathbb{Z}_2[a_1, \dots, a_n]$. Then, there exists a differential $\mathrm{d} : A \to A$ such that $\mathrm{d}(fg) = (\mathrm{d}f)g + f(\mathrm{d}g)$ and $\mathrm{d}^2 = 0$. This is defined by taking every vertex and putting a sign by going counterclockwise; if the curve jumps up, then it's positive, and if it jumps down, it's negative. For some vertex q_i, if q_j and q_k are the other vertices in all polygons including q_i, then $\mathrm{d}q_i = \sum q_j q_k$ for all such j, k. Then, taking some homology, $\mathrm{Ker}(d)/\mathrm{Im}(d)$ is another algebra, and is the invariant for the knot.

[44] That is, nonsingular and non-self-intersecting. Such a curve is always nonsingular (immersed), by the definition of a manifold.