2. Hamiltonian Systems.

This lecture contains some of the basic facts about Hamiltonian systems. Some additional material will be brought in later as it is needed.

Motivation.

To motivate the development, let us briefly consider Hamilton's equations. The starting point is Newton's second law which states that a particle of mass $m > 0$ moving in a potential $V(x)$, $x \in R^3$ moves along a curve $x(t)$ such that $m\ddot{x} = -\text{grad } V(x)$. If we introduce the momentum $p = m\dot{x}$ and the energy $H(x, p) = \frac{1}{2m} \|p\|^2 + V(x)$ then Newton's law becomes Hamilton's Equations

$$\begin{cases} \dot{x}^i = \partial H/\partial p_i \\ \dot{p}_i = -\partial H/\partial q^i \qquad i = 1, 2, 3. \end{cases}$$

One now is interested in studying this system of first order equations for given H . To do this, we introduce the matrix $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ where I is the $n \times n$ identity and note that the equations become $\dot{\xi} = J \text{ grad } H(\xi)$ where $\xi = (x, p)$. (In complex notation, setting $z = x + ip$, they may be written as $\dot{z} = 2i\partial H/\partial \bar{z}$).

Suppose we make a change of coordinates $w = f(\xi)$ where $f : R^{2n} \to R^{2n}$ is smooth. If $\xi(t)$ satisfies Hamilton's equations, the equations satisfied by $w(t)$ are $\dot{w} = A\dot{\xi} = AJ \text{ grad}_\xi H(\xi) = AJA^* \text{grad}_w H(\xi(w))$ where $A = (\partial w^i/\partial \xi^j)$ is the Jacobian of f . The equations for w will be Hamiltonian with energy $K(w) = H(\xi(w))$ if $AJA^* = J$. A transformation satisfying this condition is called canonical or symplectic.

The space $R^3 \times R^3$ of the ξ's is called the phase space. For a system of N particles we would use $R^{3N} \times R^{3N}$.

We wish to point out that for many fundamental physical systems, the phase space is a manifold rather than Euclidean space. These arize when constraints are present. For example the phase space for the motion of the rigid body is the tangent bundle of the group $SO(3)$ of 3×3 orthogonal matrices with determinant $+1$.

To generalize the notion of a Hamiltonian system, we first need to geometrize the symplectic matrix J above. In infinite dimensions there is a technical point however which is important. We give a discussion of this in the following.

Strong and Weak Nondegenerate Bilinear Forms.

Let E be a Banach space and $B : E \times E \to R$ a continuous bilinear mapping. Then B induces a continuous map $B^b : E \to E^*$, $e \mapsto B^b(e)$ defined through $B^b(e)f = B(e, f)$. We call B weakly nondegenerate if B^b is injective; i.e. $B(e, f) = 0$ for all $f \in E$ implies $e = 0$. We call B nondegenerate or strongly nondegenerate if B^b is an isomorphism. By the open mapping theorem it follows that B is nondegenerate iff B is weakly nondegenerate and B^b is onto.

If E is finite dimensional there is no difference between strong and weak nondegeneracy. However in infinite dimensions the distinction is important to bear in mind.

Let M be a Banach manifold. By a weak Riemannian structure we mean a smooth assignment $x \mapsto \langle , \rangle_x$ of a weakly nondegenerate inner product (not necessarily complete) to each tangent space T_xM. Here smooth means that in local charts $x \in U \subset E \mapsto \langle , \rangle_x \in L_2(E \times E, R)$ is smooth where $L_2(E \times E, R)$ denotes the Banach space of bilinear maps of $E \times E$ to R. Equivalently $\langle , \rangle_x$ is a smooth section of the vector bundle whose fiber at $x \in M$ is $L_2(T_xM \times T_xM, R)$.

By a Riemannian manifold we mean a weak Riemannian manifold in which $\langle , \rangle_x$ is nondegenerate. Equivalently, the topology of $\langle , \rangle_x$ is complete on T_xM, so that the model space E must be isomorphic to a Hilbert space.

For example the L_2 inner product $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$ on $E = C([0,1], R)$ is a weak Riemannian metric on E but not a Riemannian metric.

Symplectic Forms.

Let P be a manifold modelled on a Banach space E. By a symplectic form we mean a two form ω on P such that

(a) ω is closed; $d\omega = 0$

(b) for each $x \in P$, $\omega_x : T_xP \times T_xP \to R$ is nondegenerate.

If ω_x in (b) is weakly nondegenerate, we speak of a <u>weak symplectic form</u>.

The need for weak symplectic forms will be clear from examples given below. For the moment the reader may wish to assume P is finite dimensional in which case the distinction vanishes.

If (b) is dropped we refer to ω as a presymplectic form. This case will be referred to later. The first result is referred to as <u>Darboux's theorem</u>. Our proof follows Weinstein [1]. The method is also useful in Morse theory; see Palais [5].

<u>Theorem</u>. <u>Let ω be a symplectic form on the Banach manifold P. For each $x \in P$ there is a local coordinate chart about x in which ω is constant</u>.

<u>Proof</u>. We can assume $P = E$ and $x = 0 \in E$. Let ω_1 be the constant form equalling $\omega_0 = \omega(0)$. Let $\tilde{\omega} = \omega_1 - \omega$ and $\omega_t = \omega + t\tilde{\omega}$, $0 \leq t \leq 1$. For each t, $\omega_t(0) = \omega(0)$ is nondegenerate. Hence by openness of the set of linear isomorphisms of E to E^*, there is a neighborhood of 0 on which ω_t is nondegenerate for all $0 \leq t \leq 1$. We can assume that this neighborhood is a ball. Thus by the Poincaré lemma (appendix 1) $\tilde{\omega} = d\alpha$ for some one form α. We can suppose $\alpha(0) = 0$.

Define a vector field X_t by $i_{X_t}\omega_t = -\alpha$ which is possible since ω_t is nondegenerate. Moreover, X_t will be smooth. Since

$X_t(0) = 0$ we can, from the local existence theory restrict to a sufficiently small ball on which the integral curves will be defined for a time at least one.

Now let F_t be the flow of X_t . The connection between Lie derivatives and flows still holds for time dependent vector fields, so we have

$$\frac{d}{dt}(F_t^* \omega_t) = F_t^*(L_{X_t} \omega_t) + F_t^* \frac{d}{dt} \omega_t$$

$$= F_t^* d i_{X_t} \omega_t + F_t^* \tilde{\omega}$$

$$= F_t^*(d(-\alpha) + \tilde{\omega}) = 0 \ .$$

Therefore, $F_1^* \omega_1 = F_0^* \omega_0 = \omega$, so F_1 provides the chart transforming ω to the constant form ω_1 . □

Of course such a result cannot be true for riemannian structures (otherwise they would be flat). Darboux's theorem is not true for weak symplectic forms. See Marsden [4]. Recently A. Tromba has found some useful sufficient conditions to cover the weak case.

Corollary. If P is finite dimensional and ω is a symplectic form then

(a) P is even dimensional, say $\dim P = m = 2n$

(b) locally about each point there are coordinates $x^1, \ldots, x^n, y^1, \ldots, y^n$ such that

$$\omega = \sum_{1}^{n} dx^i \wedge dy^i .$$

Such coordinates are called <u>canonical</u>.

<u>Proof</u>. By elementary linear algebra, any skew symmetric bilinear form which is nondegenerate has the canonical form $\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ where I is the $n \times n$ identity. This is the matrix version of (b) pointwise on M. The result now follows from Darboux's theorem. □

The corollary actually has a generalization to infinite dimensions. Clearly it is just a result on the canonical form of a skew symmetric bilinear mapping. First some notation. Let E be a real vector space. By a <u>complex structure</u> on E we mean a linear map $J : E \to E$ such that $J^2 = -I$. By setting $ie = J(e)$ one then gives E the structure of a complex vector space. We now show that a symplectic form is the imaginary part of an inner product. (cf. Cook [1]).

<u>Proposition</u>. <u>Let</u> H <u>be a real Hilbert space and</u> B <u>a skew symmetric weakly nondegenerate bilinear form on</u> H. <u>Then there exists a complex structure</u> J <u>on</u> H <u>and a real inner product</u> s <u>such that</u>

$$s(x, y) = B(Jx, y) .$$

<u>Setting</u>

$$h(x, y) = s(x, y) + iB(x, y) ,$$

h <u>is a hermetian inner product</u>. <u>Finally</u>, h <u>or</u> s <u>is complete on</u>

H *iff* B *is nondegenerate*.

Proof. Let $\langle , \rangle$ be the given complete inner product on H. By the Riesz theorem, $B(x, y) = \langle Ax, y\rangle$ for a bounded linear operator $A : H \to H$. Since B is skew, we find $A^* = -A$.

Since B is weakly nondegenerate, A is injective. Now $-A^2 \geq 0$, and from $A = -A^*$ we see that A^2 is injective. Let P be a symmetric non-negative square root of $-A^2$. Hence P is injective. Since $P = P^*$, P has dense range. Thus P^{-1} is a well defined unbounded operator. Set $J = AP^{-1}$, so that $A = JP$. From $A = -A^*$ and $P^2 = -A^2$, we find that J is orthogonal and $J^2 = -1$. Thus we may assume J is a bounded operator. Moreover J is symplectic in the sense that $B(Jx, Jy) = B(x, y)$. Define $s(x, y) = B(Jx, y) = \langle Px, y\rangle$ since $A = JP = PJ$. Thus s is an inner product on H. Finally, it is a straightforward check to see that h is a hermetian inner product. For example; $h(ix, y) = s(Jx, y) + iB(Jx, y) = -B(x, y) + is(x, y) = ih(x, y)$. The proposition follows. □

Canonical Symplectic Forms.

We recall that a Banach space E is reflexive iff the canonical injection $E \to E^{**}$ is onto. For instance any finite dimensional or Hilbert space is reflexive. The L_p spaces, $1 < p < \infty$ are reflexive, but $C([0, 1], R)$ with the sup norm is not.

Let M be a manifold modelled on a Banach space E. Let

T^*M be its cotangent bundle, and $\tau^* : T^*M \to M$ the projection. Define the canonical one form θ on T^*M by

$$\theta(\alpha_m)W = -\alpha_m T\tau^*(W)$$

where $\alpha_m \in T_m^*M$ and $W \in T_{\alpha m}(T^*M)$. In a chart $U \subset E$, this formula is the same as saying

$$\theta(x, \alpha) \cdot (e, \beta) = -\alpha(e)$$

where $(x, \alpha) \in U \times E^*$, $(e, \beta) \in E \times E^*$. If M is finite dimensional, this says

$$\theta = -\Sigma\, p_i dq^i$$

where $q^1,\ldots,q^n, p_1,\ldots,p_n$ are coordinates for T^*M .

The canonical two form is defined by $\omega = d\theta$. Locally, using the formula for d from table one, p. 19,

$$\omega(x, \alpha)\cdot((e_1, \alpha_1), (e_2, \alpha_2)) = \{\alpha_2(e_1) - \alpha_1(e_2)\}$$

or, in the finite dimensional case,

$$\omega = \Sigma\, dq^i \wedge dp_i \ .$$

Proposition (a) The form ω is a weak symplectic form on $P = T^*M$

(b) ω is symplectic iff E is reflexive .

Proof. (a) Suppose $\omega(x, \alpha)((e_1, \alpha_1),(e_2, \alpha_2)) = 0$ for all (e_2, α_2) .

Setting $e_2 = 0$ we get $\alpha_2(e_1) = 0$ for all $\alpha_2 \in E^*$. By the Hahn-Banach theorem, this implies $e_1 = 0$. Setting $\alpha_2 = 0$ we get $\alpha_1(e_2) = 0$ for all $e_2 \in E$, so $\alpha_1 = 0$.

(b) Suppose E is reflexive. We must show that the map $\omega^b : E \times E^* \to (E \times E^*)^* = E^* \times E^{**}$, $\omega^b(e_1, \alpha_1)\cdot(c_2, \alpha_2) = \{\alpha_2(e_1) - \alpha_1(e_2)\}$ is onto. Let $(\beta, f) \in E^* \times E^{**} \approx E^* \times E$. We can take $e_1 = f$, $\alpha_1 = -\beta$; then (e_1, α_1) is mapped to (β, f) under $2\omega^b$. Conversely if ω^b is onto, then for $(\beta, f) \in E^* \times E^{**}$, there is (e_1, α_1) such that $f(\alpha_2) + \beta(e_2) = \alpha_2(e_1) - \alpha_1(e_2)$ for all e_2, α_2. Setting $e_2 = 0$ we see $f(\alpha_2) = \alpha_2(e_1)$, so $i:E \to E^{**}$ is onto. □

Symplectic Forms induced by Metrics.

If $\langle , \rangle_x$ is a weak Riemannian metric on M, we have a smooth map $\varphi : TM \to T^*M$ defined by $\varphi(v_x)w_x = \langle v_x, w_x\rangle_x$, $x \in M$. If $\langle , \rangle$ is a (strong) Riemannian metric it follows from the implicit function theorem that φ is a diffeomorphism of TM onto T^*M. In any case, set $\Omega = \varphi^*(\omega)$ where ω is the canonical form on T^*M. Clearly Ω is exact since $\Omega = d(\varphi^*(\theta))$.

Proposition. (a) _If_ $\langle , \rangle_x$ _is a weak metric, then_ Ω _is a weak symplectic form. In a chart_ U _for_ M _we have_

$$\Omega(x,e)((e_1,e_2), (e_3,e_4)) = D_x\langle e,e_1\rangle_x e_3 - D_x\langle e,e_3\rangle_x e_1 + \langle e_4,e_1\rangle_x - \langle e_2,e_3\rangle_x$$

where D_x _denotes the derivative with respect to_ x.

(b) <u>If</u> $<,>_x$ <u>is a strong metric and</u> M <u>is modelled on a reflexive space, then</u> Ω <u>is a symplectic form</u>.

(c) $\Omega = d\theta$ <u>where, locally</u>, $\theta(x, e)(e_1, e_2) = -<e, e_1>_x$.

<u>Note</u>. In the finite dimensional case, the formula for Ω becomes

$$\Omega = \Sigma\, g_{ij} dq^i \wedge d\dot{q}^i + \Sigma \frac{\partial g_{ij}}{\partial q^k} \dot{q}^i dq^j \wedge dq^k$$

where $q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n$ are coordinates for TM .

<u>Proof</u>. By definition of pull-back, $\Omega(x, e)((e_1, e_2), (e_3, e_4)) = \omega(x, e)(D\varphi_{(x,e)}(e_1, e_2), D\varphi_{(x,e)}(e_3, e_4))$. But clearly $D\varphi_{(x,e)}(e_1, e_2) = (e_1, D_x<e, \cdot>_x e_1 + <e_2, \cdot>_x)$, so the formula for Ω follows from that for ω . To check weak nondegeneracy, suppose $\Omega_{(x,e)}((e_1, e_2),(e_3, e_4)) = 0$ for all (e_3, e_4) . Setting $e_3 = 0$ we find $<e_4, e_1>_x = 0$ for all e_4 , whence $e_1 = 0$. Then we obtain $<e_2, e_3>_x = 0$, so $e_2 = 0$. Part (b) follows from the easy fact that the transform of a symplectic form by a diffeomorphism is still symplectic. □

The above result holds equally well for pseudo-Riemannian manifolds.

Note that if M = H is a Hilbert space with the constant inner product, then ω is, on H × H which we may identify with $\mathcal{H}$ – the complexified Hilbert space, equal to the imaginary part of the inner product: Let $e = e_1 + ie_2$, $f = f_1 + if_2$. Then

$$\langle e, f\rangle = (\langle e_1, f_1\rangle + \langle e_2, f_2\rangle) - i(\langle e_1, f_2\rangle - \langle e_2, f_1\rangle)$$

so

$$\omega(e, f) = -\mathrm{Im}\langle e, f\rangle .$$

Canonical Transformations.

Let P, ω be a weak symplectic manifold; i.e. ω is a weak symplectic form on P. A (smooth) map $f : P \to P$ is called canonical or symplectic when $f^*\omega = \omega$. It follows that $f^*(\omega \wedge \ldots \wedge \omega) = \omega \wedge \ldots \wedge \omega$ (k times). If P is 2n dimensional, $\mu = \omega \wedge \ldots \wedge \omega$ (n times) is nowhere vanishing; by a computation one finds μ to be a multiple of the Lebesgue measure in canonical coordinates. We call μ the phase volume or the Liouville form. Thus a symplectic map preserves the phase volume, and is necessarily a local diffeomorphism.

We briefly discuss symplectic maps induced by maps on the base space of a cotangent bundle.

Theorem. Let M be a manifold and $f : M \to M$ a diffeomorphism, define the lift of f by

$$T^*f : T^*M \to T^*M \; ; \; T^*f(\alpha_m)v = \alpha_m(Tf \cdot v) \;, \; v \in T^*_{f^{-1}(m)}M .$$

Then T^*f is symplectic and in fact $(T^*f)^*\theta = \theta$, where θ is the canonical one form. (We could, equally well consider diffeomorphisms from one manifold to another.)

Proof. By definition, $(T^*f)^*\theta(W) = \theta(TT^*f\cdot W) =$

$$-T^*f(\alpha_m)\cdot(T\tau^*TT^*f\cdot W) = -T^*f(\alpha_m)\cdot(T(\tau^*\circ T^*f)\cdot W)$$
$$= -\alpha_m\cdot(Tf\cdot T(\tau^*\circ T^*f)\cdot W)$$
$$= -\alpha_m\cdot(T(f\circ\tau^*\circ T^*f)\cdot W)$$
$$= -\alpha_m\cdot(T\tau^*\cdot W) = \theta(W)$$

since, by construction, $f\circ\tau^*\circ T^*f = \tau^*$. □

One can show conversely that any diffeomorphism of $P = T^*M$ which preserves θ is the lift of some diffeomorphism of M . But, on the other hand, there are many other symplectic maps of P which are not lifts.

Corollary. *Let* M *be a weak Riemannian manifold and* Ω *the corresponding weak symplectic form. Let* $f : M \to M$ *be a diffeomorphism which is an isometry:* $\langle v, w\rangle_x = \langle Tf\cdot v, Tf\cdot w\rangle_{f(x)}$. *Then* $Tf : TM \to TM$ *is symplectic.*

Proof. The result is immediate from the above and the fact that $T^*f \circ \varphi \circ Tf = \varphi$ where $\varphi : TM \to T^*M$ is as on p.34. □

Hamiltonian Vector Fields and Poisson Brackets.

Definition. Let P , ω be a weak symplectic manifold. A vector field $X : D \to TP$ with manifold domain D is called *Hamiltonian* if there is a C^1 function $H : D \to R$ such that

$$i_X\omega = dH$$

as 1-forms on D . We say X is <u>locally Hamiltonian</u> if $i_X\omega$ is closed.

We write $X = X_H$ because usually in examples one is given H and then one constructs the Hamiltonian vector field X_H .

Because ω is only weak, given $H : D \to R$, X_H need not exist. Also, even if H is smooth on all of P , X_H will in general be defined only on a certain subset of P , but where it is defined, it is unique.

The condition $i_{X_H}\omega = dH$ reads

$$\omega_x(X_H(x), v) = dH(x)\cdot v \ ,$$

$x \in D$, $v \in T_xD \subset T_xM$. From this we note that, necessarily, for each $x \in D$, $dH(x) : T_xD \to R$ is extendable to a bounded linear functional on T_xP .

The relation $\omega(X_H, v) = dH\cdot v$ is the geometrical formulation of the same condition $X_H(\xi) = J\cdot \mathrm{grad}\, H(\xi)$ with which we motivated the discussion.

<u>Some Properties of Hamiltonian Systems</u>.

We now give a couple of simple properties of Hamiltonian systems. The proofs are a bit more technical for densely defined vector fields so for purposes of these theorems we work with C^r vector fields.

Theorem. *Let X_H be a Hamiltonian vector field on the symplectic manifold P, ω and let F_t be the flow of X_H. Then*

(i) F_t *is symplectic*, $F_t^*\omega=\omega$

and (ii) *energy is conserved*; $H \circ F_t = H$.

Proof. (i) Since F_0 = identity, it suffices to show that $\frac{d}{dt} F_t^*\omega = 0$. But by lecture 1,

$$\frac{d}{dt} F_t^*\omega(x) = F_t^*(L_{X_H}\omega)(x)$$

$$= F_t^*[di_{X_H}\omega](x) + F_t^*[i_{X_H}d\omega](x)$$

The first term is zero because it is ddH and the second is zero because $d\omega = 0$.

(ii) By the chain rule,

$$\frac{d}{dt} (H\circ F_t)(x) = dH(F_t(x)) \cdot X_H(F_t(x))$$

$$= \omega_{F_t(x)}(X_H(F_t(x)) , X_H(F_t(x)))$$

but this is zero in view of the skew symmetry of ω. □

An immediate corollary of (i) is *Liouville's theorem*: F_t *preserves the phase volume*. It seems likely that a version of Liouville's theorem holds in infinite dimensions as well. The phase volume would be a Wiener measure induced by the symplectic form.

More generally than (ii) one can show that for any function $f : P \to R$,

$$\frac{d}{dt} f \circ F_t = \{f, H\} \circ F_t$$

where $\{f, g\} = \omega(X_f, X_g)$ is the Poisson bracket; in fact it is easy to see that

$$\{f, g\} = L_{X_g} f .$$

(Note that $F_t^* f = f \circ F_t$ for functions.)

<u>The Wave Equation as a Hamiltonian System.</u>

The wave equation for a function $u(x, t)$, $x \in R^n$, $t \in R$ is given by

$$\frac{\partial^2 u}{\partial t^2} = \Delta u + m^2 u \ , \quad m \geq 0$$

with u given at $t=0$, We consider

$$P = H^1(R^n) \times L_2(R^n)$$

where H^1 consists of functions in L_2 whose first derivatives are also in L_2 . Let

$$D = H^2 \times H^1$$

and

$$X_H(u, \dot{u}) = (\dot{u}, \Delta u + m^2 u)$$

with symplectic form that associated with the L_2 metric

$$\omega((u, \dot{u}), (v, \dot{v})) = \int \dot{v}u - \int \dot{u}v \quad .$$

(Recall that there is always an associated complex structure -- in this case that of $L_2(R, C)$; in fact there is also one making the flow of X_H unitary as in Cook [1], at least if $m > 0$) . Define

$$H(u, \dot{u}) = \frac{1}{2} \int \dot{u}^2 + \frac{1}{2} \int \|\nabla u\|^2 \quad .$$

It is an easy verification (integration by parts) that X_H , ω and H are in the proper relation, so in this sense the wave equation is Hamiltonian.

That this equation has a flow on P follows from the hyperbolic version of the Hille-Yosida theorem stated in lecture 1.

The Schrodinger Equation.

Let $P = \mathcal{H}$ a complex Hilbert space with $\omega = \mathrm{Im} \langle , \rangle$. Let H be a self adjoint operator with domain D and let

$$X_H(\varphi) = iH\cdot\varphi$$

and

$$H(\varphi) = \langle H\varphi, \varphi\rangle/2 \quad , \varphi \in D \ .$$

Again it is easy to check that ω , X_H and H are in the correct relation.

In this sense X_H is Hamiltonian. Note that $\psi(t)$ is an integral curve of X_H if

$$\frac{1}{i}\frac{d\psi}{dt} = H\psi \ ,$$

the abstract Schrodinger equation of quantum mechanics.

That X_H has a flow is another case of the Hille-Yosida theorem called Stone's theorem; i.e. if H is self adjoint, then iH generates a one parameter unitary group, denoted e^{itH}.

We know from general principles that the flow e^{itH} will be symplectic. The additional structure needed for unitarity is exactly complex linearity.

We shall return to quantum mechanical systems in a later lecture.

We next turn our attention to geodesics and more generally to Lagrangian systems.

The Spray of a Metric.

Let M be a weak Riemannian manifold with metric $\langle , \rangle_x$ on the tangent space T_xM. We now wish to define the spray S of the metric $\langle , \rangle_x$. This should be a vector field on TM; $S : TM \to T^2M$ whose integral curves project onto geodesics. Locally, if $(x, v) \in T_xM$, write $S(x, v) = ((x, v), (v, \gamma(x, v)))$. If M is finite dimensional, the geodesic spray is given by putting $\gamma^i(x, v) = -\Gamma^i_{jk}(x)v^i v^k$. In the general case, the correct definition for γ is

$$(1) \qquad \langle \gamma(x, v), w \rangle_x = \frac{1}{2} D_x \langle v, v \rangle_x \cdot w - D_x \langle v, w \rangle_x \cdot v$$

where $D_x\langle v, v\rangle_x \cdot w$ means the derivative of $\langle v, v\rangle_x$ with respect to x in the direction of w. In the finite-dimensional case, the right hand side of (1) is given by

$$\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k} v^i v^j w^k - \frac{\partial g_{ij}}{\partial x^k} v^i w^j v^k ,$$

which is the same as $-\Gamma^i_{jk} v^j v^k w_i$. So with this definition of γ, S is taken to be the spray. The verification that S is well-defined independent of the charts is not too difficult. Notice that γ is quadratic in v. One can also show that S is just the Hamiltonian vector field on TM associated with the kinetic energy $\frac{1}{2}\langle v, v\rangle$. This will actually be done below; cf. Abraham [2] and Chernoff-Marsden [1].

The point is that the definition of γ in (1) makes sense in the infinite as well as the finite dimensional case, whereas the usual definition of Γ^i_{jk} makes sense only in finite dimensions. This then gives us a way to deal with geodesics in infinite dimensional spaces.

Equations of Motion in a Potential.

Let $t \mapsto (x(t), v(t))$ be an integral curve of S. That is:

$$\dot{x}(t) = v(t) \ ; \quad \dot{v}(t) = \gamma(x(t), v(t)) . \tag{2}$$

These are the equations of motion in the absence of a potential. Now let $V : M \to R$ (the potential energy) be given. At each x, we have the differential of V, $dV(x) \in T_x^*M$, and we define grad $V(x)$ by:

(3) $$\langle \text{grad } V(x), w\rangle_x \equiv dV(x)\cdot w \ .$$

It is a definite assumption that grad V exists, since the map $T_xM \to T_x^*M$ induced by the metric is not necessarily bijective.

The equation of motion in the potential field V is given by:

(4) $$\dot{x}(t) = v(t) \ ; \quad \dot{v}(t) = \gamma(x(t), v(t)) - \text{grad } V(x(t)) \ .$$

The total energy, kinetic plus potential, is given by $H(v_x) = \frac{1}{2}\|v_x\|^2 + V(x)$. It is actually true that the vector field X_H determined by H and the symplectic structure on TM induced by the metric is given by (4). This will be part of a more general derivation of Lagrange's equations below.

Lagrangian Systems.

We now want to generalize the idea of motion in a potential to that of a Lagrangian system; these are, however, still special types of Hamiltonian systems. See Abraham [2] for an alternative exposition of the finite dimensional case, and Marsden [1], and Chernoff-Marsden [1] for additional results.

We begin with a manifold M and a given function $L : TM \to R$ called the Lagrangian. In case of motion in a potential, one takes

$$L(v_x) = \frac{1}{2} \langle v_x, v_x \rangle - V(x)$$

which differs from the energy in that we use $-V$ rather than $+V$.

Now L defines a map, called the <u>fiber derivative</u>, $FL : TM \to T^*M$ as follows: let $v, w \in T_xM$. Then set

$$FL(v) \cdot w \equiv \frac{d}{dt} L(v + tw) \Big|_{t=0}$$

That is, $FL(V) \cdot w$ is the derivative of L along the fiber in direction w .

In case of $L(v_x) = \frac{1}{2} \langle v_x, v_x \rangle_x - V(x)$, we see that $FL(v_x) \cdot w_x = \langle v_x, w_x \rangle_x$ so we recover the usual map of $TM \to T^*M$ associated with the bilinear form $\langle , \rangle_x$.

As we saw above, T^*M carries a canonical symplectic form ω . Using FL we obtain a closed two form ω_L on TM by

$$\omega_L = (FL)^* \omega \, .$$

In fact a straightforward computation yields the following local formula for ω_L : if M is modeled on a linear space E , so locally TM looks like $U \times E$ where $U \subset E$ is open, then $\omega_L(u, e)$ for $(u, e) \in U \times E$ is the skew symmetric bilinear form on $E \times E$ given by

$$\omega_L(u, e) \cdot ((e_1, e_2), (e_3, e_4)) = D_1(D_2 L(u, e) \cdot e_3) \cdot e_1$$

$$- D_1(D_2 L(u, e) \cdot e_3) \cdot e_1 + D_2 D_2 L(u, e) \cdot e_4 \cdot e_1 - D_2 D_2 L(u, e) \cdot e_2 \cdot e_3$$

where D_1 , D_2 denote the indicated partial derivatives of L .

It is easy to see that ω_L is (weakly) nondegenerate if $D_2D_2L(u, e)$ is (weakly) nondegenerate. But we want to also allow degenerate cases for later purposes. In case of motion in a potential, nondegeneracy of ω_L amounts to nondegeneracy of the metric $\langle , \rangle_x$. The action of L is defined by $A : TB \to R$, $A(v) = FL(v)\cdot v$, and the energy of L is $E = A - L$. In charts,

$$E(u, e) = D_2L(u, e)\cdot e - L(u, e)$$

and in finite dimensions it is the expression

$$E(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i - L(q, \dot{q}) ,$$

(summation convention!)

Now given L , we say that a vector field Z on TM is a Lagrangian vector field or a Lagrangian system for L if the Lagrangian condition holds:

$$\omega_L(v)(Z(v), w) = dE(v)\cdot w$$

for all $v \in T_bM$, and $w \in T_v(TM)$. Here, dE denotes the differential of E .

Below we shall see that for motion in a potential, this leads to the same equations of motion which we found above.

If ω_L were a weak symplectic form there would be at most one such Z . The fact that ω_L may be degenerate however means that Z is not uniquely determined by L so that there is some arbitrariness in what we may choose for Z . Also if ω_L is degenerate, Z may not even exist. If it does, we say that we can define consistent equations of motion. These ideas have been discussed in the finite dimensional case by Dirac [1] and Kunzle [1].

The dynamics is obtained by finding the integral curves of Z ; that is the curves $v(t)$ such that $v(t) \in TM$ satisfies $(dv/dt)(t) = Z(v(t))$. From the Lagrangian condition it is trivial to check that energy is conserved even though L may be degenerate:

Proposition. *Let* Z *be a Lagrangian vector field for* L *and let* $v(t) \in TM$ *be an integral curve of* Z . *Then* $E(v(t))$ *is constant in* t .

Proof. By the chain rule,

$$\frac{d}{dt} E(v(t)) = dE(v(t)) \cdot v'(t) = dE(v(t)) \cdot Z(v(t))$$

$$= 2\omega_L(v(t))(Z(v(t)), Z(v(t)) = 0$$

by the skew symmetry of ω_L . □

We now want to generalize our previous local expression for the spray of a metric, and the equations of motion in the presence of a potential. In the general case the equations are called "Lagrange's equations".

Proposition. Let Z *be a Lagrangian system for* L *and suppose* Z *is a second order equation (that is, in a chart* $U \times E$ *for* TM, $Z(u, e) = (e, Z_2(u, e))$ *for some map* $Z_2 : U \times E \to E$). *Then in the chart* $U \times E$, *an integral curve* $(u(t), v(t)) \in U \times E$ *of* Z *satisfies Lagrange's equations:*

$$(1) \qquad \begin{cases} \dfrac{du}{dt}(t) = v(t) \\ \\ \dfrac{d}{dt}(D_2L(u(t), v(t))\cdot w = D_1L(u(t), v(t))\cdot w \end{cases}$$

for all $w \in E$. *In case* L *is nondegenerate we have*

$$(2) \qquad \frac{dv}{dt} = \{D_2D_2L(u, v)\}^{-1}\{D_1L(u, v) - D_1D_2L(u, v)\cdot v\} .$$

In case of motion in a potential, (2) reduces readily to the equations we found previously defining the spray and gradient.

Proof. From the definition of the energy E we have
$dE(u, e)\cdot(e_1, e_2) = D_1(D_2L(u, e)\cdot e_1 + D_2D_2L(u, e)\cdot e\cdot e_2 - D_1L(u, e)\cdot e_1$.
Locally we may write $Z(u, e) = (e, Y(u, e))$ as Z is a second order equation. Using the formula for ω_L, the condition on Y may be written, after a short computation:

$$D_1L(u, e)\cdot e_1 = D_1(D_2L(u, e)\cdot e_1)\cdot e + D_2(D_2L(u, e)\cdot Y(u, e))\cdot e_1$$
$$\text{for all } e_1 \in E .$$

This is the formula (2) above. Then, if $(u(t), v(t))$ is an integral curve of Z we obtain, using dots to denote time differentiation,

$$D_1L(u, e)\cdot e_1 = D_1(D_2L(u, \dot{u})\cdot e_1\cdot \dot{u} + D_2D_2L(u, \dot{u})\cdot \ddot{u}\cdot e_1$$

$$= \frac{d}{dt}D_2L(u, \dot{u})\cdot e_1$$

by the chain rule. □

From these calculations one sees that if ω_L is nondegenerate Z is automatically a second order equation (cf. Abraham [2]). Also, the condition of being second order is intrinsic; Z is second order if $T\pi \circ Z =$ identity , where $\pi : TM \to M$ is the projection. See Abraham [2], or Lang [1].

Often L is obtained in the form

$$L(u, \dot{u}) = \int_Q \mathcal{L}(u, \frac{\partial u}{\partial x^k}, \dot{u})d\mu$$

for a Lagrangian density $\mathcal{L}$ and μ some volume element on some manifold Q . Then M is a space of functions on Q or more generally sections of a vector bundle over Q . In this case, Lagrange's equations may be converted to the usual form of Lagrange's equations for a density $\mathcal{L}$. We shall see how this is done in a couple of special cases in later lectures. (See also Marsden [1]).