Symplectic Geometry
Lecture 9

Hamilton's principle
Legendre Transformations
Thermodynamics
Generating functions

1 Variational principles.

2 Hamilton's principle.

3 The equations for geodesics - the Christoffel symbols, parallel transport.

4 The Legendre transform.

5 Lagrangian submanifolds and equilibrium thermodynamics.

6 Canonical relations and their generating functions.

We did some heavy lifting during the past two lectures. want to spend today tying up some loose ends.

Hamilton's principle

We saw in Lecture 7 that if L is a function on TQ and if

$$I[C] = \int_C L(C(t), C'(t))dt$$

then

$$I[C] = \int_{\overline{C}} \alpha - \int_{t_1}^{t_2} H(\overline{C})t.$$

Here α is the canonical one form on T^*M, $\overline{C}$ is the image in T^*Q of the curve $C_T = (C, C')$ in TQ via the Legendre transformation

$$\mathfrak{L} : TQ \to T^*Q$$

determined by L, and H is the Hamiltonian function on T^*Q associated to L.

$$I[C] = \int_C L(C(t), C'(t))dt \quad = \int_{\overline{C}} \alpha - \int_{t_1}^{t_2} H(\overline{C})t.$$

From this we derived the fact that a curve γ in T^*Q is a trajectory of the vector field X_H if and only if two conditions are satisfied:

- $\gamma = \mathfrak{L}(C_T) = \mathfrak{L}(C, C')$ and
- C is an extremal of I (relative to variations with fixed end points).

In the proof of this assertion we used the fact that a curve γ in T^*Q was of the form $\gamma = \mathfrak{L}(C_T)$ if and only if the first half of Hamilton's equations were satisfied, and then verified the second assertion above by varying the curve C on Q.

But we can consider a different variational principle - where we vary among *all* possible smooth curves on T^*Q, not necessarily curves on T^*Q arising from curves on Q. We will formulate this in the slightly more general setting of exact symplectic manifolds. We recall that a symplectic manifold (M, ω) is called **exact** if there is a one form α on M such that $d\alpha = -\omega$, and that we have chosen such a one form.

Let H be a smooth function on an exact symplectic manifold (M, ω, α), $\omega = -d\alpha$. For any smooth curve $\gamma : [a, b] \to M$ define

$$\mathcal{A}_H(\gamma) := \int_\gamma \alpha - \int_a^b H \circ \gamma(t) dt.$$

Let Λ_1 and Λ_2 be submanifolds of M such that $\alpha = 0$ when pulled back to Λ_1 or Λ_2.

Theorem 1 *For a curve $\gamma : [a,b] \to M$ with $\gamma(a) \in \Lambda_1$ and $\gamma(b) \in \Lambda_2$ the following are equivalent:*

- *γ is an extremal of $\mathcal{A}_H$ relative to variations with $\gamma(a) \in \Lambda_1$ and $\gamma(b) \in \Lambda_2$.*
- *γ is an extremal of $\mathcal{A}_H$ relative to variations with fixed end points.*
- *γ is a trajectory of X_H.*

Proof. Let Y be a vector field defined in a neighborhood of γ with $Y(\gamma(a)) \in T_{\gamma(a)}\Lambda_1$ and $Y(\gamma(b)) \in T_{\gamma(b)}\Lambda_2$. Let ϕ_s be the flow generated by Y. Then

$$\frac{d}{ds}\mathcal{A}_H(\phi_s \circ \gamma)_{|s=0} = \int_\gamma D_Y\alpha - \int_a^b [D_Y H) \circ \gamma](t)dt.$$

By Weil's formula the first integral is

$$\int_\gamma i(Y)d\alpha + \int_\gamma d\alpha(Y) = \int_\gamma i(Y)d\alpha + \alpha(Y(\gamma(b))) - \alpha(Y(\gamma(a))).$$

Under either of the first two assumptions, $\alpha(Y(\gamma(b))) = \alpha(Y(\gamma(a))) = 0$.

So the vanishing of this expression for all Y implies and is equivalent to

$$i(Y)d\alpha(\dot{\gamma}(t)) + (i(Y)dH)(\gamma(t)) = 0$$

for all t and all Y which is the same as saying that γ is a trajectory

The equations for geodesics - the Christoffel symbols, parallel transport.

The Euler-Lagrange equations for geodesics are

$$\frac{dq_i}{dt} = \dot{q}_i, \qquad \frac{d\left(\frac{\partial L}{\partial \dot{q}_i}\right)}{dt} = \frac{\partial L}{\partial q_i}$$

where $L = \frac{1}{2}\sum_j g_{ij}(q_1, \ldots, q_n)\dot{q}_i)\dot{q}_j$ so that

$$\frac{\partial L}{\partial \dot{q}_i} = \sum_j g_{ij}(q_q, \ldots, q_n)\dot{q}_j$$

and so the Euler-Lagrange equations become

$$\frac{d}{dt}\left(\sum_{ij} g_{ij}\dot{q}_j\right) = \frac{1}{2}\sum_{k\ell} \frac{\partial g_{k\ell}}{\partial q_i}\dot{q}_k\dot{q}_\ell \qquad .$$

and so the Euler-Lagrange equations become

$$\frac{d}{dt}\left(\sum_{ij} g_{ij}\dot{q}_j\right) = \frac{1}{2}\sum_{k\ell}\frac{\partial g_{k\ell}}{\partial q_i}\dot{q}_k\dot{q}_\ell$$

or

$$\sum_j \frac{d^2 q_i}{dt^2} + \sum_{jk}\frac{\partial g_{ij}}{\partial q_k}\frac{dq_k}{d}\frac{q_j}{dt} = \frac{1}{2}\sum_{k\ell}\frac{dq_k}{dt}\frac{dq_\ell}{dt}.$$

But

$$\sum_{k\ell}\frac{\partial g_{ik}}{\partial q_\ell}\dot{q}_k\dot{q}_\ell = \sum_{k\ell}\frac{\partial g_{i\ell}}{\partial q_k}\dot{q}_k\dot{q}_\ell$$

so the preceding equation becomes

$$\sum_j g_{ij}\frac{d^2 q_j}{dt^2} = \frac{1}{2}\sum_{l\ell}\left(\frac{\partial g_{k\ell}}{\partial q_i} - \frac{\partial g_{ik}}{\partial q_\ell} - \frac{\partial g_{i\ell}}{\partial q_k}\right)\frac{dq_k}{dt}\frac{dq_\ell}{dt}.$$

$$\sum_j g_{ij} \frac{d^2 q_j}{dt^2} = \frac{1}{2} \sum_{l\ell} \left(\frac{\partial g_{k\ell}}{\partial q_i} - \frac{\partial g_{ik}}{\partial q_\ell} - \frac{\partial g_{i\ell}}{\partial q_k} \right) \frac{dq_k}{dt} \frac{dq_\ell}{dt}.$$

Let (g^{ij}) be the inverse matrix to (g_{ij}) and define the **Christoffel symbols**

$$\Gamma^j_{k\ell} := \frac{1}{2} \sum_i g^{ij} \left(\frac{\partial g_{k\ell}}{\partial q_i} - \frac{\partial g_{ik}}{\partial q_\ell} - \frac{\partial g_{i\ell}}{\partial q_k} \right). \qquad (1)$$

Then the equations for geodesics can be written as

$$\frac{d^2 q_j}{dt^2} = \sum_{k\ell} \Gamma^j_{k\ell} \frac{dq_k}{dt} \frac{dq_\ell}{dt}. \qquad (2)$$

$$\frac{d^2 q_j}{dt^2} = \sum_{k\ell} \Gamma^j_{k\ell} \frac{dq_k}{dt} \frac{dq_\ell}{dt}. \tag{2}$$

In Riemannian geometry courses it is taught that a Riemann metric determines a unique notion of parallel transport along curves (the Levi-Civita connection) and that the preceding equations are the equations for a curve to be "self-parallel", that is that its tangent vector is transported parallelly along the curve. (See for example my notes available on the web.) So another characterization of geodesics is that they are self parallel curves.

A slightly different way of putting this, due to Cartan, is: Given any curve on a manifold, there is a notion of "development along this curve"; intuitively we "roll the manifold onto a Euclidean space along the curve". This then gives a curve in Euclidean space. The geodesics are those curves which roll out to straight lines.

So we have many (equivalent) definitions of geodesics. They are curves on Q which

- are projections of trajectories on T^*Q of Q_K where K is the kinetic energy. (This is the definition which we will principally use.)

- projections on Q of the Euler-Lagrange equations for $L = \frac{1}{2}\|v\|^2$ on TQ.

- curves (on Q) which locally minimize arc length.

- curves which locally minimize the kinetic energy.

- curves which are self-parallel.

- curves roll out onto straight lines under development.

- curves which are determined by Einstein's "principle of general covariance".

For this last item, which is perhaps the most beautiful characterization, and not very well known, see my notes on the web.

The Legendre transform.

Let L be a smooth function on a vector space V. Then dL is a linear differential form, i.e. a map which assigns to each $x \in V$ the element $dL_x \in T_x^*V$. Since V is a vector space, we have a canonical identification of T_xV^* with V^* for each $x \in V$. So we can think of dL as defining a map

$$\mathfrak{L} : V \to V^*.$$

The space $T^*V \cong V \oplus V^*$ carries a canonical symplectic structure (determined by the canonical two form on T^*V). So if we think of dL as defining a section $dL : V \to T^*V$ the image of this section is a Lagrangian submanifold Λ of T^*V. Under the identification $T^*V \cong V \oplus V^*$ the submanifold Λ becomes identified as

$$\Lambda = \text{graph } \mathfrak{L}.$$

$$\Lambda = \text{graph } \mathfrak{L}.$$

Conversely, suppose that Λ is a Lagrangian submanifold of T^*V. This means that the canonical one form α_V of T^*V becomes closed when restricted to Λ. If Λ is simply connected, this means that the restriction of α to Λ is equal to $d\chi$ for some function χ. If in addition Λ is connected, then χ is determined up to an additive constant. Let us assume that we have made such a choice of constant. (For a fixed vector space V this choice of constant is not very relevant. But when we have a family of vector spaces, such as T_xQ as x varies over a manifold Q, this choice becomes important.)

If π_V denotes the projection $\pi_V : V \oplus V^* \to V$ and if the restriction of π_V to Λ

$$(\pi_V)_{|\Lambda} : \ \Lambda \to V$$

is a diffeomorphism, then the function L on V given by

$$L = \chi \circ \left((\pi_V)_{|\Lambda}\right)^{-1}$$

gives the function L such that $\Lambda = \text{graph } \mathfrak{L}$ where $\mathfrak{L}$ is the Legendre transformation associated with L.

Now $(V^*)^* \cong V$. In other words, we can also think of $V \oplus V^*$ as the cotangent bundle of V^*. Under this identification, the symplectic form ω gets replaced by its negative. Indeed, if $x_1, \ldots, x_n$ are linear coordinates on V and $y_1, \ldots, y_n$ are the dual coordinates on V^* then thinking of $V \oplus V^*$ as T^*V the canonical two form is

$$\omega = \sum_i dx_i \wedge dy_i$$

whereas the canonical two form for $T^*(V^*)$ is

$$\sum_i dy_i \wedge dx_i.$$

Of course, Λ is a Lagangian submanifold relative to ω if and only if it is a Lagrangian submanifold relative to $-\omega$. The canonical one form α_V is given by

$$\alpha_V = \sum_i y_i dx_i$$

while the canonical one form α_{V^*} is

$$\alpha_{V^*} = \sum_i x_i dy_i.$$

Now we have

$$\alpha_{V^*} = d\left(\sum_i x_i y_y\right) - \alpha_V$$

(on all of $V \oplus V^*$).

$$\alpha_{V^*} = d\left(\sum_i x_i y_y\right) - \alpha_V$$

(on all of $V \oplus V^*$). So the restriction of α_{V^*} to Λ is

$$d\left(\sum_i x_i dy_i - \chi\right).$$

So if the projection of Λ to V^* is also a diffeomorphism, the inverse of $\mathfrak{L}^{-1}$ of the Legendre transformation is also a Legendre transformation given by the function

$$H = \sum_i x_i y_i - L.$$

In short, the "correct" definition of the Legendre transformation is that it is a Lagrangian submanifold of $V \oplus V^*$ (satisfying conditions relative to projections onto each summand).

Lagrangian submanifolds and equilibrium thermodynamics.

In this section I want to briefly explain how the combined first and second laws of thermodynamics can be formulated as saying that the manifold of equilibrium states of a thermodynamical system is a Lagrangian submanifold of a certain symplectic manifold. For the background to this, and for a formulation of equilibrium thermodynamics according to the approach of Born and Caratheodory, see Chapter 22 of volume II of my book with Bamberg, *A course in mathematics for students of physics*. Here is a rapid review:

Systems and states.

The first basic assumption of thermodynamics, common to almost all physical theories, is that we can isolate a portion of the universe which we call a *system*. The system can exist in various *states*. The description of all possible states of the system is usually enormously complicated. For instance, if our system consisted of a gas in some enclosure, the gas might be in turbulent motion in which case we would have to specify the local velocity of each small portion of the gas as part of our description of a state.

Equilibrium states.

The next basic assumption is that among all possible states there is a distinguished class of states called *equilibrium states*. If all interactions of the system with the rest of the universe were held constant, then the system would pass through various states, but tend to a definite equilibrium state. Although the general states of the system are very complicated, the equilibrium states are relatively simple to describe and the set of equilibrium states has the structure of a finite dimensional manifold.

Diathermal contact.

The next assumptions single out certain types of interactions: There exists a special form of interaction between two systems called *diathermal contact*. In this form of interaction there is no observed macroscopic motion or exchange of material. The changes which do occur can be described as follows: Consider the combined system of the two original systems as a new system. In diathermal contact, the equilibrium states of the combined system form a subset of the set of pairs of equilibrium states of the original system.

Thermal equilibrium.

In diathermal contact, the equilibrium states of the combined system form a subset of the set of pairs of equilibrium states of the original system. In other words, if p_1 is an equilibrium sate of the first system and p_2 is an equilibrium state of the second system and the two systems are brought into diathermal contact, the combined system will not necessarily be in equilibrium, but will tend to a definite equilibrium state (q_1, q_2) where q_1 is a new equilibrium state of the first system and q_2 is a new equilibrium state of the second system. We say that q_1 is in thermal equilibrium with q_2.

The zero-th law of thermodynamics.

It is an observed law of nature (sometimes called the *zeroth law of thermodynamics* that thermal equilibrium is an equivalence relation. The equivalence class of all systems at a definite equilibrium is called an abstract temperature. If we fix one definite system and some numerical function of the equilibrium states of this system which takes on different values at different equilibrium states we obtain a numerical function θ on equilibrium states of all systems called an *empirical temperature*. It is part of the assertion of the "zeroth law" that we can choose this function to be a differentiable function on the manifold of equilibrium states of any system, and that two states have the same abstract temperature if and only if they have the same value of θ.

Configurational variables.

In the description of any system, there are certain functions that play an important role. For example, if the system consisted of gas in a flexible container, the total volume would be such a function, as would the total mass of each of the various chemical constituents of the gas. These functions are called *configurational variables.* The configurational variables when restricted to the manifold of equilibrium states are differentiable functions, and these functions, together with θ form a system of local coordinates. In many examples the total volume V is one of the configurational variables. So that $V, x_2, \ldots, x_n$ would be a system of coordinates where $x_2, \ldots, x_n$ are the remaining configurational variables.

Adiabatic interactions.

There exists another important class of interactions of the system with the rest of the universe called *adiabatic interactions*. These interactions have the property that thermal equilibrium is not disturbed unless there is some change in the configurational variables. If, for a period of time, all the interactions of a system are adiabatic, the system is said to be in an adiabatic enclosure. A curve joining two states ρ and ρ' while the system is in an adiabatic enclosure is called an *adiabatic curve*.

The first law.

A curve joining two states ρ and ρ' while the system is in an adiabatic enclosure is called an *adiabatic curve*. The **first law of thermodynamics** is a generalization from experience which asserts that if a system in an adiabatic enclosure is brought from one state ρ to another state ρ' by applying external work, the amount of this work is always the same no matter how this work is applied. There is therefore a function U on the space of states of the system called the *internal energy* such that $W = U(\rho') - U(\rho)$ is the amount of work done on the system when the system is moved adiabatically from ρ to ρ'.

Heat

If the system is *not* in an adiabatic enclosure and is brought from ρ to ρ' by some curve of interactions, the total external work applied need no longer equal $U(\rho') - U(\rho)$. The difference $U(\rho') - U(\rho) - W = Q$ is called the *heat* supplied by the process. Here the work done and the heat supplied depend on the process and not merely on the initial and final states.

Reversible paths.

A curve γ of states is called *reversible* if for every t, the state $\gamma(t)$ is (sufficiently close to) an equilibrium curve. The intuitive idea is that the interactions are proceeding very slowly relative to the "relaxation time" of the system. For example, suppose that the interaction takes place through some piston and rod arrangement which changes the volume of a gas. If the gas is in an equilibrium state and we suddenly push in the piston, the new state of the gas will not be in equilibrium. There will be eddies or shock waves etc. If we then hold the piston in its new position, the gas will settle down to a new equilibrium state. The idea of a reversible curve is that if we move the piston sufficiently slowly, the states will be approximately equilibrium states for all time.

Pressure.

For the case where only the volume of a gas is changing, it is a familiar observation that there exists a function p called the pressure, on the set of equilibrium states so that the work done on the system along any reversible curve is given by $-\int_\gamma p dV$. More generally, there exists a linear differential form τ on the manifold of equilibrium states such that the amount of work done along a reversible curve is $\int_\gamma \tau$ and that τ has the form

$$\tau = \nu_1 dx_1 + \cdots + \nu_n dx_n$$

where the $x_1, \ldots, x_n$ are the configurational variables.

The heat form.

Therefore the heat supplied along the path γ is given by the integral

$$Q(\gamma) = \int_\gamma \sigma$$

where

$$\sigma = dU - \tau.$$

In case $x_1 = V$ is the volume so that $\nu_1 = -p$ is the pressure this reads

$$\sigma = dU + pdV - \nu_2 dx_2 - \cdots - \nu_n dx_n.$$

In particular, by definition, an adiabatic reversible curve is a curve on the manfold of equilibrium states which is a null curve for σ, i.e it satisfies $\langle \sigma, \dot{\gamma}(t) \rangle = 0$ at all points where γ is differentiable.

The second law.

Now to the **second law of thermodynamics**. It is an everyday experience that certain types of work done on a system can not be recovered: Given an system in an adiabatic enclosure, there are certain types of work that can be done on the system such as violent stirring, which are such that you cannot get your work back by any reversible adiabatic curve. More precisely:

> Near any equilibrium state ρ of any system there are arbitrarily close equilibrium states which can not be joined to ρ be any reversible adiabatic curve.

Entropy.

By a certain amount of argumentation, including a non-trivial theorem of Caratheodory (closely related to Darboux's theorem!) this has the following consequence:

> There exists a universal *absolute temperature scale* T (that is a choice of empirical temperature) determined up to arbitrary multiplicative constant with the following property: Fixing this constant (thus choosing T) determines a function S on the set of equilibrium states of any system so that the heat form σ is given by
> $$\sigma = TdS.$$
> The function S is determined up to additive constant (one for each system) and is called the *entropy*.

The combined laws.

I refer to B&S for a proof of this fact. We can combine the first law $\sigma = dU - \tau$ and the second law $TdS = \sigma$ as the combined law

$$dU = TdS + \tau.$$

For notational simplicity I will now limit myself to the case of one configurational variable V so $\tau = -pdV$. The general case differs only in the need for more complicated notation. So now the combined laws read as

$$dU = TdS - pdV. \tag{3}$$

The gas constant.

$$dU = TdS - pdV. \tag{3}$$

Let us set

$$\beta := (RT)^{-1}$$

$$\nu = \beta p.$$

Here R is a conversion factor from units of temperature to units of energy known as the "gas constant". So β has units of inverse energy and ν has units of inverse volume so that βdU and νdV are "dimensionless".

The combined laws and symplectic geometry.

We can consider a four dimensional space $V \oplus V^*$ where V is two dimensional and $(\beta, \nu) \in V$ and $(U, V) \in V^*$ with the symplectic form

$$\omega = d\beta \wedge dU + d\nu \wedge dV.$$

We can now formulate the combined laws as saying that

> The manifold of equilibrium states is a Lagrangian submanifold Λ of $V \oplus V^*$.

The entropy as a generating function.

If we (can) use U and V as coordinates on Λ then the fact that

$$\omega = d(\beta U + \nu dV)$$

vanishes on Λ says that there is a function S on Λ such that

$$dS = \beta dU + \nu dV$$

(if Λ is simply connected). This is just a rewrite of (3). If we introduce U and V as coordinates, this says that S is a generating function for Λ. In other words, if we are given S as an explicit function of U and V then

$$\beta = \frac{\partial S}{\partial U}, \quad \nu = \frac{\partial S}{\partial V}.$$

Planck's function as a generating function.

If we (can) use β and ν as coordinates on Λ , then from

$$\omega = d(-U\,d\beta - V\,d\nu)$$

we see that

$$Y = S - \beta U - \nu V$$

is a generating function for Λ and we recover U and V as functions of β and ν by

$$U = -\frac{\partial Y}{\partial \beta}, \quad V = -\frac{\partial Y}{\partial \nu}.$$

Massieu's function as a generating function.

Most importantly, we can write (on all of $V \oplus V^*$)

$$\omega = d(U d\beta + \nu dV)$$

so that on Λ we have

$$dZ = -U d\beta + \nu dV \tag{4}$$

where the *Massieu function* Z is defined by

$$Z = S - \beta U. \tag{5}$$

The Massieu function and the partition function.

$$dZ = -U d\beta + \nu dV \tag{4}$$

where the *Massieu function* Z is defined by

$$Z = S - \beta U. \tag{5}$$

Thus the Massieu function is a generating function for Λ in terms of the variables β and V. So (4) and (5) give the first and second laws of thermodynamics together with a complete description of the equilibrium states of the system. In statistical mechanics, the Massieu function Z is identified with the "partition function", and thus provides a link between microscopic theory and observed macroscopic phenomena.

Canonical relations and their generating functions.

Let M be a symplectic manifold with symplectic form ω. Then $-\omega$ is also a symplectic form on M. We will frequently write M instead of (M, ω) and by abuse of notation we will let M^- denote the manifold M with the symplectic form $-\omega$.

Let (M_i, ω_i) $i = 1, 2$ be symplectic manifolds. A Lagrangian submanifold Γ of

$$\Gamma \subset M_1^- \times M_2$$

is called a **canonical relation**. So Γ is a subset of $M_1 \times M_2$ which is a Lagrangian submanifold relative to the symplectic form $\omega_2 - \omega_1$ in the obvious notation. So a canonical relation is a relation which is a Lagrangian submanifold.

The graph of a symplectomorphism.

For example, if $f : M_1 \to M_2$ is a symplectomorphism, then $\Gamma_f = \text{graph } f$ is a canonical relation.

Let us put this in the language of fiber products: Let

$$\pi : \Gamma_1 \to M_2$$

denote the restriction to Γ_1 of the projection of $M_1 \times M_2$ onto the second factor. Let

$$\rho : \Gamma_2 \to M_2$$

denote the restriction to Γ_2 of the projection of $M_2 \times M_3$ onto the first factor. Let

$$F \subset M_1 \times M_2 \times M_2 \times M_3$$

be defined by

$$F = (\pi \times \rho)^{-1} \Delta_{M_2}.$$

In other words, F is defined as the fiber product (or exact square)

$$\begin{array}{ccc} F & \xrightarrow{\iota_1} & \Gamma_1 \\ {\scriptstyle \iota_2}\downarrow & & \downarrow{\scriptstyle \pi} \\ \Gamma_2 & \xrightarrow[\rho]{} & M_2 \end{array} . \tag{6}$$

so

$$F \subset \Gamma_1 \times \Gamma_2 \subset M_1 \times M_2 \times M_2 \times M_3.$$

Let pr_{13} denote the projection of $M_1 \times M_2 \times M_2 \times M_3$ onto $M_1 \times M_3$ (projection onto the first and last components). Let π_{13} denote the restriction of pr_{13} to F. Then, as a *set*,

$$\Gamma_2 \circ \Gamma_1 = \pi_{13}(F). \tag{7}$$

The map pr_{13} is smooth, and hence its restriction to any submanifold is smooth. The problems are that

Problems with the definition of composition.

1. F defined as
$$F = (\pi \times \rho)^{-1}\Delta_{M_2},$$
i.e. by (6), need not be a submanifold, and

2. that the restriction π_{13} of pr_{13} to F need not be an embedding.

So we need some additional hypotheses to ensure that $\Gamma_2 \circ \Gamma_1$ is a submanifold of $M_1 \times M_3$. Once we impose these hypotheses we will find it easy to check that $\Gamma_2 \circ \Gamma_1$ is a Lagrangian submanifold of $M_1^- \times M_3$ and hence a canonical relation. I do not want to go into these conditions at the moment.

The cotangent case.

If $M_i = T^*Q_i$ let

$$\varsigma_1 : T^*Q_1 \to T^*Q_1$$

be defined by

$$\varsigma_1(x, \xi) = (x, -\xi).$$

Then $\varsigma_1^*(\alpha_{Q_1}) = -\alpha_{Q_1}$ and hence

$$\varsigma_1^*(\omega_{Q_1}) = -\omega_{Q_1}.$$

We can think of this as saying that ς_1 is a symplectomorphism of M_1 with M_1^- and hence

$$\varsigma_1 \times \mathrm{id}$$

is a symplectomorphism of $M_1 \times M_2$ with $M_1^- \times M_2$. So

$$\Gamma \subset M_1^- \times M_2$$

is a canonical relation if and only if

$$\Lambda := (\varsigma_1 \times \mathrm{id})(\Gamma)$$

is a Lagrangian submanifold of $T^*(Q_1 \times Q_2)$. If Λ projects diffeomorphically onto $Q_1 \times Q_2$ then it will have a generating function ϕ. We then say that $\phi \circ (\varsigma_1 \times \mathrm{id})$ is a generating function for Γ.

For example, let Q be a Riemannian manifold and let

$\phi_t \in C^\infty(Q \times Q)$ be defined by

$$\phi_t(x, y) := \frac{1}{2t} d(x, y)^2, \tag{8}$$

where

$$t \neq 0.$$

Let us compute Λ_ϕ and $(\varsigma_1 \times \mathrm{id})(\Lambda_\phi)$. We first do this computation under the assumption that $Q = \mathbb{R}^n$ and the metric occurring in (8) is the Euclidean metric so that

$$\phi(x, y, t) \quad = \quad \frac{1}{2t} \sum_i (x_i - y_i)^2$$

$$
\begin{aligned}
\phi(x,y,t) &= \frac{1}{2t}\sum_i (x_i - y_i)^2 \\
\frac{\partial \phi}{\partial x_i} &= \frac{1}{t}(x_i - y_i) \\
\frac{\partial \phi}{\partial y_i} &= \frac{1}{t}(y_i - x_i) \quad \text{so} \\
\Lambda_\phi &= \{(x, \frac{1}{t}(x-y), y, \frac{1}{t}(y-x)\} \text{ and} \\
(\varsigma_1 \times \mathrm{id})(\Lambda_\phi) &= \{(x, \frac{1}{t}(y-x), y, \frac{1}{t}(y-x)\}.
\end{aligned}
$$

$$(\varsigma_1 \times \mathrm{id})(\Lambda_\phi) \quad = \quad \{(x, \frac{1}{t}(y-x), y, \frac{1}{t}(y-x)\}.$$

In this last equation let us set $y - x = t\xi$, i.e.

$$\xi = \frac{1}{t}(y-x)$$

which is possible since $t \neq 0$. Then

$$(\varsigma_1 \times \mathrm{id})(\Lambda_\phi) = \{(x, \xi, x + t\xi, \xi)\}$$

which is the graph of the symplectic map

$$(x, \xi) \mapsto (x, x + t\xi).$$

If we identify cotangent vectors with tangent vectors (using the Eulidean metric) then $x + t\xi$ is the point along the line passing through x with tangent vector ξ a distance $t\|\xi\|$ out. The one parameter family of maps $(x, \xi) \mapsto (x, x + t\xi)$ is known as the geodesic flow.

More generally, this same computation works on any geodesically convex Riemannian manifold:

A Riemannian manifold Q is called **geodesically convex** if, given any two points x and y in Q, there is a unique geodesic which joins them. We will show that the above computation of the generating function works for any geodesically convex Riemannian manifold. In fact, we will prove a more general result. Recall that geodesics on a Riemannian manifold can be described as follows: A Riemann metric on a manifold Q is the same as a scalar product on each tangent space T_xQ which varies smoothly with x. This induces an identification of TQ with T^*Q an hence a scalar product $\langle\ ,\ \rangle_x$ on each T_x^*Q. This in turn induces the "kinetic energy" Hamiltonian

$$H(x,\xi) := \frac{1}{2}\langle\xi,\xi\rangle_x.$$

The principle of least action says that the solution curves of the corresponding vector field v_H project under $\pi : T^*Q \to Q$ to geodesics of Q and every geodesic is the projection of such a trajectory. An important property of the kinetic energy Hamilonian is that it is quadratic of degree two in the fiber variables. We will prove a theorem (see Theorem 2 below) which generalizes the above computation and is valid for any Hamiltonian which is homogeneous of degree $k \neq 1$ in the fiber variables and which satisfies a condition analogous to the geodesic convexity theorem. We first recall some facts about homogeneous functions and Euler's theorem.

Let $\alpha = \alpha_Q$ be the canonical one form on T^*Q. From its very definition it follows that $\mathfrak{d}(t)^*\alpha = e^t\alpha$ and hence that $D_{\mathcal{E}}\alpha = \alpha$. Since $\mathcal{E}$ is everywhere tangent to the fiber, $i(\mathcal{E})\alpha = 0$

and hence that

$$\alpha = D_{\mathcal{E}}\alpha = i(\mathcal{E})d\alpha = -i(\mathcal{E})\omega$$

where $\omega = \omega_Q = -d\alpha$.

Now let H be a function on T^*Q which is homogeneous of degree k in the fiber variables. Then

$$\begin{aligned} kH = \mathcal{E}H &= i(\mathcal{E})dH \\ &= i(\mathcal{E})i(v_H)\omega \\ &= -i(v_H)i(\mathcal{E})\omega \\ &= i(v_H)\alpha \quad \text{so} \\ (\exp v_H)^*\alpha - \alpha &= \int_0^1 \frac{d}{dt}(\exp tv_H)^*\alpha dt \ \text{ and} \end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}(\exp tv_H)^*\alpha &= (\exp tv_H)^*\left(i(v_H)d\alpha + di(v_H)\alpha\right) \\
&= (\exp tv_H)^*\left(-i(v_H)\omega + di(v_H)\alpha\right) \\
&= (\exp tv_H)^*\left(-dH + kdH\right) \\
&= (k-1)(\exp tv_H)^*dH \\
&= (k-1)d(\exp tv_H)^*H \\
&= (k-1)dH
\end{aligned}$$

since H is constant along the trajectories of v_H. So

$$(\exp v_H)^*\alpha - \alpha = (k-1)dH. \tag{9}$$

Remark. In the above calculation we assumed that H was smooth on all of T^*Q including the zero section, effectively implying that H is a polynomial in the fiber variables. But the same argument will go through (if $k > 0$) if all we assume is that H (and hence v_H) are defined on $T^*Q\setminus$ the zero section, in which case H can be a more general homogeneous function on $T^*Q\setminus$ the zero section.

Now $\exp v_H \ : T^*Q \to T^*Q$ is symplectic map. Let

$$\Gamma := \text{graph } (\exp v_H),$$

so $\Gamma \subset T^*Q^- \times T^*Q$ is a Lagrangian submanifold. Suppose that the projection $\pi_{Q\times Q}$ of Γ onto $Q \times Q$ is a diffeomorphism, i.e. suppose that Γ is horizontal. This says precisely that for every $(x, y) \in Q \times Q$ there is a unique $\xi \in T^*_Q Q$ such that

$$\pi \exp v_H(x, \xi) = y.$$

In the case of the geodesic flow, this is guaranteed by the condition of geodesic convexity.

Since Γ is horizontal, it has a generating function ϕ such that

$$d\phi = \mathrm{pr}_2^* \alpha - \mathrm{pr}_1^* \alpha$$

where $\mathrm{pr}_i, \quad i = 1, 2$ are the projections of $T^*(Q \times Q) = T^*Q \times T^*Q$ onto the first and second factors. On the other hand pr_1 is a diffeomorphism of Γ onto T^*Q. So

$$\mathrm{pr}_1 \circ (\pi_{Q \times Q|\Lambda})^{-1}$$

is a diffeomorphism of $Q \times Q$ with T^*Q.

Theorem 2 *Assume the above hypotheses. Then up to an additive constant we have*

$$\left(\mathrm{pr}_1 \circ (\pi_{Q \times Q|\Lambda})^{-1}\right)^* [(k-1)H] = \phi.$$

$$\phi_t(x,y) := \frac{1}{2t} d(x,y)^2, \tag{8}$$

$$(\exp v_H)^* \alpha - \alpha = (k-1) dH. \tag{9}$$

Theorem 2 *Assume the above hypotheses. Then up to an additive constant we have*

$$\left(\mathrm{pr}_1 \circ (\pi_{Q\times Q|\Lambda})^{-1}\right)^* [(k-1)H] = \phi.$$

Indeed, this follows immediately from(9). An immediate corollary is that (8) is the generating function for the time t flow on a geodesically convex Riemannian manifold.

As mentioned in the above remark, the same theorem will hold if H is only defined on $T^* \setminus \{0\}$ and the same hypotheses hold with $Q \times Q$ replaced by $Q \times Q \setminus \Delta$.