1 Topic 2: Canonical Transformations and the Hamilton-Jacobi Equation

Reading: Hand & Finch Chapter 6 (required), Goldstein 391-396 (supplemental, but I will cover this material in the notes).

The first part of this topic includes the material in Chapter 6 with some supplementary reading from Goldstein. At the end, we will cover in a bit more detail the relationship to wave mechanics.

1.1 Canonical Transformations

We have considered coordinate transformations between sets of space co-ordinates

$$Q_k = Q_k\left(\mathbf{q}(t), t\right) \tag{1}$$

Where $\mathbf{q}$ has N components in general. These are called *point transformations.* Obvious examples of this are going from Cartesian to cylindrical or spherical co-ordinates.

One advantage of the Lagrange formulation of mechanics is that it easily lets us choose any invertable function of the Cartesian coordinates as generalized coordinates and then write down Lagrange's equation *directly*:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \quad i = 1, ...N \tag{2}$$

so we get the equations of motion with no contortions.

Is there a similar situation for the coordinates of phase space and Hamilton's formulation? *Not in general* – you can easily invent transformations from (q_i, p_i) to a new set $(Q_i(q_i, p_i), P_i(q_i, p_i))$ (such a transformation containing both q and p is called a *contact* transformation) so that the equations of motion for (Q_i, P_i) *do not* follow from a Hamiltonian.

A simple example:

$$H = \frac{p^2}{2m}, \;\; \dot{p} = 0, \;\; \dot{q} = \frac{p}{m} \tag{3}$$

Transform

$$P = pt \;\; \text{so} \;\; \dot{P} = \dot{p}t + p = p = \frac{P}{t} \tag{4}$$

$$Q = qt \;\; \text{so} \; \dot{Q} = \dot{q}t + q = \frac{pt}{m} + q = \frac{P}{m} + \frac{Q}{t} \tag{5}$$

Does a function $K\left(P, Q, t\right)$ exist such that

$$\dot{P} = -\frac{\partial K}{\partial Q} \tag{6}$$

$$\dot{Q} = \frac{\partial K}{\partial P} \tag{7}$$

? If so, then

$$\frac{\partial \dot{P}}{\partial P} = -\frac{\partial^2 K}{\partial P \partial Q} = -\frac{\partial \dot{Q}}{\partial Q} \tag{8}$$

But

$$\frac{\partial \dot{P}}{\partial P} = \frac{1}{t} \tag{9}$$

$$\frac{\partial \dot{Q}}{\partial Q} = \frac{1}{t} \tag{10}$$

So this transformation does not lead to Hamilton's equation, and area in (P, Q) space is not preserved under motion. A transformation like this is said to be *non-canonical*. This has nothing to do with finding the motion in the new variable. In particular if

$$\frac{dP}{dt} = \frac{P}{t} \rightarrow \frac{dP}{P} = \frac{dt}{t} \rightarrow P = kt \tag{11}$$

$$\frac{dQ}{dt} = \frac{P}{m} + \frac{Q}{t} = \frac{kt}{m} + \frac{Q}{t} \tag{12}$$

or

$$\frac{1}{t}\frac{dQ}{dt} - \frac{Q}{t^2} = \frac{k}{m} \tag{13}$$

$$\frac{d}{dt}\left(\frac{Q}{t}\right) = \frac{k}{m} \tag{14}$$

so

$$\begin{aligned} \frac{Q}{t} &= \frac{k}{m}t + a \\ Q &= \frac{k}{m}t^2 + at \rightarrow q = \frac{k}{m}t + a \end{aligned}$$

However any result that follows from Hamilton's equation does not apply in the non-canonical space of (P, Q).

We can, however, find a class of transformations that are said to be canonical, in that they do follow from a new Hamiltonian. These transformations (involving q_i, p_i) are called contact (as opposed to point) transformations. These phase space of the new variables has the same properties as the old (e.g. Liouville's theorem holds). Two examples for the free particle:

$$H = \frac{p^2}{2m} \tag{15}$$

and apply the transformation

$$\begin{aligned} P &= ap + bq \\ Q &= cp + dq \end{aligned}$$

and the inverse

$$\begin{aligned} p &= \frac{1}{\Delta}(dP - bQ) \\ q &= \frac{1}{\Delta}(-cP + aQ) \end{aligned}$$

where we assume the determinant of the coefficients $\Delta = ad - bc \neq 0$.

So

$$\begin{aligned} \dot{P} &= b\frac{p}{m} = \frac{b}{m\Delta}(dP - bQ) \\ \dot{Q} &= d\frac{p}{m} = \frac{d}{m\Delta}(dP - bQ) \end{aligned}$$

so we want to know if there is a K such that Hamilton's equations hold. We saw above that this requires $\frac{\partial \dot{P}}{\partial P} = -\frac{\partial \dot{Q}}{\partial Q}$. In fact

$$\begin{aligned} \frac{\partial \dot{P}}{\partial P} &= \frac{bd}{m\Delta} \\ \frac{\partial \dot{Q}}{\partial Q} &= \frac{-db}{m\Delta} \end{aligned}$$

So there is a new Hamiltonian, K:

$$K(P, Q) = \Delta H(p(P, Q), q(P, Q)) = \frac{1}{2m\Delta}(dP - bQ)^2 \tag{16}$$

so

$$\begin{aligned} \dot{P} &= \frac{-\partial K}{\partial Q} = \frac{b}{m\Delta}(dP - bQ) \\ \dot{Q} &= \frac{\partial K}{\partial P} = \frac{d}{m\Delta}(dP - bQ) \end{aligned}$$

as above. Except for the factor Δ, this looks just like what you might expect from Lagrangian mechanics – just write the Hamiltonian in the new coordinates. If $\Delta = 1$, that's just what happens, however for a general case "just evaluate the old Hamiltonian in the new variable" doesn't work.

To show that the contact transformations are really very different from the point tranformations of Lagrangian mechanics consider this example – again the free particle:

$$H = \frac{p^2}{2m} \tag{17}$$

and the transform

$$\begin{aligned} P &= p\cos q \\ Q &= p\sin q \end{aligned}$$

so $p^2 = P^2 + Q^2$ and

$$\begin{aligned}
\dot{P} &= \dot{p}\cos q - p\dot{q}\sin q = -\frac{p^2}{m}\sin q = -\sqrt{P^2+Q^2}\frac{Q}{m} \\
\dot{Q} &= \dot{p}\sin q + p\dot{q}\cos q = \frac{p^2}{m}\cos q = \frac{p^2}{m}\cos q = \sqrt{P^2+Q^2}\frac{P}{m}
\end{aligned}$$

so

$$\begin{aligned}
\frac{\partial\dot{P}}{\partial P} &= -\frac{P}{\sqrt{P^2+Q^2}}\frac{Q}{m} \\
\frac{\partial\dot{Q}}{\partial Q} &= \frac{Q}{\sqrt{P^2+Q^2}}\frac{P}{m}
\end{aligned}$$

so like in the previous example K does exist and

$$K = \frac{1}{3m}\left(P^2+Q^2\right)^{\frac{3}{2}} \tag{18}$$

Note this is not even close to $H\left(p\left(P,Q\right),q\left(P,Q\right)\right)$! This transformation is "cononical" for the free particle, but it is not canonical for other Hamiltonians. For example

$$H = \frac{p^2}{2m} + mgq \tag{19}$$

$$\dot{p} = -mg, \quad \dot{q} = \frac{p}{m} \tag{20}$$

$$\begin{aligned}
\dot{P} &= -mg\frac{P}{\sqrt{P^2+Q^2}} - \frac{Q}{m}\sqrt{P^2+Q^2} \\
\dot{Q} &= -mg\frac{Q}{\sqrt{P^2+Q^2}} + \frac{P}{m}\sqrt{P^2+Q^2}
\end{aligned}$$

and

$$\frac{\partial\dot{P}}{\partial P} \neq \frac{-\partial\dot{Q}}{\partial Q} \tag{21}$$

So, this is a "cononical-like" transformation.

For a contact transformation to be a general canonical transformation, it must yield Hamilton's equations *for any Hamiltonian.*

Now we're clear on what canonical transformations are, lets see how to construct them. As your book does, I'm going to stick to 2-D phase space (1-D configuration space) for the following discussions. I'll also note that I will like your book take the generating function approach to constructing canonical transformations. There is another, seemingly unrelated approach that can be derived in terms of the matrix, or sympletic formalism of Hamilton's equations. We'll cover this later (see also appendix A of chapter 6 and Goldstein pp 391ff).

1.2 Generating Function Approach to Canonical Transformations

We saw last term (proven in a presentation problem) that two Lagrangians differing by a time derivative of the form $\frac{dF(q,t)}{dt}$ both are valid descriptions of the same physical system. Note that F does not depend on $\dot{q}$. Consider two different descriptions of the same system $L'\left(Q,\dot{Q},t\right)$ and $L(q,\dot{q},t)$. These refer to the same physical system if

$$L'\left(Q,\dot{Q},t\right) = L(q,\dot{q},t) - \frac{dF(q,Q,t)}{dt} \tag{22}$$

Note here F can be a function of q, Q but not their time derivatives.

We want the Euler-Lagrange equations to hold in terms of the new variables. Integrating both sides we have

$$\int_{t_1}^{t_2} L' dt = \int_{t_1}^{t_2} L dt + F(q(t_1), Q(t_1), t_1) - F(q(t_2), Q(t_2), t_2) \tag{23}$$

We can see that Hamilton's principle will hold in the new system if it holds in the old if we take the variation of the above equation and assume that arbitrary variations in δq imply arbitrary variations in δQ (assuming δF vanishes at the end points). So, F can be used to generate a new Lagrangian in terms of new variables for which Hamilton's principle holds.

Now we have to figure out how to construct the new canonical momentum and the new Hamiltonian to get the new configuration space (in which everything we derive from Hamilton's formalism applies). Note we have specified a generating function, F, not a set of transformation equations. We have to get the specific form of the transformation equations from F.

To to the above, take the time derivative of $F(Q,q,t)$:

$$\frac{dF}{dt} = \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial Q}\dot{Q} + \frac{\partial F}{\partial t} \tag{24}$$

Now $L' = L\left(Q,\dot{Q},t\right)$, so

$$\frac{\partial L'}{\partial \dot{q}} = \frac{\partial L}{\partial \dot{q}} - \frac{\partial}{\partial \dot{q}}\left(\frac{dF}{dt}\right) = \frac{\partial L}{\partial \dot{q}} - \frac{\partial F}{\partial q} = 0 \tag{25}$$

and

$$p = \frac{\partial F}{\partial q} \tag{26}$$

By definition the momentum canonical to Q is

$$P = \frac{\partial L'}{\partial \dot{Q}} = -\frac{\partial F}{\partial Q} \tag{27}$$

To get the tranformation explicity, we solve $p = \frac{\partial F}{\partial q}$ for $Q = Q(q,p,t)$, then we solve $P = -\frac{\partial F}{\partial Q}$ for $P = P(q,p,t)$ (substituting our $Q(q,p,t)$ in for Q).

To find the new Hamiltonian, $K(Q,P)$, we construct it from our Q,P and L' :

$$\begin{aligned}K(Q,P,t) &\equiv P\dot{Q}-L' \\ &= -\frac{\partial F}{\partial Q}\dot{Q}-L+\frac{\partial F}{\partial q}\dot{q}+\frac{\partial F}{\partial Q}\dot{Q}+\frac{\partial F}{\partial t} \\ &= P\dot{Q}-L+p\dot{q}-P\dot{Q}+\frac{\partial F}{\partial t} \\ &= p\dot{q}-L+\frac{\partial F}{\partial t}\end{aligned}$$

and finally

$$K(Q,P,t)=H(q(Q,P),p(Q,P),t)+\frac{\partial F(q(Q,P),Q,t)}{\partial t} \tag{28}$$

Now can we use any $F(Q,q,t)$ we want? Without proof, it is also necessary and sufficient that $\frac{\partial^2 F}{\partial q\partial Q}\neq 0$. If this second derivative vanishes, the transformation will not be invertable. Suppose we are given a set of transformation equations, how do we know if they are cononical? First, express p,P as functions of q,Q,t. Then solve for F using $P=-\frac{\partial F}{\partial Q}, p=\frac{\partial F}{\partial q}$, and see if our conditions apply ($F=F(q,Q,t)$ and $\frac{\partial^2 F}{\partial q\partial Q}\neq 0$). This may or may not be possible to solve - and we saw that not all contact transformations are canonical.

1.2.1 Types of Generating Function

We have considered generating functions of the form $F(q,Q,t)$, where we assume $\delta F=0$ at $t=t_{1,}t_2$, and F satisfies the condition on the double partial derivative above. Now we can perform a Legendre transformation on either q or Q to replace them with either p or P, so that we can express *the same* canonical transformation by any of four generating functions, $F=F_1(q,Q,t), F_2(q,P,t), F_3(p,Q,t)$ or $F_4(p,P,t)$. Why this is useful will be clear later.

We therefore get any generating function from any other one by a series of transformations. For example, to get $F_3(p,Q,t)$, transform F_1

$$F_3(p,Q,t)=F_1(q,Q,t)-qp \tag{29}$$

and from

$$\frac{\partial F_3}{\partial p}=\frac{\partial F_1}{\partial p}-q=0-q \tag{30}$$

so

$$q=-\frac{\partial F_3}{\partial p} \tag{31}$$

and

$$P=-\frac{\partial F_1}{\partial Q}=-\frac{\partial F_3}{\partial Q} \tag{32}$$

Note that we have assumed that q,p,Q,P are all independent variables. This is fine for the purposes of the derivation, as we know they form an independent set dynamically. To get the transformation equations which give us the *functional* dependence, we use $q=-\frac{\partial F_3}{\partial p}$, and $P=-\frac{\partial F_3}{\partial Q}$.

We can go through the same exercise to get F_2 and F_4, but we have to change the sign on the transformation (because of the asymmetry of the minus sign in Hamilton's equations).

$$\begin{aligned} F_4(p,P,t) &= F_3(p,Q,t) + PQ \\ F_2(q,P,t) &= F_1(q,Q,t) + QP \end{aligned}$$

and (without going through the straightforward steps) we get the following transformation equations:

$F_1(q,Q,t)$ $\quad p = \frac{\partial F_1}{\partial q}$; $P = -\frac{\partial F_1}{dQ}$
$F_2(q,P,t)$ $\quad p = \frac{\partial F_2}{\partial q}$; $Q = \frac{\partial F_2}{\partial P}$
$F_3(p,Q,t)$ $\quad q = -\frac{\partial F_3}{\partial p}$; $P = -\frac{\partial F_3}{\partial Q}$
$F_4(p,P,t)$ $\quad q = -\frac{\partial F_4}{\partial p}$; $Q = \frac{\partial F_4}{\partial P}$

1.3 Poisson Brackets

The Poisson bracket of two (arbitrary) functions, F, G with respect to canonically conjugate pair is defined by

$$[F,G]_{q,p} \equiv \sum_{k=1}^{N} \left(\frac{\partial F}{\partial q_k}\frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q_k} \right) \tag{33}$$

The value of the Poisson bracket is independent of which set of conjugate variables we use to evaluate the partials, so long as they are related by a canonical transformation

$$[F,G]_{q,p} = [F',G']_{Q,P} \tag{34}$$

where F', G' are the transformed functions.

If we let $F = Q, G = P$, then

$$[Q,P]_{Q,P} = [Q(q,p), P(q,p)]_{q,p} = 1 \tag{35}$$

The significance of this is that without knowing the form of the generating function for the canonical tranformation $q,p \to Q,P$ we can test whether a given relationship is canonical or not. If this holds, then the transformation must be canonical (it is a sufficient and necessary condition). You can demonstrate this directly using the formulae of the canonical transformation.

If we consider an arbitrary number of dimensions, $[F,G]_{q,p} = [F',G']_{Q,P}$, and $\frac{\partial q_k}{\partial ql} = \delta_{lk}$, $\frac{\partial q_k}{\partial p_l} = \delta_{lk}$,... then we have

$$[Q_i, Q_k]_{p,q} = 0,\ [P_i, P_k]_{p,q} = 0,\ [P_i, Q_k]_{p,q} = \delta_{ik} \tag{36}$$

1.4 The Sympletic Property of General Canonical Transformations

(See Appendix A of your book or Goldstein ch. 9)

We can work out a sufficient condition for a general canonical transformation easily using a separate appoach. First introduce a systematic notation. Let

$$\boldsymbol{\eta} = \begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_n \\ q_1 \\ q_2 \\ \cdots \\ q_n \end{pmatrix} = \begin{pmatrix} \vec{p} \\ \vec{q} \end{pmatrix} \tag{37}$$

where $\boldsymbol{\eta}$ is a 2N dimensional vector, $\vec{p}, \vec{q}$ are N-dimensional vectors, and we define

$$\mathbf{J} = \begin{pmatrix} \mathbf{0} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \tag{38}$$

where $\mathbf{0}, \mathbf{1}$ are $N \times N$ matrices and $\mathbf{J}$ is a $2N \times 2N$ matrix. Hamilton's equations using this notation are

$$\dot{\eta}_i = J_{ij} \frac{\partial H}{\partial \eta_j} \tag{39}$$

Now a few properties of the matrix $\mathbf{J}$: $\tilde{\mathbf{J}} = -\mathbf{J}$, $\mathbf{J}^{-1} = -\mathbf{J}$, $\rightarrow \mathbf{J}^2 = -\mathbf{1}$ so

$$(\det \mathbf{J})^2 = (\det(-\mathbf{1}))^2 = (-1)^{2N} = +1 \tag{40}$$

so

$$\det \mathbf{J} = \pm 1 \tag{41}$$

Its value is actually $+1$ as can be seen by row interchange (I won't go through it but you can write it out for yourself) - anyway you can see for the simple example

$$\begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} = 1 \tag{42}$$

So in summary, $\mathbf{J}$ is antisymmetric, has $-\mathbf{1}$ as its square, and has unit determinant. Now let $\xi_i = \xi_i(\eta_j, t)$ be invertible so that $\eta_i = \eta_i(\xi_j, t)$ (this implies det $\left(\frac{\partial \xi_i}{\partial \eta_j}\right)$ is not identically zero). Then

$$H_\eta(\eta_j, t) = H_\eta(\eta_j(\xi_i, t), t) = H_\xi(\xi_i, t) = H_\xi(\xi_i(\eta_j, t), t) \tag{43}$$

$$\frac{\partial H_\eta}{\partial \eta_k} = \frac{\partial H_\xi}{\partial \xi_i} \frac{\partial \xi_i}{\partial \eta_k} \tag{44}$$

and

$$\dot{\xi}_i = \frac{\partial \xi_i}{\partial \eta_j} \dot{\eta}_j + \frac{\partial \xi_i}{\partial t} = \frac{\partial \xi_i}{\partial \eta_j} J_{jk} \frac{\partial H_\eta}{\partial \eta_k} + \frac{\partial \xi_i}{\partial t} \tag{45}$$

so

$$\begin{aligned}\dot{\xi}_i &= \frac{\partial \xi_i}{\partial \eta_j} J_{jk} \frac{\partial H_\eta}{\partial \eta_k} + \frac{\partial \xi_i}{\partial t}|_\eta \\ &= \frac{\partial \xi_i}{\partial \eta_j} J_{jk} \frac{\partial H_\xi}{\partial \xi_l} \frac{\partial \xi_l}{\partial \eta_k} + \frac{\partial \xi_i}{\partial t}|_\eta\end{aligned}$$

If it is the case that

$$(1) \qquad \frac{\partial \xi_i}{\partial \eta_j} J_{jk} \frac{\partial \xi_l}{\partial \eta_k} = k J_{il} \quad \text{for constant } k \neq 0 \tag{46}$$

and that there is a function $F(\xi_i, t)$ such that

$$\frac{\partial \xi_i}{\partial t}|_\eta = J_{il} \frac{\partial F}{\partial \xi_l} \tag{47}$$

then

$$\frac{\partial F}{\partial \xi_l}|_t = -J_{li} \frac{\partial \xi_i}{\partial t}|_\eta \tag{48}$$

Then we have

$$\dot{\xi}_i = J_{il} \frac{\partial}{\partial \xi_l} (k H_\xi + F) \tag{49}$$

and the new Hamiltonian is

$$K(\xi_j, t) = k H_\xi (\xi_j, t) + F(\xi_j, t) \tag{50}$$

if the constant k is 1, then we have an ordinary canonical tranformation, and otherwise what some people call a "general canonical transformation".

In matrix notation if we let

$$(\mathbf{X})_{ij} = \frac{\partial \xi_i}{\partial \eta_j} \tag{51}$$

then we have

$$(\mathbf{X})_{ij} (\mathbf{J})_{jk} (\mathbf{X})_{lk} = k (\mathbf{J})_{il} \tag{52}$$

or

$$\begin{aligned}\mathbf{X}\mathbf{J}\tilde{\mathbf{X}} &= k\mathbf{J} \quad \text{general canonical} \\ \mathbf{X}\mathbf{J}\tilde{\mathbf{X}} &= \mathbf{J} \quad \text{canonical } (k=1)\end{aligned}$$

example:

$$\begin{aligned}P &= p + mv \quad (P = \xi, p = \eta) \\ Q &= q + vt\end{aligned}$$

(a Galilean transformation). Then

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{53}$$

so this is an ordinary cononical transfomation (but time dependent) with $k = 1$ if we can find an F so that

$$\begin{pmatrix} \frac{\partial F}{\partial P} \\ \frac{\partial F}{\partial Q} \end{pmatrix} = -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ v \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix} \tag{54}$$

so $F = Pv$ is sufficient. We then construct

$$K\left(P, Q, t\right) = H_{p,q}((P - mv), (Q - vt), t) + Pv + const \tag{55}$$

For a free particle

$$K = \frac{(p - mv)^2}{2m} + Pv + const = \frac{P^2}{2m} \tag{56}$$

(a symmetry). For the SHO

$$K = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 \left(Q - vt\right)^2 \tag{57}$$

and the solution is

$$Q = a\cos\left(\omega t + \phi\right) + vt \tag{58}$$

We supplement the above with two facts that I will not prove:

(1) The sufficient conditions given above can be shown to be necessary also

(2) Any general canonical tranformation can always be compounded of a linear transformation

$$\xi_i = \omega_{ij}\eta_j \tag{59}$$

where ω_{ij} is a suitably chosen symmetric matrix followed by an ordinary canonical transformation. Thus, all the interesting transformations are the ordinary ones. In summary;

$$\xi_i = \xi_i\left(\eta_j, t\right) \tag{60}$$

$$\mathbf{XJ}\tilde{\mathbf{X}} = \mathbf{J} \tag{61}$$

$$\left(\mathbf{X}\right)_{ij} = \frac{\partial \xi_i}{\partial \eta_j} \tag{62}$$

$$K = H + F \tag{63}$$

$$\frac{\partial F}{\partial \xi_i}|_{t,\text{all other } \xi_i\text{'s}} = -J_{ij}\frac{\partial \xi_j}{\partial t}|_\eta \qquad \text{(be careful to note what is held fixed)} \tag{64}$$

$$\dot{\xi}_i = J_{ij}\frac{\partial K}{\partial \xi_j} \tag{65}$$

Note, if the transformation doesn not depend on time, take $F = 0$ so

$$K\left(\xi_i, t\right) = H\left(\eta_j\left(\xi_i\right), t\right) \tag{66}$$

in other words, just the analog of what happens in Lagrangian theory.

1.5 The Hamilton-Jacobi Equation

The Hamilton-Jacobi Equation is an important example of how new information about mechanics can come out of the action by considering various kinds of variation of the trajectories.

From the definition of the action, it is easy to see that it can be considered to be an ordinary function of the variables $q_i^{(2)}, t^{(2)}, q_i^{(1)}, t^{(1)}$ by using actual motions connecting these points in time-configuration space:

$$S\left(q_i^{(2)}, t^{(2)}, q_i^{(1)}, t^{(1)}\right) = \int_{t^{(1)}}^{t^{(2)}} L\left(q_i, \dot{q}_i, t\right) dt \tag{67}$$

where $q_i\left(t\right), \dot{q}_i\left(t\right)$ are the actual motions satisfying Lagrange's equations with the boundary condition $q_i^{(2)} = q_i\left(t^{(2)}\right), q_i^{(1)} = q_i\left(t^{(1)}\right)$.

As an example:

$$\begin{aligned} L &= \frac{1}{2}m\dot{x}^2 \\ H &= \frac{p^2}{2m}, \quad p = m\dot{x} \end{aligned}$$

then

$$x\left(t\right) = x^{(1)} + \frac{x^{(2)} - x^{(1)}}{t^{(2)} - t^{(1)}}\left(t - t^{(1)}\right) \tag{68}$$

So

$$\begin{aligned} S &= \int_{t_1}^{t_2} \frac{1}{2}m\left(\frac{x^{(2)} - x^{(1)}}{t^{(2)} - t^{(1)}}\right)^2 dt \\ &= \frac{1}{2}m\frac{(x^{(2)} - x^{(1)})^2}{t^{(2)} - t^{(1)}} \end{aligned}$$

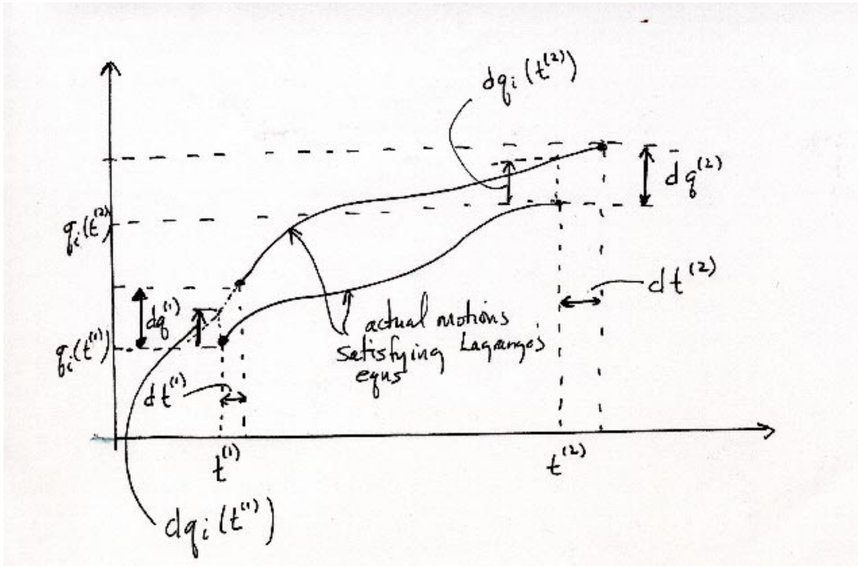
Now it is interesting to notice that

$$\begin{aligned} \frac{\partial S}{\partial x^{(2)}} &= m\frac{x^{(2)} - x^{(1)}}{t^{(2)} - t^{(1)}} = p^{(2)} \\ \frac{\partial S}{\partial x^{(1)}} &= -m\frac{x^{(2)} - x^{(1)}}{t^{(2)} - t^{(1)}} = -p^{(1)} \end{aligned}$$

and also

$$\begin{aligned} \frac{\partial S}{\partial t^{(2)}} &= \frac{-1}{2}m\left(\frac{x^{(2)} - x^{(1)}}{t^{(2)} - t^{(1)}}\right)^2 = -H^{(2)} \\ \frac{\partial S}{\partial t^{(1)}} &= \frac{1}{2}m\left(\frac{x^{(2)} - x^{(1)}}{t^{(2)} - t^{(1)}}\right)^2 = H^{(1)} \end{aligned}$$

(note that in the above the H's could have been L's in this simple, ambiguous example).



Let's work out the general case of a change in S when we wiggle each of the $2n+2$ variables $q_i^{(1)}, q_i^{(2)}, t^{(1)}, t^{(2)}$. Consider the following picture

Then

$$
\begin{aligned}
dS &= \int_{t^{(1)}+dt^{(1)}}^{t^{(2)}+dt^{(2)}} L\left(q_i\left(t\right)+dq_i\left(t\right), \dot{q}_i+d\dot{q}_i, t\right) dt - \int_{t^{(1)}}^{t^{(2)}} L dt \\
&= L^{(2)} dt^{(2)} - L^{(1)} dt^{(1)} + p_i^{(2)} dq_i\left(t^{(2)}\right) - p_i^{(1)} dq_i\left(t^{(1)}\right)
\end{aligned}
$$

The term $p_i^{(2)} dq_i\left(t^{(2)}\right) - p_i^{(1)} dq_i\left(t^{(1)}\right)$ comes from the case when the ends $t^{(1,2)}$ are fixed and so the dq's are the changes at these times.

But $dq_i\left(t^{(1)}\right) \neq dq_i^{(1)}$ (see the picture). In fact,

$$
dq_i^{(1)} = dq_i\left(t^{(1)}\right) + \dot{q}_i\left(t^{(1)}\right) dt^{(1)} \tag{69}
$$

and similary for (2). So

$$
\begin{aligned}
dS &= L^{(2)} dt^{(2)} + p_i^{(2)} \left(dq_i^{(2)} - \dot{q}_i\left(t^{(2)}\right) dt^{(2)}\right) - L^{(1)} dt^{(1)} - p_i^{(1)} \left(dq_i^{(1)} - \dot{q}_i^{(1)} dt^{(1)}\right) \\
&= \left(p_i dq_i - H dt\right)\Big|_{t^{(1)}}^{t^{(2)}}
\end{aligned}
$$

So we have, with a slight change of notation in which you can think of (q_i, t) as the free variable and $\left(q_i^{(0)}, t^{(0)}\right)$ as initial constants

$$
dS = p_i dq_i - H dt - \left(p_i^{(0)} dq_i^{(0)} - H^{(0)} dt^{(0)}\right) \tag{70}
$$

We can get the result

$$
\frac{\partial S}{\partial q_i} = p_i \, ; \; \frac{\partial S}{dt} = -H \tag{71}
$$

and

$$\frac{\partial S}{\partial q_i^{(0)}} = p_i^{(0)}; \ \frac{\partial S}{dt^{(0)}} = H^{(0)} \tag{72}$$

very easily. Think of S as an indefinite integral (i.e. t as a variable)

$$S = \int^t L dt = \int^t (-H + p_i \dot{q}_i)\, dt \tag{73}$$

so

$$\frac{dS}{dt} = (-H + p_i \dot{q}_i) \tag{74}$$

But $S = S(q_i, t)$, so

$$\frac{dS}{dt} = \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t} = -H + p_i \dot{q}_i \tag{75}$$

so

$$\frac{\partial S}{\partial t} = -H; \ \frac{\partial S}{\partial q_i} = p_i \tag{76}$$

and we also have $\frac{\partial S}{\partial q_i^{(0)}} = p_i^{(0)}$; $\frac{\partial S}{dt^{(0)}} = H^{(0)}$.

Hamilton was intrigued by these equations, and he noticed the following. Remember that

$$H = H(p_i, q_i, t) \tag{77}$$

and so

$$\frac{\partial S}{\partial t} + H\left(\frac{\partial S}{\partial q_i}, q_i, t\right) = 0 \tag{78}$$

which is a non-linear partial differential equation for the function $S(q_i, t)$. (since its non-linear, the sums of solutions are not necessarily solutions). Lets see how this works for the simplest possible case:

$$H = \frac{p^2}{2m} \tag{79}$$

so

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 = 0 \tag{80}$$

S for the free particle is $S = \frac{1}{2}m\frac{(x-x_o)^2}{t-t_o}$

$$\begin{aligned} \frac{\partial S}{\partial t} &= -\frac{1}{2}m\left(\frac{x-x_o}{t-t_o}\right)^2 \\ \frac{\partial S}{\partial x} &= m\frac{x-x_o}{t-t_o} \end{aligned}$$

so combining

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 = \frac{1}{2}m\left(\frac{x-x_0}{t-t_o}\right)^2 \tag{81}$$

and Hamilton's partial differential equation works.

But what about the converse. Can you use Hamilton's PDE to calculate an S and use it to solve mechanics problems? The answer is yes *if* you happen to find the "right" solution. But, PDE's have an infinity of solutions . For example, the above solution to the PDE can be found by a familiar separation of variables procedure. Assume

$$S = X(x)\,T(t) \tag{82}$$

so Hamilton's PDE becomes

$$XT' + \frac{1}{2m}(X'T)^2 = 0 \tag{83}$$

where prime means differentiation wrt the function's argument. Then divide by XT^2 (assumed non-zero) to get

$$\frac{T'}{T^2} + \frac{1}{2m}\frac{X'^2}{X} = 0 \tag{84}$$

Since x, t are independent and we have $f(t) + \frac{g(x)}{2m} = 0$ we must have

$$\begin{aligned} f(t) &= -k\frac{1}{2m} = const \\ g(x) &= +k \end{aligned}$$

so

$$2mT' = -kT^2,\ \ \frac{dT}{T^2} = -kdt\frac{1}{2m} \tag{85}$$

so

$$\begin{aligned} \frac{1}{T} &= \frac{kt}{2m} + a \\ T &= \frac{1}{\frac{k}{2m}(t - t_o)} \end{aligned}$$

$$\begin{aligned} X' &= \sqrt{kX} \\ \frac{dX}{\sqrt{X}} &= \sqrt{k}dx \\ 2\sqrt{X} &= \sqrt{k}x + b \end{aligned}$$

or

$$X = \frac{1}{4}k(x - x_o)^2 \tag{86}$$

and

$$S = XT = \frac{\frac{1}{4}k(x - x_o)^2}{\frac{k}{2m}(t - t_o)} = \frac{m}{2}\frac{(x - x_o)^2}{(t - t_0)} \tag{87}$$

which is just the same as we got by direct integration.

But what if we had assumed

$$S = T + X \tag{88}$$

so

$$\begin{aligned} T' + \frac{1}{2m}\left(X'\right)^2 &= 0 \\ \frac{1}{2m}\left(X'\right)^2 &= k \end{aligned}$$

$$\rightarrow X = \sqrt{2mk}x + a \tag{89}$$

so

$$T' = -k,\, T = -kt + b \tag{90}$$

so

$$S = \sqrt{2mk}x - kt + const \tag{91}$$

which is a whole lot different than the previous solution! Note that this function is in fact an indefinite integral of the Lagrangian of a free particle with motion

$$x = vt + x_o \tag{92}$$

so

$$\dot{x} = v \tag{93}$$

$$S = \int \frac{1}{2} m\dot{x}^2 dt = \frac{1}{2} mv^2 t + const \tag{94}$$

which is equal to $\sqrt{2mk}x - kt = \sqrt{2mk}vt - kt + \sqrt{2mk}x_o = \frac{1}{2}mv^2 t + const$ if $\frac{1}{2}mv^2 = \sqrt{2mk}v - k$ and $const = \sqrt{2mk}x_o$

This simple example raises the question – How do we figure out which solutions of Hamilton's PDE are associated with actual motion and how do you extract that motion from the function S?

Hamilton did not manage to solve the problem, but Jacobi did, and so the PDE is now known as the *Hamilton-Jacobi equation.* We can understand what Jacobi produced if we remember that if we write

$$S\left(q_i, t; q_i^{(0)}, t^{(0)}\right) \tag{95}$$

(think of $q_i^{(0)}, t^{(0)}$ as constants) we also get

$$\begin{aligned} \frac{\partial S}{\partial t^{(0)}} &= H\left(p_i^{(0)}, q_i^{(0)}, t^{(0)}\right) \\ \frac{\partial S}{\partial q_i^{(0)}} &= -p_i^{(0)} \end{aligned}$$

or when we differentiate S with respect to constants $t^0, q_i^{(0)}$ we get other constants, $p_i^{(0)}, H^0$. So Jacobi's rule for a system with N degrees of freedom:

(1) Find any solution S of the HJ equation which depends on $N+1$ arbitrary algebraically independent (see definition below) constants,

$$S = S\left(q_i, t; a_i\right) + A \quad i = 1, ...N \tag{96}$$

(one of the constants is always additive and ignorable).

(2) Obtain the EOM by setting

$$\frac{\partial S}{\partial a_i} = b_i \tag{97}$$

where the b_i are N more constants, and solve for the

$$q_i = q_i\left(t; a_j, b_k\right) \quad i, j, k = 1....N \tag{98}$$

By alegebraic independence we mean that the $N \times N$ matrix

$$\frac{\partial^2}{\partial q_i \partial a_j} \tag{99}$$

has a determinant that is not identically zero so that the matrix "usually" has an inverse (there may be singular points where there is no inverse – i.e. the determinant is zero). So if you get a solution with a_k, $k = 1, ...N-1$ you cannot fill out the set of constants by for example splitting one of these into the sum of two new ones, so

$$a^{new} = \left(a_1, a_2, ..., a'_{n-1}, a''_{n-1}\right) \tag{100}$$

where $a'_{n-1} + a''_{n-1} = a^{old}_{n-1}$. Clearly in this case two rows (if j determines rows) of $\frac{\partial S}{\partial q_i \partial a_j^{new}}$ will be identical and the determinant is identically zero.

Check the theorem in two simple cases: $L = \frac{1}{2}m\dot{q}^2$ so

$$S = \frac{1}{2}m\frac{\left(x - x_o\right)^2}{t - t_o} \tag{101}$$

take x_o as a_1: then

$$\frac{\partial S}{\partial a_1} = \frac{\partial S}{\partial x_o} = -m\frac{x - x_o}{t - t_o} = b_1 \tag{102}$$

Which gives a solution.

If we take t_o as a_1: then

$$\frac{\partial S}{\partial a_1} = \frac{\partial S}{\partial t_o} = \frac{1}{2}m\left(\frac{x - x_0}{t - t_o}\right)^2 = b_1 \tag{103}$$

this also gives a solution

$$S = \sqrt{2mk}x - kt + const \tag{104}$$

take k as a_1, so

$$\frac{\partial S}{\partial a_1} = \frac{\partial S}{\partial k} = \frac{1}{2}\sqrt{\frac{2m}{k}}x - t = b_1 \tag{105}$$

or

$$x = 2\sqrt{\frac{k}{2m}}t + const \tag{106}$$

– a solution.

We can easily prove Jacobi's theorem:

$$\frac{d}{dt}\frac{\partial S}{\partial a_i} = 0 = \frac{\partial^2 S}{\partial a_i \partial q_j}\dot{q}_j + \frac{\partial^2 S}{\partial a_i \partial t} \tag{107}$$

But

$$\frac{\partial S}{\partial t} = -H\left(\frac{\partial S}{\partial q_i}, q_i, t\right) \tag{108}$$

only $p_i = \frac{\partial S}{\partial q_i}$ depends on the a's, so

$$\frac{\partial^2 S}{\partial a_i \partial t} = -\frac{\partial H}{\partial p_j}\frac{\partial^2 S}{\partial q_j \partial a_i} \tag{109}$$

so we get

$$\frac{\partial^2 s}{\partial a_i \partial q_j}\left(\dot{q}_j - \frac{\partial H}{\partial p_j}\right) = 0 = \mathbf{S}\left(\dot{\mathbf{q}} - \frac{\partial H}{\partial \mathbf{p}}\right) \tag{110}$$

If the matrix $\frac{\partial^2 S}{\partial a_i \partial q_j}$ is invertible, then the only solution to these equations is $\dot{q}_j = \frac{\partial H}{\partial p_j}$, half of of Hamilton's equations.

Next from

$$p_i = \frac{\partial S}{\partial q_i} \tag{111}$$

get

$$\begin{aligned}
\dot{p}_i &= \frac{d}{dt}\left(\frac{\partial S}{\partial q_i}\right) \\
&= \frac{\partial^2 S}{\partial q_i \partial q_j}\dot{q}_j + \frac{\partial^2 S}{\partial q_i \partial t} \\
&= \frac{\partial}{\partial q_i}\left(\frac{\partial S}{\partial q_j}\dot{q}_j + \frac{\partial S}{\partial t}\right) \\
&= \frac{\partial}{\partial q_i}\left(p_j \dot{q}_j - H\right)
\end{aligned}$$

and

$$\dot{p}_i = \frac{\partial L}{\partial q_i} \tag{112}$$

which is just Lagrange's equation.

A very important special case is that in which the Hamiltonian is conserved (i.e. independent of t) and since it is often the energy, call the conserved value of H the constant E, and then S must be linear in t. Then

$$S = -Et + W\left(q_i\right) \tag{113}$$

satisfies the H-J equation if we have

$$-E + H\left(\frac{\partial W}{\partial q_i}, q_i\right) = 0 \tag{114}$$

here E is a constant. You can consider this another proof of $H = \mathit{const}$ if $\frac{\partial H}{\partial t} = 0$.

1.5.1 The Relationship of the H-J Equation to Quantum Mechanics

The H-J equation is an important step to formulating a wave equation for particles. Write a solution of the H-J equation for a free particle in 3-D as

$$S = \mathbf{p} \cdot \mathbf{x} - Et \tag{115}$$

where the three components of $\mathbf{p}$ are the three constants. By the H-J equation

$$\frac{\partial S}{\partial t} + H\left(\frac{\partial S}{\partial x_i}, x_i\right) = -E + \frac{\mathbf{p} \cdot \mathbf{p}}{2m} = 0 \tag{116}$$

we get the expected $E = \frac{p^2}{2m}$. We get the EOM by

$$\frac{\partial S}{\partial p_i} = x_i - \frac{\partial E}{\partial p_i} t = x_{io} \quad \text{(constant)} \tag{117}$$

or

$$x_i = \frac{p_i}{m} t + x_{i0} \tag{118}$$

However, if you look at S you see that at a fixed moment in time, a fixed value of S defines a plane and the vector $\mathbf{p}$ is normal to it. At a later time $t + \Delta t$ a plane of constant S will still have $\mathbf{p}$ as its normal and it will be displaced along $\mathbf{p}$ by and amount $\frac{E}{|\mathbf{p}|}\Delta t$. If this is not obvious, adopt the coordinate system so $\mathbf{p} = p\hat{\mathbf{e}}_x$ and we have

$$S = px - \frac{p^2}{2m} t = p\left(x + \Delta x\right) - \frac{p^2}{2m}\left(t + \Delta t\right) \tag{119}$$

so $\frac{\Delta x}{\Delta t}$ is the volocity at which the plane moves $= \frac{p}{2m}$. So we see that the particle trajectories are normal to surfaces of constant S (although the planes of constant S travel at half the speed of the particle). This may seem disappointing at first, until you notice that the group velocity of the wave is $\frac{\dot{p}}{m}$. The particle trajectories have the same geometric relation to the planes of constant S as the rays of optics do to planes of constant wave phase.

This suggests considering S as a wave phase factor

$$\psi = \psi_o e^{i\frac{S}{\hbar}} \tag{120}$$

where $\hbar$ is a scale factor with dimensions of $S =$ (energy $\times$ time) or (momentum $\times$ distance). No prejudice about its value follows from classical mechanics except that if in fact classical physics is the ray approximation to some underlying wave mechanics, then $\hbar$ must be very tiny, since it took almost 100 years from the time that Hamilton did the above for optics before any wave character was experimentally detected for a material particle.

From

$$\psi = \psi_o e^{i\frac{S}{\hbar}} \quad \text{and } S = \mathbf{p} \cdot \mathbf{x} - Et \tag{121}$$

what equation does ψ satisfy? We get

$$\boldsymbol{\nabla\psi} = \frac{i}{\hbar}\psi\boldsymbol{\nabla}\mathbf{S} \tag{122}$$

take the divergence of both sides

$$\nabla \cdot (\boldsymbol{\nabla\psi}) = \nabla^2\psi = \frac{i}{\hbar}\boldsymbol{\nabla\psi}\cdot\boldsymbol{\nabla}\mathbf{S} + \frac{i}{\hbar}\psi\nabla^2 S \tag{123}$$

for the free particle $\frac{i}{\hbar}\psi\nabla^2 S = 0$, and

$$\nabla^2\psi = \frac{i}{\hbar}\boldsymbol{\nabla\psi}\cdot\boldsymbol{\nabla}\mathbf{S} = \left(\frac{i}{\hbar}\right)^2 \psi\left(\boldsymbol{\nabla}\mathbf{S}\right)^2 \tag{124}$$

and

$$\frac{\partial\psi}{\partial t} = \frac{i}{\hbar}\psi\frac{\partial S}{\partial t} \tag{125}$$

using the H-J equation with $H = \frac{p^2}{2m}$

$$\frac{\partial\psi}{\partial t} = \frac{i}{\hbar}\psi\frac{1}{2m}\left(\boldsymbol{\nabla}\mathbf{S}\right)^2 = \frac{i}{\hbar}\frac{1}{2m}\left(-\hbar^2\nabla^2\psi\right) \tag{126}$$

or

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi \tag{127}$$

A familiar equation!

In fact the relationship between the wave equation and the H-J equation was well known to Hamilton, his contemporaries and his successors. In optics, the analog to the H-J equatino is called the *eikonal equation.* The S function is the icon or representative of the wave function. The function $e^{i\frac{S}{\hbar}}$ for suitable "$\hbar$" in the optics case gives an approximation to the wave equation in certain cases. It is the phase of a wave that carries most of the really wave-like character of a wave (think of interference) and so it is usually the most interesting physically.

It is interesting to inquire under what conditions the solution S to the $H - J$ equation makes a good approximation to the *phase* of the solution of the Schroedinger equation. Without proof, if

$$-\frac{\hbar}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial\psi}{\partial t} \tag{128}$$

and we take

$$\psi = A\left(\mathbf{r}\right)e^{i\frac{S(r,t)}{\hbar}} = A\left(\mathbf{r}\right)e^{\frac{i}{\hbar}(W(\mathbf{r})-Et)} \tag{129}$$

then

$$\left(\frac{1}{2m}\left(\boldsymbol{\nabla}\mathbf{S}\right)^2 + V - E\right) = \frac{\hbar}{2m}\left[\hbar\frac{\nabla^2 A}{A} + 2i\ \boldsymbol{\nabla}\mathbf{W}\cdot\frac{\boldsymbol{\nabla}\mathbf{A}}{A} + i\nabla^2 W\right] \tag{130}$$

The r.h.s. can be approximated by zero when $\hbar$ is very small compared to other quantities with the same dimensions in the physical problem. Thus for macroscopic bodies this is the case since

$$\hbar\tilde{}\ \text{(mass of an electron)}\cdot\frac{cm}{s}\cdot cm \tag{131}$$

It is useful to associate a wavelength with $W(x)$, which in 1-D

$$\begin{aligned}\frac{1}{\hbar}W(x) &= \frac{1}{\hbar}W(x_o) + \frac{1}{\hbar}(x - x_o)\frac{dW}{dx}|_{x_o} + ... \\ &= W(x_o) + \frac{x - x_o}{\lambda_o}\end{aligned}$$

where $\frac{1}{\lambda_o} = \frac{1}{\hbar}\frac{dW}{dx}$ so

$$e^{i\frac{1}{\hbar}(W(x)-Et)} \cong e^{i\left(\frac{x-x_o}{\lambda_o}-\omega t\right)} \tag{132}$$

where $\omega = \frac{E}{\hbar}$. Then the r.h.s. can be shown to be well approximated by zero if

$$\lambda_o\left|\frac{dV}{dx}\right| \ll \text{ kinetic energy} \tag{133}$$

i.e. the potential is slowly varying in space. In this case, the classical action S yields a reasonable approximation to the phase function of the quantum mechanical wave function. This corresponds to the WKB limit where the potential is essentially constant over many de Broglie wavelengths.

1.6 Adiabatic Invariants

This section follows the derivation of Landau and Lifshitz (see pp. 154ff).

Adiabatic invariants are quantities that remain essentially constant in a system that is not closed, but where some parameter varies slowly. We're going to consider systems that are strictly periodic and conservative when they are "closed" (i.e. when we keep all the system parameters constant), and consider what we can say about the system as we slowly vary one parameter. Specific examples of such systems would be a pendulum where we slowly change the length of the string (by slow we mean slow compared to the natural frequency), or a mass on a spring with a slowly changing spring constant. When you change the length of the string in a pendulum you know ω increases, and E changes, but can we construct some combination of parameters that stays essentially constant.

Call the parameter we vary λ, and let it vary slowly (adiabatically) with time as the result of some external action. If the period is T, we require

$$T\frac{d\lambda}{dt} \ll \lambda \tag{134}$$

The energy E changes slowly (if we average over the rapid oscillations of the system) with time as λ changes, and $\frac{dE}{dt}$ is then some function of λ. This dependence can be expressed as the constancy of some combination of E and λ (called an *adiabatic invariant*) which remains constant during the motion of a system with slowly varying parameters.

We can write the Hamiltonian as $H(q, p; \lambda)$, and the rate of change of the energy is

$$\frac{dE}{dt} = \frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda}\frac{\partial \lambda}{\partial t} \tag{135}$$

Now $\frac{\partial H}{\partial \lambda}\frac{\partial \lambda}{\partial t}$ depends on the rapidly varying q, p as well as on the slowly varying λ. We want to average over the rapid (periodic) variations resulting from the oscillatory motion to isolate the slow variations in λ

$$\overline{\frac{dE}{dt}} = \frac{\partial \lambda}{\partial t}\overline{\frac{\partial H}{\partial \lambda}} \tag{136}$$

where we average over the rapidly varying (oscillating) H, but during our averaging time, λ remains essentially constant, so we can pull the $\frac{\partial \lambda}{\partial t}$ out of the averaging. Furthermore, in averaging H we consider q, p to vary and λ to be constant. We are essentially averaging over the motion of the closed system (what would happen with λ constant).

The average is

$$\overline{\frac{\partial H}{\partial \lambda}} = \frac{1}{T}\int_0^T \frac{\partial H}{\partial \lambda} dt \tag{137}$$

From Hamilton's equations, $\dot{q} = \frac{\partial H}{\partial p}$, and we can change the integral over time to one over q: $dt = dq\left(\frac{\partial H}{\partial p}\right)^{-1}$

$$T = \int_0^T dt = \oint dq \left(\frac{\partial H}{\partial p}\right)^{-1} \tag{138}$$

where the integral over q is taken over the complete range of variation of the coordinate during one cycle.

So

$$\overline{\frac{dE}{dt}} = \frac{\partial \lambda}{\partial t}\overline{\frac{\partial H}{\partial \lambda}} = \frac{\partial \lambda}{\partial t}\frac{\oint dq\left(\frac{\partial H}{\partial \lambda}\right)\left(\frac{\partial H}{\partial p}\right)^{-1}}{\oint dq\left(\frac{\partial H}{\partial p}\right)^{-1}} \tag{139}$$

Since we are taking the averages for constant λ, $H = E$=const also over the integral, and p is a defined function of q, E, λ, $p = p\,(q; E, \lambda)$. So

$$H\,(q, p, \lambda) = E \tag{140}$$

and

$$\frac{dH}{d\lambda} = 0 = \frac{\partial H}{\partial \lambda} + \left(\frac{\partial H}{\partial p}\right)\left(\frac{\partial p}{\partial \lambda}\right) \tag{141}$$

$$\frac{\frac{\partial H}{\partial \lambda}}{\frac{\partial H}{\partial p}} = -\frac{\partial p}{\partial \lambda} \tag{142}$$

if we substitute this into the expression for $\overline{\frac{dE}{dt}}$

$$\overline{\frac{dE}{dt}} = -\frac{d\lambda}{dt}\frac{\oint\left(\frac{\partial p}{\partial \lambda}\right)dq}{\oint\left(\frac{\partial p}{dE}\right)dq} \tag{143}$$

so

$$\oint\left(\frac{\partial p}{\partial E}\overline{\frac{dE}{dt}} + \frac{\partial p}{\partial \lambda}\frac{d\lambda}{dt}\right)dq = 0 \tag{144}$$

or exchanging the order of integration and differentiation

$$\frac{d}{dt}\oint pdq = 0 \tag{145}$$

If we define

$$I \equiv \oint \frac{1}{2\pi} pdq \tag{146}$$

then

$$\overline{\frac{dI}{dt}} = 0 \tag{147}$$

Remember the integral is over a period with constant E, λ. So in this approximation I remains constant when λ varies.

I is the adiabatic invariant we have been looking for, and is a function of E, λ. If we look at the partial derivative with respect to E

$$2\pi\frac{\partial I}{\partial E} = \oint \frac{\partial p}{\partial E} dq = T = \frac{\omega}{2\pi} \tag{148}$$

we see that the partial is related to the period.

The geometrical significance of I is that it is related to the area of phase space enclosed by the curve

$$I = \frac{1}{2\pi}\oint pdq = \frac{1}{2\pi}\int\int dp\ dq \tag{149}$$

As a simple example, compute I for a simple harmonic oscillator with natural frequency ω:

$$H = \frac{1}{2}\frac{p^2}{m} + \frac{1}{2}m\omega^2 q^2 \tag{150}$$

Since $H = E = const$, the phase path is an ellipse with semi-axes $\sqrt{2mE}$ on the p axis and $\sqrt{2E/(m\omega^2)}$ on the q axis, and the area divided by 2π is

$$I = E/\omega \tag{151}$$

The significance of this is that when the parameters of the oscillator vary slowly, the energy is proportional to the frequency.

Suppose the SHO starts with maximum amplitude q_o and mass m_o at $t = 0$. Assume its mass is gradually ("adiabatically") increased and its spring constant k is held fixed. What is the amplitude q when the mass is m?

$$E_o = \frac{1}{2}kq_o^2$$
$$\omega_o = \sqrt{\frac{k}{m_o}}$$

so

$$I = \frac{1}{2}kq_o^2\sqrt{\frac{m_o}{k}} = \frac{1}{2}kq^2\sqrt{\frac{m}{k}} \tag{152}$$

or

$$\begin{aligned} q^2 &= q_o^2 \sqrt{\frac{m_o}{m}} \\ q &= q_o \left(\frac{m_o}{m}\right)^{1/4} \end{aligned}$$

In fact

$$q(t) \cong q_o \left(\frac{m_o}{m(t)}\right)^{1/4} \cos\left(\int_0^t \sqrt{\frac{k}{m(t)}} dt + \phi_o\right) \tag{153}$$

Historically the adiabatic invariance of $\oint pdq$ was taken to be of great significance during the early development of quantum mechanics (the "old" quantum mechanics). Remember Plank's famous hypothesis that for an SHO

$$E_n = n\hbar\omega \tag{154}$$

(where n is an integer and $\hbar$ a constant) are the only allowed energies. As we see above

$$2\pi \frac{E}{\omega} = \oint pdq = nh \text{ (by postulate)} \tag{155}$$

is an adiabatic invariant and so remains constant under any kind of variation slow compared to the period of the oscillation. Think of $T = \frac{2\pi}{\omega}$ for optical light frequencies

$$\nu = \frac{c}{\lambda} \tag{156}$$

$$T = \frac{\lambda}{c} = \frac{500 \cdot 10^{-9} m}{3 \cdot 10^8 \ m/s} = 200 \cdot 10^{-17} \ s \sim 10^{-15} s \tag{157}$$

so fast macroscopic changes might occur in 10^{-9} s and still be adiabatic.

Ehrenfest adopted the adhoc principle that quantum conditions should be applied to adiabatic invariants which Sommerfeld generalized to an "action variable", $\oint pdq$.

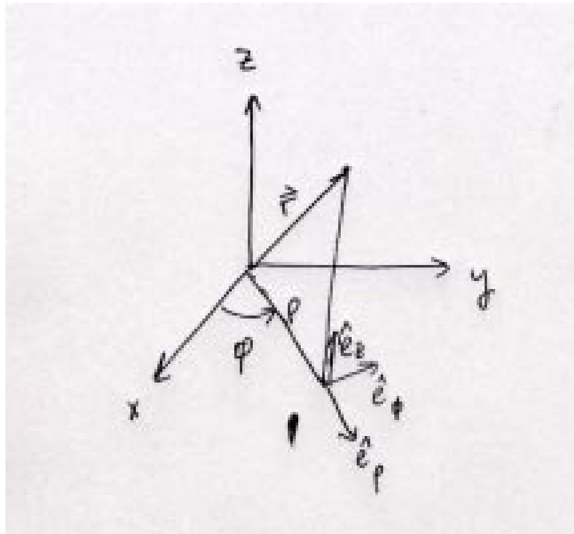
1.6.1 Adiabatic Invariant for a Charged Particle in a Magnetic Field

There is a very important adiabatic invariant used in many fields involving charged particles moving in magnetic fields. It can be expressed in many ways, one of which is that the magnetic moment of a charged particle circulating in a magnetic field that changes slowly in time is invariant. This also applies if a particle drifts along circulating around magnetic field lines in an inhomogeneous magnetic field so that the average motion is through a changing magnetic field.

Let's derive this using Hamilton's methods and the definition of an adiabatic invariant.

For constant **B** we need **A**. Although it is not unique,

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r} \tag{158}$$



works. Take $\mathbf{B} = B\hat{\mathbf{e}}_z$ and use cylindrical coordinates

$$\mathbf{v} = \frac{dz}{dt}\hat{\mathbf{e}}_z + \left(\rho\dot{\phi}\right)\hat{\mathbf{e}}_\phi + \dot{\rho}\hat{\mathbf{e}}_\rho \tag{159}$$

So

$$A = \frac{1}{2}B\hat{\mathbf{e}}_z \times (\rho\hat{\mathbf{e}}_\rho + z\hat{\mathbf{e}}_z) = \frac{1}{2}B\rho\hat{\mathbf{e}}_\phi \tag{160}$$

so $\mathbf{A}$ circulates around the z axis and increases linearly in ρ.

The Lagrangian (refer to our previous discussion of L for a charged particle in a field) is

$$L = \frac{1}{2}m\left(\dot{\rho}^2 + \left(\rho\dot{\phi}\right)^2 + \dot{z}^2\right) + \frac{e}{c}\frac{1}{2}B\rho^2\dot{\phi} \tag{161}$$

The canonical momenta are

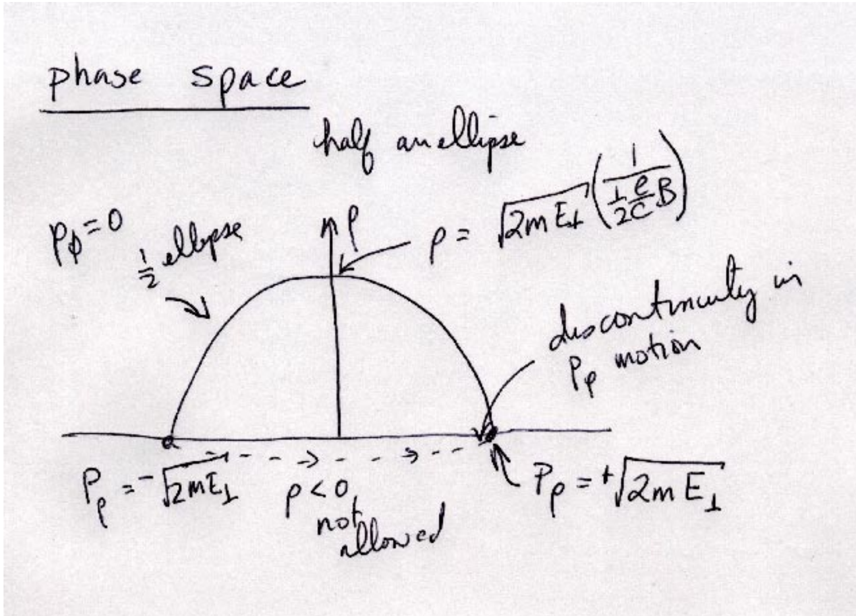
$$\begin{aligned} p_\rho &= m\dot{\rho} \\ p_\phi &= m\rho^2\dot{\phi} + \frac{1}{2}\frac{e}{c}B\rho^2 \\ p_z &= m\frac{dz}{dt} \end{aligned}$$

z is cyclic, ϕ is cyclic, and $\frac{\partial L}{\partial t} = 0$, so $p_z = const$, $p_\phi = const$, $H = E = const$. The Hamiltonian is

$$H = \frac{p_\rho^2}{2m} + \frac{p_z^2}{2m} + \frac{\left(p_\phi - \frac{1}{2}\frac{e}{c}B\rho^2\right)^2}{2m\rho^2} = E \tag{162}$$

Thus the motion, treated as a trajectory in phase space moves on "the energy shell" defined by

$$E - \frac{p_z^2}{2m} = E_\perp = \frac{p_\rho^2}{2m} + \frac{1}{2m}\left(\frac{p_\phi}{\rho} - \frac{1}{2}\frac{e}{c}B\rho\right)^2 = const \tag{163}$$



This algebraic equation gives a closed curve in the (ρ, p_ρ) plane that is quite simple in just one case: the special case in which $p_\phi = 0$. It is the only case in which $\rho = 0$ is allowed, or the orbit passes through the z-axis.

For $p_\phi = 0$,

$$E_\perp = \frac{p_\rho^2}{2m} + \frac{1}{2m}\left(\frac{1}{2}\frac{e}{c}B\rho\right)^2 \qquad (\rho \geq 0) \tag{164}$$

This is the equation for half an ellipse, so we have the following geometry in phase space:

and in configuration space

and

$$p_\phi = 0 \rightarrow \dot{\phi} = -\frac{1}{2}\frac{eB}{mc} \tag{165}$$

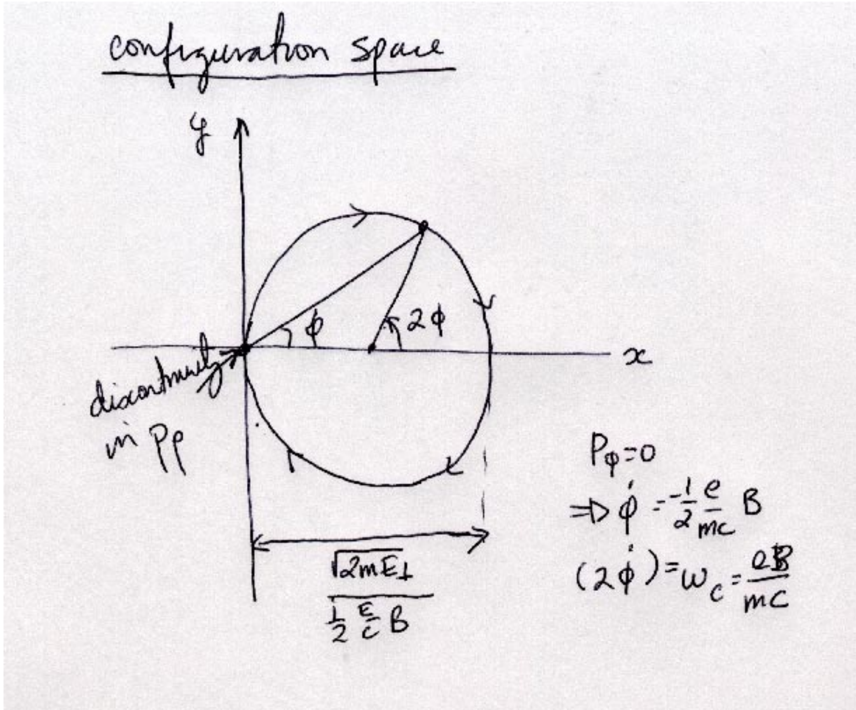
and

$$2\dot{\phi} = \omega_c = \frac{eB}{mc} \tag{166}$$

In this simple case we get out the cyclotron frequency, and we also get the adiabatic invariant:

$$\begin{aligned} 2\pi I &= \frac{1}{2}\pi\rho_{\max}p_{\rho\,\max} \\ &= \frac{1}{2}\pi\sqrt{2mE_\perp}\frac{1}{\frac{1}{2}\omega_c m}\sqrt{2mE_\perp} \\ &= 2\pi\frac{E_\perp}{\omega_c} \end{aligned}$$

Since it is always possible to choose the origin of the coordinates so that the elliptical orbit of the particle passes through the z-axis, the above result is in fact the general case for the



adiabatic invariant of a charged particle in a $B-$field. We can express it in various ways (see HW):

a) the radius of a particle's orbit changes inversely as the square root of the magnetic field.

b) the amount of magnetic flux linking the particle's orbit is a constant

c) the magnetic momentof the circulating charged particle is a constant.

1.7 Action-Angle Variables

We'll now consider periodic systems where λ is constant, so that the system is closed. We want to perform a cononical transformation from the old q, p to a conjugate pair where I is the new "momentum". The generating function that will get us there is

$$F = W(q, E; \lambda) = \int p(q, E; \lambda)\, dq \tag{167}$$

taken for a given constant E and λ. For a closed system, we can replace E with I, since it is a function of the energy, and we can write $W = W(q, I; \lambda)$ and

$$\frac{\partial W}{\partial q}|_E = \frac{\partial W}{\partial q}|_I \tag{168}$$

and so

$$p = \frac{\partial W(q, I; \lambda)}{\partial q} \tag{169}$$

(from the formulae for canonical transformations). We can use the other relation for canonical transformation to get the "coordinate" variable

$$\psi = \frac{\partial W(q, I; \lambda)}{\partial I} \tag{170}$$

So I and ψ are canonical variables, where I is called the *action variable*, and ψ the *angle variable*.

We are considering a conservative system with a time-independent Hamiltonian, and we have therefore used a generating not explicitly dependent on time. The new H' is therefore just H expressed in terms of the new variables. H' is just $E(I)$ expressed as a function of the action variable. Hamilton's equations in the action-angle variables are

$$\frac{dI}{dt} = 0, \quad \dot{\psi} = \frac{dE(I)}{dI} \tag{171}$$

The first shows that I is constant (as we knew). The second equation shows that ψ is linearly increasing with time

$$\psi = \frac{dE}{dI} t + const = \omega(I) t + const \tag{172}$$

and we equate it with the phase of the oscillations.

$W(q, I)$ is a many-valued function of the coordinates which increases each period by

$$\Delta W = 2\pi I \tag{173}$$

We see this from the definition $W = \int p dq$, and $I = \frac{1}{2\pi} \oint p dq$.

If we express q, p in terms of the action-angle variables, these must remain unchanged when $\psi \to \psi + 2\pi$ (with I const). So q, p are periodic functions of ψ with period 2π.

The action-angle variables may seem to be of very limited utility given the restrictions, but they are important historically and also in numerous astronomical contexts. These action-angle variables can also be used to formulate the EOM when the system is not closed, and $\lambda = \lambda(t)$. Then we still have $p = \frac{\partial W(q,I;\lambda)}{\partial q}$ and $\psi = \frac{\partial W(q,I;\lambda)}{\partial I}$, and

$$W(q, E; \lambda) = \int p(q, E; \lambda) dq \tag{174}$$

In the same approximation we used to get the adiabatic invariant, we calculate $W(q, E; \lambda) = \int p(q, E; \lambda) dq$ and $I = \frac{1}{2\pi} \oint p dq$ taking λ to have a fixed value, so that $W(q, E; \lambda)$ is the same function it was before, but we then allow λ to be $\lambda(t)$.

The generating function is now an explicit function of time, so we get H' from

$$\begin{aligned} H' &= E(I; \lambda) + \frac{\partial W}{\partial t} \\ &= E(I; \lambda) + \frac{\partial W}{\partial \lambda}|_{q,I} \dot{\lambda} \end{aligned}$$

where we express $\frac{\partial W}{\partial \lambda}|_{q,I}$ in terms of I and ψ after differentiating with respect to λ.

Hamilton's equations are then

$$\frac{dI}{dt} = -\frac{\partial H'}{\partial \psi} = -\frac{\partial}{\partial \psi}\frac{\partial W}{\partial \lambda}|_{q,I}\ \ \dot{\lambda} \tag{175}$$

$$\dot{\psi} = \frac{\partial H'}{\partial I} = \omega\left(I;\lambda\right) + \left(\frac{\partial}{\partial I}\left(\frac{\partial W}{\partial \lambda}|_{q,I}\right)|_{\psi,\lambda}\right)\dot{\lambda} \tag{176}$$

where $\omega = \left(\frac{\partial E}{\partial I}\right)_{\lambda}$ is the oscillation frequency calculated as if λ were constant.