Chapter 4

Canonical Transformations, Hamilton-Jacobi Equations, and Action-Angle Variables

We've made good use of the Lagrangian formalism. Here we'll study dynamics with the Hamiltonian formalism. Problems can be greatly simplified by a good choice of generalized coordinates. How far can we push this?

Example: Let us imagine that we find coordinates q_i that are all cyclic. Then $\dot{p}_i = 0$, so $p_i = \alpha_i$ are all constant. If H is conserved, then:

$$H = H(\alpha_1, \ldots, \alpha_n) \tag{4.1}$$

is also constant in time. In such a case the remaining equations of motion:

$$\dot{q}_i = \frac{\partial H}{\partial \alpha_i} = \omega_i(\alpha) \quad \Rightarrow \quad q_i = \omega_i\, t + \delta_i \tag{4.2}$$

All coordinates are linear in time and the motion becomes very simple.

We might imagine searching for a variable transformation to make as many coordinates as possible cyclic. Before proceeding along this path, we must see what transformations are allowed.

4.1 Generating Functions for Canonical Transformations

Recall the the Euler-Lagrange equations are invariant when:

- a point transformation occurs $Q = Q(q,t)$ with $L[q,t] = L'[Q,t]$;
- a total derivative is summed to the Lagrangian $L' = L + \frac{dF[q,t]}{dt}$.

For H we consider point transformations in phase space:

$$Q_i = Q_i(q,p,t) \quad \text{and} \quad P_i = P_i(q,p,t), \tag{4.3}$$

where the Hamilton's equations for the evolution of the canonical variables (q,p) are satisfied:

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{4.4}$$

Generally, not all transformations preserve the equations of motion. However, the transformation is *canonical* if there exists a new Hamiltonian:

$$K = K(Q,P,t), \tag{4.5}$$

where

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} \quad \text{and} \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i}. \tag{4.6}$$

For notational purposes let repeated indices be summed over implicitly.
Hamilton's principle can be written as:

$$\delta \int_{t_1}^{t_2} \left(p_i\dot{q}_i - H(q,p,t)\right) dt = 0, \tag{4.7}$$

or in the new Hamiltonian as:

$$\delta \int_{t_1}^{t_2} \left(P_i\dot{Q}_i - K(Q,P,t)\right) dt = 0. \tag{4.8}$$

For the Eq.(4.7) to imply Eq.(4.8), then we need:

$$\lambda(p_i\dot{q}_i - H) = P_i\dot{Q}_i - K + \dot{F}. \tag{4.9}$$

Since $\dot{F}$ is a total derivative and the ends of the path are fixed:

$$\delta q|_{t_1}^{t_2} = 0 \quad \text{and} \quad \delta p|_{t_1}^{t_2} = 0 \quad \Rightarrow \quad \delta F|_{t_1}^{t_2} = 0 \tag{4.10}$$

There are a few things to be said about transformations and λ.

- If $\lambda = 1$ then the transformation is *canonical*, which is what we will study.
- If $\lambda \neq 1$ then the transformation is *extended canonical*, and the results from $\lambda = 1$ can be recovered by rescaling q and p appropriately.

- If $Q_i = Q_i(q,p)$ and $P_i = P_i(q,p)$ without explicit dependence on time, then the transformation is *restricted canonical*.

We will always take transformations $Q_i = Q_i(q,p,t)$ and $P_i = P_i(q,p,t)$ to be invertible in any of the canonical variables. If F depends on a mix of old and new phase space variables, it is called a *generating function* of the canonical transformation. There are four important cases of this.

1. Let us take

$$F = F_1(q,Q,t) \tag{4.11}$$

where the old coordinates q_i and the new coordinates Q_i are independent. Then:

$$p_i\dot{q}_i - H = P_i\dot{Q}_i - K + \dot{F}_1 = P_i\dot{Q}_i - K + \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial q_i}\dot{q}_i + \frac{\partial F_1}{\partial Q_i}\dot{Q}_i \tag{4.12}$$

from this we see that $P_i\dot{Q}_i$ cancels and equating the terms with a $\dot{q}_i$, a $\dot{Q}_i$ and the remaining terms gives:

$$p_i = \frac{\partial F_1}{\partial Q_i} \quad , \quad P_i = -\frac{\partial F_1}{\partial Q_i} \quad \text{and} \quad K = H + \frac{\partial F_1}{\partial t}, \tag{4.13}$$

which gives us formula for a transformation:

$$p_i = p_i(q,Q,t) \quad \text{and} \quad P_i = P_i(q,Q,t) \tag{4.14}$$

and connects K to an initial H.

Example: if

$$F_1 = -\frac{Q}{q}, \tag{4.15}$$

then:

$$p = \frac{\partial F_1}{\partial q} = \frac{Q}{q^2} \quad \text{and} \quad P = -\frac{\partial F_1}{\partial Q} = \frac{1}{q}. \tag{4.16}$$

Writing the new coordinates as function of the old ones yields

$$Q = pq^2 \quad \text{and} \quad P = \frac{1}{q} \tag{4.17}$$

Example: Given the transformations

$$Q = \ln\left(\frac{p}{q}\right) \quad \text{and} \quad P = -\left(\frac{q^2}{2} + 1\right)\frac{p}{q}, \tag{4.18}$$

we can prove they are canonical by finding a corresponding generating function. We know:

$$\frac{\partial F_1}{\partial q} = p = qe^Q, \tag{4.19}$$

which gives us

$$F_1 = \int qe^Q \, dq + g(Q) = \frac{q^2}{2}e^Q + g(Q), \tag{4.20}$$

and

$$\begin{aligned} P = -\frac{\partial F_1}{\partial Q} = -\frac{q^2}{2}e^Q - \frac{dg}{dQ} = -\left(\frac{q^2}{2}+1\right)\frac{p}{q} = -\left(\frac{q^2}{2}+1\right)e^Q \\ \Rightarrow g(Q) = e^Q \,. \end{aligned} \tag{4.21}$$

Thus F_1 is given by:

$$F_1 = \left(\frac{q^2}{2}+1\right)e^Q. \tag{4.22}$$

2. Let:

$$F = F_2(q, P, t) - Q_i P_i \tag{4.23}$$

where we wish to treat the old coordinates q_i and new momenta P_i as independent variables. Then:

$$\dot{q}_i p_i - H = \dot{Q}_i P_i - K + \dot{F}_2 - \dot{Q}_i P_i - Q_i \dot{P}_i = -Q_i \dot{P}_i - K + \frac{\partial F_2}{\partial t} + \frac{\partial F_2}{\partial q_i}\dot{q}_i + \frac{\partial F_2}{\partial P_i}\dot{P}_i \tag{4.24}$$

This corresponds to

$$p_i = \frac{\partial F_2}{\partial q_i} \quad ; \quad Q_i = \frac{\partial F_2}{\partial P_i} \quad \text{and} \quad K = H + \frac{\partial F_2}{\partial t}. \tag{4.25}$$

3. We could also take

$$F = F_3(p, Q, t) + q_i p_i \tag{4.26}$$

with the new coordinates Q_i and the old momenta p_i as independent variables.

4. Finally we could take

$$F = F_4(p, P, t) + q_i p_i - Q_i P_i \tag{4.27}$$

with the old momenta p_i and new momenta P_i as independent variables.

This can be summarized in the table below.

Function	Transformations	Simplest case	
$F_1(q, Q, t)$	$p_i = \frac{\partial F_1}{\partial q_i}, P_i = -\frac{\partial F_1}{\partial Q_i}$	$F_1 = q_i Q_i$	$Q_i = p_i, P_i = -q_i$
$F_2(q, P, t)$	$p_i = \frac{\partial F_2}{\partial q_i}, Q_i = \frac{\partial F_2}{\partial P_i}$	$F_2 = q_i P_i$	$Q_i = q_i, P_i = p_i$
$F_3(p, Q, t)$	$q_i = -\frac{\partial F_3}{\partial p_i}, P_i = -\frac{\partial F_3}{\partial Q_i}$	$F_3 = p_i Q_i$	$Q_i = -q_i, P_i = -p_i$
$F_4(p, P, t)$	$q_i = -\frac{\partial F_4}{\partial p_i}, Q_i = \frac{\partial F_4}{\partial P_i}$	$F_4 = p_i P_i$	$Q_i = p_i, P_i = -q_i$

The simplest case of the 2^{nd} (F_2) transformation is just an identity transformation. For any of these F_i cases we also have:

$$K = H + \frac{\partial F_i}{\partial t}. \tag{4.28}$$

If F_i is independent of time then this implies

$$K = H \tag{4.29}$$

Mixed cases may also occur when more than two old canonical coordinates are present. (In this chapter we will be using Einstein's repeated index notation for implicit summation, unless otherwise stated.)

Example: consider

$$F_2 = f_i(q, t) P_i \tag{4.30}$$

for some functions f_i where $i \in \{1, \dots, n\}$. Then

$$Q_i = \frac{\partial F_2}{\partial P_i} = f_i(q, t) \tag{4.31}$$

is a coordinate point transformation. It is canonical with

$$p_i = \frac{\partial f_i}{\partial q_j} P_j, \tag{4.32}$$

which can be inverted to get $P_j = P_j(q, p, t)$.

Example: Consider the harmonic oscillator:

$$H = \frac{p^2}{2m} + \frac{kq^2}{2} \quad \text{where} \quad k = m\omega^2 \tag{4.33}$$

Let us try the transformation:

$$p = \alpha\sqrt{2mP}\cos(Q) \tag{4.34}$$
$$q = \frac{\alpha}{m\omega}\sqrt{2mP}\sin(Q)$$

for α constant. Then:

$$K = H = P\alpha^2\left(\cos^2(Q) + \sin^2(Q)\right) = P\alpha^2, \tag{4.35}$$

so the new momentum

$$P = \frac{E}{\alpha^2} \tag{4.36}$$

is just proportional to the energy, while Q is a cyclic variable.

Is this transformation canonical? We can find a generating function $F = F_1(q, Q)$ by dividing the old variables:

$$\frac{p}{q} = m\omega\cot(Q). \tag{4.37}$$

This gives us:

$$p = \frac{\partial F_1}{\partial q} \Rightarrow F_1 = \int p(q,Q)\,dq + g(Q) = \frac{1}{2}m\omega q^2\cot(Q) + g(Q) \tag{4.38}$$
$$P = -\frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2\sin^2(Q)} - \frac{dg}{dQ}$$

Setting:

$$\frac{dg}{dQ} = 0 \Rightarrow q^2 = \frac{2P}{m\omega}\sin^2(Q), \tag{4.39}$$

which tells us the transformation is canonical if $\alpha = \sqrt{\omega}$. This means:

$$P = \frac{E}{\omega} \tag{4.40}$$

By Hamilton's equations Eq.(4.4):

$$\dot{Q} = \frac{\partial K}{\partial P} = \omega \Rightarrow Q = \omega t + \delta. \tag{4.41}$$

Putting this altogether, this gives the familiar results:

$$q = \sqrt{\frac{2E}{m\omega^2}}\sin(\omega t + \delta) \tag{4.42}$$
$$p = \sqrt{2mE}\cos(\omega t + \delta).$$

Lets record for future use our final canonical transformation here:

$$q = \sqrt{\frac{2P}{m\omega}}\sin(Q)\,, \qquad\qquad p = \sqrt{2m\omega P}\cos(Q)\,.$$

So far, a transformation $Q = Q(q, p, t)$ and $P = P(q, p, t)$ is canonical if we can find a generating function F. This involves integration, which could be complicated, so it would be nice to have a test that only involves differentiation. There is one!

4.2 Poisson Brackets and the Symplectic Condition

In Classical Mechanics II (8.223) the *Poisson bracket* of the quantities u and v was defined as

$$\{u, v\}_{q,p} = \sum_i \left(\frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial v}{\partial q_i} \frac{\partial u}{\partial p_i} \right) \tag{4.43}$$

It is easy to check that the following fundamental Poisson bracket relations are satisfied:

$$\{q_i, q_j\}_{q,p} = \{p_i, p_j\}_{q,p} = 0 \quad \text{and} \quad \{q_i, p_j\}_{q,p} = \delta_{ij}. \tag{4.44}$$

There are a few other properties of note. These include:

$$\{u, u\} = 0, \tag{4.45}$$

$$\{u, v\} = -\{v, u\}, \tag{4.46}$$

$$\{au + bv, w\} = a\{u, w\} + b\{v, w\}, \tag{4.47}$$

$$\{uv, w\} = u\{v, w\} + \{u, w\}v, \tag{4.48}$$

$$\{u, \{v, w\}\} + \{v, \{w, u\}\} + \{w, \{u, v\}\} = 0, \tag{4.49}$$

for a, b constants. Eq.(4.49) is the Jacobi identity.

The above looks a lot like the commutators of operators in quantum mechanics, such as:

$$[\hat{x}, \hat{p}] = \mathrm{i}\hbar \tag{4.50}$$

Indeed, quantizing a classical theory by replacing Poisson brackets with commutators through:

$$[u, v] = i\hbar\{u, v\} \tag{4.51}$$

is a popular approach (first studied by Dirac). It is also the root of the name "canonical quantization". (Note that Eq.(4.48) was written in a manner to match the analogous formula in quantum mechanics where the operator ordering is important, just in case its familiar. Here we can multiply functions in either order.)

Now we can state the desired criteria that only involves derivatives.

Theorem: A transformation $Q_j = Q_j(q, p, t)$ and $P_j = P_j(q, p, t)$ is canonical if and only if:

$$\{Q_i, Q_j\}_{q,p} = \{P_i, P_j\}_{q,p} = 0 \quad \text{and } \{Q_i, P_j\}_{q,p} = \delta_{ij}. \tag{4.52}$$

To prove it, we'll need some more notation. Let's get serious about treating q_i and p_i on an equal footing together, defining the following two quantities:

$$\vec{\eta} = \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ p_1 \\ \vdots \\ p_n \end{pmatrix} \quad \text{and} \quad J = \begin{bmatrix} 0_{n\times n} & \mathbb{1}_{n\times n} \\ -\mathbb{1}_{n\times n} & 0_{n\times n} \end{bmatrix} \tag{4.53}$$

where $0_{n\times n}$ is the $n \times n$ zero matrix, $\mathbb{1}_{n\times n}$ is the $n \times n$ identity matrix. The following properties of J will be useful:

$$J^\top = -J \quad , \quad J^2 = -\mathbb{1}_{2n\times 2n} \quad \text{and} \quad J^\top J = JJ^\top = \mathbb{1}_{2n\times 2n}\,. \tag{4.54}$$

We also note that $\det(J) = 1$.

With this notation Hamilton's equations, Eq.(4.4), can be rewritten as:

$$\dot{\vec{\eta}} = J\frac{\partial H}{\partial \vec{\eta}} \quad \text{or} \quad \dot{\vec{\eta}} = J\nabla_{\vec{\eta}} H\,. \tag{4.55}$$

The notation $\nabla_{\vec{\eta}} H$ better emphasizes that this quantity is a vector, but we will stick to using the first notation for this vector, $\partial H/\partial \vec{\eta}$, below.

Although the Theorem is true for time dependent transformations, lets carry out the proof for the simpler case of time independent transformations $Q_i = Q_i(q,p)$ and $P_i = P_i(q,p)$. This implies $K = H$. Let us define:

$$\vec{\xi} = \begin{pmatrix} Q_1 \\ \vdots \\ Q_n \\ P_1 \\ \vdots \\ P_n \end{pmatrix} \tag{4.56}$$

which is a function of the original coordinates, so we can write:

$$\vec{\xi} = \vec{\xi}(\vec{\eta}) \tag{4.57}$$

Now consider the time derivative of $\vec{\xi}$:

$$\dot{\xi}_i = \frac{\partial \xi_i}{\partial \eta_j}\dot{\eta}_j \quad \Leftrightarrow \quad \dot{\vec{\xi}} = M\dot{\vec{\eta}} \quad \text{where} \quad M_{ij} = \frac{\partial \xi_i}{\partial \eta_j}. \tag{4.58}$$

Here M corresponds to the *Jacobian* of the transformation.

From the Hamilton's equations, we know that

$$\dot{\vec{\eta}} = J\frac{\partial H}{\partial \vec{\eta}}. \tag{4.59}$$

We want to show that :

$$\dot{\vec{\xi}} = J\frac{\partial H}{\partial \vec{\xi}} \quad \text{for} \quad \vec{\xi} = \vec{\xi}(\vec{\eta}) \text{ a canonical transformation.} \tag{4.60}$$

Let us now consider:

$$\dot{\xi}_i = \frac{\partial \xi_i}{\partial \eta_j}\dot{\eta}_j = \frac{\partial \xi_i}{\partial \eta_j} J_{jk} \frac{\partial H}{\partial \eta_k} = \frac{\partial \xi_i}{\partial \eta_j} J_{jk} \frac{\partial \xi_l}{\partial \eta_k}\frac{\partial H}{\partial \xi_l} \quad \Leftrightarrow \quad \dot{\vec{\xi}} = MJM^\top \frac{\partial H}{\partial \vec{\xi}} \tag{4.61}$$

for any H. Then

$$\vec{\xi} = \vec{\xi}(\vec{\eta}) \text{ is a } \textit{canonical transformation} \quad \text{iff} \quad MJM^\top = J \tag{4.62}$$

is satisfied. This is known as the "symplectic condition". Moreover, since

$$MJ = J(M^\top)^{-1} \quad \text{and} \quad J^2 = -\mathbb{1}, \tag{4.63}$$

we can write:

$$J(MJ)J = -JM = J(JM^{\top-1})J = -M^{\top-1}J \Rightarrow JM = M^{\top-1}J\,. \tag{4.64}$$

Thus we see that $MJM^\top = J$ is equivalent to:

$$M^\top JM = J\,. \tag{4.65}$$

Now consider Poisson brackets in this matrix notation:

$$\{u,v\}_{q,p} = \{u,v\}_{\vec{\eta}} = \left(\frac{\partial u}{\partial \vec{\eta}}\right)^\top J\,\frac{\partial v}{\partial \vec{\eta}} \tag{4.66}$$

and the fundamental Poisson brackets are:

$$\{\eta_i,\eta_j\}_{\vec{\eta}} = J_{ij} \tag{4.67}$$

Then we can calculate the Poisson brackets that appeared in the theorem we are aiming to prove as

$$\{\xi_i,\xi_j\}_{\vec{\eta}} = \left(\frac{\partial \xi_i}{\partial \vec{\eta}}\right)^\top J\frac{\partial \xi_j}{\partial \vec{\eta}} = (M^\top JM)_{ij} \tag{4.68}$$

This last equation is the same as Eq.(4.65). The new variables satisfy the Poisson bracket relationships Eq.(4.67):

$$\{\xi_i,\xi_j\}_{\vec{\eta}} = J_{ij} \tag{4.69}$$

if and only if

$$M^\top JM = J \tag{4.70}$$

which itself is true if, and only if, $\vec{\xi} = \vec{\xi}(\vec{\eta})$ is canonical, Eq.(4.65), completing the proof.

There are two facts that arise from this.

- Poisson brackets are canonical invariants

$$\{u, v\}_{\vec{\eta}} = \{u, v\}_{\vec{\xi}} = \{u, v\}. \tag{4.71}$$

This is true because:

$$\{u, v\}_{\vec{\eta}} = \left(\frac{\partial u}{\partial \vec{\eta}}\right)^{\top} J \frac{\partial v}{\partial \vec{\eta}} = \left(M^{\top} \frac{\partial u}{\partial \vec{\xi}}\right)^{\top} J \left(M^{\top} \frac{\partial v}{\partial \vec{\xi}}\right) \tag{4.72}$$

$$= \left(\frac{\partial u}{\partial \vec{\xi}}\right)^{\top} M J M^{\top} \frac{\partial v}{\partial \vec{\xi}} = \left(\frac{\partial u}{\partial \vec{\xi}}\right)^{\top} J \frac{\partial v}{\partial \vec{\xi}} = \{u, v\}_{\vec{\xi}} \tag{4.73}$$

- Phase space volume elements are preserved by canonical transformations, as discussed in 8.223. Phase space volume is given by:

$$V_{\vec{\xi}} = \prod_i dQ_i dP_i = |\det(M)| \prod_j dq_j dp_j = |\det(M)| V_{\vec{\eta}}. \tag{4.74}$$

However, we also have:

$$\det(M^{\top} J M) = \det(J) = (\det(M))^2 \det(J) \Rightarrow |\det(M)| = 1. \tag{4.75}$$

4.3 Equations of Motion & Conservation Theorems

Let us consider a function:

$$u = u(q, p, t) \tag{4.76}$$

Then:

$$\dot{u} = \frac{\partial u}{\partial q_i} \dot{q}_i + \frac{\partial u}{\partial p_i} \dot{p}_i + \frac{\partial u}{\partial t} = \frac{\partial u}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial u}{\partial t}, \tag{4.77}$$

which can be written more concisely as

$$\dot{u} = \{u, H\} + \frac{\partial u}{\partial t} \tag{4.78}$$

for any canonical variables (q, p) and corresponding Hamiltonian H. Performing canonical quantization on this yields the Heisenberg equation of time evolution in quantum mechanics. There are a few easy cases to check.

- If $u = q_i$ then:

$$\dot{q}_i = \{q_i, H\} + \frac{\partial q_i}{\partial t} = \frac{\partial H}{\partial p_i} \tag{4.79}$$

- If $u = p_i$ then:

$$\dot{p}_i = \{p_i, H\} + \frac{\partial p_i}{\partial t} = -\frac{\partial H}{\partial q_i} \tag{4.80}$$

  Together the above two cases yield Hamilton's equations of motion.

- Also, if $u = H$ then:

$$\dot{H} = \{H, H\} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \tag{4.81}$$

Next consider what we can say if a quantity u is conserved. Here:

$$\dot{u} = 0 = \{u, H\} + \frac{\partial u}{\partial t}. \tag{4.82}$$

As a corollary, if

$$\frac{\partial u}{\partial t} = 0, \tag{4.83}$$

then

$$\{u, H\} = 0 \Rightarrow u \text{ is conserved.} \tag{4.84}$$

(In quantum mechanics this the analog of saying that u is conserved if u commutes with H.)

Another fact, is that if u and v are conserved then so is $\{u, v\}$. This could potentially provide a way to compute a new constant of motion. To prove it, first consider the special case where:

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial t} = 0 \tag{4.85}$$

then using the Jacobi identity we have:

$$\{H, \{u, v\}\} = -\{u, \{v, H\}\} - \{v, \{H, u\}\} = -\{u, 0\} - \{v, 0\} = 0 \tag{4.86}$$

For the most general case we proceed in a similar manner:

$$\begin{aligned}\{\{u, v\}, H\} &= \{u, \{v, H\}\} + \{v, \{H, u\}\} = -\left\{u, \frac{\partial v}{\partial t}\right\} + \left\{v, \frac{\partial u}{\partial t}\right\} \\ &= -\frac{\partial}{\partial t}\{u, v\} \quad \Rightarrow \quad \frac{d}{dt}\{u, v\} = 0\end{aligned} \tag{4.87}$$

<u>Infinitesimal Canonical Transformations</u>

Let us now consider the generating function:

$$F_2(q, P, t) = q_i P_i + \epsilon G(q, P, t), \tag{4.88}$$

where $F_2 = q_i P_i$ is an identity transformation, and $|\epsilon| \ll 1$ is infinitesimal. The function $G(q, P, t)$ is known as the generating function of an infinitesimal canonical transformation. Using the properties of an F_2 generating function we have:

$$p_j = \frac{\partial F_2}{\partial q_j} = P_j + \epsilon \frac{\partial G}{\partial q_j} \quad \Rightarrow \quad \delta p_j = P_j - p_j = -\epsilon \frac{\partial G}{\partial q_j} \tag{4.89}$$

giving the infinitesimal transformation in the momentum. Likewise:

$$Q_j = \frac{\partial F_2}{\partial P_j} = q_j + \epsilon \frac{\partial G}{\partial P_j} \tag{4.90}$$

Since $P_j = p_j + O(\epsilon)$ and ϵ is infinitesimal we can replace $\partial G(q, P, t)/\partial P_j = \partial G(q, p, t)/\partial p_j + O(\epsilon)$. Therefore we have:

$$Q_j = q_j + \epsilon \frac{\partial G}{\partial p_j} + O(\epsilon^2) \quad \Rightarrow \quad \delta q_j = Q_j - q_j = \epsilon \frac{\partial G}{\partial p_j} \tag{4.91}$$

where now we note that we can consider $G = G(q, p, t)$, a function of q and p, to this order. Returning to the combined notation of $\vec{\eta}^{\top} = (q_1, \ldots, q_n, p_1, \ldots, p_n)$, Eq.(4.89) and Eq.(4.90) can be consisely written as the following Poisson bracket:

$$\delta \vec{\eta} = \epsilon \{\vec{\eta}, G\} \tag{4.92}$$

Example: if $G = p_i$ then $\delta p_i = 0$ and $\delta q_j = \epsilon \delta_{ij}$, which is why *momentum is the generator of spatial translations.*

Example: if G is the z component of the angular momentum:

$$G = L_z = \sum_i (x_i p_{iy} - y_i p_{ix}) \quad \text{and} \quad \epsilon = \delta\theta \tag{4.93}$$

then the infinitesimal change correponds to a rotation

$$\delta x_i = -y_i \delta\theta \quad , \quad \delta y_i = x_i \delta\theta \quad , \quad \delta z_i = 0 \tag{4.94}$$

$$\delta p_{ix} = -p_{iy} \delta\theta \quad , \quad \delta p_{iy} = p_{ix} \delta\theta \quad , \quad \delta p_{iz} = 0 \tag{4.95}$$

which is why *angular momentum is the generator of rotations.*

Important Example: if $G = H$ and $\epsilon = dt$ then

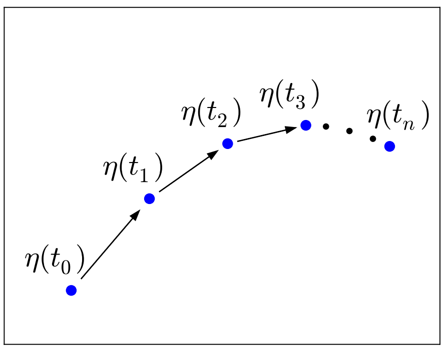
$$\epsilon\{\vec{\eta}, G\} = \{\vec{\eta}, H\} dt = \dot{\vec{\eta}} dt = d\vec{\eta}$$

On the left hand side we have the change to the phase space coordinates due to our transformation. On the right hand side we have the physical increment to the phase space variables that takes place in time dt. The fact that these are equivalent tells us that *the Hamiltonian is the generator of time evolution*. The infinitesimal transformation generated by the Hamiltonian is corresponds with the physical motion.

Rather than trying to think of this as a passive coordinate change $(q,p) \to (Q,P)$, it is useful if we can take an *active view* of the infinitesimal canonical transformation generated by H. Let the time t be a parameter for the family of transformations with $\epsilon = dt$: the initial conditions are:

$$\vec{\eta}_0(t_0) = \vec{\eta}_0 \tag{4.96}$$

The result is a series of transformations of $\vec{\eta}$ that move us in a fixed set of phase space coordinates from one point to another:



Phase Space $\vec{\eta}$

$$\vec{\eta}_0(t_0) \to \vec{\eta}_1(\vec{\eta}_0, t_1) \to \ldots \to \vec{\eta}_n(\vec{\eta}_{n-1}, t_n)$$
$$\text{where } \ t_n = t \text{ is the final time} \tag{4.97}$$

All together, combining an infinite number of infinitesimal transformations allows us to make a finite transformation, resulting in:

$$\vec{\eta} = \vec{\eta}(\vec{\eta}_0, t) \quad \text{ or } \quad \vec{\eta}_0 = \vec{\eta}_0(\vec{\eta}, t) \tag{4.98}$$

This is a canonical transformation that yields a solution for the motion!

How could we directly find this transformation, without resorting to stringing together infinitesimal transformations? We can simply look for a canonical transformation with new coordinates Q_i and new momenta P_i that are all constants, implying an equation of the type:

$$\vec{\eta}_0 = \vec{\eta}_0(\vec{\eta}, t) \tag{4.99}$$

Inverting this then gives the solution for the motion.

This logic can be used to extend our proof of the Theorem in Section 4.2 to fully account for time dependent transformations. (see Goldstein). Using $K = H + \epsilon \partial G/\partial t$, Goldstein also describes in some detail how the change to the Hamiltonian ΔH under an active infinitesimal canonical transformation satisfies:

$$\Delta H = -\epsilon\{G, H\} - \epsilon\frac{\partial G}{\partial t} = -\epsilon\dot{G} \tag{4.100}$$

This says "the constants of motion are generating functions G of the infinitesimal canonical transformation that leave H invariant"; that is, $\dot{G} = 0$ if and only if $\Delta H = 0$ under the transformation. Thus a conservation law exists if and only if there is a symmetry present.

4.4 Hamilton-Jacobi Equation

Let us take the suggestion from the end of the previous section seriously and look for new canonical variables that are all cyclic, such that:

$$\dot{Q}_i = \dot{P}_i = 0 \quad \Rightarrow \quad (Q, P) \text{ are all constants} . \tag{4.101}$$

If the new Hamiltonian K is independent of (Q, P) then:

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} = 0 \quad \text{and} \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0 . \tag{4.102}$$

We could look for a constant K, but it it is simplest to simply look for $K = 0$.

Using a generating function $F = F_2(q, P, t)$, then we need

$$K = H(q, p, t) + \frac{\partial F_2}{\partial t} = 0 . \tag{4.103}$$

Because $p_i = \frac{\partial F_2}{\partial q_i}$, then we can rewrite this condition as

$$H\left(q_1, \ldots, q_n, \frac{\partial F_2}{\partial q_1}, \ldots, \frac{\partial F_2}{\partial q_n}, t\right) + \frac{\partial F_2}{\partial t} = 0, \tag{4.104}$$

which is the time dependent *Hamilton-Jacobi equation* (henceforth abbreviated as the H-J equation). This is a 1^{st} order partial differential equation in $n + 1$ variables $(q_1, \ldots, q_n, t)$ for F_2. The solution for F_2 has $n + 1$ independent constants of integration, One of these constants is trivial ($F_2 \to F_2 + C$ for a pure constant C), so we'll ignore this one. Hence, suppose the solution is:

$$F_2 \equiv S = S(q_1, \ldots, q_n, \alpha_1, \ldots \alpha_n, t), \tag{4.105}$$

where S is called Hamilton's principal function and each α_i is an independent constant. We can pick our new momenta to be the constants of integration $P_i = \alpha_i$ for $i \in \{1, \ldots, n\}$ (so that $\dot{P}_i = 0$), thus specifying $F_2 = F_2(q, P, t)$ as desired. Then, using again the property of an F_2 generating function (and $K = 0$), we have that the new constant variables are:

$$P_i \equiv \alpha_i \quad \text{and} \quad Q_i \equiv \beta_i = \frac{\partial S(q, \alpha, t)}{\partial \alpha_i} . \tag{4.106}$$

We introduce the notation β_i to emphasize that these are constants.

From these results we can obtain a solution for the motion as follows. From the invertibility of our transformations we have:

$$\beta_i(q, \alpha, t) = \frac{\partial S}{\partial \alpha_i} \quad \Rightarrow \quad q_i = q_i(\alpha, \beta, t) , \tag{4.107}$$

$$p_i(q, \alpha, t) = \frac{\partial S}{\partial q_i} \quad \Rightarrow \quad p_i = p_i(q, \alpha, t) = p_i(q(\alpha, \beta, t), \alpha, t) = p_i(\alpha, \beta, t) .$$

(Note that function argument notation has been abused slightly here since $p_i(q, \alpha, t)$ and $p_i(\alpha, \beta, t)$ are technically different functions of their three arguments. Since we are always sticking explicit variables into the slots this should not cause confusion.) If desired, we can also swap our $2n$ constants α_i and β_i for $2n$ initial conditions q_{i0} and p_{i0}, to obtain a solution for the initial value problem. We obtain one set of constants in terms of the other set by solving the $2n$ equations obtained from the above results at $t = t_0$:

$$q_{i0} = q_i(\alpha, \beta, t_0)\,, \qquad\qquad p_{i0} = p_i(\alpha, \beta, t_0)\,. \tag{4.108}$$

Thus we see that Hamilton's principal function S is the generator of canonical transformations of constant (Q, P), and provides a method of obtaining solutions to classical mechanics problems by way of finding a transformation.

There are a few comments to be made about this.

1. The choice of constants α_i is somewhat arbitrary, as any other independent choice $\gamma_i = \gamma_i(\alpha)$ is equally good. Thus, when solving the H-J equation, we introduce the constants α_i in whatever way is most convenient.

2. What is S? We know that:

$$\dot{S} = \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial P_i}\dot{P}_i + \frac{\partial S}{\partial t}, \tag{4.109}$$

   but we also know that:

$$\frac{\partial S}{\partial q_i} = p_i \quad , \quad \dot{P}_i = 0 \quad \text{and} \quad \frac{\partial S}{\partial t} = -H \tag{4.110}$$

   Putting Eq.(4.109) and Eq.(4.110) together we have:

$$\dot{S} = p_i\dot{q}_i - H = L \quad \Rightarrow \quad S = \int L\, dt \tag{4.111}$$

   Thus S is the classical action which is an indefinite integral over time of the Lagrangian (so it is no coincidence that the same symbol is used).

3. The H-J equation is also the semiclassical limit of the quantum mechanical Schrödinger equation (0'th order term in the WKB approximation). To see this consider the Schrödinger equation

$$i\hbar\frac{\partial \psi}{\partial t} = H\psi = \Big(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q)\Big)\psi\,, \tag{4.112}$$

with the wavefunction $\psi = \exp(iS/\hbar)$. At this point we are just making a change of variable, without loss of generality, and $S(q,t)$ is complex. Plugging it in, and canceling an exponential after taking the derivative, we find

$$-\frac{\partial S}{\partial t} = -\frac{i\hbar}{2m}\frac{\partial^2 S}{\partial q^2} + \frac{1}{2m}\Big(\frac{\partial S}{\partial q}\Big)^2 + V(q)\,. \tag{4.113}$$

This equation is just another way of writing the Schrödinger equation, to solve for a complex S instead of ψ. If we now take $\hbar \to 0$ then we find that the imaginary term goes away leaving

$$0 = \frac{\partial S}{\partial t} + \frac{1}{2m}\Big(\frac{\partial S}{\partial q}\Big)^2 + V(q) = \frac{\partial S}{\partial t} + H\Big(q, \frac{\partial S}{\partial q}\Big)\,, \tag{4.114}$$

which is the Hamilton-Jacobi equation for S with a standard $p^2/2m$ kinetic term in H.

Having set things up, it is always good for us to test a new formalism on an example where we know the solution.

Example: let us consider the harmonic oscillator Eq.(4.33):

$$H = \frac{1}{2m}\left(p^2 + (m\omega q)^2\right) = E \tag{4.115}$$

Here we will look for one constant $P = \alpha$ and one constant $Q = \beta$. The H-J equation says

$$\frac{1}{2m}\left(\left(\frac{\partial S}{\partial q}\right)^2 + (m\omega q)^2\right) + \frac{\partial S}{\partial t} = 0. \tag{4.116}$$

In solving this, we note that the dependence of S on q and t is *separable*

$$S(q,\alpha,t) = W(q,\alpha) + g(\alpha,t), \tag{4.117}$$

which gives:

$$\frac{1}{2m}\left(\left(\frac{\partial W}{\partial q}\right)^2 + (m\omega q)^2\right) = -\frac{\partial g}{\partial t} = \alpha\,. \tag{4.118}$$

Since the left side is independent of t and the right hand side is independent of q, then the result must be equal to a separation constant α that is independent of q and t. We will choose our new $P = \alpha$. Now we have

$$\frac{\partial g}{\partial t} = -\alpha \Rightarrow g = -\alpha t \tag{4.119}$$

where we have avoided the addition of a further additive constant (since our convention was to always drop an additive constant when determining S). To identify what α is note that

$$H = -\frac{\partial S}{\partial t} = -\frac{\partial g}{\partial t} = \alpha, \tag{4.120}$$

which corresponds to the constant energy,

$$\alpha = E. \tag{4.121}$$

The other equation we have to solve is

$$\frac{1}{2m}\left(\left(\frac{\partial W}{\partial q}\right)^2 + (m\omega q)^2\right) = \alpha, \tag{4.122}$$

so rearranging and integrating this gives the indefinite integral

$$W = \pm \int \sqrt{2m\alpha - (m\omega q)^2}\, dq\,, \tag{4.123}$$

which we will leave unintegrated until we must do so. The full solution is then given by:

$$S = -\alpha t \pm \int \sqrt{2m\alpha - (m\omega q)^2}\, dq. \tag{4.124}$$

With this result for Hamilton's Principal function in hand we can now solve for the equations of motion. The equations of motion come from (we now do the integral, after taking the partial derivative):

$$\beta = \frac{\partial S}{\partial \alpha} = -t \pm m \int \frac{dq}{\sqrt{2m\alpha - (m\omega q)^2}} \Rightarrow t + \beta = \pm\frac{1}{\omega}\arcsin\left(\sqrt{\frac{m\omega^2}{2\alpha}}q\right). \tag{4.125}$$

Inverting gives:

$$q = \pm\sqrt{\frac{2\alpha}{m\omega^2}}\sin(\omega(t+\beta)), \tag{4.126}$$

so β is related to the phase. Next we consider p and use this result to obtain:

$$p = \frac{\partial S}{\partial q} = \pm\sqrt{2m\alpha - (m\omega q)^2} = \pm\sqrt{2m\alpha}\cos(\omega(t+\beta)) \tag{4.127}$$

These results are as expected. We can trade (α, β) for the initial conditions (q_0, p_0) at $t = 0$. The choice of phase (from shifting β so that $\omega\beta \to \omega\beta + \pi$) allows taking the positive sign of each square root in the solutions above.

Separation of variables is the main technique to solve the H-J equation. In particular, for a time independent H where

$$\dot{H} = \frac{\partial H}{\partial t} = 0 \tag{4.128}$$

we can always separate time by taking:

$$S(q, \alpha, t) = W(q, \alpha) - \alpha_1\, t, \tag{4.129}$$

where α_1 has been chosen as the separation constant, then plugging this into the time dependent H-J equation yields (just as in our Harmonic Oscillator example):

$$H\left(q_1, \ldots, q_n, \frac{\partial W}{\partial q_1}, \ldots, \frac{\partial W}{\partial q_n}\right) = \alpha_1 \,. \tag{4.130}$$

This result is referred to as the *time independent Hamilton-Jacobi equation*. Since $\dot{H} = 0$, H is conserved, and equal to a constant α_1. If $H = E$ then this constant is energy, $\alpha_1 = E$. The function W is called Hamilton's characteristic function.

The idea is now to solve the time independent H-J equation for $W = W(q, \alpha)$ where $P = \alpha$ still. If we follow the setup from our time dependent solution above then the equations of motion are obtained from the following prescription for identifying variables:

$$\begin{aligned} p_i &= \frac{\partial W}{\partial q_i} \quad \text{for} \quad i \in \{1, \ldots, n\}\,, \\ Q_1 &= \beta_1 = \frac{\partial S}{\partial \alpha_1} = \frac{\partial W}{\partial \alpha_1} - t\,, \\ Q_j &= \beta_j = \frac{\partial W}{\partial \alpha_j} \quad \text{for } j \in \{2, \ldots, n\} \text{ for } n > 1\,. \end{aligned} \tag{4.131}$$

Here all the Q_i are constants.

There is an alternative to the above setup, which allows us to not refer to the time dependent solution. The alternative is to consider $W = F_2(q, P)$ as the generating function, instead of S and only demand that all the new momenta P_i are constants with $P_1 = \alpha_1 = H$ for a time independent Hamiltonian H. At the start of chapter 4 we saw that this less restrictive scenario would lead to Qs that could have a linear time dependence, which is still pretty simple.

This is almost identical to the above setup but we rename and reshuffle a few things. The following three equations are the same as before:

$$p_i = \frac{\partial W}{\partial q_i} \quad , \quad P_i = \alpha_i \quad \text{and} \quad H\left(q, \frac{\partial W}{\partial q}\right) = \alpha_1 \tag{4.132}$$

However, now we have a non-zero K and different equation for Q_1:

$$K = H = \alpha_1 \quad \text{and} \quad Q_i = \frac{\partial W}{\partial \alpha_i} \text{ for all } i \in \{1, \ldots, n\}. \tag{4.133}$$

This means:

$$\dot{Q}_1 = \frac{\partial K}{\partial \alpha_1} = 1 \quad \Rightarrow \quad Q_1 = t + \beta_1 = \frac{\partial W}{\partial \alpha_1} \tag{4.134}$$

which is Eq. (4.131) but rearranged from the perspective of Q_1. For $j > 1$, the equations are the same as before Eq.(4.131):

$$\dot{Q}_j = \frac{\partial K}{\partial \alpha_j} = 0 \quad \Rightarrow \quad Q_j = \beta_j = \frac{\partial W}{\partial \alpha_j} \tag{4.135}$$

In this language we do not need to mention S only W. There are a few comments to be made:

1. Again, the choice of α is arbitrary, and $\alpha_i = \alpha_i(\gamma)$ is fine. If we do replace $\alpha_1 = \alpha_1(\gamma)$ then $\dot{Q}_i = \frac{\partial K}{\partial \gamma_i} = v_i$ is a constant so that (potentially) all of the Q_i become linear in time:
$$Q_i = v_i t + \beta_i \text{ for all } i \in \{1, \ldots, n\} \tag{4.136}$$
2. What is W? We know that:
$$\dot{W} = \frac{\partial W}{\partial q_i}\dot{q}_i = p_i \dot{q}_i \Rightarrow W = \int p_i \dot{q}_i \, dt = \int p_i \, dq_i, \tag{4.137}$$
which is a different sort of "action".
3. The time independent H-J equation has some similarity to the time-independent Schrödinger energy eigenvalue equation (both involve H and and constant E, but the former is a non-linear equation for W, while the latter is a linear equation for the wavefunction ψ).

The H-J method is most useful when there is a separation of variables in H.

Example: if
$$H = h_1(q_1, q_2, p_1, p_2) + h_2(q_1, q_2, p_1, p_2) f(q_3, p_3) = \alpha_1, \tag{4.138}$$
so that q_3 is separable, then
$$f(q_3, p_3) = \frac{\alpha_1 - h_1}{h_2} \tag{4.139}$$
is a constant because the right hand side is independent of q_3 and p_3. Thus we assign
$$f(q_3, p_3) = \alpha_2 \tag{4.140}$$
for convenience. We can then write:
$$W = W'(q_1, q_2, \alpha) + W_3(q_3, \alpha) \quad \Rightarrow \quad f\left(q_3, \frac{\partial W_3}{\partial q_3}\right) = \alpha_2 \quad \text{and} \tag{4.141}$$
$$h_1\left(q_1, q_2, \frac{\partial W'}{\partial q_1}, \frac{\partial W'}{\partial q_2}\right) + \alpha_2 h_2\left(q_1, q_2, \frac{\partial W'}{\partial q_1}, \frac{\partial W'}{\partial q_2}\right) = \alpha_1 \tag{4.142}$$

Here, q_1 and q_2 may or may not be separable.

If all variables are separable then we use the solution:
$$W = \sum_i W_i(q_i, \alpha) \tag{4.143}$$

We can simply try a solution of this form to test for separability.

Note that cyclic coordinates are always separable.

Proof: let us say that q_1 is cyclic. Then

$$p_1 \equiv \gamma \quad \text{and} \quad H\left(q_2, \dots, q_n, \frac{\partial W}{\partial q_1}, \frac{\partial W}{\partial q_2}, \dots, \frac{\partial W}{\partial q_n}\right) = \alpha_1, \tag{4.144}$$

where γ is constant. Let us now write

$$W(q, \alpha) = W_1(q_1, \alpha) + W'(q_2, \dots, q_n, \alpha). \tag{4.145}$$

This gives us:

$$p_1 = \frac{\partial W_1}{\partial q_1} = \gamma \Rightarrow W_1 = \gamma q_1. \tag{4.146}$$

Which gives us:

$$W(q, \alpha) = \gamma q_1 + W'(q_2, \dots, q_n, \alpha) \tag{4.147}$$

This procedure can be repeated for any remaining cyclic variables.

Note that the choice of variables is often important in finding a result that separates. A problem with spherical symmetry may separate in spherical coordinates but not Cartesian coordinates.

4.5 Kepler Problem

As an extended example, let us consider the Kepler problem of two masses m_1 and m_2 in a central potential (with the CM coordinate $\mathbf{R} = 0$). The Lagrangian is:

$$L = \frac{1}{2} m \dot{\mathbf{r}}^2 - V(r) \quad \text{where} \quad \frac{1}{m} \equiv \frac{1}{m_1} + \frac{1}{m_2}, \tag{4.148}$$

and here m is the reduced mass. *Any* $V(r)$ conserves $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, so the motion of $\mathbf{r}$ and $\mathbf{p}$ is in a plane perpendicular to $\mathbf{L}$. The coordinates in the plane can be taken as (r, ψ), so:

$$L = \frac{m}{2}\left(\dot{r}^2 + r^2 \dot{\psi}^2\right) - V(r), \tag{4.149}$$

with ψ being cyclic, which implies:

$$p_\psi = m r^2 \dot{\psi} \text{ is a constant.} \tag{4.150}$$

In fact $p_\psi = |\mathbf{L}| \equiv \ell$. Notationally, we use ℓ for the magnitude of the angular momentum $\mathbf{L}$ to distinguish it from the Lagrangian L.

The energy is then:

$$E = \frac{m}{2}\dot{r}^2 + \frac{\ell^2}{2mr^2} + V(r), \tag{4.151}$$

which is constant, and this can be rewritten as:

$$E = \frac{m}{2}\dot{r}^2 + V_{\text{eff}}(r) \quad \text{where} \quad V_{\text{eff}}(r) = V(r) + \frac{\ell^2}{2mr^2}, \tag{4.152}$$

where V_{eff} is the effective potential, as plotted below for the gravitational potential.

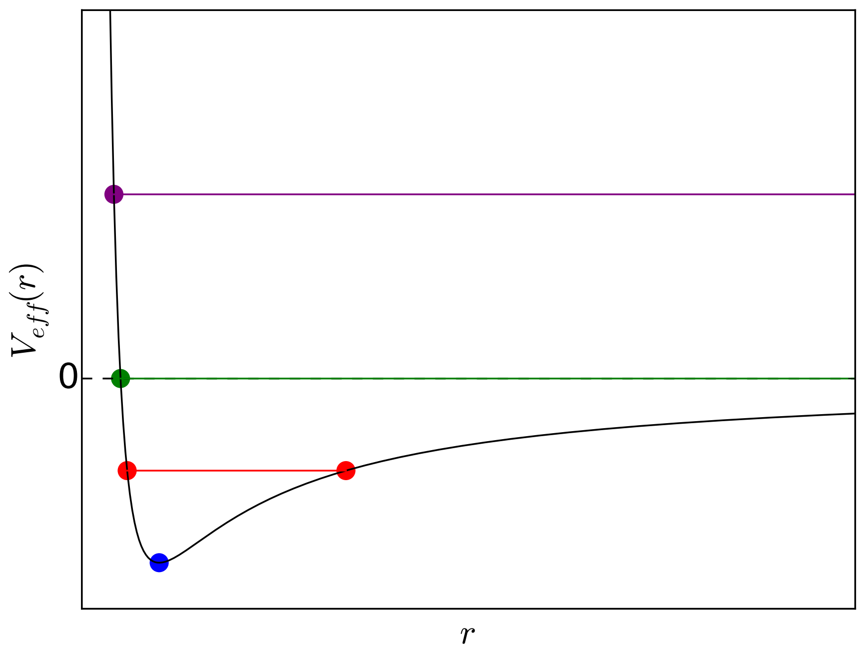


Figure 4.1: Plot of the effective potential V_{eff} along with the different qualitative orbits allowed in a gravity-like potential. The points correspond to turning points of the orbit.

Writing the E-L equation for $\dot{r} = dr/dt = \ldots$ and then solving for it as $dt = dr/(\ldots)$, and integrating yields

$$t = t(r) = \int_{r_0}^{r} \frac{dr'}{\sqrt{\frac{2}{m}\left(E - V(r') - \frac{\ell^2}{2mr'^2}\right)}} \tag{4.153}$$

as an implicit solution to the radial motion.

The orbital motion comes as $r = r(\psi)$ or $\psi = \psi(r)$ by using Eq.(4.150) and substituting, in Eq.(4.153). We have $\dot{\psi} = d\psi/dt = \ell/(mr^2)$, so we can use this to replace dt by $d\psi$ in $dt = dr/(\ldots)$ to get an equation of the form $d\psi = dr/(\ldots)$. The result is given by

$$\psi - \psi_0 = \ell \int_{r_0}^{r} \frac{dr'}{r'^2\sqrt{2m\left(E - V(r') - \frac{\ell^2}{2mr'^2}\right)}} \tag{4.154}$$

In the particular case of $V(r) = -\frac{k}{r}$, the solution of the orbital equation is:

$$\frac{1}{r(\psi)} = \frac{mk}{\ell^2}\left(1 + \varepsilon\cos(\psi - \psi')\right) \tag{4.155}$$

where the eccentricity ε is given by:

$$\varepsilon \equiv \sqrt{1 + \frac{2E\ell^2}{mk^2}} \tag{4.156}$$

Below are plotted the different qualitative orbits for the gravitic potential, with different ε or E (circular, elliptical, parabolic, and hyperbolic respectively).

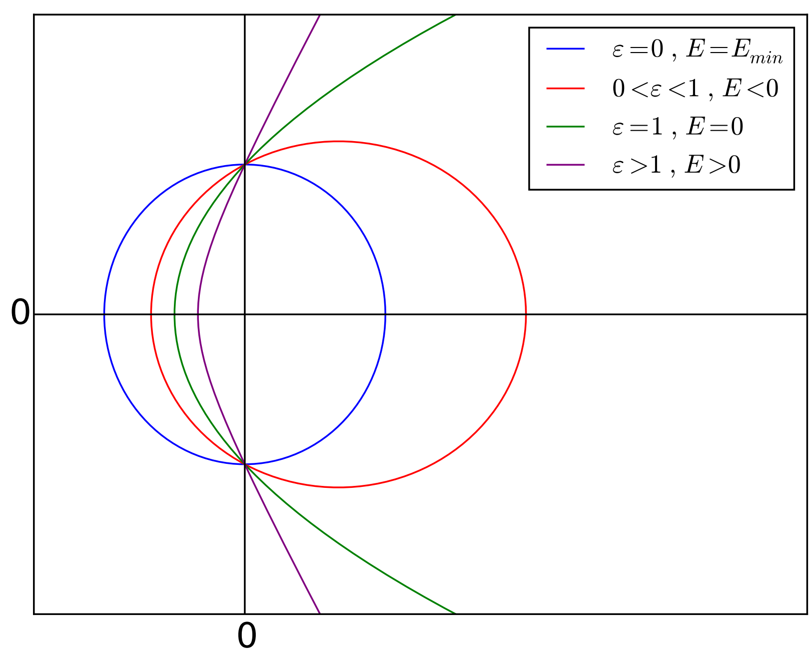


Figure 4.2: Different Orbits for the gravity-like potential. The orbits' colors match those of Fig.(4.1). The unbounded orbits occur for $E \geq 0$. The different curves correspond to the different possible conic sections.

Consider solving this problem instead by the H-J method. Lets start by considering as the variables (r, ψ) so that we assume that the motion of the orbit is in a plane. Here

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\psi^2}{r^2}\right) + V(r) = \alpha_1 = E\,. \tag{4.157}$$

As ψ is cyclic, then $p_\psi \equiv \alpha_\psi$ is constant. Using:

$$W = W_1(r) + \alpha_\psi \psi, \tag{4.158}$$

then the time independent H-J equation is:

$$\frac{1}{2m}\left(\left(\frac{\partial W_1}{\partial r}\right)^2 + \frac{\alpha_\psi^2}{r^2}\right) + V(r) = \alpha_1. \tag{4.159}$$

This is simplified to

$$\frac{\partial W_1}{\partial r} = \sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\psi^2}{r^2}} \tag{4.160}$$

and solved by

$$W = \alpha_\psi \psi + \int \sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\psi^2}{r^2}}\, dr. \tag{4.161}$$

The transformation equations are:

$$\begin{aligned} t + \beta_1 &= \frac{\partial W}{\partial \alpha_1} = m \int \frac{dr}{\sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\psi^2}{r^2}}}, \\ \beta_2 &= \frac{\partial W}{\partial \alpha_\psi} = \psi - \alpha_\psi \int \frac{dr}{r^2\sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\psi^2}{r^2}}}. \end{aligned} \tag{4.162}$$

Thus we immediately get the radial equation $t = t(r)$ and orbital equation $\psi = \psi(r)$ from this, with $\alpha_\psi = \ell$ and $\alpha_1 = E$, showing that the constants are physically relevant parameters.

Let's solve this problem again, but suppose the motion is in 3 dimensions (as if we did not know the plane of the orbit). Using spherical coordinates (r, θ, φ) this corresponds to

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2\theta}\right) + V(r) = \alpha_1\,. \tag{4.163}$$

Lets try a separable solution

$$W = W_r(r) + W_\theta(\theta) + W_\varphi(\varphi)\,. \tag{4.164}$$

Since φ is cyclic we know it is separable and that:

$$W_\varphi(\varphi) = \alpha_\varphi\, \varphi\,. \tag{4.165}$$

Together, this leaves;

$$\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{1}{r^2}\left(\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{\alpha_\varphi^2}{\sin^2(\theta)}\right) + 2mV(r) = 2m\alpha_1\,. \tag{4.166}$$

Because the term:

$$\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{\alpha_\varphi^2}{\sin^2(\theta)} \tag{4.167}$$

only depends on θ while the rest of the equation depends on r, it must be a constant so we can say:

$$\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{\alpha_\varphi^2}{\sin^2(\theta)} \equiv \alpha_\theta^2 \tag{4.168}$$

and the separation works. This then gives:

$$\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{\alpha_\theta^2}{r^2} = 2m(\alpha_1 - V(r)), \tag{4.169}$$

which is the same equation we considered before when assuming the motion was in a plane, with $\alpha_1 = E$ and $\alpha_\theta = \ell$. Eq. (4.168) says that

$$p_\theta^2 + \frac{p_\varphi^2}{\sin^2(\theta)} = \ell^2 \,. \tag{4.170}$$

Here p_φ is the constant angular momentum about the $\hat{z}$ axis.

4.6 Action-Angle Variables

For many problems, we may not be able to solve analytically for the exact motion or for orbital equations, but we can still characterize the motion. For *periodic* systems we can find the frequency by exploiting *action-angle variables*.

The simplest case is for a single dimension of canonical coordinates (q, p). If $H(q, p) = \alpha_1$ then $p = p(q, \alpha_1)$. There are two types of periodic motion to consider.

1. Libration (oscillation) is characterized by a closed phase space orbit, so that q and p evolve periodically with the same frequency.

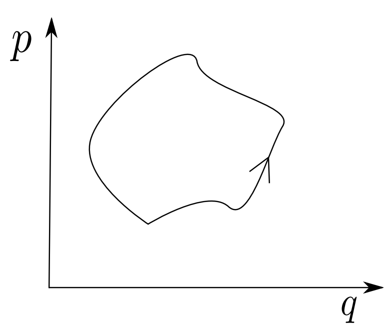


Figure 4.3: Phase space orbit of a libration (oscilation). The trajectory closes on itself, the state returns to the same position after some time τ.

2. Rotation is characterized by an open phase space path, so p is periodic while q evolves without bound.

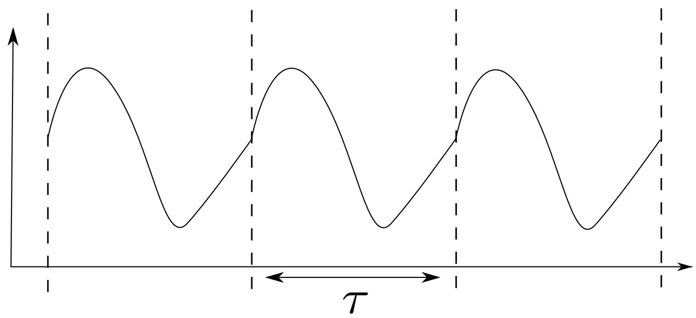


Figure 4.4: Phase space orbit of a rotation. Although the orbit is not closed, each period the evolution of the system is the same, leading to a orbit that repeats itself with a translation.

Example: a pendulum of length a may be characterized by canonical coordinates (θ, p_θ), where:

$$E = H = \frac{p_\theta^2}{2ma^2} - mga\cos\theta \tag{4.171}$$

This means:

$$p_\theta = \pm\sqrt{2ma^2(E + mga\cos\theta)} \tag{4.172}$$

must be real. A rotation occurs when $E > mga$, and oscillations occur when $E < mga$. The critical point in between (when the pendulum just makes it to the top) is when $E = mga$ exactly, and is depicted by a dashed line in the figure below.

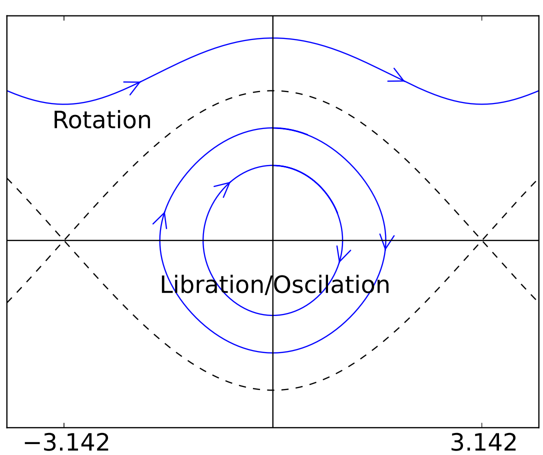


Figure 4.5: The pendulum exhibits both librations and rotations depending on the initial conditions.

For either type of periodic motion, it is useful to replace $P = \alpha_1$ by a different (constant) choice called the *action variable*

$$J = \oint p\, dq, \tag{4.173}$$

where $\oint$ refers to a definite integral over one period in phase space. To see that J is constant, recall that $p = p(q, \alpha_1)$, so plugging this into the definite integral we are left with $J = J(\alpha_1)$. Also, we have the inverse, $\alpha_1 = H = H(J)$, and can rewrite Hamilton's characteristic function in terms of J by $W = W(q, \alpha_1) = W(q, H(J)) = W(q, J)$ (where again the argument notation is abused slightly).

The coordinate conjugate to J is the *angle variable*

$$\omega = \frac{\partial W}{\partial J} \tag{4.174}$$

(where ω is not meant to imply an angular velocity). This means

$$\dot{\omega} = \frac{\partial H(J)}{\partial J} = \nu(J) \text{ is a constant.} \tag{4.175}$$

As a result the angle variable has linear time dependence,

$$\omega = \nu t + \beta\,, \tag{4.176}$$

for some initial condition β. Dimensionally, J has units of angular momentum, while ω has no dimensions (like an angle or a phase).

To see why it is useful to use the canonical variables (ω, J), let us consider the change in ω when q goes through a complete cycle.

$$\Delta\omega = \oint \frac{\partial \omega}{\partial q}\, dq = \oint \frac{\partial^2 W}{\partial q \partial J}\, dq = \frac{\partial}{\partial J} \oint \frac{\partial W}{\partial q}\, dq = \frac{\partial}{\partial J} \oint p\, dq = 1 \tag{4.177}$$

where in the last equality we used the definition of J in Eq.(4.173). Also, we have $\Delta\omega = \nu\tau$ where τ is the period. Thus

$$\nu = \frac{1}{\tau} \tag{4.178}$$

is the *frequency* of periodic motion. If we find $H = H(J)$ then

$$\nu = \frac{\partial H(J)}{\partial J} \tag{4.179}$$

immediately gives the frequency $\nu = \nu(J)$ for the system. Often, we then $J = J(E)$ to get $\nu = \nu(E)$ the frequency at a given energy. This is a very efficient way of finding the frequency of the motion without solving for extraneous information.

Example: let us consider a pendulum with action-angle variables. We define:

$$\tilde{E} = \frac{E}{mga} \tag{4.180}$$

so that $\tilde{E} > 1$ corresponds to rotation and $\tilde{E} < 1$ corresponds to oscillation. This means

$$p_\theta = \pm\sqrt{2m^2ga^3}\sqrt{\tilde{E} + \cos\theta}. \tag{4.181}$$

For $\tilde{E} > 1$:

$$J = \sqrt{2m^2ga^3}\int_{-\pi}^{\pi} d\theta\,\sqrt{\tilde{E} + \cos\theta}\,, \tag{4.182}$$

which is an elliptic integral. For $\tilde{E} < 1$:

$$\begin{aligned} J &= \sqrt{2m^2ga^3}\int_{-\theta_0}^{+\theta_0} d\theta\,\sqrt{\tilde{E} + \cos\theta} + \sqrt{2m^2ga^3}\int_{\theta_0}^{-\theta_0} d\theta\,\Big[-\sqrt{\tilde{E} + \cos\theta}\,\Big] \\ &= 4\sqrt{2m^2ga^3}\int_{0}^{\theta_0} d\theta\,\sqrt{\tilde{E} + \cos\theta}\,, \end{aligned} \tag{4.183}$$

as the contributions from the four intervals that the pendulum swings through in one period are all equivalent. Here θ_0 is the turning point of the oscillation, and $\tilde{E} = -\cos(\theta_0)$.

From this:

$$\nu = \frac{\partial E}{\partial J} = \left(\frac{\partial J}{\partial E}\right)^{-1} \tag{4.184}$$

which we can solve graphically by making a plot of J vs E, then dJ/dE versus E, and finally the inverse $\nu = dE/dJ$ versus E.

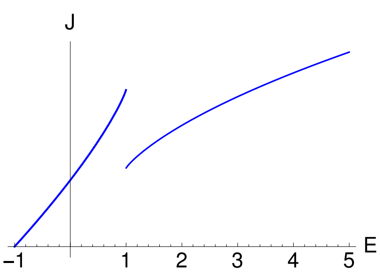


Figure 4.6: Plot of $J(E)$ versus $\tilde{E}$. The discontinuity corresponds to the transition from Oscilation to Rotation.

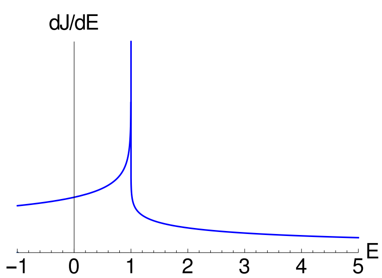


Figure 4.7: Plot of $\dfrac{dJ}{dE}$. The discontinuity is logarithmic divergent so it is integrable.

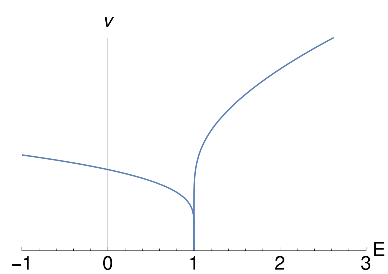


Figure 4.8: Plot of the frequency of oscillation $\nu(E)$ versus $\tilde{E}$. As $\tilde{E} \to -1$ we approach the small amplitude limit, where $\nu = (2\pi)^{-1}\sqrt{g/a}$.

Example: let us consider the limit $|\theta| \ll 1$ of a pendulum, so:

$$H = \frac{p_\theta^2}{2ma^2} + \frac{mga}{2}\theta^2 - mga \tag{4.185}$$

We can actually consider this in the context of a general harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{m\Omega^2}{2}x^2, \tag{4.186}$$

where:

$$\Omega = \sqrt{\frac{g}{a}} \quad , \quad x = a\theta \quad \text{and} \quad p = \frac{p_\theta}{a} \tag{4.187}$$

Notationally, Ω is used for the harmonic oscillator frequency to distinguish from the transformed angle variable ω. We then have:

$$J = \oint p\, dq = \oint \pm\sqrt{2mE - m^2\Omega^2 x^2}\, dx \tag{4.188}$$

Note that the coordinate does not need to be an angle, as may be the case for general x. This gives:

$$J = 4\sqrt{2mE}\int_0^{x_0} \sqrt{1 - \frac{x^2}{x_0^2}}\, dx \quad \text{where} \quad x_0 \equiv \frac{\sqrt{2mE}}{m\Omega}. \tag{4.189}$$

Solving the integral yields

$$J = \pi\sqrt{2mE}x_0 = \frac{2\pi mE}{m\Omega} = \frac{2\pi E}{\Omega}, \tag{4.190}$$

which gives us

$$\frac{\partial E}{\partial J} = \frac{\Omega}{2\pi}, \tag{4.191}$$

the expected cyclic frequency for the harmonic oscillator.

Multiple Variables: We can treat multiple variables $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ with the action-angle formalism if *each* pair (q_i, p_i) has an oscillatory or rotating orbit. Lets also assume that the H-J equation is completely separable into:

$$W = \sum_j W_j(q_j, \alpha). \tag{4.192}$$

Here we have

$$p_i = \frac{\partial W_i}{\partial q_i} = p_i(q_i, \alpha_1, \ldots, \alpha_n) \Rightarrow J_i = \oint p_i\, dq_i = J_i(\alpha_1, \ldots, \alpha_n) \tag{4.193}$$

where repeated indices do not correspond to implicit sums here. This implies that the inverse will be $\alpha_j = \alpha_j(J_1, \ldots, J_n)$ and thus $\alpha_1 = H = H(J_1, \ldots, J_n)$. Likewise:

$$\omega_i = \frac{\partial W}{\partial J_i} = \sum_j \frac{\partial W_j}{\partial J_i} = \omega_i(q_1, \ldots, q_n, J_1, \ldots, J_n)\,. \tag{4.194}$$

Just as in the one dimensional case the time derivative of the angle variables is a constant

$$\dot{\omega}_i = \frac{\partial H}{\partial J_i} = \nu_i(J_1, \dots, J_n) \tag{4.195}$$

which are the frequencies describing motion in this "multi-periodic" system. Due to the presence of multiple frequencies, the motion through the whole $2n$-dimensional phase space need not be periodic in time.

Example: in the 2-dimensional harmonic oscillator:

$$x = A\cos(2\pi\nu_1 t) \text{ and } y = B\cos(2\pi\nu_2 t) \tag{4.196}$$

$$p_x = m\dot{x} \text{ and } p_y = m\dot{y} \tag{4.197}$$

The overall motion is not periodic in time unless $\frac{\nu_1}{\nu_2}$ is a rational number.

Kepler Problem Example:

Let us do a more extended and detailed example. Returning to the Kepler problem:

$$V(r) = -\frac{k}{r} \tag{4.198}$$

with its separable W:

$$W = W_r(r,\alpha) + W_\theta(\theta,\alpha) + W_\varphi(\varphi,\alpha)\,. \tag{4.199}$$

If we take $E < 0$, we have oscillation in r and θ, along with a rotation in φ. In particular from solving our earlier differential equations for W_θ and W_r, we have

$$\begin{aligned} W_\varphi &= \alpha_\varphi \varphi \\ W_\theta &= \pm\int \sqrt{\alpha_\theta^2 - \frac{\alpha_\varphi^2}{\sin^2(\theta)}}\, d\theta \\ W_r &= \pm\sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\theta^2}{r^2}}\, dr \end{aligned}$$

Here we have

$$J_\varphi = \oint p_\varphi\, d\varphi = \oint \frac{\partial W}{\partial \varphi}\, d\varphi = \oint \alpha_\varphi\, d\varphi \tag{4.200}$$

For the cyclic variable φ, we still call the constant p_φ periodic and will take the period to be 2π (arbitrarily since any period would work), which corresponds to particle returning to the original point in space. Thus

$$J_\varphi = 2\pi\alpha_\varphi, \tag{4.201}$$

where α_φ is the angular momentum about $\hat{z}$.

Continuing, in a similar manner we have

$$J_\theta = \oint p_\theta\, d\theta = \oint \frac{\partial W}{\partial \theta}\, d\theta = \oint \pm\sqrt{\alpha_\theta^2 - \frac{\alpha_\varphi^2}{\sin^2(\theta)}}\, d\theta \tag{4.202}$$

Let us call:

$$\cos(\gamma) \equiv \frac{\alpha_\varphi}{\alpha_\theta} \tag{4.203}$$

the angular momentum fraction. Then:

$$J_\theta = \alpha_\theta \oint \pm\sqrt{1 - \frac{\cos^2(\gamma)}{\sin^2(\theta)}}\, d\theta \tag{4.204}$$

If we let $\sin(\theta_0) = \cos(\gamma)$, then $p_\theta = 0$ at the turning points, $\theta \in \{\theta_0, \pi - \theta_0\}$, as expected.

Here one oscillator goes from $\pi - \theta_0 \to \theta_0$ when $p_\theta > 0$, and in reverse for $p_\theta < 0$. Moreover, $\sin(\theta)^{-2}$ is even about $\theta = \frac{\pi}{2}$. This gives

$$J_\theta = 4\alpha_\theta \int_{\frac{\pi}{2}}^{\theta_0} \sqrt{1 - \frac{\cos^2(\gamma)}{\sin^2(\theta)}}\, d\theta. \tag{4.205}$$

Making two more substitutions

$$\cos(\theta) \equiv \sin(\gamma)\sin(\psi)\,, \text{ and then } u \equiv \tan(\psi)\,, \tag{4.206}$$

after some work the expression becomes

$$\begin{aligned} J_\theta &= 4\alpha_\theta \int_0^\infty \left(\frac{1}{1+u^2} - \frac{\cos^2(\gamma)}{1+u^2\cos^2(\gamma)}\right) du = 2\pi\alpha_\theta(1-\cos(\gamma)) \\ &= 2\pi(\alpha_\theta - \alpha_\varphi), \end{aligned} \tag{4.207}$$

which gives

$$J_\theta + J_\varphi = 2\pi\alpha_\theta. \tag{4.208}$$

Finally we can consider

$$J_r = \oint \sqrt{2mE - 2mV(r) - \frac{(J_\theta + J_\varphi)^2}{4\pi^2 r^2}}\, dr \tag{4.209}$$

We can immediately make some observations. We observe that $J_r = J_r(E, J_\theta + J_\varphi)$ is a function of two variables for any $V = V(r)$, and thus if we invert $E = E(J_r, J_\theta + J_\varphi)$. This implies:

$$\frac{\partial E}{\partial J_\theta} = \frac{\partial E}{\partial J_\varphi} \quad \Rightarrow \quad \nu_\theta = \nu_\varphi \tag{4.210}$$

The two frequencies are degenerate for any $V = V(r)$.

For the $V(r) = -kr^{-1}$ potential, the integration can be performed (for example, by contour integration) to give (for $E < 0$)

$$J_r = -(J_\theta + J_\varphi) + \pi k\sqrt{\frac{2m}{-E}} \Rightarrow J_r + J_\theta + J_\varphi = \pi k\sqrt{\frac{2m}{-E}}\,. \tag{4.211}$$

This means:

$$E = -\frac{2\pi^2 k^2 m}{(J_r + J_\theta + J_\varphi)^2} \quad \Rightarrow \quad \nu_\theta = \nu_\varphi = \nu_r \tag{4.212}$$

In particular:

$$\nu_r = \frac{\partial E}{\partial J_r} = 4\pi^2 k^2 (J_r + J_\theta + J_\varphi)^{-3} = \frac{1}{\pi k}\sqrt{\frac{-2E^3}{m}} \tag{4.213}$$

which is the correct orbital frequency in a bound Kepler orbit.

Using the relations between $\{\alpha_1 = E, \alpha_\theta, \alpha_\varphi\}$ and $\{J_r, J_\theta, J_\varphi\}$, we can also get Hamilton's characteristic function for this system as

$$\begin{aligned} W &= W_\varphi + W_\theta + W_r \\ &= \frac{\varphi J_\varphi}{2\pi} \pm \int \sqrt{(J_\theta + J_\varphi)^2 - \frac{J_\varphi^2}{\sin^2(\theta)}}\,\frac{d\theta}{2\pi} \pm \int \sqrt{\frac{-(2\pi m k)^2}{(J_r + J_\theta + J_\varphi)^2} + \frac{2mk}{r} - \frac{(J_\theta + J_\varphi)^2}{(2\pi r)^2}}\,dr. \end{aligned}$$

This then gives the angle variables:

$$\begin{aligned} \omega_r &= \frac{\partial W}{\partial J_r} = \omega_r(r, J_r + J_\theta + J_\varphi, J_\theta + J_\varphi) \\ \omega_\theta &= \frac{\partial W}{\partial J_\theta} = \omega_\theta(r, \theta, J_r + J_\theta + J_\varphi, J_\theta + J_\varphi, J_\varphi) \\ \omega_\varphi &= \frac{\partial W}{\partial J_\varphi} = \omega_\varphi(r, \theta, \varphi, J_r + J_\theta + J_\varphi, J_\theta + J_\varphi, J_\varphi) \end{aligned} \tag{4.214}$$

where $\dot\omega_r = \nu_r$, $\dot\omega_\theta = \nu_\theta$, and $\dot\omega_\varphi = \nu_\varphi$. Of course, in this case, $\nu_r = \nu_\theta = \nu_\varphi$.

At this point we can identify five constants for the Kepler problem

$$\begin{aligned} J_1 &= J_\varphi \\ J_2 &= J_\theta + J_\varphi \\ J_3 &= J_r + J_\theta + J_\varphi \\ \omega_1 &= \omega_\varphi - \omega_\theta \\ \omega_2 &= \omega_\theta - \omega_r\,. \end{aligned} \tag{4.215}$$

(These 5 constants could also be identified from the angular momentum $\vec{L}$, energy E, and Laplace-Runge-Lenz vector $\vec{A}$.) What are they? There are two constants specifying the plane of the orbit (the $x'y'$-plane), which are the inclination angle i and the longitude of the ascending node Ω. There are three constants specifying the form of the ellipse, which are the semi-major axis a (giving the size), the eccentricity ε (giving the shape), and the angle ω (giving the orientation within the plane). These are all shown in Fig. 4.9.

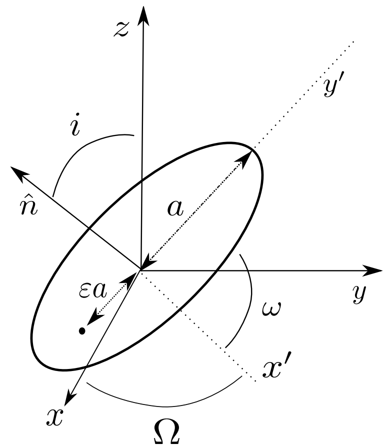


Figure 4.9: Picture of an orbit in 3d and the five parameters necessary to fully specify it. The angles i, Ω and ω provide the orientation in space while a and ε provide the size and shape of the conic section.

It can be shown that the relations between these constants and the ones above are

$$\cos(i) = \frac{J_1}{J_2} \qquad a = -\frac{k}{2E} = \frac{J_3^2}{4\pi^2 mk} \qquad \epsilon = \sqrt{1 - \left(\frac{J_2}{J_3}\right)^2}$$
$$\Omega = 2\pi\omega_1 \qquad \omega = 2\pi\omega_2$$

providing a fairly simple physical interpretations to $(J_1, J_2, J_3, \omega_1, \omega_2)$. Also recall that $J_2 = 2\pi\alpha_\theta = 2\pi\ell$.

When the orbit is perturbed by additional forces (like a moon, another planet, general relativistic corrections, and so on), these action-angle variables provide a natural way to describe the modified orbit. We will soon see that they become functions of time. For example, from the above list of constants, the perturbed $\omega = \omega(t)$ is the precession of the perihelion of an orbit. We will learn how to derive this time dependence next, in our chapter on Perturbation Theory.