Chapter 3

Hamiltonian Mechanics

3.1 The Hamiltonian

Recall that $L = L(q, \dot{q}, t)$, and

$$p_\sigma = \frac{\partial L}{\partial \dot{q}_\sigma} \ . \tag{3.1}$$

The Hamiltonian, $H(q,p)$ is obtained by a Legendre transformation,

$$H(q,p) = \sum_{\sigma=1}^{n} p_\sigma \, \dot{q}_\sigma - L \ . \tag{3.2}$$

Note that

$$\begin{aligned} dH &= \sum_{\sigma=1}^{n} \bigg(p_\sigma \, d\dot{q}_\sigma + \dot{q}_\sigma \, dp_\sigma - \frac{\partial L}{\partial q_\sigma} \, dq_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma} \, d\dot{q}_\sigma \bigg) - \frac{\partial L}{\partial t} \, dt \\ &= \sum_{\sigma=1}^{n} \bigg(\dot{q}_\sigma \, dp_\sigma - \frac{\partial L}{\partial q_\sigma} \, dq_\sigma \bigg) - \frac{\partial L}{\partial t} \, dt \ . \end{aligned} \tag{3.3}$$

Thus, we obtain Hamilton's equations of motion,

$$\frac{\partial H}{\partial p_\sigma} = \dot{q}_\sigma \quad , \quad \frac{\partial H}{\partial q_\sigma} = -\frac{\partial L}{\partial q_\sigma} = -\dot{p}_\sigma \tag{3.4}$$

and

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \ . \tag{3.5}$$

Some remarks:

- As an example, consider a particle moving in three dimensions, described by spherical polar coordinates (r, θ, ϕ). Then

$$L = \tfrac{1}{2} m \left(\dot{r}^2 + r^2 \, \dot{\theta}^2 + r^2 \sin^2\!\theta \, \dot{\phi}^2 \right) - U(r, \theta, \phi) \ . \tag{3.6}$$

  We have

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \quad , \quad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2 \, \dot{\theta} \quad , \quad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2 \, \sin^2\!\theta \, \dot{\phi} \ , \tag{3.7}$$

and thus

$$H = p_r\,\dot r + p_\theta\,\dot\theta + p_\phi\,\dot\phi - L$$

$$= \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2\sin^2\!\theta} + U(r,\theta,\phi)\ . \tag{3.8}$$

Note that H is time-independent, hence $\frac{\partial H}{\partial t} = \frac{dH}{dt} = 0$, and therefore H is a constant of the motion.

- In order to obtain $H(q,p)$ we must invert the relation $p_\sigma = \frac{\partial L}{\partial \dot q_\sigma} = p_\sigma(q,\dot q)$ to obtain $\dot q_\sigma(q,p)$. This is possible if the Hessian,

$$\frac{\partial p_\alpha}{\partial \dot q_\beta} = \frac{\partial^2 L}{\partial \dot q_\alpha\,\partial \dot q_\beta} \tag{3.9}$$

is nonsingular. This is the content of the 'inverse function theorem' of multivariable calculus.

- Define the rank $2n$ vector, ξ, by its components,

$$\xi_i = \begin{cases} q_i & \text{if } 1 \le i \le n \\ p_{i-n} & \text{if } n \le i \le 2n\ . \end{cases} \tag{3.10}$$

Then we may write Hamilton's equations compactly as

$$\dot\xi_i = J_{ij}\,\frac{\partial H}{\partial \xi_j}\ , \tag{3.11}$$

where

$$J = \begin{pmatrix} \mathbf{0}_{n\times n} & \mathbf{1}_{n\times n} \\ -\mathbf{1}_{n\times n} & \mathbf{0}_{n\times n} \end{pmatrix} \tag{3.12}$$

is a rank $2n$ matrix. Note that $J^{\rm t} = -J$, *i.e.* J is antisymmetric, and that $J^2 = -\mathbf{1}_{2n\times 2n}$. We shall utilize this 'symplectic structure' to Hamilton's equations shortly.

3.2 Modified Hamilton's Principle

We have that

$$0 = \delta\int\limits_{t_a}^{t_b}\!dt\,L = \delta\int\limits_{t_a}^{t_b}\!dt\,\big(p_\sigma\,\dot q_\sigma - H\big) \tag{3.13}$$

$$= \int\limits_{t_a}^{t_b}\!dt\,\bigg\{p_\sigma\,\delta\dot q_\sigma + \dot q_\sigma\,\delta p_\sigma - \frac{\partial H}{\partial q_\sigma}\,\delta q_\sigma - \frac{\partial H}{\partial p_\sigma}\,\delta p_\sigma\bigg\}$$

$$= \int\limits_{t_a}^{t_b}\!dt\,\Bigg\{-\bigg(\dot p_\sigma + \frac{\partial H}{\partial q_\sigma}\bigg)\,\delta q_\sigma + \bigg(\dot q_\sigma - \frac{\partial H}{\partial p_\sigma}\bigg)\,\delta p_\sigma\Bigg\} + \big(p_\sigma\,\delta q_\sigma\big)\Big|_{t_a}^{t_b}\ ,$$

assuming $\delta q_\sigma(t_a) = \delta q_\sigma(t_b) = 0$. Setting the coefficients of δq_σ and δp_σ to zero, we recover Hamilton's equations.

3.3 Phase Flow is Incompressible

A flow for which $\boldsymbol{\nabla}\cdot\boldsymbol{v}=0$ is *incompressible* – we shall see why in a moment. Let's check that the divergence of the phase space velocity does indeed vanish:

$$
\begin{aligned}
\boldsymbol{\nabla}\cdot\dot{\boldsymbol{\xi}} &= \sum_{\sigma=1}^{n}\left\{\frac{\partial\dot{q}_\sigma}{\partial q_\sigma}+\frac{\partial\dot{p}_\sigma}{\partial p_\sigma}\right\} \\
&= \sum_{i=1}^{2n}\frac{\partial\dot{\xi}_i}{\partial\xi_i}=\sum_{i,j}J_{ij}\,\frac{\partial^2 H}{\partial\xi_i\,\partial\xi_j}=0\ .
\end{aligned}
\tag{3.14}
$$

Now let $\rho(\boldsymbol{\xi},t)$ be a distribution on phase space. Continuity implies

$$
\frac{\partial\rho}{\partial t}+\boldsymbol{\nabla}\cdot(\rho\,\dot{\boldsymbol{\xi}})=0\ . \tag{3.15}
$$

Invoking $\boldsymbol{\nabla}\cdot\dot{\boldsymbol{\xi}}=0$, we have that

$$
\frac{D\rho}{Dt}=\frac{\partial\rho}{\partial t}+\dot{\boldsymbol{\xi}}\cdot\boldsymbol{\nabla}\rho=0\ , \tag{3.16}
$$

where $D\rho/Dt$ is sometimes called the *convective derivative* – it is the total derivative of the function $\rho\big(\boldsymbol{\xi}(t),t\big)$, evaluated at a point $\xi(t)$ in phase space which moves according to the dynamics. This says that the density in the "comoving frame" is locally constant.

3.4 Poincaré Recurrence Theorem

Let g_τ be the 'τ-advance mapping' which evolves points in phase space according to Hamilton's equations

$$
\dot{q}_i=+\frac{\partial H}{\partial p_i}\qquad,\qquad \dot{p}_i=-\frac{\partial H}{\partial q_i} \tag{3.17}
$$

for a time interval $\varDelta t=\tau$. Consider a region $\varOmega$ in phase space. Define $g_\tau^n\varOmega$ to be the $n^{\rm th}$ image of $\varOmega$ under the mapping g_τ. Clearly g_τ is invertible; the inverse is obtained by integrating the equations of motion backward in time. We denote the inverse of g_τ by g_τ^{-1}. By Liouville's theorem, g_τ is volume preserving when acting on regions in phase space, since the evolution of any given point is Hamiltonian. This follows from the continuity equation for the phase space density,

$$
\frac{\partial\varrho}{\partial t}+\nabla\cdot(\boldsymbol{u}\varrho)=0 \tag{3.18}
$$

where $\boldsymbol{u}=\{\dot{\boldsymbol{q}},\dot{\boldsymbol{p}}\}$ is the velocity vector in phase space, and Hamilton's equations, which say that the phase flow is incompressible, *i.e.* $\nabla\cdot\boldsymbol{u}=0$:

$$
\begin{aligned}
\nabla\cdot\boldsymbol{u} &= \sum_{i=1}^{n}\left\{\frac{\partial\dot{q}_i}{\partial q_i}+\frac{\partial\dot{p}_i}{\partial p_i}\right\} \\
&= \sum_{i=1}^{n}\left\{\frac{\partial}{\partial q_i}\bigg(\frac{\partial H}{\partial p_i}\bigg)+\frac{\partial}{\partial p_i}\bigg(-\frac{\partial H}{\partial q_i}\bigg)\right\}=0\ .
\end{aligned}
\tag{3.19}
$$

Thus, we have that the convective derivative vanishes, *viz.*

$$\frac{D\varrho}{Dt} \equiv \frac{\partial\varrho}{\partial t} + \boldsymbol{u}\cdot\nabla\varrho = 0 \ , \tag{3.20}$$

which guarantees that the density remains constant in a frame moving with the flow.

The proof of the recurrence theorem is simple. Assume that g_τ is invertible and volume-preserving, as is the case for Hamiltonian flow. Further assume that phase space volume is finite. Since the energy is preserved in the case of time-independent Hamiltonians, we simply ask that the volume of phase space *at fixed total energy* E be finite, *i.e.*

$$\int d\mu\, \delta\big(E - H(\boldsymbol{q},\boldsymbol{p})\big) < \infty \ , \tag{3.21}$$

where $d\mu = d\boldsymbol{q}\, d\boldsymbol{p}$ is the phase space uniform integration measure.

Theorem: In any finite neighborhood Ω of phase space there exists a point $\boldsymbol{\varphi}_0$ which will return to Ω after n applications of g_τ, where n is finite.

Proof: Assume the theorem fails; we will show this assumption results in a contradiction. Consider the set Υ formed from the union of all sets $g_\tau^m\,\Omega$ for all m:

$$\Upsilon = \bigcup_{m=0}^{\infty} g_\tau^m\,\Omega \tag{3.22}$$

We assume that the set $\{g_\tau^m\,\Omega \,|\, m \in \mathbf{Z}\ , m \geq 0\}$ is disjoint. The volume of a union of disjoint sets is the sum of the individual volumes. Thus,

$$\begin{aligned}\text{vol}(\Upsilon) &= \sum_{m=0}^{\infty} \text{vol}(g_\tau^m\,\Omega) \\ &= \text{vol}(\Omega)\cdot\sum_{m=1}^{\infty} 1 = \infty \ ,\end{aligned} \tag{3.23}$$

since $\text{vol}(g_\tau^m\,\Omega) = \text{vol}(\Omega)$ from volume preservation. But clearly Υ is a subset of the entire phase space, hence we have a contradiction, because by assumption phase space is of finite volume.

Thus, the assumption that the set $\{g_\tau^m\,\Omega \,|\, m \in \mathbf{Z}\ , m \geq 0\}$ is disjoint fails. This means that there exists some pair of integers k and l, with $k \neq l$, such that $g_\tau^k\,\Omega \cap g_\tau^l\,\Omega \neq \emptyset$. Without loss of generality we may assume $k > l$. Apply the inverse g_τ^{-1} to this relation l times to get $g_\tau^{k-l}\,\Omega \cap \Omega \neq \emptyset$. Now choose any point $\boldsymbol{\varphi} \in g_\tau^n\,\Omega \cap \Omega$, where $n = k - l$, and define $\boldsymbol{\varphi}_0 = g_\tau^{-n}\boldsymbol{\varphi}$. Then by construction both $\boldsymbol{\varphi}_0$ and $g_\tau^n\,\boldsymbol{\varphi}_0$ lie within Ω and the theorem is proven.

Each of the two central assumptions – invertibility and volume preservation – is crucial. Without either of them, the proof fails. Consider, for example, a volume-preserving map which is not invertible. An example might be a mapping $f: \mathrm{R} \to \mathrm{R}$ which takes any real number to its fractional part. Thus, $f(\pi) = 0.14159265\ldots$. Let us restrict our attention to intervals of width less than unity. Clearly f is then volume preserving. The action of f on the interval $[2,3)$ is to map it to the interval $[0,1)$. But $[0,1)$ remains fixed under the action of f, so no point within the interval $[2, 3)$ will ever return under repeated iterations of f. Thus, f does not exhibit Poincaré recurrence.

Consider next the case of the damped harmonic oscillator. In this case, phase space volumes contract. For a one-dimensional oscillator obeying $\ddot{x} + 2\beta\dot{x} + \Omega_0^2 x = 0$ one has $\nabla \cdot \boldsymbol{u} = -2\beta < 0$ ($\beta > 0$ for damping). Thus the convective derivative obeys $D_t \varrho = -(\nabla \cdot \boldsymbol{u})\varrho = +2\beta\varrho$ which says that the density increases exponentially in the comoving frame, as $\varrho(t) = e^{2\beta t}\,\varrho(0)$. Thus, phase space volumes collapse, and are not preserved by the dynamics. In this case, it is possible for the set Υ to be of finite volume, even if it is the union of an infinite number of sets $g_\tau^n\,\Omega$, because the volumes of these component sets themselves decrease exponentially, as $\text{vol}(g_\tau^n\,\Omega) = e^{-2n\beta\tau}\,\text{vol}(\Omega)$. A damped pendulum, released from rest at some small angle θ_0, will not return arbitrarily close to these initial conditions.

3.5 Kac Ring Model

The implications of the Poincaré recurrence theorem are surprising – even shocking. If one takes a bottle of perfume in a sealed, evacuated room and opens it, the perfume molecules will diffuse throughout the room. The recurrence theorem guarantees that after some finite time T all the molecules will go back inside the bottle (and arbitrarily close to their initial velocities as well). The hitch is that this could take a very long time, *e.g.* much much longer than the age of the Universe.

On less absurd time scales, we know that most systems come to thermodynamic equilibrium. But how can a system both exhibit equilibration *and* Poincaré recurrence? The two concepts seem utterly incompatible!

A beautifully simple model due to Kac shows how a recurrent system can exhibit the phenomenon of equilibration. Consider a ring with N sites. On each site, place a 'spin' which can be in one of two states: up or down. Along the N links of the system, F of

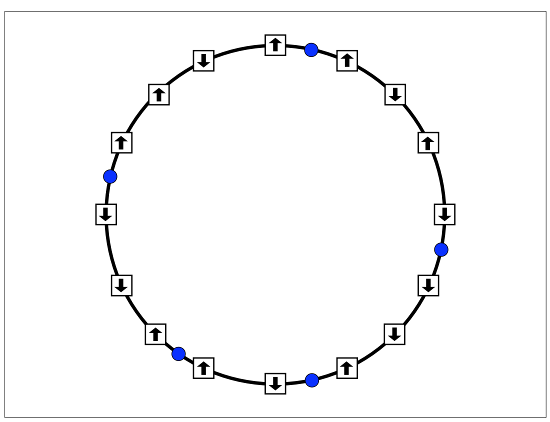

Figure 3.1: A configuration of the Kac ring with $N = 16$ sites and $F = 4$ flippers. The flippers, which live on the links, are represented by blue dots.

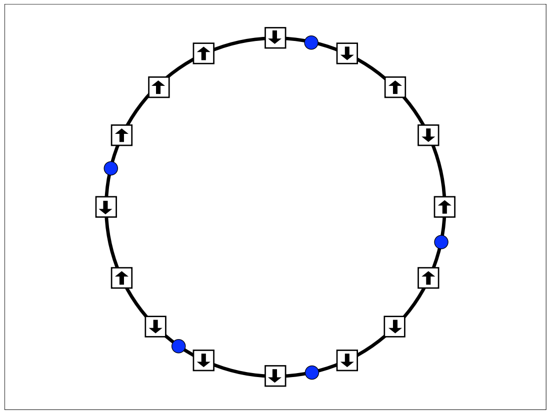

Figure 3.2: The ring system after one time step. Evolution proceeds by clockwise rotation. Spins passing through flippers are flipped.

them contain 'flippers'. The configuration of the flippers is set at the outset and never changes. The dynamics of the system are as follows: during each time step, every spin moves clockwise a distance of one lattice spacing. Spins which pass through flippers reverse their orientation: up becomes down, and down becomes up.

The 'phase space' for this system consists of 2^N discrete configurations. Since each configuration maps onto a unique image under the evolution of the system, phase space 'volume' is preserved. The evolution is invertible; the inverse is obtained simply by rotating the spins counterclockwise. Figures 3.1 and 3.2 depict an example configuration for the system, and its first iteration under the dynamics.

Suppose the flippers were not fixed, but moved about randomly. In this case, we could focus on a single spin and determine its configuration probabilistically. Let p_n be the probability that a given spin is in the up configuration at time n. The probability that it is up at time $(n+1)$ is then

$$p_{n+1} = (1-x)\, p_n + x\,(1-p_n) \ , \tag{3.24}$$

where $x = F/N$ is the fraction of flippers in the system. In words: a spin will be up at time $(n+1)$ if it was up at time n and did not pass through a flipper, or if it was down at time n and did pas through a flipper. If the flipper locations are randomized at each time step, then the probability of flipping is simply $x = F/N$. Equation 3.24 can be solved immediately:

$$p_n = \tfrac{1}{2} + (1-2x)^n\,(p_0 - \tfrac{1}{2}) \ , \tag{3.25}$$

which decays exponentially to the equilibrium value of $p_{\rm eq} = \frac{1}{2}$ with time scale $\tau = -1/\ln|1-2x|$. If we define the magnetization $m \equiv (N_\uparrow - N_\downarrow)/N$, then $m = 2p-1$, so $m_n = (1-2x)^n\, m_0$. The equilibrium magnetization is $m_{\rm eq} = 0$. Note that for $\frac{1}{2} < x < 1$ that the magnetization reverses sign each time step, as well as decreasing exponentially in magnitude.

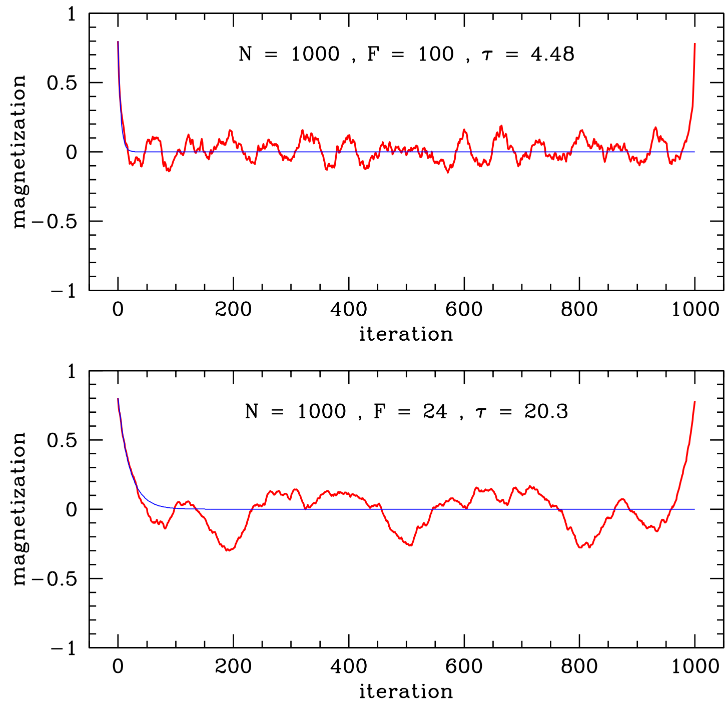

Figure 3.3: Two simulations of the Kac ring model, each with $N = 1000$ sites and with $F = 100$ flippers (top panel) and $F = 24$ flippers (bottom panel). The red line shows the magnetization as a function of time, starting from an initial configuration in which 90% of the spins are up. The blue line shows the prediction of the *Stosszahlansatz*, which yields an exponentially decaying magnetization with time constant τ.

The assumption that leads to equation 3.24 is called the *Stosszahlansatz*. The resulting dynamics are irreversible: the magnetization inexorably decays to zero. However, the Kac ring model is purely deterministic, and the *Stosszahlansatz* can at best be an approximation to the true dynamics. Clearly the *Stosszahlansatz* fails to account for correlations such as the following: if spin i is flipped at time n, then spin $i+1$ will have been flipped at time $n-1$. Indeed, since the dynamics of the Kac ring model are invertible and volume preserving, it must exhibit Poincaré recurrence.

The model is trivial to simulate. The results of such a simulation are shown in figure 3.3 for a ring of $N = 1000$ sites, with $F = 100$ and $F = 24$ flippers. Note how the magnetization decays and fluctuates about the equilibrium value $_{\mathrm{eq}} = 0$, but that after N iterations m recovers its initial value: $m_N = m_0$. The recurrence time for this system is simply N if F is even, and $2N$ if F is odd, since every spin will then have flipped an even number of times.

In figure 3.4 we plot two other simulations. The top panel shows what happens when $x > \frac{1}{2}$, so that the magnetization wants to reverse its sign with every iteration. The bottom panel shows a simulation for a larger ring, with $N = 25000$ sites. Note that the fluctuations in m about equilibrium are smaller than in the cases with $N = 1000$ sites. Why?

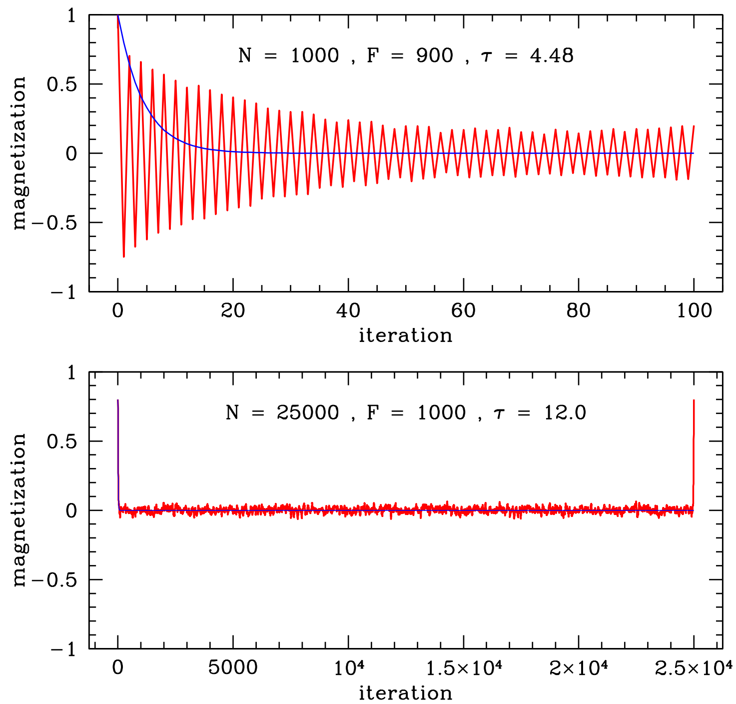

Figure 3.4: Simulations of the Kac ring model. Top: $N = 1000$ sites with $F = 900$ flippers. The flipper density $x = F/N$ is greater than $\frac{1}{2}$, so the magnetization reverses sign every time step. Only 100 iterations are shown, and the blue curve depicts the absolute value of the magnetization within the *Stosszahlansatz*. Bottom: $N = 25,000$ sites with $F = 1000$ flippers. Note that the fluctuations about the 'equilibrium' magnetization $m = 0$ are much smaller than in the $N = 1000$ site simulations.

3.6 Poisson Brackets

The time evolution of any function $F(q,p)$ over phase space is given by

$$\begin{aligned}\frac{d}{dt}\,F\big(q(t),p(t),t\big) &= \frac{\partial F}{\partial t} + \sum_{\sigma=1}^{n}\left\{\frac{\partial F}{\partial q_\sigma}\,\dot q_\sigma + \frac{\partial F}{\partial p_\sigma}\,\dot p_\sigma\right\} \\ &\equiv \frac{\partial F}{\partial t} + \big\{F,H\big\}\ , \end{aligned} \tag{3.26}$$

where the *Poisson bracket* $\{\cdot\,,\cdot\}$ is given by

$$\big\{A,B\big\} \equiv \sum_{\sigma=1}^{n}\left(\frac{\partial A}{\partial q_\sigma}\,\frac{\partial B}{\partial p_\sigma} - \frac{\partial A}{\partial p_\sigma}\,\frac{\partial B}{\partial q_\sigma}\right) \tag{3.27}$$

$$= \sum_{i,j=1}^{2n} J_{ij}\,\frac{\partial A}{\partial \xi_i}\,\frac{\partial B}{\partial \xi_j}\ . \tag{3.28}$$

Properties of the Poisson bracket:

- Antisymmetry:
$$\big\{f,g\big\} = -\big\{g,f\big\} \ . \tag{3.29}$$

- Bilinearity: if λ is a constant, and f, g, and h are functions on phase space, then
$$\big\{f+\lambda\, g,h\big\} = \big\{f,h\big\} + \lambda\big\{g,h\big\} \ . \tag{3.30}$$
Linearity in the second argument follows from this and the antisymmetry condition.

- Associativity:
$$\big\{fg,h\big\} = f\big\{g,h\big\} + g\big\{f,h\big\} \ . \tag{3.31}$$

- Jacobi identity:
$$\big\{f,\{g,h\}\big\} + \big\{g,\{h,f\}\big\} + \big\{h,\{f,g\}\big\} = 0 \ . \tag{3.32}$$

Some other useful properties:

- If $\{A,H\}=0$ and $\frac{\partial A}{\partial t}=0$, then $\frac{dA}{dt}=0$, *i.e.* $A(q,p)$ is a constant of the motion.
- If $\{A,H\}=0$ and $\{B,H\}=0$, then $\big\{\{A,B\},H\big\}=0$. If in addition A and B have no explicit time dependence, we conclude that $\{A,B\}$ is a constant of the motion.
- It is easily established that
$$\{q_\alpha,q_\beta\}=0 \quad , \quad \{p_\alpha,p_\beta\}=0 \quad , \quad \{q_\alpha,p_\beta\}=\delta_{\alpha\beta} \ . \tag{3.33}$$

3.7 Canonical Transformations

3.7.1 Point Transformations in Lagrangian Mechanics

In Lagrangian mechanics, we are free to redefine our generalized coordinates, *viz.*
$$Q_\sigma = Q_\sigma(q_1,\ldots,q_n,t) \ . \tag{3.34}$$
This is called a "point transformation." The transformation is invertible if
$$\det\left(\frac{\partial Q_\alpha}{\partial q_\beta}\right) \neq 0 \ . \tag{3.35}$$
The transformed Lagrangian, $\tilde L$, written as a function of the new coordinates Q and velocities $\dot Q$, is
$$\tilde L\big(Q,\dot Q,t\big) = L\big(q(Q,t),\dot q(Q,\dot Q,t)\big) \ . \tag{3.36}$$
Finally, Hamilton's principle,
$$\delta\int\limits_{t_1}^{t_b} dt\,\tilde L(Q,\dot Q,t)=0 \tag{3.37}$$

with $\delta Q_\sigma(t_a) = \delta Q_\sigma(t_b) = 0$, still holds, and the form of the Euler-Lagrange equations remains unchanged:

$$\frac{\partial \tilde{L}}{\partial Q_\sigma} - \frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{Q}_\sigma}\right) = 0 \ . \tag{3.38}$$

The invariance of the equations of motion under a point transformation may be verified explicitly. We first evaluate

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{Q}_\sigma}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\alpha}\,\frac{\partial \dot{q}_\alpha}{\partial \dot{Q}_\sigma}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\alpha}\,\frac{\partial q_\alpha}{\partial Q_\sigma}\right) \ , \tag{3.39}$$

where the relation

$$\frac{\partial \dot{q}_\alpha}{\partial \dot{Q}_\sigma} = \frac{\partial q_\alpha}{\partial Q_\sigma} \tag{3.40}$$

follows from

$$\dot{q}_\alpha = \frac{\partial q_\alpha}{\partial Q_\sigma}\,\dot{Q}_\sigma + \frac{\partial q_\alpha}{\partial t} \ . \tag{3.41}$$

Now we compute

$$\begin{aligned}
\frac{\partial \tilde{L}}{\partial Q_\sigma} &= \frac{\partial L}{\partial q_\alpha}\,\frac{\partial q_\alpha}{\partial Q_\sigma} + \frac{\partial L}{\partial \dot{q}_\alpha}\,\frac{\partial \dot{q}_\alpha}{\partial Q_\sigma} \\
&= \frac{\partial L}{\partial q_\alpha}\,\frac{\partial q_\alpha}{\partial Q_\sigma} + \frac{\partial L}{\partial \dot{q}_\alpha}\left(\frac{\partial^2 q_\alpha}{\partial Q_\sigma\,\partial Q_{\sigma'}}\,\dot{Q}_{\sigma'} + \frac{\partial^2 q_\alpha}{\partial Q_\sigma\,\partial t}\right) \\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right)\frac{\partial q_\alpha}{\partial Q_\sigma} + \frac{\partial L}{\partial \dot{q}_\alpha}\,\frac{d}{dt}\left(\frac{\partial q_\alpha}{\partial Q_\sigma}\right) \\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\,\frac{\partial q_\alpha}{\partial Q_\sigma}\right) = \frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{Q}_\sigma}\right) \ ,
\end{aligned} \tag{3.42}$$

where the last equality is what we obtained earlier in eqn. 3.39.

3.7.2 Canonical Transformations in Hamiltonian Mechanics

In Hamiltonian mechanics, we will deal with a much broader class of transformations – ones which mix all the $q's$ and $p's$. The general form for a canonical transformation (CT) is

$$q_\sigma = q_\sigma\big(Q_1, \ldots, Q_n; P_1, \ldots, P_n; t\big) \tag{3.43}$$

$$p_\sigma = p_\sigma\big(Q_1, \ldots, Q_n; P_1, \ldots, P_n; t\big) \ , \tag{3.44}$$

with $\sigma \in \{1, \ldots, n\}$. We may also write

$$\xi_i = \xi_i\big(\Xi_1, \ldots, \Xi_{2n}; t\big) \ , \tag{3.45}$$

with $i \in \{1, \ldots, 2n\}$. The transformed Hamiltonian is $\tilde{H}(Q, P, t)$.

What sorts of transformations are allowed? Well, if Hamilton's equations are to remain invariant, then

$$\dot{Q}_\sigma = \frac{\partial \tilde{H}}{\partial P_\sigma} \quad , \quad \dot{P}_\sigma = -\frac{\partial \tilde{H}}{\partial Q_\sigma} \ , \tag{3.46}$$

which gives

$$\frac{\partial \dot{Q}_\sigma}{\partial Q_\sigma} + \frac{\partial \dot{P}_\sigma}{\partial P_\sigma} = 0 = \frac{\partial \dot{\Xi}_i}{\partial \Xi_i} \ . \tag{3.47}$$

I.e. the flow remains incompressible in the new (Q, P) variables. We will also require that phase space volumes are preserved by the transformation, *i.e.*

$$\det\left(\frac{\partial \Xi_i}{\partial \xi_j}\right) = \left|\left|\frac{\partial(Q,P)}{\partial(q,p)}\right|\right| = 1 \ . \tag{3.48}$$

Additional conditions will be discussed below.

3.7.3 Hamiltonian Evolution

Hamiltonian evolution itself defines a canonical transformation. Let $\xi_i = \xi_i(t)$ and $\xi_i' = \xi_i(t+dt)$. Then from the dynamics $\dot{\xi}_i = J_{ij} \frac{\partial H}{\partial \xi_j}$, we have

$$\xi_i(t+dt) = \xi_i(t) + J_{ij} \frac{\partial H}{\partial \xi_j} \, dt + \mathcal{O}\big(dt^2\big) \ . \tag{3.49}$$

Thus,

$$\begin{aligned} \frac{\partial \xi_i'}{\partial \xi_j} &= \frac{\partial}{\partial \xi_j}\left(\xi_i + J_{ik} \frac{\partial H}{\partial \xi_k} \, dt + \mathcal{O}\big(dt^2\big)\right) \\ &= \delta_{ij} + J_{ik} \frac{\partial^2 H}{\partial \xi_j \, \partial \xi_k} \, dt + \mathcal{O}\big(dt^2\big) \ . \end{aligned} \tag{3.50}$$

Now, using the result

$$\det\big(1 + \epsilon M\big) = 1 + \epsilon \, \mathrm{Tr}\, M + \mathcal{O}(\epsilon^2) \ , \tag{3.51}$$

we have

$$\left|\left|\frac{\partial \xi_i'}{\partial \xi_j}\right|\right| = 1 + J_{jk} \frac{\partial^2 H}{\partial \xi_j \, \partial \xi_k} \, dt + \mathcal{O}\big(dt^2\big) \tag{3.52}$$

$$= 1 + \mathcal{O}\big(dt^2\big) \ . \tag{3.53}$$

3.7.4 Symplectic Structure

We have that

$$\dot{\xi}_i = J_{ij} \frac{\partial H}{\partial \xi_j} \ . \tag{3.54}$$

Suppose we make a time-independent canonical transformation to new phase space coordinates, $\Xi_a = \Xi_a(\xi)$. We then have

$$\dot{\Xi}_a = \frac{\partial \Xi_a}{\partial \xi_j} \, \dot{\xi}_j = \frac{\partial \Xi_a}{\partial \xi_j} \, J_{jk} \frac{\partial H}{\partial \xi_k} \ . \tag{3.55}$$

But if the transformation is canonical, then the equations of motion are preserved, and we also have

$$\dot{\Xi}_a = J_{ab}\,\frac{\partial\tilde{H}}{\partial\Xi_b} = J_{ab}\,\frac{\partial\xi_k}{\partial\Xi_b}\,\frac{\partial H}{\partial\xi_k}\ . \tag{3.56}$$

Equating these two expressions, we have

$$M_{aj}\,J_{jk}\,\frac{\partial H}{\partial\xi_k} = J_{ab}\,M^{-1}_{kb}\,\frac{\partial H}{\partial\xi_k}\ , \tag{3.57}$$

where

$$M_{aj} \equiv \frac{\partial\Xi_a}{\partial\xi_j} \tag{3.58}$$

is the Jacobian of the transformation. Since the equality must hold for all ξ, we conclude

$$MJ = J\big(M^{\rm t}\big)^{-1} \quad\Longrightarrow\quad MJM^{\rm t} = J\ . \tag{3.59}$$

A matrix M satisfying $MM^{\rm t} = \mathbf{1}$ is of course an *orthogonal* matrix. A matrix M satisfying $MJM^{\rm t} = J$ is called *symplectic*. We write $M \in {\rm Sp}(2n)$, *i.e.* M is an element of the group of *symplectic matrices*[1] of rank $2n$.

The symplectic property of M guarantees that the Poisson brackets are preserved under a canonical transformation:

$$\begin{aligned}
\big\{A,B\big\}_\xi &= J_{ij}\,\frac{\partial A}{\partial\xi_i}\,\frac{\partial B}{\partial\xi_j}\\
&= J_{ij}\,\frac{\partial A}{\partial\Xi_a}\,\frac{\partial\Xi_a}{\partial\xi_i}\,\frac{\partial B}{\partial\Xi_b}\,\frac{\partial\Xi_b}{\partial\xi_j}\\
&= \big(M_{ai}\,J_{ij}\,M^{\rm t}_{jb}\big)\,\frac{\partial A}{\partial\Xi_a}\,\frac{\partial B}{\partial\Xi_b}\\
&= J_{ab}\,\frac{\partial A}{\partial\Xi_a}\,\frac{\partial B}{\partial\Xi_b}\\
&= \big\{A,B\big\}_\Xi\ .
\end{aligned} \tag{3.60}$$

3.7.5 Generating Functions for Canonical Transformations

For a transformation to be canonical, we require

$$\delta\int\limits_{t_a}^{t_b}\!dt\,\Big\{p_\sigma\,\dot{q}_\sigma - H(q,p,t)\Big\} = 0 = \delta\int\limits_{t_a}^{t_b}\!dt\,\Big\{P_\sigma\,\dot{Q}_\sigma - \tilde{H}(Q,P,t)\Big\}\ . \tag{3.61}$$

This is satisfied provided

$$\Big\{p_\sigma\,\dot{q}_\sigma - H(q,p,t)\Big\} = \lambda\,\bigg\{P_\sigma\,\dot{Q}_\sigma - \tilde{H}(Q,P,t) + \frac{dF}{dt}\bigg\}\ , \tag{3.62}$$

[1] Note that the rank of a symplectic matrix is always even. Note also $MJM^{\rm t} = J$ implies $M^{\rm t}JM = J$.

where λ is a constant. For canonical transformations, $\lambda = 1$.[2] Thus,

$$\tilde{H}(Q,P,t) = H(q,p,t) + P_\sigma\,\dot{Q}_\sigma - p_\sigma\,\dot{q}_\sigma + \frac{\partial F}{\partial q_\sigma}\,\dot{q}_\sigma + \frac{\partial F}{\partial Q_\sigma}\,\dot{Q}_\sigma + \frac{\partial F}{\partial p_\sigma}\,\dot{p}_\sigma + \frac{\partial F}{\partial P_\sigma}\,\dot{P}_\sigma + \frac{\partial F}{\partial t}\ . \qquad (3.63)$$

Thus, we require

$$\frac{\partial F}{\partial q_\sigma} = p_\sigma \quad , \quad \frac{\partial F}{\partial Q_\sigma} = -P_\sigma \quad , \quad \frac{\partial F}{\partial p_\sigma} = 0 \quad , \quad \frac{\partial F}{\partial P_\sigma} = 0\ . \qquad (3.64)$$

The transformed Hamiltonian is

$$\tilde{H}(Q,P,t) = H(q,p,t) + \frac{\partial F}{\partial t}\ . \qquad (3.65)$$

There are four possibilities, corresponding to the freedom to make Legendre transformations with respect to each of the arguments of $F(q,Q)$:

$$F(q,Q,t) = \begin{cases} F_1(q,Q,t) & ; \quad p_\sigma = +\frac{\partial F_1}{\partial q_\sigma} \quad , \quad P_\sigma = -\frac{\partial F_1}{\partial Q_\sigma} \quad \text{(type I)} \\[2ex] F_2(q,P,t) - P_\sigma\,Q_\sigma & ; \quad p_\sigma = +\frac{\partial F_2}{\partial q_\sigma} \quad , \quad Q_\sigma = +\frac{\partial F_2}{\partial P_\sigma} \quad \text{(type II)} \\[2ex] F_3(p,Q,t) + p_\sigma\,q_\sigma & ; \quad q_\sigma = -\frac{\partial F_3}{\partial p_\sigma} \quad , \quad P_\sigma = -\frac{\partial F_3}{\partial Q_\sigma} \quad \text{(type III)} \\[2ex] F_4(p,P,t) + p_\sigma\,q_\sigma - P_\sigma\,Q_\sigma & ; \quad q_\sigma = -\frac{\partial F_4}{\partial p_\sigma} \quad , \quad Q_\sigma = +\frac{\partial F_4}{\partial P_\sigma} \quad \text{(type IV)} \end{cases}$$

In each case ($\gamma = 1,2,3,4$), we have

$$\tilde{H}(Q,P,t) = H(q,p,t) + \frac{\partial F_\gamma}{\partial t}\ . \qquad (3.66)$$

Let's work out some examples:

- Consider the type-II transformation generated by

  $$F_2(q,P) = A_\sigma(q)\,P_\sigma\ , \qquad (3.67)$$

  where $A_\sigma(q)$ is an arbitrary function of the $\{q_\sigma\}$. We then have

  $$Q_\sigma = \frac{\partial F_2}{\partial P_\sigma} = A_\sigma(q) \qquad , \qquad p_\sigma = \frac{\partial F_2}{\partial q_\sigma} = \frac{\partial A_\alpha}{\partial q_\sigma}\,P_\alpha\ . \qquad (3.68)$$

  Thus,

  $$Q_\sigma = A_\sigma(q) \qquad , \qquad P_\sigma = \frac{\partial q_\alpha}{\partial Q_\sigma}\,p_\alpha\ . \qquad (3.69)$$

[2] Solutions of eqn. 3.62 with $\lambda \neq 1$ are known as *extended* canonical transformations. We can always rescale coordinates and/or momenta to achieve $\lambda = 1$.

This is a general point transformation of the kind discussed in eqn. 3.34. For a general linear point transformation, $Q_\alpha = M_{\alpha\beta}\, q_\beta$, we have $P_\alpha = p_\beta\, M^{-1}_{\beta\alpha}$, *i.e.* $Q = Mq$, $P = p\, M^{-1}$. If $M_{\alpha\beta} = \delta_{\alpha\beta}$, this is the identity transformation. $F_2 = q_1 P_3 + q_3 P_1$ interchanges labels 1 and 3, *etc.*

- Consider the type-I transformation generated by

$$F_1(q,Q) = A_\sigma(q)\, Q_\sigma \ . \tag{3.70}$$

We then have

$$p_\sigma = \frac{\partial F_1}{\partial q_\sigma} = \frac{\partial A_\alpha}{\partial q_\sigma}\, Q_\alpha \tag{3.71}$$

$$P_\sigma = -\frac{\partial F_1}{\partial Q_\sigma} = -A_\sigma(q) \ . \tag{3.72}$$

Note that $A_\sigma(q) = q_\sigma$ generates the transformation

$$\begin{pmatrix} q \\ p \end{pmatrix} \longrightarrow \begin{pmatrix} -P \\ +Q \end{pmatrix} \ . \tag{3.73}$$

- A mixed transformation is also permitted. For example,

$$F(q,Q) = q_1\, Q_1 + (q_3 - Q_2)\, P_2 + (q_2 - Q_3)\, P_3 \tag{3.74}$$

is of type-I with respect to index $\sigma = 1$ and type-II with respect to indices $\sigma = 2, 3$. The transformation effected is

$$Q_1 = p_1 \qquad\qquad Q_2 = q_3 \qquad\qquad Q_3 = q_2 \tag{3.75}$$

$$P_1 = -q_1 \qquad\qquad P_2 = p_3 \qquad\qquad P_3 = p_2 \ . \tag{3.76}$$

- Consider the harmonic oscillator,

$$H(q,p) = \frac{p^2}{2m} + \tfrac{1}{2} k q^2 \ . \tag{3.77}$$

If we could find a time-independent canonical transformation such that

$$p = \sqrt{2mf(P)}\, \cos Q \qquad , \qquad q = \sqrt{\frac{2\, f(P)}{k}}\, \sin Q \ , \tag{3.78}$$

where $f(P)$ is some function of P, then we'd have $\tilde{H}(Q,P) = f(P)$, which is cyclic in Q. To find this transformation, we take the ratio of p and q to obtain

$$p = \sqrt{mk}\, q\, \operatorname{ctn} Q \ , \tag{3.79}$$

which suggests the type-I transformation

$$F_1(q,Q) = \tfrac{1}{2}\sqrt{mk}\, q^2\, \operatorname{ctn} Q \ . \tag{3.80}$$

This leads to

$$p = \frac{\partial F_1}{\partial q} = \sqrt{mk}\, q \,\text{ctn}\, Q \quad , \quad P = -\frac{\partial F_1}{\partial Q} = \frac{\sqrt{mk}\, q^2}{2\,\sin^2\! Q} \ . \tag{3.81}$$

Thus,

$$q = \frac{\sqrt{2P}}{\sqrt[4]{mk}} \sin Q \quad \Longrightarrow \quad f(P) = \sqrt{\frac{k}{m}}\, P = \omega P \ , \tag{3.82}$$

where $\omega = \sqrt{k/m}$ is the oscillation frequency. We therefore have

$$\tilde{H}(Q,P) = \omega P \ , \tag{3.83}$$

whence $P = E/\omega$. The equations of motion are

$$\dot{P} = -\frac{\partial \tilde{H}}{\partial Q} = 0 \quad , \quad \dot{Q} = \frac{\partial \tilde{H}}{\partial P} = \omega \ , \tag{3.84}$$

which yields

$$Q(t) = \omega t + \varphi_0 \quad , \quad q(t) = \sqrt{\frac{2E}{m\omega^2}}\, \sin\big(\omega t + \varphi_0\big) \ . \tag{3.85}$$

3.8 Hamilton-Jacobi Theory

We've stressed the great freedom involved in making canonical transformations. Coordinates and momenta, for example, may be interchanged – the distinction between them is purely a matter of convention! We now ask: is there any specially preferred canonical transformation? In this regard, one obvious goal is to make the Hamiltonian $\tilde{H}(Q,P,t)$ and the corresponding equations of motion as simple as possible.

Recall the general form of the canonical transformation:

$$\tilde{H}(Q,P) = H(q,p) + \frac{\partial F}{\partial t} \ , \tag{3.86}$$

with

$$\frac{\partial F}{\partial q_\sigma} = p_\sigma \qquad\qquad \frac{\partial F}{\partial p_\sigma} = 0 \tag{3.87}$$

$$\frac{\partial F}{\partial Q_\sigma} = -P_\sigma \qquad\qquad \frac{\partial F}{\partial P_\sigma} = 0 \ . \tag{3.88}$$

We now demand that this transformation result in the simplest Hamiltonian possible, that is, $\tilde{H}(Q,P,t) = 0$. This requires we find a function F such that

$$\frac{\partial F}{\partial t} = -H \quad , \quad \frac{\partial F}{\partial q_\sigma} = p_\sigma \ . \tag{3.89}$$

The remaining functional dependence may be taken to be either on Q (type I) or on P (type II). As it turns out, the generating function F we seek is in fact the action, S, which is the integral of L with respect to time, expressed as a function of its endpoint values.

3.8.1 The Action as a Function of Coordinates and Time

We have seen how the action $S[\eta(\tau)]$ is a *functional* of the path $\eta(\tau)$ and a *function* of the endpoint values $\{q_a, t_a\}$ and $\{q_b, t_b\}$. Let us define the action *function* $S(q,t)$ as

$$S(q,t) = \int_{t_a}^{t} d\tau\, L\big(\eta, \dot\eta, \tau\big) \ , \tag{3.90}$$

where $\eta(\tau)$ starts at (q_a, t_a) and ends at (q,t). We also require that $\eta(\tau)$ satisfy the Euler-Lagrange equations,

$$\frac{\partial L}{\partial \eta_\sigma} - \frac{d}{d\tau}\bigg(\frac{\partial L}{\partial \dot\eta_\sigma}\bigg) = 0 \tag{3.91}$$

Let us now consider a new path, $\tilde\eta(\tau)$, also starting at (q_a, t_a), but ending at $(q+dq, t+dt)$, and also satisfying the equations of motion. The differential of S is

$$\begin{aligned} dS &= S\big[\tilde\eta(\tau)\big] - S\big[\eta(\tau)\big] \\ &= \int_{t_a}^{t+dt} d\tau\, L(\tilde\eta, \dot{\tilde\eta}, \tau) - \int_{t_a}^{t} d\tau\, L\big(\eta, \dot\eta, \tau\big) \qquad (3.92)\\ &= \int_{t_a}^{t} d\tau \left\{ \frac{\partial L}{\partial \eta_\sigma}\Big[\tilde\eta_\sigma(\tau) - \eta_\sigma(\tau)\Big] + \frac{\partial L}{\partial \dot\eta_\sigma}\Big[\dot{\tilde\eta}_\sigma(\tau) - \dot\eta_\sigma(\tau)\Big] \right\} + L\big(\tilde\eta(t), \dot{\tilde\eta}(t), t\big)\, dt \\ &= \int_{t_a}^{t} d\tau \left\{ \frac{\partial L}{\partial \eta_\sigma} - \frac{d}{d\tau}\bigg(\frac{\partial L}{\partial \dot\eta_\sigma}\bigg) \right\} \Big[\tilde\eta_\sigma(\tau) - \eta_\sigma(\tau)\Big] \\ &\qquad\qquad + \frac{\partial L}{\partial \dot\eta_\sigma}\bigg|_t \Big[\tilde\eta_\sigma(t) - \eta_\sigma(t)\Big] + L\big(\tilde\eta(t), \dot{\tilde\eta}(t), t\big)\, dt \\ &= 0 + \pi_\sigma(t)\, \delta\eta_\sigma(t) + L\big(\eta(t), \dot\eta(t), t\big)\, dt + \mathcal{O}(\delta q \cdot dt) \ , \qquad (3.93) \end{aligned}$$

where we have defined

$$\pi_\sigma = \frac{\partial L}{\partial \dot\eta_\sigma} \ , \tag{3.94}$$

and

$$\delta\eta_\sigma(\tau) \equiv \tilde\eta_\sigma(\tau) - \eta_\sigma(\tau) \ . \tag{3.95}$$

Note that the differential dq_σ is given by

$$\begin{aligned} dq_\sigma &= \tilde\eta_\sigma(t+dt) - \eta_\sigma(t) \qquad (3.96)\\ &= \tilde\eta_\sigma(t+dt) - \tilde\eta_\sigma(t) + \tilde\eta_\sigma(t) - \eta_\sigma(t) \\ &= \dot{\tilde\eta}_\sigma(t)\, dt + \delta\eta_\sigma(t) \\ &= \dot q_\sigma(t)\, dt + \delta\eta_\sigma(t) + \mathcal{O}(\delta q \cdot dt) \ . \qquad (3.97) \end{aligned}$$

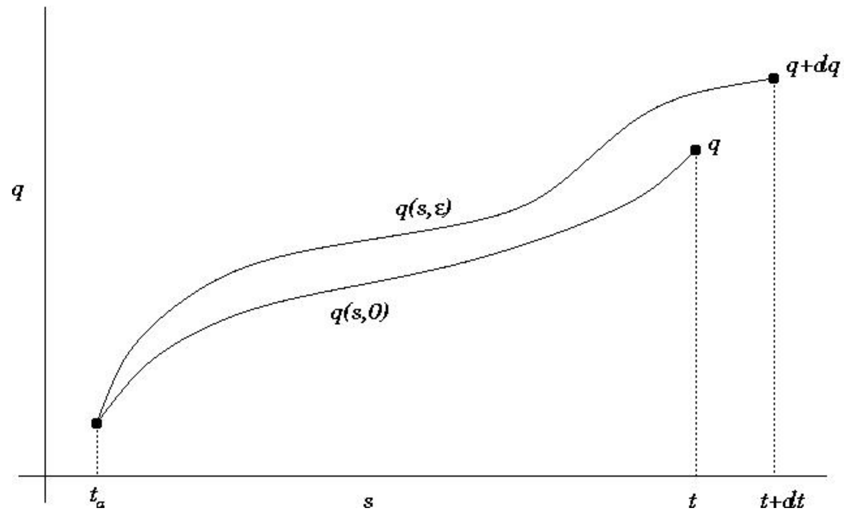

Figure 3.5: A one-parameter family of paths $q(s;\epsilon)$.

Thus, with $\pi_\sigma(t) \equiv p_\sigma$, we have

$$\begin{aligned} dS &= p_\sigma\, dq_\sigma + \big(L - p_\sigma\, \dot{q}_\sigma\big)\, dt \\ &= p_\sigma\, dq_\sigma - H\, dt\ . \end{aligned} \tag{3.98}$$

We therefore obtain

$$\frac{\partial S}{\partial q_\sigma} = p_\sigma \quad , \quad \frac{\partial S}{\partial t} = -H \quad , \quad \frac{dS}{dt} = L\ . \tag{3.99}$$

What about the lower limit at t_a? Clearly there are $n+1$ constants associated with this limit: $\big\{q_1(t_a), \ldots, q_n(t_a); t_a\big\}$. Thus, we may write

$$S = S(q_1, \ldots, q_n; \Lambda_1, \ldots, \Lambda_n, t) + \Lambda_{n+1}\ , \tag{3.100}$$

where our $n+1$ constants are $\{\Lambda_1, \ldots, \Lambda_{n+1}\}$. If we regard S as a mixed generator, which is type-I in some variables and type-II in others, then each Λ_σ for $1 \le \sigma \le n$ may be chosen to be either Q_σ or P_σ. We will define

$$\Gamma_\sigma = \frac{\partial S}{\partial \Lambda_\sigma} = \begin{cases} +Q_\sigma & \text{if } \Lambda_\sigma = P_\sigma \\ -P_\sigma & \text{if } \Lambda_\sigma = Q_\sigma \end{cases} \tag{3.101}$$

For each σ, the two possibilities $\Lambda_\sigma = Q_\sigma$ or $\Lambda_\sigma = P_\sigma$ are of course rendered equivalent by a canonical transformation $(Q_\sigma, P_\sigma) \rightarrow (P_\sigma, -Q_\sigma)$.

3.8.2 The Hamilton-Jacobi Equation

Since the action $S(q, \Lambda, t)$ has been shown to generate a canonical transformation for which $\tilde{H}(Q, P) = 0$. This requirement may be written as

$$H\Big(q_1, \ldots, q_n, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}, t\Big) + \frac{\partial S}{\partial t} = 0\ . \tag{3.102}$$

This is the *Hamilton-Jacobi equation* (HJE). It is a first order partial differential equation in $n+1$ variables, and in general is nonlinear (since kinetic energy is generally a quadratic function of momenta). Since $\tilde{H}(Q,P,t)=0$, the equations of motion are trivial, and

$$Q_\sigma(t) = \text{const.} \quad , \quad P_\sigma(t) = \text{const.} \tag{3.103}$$

Once the HJE is solved, one must invert the relations $\Gamma_\sigma = \partial S(q,\Lambda,t)/\partial \Lambda_\sigma$ to obtain $q(Q,P,t)$. This is possible only if

$$\det\left(\frac{\partial^2 S}{\partial q_\alpha\,\partial \Lambda_\beta}\right) \neq 0 \ , \tag{3.104}$$

which is known as the *Hessian condition*.

It is worth noting that the HJE may have several solutions. For example, consider the case of the free particle, with $H(q,p)=p^2/2m$. The HJE is

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{\partial S}{\partial t} = 0 \ . \tag{3.105}$$

One solution of the HJE is

$$S(q,\Lambda,t) = \frac{m\,(q-\Lambda)^2}{2t} \ . \tag{3.106}$$

For this we find

$$\Gamma = \frac{\partial S}{\partial \Lambda} = -\frac{m}{t}\,(q-\Lambda) \quad \Rightarrow \quad q(t) = \Lambda - \frac{\Gamma}{m}\,t \ . \tag{3.107}$$

Here $\Lambda = q(0)$ is the initial value of q, and $\Gamma = -p$ is minus the momentum.

Another equally valid solution to the HJE is

$$S(q,\Lambda,t) = q\,\sqrt{2m\Lambda} - \Lambda\, t \ . \tag{3.108}$$

This yields

$$\Gamma = \frac{\partial S}{\partial \Lambda} = q\sqrt{\frac{2m}{\Lambda}} - t \quad \Rightarrow \quad q(t) = \sqrt{\frac{\Lambda}{2m}}\,(t+\Gamma) \ . \tag{3.109}$$

For this solution, Λ is the energy and Γ may be related to the initial value of $q(t) = \Gamma\sqrt{\Lambda/2m}$.

3.8.3 Time-Independent Hamiltonians

When H has no explicit time dependence, we may reduce the order of the HJE by one, writing

$$S(q,\Lambda,t) = W(q,\Lambda) + T(\Lambda,t) \ . \tag{3.110}$$

The HJE becomes

$$H\left(q,\frac{\partial W}{\partial q}\right) = -\frac{\partial T}{\partial t} \ . \tag{3.111}$$

Note that the LHS of the above equation is independent of t, and the RHS is independent of q. Therefore, each side must only depend on the constants Λ, which is to say that each side must be a constant, which, without loss of generality, we take to be Λ_1. Therefore

$$S(q,\Lambda,t) = W(q,\Lambda) - \Lambda_1 t \ . \tag{3.112}$$

The function $W(q,\Lambda)$ is called *Hamilton's characteristic function*. The HJE now takes the form

$$H\left(q_1,\ldots,q_n,\frac{\partial W}{\partial q_1},\ldots,\frac{\partial W}{\partial q_n}\right)=\Lambda_1\ . \tag{3.113}$$

Note that adding an arbitrary constant C to S generates the same equation, and simply shifts the last constant $\Lambda_{n+1}\to\Lambda_{n+1}+C$. This is equivalent to replacing t by $t-t_0$ with $t_0=C/\Lambda_1$, *i.e.* it just redefines the zero of the time variable.

3.8.4 Example: One-Dimensional Motion

As an example of the method, consider the one-dimensional system,

$$H(q,p)=\frac{p^2}{2m}+U(q)\ . \tag{3.114}$$

The HJE is

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2+U(q)=\Lambda\ . \tag{3.115}$$

which may be recast as

$$\frac{\partial S}{\partial q}=\sqrt{2m\big[\Lambda-U(q)\big]}\ , \tag{3.116}$$

with solution

$$S(q,\Lambda,t)=\sqrt{2m}\int^{q}\!dq'\,\sqrt{\Lambda-U(q')}-\Lambda\,t\ . \tag{3.117}$$

We now have

$$p=\frac{\partial S}{\partial q}=\sqrt{2m\big[\Lambda-U(q)\big]}\ , \tag{3.118}$$

as well as

$$\Gamma=\frac{\partial S}{\partial\Lambda}=\sqrt{\frac{m}{2}}\int^{q(t)}\frac{dq'}{\sqrt{\Lambda-U(q')}}-t\ . \tag{3.119}$$

Thus, the motion $q(t)$ is given by quadrature:

$$\Gamma+t=\sqrt{\frac{m}{2}}\int^{q(t)}\frac{dq'}{\sqrt{\Lambda-U(q')}}\ , \tag{3.120}$$

where Λ and Γ are constants. The lower limit on the integral is arbitrary and merely shifts t by another constant. Note that Λ is the total energy.

3.8.5 Separation of Variables

It is convenient to first work an example before discussing the general theory. Consider the following Hamiltonian, written in spherical polar coordinates:

$$H=\frac{1}{2m}\left(p_r^2+\frac{p_\theta^2}{r^2}+\frac{p_\phi^2}{r^2\sin^2\!\theta}\right)+\overbrace{A(r)+\frac{B(\theta)}{r^2}+\frac{C(\phi)}{r^2\sin^2\!\theta}}^{\text{potential }U(r,\theta,\phi)}\ . \tag{3.121}$$

We seek a solution with the characteristic function

$$W(r,\theta,\phi) = W_r(r) + W_\theta(\theta) + W_\phi(\phi) \ . \tag{3.122}$$

The HJE is then

$$\frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{1}{2mr^2}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{1}{2mr^2\sin^2\!\theta}\left(\frac{\partial W_\phi}{\partial \phi}\right)^2 + A(r) + \frac{B(\theta)}{r^2} + \frac{C(\phi)}{r^2\sin^2\!\theta} = \Lambda_1 = E \ . \tag{3.123}$$

Multiply through by $r^2\sin^2\!\theta$ to obtain

$$\frac{1}{2m}\left(\frac{\partial W_\phi}{\partial \phi}\right)^2 + C(\phi) = -\sin^2\!\theta\left\{\frac{1}{2m}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + B(\theta)\right\} - r^2\sin^2\!\theta\left\{\frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2 + A(r) - \Lambda_1\right\} \ . \tag{3.124}$$

The LHS is independent of (r,θ), and the RHS is independent of ϕ. Therefore, we may set

$$\frac{1}{2m}\left(\frac{\partial W_\phi}{\partial \phi}\right)^2 + C(\phi) = \Lambda_2 \ . \tag{3.125}$$

Proceeding, we replace the LHS in eqn. 3.124 with Λ_2, arriving at

$$\frac{1}{2m}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + B(\theta) + \frac{\Lambda_2}{\sin^2\!\theta} = -r^2\left\{\frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2 + A(r) - \Lambda_1\right\} \ . \tag{3.126}$$

The LHS of this equation is independent of r, and the RHS is independent of θ. Therefore,

$$\frac{1}{2m}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + B(\theta) + \frac{\Lambda_2}{\sin^2\!\theta} = \Lambda_3 \ . \tag{3.127}$$

We're left with

$$\frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2 + A(r) + \frac{\Lambda_3}{r^2} = \Lambda_1 \ . \tag{3.128}$$

The full solution is therefore

$$S(q,\Lambda,t) = \sqrt{2m}\int^{r}\!dr'\,\sqrt{\Lambda_1 - A(r') - \frac{\Lambda_3}{r'^2}} \tag{3.129}$$

$$+\sqrt{2m}\int^{\theta}\!d\theta'\,\sqrt{\Lambda_3 - B(\theta') - \frac{\Lambda_2}{\sin^2\!\theta'}}$$

$$+\sqrt{2m}\int^{\phi}\!d\phi'\,\sqrt{\Lambda_2 - C(\phi')} - \Lambda_1 t \ . \tag{3.130}$$

We then have

$$\varGamma_1 = \frac{\partial S}{\partial \varLambda_1} = \int^{r(t)} \frac{\sqrt{\frac{m}{2}}\,dr'}{\sqrt{\varLambda_1 - A(r') - \varLambda_3\, r'^{-2}}} - t \tag{3.131}$$

$$\varGamma_2 = \frac{\partial S}{\partial \varLambda_2} = -\int^{\theta(t)} \frac{\sqrt{\frac{m}{2}}\,d\theta'}{\sin^2\!\theta' \sqrt{\varLambda_3 - B(\theta') - \varLambda_2 \csc^2\!\theta'}} + \int^{\phi(t)} \frac{\sqrt{\frac{m}{2}}\,d\phi'}{\sqrt{\varLambda_2 - C(\phi')}} \tag{3.132}$$

$$\varGamma_3 = \frac{\partial S}{\partial \varLambda_3} = -\int^{r(t)} \frac{\sqrt{\frac{m}{2}}\,dr'}{r'^2 \sqrt{\varLambda_1 - A(r') - \varLambda_3\, r'^{-2}}} + \int^{\theta(t)} \frac{\sqrt{\frac{m}{2}}\,d\theta'}{\sqrt{\varLambda_3 - B(\theta') - \varLambda_2 \csc^2\!\theta'}} \ . \tag{3.133}$$

The game plan here is as follows. The first of the above trio of equations is inverted to yield $r(t)$ in terms of t and constants. This solution is then invoked in the last equation (the upper limit on the first integral on the RHS) in order to obtain an implicit equation for $\theta(t)$, which is invoked in the second equation to yield an implicit equation for $\phi(t)$. The net result is the motion of the system in terms of time t and the six constants $(\varLambda_1, \varLambda_2, \varLambda_3, \varGamma_1, \varGamma_2, \varGamma_3)$. A seventh constant, associated with an overall shift of the zero of t, arises due to the arbitrary lower limits of the integrals.

In general, the separation of variables method begins with[3]

$$W(q, \varLambda) = \sum_{\sigma=1}^{n} W_\sigma(q_\sigma, \varLambda) \ . \tag{3.134}$$

Each $W_\sigma(q_\sigma, \varLambda)$ may be regarded as a function of the single variable q_σ, and is obtained by satisfying an ODE of the form[4]

$$H_\sigma\bigg(q_\sigma, \frac{dW_\sigma}{dq_\sigma}\bigg) = \varLambda_\sigma \ . \tag{3.135}$$

We then have

$$p_\sigma = \frac{\partial W_\sigma}{\partial q_\sigma} \quad , \quad \varGamma_\sigma = \frac{\partial W}{\partial \varLambda_\sigma} + \delta_{\sigma,1}\, t \ . \tag{3.136}$$

Note that while each W_σ depends on only a single q_σ, it may depend on several of the $\varLambda_\sigma$.

3.8.6 Example #2 : Point Charge plus Electric Field

Consider a potential of the form

$$U(r) = \frac{k}{r} - Fz \ , \tag{3.137}$$

which corresponds to a charge in the presence of an external point charge plus an external electric field. This problem is amenable to separation in parabolic coordinates, (ξ, η, φ):

$$x = \sqrt{\xi\eta}\,\cos\varphi \quad , \quad y = \sqrt{\xi\eta}\,\sin\varphi \quad , \quad z = \tfrac{1}{2}(\xi - \eta) \ . \tag{3.138}$$

[3] Here we assume *complete separability*. A given system may only be *partially* separable.

[4] $H_\sigma(q_\sigma, p_\sigma)$ may also depend on several of the $\varLambda_\alpha$. See *e.g.* eqn. 3.128, which is of the form $H_r\big(r, \partial_r W_r, \varLambda_3\big) = \varLambda_1$.

Note that

$$\rho \equiv \sqrt{x^2+y^2} = \sqrt{\xi\eta} \tag{3.139}$$

$$r = \sqrt{\rho^2+z^2} = \tfrac{1}{2}(\xi+\eta)\ . \tag{3.140}$$

The kinetic energy is

$$\begin{aligned} T &= \tfrac{1}{2}m\big(\dot\rho^2+\rho^2\,\dot\varphi^2+\dot z^2\big) \\ &= \tfrac{1}{8}m\,(\xi+\eta)\bigg(\frac{\dot\xi^2}{\xi}+\frac{\dot\eta^2}{\eta}\bigg)+\tfrac{1}{2}m\,\xi\eta\,\dot\varphi^2\ , \end{aligned} \tag{3.141}$$

and hence the Lagrangian is

$$L = \tfrac{1}{8}m\,(\xi+\eta)\bigg(\frac{\dot\xi^2}{\xi}+\frac{\dot\eta^2}{\eta}\bigg)+\tfrac{1}{2}m\,\xi\eta\,\dot\varphi^2-\frac{2k}{\xi+\eta}+\tfrac{1}{2}F\,(\xi-\eta)\ . \tag{3.142}$$

Thus, the conjugate momenta are

$$p_\xi = \frac{\partial L}{\partial\dot\xi} = \tfrac{1}{4}m\,(\xi+\eta)\,\frac{\dot\xi}{\xi} \tag{3.143}$$

$$p_\eta = \frac{\partial L}{\partial\dot\eta} = \tfrac{1}{4}m\,(\xi+\eta)\,\frac{\dot\eta}{\eta} \tag{3.144}$$

$$p_\varphi = \frac{\partial L}{\partial\dot\varphi} = m\,\xi\eta\,\dot\varphi\ , \tag{3.145}$$

and the Hamiltonian is

$$H = p_\xi\,\dot\xi+p_\eta\,\dot\eta+p_\varphi\,\dot\varphi \tag{3.146}$$

$$= \frac{2}{m}\bigg(\frac{\xi\,p_\xi^2+\eta\,p_\eta^2}{\xi+\eta}\bigg)+\frac{p_\varphi^2}{2m\xi\eta}+\frac{2k}{\xi+\eta}-\tfrac{1}{2}F\,(\xi-\eta)\ . \tag{3.147}$$

Notice that $\partial H/\partial t = 0$, which means $dH/dt = 0$, *i.e.* $H = E \equiv \Lambda_1$ is a constant of the motion. Also, φ is cyclic in H, so its conjugate momentum p_φ is a constant of the motion.

We write

$$S(q,\Lambda) = W(q,\Lambda)-Et \tag{3.148}$$

$$= W_\xi(\xi,\Lambda)+W_\eta(\eta,\Lambda)+W_\varphi(\varphi,\Lambda)-Et\ . \tag{3.149}$$

with $E = \Lambda_1$. Clearly we may take

$$W_\varphi(\varphi,\Lambda) = P_\varphi\,\varphi\ , \tag{3.150}$$

where $P_\varphi = \Lambda_2$. Multiplying the Hamilton-Jacobi equation by $\frac{1}{2}m\,(\xi+\eta)$ then gives

$$\begin{aligned} &\xi\bigg(\frac{dW_\xi}{d\xi}\bigg)^2+\frac{P_\varphi^2}{4\xi}+mk-\tfrac{1}{4}F\xi^2-\tfrac{1}{2}mE\xi \\ &\qquad = -\eta\bigg(\frac{dW_\eta}{d\eta}\bigg)^2-\frac{P_\varphi^2}{4\eta}-\tfrac{1}{4}F\eta^2+\tfrac{1}{2}mE\eta \equiv \Upsilon\ , \end{aligned} \tag{3.151}$$

where $\Upsilon = \Lambda_3$ is the third constant: $\Lambda = (E, P_\varphi, \Upsilon)$. Thus,

$$S(\overbrace{\xi, \eta, \varphi}^{q}; \underbrace{E, P_\varphi, \Upsilon}_{\Lambda}) = \int^{\xi} d\xi' \sqrt{\tfrac{1}{2}mE + \frac{\Upsilon - mk}{\xi'} + \tfrac{1}{4}mF\xi' - \frac{P_\varphi^2}{4\xi'^2}}$$
$$+ \int^{\eta} d\eta' \sqrt{\tfrac{1}{2}mE - \frac{\Upsilon}{\eta'} - \tfrac{1}{4}mF\eta' - \frac{P_\varphi^2}{4\eta'^2}}$$
$$+ P_\varphi \varphi - Et \ . \tag{3.152}$$

3.8.7 Example #3 : Charged Particle in a Magnetic Field

The Hamiltonian is

$$H = \frac{1}{2m}\Big(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\Big)^2 \ . \tag{3.153}$$

We choose the gauge $\boldsymbol{A} = Bx\hat{\boldsymbol{y}}$, and we write

$$S(x, y, P_1, P_2) = W_x(x, P_1, P_2) + W_y(y, P_1, P_2) - P_1\, t \ . \tag{3.154}$$

Note that here we will consider S to be a function of $\{q_\sigma\}$ and $\{P_\sigma\}$.

The Hamilton-Jacobi equation is then

$$\left(\frac{\partial W_x}{\partial x}\right)^2 + \left(\frac{\partial W_y}{\partial y} - \frac{eBx}{c}\right)^2 = 2mP_1 \ . \tag{3.155}$$

We solve by writing

$$W_y = P_2\, y \qquad \Rightarrow \qquad \left(\frac{dW_x}{dx}\right)^2 + \left(P_2 - \frac{eBx}{c}\right)^2 = 2mP_1 \ . \tag{3.156}$$

This equation suggests the substitution

$$x = \frac{cP_2}{eB} + \frac{c}{eB}\sqrt{2mP_1}\,\sin\theta \ . \tag{3.157}$$

in which case

$$\frac{\partial x}{\partial \theta} = \frac{c}{eB}\sqrt{2mP_1}\,\cos\theta \tag{3.158}$$

and

$$\frac{\partial W_x}{\partial x} = \frac{\partial W_x}{\partial \theta}\cdot\frac{\partial \theta}{\partial x} = \frac{eB}{c\sqrt{2mP_1}}\,\frac{1}{\cos\theta}\,\frac{\partial W_x}{\partial \theta} \ . \tag{3.159}$$

Substitution this into eqn. 3.156, we have

$$\frac{\partial W_x}{\partial \theta} = \frac{2mcP_1}{eB}\,\cos^2\!\theta \ , \tag{3.160}$$

with solution

$$W_x = \frac{mcP_1}{eB}\,\theta + \frac{mcP_1}{2eB}\,\sin(2\theta) \ . \tag{3.161}$$

We then have

$$p_x = \frac{\partial W_x}{\partial x} = \frac{\partial W_x}{\partial \theta} \bigg/ \frac{\partial x}{\partial \theta} = \sqrt{2mP_1}\,\cos\theta \tag{3.162}$$

and

$$p_y = \frac{\partial W_y}{\partial y} = P_2 \ . \tag{3.163}$$

The type-II generator we seek is then

$$S(q,P,t) = \frac{mcP_1}{eB}\,\theta + \frac{mcP_1}{2eB}\,\sin(2\theta) + P_2\,y - P_1\,t \ , \tag{3.164}$$

where

$$\theta = \frac{eB}{c\sqrt{2mP_1}}\,\sin^{-1}\!\left(x - \frac{cP_2}{eB}\right) . \tag{3.165}$$

Note that, from eqn. 3.157, we may write

$$dx = \frac{c}{eB}\,dP_2 + \frac{mc}{eB}\,\frac{1}{\sqrt{2mP_1}}\,\sin\theta\,dP_1 + \frac{c}{eB}\,\sqrt{2mP_1}\,\cos\theta\,d\theta \ , \tag{3.166}$$

from which we derive

$$\frac{\partial\theta}{\partial P_1} = -\frac{\tan\theta}{2P_1} \qquad , \qquad \frac{\partial\theta}{\partial P_2} = -\frac{1}{\sqrt{2mP_1}\,\cos\theta} \ . \tag{3.167}$$

These results are useful in the calculation of Q_1 and Q_2:

$$\begin{aligned} Q_1 &= \frac{\partial S}{\partial P_1} \\ &= \frac{mc}{eB}\,\theta + \frac{mcP_1}{eB}\,\frac{\partial\theta}{\partial P_1} + \frac{mc}{2eB}\,\sin(2\theta) + \frac{mcP_1}{eB}\,\cos(2\theta)\,\frac{\partial\theta}{\partial P_1} - t \\ &= \frac{mc}{eB}\,\theta - t \end{aligned} \tag{3.168}$$

and

$$\begin{aligned} Q_2 &= \frac{\partial S}{\partial P_2} \\ &= y + \frac{mcP_1}{eB}\big[1 + \cos(2\theta)\big]\,\frac{\partial\theta}{\partial P_2} \\ &= y - \frac{c}{eB}\,\sqrt{2mP_1}\,\cos\theta \ . \end{aligned} \tag{3.169}$$

Now since $\tilde{H}(P,Q) = 0$, we have that $\dot{Q}_\sigma = 0$, which means that each Q_σ is a constant. We therefore have the following solution:

$$x(t) = x_0 + A\,\sin(\omega_c t + \delta) \tag{3.170}$$

$$y(t) = y_0 + A\,\cos(\omega_c t + \delta) \ , \tag{3.171}$$

where $\omega_c = eB/mc$ is the 'cyclotron frequency', and

$$x_0 = \frac{cP_2}{eB} \quad , \quad y_0 = Q_2 \quad , \quad \delta \equiv \omega_c\,Q_1 \quad , \quad A = \frac{c}{eB}\,\sqrt{2mP_1} \ . \tag{3.172}$$

3.9 Action-Angle Variables

3.9.1 Circular Phase Orbits: Librations and Rotations

In a completely integrable system, the Hamilton-Jacobi equation may be solved by separation of variables. Each momentum p_σ is a function of only its corresponding coordinate q_σ plus constants – no other coordinates enter:

$$p_\sigma = \frac{\partial W_\sigma}{\partial q_\sigma} = p_\sigma(q_\sigma, \Lambda) \ . \tag{3.173}$$

The motion satisfies

$$H_\sigma(q_\sigma, p_\sigma) = \Lambda_\sigma \ . \tag{3.174}$$

The level sets of H_σ are curves $\mathcal{C}_\sigma$. In general, these curves each depend on all of the constants Λ, so we write $\mathcal{C}_\sigma = \mathcal{C}_\sigma(\Lambda)$. The curves $\mathcal{C}_\sigma$ are the *projections* of the full motion onto the (q_σ, p_σ) plane. In general we will assume the motion, and hence the curves $\mathcal{C}_\sigma$, is *bounded*. In this case, two types of projected motion are possible: librations and rotations. Librations are periodic oscillations about an equilibrium position. Rotations involve the advancement of an angular variable by 2π during a cycle. This is most conveniently illustrated in the case of the simple pendulum, for which

$$H(p_\phi, \phi) = \frac{p_\phi^2}{2I} + \tfrac{1}{2} I\omega^2 \big(1 - \cos\phi\big) \ . \tag{3.175}$$

- When $E < I\,\omega^2$, the momentum p_ϕ vanishes at $\phi = \pm\cos^{-1}(2E/I\omega^2)$. The system executes librations between these extreme values of the angle ϕ.
- When $E > I\,\omega^2$, the kinetic energy is always positive, and the angle advances monotonically, executing rotations.

In a completely integrable system, each $\mathcal{C}_\sigma$ is either a libration or a rotation[5]. Both librations and rotations are closed curves. Thus, each $\mathcal{C}_\sigma$ is in general homotopic to (= "can be continuously distorted to yield") a circle, S^1. For n freedoms, the motion is therefore confined to an n-torus, T^n:

$$T^n = \overbrace{S^1 \times S^1 \times \cdots \times S^1}^{n \text{ times}} \ . \tag{3.176}$$

These are called *invariant tori* (or *invariant manifolds*). There are many such tori, as there are many $\mathcal{C}_\sigma$ curves in each of the n two-dimensional submanifolds.

Invariant tori never intersect! This is ruled out by the uniqueness of the solution to the dynamical system, expressed as a set of coupled ordinary differential equations.

Note also that phase space is of dimension $2n$, while the invariant tori are of dimension n. Phase space is 'covered' by the invariant tori, but it is in general difficult to conceive of how this happens. Perhaps the most accessible analogy is the $n = 1$ case, where the '1-tori' are just circles. Two-dimensional phase space is covered noninteracting circular orbits. (The orbits are *topologically* equivalent to circles, although *geometrically* they may be distorted.) It is challenging to think about the $n = 2$ case, where a four-dimensional phase space is filled by nonintersecting 2-tori.

[5] $\mathcal{C}_\sigma$ may correspond to a separatrix, but this is a nongeneric state of affairs.

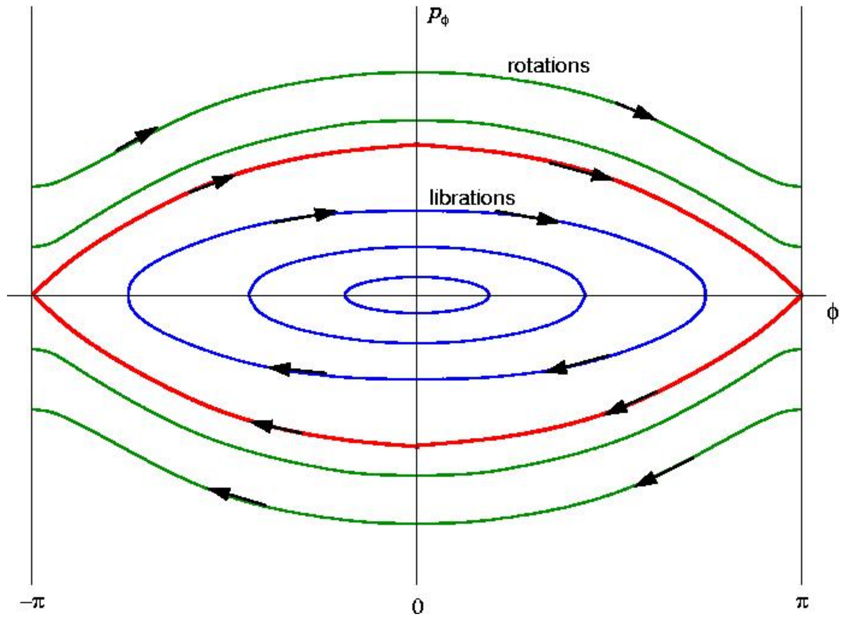

Figure 3.6: Phase curves for the simple pendulum, showing librations (in blue), rotations (in green), and the separatrix (in red). This phase flow is most correctly viewed as taking place on a cylinder, obtained from the above sketch by identifying the lines $\phi = \pi$ and $\phi = -\pi$.

3.9.2 Action-Angle Variables

For a completely integrable system, one can transform canonically from (q, p) to new coordinates (ϕ, J) which specify a particular n-torus T^n as well as the location on the torus, which is specified by n angle variables. The $\{J_\sigma\}$ are 'momentum' variables which specify the torus itself; they are constants of the motion since the tori are invariant. They are called *action variables*. Since $\dot{J}_\sigma = 0$, we must have

$$\dot{J}_\sigma = -\frac{\partial H}{\partial \phi_\sigma} = 0 \quad \Longrightarrow \quad H = H(J) \ . \tag{3.177}$$

The $\{\phi_\sigma\}$ are the *angle variables*.

The coordinate ϕ_σ describes the projected motion along $\mathcal{C}_\sigma$, and is normalized by

$$\oint\limits_{\mathcal{C}_\sigma} d\phi_\sigma = 2\pi \qquad \text{(once around } \mathcal{C}_\sigma) \ . \tag{3.178}$$

The dynamics of the angle variables are given by

$$\dot{\phi}_\sigma = \frac{\partial H}{\partial J_\sigma} \equiv \nu_\sigma(J) \ . \tag{3.179}$$

Thus,

$$\phi_\sigma(t) = \phi_\sigma(0) + \nu_\sigma(J)\, t \ . \tag{3.180}$$

The $\big\{\nu_\sigma(J)\big\}$ are *frequencies* describing the rate at which the $\mathcal{C}_\sigma$ are traversed; $T_\sigma(J) = 2\pi/\nu_\sigma(J)$ is the period.

3.9.3 Canonical Transformation to Action-Angle Variables

The $\{J_\sigma\}$ determine the $\{\mathcal{C}_\sigma\}$; each q_σ determines a point on $\mathcal{C}_\sigma$. This suggests a type-II transformation, with generator $F_2(q,J)$:

$$p_\sigma = \frac{\partial F_2}{\partial q_\sigma} \quad , \quad \phi_\sigma = \frac{\partial F_2}{\partial J_\sigma} \ . \tag{3.181}$$

Note that[6]

$$2\pi = \oint\limits_{\mathcal{C}_\sigma} d\phi_\sigma = \oint\limits_{\mathcal{C}_\sigma} d\bigg(\frac{\partial F_2}{\partial J_\sigma}\bigg) = \oint\limits_{\mathcal{C}_\sigma} \frac{\partial^2 F_2}{\partial J_\sigma\,\partial q_\sigma}\,dq_\sigma = \frac{\partial}{\partial J_\sigma}\oint\limits_{\mathcal{C}_\sigma} p_\sigma\,dq_\sigma \ , \tag{3.182}$$

which suggests the definition

$$J_\sigma = \frac{1}{2\pi}\oint\limits_{\mathcal{C}_\sigma} p_\sigma\,dq_\sigma \ . \tag{3.183}$$

I.e. J_σ is $(2\pi)^{-1}$ times the area enclosed by $\mathcal{C}_\sigma$.

If, separating variables,

$$W(q,\Lambda) = \sum_\sigma W_\sigma(q_\sigma,\Lambda) \tag{3.184}$$

is Hamilton's characteristic function for the transformation $(q,p) \to (Q,P)$, then

$$J_\sigma = \frac{1}{2\pi}\oint\limits_{\mathcal{C}_\sigma} \frac{\partial W_\sigma}{\partial q_\sigma}\,dq_\sigma = J_\sigma(\Lambda) \tag{3.185}$$

is a function only of the $\{\Lambda_\alpha\}$ and not the $\{\Gamma_\alpha\}$. We then invert this relation to obtain $\Lambda(J)$, to finally obtain

$$F_2(q,J) = W\big(q,\Lambda(J)\big) = \sum_\sigma W_\sigma\big(q_\sigma,\Lambda(J)\big) \ . \tag{3.186}$$

Thus, the recipe for canonically transforming to action-angle variable is as follows:

(1) Separate and solve the Hamilton-Jacobi equation for $W(q,\Lambda) = \sum_\sigma W_\sigma(q_\sigma,\Lambda)$.

(2) Find the orbits $\mathcal{C}_\sigma$ – the level sets of satisfying $H_\sigma(q_\sigma,p_\sigma) = \Lambda_\sigma$.

(3) Invert the relation $J_\sigma(\Lambda) = \frac{1}{2\pi}\oint\limits_{\mathcal{C}_\sigma} \frac{\partial W_\sigma}{\partial q_\sigma}\,dq_\sigma$ to obtain $\Lambda(J)$.

(4) $F_2(q,J) = \sum_\sigma W_\sigma\big(q_\sigma,\Lambda(J)\big)$ is the desired type-II generator[7].

---

[6] In general, we should write $d\big(\frac{\partial F_2}{\partial J_\sigma}\big) = \frac{\partial^2 F_2}{\partial J_\sigma\,\partial q_\alpha}\,dq_\alpha$ with a sum over α. However, in eqn. 3.182 all coordinates and momenta other than q_σ and p_σ are held fixed. Thus, $\alpha=\sigma$ is the only term in the sum which contributes.

[7] Note that $F_2(q,J)$ is time-independent. *I.e.* we are not transforming to $\tilde{H}=0$, but rather to $\tilde{H}=\tilde{H}(J)$.

3.9.4 Example : Harmonic Oscillator

The Hamiltonian is

$$H = \frac{p^2}{2m} + \tfrac{1}{2} m\omega_0^2 q^2 \ , \tag{3.187}$$

hence the Hamilton-Jacobi equation is

$$\left(\frac{dW}{dq}\right)^2 + m^2\omega_0^2 q^2 = 2m\Lambda \ . \tag{3.188}$$

Thus,

$$p = \frac{dW}{dq} = \pm\sqrt{2m\Lambda - m^2\omega_0^2 q^2} \ . \tag{3.189}$$

We now define

$$q \equiv \left(\frac{2\Lambda}{m\omega_0^2}\right)^{1/2} \sin\theta \quad \Rightarrow \quad p = \sqrt{2m\Lambda}\,\cos\theta \ , \tag{3.190}$$

in which case

$$J = \frac{1}{2\pi}\oint p\,dq = \frac{1}{2\pi}\cdot\frac{2\Lambda}{\omega_0}\cdot\int_0^{2\pi}\!d\theta\,\cos^2\!\theta = \frac{\Lambda}{\omega_0} \ . \tag{3.191}$$

Solving the HJE, we write

$$\frac{dW}{d\theta} = \frac{\partial q}{\partial\theta}\cdot\frac{dW}{dq} = 2J\cos^2\!\theta \ . \tag{3.192}$$

Integrating,

$$W = J\theta + \tfrac{1}{2} J\sin 2\theta \ , \tag{3.193}$$

up to an irrelevant constant. We then have

$$\phi = \frac{\partial W}{\partial J}\bigg|_q = \theta + \tfrac{1}{2}\sin 2\theta + J\big(1+\cos 2\theta\big)\,\frac{\partial\theta}{\partial J}\bigg|_q \ . \tag{3.194}$$

To find $(\partial\theta/\partial J)_q$, we differentiate $q = \sqrt{2J/m\omega_0}\,\sin\theta$:

$$dq = \frac{\sin\theta}{\sqrt{2m\omega_0 J}}\,dJ + \sqrt{\frac{2J}{m\omega_0}}\cos\theta\,d\theta \quad \Rightarrow \quad \frac{\partial\theta}{\partial J}\bigg|_q = -\frac{1}{2J}\,\tan\theta \ . \tag{3.195}$$

Plugging this result into eqn. 3.194, we obtain $\phi = \theta$. Thus, the full transformation is

$$q = \left(\frac{2J}{m\omega_0}\right)^{1/2} \sin\phi \quad , \quad p = \sqrt{2m\omega_0 J}\,\cos\phi \ . \tag{3.196}$$

The Hamiltonian is

$$H = \omega_0\, J \ , \tag{3.197}$$

hence $\dot\phi = \frac{\partial H}{\partial J} = \omega_0$ and $\dot J = -\frac{\partial H}{\partial\phi} = 0$, with solution $\phi(t) = \phi(0) + \omega_0\, t$ and $J(t) = J(0)$.

3.9.5 Example : Particle in a Box

Consider a particle in an open box of dimensions $L_x \times L_y$ moving under the influence of gravity. The bottom of the box lies at $z = 0$. The Hamiltonian is

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} + mgz \ . \tag{3.198}$$

Step one is to solve the Hamilton-Jacobi equation via separation of variables. The Hamilton-Jacobi equation is written

$$\frac{1}{2m}\bigg(\frac{\partial W_x}{\partial x}\bigg)^2 + \frac{1}{2m}\bigg(\frac{\partial W_y}{\partial y}\bigg)^2 + \frac{1}{2m}\bigg(\frac{\partial W_z}{\partial z}\bigg)^2 + mgz = E \equiv \Lambda_z \ . \tag{3.199}$$

We can solve for $W_{x,y}$ by inspection:

$$W_x(x) = \sqrt{2m\Lambda_x}\, x \quad , \quad W_y(y) = \sqrt{2m\Lambda_y}\, y \ . \tag{3.200}$$

We then have[8]

$$W_z'(z) = -\sqrt{2m\big(\Lambda_z - \Lambda_x - \Lambda_y - mgz\big)} \tag{3.201}$$

$$W_z(z) = \frac{2\sqrt{2}}{3\sqrt{m}\, g}\big(\Lambda_z - \Lambda_x - \Lambda_y - mgz\big)^{3/2} \ . \tag{3.202}$$

Step two is to find the $\mathcal{C}_\sigma$. Clearly $p_{x,y} = \sqrt{2m\Lambda_{x,y}}$. For fixed p_x, the x motion proceeds from $x = 0$ to $x = L_x$ and back, with corresponding motion for y. For x, we have

$$p_z(z) = W_z'(z) = \sqrt{2m\big(\Lambda_z - \Lambda_x - \Lambda_y - mgz\big)} \ , \tag{3.203}$$

and thus $\mathcal{C}_z$ is a truncated parabola, with $z_{\rm max} = (\Lambda_z - \Lambda_x - \Lambda_y)/mg$.

Step three is to compute $J(\Lambda)$ and invert to obtain $\Lambda(J)$. We have

$$J_x = \frac{1}{2\pi}\oint\limits_{\mathcal{C}_x} p_x\, dx = \frac{1}{\pi}\int\limits_0^{L_x} dx \sqrt{2m\Lambda_x} = \frac{L_x}{\pi}\sqrt{2m\Lambda_x} \tag{3.204}$$

$$J_y = \frac{1}{2\pi}\oint\limits_{\mathcal{C}_y} p_y\, dy = \frac{1}{\pi}\int\limits_0^{L_y} dy \sqrt{2m\Lambda_y} = \frac{L_y}{\pi}\sqrt{2m\Lambda_y} \tag{3.205}$$

and

$$\begin{aligned} J_z &= \frac{1}{2\pi}\oint\limits_{\mathcal{C}_z} p_z\, dz = \frac{1}{\pi}\int\limits_0^{z_{\rm max}} dx \sqrt{2m\big(\Lambda_z - \Lambda_x - \Lambda_y - mgz\big)} \\ &= \frac{2\sqrt{2}}{3\pi\sqrt{m}\, g}\big(\Lambda_z - \Lambda_x - \Lambda_y\big)^{3/2} \ . \end{aligned} \tag{3.206}$$

[8]Our choice of signs in taking the square roots for W_x', W_y', and W_z' is discussed below.

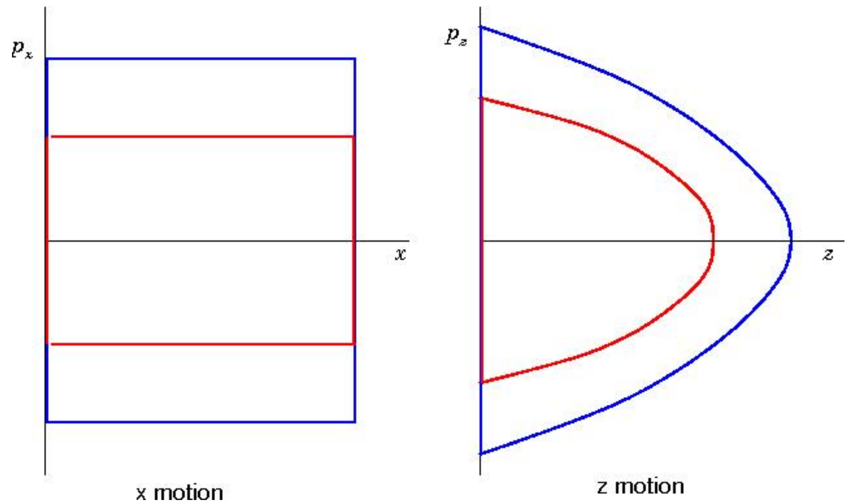

Figure 3.7: The librations $\mathcal{C}_z$ and $\mathcal{C}_x$. Not shown is $\mathcal{C}_y$, which is of the same shape as $\mathcal{C}_x$.

We now invert to obtain

$$\Lambda_x = \frac{\pi^2}{2mL_x^2}\,J_x^2 \qquad , \qquad \Lambda_y = \frac{\pi^2}{2mL_y^2}\,J_y^2 \tag{3.207}$$

$$\Lambda_z = \left(\frac{3\pi\sqrt{m}\,g}{2\sqrt{2}}\right)^{2/3} J_z^{2/3} + \frac{\pi^2}{2mL_x^2}\,J_x^2 + \frac{\pi^2}{2mL_y^2}\,J_y^2 \ . \tag{3.208}$$

$$F_2(x,y,z,J_x,J_y,J_z) = \frac{\pi x}{L_x}\,J_x + \frac{\pi y}{L_y}\,J_y + \pi\left(J_z^{2/3} - \frac{2m^{2/3}g^{1/3}z}{(3\pi)^{2/3}}\right)^{3/2} \ . \tag{3.209}$$

We now find

$$\phi_x = \frac{\partial F_2}{\partial J_x} = \frac{\pi x}{L_x} \quad , \quad \phi_y = \frac{\partial F_2}{\partial J_y} = \frac{\pi y}{L_y} \tag{3.210}$$

and

$$\phi_z = \frac{\partial F_2}{\partial J_z} = \pi\sqrt{1 - \frac{2m^{2/3}g^{1/3}z}{(3\pi J_z)^{2/3}}} = \pi\sqrt{1 - \frac{z}{z_{\rm max}}} \ , \tag{3.211}$$

where

$$z_{\rm max}(J_z) = \frac{(3\pi J_z)^{2/3}}{2m^{2/3}g^{1/3}} \ . \tag{3.212}$$

The momenta are

$$p_x = \frac{\partial F_2}{\partial x} = \frac{\pi J_x}{L_x} \quad , \quad p_y = \frac{\partial F_2}{\partial y} = \frac{\pi J_y}{L_y} \tag{3.213}$$

and

$$p_z = \frac{\partial F_2}{\partial z} = -\sqrt{2m}\left(\left(\frac{3\pi\sqrt{m}\,g}{2\sqrt{2}}\right)^{2/3} J_z^{2/3} - mgz\right)^{1/2} \ . \tag{3.214}$$

We note that the angle variables $\phi_{x,y,z}$ seem to be restricted to the range $[0,\pi]$, which seems to be at odds with eqn. 3.182. Similarly, the momenta $p_{x,y,z}$ all seem to be positive, whereas we know the momenta reverse sign when the particle bounces off a wall. The origin of the apparent discrepancy is that when we solved for the functions $W_{x,y,z}$, we had to take a square root in each case, and we chose a particular branch of the square root. So rather than $W_x(x) = \sqrt{2m\Lambda_x}\, x$, we should have taken

$$W_x(x) = \begin{cases} \sqrt{2m\Lambda_x}\, x & \text{if } p_x > 0 \\ \sqrt{2m\Lambda_x}\,(2L_x - x) & \text{if } p_x < 0\ . \end{cases} \tag{3.215}$$

The relation $J_x = (L_x/\pi)\sqrt{2m\Lambda_x}$ is unchanged, hence

$$W_x(x) = \begin{cases} (\pi x/L_x)\, J_x & \text{if } p_x > 0 \\ 2\pi J_x - (\pi x/L_x)\, J_x & \text{if } p_x < 0\ . \end{cases} \tag{3.216}$$

and

$$\phi_x = \begin{cases} \pi x/L_x & \text{if } p_x > 0 \\ \pi(2L_x - x)/L_x & \text{if } p_x < 0\ . \end{cases} \tag{3.217}$$

Now the angle variable ϕ_x advances by 2π during the cycle $\mathcal{C}_x$. Similar considerations apply to the y and z sectors.

3.9.6 Kepler Problem in Action-Angle Variables

This is discussed in detail in standard texts, such as Goldstein. The potential is $V(r) = -k/r$, and the problem is separable. We write[9]

$$W(r,\theta,\phi) = W_r(r) + W_\theta(\theta) + W_\varphi(\varphi)\ , \tag{3.218}$$

hence

$$\frac{1}{2m}\bigg(\frac{\partial W_r}{\partial r}\bigg)^2 + \frac{1}{2mr^2}\bigg(\frac{\partial W_\theta}{\partial \theta}\bigg)^2 + \frac{1}{2mr^2\sin^2\!\theta}\bigg(\frac{\partial W_\varphi}{\partial \varphi}\bigg)^2 + V(r) = E \equiv \Lambda_r\ . \tag{3.219}$$

Separating, we have

$$\frac{1}{2m}\bigg(\frac{dW_\varphi}{d\varphi}\bigg)^2 = \Lambda_\varphi \quad \Rightarrow \quad J_\varphi = \oint\limits_{\mathcal{C}_\varphi} d\varphi\, \frac{dW_\varphi}{d\varphi} = 2\pi\sqrt{2m\Lambda_\varphi}\ . \tag{3.220}$$

Next we deal with the θ coordinate:

$$\begin{aligned} \frac{1}{2m}\bigg(\frac{dW_\theta}{d\theta}\bigg)^2 &= \Lambda_\theta - \frac{\Lambda_\varphi}{\sin^2\!\theta} \qquad \Rightarrow \\ J_\theta &= 4\sqrt{2m\Lambda_\theta}\int\limits_0^{\theta_0} d\theta\, \sqrt{1 - \big(\Lambda_\varphi/\Lambda_\theta\big)\,\sin^{-2}\!\theta} \\ &= 2\pi\sqrt{2m}\Big(\sqrt{\Lambda_\theta} - \sqrt{\Lambda_\varphi}\Big)\ , \end{aligned} \tag{3.221}$$

[9] We denote the azimuthal angle by φ to distinguish it from the AA variable ϕ.

where $\theta_0 = \sin^{-1}(\Lambda_\varphi/\Lambda_\theta)$. Finally, we have[10]

$$\begin{aligned}
\frac{1}{2m}\left(\frac{dW_r}{dr}\right)^2 &= E + \frac{k}{r} - \frac{\Lambda_\theta}{r^2} \quad \Rightarrow \\
J_r &= \oint\limits_{\mathcal{C}_r} dr \, \sqrt{2m\left(E + \frac{k}{r} - \frac{\Lambda_\theta}{r^2}\right)} \\
&= -(J_\theta + J_\varphi) + \pi k \sqrt{\frac{2m}{|E|}} \ ,
\end{aligned} \tag{3.222}$$

where we've assumed $E < 0$, *i.e.* bound motion.

Thus, we find

$$H = E = -\frac{2\pi^2 m k^2}{\big(J_r + J_\theta + J_\varphi\big)^2} \ . \tag{3.223}$$

Note that the frequencies are completely degenerate:

$$\nu \equiv \nu_{r,\theta,\varphi} = \frac{\partial H}{\partial J_{r,\theta,\varphi}} = \frac{4\pi^2 m k^2}{\big(J_r + J_\theta + J_\varphi\big)^3} = \left(\frac{\pi^2 m k^2}{2|E|^3}\right)^{1/2} \ . \tag{3.224}$$

This threefold degeneracy may be removed by a transformation to new AA variables,

$$\Big\{(\phi_r, J_r),\ (\phi_\theta, J_\theta),\ (\phi_\varphi, J_\varphi)\Big\} \longrightarrow \Big\{(\phi_1, J_1),\ (\phi_2, J_2),\ (\phi_3, J_3)\Big\} \ , \tag{3.225}$$

using the type-II generator

$$F_2(\phi_r, \phi_\theta, \phi_\varphi; J_1, J_2, J_3) = (\phi_\varphi - \phi_\theta)\, J_1 + (\phi_\theta - \phi_r)\, J_2 + \phi_r\, J_3 \ , \tag{3.226}$$

which results in

$$\phi_1 = \frac{\partial F_2}{\partial J_1} = \phi_\varphi - \phi_\theta \qquad\qquad J_r = \frac{\partial F_2}{\partial \phi_r} = J_3 - J_2 \tag{3.227}$$

$$\phi_2 = \frac{\partial F_2}{\partial J_2} = \phi_\theta - \phi_r \qquad\qquad J_\theta = \frac{\partial F_2}{\partial \phi_\theta} = J_2 - J_1 \tag{3.228}$$

$$\phi_3 = \frac{\partial F_2}{\partial J_3} = \phi_r \qquad\qquad J_\varphi = \frac{\partial F_2}{\partial \phi_\varphi} = J_1 \ . \tag{3.229}$$

The new Hamiltonian is

$$H(J_1, J_2, J_3) = -\frac{2\pi^2 m k^2}{J_3^2} \ , \tag{3.230}$$

whence $\nu_1 = \nu_2 = 0$ and $\nu_3 = \nu$.

[10]The details of performing the integral around $\mathcal{C}_r$ are discussed in *e.g.* Goldstein.

3.9.7 Charged Particle in a Magnetic Field

For the case of the charged particle in a magnetic field, studied above in section 3.8.7, we found

$$x = \frac{cP_2}{eB} + \frac{c}{eB}\sqrt{2mP_1}\,\sin\theta \tag{3.231}$$

and

$$p_x = \sqrt{2mP_1}\,\cos\theta \qquad , \qquad p_y = P_2 \ . \tag{3.232}$$

The action variable J is then

$$J = \oint p_x\,dx = \frac{2mcP_1}{eB}\int\limits_0^{2\pi} d\theta\,\cos^2\!\theta = \frac{mcP_1}{eB} \ . \tag{3.233}$$

We then have

$$W = J\theta + \tfrac{1}{2}J\sin(2\theta) + Py \ , \tag{3.234}$$

where $P \equiv P_2$. Thus,

$$\begin{aligned}\phi &= \frac{\partial W}{\partial J} \\ &= \theta + \tfrac{1}{2}\sin(2\theta) + J\big[1+\cos(2\theta)\big]\,\frac{\partial\theta}{\partial J} \\ &= \theta + \tfrac{1}{2}\sin(2\theta) + 2J\cos^2\!\theta\cdot\bigg(-\frac{\tan\theta}{2J}\bigg) \\ &= \theta \ .\end{aligned} \tag{3.235}$$

The other canonical pair is (Q,P), where

$$Q = \frac{\partial W}{\partial P} = y - \sqrt{\frac{2cJ}{eB}}\,\cos\phi \ . \tag{3.236}$$

Therefore, we have

$$x = \frac{cP}{eB} + \sqrt{\frac{2cJ}{eB}}\,\sin\phi \qquad , \qquad y = Q + \sqrt{\frac{2cJ}{eB}}\,\cos\phi \tag{3.237}$$

and

$$p_x = \sqrt{\frac{2eBJ}{c}}\,\cos\phi \qquad , \qquad p_y = P \ . \tag{3.238}$$

The Hamiltonian is

$$\begin{aligned}H &= \frac{p_x^2}{2m} + \frac{1}{2m}\bigg(p_y - \frac{eBx}{c}\bigg)^2 \\ &= \frac{eBJ}{mc}\,\cos^2\!\phi + \frac{eBJ}{mc}\,\sin^2\!\phi \\ &= \omega_c\,J \ ,\end{aligned} \tag{3.239}$$

where $\omega_c = eB/mc$. The equations of motion are

$$\dot{\phi} = \frac{\partial H}{\partial J} = \omega_c \qquad , \qquad \dot{J} = -\frac{\partial H}{\partial \phi} = 0 \tag{3.240}$$

and

$$\dot{Q} = \frac{\partial H}{\partial P} = 0 \qquad , \qquad \dot{P} = -\frac{\partial H}{\partial Q} = 0 \ . \tag{3.241}$$

Thus, Q, P, and J are constants, and $\phi(t) = \phi_0 + \omega_c \, t$.

3.9.8 Motion on Invariant Tori

The angle variables evolve as

$$\phi_\sigma(t) = \nu_\sigma(J)\, t + \phi_\sigma(0) \ . \tag{3.242}$$

Thus, they wind around the invariant torus, specified by $\{J_\sigma\}$ at constant rates. In general, while each ϕ_σ executed periodic motion around a circle, the motion of the system as a whole is not periodic, since the frequencies $\nu_\sigma(J)$ are not, in general, commensurate. In order for the motion to be periodic, there must exist a set of integers, $\{l_\sigma\}$, such that

$$\sum_{\sigma=1}^{n} l_\sigma \, \nu_\sigma(J) = 0 \ . \tag{3.243}$$

This means that the ratio of any two frequencies ν_σ/ν_α must be a rational number. On a given torus, there are several possible orbits, depending on initial conditions $\phi(0)$. However, since the frequencies are determined by the action variables, which specify the tori, on a given torus either all orbits are periodic, or none are.

In terms of the original coordinates q, there are two possibilities:

$$\begin{aligned} q_\sigma(t) &= \sum_{l_1=-\infty}^{\infty} \cdots \sum_{l_n=-\infty}^{\infty} A^{(\sigma)}_{l_1 l_2 \cdots l_n} \, e^{i l_1 \phi_1(t)} \cdots e^{i l_n \phi_n(t)} \\ &\equiv \sum_{\boldsymbol{l}} A^\sigma_{\boldsymbol{l}} \, e^{i \boldsymbol{l} \cdot \boldsymbol{\phi}(t)} \qquad \text{(libration)} \end{aligned} \tag{3.244}$$

or

$$q_\sigma(t) = q^\circ_\sigma \, \phi_\sigma(t) \ + \sum_{\boldsymbol{l}} B^\sigma_{\boldsymbol{l}} \, e^{i \boldsymbol{l} \cdot \boldsymbol{\phi}(t)} \qquad \text{(rotation)} \ . \tag{3.245}$$

For rotations, the variable $q_\sigma(t)$ increased by $\Delta q_\sigma = 2\pi \, q^\circ_\sigma$.

3.10 Canonical Perturbation Theory

3.10.1 Canonical Transformations and Perturbation Theory

Suppose we have a Hamiltonian

$$H(\xi, t) = H_0(\xi, t) + \epsilon \, H_1(\xi, t) \ , \tag{3.246}$$

where ϵ is a small dimensionless parameter. Let's implement a type-II transformation, generated by $S(q,P,t)$:[11]

$$\tilde{H}(Q,P,t) = H(q,p,t) + \frac{\partial}{\partial t}\,S(q,P,t)\ . \tag{3.247}$$

Let's expand everything in powers of ϵ:

$$q_\sigma = Q_\sigma + \epsilon\, q_{1,\sigma} + \epsilon^2\, q_{2,\sigma} + \ldots \tag{3.248}$$

$$p_\sigma = P_\sigma + \epsilon\, p_{1,\sigma} + \epsilon^2\, p_{2,\sigma} + \ldots \tag{3.249}$$

$$\tilde{H} = \tilde{H}_0 + \epsilon\,\tilde{H}_1 + \epsilon^2\,\tilde{H}_2 + \ldots \tag{3.250}$$

$$S = \underbrace{q_\sigma\, P_\sigma}_{\substack{\text{identity}\\ \text{transformation}}} + \epsilon\, S_1 + \epsilon^2\, S_2 + \ldots\ . \tag{3.251}$$

Then

$$\begin{aligned} Q_\sigma = \frac{\partial S}{\partial P_\sigma} &= q_\sigma + \epsilon\,\frac{\partial S_1}{\partial P_\sigma} + \epsilon^2\,\frac{\partial S_2}{\partial P_\sigma} + \ldots \\ &= Q_\sigma + \bigg(q_{1,\sigma} + \frac{\partial S_1}{\partial P_\sigma}\bigg)\epsilon + \bigg(q_{2,\sigma} + \frac{\partial S_2}{\partial P_\sigma}\bigg)\epsilon^2 + \ldots \end{aligned} \tag{3.252}$$

and

$$p_\sigma = \frac{\partial S}{\partial q_\sigma} = P_\sigma + \epsilon\,\frac{\partial S_1}{\partial q_\sigma} + \epsilon^2\,\frac{\partial S_2}{\partial q_\sigma} + \ldots \tag{3.253}$$

$$= P_\sigma + \epsilon\, p_{1,\sigma} + \epsilon^2\, p_{2,\sigma} + \ldots\ . \tag{3.254}$$

We therefore conclude, order by order in ϵ,

$$q_{k,\sigma} = -\frac{\partial S_k}{\partial P_\sigma} \quad , \quad p_{k,\sigma} = +\frac{\partial S_k}{\partial q_\sigma}\ . \tag{3.255}$$

Now let's expand the Hamiltonian:

$$\tilde{H}(Q,P,t) = H_0(q,p,t) + \epsilon\, H_1(q,p,t) + \frac{\partial S}{\partial t} \tag{3.256}$$

$$\begin{aligned} &= H_0(Q,P,t) + \frac{\partial H_0}{\partial Q_\sigma}\,(q_\sigma - Q_\sigma) + \frac{\partial H_0}{\partial P_\sigma}\,(p_\sigma - P_\sigma) \\ &\qquad + \epsilon H_1(Q,P,t) + \epsilon\,\frac{\partial}{\partial t}\,S_1(Q,P,t) + \mathcal{O}(\epsilon^2) \\ &= H_0(Q,P,t) + \Bigg(-\frac{\partial H_0}{\partial Q_\sigma}\,\frac{\partial S_1}{\partial P_\sigma} + \frac{\partial H_0}{\partial P_\sigma}\,\frac{\partial S_1}{\partial Q_\sigma} + \frac{\partial S_1}{\partial t} + H_1\Bigg)\,\epsilon + \mathcal{O}(\epsilon^2) \\ &= H_0(Q,P,t) + \bigg(H_1 + \big\{S_1, H_0\big\} + \frac{\partial S_1}{\partial t}\bigg)\,\epsilon + \mathcal{O}(\epsilon^2)\ . \end{aligned} \tag{3.257}$$

[11] $Here, S(q,P,t)$ is not meant to signify Hamilton's principal function.

In the above expression, we evaluate $H_k(q,p,t)$ and $S_k(q,P,t)$ at $q = Q$ and $p = P$ and expand in the differences $q - Q$ and $p - P$. Thus, we have derived the relation

$$\tilde{H}(Q,P,t) = \tilde{H}_0(Q,P,t) + \epsilon \tilde{H}_1(Q,P,t) + \ldots \tag{3.258}$$

with

$$\tilde{H}_0(Q,P,t) = H_0(Q,P,t) \tag{3.259}$$

$$\tilde{H}_1(Q,P,t) = H_1 + \big\{S_1, H_0\big\} + \frac{\partial S_1}{\partial t} \ . \tag{3.260}$$

The problem, though, is this: we have one equation, eqn, 3.260, for the two unknowns $\tilde{H}_1$ and S_1. Thus, the problem is underdetermined. Of course, we could choose $\tilde{H}_1 = 0$, which basically recapitulates standard Hamilton-Jacobi theory. But we might just as well demand that $\tilde{H}_1$ satisfy some other requirement, such as that $\tilde{H}_0 + \epsilon\, \tilde{H}_1$ being integrable.

Incidentally, this treatment is paralleled by one in quantum mechanics, where a unitary transformation may be implemented to eliminate a perturbation to lowest order in a small parameter. Consider the Schrödinger equation,

$$i\hbar \frac{\partial \psi}{\partial t} = (\mathcal{H}_0 + \epsilon\, \mathcal{H}_1)\, \psi \ , \tag{3.261}$$

and define χ by

$$\psi \equiv e^{iS/\hbar}\, \chi \ , \tag{3.262}$$

with

$$S = \epsilon\, S_1 + \epsilon^2\, S_2 + \ldots \ . \tag{3.263}$$

As before, the transformation $U \equiv \exp(iS/\hbar)$ collapses to the identity in the $\epsilon \to 0$ limit. Now let's write the Schrödinger equation for χ. Expanding in powers of ϵ, one finds

$$i\hbar \frac{\partial \chi}{\partial t} = \mathcal{H}_0\, \chi + \epsilon \left(\mathcal{H}_1 + \frac{1}{i\hbar} \big[S_1, \mathcal{H}_0\big] + \frac{\partial S_1}{\partial t} \right) \chi + \ldots \equiv \tilde{\mathcal{H}}\, \chi \ , \tag{3.264}$$

where $[A,B] = AB - BA$ is the commutator. Note the classical-quantum correspondence,

$$\{A,B\} \longleftrightarrow \frac{1}{i\hbar}\, [A,B] \ . \tag{3.265}$$

Again, what should we choose for S_1? Usually the choice is made to make the $\mathcal{O}(\epsilon)$ term in $\tilde{\mathcal{H}}$ vanish. But this is not the only possible simplifying choice.

3.10.2 Canonical Perturbation Theory for $n = 1$ Systems

Henceforth we shall assume $H(\xi,t) = H(\xi)$ is time-independent, and we write the perturbed Hamiltonian as

$$H(\xi) = H_0(\xi) + \epsilon H_1(\xi) \ . \tag{3.266}$$

Let (ϕ_0, J_0) be the action-angle variables for H_0. Then

$$\tilde{H}_0(\phi_0, J_0) = H_0\big(q(\phi_0, J_0), p(\phi_0, J_0)\big) = \tilde{H}_0(J_0) \ . \tag{3.267}$$

We define

$$\tilde{H}_1(\phi_0, J_0) = H_1\big(q(\phi_0, J_0), p(\phi_0, J_0)\big) \ . \tag{3.268}$$

We assume that $\tilde{H} = \tilde{H}_0 + \epsilon\, \tilde{H}_1$ is integrable[12], so it, too, possesses action-angle variables, which we denote by (ϕ, J)[13]. Thus, there must be a canonical transformation taking $(\phi_0, J_0) \to (\phi, J)$, with

$$\tilde{H}\big(\phi_0(\phi, J), J_0(\phi, J)\big) \equiv K(J) = E(J) \ . \tag{3.269}$$

We solve via a type-II canonical transformation:

$$S(\phi_0, J) = \phi_0 J + \epsilon\, S_1(\phi_0, J) + \epsilon^2\, S_2(\phi_0, J) + \ldots \ , \tag{3.270}$$

where $\phi_0 J$ is the identity transformation. Then

$$J_0 = \frac{\partial S}{\partial \phi_0} = J + \epsilon\, \frac{\partial S_1}{\partial \phi_0} + \epsilon^2\, \frac{\partial S_2}{\partial \phi_0} + \ldots \tag{3.271}$$

$$\phi = \frac{\partial S}{\partial J} = \phi_0 + \epsilon\, \frac{\partial S_1}{\partial J} + \epsilon^2\, \frac{\partial S_2}{\partial J} + \ldots \ , \tag{3.272}$$

and

$$E(J) = E_0(J) + \epsilon\, E_1(J) + \epsilon^2\, E_2(J) + \ldots \tag{3.273}$$

$$= \tilde{H}_0(\phi_0, J_0) + \tilde{H}_1(\phi_0, J_0) \ . \tag{3.274}$$

We now expand $\tilde{H}(\phi_0, J_0)$ in powers of $J_0 - J$:

$$\begin{aligned}
\tilde{H}(\phi_0, J_0) &= \tilde{H}_0(\phi_0, J_0) + \epsilon\, \tilde{H}_1(\phi_0, J_0) && (3.275)\\
&= \tilde{H}_0(J) + \frac{\partial \tilde{H}_0}{\partial J}\,(J_0 - J) + \tfrac{1}{2}\,\frac{\partial^2 \tilde{H}_0}{\partial J^2}\,(J_0 - J)^2 + \ldots \\
&\qquad + \epsilon\, \tilde{H}_1(\phi_0, J_0) + \epsilon\, \frac{\partial \tilde{H}_1}{\partial J}\,(J_0 - J) + \ldots \\
&= \tilde{H}_0(J) + \left(\tilde{H}_1(\phi_0, J_0) + \frac{\partial \tilde{H}_0}{\partial J}\,\frac{\partial S_1}{\partial \phi_0}\right)\epsilon && (3.276)\\
&\qquad + \left(\frac{\partial \tilde{H}_0}{\partial J}\,\frac{\partial S_2}{\partial \phi_0} + \frac{1}{2}\,\frac{\partial^2 \tilde{H}_0}{\partial J^2}\left(\frac{\partial S_1}{\partial \phi_0}\right)^2 + \frac{\partial \tilde{H}_1}{\partial J}\,\frac{\partial S_1}{\partial \phi_0}\right)\epsilon^2 + \ldots \ .
\end{aligned}$$

Equating terms, then,

$$E_0(J) = \tilde{H}_0(J) \tag{3.277}$$

$$E_1(J) = \tilde{H}_1(\phi_0, J) + \frac{\partial \tilde{H}_0}{\partial J}\,\frac{\partial S_1}{\partial \phi_0} \tag{3.278}$$

$$E_2(J) = \frac{\partial \tilde{H}_0}{\partial J}\,\frac{\partial S_2}{\partial \phi_0} + \frac{1}{2}\,\frac{\partial^2 \tilde{H}_0}{\partial J^2}\left(\frac{\partial S_1}{\partial \phi_0}\right)^2 + \frac{\partial \tilde{H}_1}{\partial J}\,\frac{\partial S_1}{\partial \phi_0} \ . \tag{3.279}$$

[12]This is always true, in fact, for $n = 1$.

[13]We assume the motion is bounded, so action-angle variables may be used.

How, one might ask, can we be sure that the LHS of each equation in the above hierarchy depends only on J when each RHS seems to depend on ϕ_0 as well? The answer is that we use the freedom to choose each S_k to make this so. We demand each RHS be independent of ϕ_0, which means it must be equal to its average, $\langle\, \mathrm{RHS}(\phi_0)\,\rangle$, where

$$\big\langle f\big(\phi_0\big)\big\rangle = \int_0^{2\pi} \frac{d\phi_0}{2\pi}\, f\big(\phi_0\big) \ . \tag{3.280}$$

The average is performed *at fixed* J and *not* at fixed J_0. In this regard, we note that holding J constant and increasing ϕ_0 by 2π also returns us to the same starting point. Therefore, J is a periodic function of ϕ_0. We must then be able to write

$$S_k(\phi_0, J) = \sum_{m=-\infty}^{\infty} S_k(J;m)\, e^{im\phi_0} \tag{3.281}$$

for each $k > 0$, in which case

$$\left\langle \frac{\partial S_k}{\partial\phi_0}\right\rangle = \frac{1}{2\pi}\big[S_k(2\pi) - S_k(0)\big] = 0 \ . \tag{3.282}$$

Let's see how this averaging works to the first two orders of the hierarchy. Since $\tilde H_0(J)$ is independent of ϕ_0 and since $\partial S_1/\partial\phi_0$ is periodic, we have

$$E_1(J) = \big\langle \tilde H_1(\phi_0, J)\big\rangle + \frac{\partial \tilde H_0}{\partial J} \overbrace{\left\langle \frac{\partial S_1}{\partial\phi_0}\right\rangle}^{\text{this vanishes!}} \tag{3.283}$$

and hence S_1 must satisfy

$$\frac{\partial S_1}{\partial \phi_0} = \frac{\langle \tilde H_1\rangle - \tilde H_1}{\nu_0(J)} \ , \tag{3.284}$$

where $\nu_0(J) = \partial\tilde H_0/\partial J$. Clearly the RHS of eqn. 3.284 has zero average, and must be a periodic function of ϕ_0. The solution is $S_1 = S_1(\phi_0, J) + g(J)$, where $g(J)$ is an arbitrary function of J. However, $g(J)$ affects only the difference $\phi - \phi_0$, changing it by a constant value $g'(J)$. So there is no harm in taking $g(J) = 0$.

Next, let's go to second order in ϵ. We have

$$E_2(J) = \left\langle \frac{\partial \tilde H_1}{\partial J}\,\frac{\partial S_1}{\partial\phi_0}\right\rangle + \tfrac{1}{2}\frac{\partial\nu_0}{\partial J}\left\langle\left(\frac{\partial S_1}{\partial\phi_1}\right)^2\right\rangle + \nu_0(J)\overbrace{\left\langle \frac{\partial S_2}{\partial\phi_0}\right\rangle}^{\text{this vanishes!}} \ . \tag{3.285}$$

The equation for S_2 is then

$$\begin{aligned}\frac{\partial S_2}{\partial\phi_0} = \frac{1}{\nu_0^2(J)}\Bigg\{ &\left\langle\frac{\partial \tilde H_1}{\partial J}\right\rangle \langle \tilde H_0\rangle - \left\langle \frac{\partial\tilde H_1}{\partial J}\,\tilde H_0\right\rangle - \frac{\partial\tilde H_1}{\partial J}\,\langle\tilde H_1\rangle + \frac{\partial \tilde H_1}{\partial J}\,\tilde H_1 \\ &+ \frac{1}{2}\,\frac{\partial\ln\nu_0}{\partial J}\left(\langle \tilde H_1^2\rangle - 2\big\langle\tilde H_1\big\rangle^2 + 2\big\langle \tilde H_1\big\rangle - \tilde H_1^2\right)\Bigg\} \ .\end{aligned} \tag{3.286}$$

The expansion for the energy $E(J)$ is then

$$E(J) = \tilde{H}_0(J) + \epsilon\big\langle \tilde{H}_1 \big\rangle + \frac{\epsilon^2}{\nu_0(J)} \Bigg\{ \bigg\langle \frac{\partial \tilde{H}_1}{\partial J} \bigg\rangle \big\langle \tilde{H}_1 \big\rangle - \bigg\langle \frac{\partial \tilde{H}_1}{\partial J}\, \tilde{H}_1 \bigg\rangle + \frac{1}{2}\,\frac{\partial \ln \nu_0}{\partial J} \Big(\big\langle \tilde{H}_1^2 - \big\langle \tilde{H}_1 \big\rangle^2 \Big) \Bigg\} + \mathcal{O}(\epsilon^3) \ . \tag{3.287}$$

Note that we don't need S to find $E(J)$! The perturbed frequencies are

$$\nu(J) = \frac{\partial E}{\partial J} \ . \tag{3.288}$$

Sometimes the frequencies are all that is desired. However, we can of course obtain the full motion of the system via the succession of canonical transformations,

$$(\phi, J) \longrightarrow (\phi_0, J_0) \longrightarrow (q, p) \ . \tag{3.289}$$

3.10.3 Example : Nonlinear Oscillator

Consider the nonlinear oscillator with Hamiltonian

$$H(q,p) = \overbrace{\frac{p^2}{2m} + \tfrac{1}{2} m \nu_0^2 q^2}^{H_0} + \tfrac{1}{4} \epsilon \alpha q^4 \ . \tag{3.290}$$

The action-angle variables for the harmonic oscillator Hamiltonian H_0 are

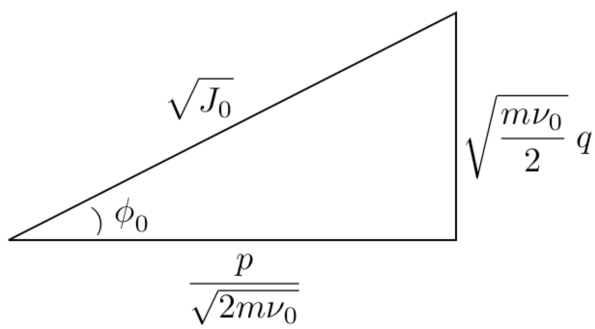

Figure 3.8: Action-angle variables for the harmonic oscillator.

$$\phi_0 = \tan^{-1}\!\big(mvq/p\big) \quad , \quad J_0 = \frac{p^2}{2m\nu_0} + \tfrac{1}{2} m \nu_0 q^2 \ , \tag{3.291}$$

and the relation between (ϕ_0, J_0) and (q, p) is further depicted in fig. 3.8. Note $H_0 = \nu_0\, J_0$. For the full Hamiltonian, we have

$$\begin{aligned} \tilde{H}(\phi_0, J_0) &= \nu_0 J_0 + \tfrac{1}{4}\epsilon\,\alpha \bigg(\sqrt{\frac{2J_0}{m\nu_0}}\, \sin\phi_0 \bigg)^4 \\ &= \nu_0 J_0 + \frac{\epsilon\alpha}{m^2 \nu_0^2}\, J_0^2\, \sin^4\!\phi_0 \ . \end{aligned} \tag{3.292}$$

We may now evaluate

$$E_1(J) = \langle \tilde{H}_1 \rangle = \frac{\alpha J^2}{m^2 \nu_0^2} \int_0^{2\pi} \frac{d\phi_0}{2\pi} \, \sin^4\!\phi_0 = \frac{3\alpha J^2}{8 m^2 \nu_0^2} \ . \tag{3.293}$$

The frequency, to order ϵ, is

$$\nu(J) = \nu_0 + \frac{3\,\epsilon\,\alpha J}{4 m^2 \nu_0^2} \ . \tag{3.294}$$

Now to lowest order in ϵ, we may replace J by $J_0 = \frac{1}{2} m \nu_0 A^2$, where A is the amplitude of the q motion. Thus,

$$\nu(A) = \nu_0 + \frac{3\epsilon\alpha}{8 m \nu_0} \ . \tag{3.295}$$

This result agrees with that obtained via heavier lifting, using the Poincaré-Lindstedt method.

Next, let's evaluate the canonical transformation $(\phi_0, J_0) \to (\phi, J)$. We have

$$\nu_0 \, \frac{\partial S_1}{\partial \phi_0} = \frac{\alpha J^2}{m^2 \nu_0^2} \left(\tfrac{3}{8} - \sin^4\!\phi_0 \right) \qquad \Rightarrow$$

$$S(\phi_0, J) = \phi_0 \, J + \frac{\epsilon \alpha J^2}{8 m^2 \nu_0^3} \left(3 + 2 \sin^2\!\phi_0\right) \, \sin\phi_0 \, \cos\phi_0 + \mathcal{O}(\epsilon^2) \ . \tag{3.296}$$

Thus,

$$\phi = \frac{\partial S}{\partial J} = \phi_0 + \frac{\epsilon \alpha J}{4 m^2 \nu_0^3} \left(3 + 2 \sin^2\!\phi_0\right) \, \sin\phi_0 \, \cos\phi_0 + \mathcal{O}(\epsilon^2) \tag{3.297}$$

$$J_0 = \frac{\partial S}{\partial \phi_0} = J + \frac{\epsilon \alpha J^2}{8 m^2 \nu_0^3} \left(4 \cos 2\phi_0 - \cos 4\phi_0\right) + \mathcal{O}(\epsilon^2) \ . \tag{3.298}$$

Again, to lowest order, we may replace J by J_0 in the above, whence

$$J = J_0 - \frac{\epsilon \alpha J_0^2}{8 m^2 \nu_0^3} \left(4 \cos 2\phi_0 - \cos 4\phi_0\right) + \mathcal{O}(\epsilon^2) \tag{3.299}$$

$$\phi = \phi_0 + \frac{\epsilon \alpha J_0}{8 m^2 \nu_0^3} \left(3 + 2 \sin^2\!\phi_0\right) \, \sin 2\phi_0 + \mathcal{O}(\epsilon^2) \ . \tag{3.300}$$

To obtain (q, p) in terms of (ϕ, J) is not analytically tractable – the relations cannot be analytically inverted.

3.10.4 $n > 1$ Systems : Degeneracies and Resonances

Generalizing the procedure we derived for $n = 1$, we obtain

$$J_0^\alpha = \frac{\partial S}{\partial \phi_0^\alpha} = J^\alpha + \epsilon \, \frac{\partial S_1}{\partial \phi_0^\alpha} + \epsilon^2 \, \frac{\partial S_2}{\partial \phi_0^\alpha} + \ldots \tag{3.301}$$

$$\phi^\alpha = \frac{\partial S}{\partial J^\alpha} = \phi_0^\alpha + \epsilon \, \frac{\partial S_1}{\partial J^\alpha} + \epsilon^2 \, \frac{\partial S_2}{\partial J^\alpha} + \ldots \tag{3.302}$$

and

$$E_0(\boldsymbol{J}) = \tilde{H}_0(\boldsymbol{J}) \tag{3.303}$$

$$E_1(\boldsymbol{J}) = \tilde{H}_0(\boldsymbol{\phi}_0, \boldsymbol{J}) + \nu_0^\alpha(\boldsymbol{J})\, \frac{\partial S_1}{\partial \phi_0^\alpha} \tag{3.304}$$

$$E_2(\boldsymbol{J}) = \frac{\partial \tilde{H}_0}{\partial J_\alpha}\, \frac{\partial S_2}{\partial \phi_0^\alpha} + \frac{1}{2}\, \frac{\partial \nu_0^\alpha}{\partial J^\beta}\, \frac{\partial S_1}{\partial \phi_0^\alpha}\, \frac{\partial S_1}{\partial \phi_0^\beta} + \nu_0^\alpha\, \frac{\partial S_1}{\partial \phi_0^\alpha}\ . \tag{3.305}$$

We now implement the averaging procedure, with

$$\big\langle f(J^1,\ldots,J^n)\big\rangle = \int\limits_0^{2\pi} \frac{d\phi_0^1}{2\pi} \,\cdots \int\limits_0^{2\pi} \frac{d\phi_0^n}{2\pi}\, f\big(\phi_0^1,\ldots,\phi_0^n,J^1,\ldots,J^n\big)\ . \tag{3.306}$$

The equation for S_1 is

$$\nu_0^\alpha\, \frac{\partial S_1}{\partial \phi_0^\alpha} = \big\langle \tilde{H}_1 \big\rangle - \tilde{H}_1 \equiv -{\sum_{\boldsymbol{l}}}' V_{\boldsymbol{l}}\, e^{i\boldsymbol{l}\cdot\boldsymbol{\phi}}\ , \tag{3.307}$$

where $\boldsymbol{l} = \{l^1, l^2, \ldots, l^n\}$, with each l^σ an integer, and with $\boldsymbol{l} \neq 0$. The solution is

$$S_1(\boldsymbol{\phi}_0, \boldsymbol{J}) = i\,{\sum_{l}}' \frac{V_{\boldsymbol{l}}}{\boldsymbol{l}\cdot\boldsymbol{\nu}_0}\, e^{i\boldsymbol{l}\cdot\boldsymbol{\phi}}\ . \tag{3.308}$$

where $\boldsymbol{l}\cdot\boldsymbol{\nu}_0 = l^\alpha \nu_0^\alpha$. When two or more of the frequencies $\nu_\alpha(J)$ are *commensurate*, there exists a set of integers l such that the denominator of $D(l)$ vanishes. But even when the frequencies are not rationally related, one can approximate the ratios $\nu_0^\alpha/\nu_0^{\alpha'}$ by rational numbers, and for large enough l the denominator can become arbitrarily small.

A similar problem arises with periodic time-dependent perturbations. Consider the system

$$H(\boldsymbol{\phi}, \boldsymbol{J}, t) = H_0(\boldsymbol{J}) + \epsilon\, V(\boldsymbol{\phi}, \boldsymbol{J}, t)\ , \tag{3.309}$$

where $V(t+T) = V(t)$. This means we may write

$$V(\boldsymbol{\phi}, \boldsymbol{J}, t) = \sum_k V_k(\boldsymbol{\phi}, \boldsymbol{J})\, e^{-ik\Omega t} \tag{3.310}$$

$$= \sum_{\boldsymbol{k}} \sum_{l} \hat{V}_{k,\boldsymbol{l}}(\boldsymbol{J})\, e^{i\boldsymbol{l}\cdot\boldsymbol{\phi}}\, e^{-ik\Omega t}\ . \tag{3.311}$$

by Fourier transforming from both time and angle variables; here $\Omega = 2\pi/T$. Note that $V(\boldsymbol{\phi}, \boldsymbol{J}, t)$ is real if $V^*_{k,\boldsymbol{l}} = V_{-\boldsymbol{k},-l}$. The equations of motion are

$$\dot{J}^\alpha = -\frac{\partial H}{\partial \phi^\alpha} = -i\epsilon \sum_{k,\boldsymbol{l}} l^\alpha\, \hat{V}_{k,\boldsymbol{l}}(\boldsymbol{J})\, e^{i\boldsymbol{l}\cdot\boldsymbol{\phi}}\, e^{-ik\Omega t} \tag{3.312}$$

$$\dot{\phi}^\alpha = +\frac{\partial H}{\partial J^\alpha} = \nu_0^\alpha(\boldsymbol{J}) + \epsilon \sum_{k,\boldsymbol{l}} \frac{\partial \hat{V}_{k,\boldsymbol{l}}(\boldsymbol{J})}{\partial J^\alpha}\, e^{i\boldsymbol{l}\cdot\boldsymbol{\phi}}\, e^{-ik\Omega t}\ . \tag{3.313}$$

We now expand in ϵ:

$$\phi^\alpha = \phi_0^\alpha + \epsilon\,\phi_1^\alpha + \epsilon^2\,\phi_2^\alpha + \ldots \tag{3.314}$$

$$J^\alpha = J_0^\alpha + \epsilon\,J_1^\alpha + \epsilon^2\,J_2^\alpha + \ldots\ . \tag{3.315}$$

To order ϵ^0, $J^\alpha = J_0^\alpha$ and $\phi_0^\alpha = \nu_0^\alpha\,t + \beta_0^\alpha$. To order ϵ^1,

$$\dot{J}_1^\alpha = -i \sum_{k,l} l^\alpha\,\hat{V}_{k,\boldsymbol{l}}(\boldsymbol{J}_0)\,e^{i(\boldsymbol{l}\cdot\boldsymbol{\nu}_0 - k\Omega)t}\,e^{i\cdot\boldsymbol{\beta_0}} \tag{3.316}$$

and

$$\dot{\phi}_1^\alpha = \frac{\partial\nu_0^\alpha}{\partial J^\beta}\,J_1^\beta + \sum_{k,\boldsymbol{l}} \frac{\partial\hat{V}_{k,\boldsymbol{l}}(\boldsymbol{J})}{\partial J^\alpha}\,e^{i(\boldsymbol{l}\cdot\boldsymbol{\nu}_0 - k\Omega)t}\,e^{i\boldsymbol{l}\cdot\boldsymbol{\beta}_0}\ , \tag{3.317}$$

where derivatives are evaluated at $\boldsymbol{J} = \boldsymbol{J}_0$. The solution is:

$$J_1^\alpha = \sum_{k,\boldsymbol{l}} \frac{l^\alpha\,\hat{V}_{k,\boldsymbol{l}}(J_0)}{k\Omega - \boldsymbol{l}\cdot\boldsymbol{\nu}_0}\,e^{i(\boldsymbol{l}\cdot\boldsymbol{\nu}_0 - k\Omega)t}\,e^{i\boldsymbol{l}\cdot\boldsymbol{\beta}_0} \tag{3.318}$$

$$\phi_1^\alpha = \left\{ \frac{\partial\nu_0^\alpha}{\partial J^\beta}\,\frac{l^\beta\,\hat{V}_{k,\boldsymbol{l}}(J_0)}{(k\Omega - \boldsymbol{l}\cdot\boldsymbol{\nu}_0)^2} + \frac{\partial\hat{V}_{k,\boldsymbol{l}}(J)}{\partial J^\alpha}\,\frac{1}{k\Omega - \boldsymbol{l}\cdot\boldsymbol{\nu}_0} \right\} e^{i(\boldsymbol{l}\cdot\boldsymbol{\nu}_0 - k\Omega)t}\,e^{i\boldsymbol{l}\cdot\boldsymbol{\beta}_0}\ . \tag{3.319}$$

When the resonance condition,

$$k\Omega = \boldsymbol{l}\cdot\boldsymbol{\nu}_0(\boldsymbol{J}_0)\ , \tag{3.320}$$

holds, the denominators vanish, and the perturbation theory breaks down.

3.10.5 Particle-Wave Interaction

Consider a particle of charge e moving in the presence of a constant magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{z}}$ and a space- and time-varying electric field $\boldsymbol{E}(\boldsymbol{x}, t)$, described by the Hamiltonian

$$H = \frac{1}{2m}\big(\boldsymbol{p} - \tfrac{e}{c}\boldsymbol{A}\big)^2 + \epsilon\,eV_0\,\cos(k_\perp x + k_z z - \omega t)\ , \tag{3.321}$$

where ϵ is a dimensionless expansion parameter. Working in the gauge $\boldsymbol{A} = Bx\hat{\boldsymbol{y}}$, from our earlier discussions in section 3.8.7, we may write

$$H = \omega_c J + \frac{p_z^2}{2m} + \epsilon\,eV_0\,\cos\left(k_z z + \frac{k_\perp P}{m\omega_c} + k_\perp\sqrt{\frac{2J}{m\omega_c}}\,\sin\phi - \omega t\right)\ . \tag{3.322}$$

Here,

$$x = \frac{P}{m\omega_c} + \sqrt{\frac{2J}{m\omega_c}}\,\sin\phi \qquad , \qquad y = Q + \sqrt{\frac{2J}{m\omega_c}}\,\cos\phi\ , \tag{3.323}$$

with $\omega_c = eB/mc$, the cyclotron frequency. We now make a mixed canonical transformation, generated by

$$F = \phi J' + \Big(k_z z + \frac{k_\perp P}{m\omega_c} - \omega t\Big)K' - PQ'\ , \tag{3.324}$$

where the new sets of conjugate variables are $\{(\phi', J'), (Q', P'), (\psi', K')\}$. We then have

$$\phi' = \frac{\partial F}{\partial J'} = \phi \qquad\qquad J = \frac{\partial F}{\partial \phi} = J' \tag{3.325}$$

$$Q = -\frac{\partial F}{\partial P} = -\frac{k_\perp K'}{m\omega_c} + Q' \qquad\qquad P' = -\frac{\partial F}{\partial Q'} = P \tag{3.326}$$

$$\psi' = \frac{\partial F}{\partial K'} = k_z z + \frac{k_\perp P}{m\omega_c} - \omega t \qquad\qquad p_z = \frac{\partial F}{\partial z} = k_z K' \ . \tag{3.327}$$

The transformed Hamiltonian is

$$\begin{aligned} H' &= H + \frac{\partial F}{\partial t} \\ &= \omega_c J' + \frac{k_z^2}{2m} K'^2 - \omega K' + \epsilon\, e V_0 \cos\left(\psi' + k_\perp \sqrt{\frac{2J'}{m\omega_c}}\, \sin\phi'\right) \ . \end{aligned} \tag{3.328}$$

We will now drop primes and simply write $H = H_0 + \epsilon\, H_1$, with

$$H_0 = \omega_c J + \frac{k_z^2}{2m} K^2 - \omega K \tag{3.329}$$

$$H_1 = e V_0 \cos\left(\psi + k_\perp \sqrt{\frac{2J}{m\omega_c}}\, \sin\phi\right) \ . \tag{3.330}$$

When $\epsilon = 0$, the frequencies associated with the ϕ and ψ motion are

$$\omega_\phi^0 = \frac{\partial H_0}{\partial \phi} = \omega_c \quad , \quad \omega_\psi^0 = \frac{\partial H_0}{\partial \psi} = \frac{k_z^2 K}{m} - \omega = k_z v_z - \omega \ , \tag{3.331}$$

where $v_z = p_z/m$ is the z-component of the particle's velocity. Now let us solve eqn. 3.307:

$$\omega_\phi^0 \frac{\partial S_1}{\partial \phi} + \omega_\psi^0 \frac{\partial S_1}{\partial \psi} = \langle\, H_1 \,\rangle - H_1 \ . \tag{3.332}$$

This yields

$$\begin{aligned} \omega_c \frac{\partial S_1}{\partial \phi} + \left(\frac{k_z^2 K}{m} - \omega\right) \frac{\partial S_1}{\partial \psi} &= -eA_0 \cos\left(\psi + k_\perp \sqrt{\frac{2J}{m\omega_c}}\, \sin\phi\right) \\ &= -eA_0 \sum_{n=-\infty}^{\infty} J_n\!\left(k_\perp \sqrt{\frac{2J}{m\omega_c}}\right) \cos(\psi + n\phi) \ , \end{aligned} \tag{3.333}$$

where we have used the result

$$e^{iz\sin\theta} = \sum_{n=-\infty}^{\infty} J_n(z)\, e^{in\theta} \ . \tag{3.334}$$

The solution for S_1 is

$$S_1 = \sum_n \frac{eV_0}{\omega - n\omega_c - k_z^2 \bar{K}/m}\, J_n\!\left(k_\perp \sqrt{\frac{2\bar{J}}{m\omega_c}}\right) \sin(\psi + n\phi) \ . \tag{3.335}$$

We then have new action variables $\bar{J}$ and $\bar{K}$, where

$$J = \bar{J} + \epsilon\,\frac{\partial S_1}{\partial\phi} + \mathcal{O}(\epsilon^2) \tag{3.336}$$

$$K = \bar{K} + \epsilon\,\frac{\partial S_1}{\partial\psi} + \mathcal{O}(\epsilon^2)\ . \tag{3.337}$$

Defining the dimensionless variable

$$\lambda \equiv k_\perp\sqrt{\frac{2J}{m\omega_c}}\ , \tag{3.338}$$

we obtain the result

$$\left(\frac{m\omega_c^2}{2eV_0k_\perp^2}\right)\bar{\lambda}^2 = \left(\frac{m\omega_c^2}{2eV_0k_\perp^2}\right)\lambda^2 - \epsilon\sum_n \frac{nJ_n(\lambda)\cos(\psi+n\phi)}{\frac{\omega}{\omega_c} - n - \frac{k_z^2K}{m\omega_c}} + \mathcal{O}(\epsilon^2)\ , \tag{3.339}$$

where $\bar{\lambda} = k_\perp\sqrt{2\bar{J}/m\omega_c}$.[14]

We see that resonances occur whenever

$$\frac{\omega}{\omega_c} - \frac{k_z^2K}{m\omega_c} = n\ , \tag{3.340}$$

for any integer n. Let us consider the case $k_z = 0$, in which the resonance condition is $\omega = n\omega_c$. We then have

$$\frac{\bar{\lambda}^2}{2\alpha} = \frac{\lambda^2}{2\alpha} - \sum_n \frac{n\,J_n(\lambda)\cos(\psi+n\phi)}{\frac{\omega}{\omega_c} - n}\ , \tag{3.341}$$

where

$$\alpha = \frac{E_0}{B}\cdot\frac{ck_\perp}{\omega_c} \tag{3.342}$$

is a dimensionless measure of the strength of the perturbation, with $E_0 \equiv k_\perp V_0$. In Fig. 3.9 we plot the level sets for the RHS of the above equation $\lambda(\psi)$ for $\phi = 0$, for two different values of the dimensionless amplitude α, for $\omega/\omega_c = 30.11$ (*i.e.* off resonance). Thus, when the amplitude is small, the level sets are far from a primary resonance, and the analytical and numerical results are very similar (left panels). When the amplitude is larger, resonances may occur which are not found in the lowest order perturbation treatment. However, as is apparent from the plots, the gross features of the phase diagram are reproduced by perturbation theory. What is missing is the existence of 'chaotic islands' which initially emerge in the vicinity of the trapping regions.

[14]Note that the argument of J_n in eqn. 3.339 is λ and not $\bar{\lambda}$. This arises because we are computing the new action $\bar{J}$ in terms of the old variables (ϕ, J) and (ψ, K).

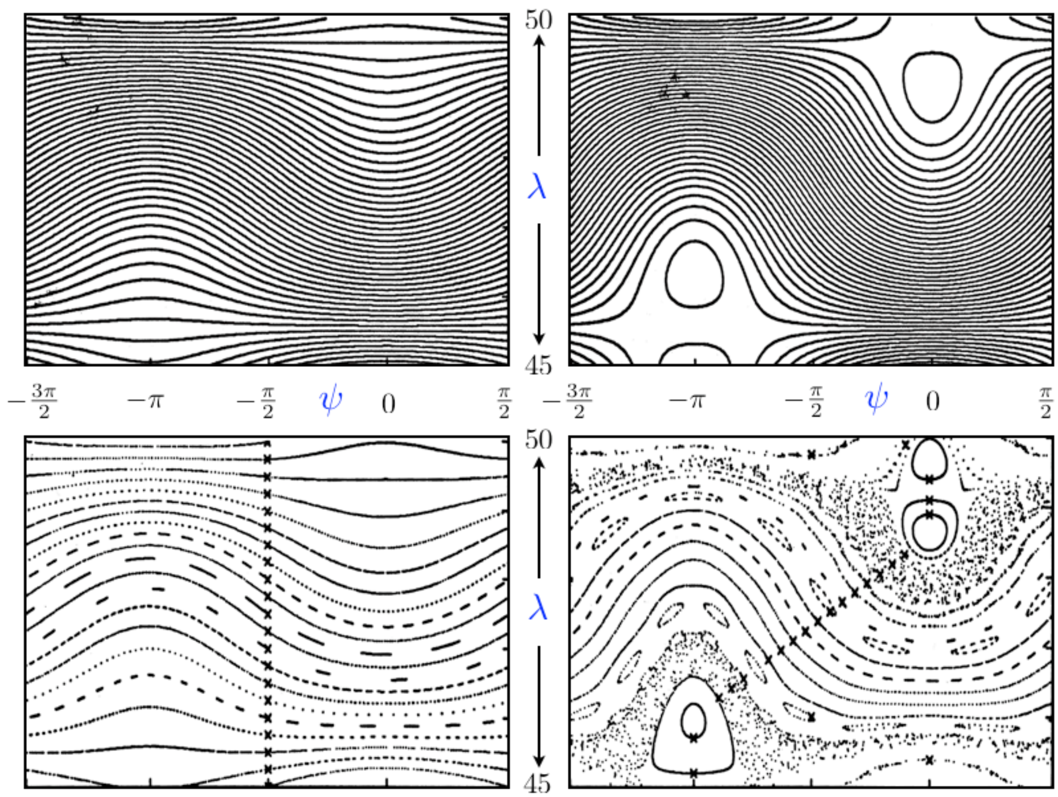

Figure 3.9: Plot of λ *versus* ψ for $\phi = 0$ (Poincaré section) for $\omega = 30.11\,\omega_c$ Top panels are nonresonant invariant curves calculated to first order. Bottom panels are exact numerical dynamics, with `x` symbols marking the initial conditions. Left panels: weak amplitude (no trapping). Right panels: stronger amplitude (shows trapping). From Lichtenberg and Lieberman (1983).

3.11 Adiabatic Invariants

Adiabatic perturbations are slow, smooth, time-dependent perturbations to a dynamical system. A classic example: a pendulum with a slowly varying length $l(t)$. Suppose $\lambda(t)$ is the adiabatic parameter. We write $H = H\big(q, p; \lambda(t)\big)$. All explicit time-dependence to H comes through $\lambda(t)$. Typically, a dimensionless parameter ϵ may be associated with the perturbation:

$$\epsilon = \frac{1}{\omega_0}\left|\frac{d\ln\lambda}{dt}\right| , \tag{3.343}$$

where ω_0 is the natural frequency of the system when λ is constant. We require $\epsilon \ll 1$ for adiabaticity.

In adiabatic processes, the action variables are conserved to a high degree of accuracy. These are the *adiabatic invariants*. For example, for the harmonix oscillator, the action is $J = E/\nu$. While E and ν may vary considerably during the adiabatic process, their ratio is very nearly fixed. As a consequence, assuming small oscillations,

$$E = \nu J = \tfrac{1}{2} mgl\,\theta_0^2 \quad \Rightarrow \quad \theta_0(l) \approx \frac{2J}{m\sqrt{g}\,l^{3/2}} \ , \tag{3.344}$$

so $\theta_0(\ell) \propto l^{-3/4}$.

Suppose that for fixed λ the Hamiltonian is transformed to action-angle variables via the generator $S(q, J; \lambda)$. The transformed Hamiltonian is

$$\tilde{H}(\phi, J, t) = H(\phi, J; \lambda) + \frac{\partial S}{\partial \lambda}\,\dot{\lambda} \ , \tag{3.345}$$

where

$$H(\phi, J; \lambda) = H\big(q(\phi, J; \lambda), p(\phi, J; \lambda); \lambda\big) \ . \tag{3.346}$$

We assume $n = 1$ here. Hamilton's equations are now

$$\dot{\phi} = +\frac{\partial \tilde{H}}{\partial J} = \nu(J; \lambda) + \frac{\partial^2 S}{\partial \lambda\,\partial J}\,\dot{\lambda} \tag{3.347}$$

$$\dot{J} = -\frac{\partial \tilde{H}}{\partial \phi} = -\frac{\partial^2 S}{\partial \lambda\,\partial \phi}\,\dot{\lambda} \ . \tag{3.348}$$

The second of these may be Fourier decomposed as

$$\dot{J} = -i\dot{\lambda} \sum_m m\,\frac{\partial S_m(J; \lambda)}{\partial \lambda}\,e^{im\phi} \ , \tag{3.349}$$

hence

$$\Delta J = J(t = +\infty) - J(t = -\infty) = -i \sum_m m \int\limits_{-\infty}^{\infty} \! dt\, \frac{\partial S_m(J; \lambda)}{\partial \lambda}\,\dot{\lambda}\,e^{im\phi} \ . \tag{3.350}$$

Since $\dot{\lambda}$ is small, we have $\phi(t) = \nu\,t + \beta$, to lowest order. We must therefore evaluate integrals such as

$$\mathcal{I} = \int\limits_{-\infty}^{\infty} \! dt \left\{ \frac{\partial S_m(J; \lambda)}{\partial \lambda}\,\dot{\lambda} \right\} e^{im\nu t} \ . \tag{3.351}$$

The term in curly brackets is a smooth, slowly varying function of t. Call it $f(t)$. We presume $f(t)$ can be analytically continued off the real t axis, and that its closest singularity in the complex t plane lies at $t = \pm i\tau$, in which case $\mathcal{I}$ behaves as $\exp(-|m|\nu\tau)$. Consider, for example, the Lorentzian,

$$f(t) = \frac{\mathcal{C}}{1 + (t/\tau)^2} \quad \Rightarrow \quad \int\limits_{-\infty}^{\infty} \! dt\, f(t)\,e^{im\nu t} = \pi\tau\,e^{-|m|\nu\tau} \ , \tag{3.352}$$

which is exponentially small in the time scale τ. Because of this, only $m = \pm 1$ need be considered. What this tells us is that the change ΔJ may be made arbitrarily small by a sufficiently slowly varying $\lambda(t)$.

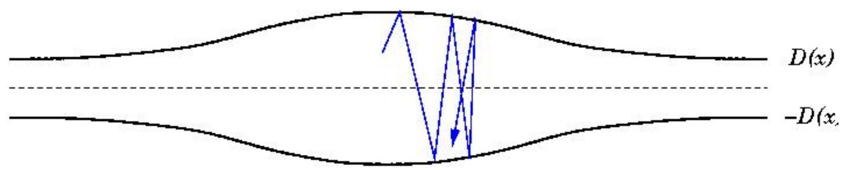

Figure 3.10: A mechanical mirror.

3.11.1 Example: Mechanical Mirror

Consider a two-dimensional version of a mechanical mirror, depicted in fig. 3.10. A particle bounces between two curves, $y = \pm D(x)$, where $|D'(x)| << 1$. The bounce time is $\tau_{\mathrm{b}\perp} = 2D/v_y$. We assume $\tau \ll L/v_x$, where $v_{x,y}$ are the components of the particle's velocity, and L is the total length of the system. There are, therefore, many bounces, which means the particle gets to sample the curvature in $D(x)$.

The adiabatic invariant is the action,

$$J = \frac{1}{2\pi} \int\limits_{-D}^{D} dy\, m\, v_y + \frac{1}{2\pi} \int\limits_{D}^{-D} dy\, m\,(-v_y) = \frac{2}{\pi}\, m v_y\, D(x) \ . \tag{3.353}$$

Thus,

$$E = \tfrac{1}{2} m \big(v_x^2 + v_y^2\big) = \tfrac{1}{2} m v_x^2 + \frac{\pi^2 J^2}{8 m D^2(x)} \ , \tag{3.354}$$

or

$$v_x^2 = \frac{2E}{m} - \left(\frac{\pi J}{2mD(x)}\right)^2 \ . \tag{3.355}$$

This means that the particle is reflected in the throat of the device at horizontal coordinate x^* such that

$$D(x^*) = \frac{\pi J}{\sqrt{8mE}} \ . \tag{3.356}$$

3.11.2 Example: Magnetic Mirror

Consider a particle of charge e moving in the presence of a uniform magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{z}}$. Recall the basic physics: velocity in the parallel direction v_z is conserved, while in the plane perpendicular to $\boldsymbol{B}$ the particle executes circular 'cyclotron orbits', satisfying

$$\frac{m v_\perp^2}{\rho} = \frac{e}{c}\, v_\perp B \quad \Rightarrow \quad \rho = \frac{m c v_\perp}{eB} \ , \tag{3.357}$$

where ρ is the radial coordinate in the plane perpendicular to $\boldsymbol{B}$. The period of the orbits is $T = 2\pi\rho . v_\perp = 2\pi mc/eB$, hence their frequency is $\omega_\mathrm{c} = eB/mc$, known as the *cyclotron frequency*.

Now assume that the magnetic field is spatially dependent. Note that a spatially varying $\boldsymbol{B}$-field cannot be unidirectional:

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = \boldsymbol{\nabla}_\perp \cdot \boldsymbol{B}_\perp + \frac{\partial B_z}{\partial z} = 0 \ . \tag{3.358}$$

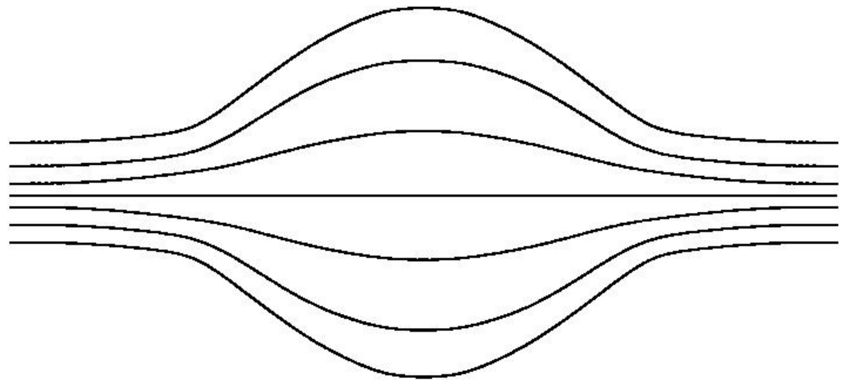

Figure 3.11: $\boldsymbol{B}$ field lines in a magnetic bottle.

The non-collinear nature of $\boldsymbol{B}$ results in the *drift* of the cyclotron orbits. Nevertheless, if the field $\boldsymbol{B}$ felt by the particle varies slowly on the time scale $T = 2\pi/\omega_c$, then the system possesses an adiabatic invariant:

$$J = \frac{1}{2\pi} \oint_{\mathcal{C}} \boldsymbol{p} \cdot d\boldsymbol{\ell} = \frac{1}{2\pi} \oint_{\mathcal{C}} \big(m\boldsymbol{v} + \tfrac{e}{c}\,\boldsymbol{A}\big) \cdot d\boldsymbol{\ell} \tag{3.359}$$

$$= \frac{m}{2\pi} \oint_{\mathcal{C}} \boldsymbol{v} \cdot d\boldsymbol{\ell} + \frac{e}{2\pi c} \oint_{\text{int}(\mathcal{C})} \boldsymbol{B} \cdot \hat{\boldsymbol{n}}\, d\Sigma \ . \tag{3.360}$$

The last two terms are of opposite sign, and one has

$$J = -\frac{m}{2\pi} \cdot \frac{\rho e B_z}{mc} \cdot 2\pi\rho + \frac{e}{2\pi c} \cdot B_z \cdot \pi\rho^2 \tag{3.361}$$

$$= -\frac{eB_z\rho^2}{2c} = -\frac{e}{2\pi c} \cdot \Phi_B(\mathcal{C}) = -\frac{m^2 v_\perp^2 c}{2eB_z} \ , \tag{3.362}$$

where $\Phi_B(\mathcal{C})$ is the magnetic flux enclosed by $\mathcal{C}$.

The energy is

$$E = \tfrac{1}{2} m v_\perp^2 + \tfrac{1}{2} m v_z^2 \ , \tag{3.363}$$

hence we have

$$v_z = \sqrt{\frac{2}{m}\big(E - MB\big)} \ . \tag{3.364}$$

where

$$M \equiv -\frac{e}{mc}\, J = \frac{e^2}{2\pi mc^2}\, \Phi_{\rm B}(\mathcal{C}) \tag{3.365}$$

is the *magnetic moment*. Note that v_z vanishes when $B = B_{\rm max} = E/M$. When this limit is reached, the particle turns around. This is the physics of the *magnetic mirror*.

A pair of magnetic mirrors may be used to confine charged particles in a *magnetic bottle*, depicted in fig. 3.11.

Let $v_{\parallel,0}$, $\boldsymbol{v}_{\perp,0}$, and $B_{\parallel,0}$ be the longitudinal particle velocity, transverse particle velocity, and longitudinal component of the magnetic field, respectively, at the point of injection. Our two conservation laws (J and E) guarantee

$$v_{\parallel}^2(z) + v_{\perp}^2(z) = v_{\parallel,0}^2 + v_{\perp,0}^2 \tag{3.366}$$

$$\frac{v_{\perp}(z)^2}{B_{\parallel}(z)} = \frac{v_{\perp,0}^2}{B_{\parallel,0}} \ . \tag{3.367}$$

This leads to reflection at a longitudinal coordinate z^*, where

$$B_{\parallel}(z^*) = B_{\parallel,0}\sqrt{1 + \frac{v_{\parallel,0}^2}{v_{\perp,0}^2}} \ . \tag{3.368}$$

The physics is quite similar to that of the mechanical mirror.

3.11.3 Resonances

When $n > 1$, we have

$$\dot{J}^{\alpha} = -i\dot{\lambda}\sum_{\boldsymbol{m}} m^{\alpha}\,\frac{\partial S_{\boldsymbol{m}}(J;\lambda)}{\partial\lambda}\,e^{i\boldsymbol{m}\cdot\boldsymbol{\phi}} \tag{3.369}$$

$$\Delta J = -i\sum_{\boldsymbol{m}} m^{\alpha}\int\limits_{-\infty}^{\infty}\!dt\,\frac{\partial S_{\boldsymbol{m}}(J;\lambda)}{\partial\lambda}\,\dot{\lambda}\;e^{i\boldsymbol{m}\cdot\boldsymbol{\nu}t}\,e^{i\boldsymbol{m}\cdot\boldsymbol{\beta}} \ . \tag{3.370}$$

Therefore, when $\boldsymbol{m}\cdot\boldsymbol{\nu}(J) = 0$ we have a resonance, and the integral grows linearly with time – a violation of the adiabatic invariance of J^{α}.

3.12 Fast Perturbations : Rapidly Oscillating Fields

Consider a free particle moving under the influence of an oscillating force,

$$m\ddot{q} = F\sin\omega t \ . \tag{3.371}$$

The motion of the system is then

$$q(t) = q_{\rm h}(t) - \frac{F\,\sin\omega t}{m\omega^2} \ , \tag{3.372}$$

where $q_{\rm h}(t) = A + Bt$ is the solution to the homogeneous (unforced) equation of motion. Note that the amplitude of the response $q - q_{\rm h}$ goes as ω^{-2} and is therefore small when ω is large.

Now consider a general $n = 1$ system, with

$$H(q,p,t) = H_0(q,p) + V_1(q)\,\sin(\omega t + \delta) \ . \tag{3.373}$$

We assume that ω is much greater than any natural oscillation frequency associated with H_0. We separate the motion $q(t)$ and $p(t)$ into slow and fast components:

$$q(t) = \bar{q}(t) + \zeta(t) \tag{3.374}$$

$$p(t) = \bar{p}(t) + \pi(t) \ , \tag{3.375}$$

where $\zeta(t)$ and $\pi(t)$ oscillate with the driving frequency ω. Since ζ and π will be small, we expand Hamilton's equations in these quantities:

$$\dot{\bar{q}} + \dot{\zeta} = \frac{\partial H_0}{\partial \bar{p}} + \frac{\partial^2 H_0}{\partial \bar{p}^2}\,\pi + \frac{\partial^2 H_0}{\partial \bar{q}\,\partial \bar{p}}\,\zeta + \frac{1}{2}\,\frac{\partial^3 H_0}{\partial \bar{q}^2\,\partial \bar{p}}\,\zeta^2 + \frac{\partial^3 H_0}{\partial \bar{q}\,\partial \bar{p}^2}\,\zeta\pi + \frac{1}{2}\,\frac{\partial^3 H_0}{\partial \bar{p}^3}\,\pi^2 + \ldots \tag{3.376}$$

$$\begin{aligned}\dot{\bar{p}} + \dot{\pi} = -\frac{\partial H_0}{\partial \bar{q}} - \frac{\partial^2 H_0}{\partial \bar{q}^2}\,\zeta - \frac{\partial^2 H_0}{\partial \bar{q}\,\partial \bar{p}}\,\pi - \frac{1}{2}\,\frac{\partial^3 H_0}{\partial \bar{q}^3}\,\zeta^2 - \frac{\partial^3 H_0}{\partial \bar{q}^2\,\partial \bar{p}}\,\zeta\pi - \frac{1}{2}\,\frac{\partial^3 H_0}{\partial \bar{q}\,\partial \bar{p}^2}\,\pi^2 \\ - \frac{\partial V}{\partial \bar{q}}\,\sin(\omega t + \delta) - \frac{\partial^2 V}{\partial \bar{q}^2}\,\zeta\,\sin(\omega t + \delta) - \ldots \ .\end{aligned} \tag{3.377}$$

We now average over the fast degrees of freedom to obtain an equation of motion for the slow variables $\bar{q}$ and $\bar{p}$, which we here carry to lowest nontrivial order in averages of fluctuating quantities:

$$\dot{\bar{q}} = \frac{\partial H_0}{\partial \bar{p}} + \frac{1}{2}\,\frac{\partial^3 H_0}{\partial \bar{q}^2\,\partial \bar{p}}\,\big\langle \zeta^2 \big\rangle + \frac{\partial^3 H_0}{\partial \bar{q}\,\partial \bar{p}^2}\,\big\langle \zeta\pi \big\rangle + \frac{1}{2}\,\frac{\partial^3 H_0}{\partial \bar{p}^3}\,\big\langle \pi^2 \big\rangle \tag{3.378}$$

$$\dot{\bar{p}} = -\frac{\partial H_0}{\partial \bar{q}} - \frac{1}{2}\,\frac{\partial^3 H_0}{\partial \bar{q}^3}\,\big\langle \zeta^2 \big\rangle - \frac{\partial^3 H_0}{\partial \bar{q}^2\,\partial \bar{p}}\,\big\langle \zeta\pi \big\rangle - \frac{1}{2}\,\frac{\partial^3 H_0}{\partial \bar{q}\,\partial \bar{p}^2}\,\big\langle \pi^2 \big\rangle - \frac{\partial^2 V}{\partial \bar{q}^2}\,\big\langle \zeta\,\sin(\omega t + \delta) \big\rangle \ . \tag{3.379}$$

The fast degrees of freedom obey

$$\dot{\zeta} = \frac{\partial^2 H_0}{\partial \bar{q}\,\partial \bar{p}}\,\zeta + \frac{\partial^2 H_0}{\partial \bar{p}^2}\,\pi \tag{3.380}$$

$$\dot{\pi} = -\frac{\partial^2 H_0}{\partial \bar{q}^2}\,\zeta - \frac{\partial^2 H_0}{\partial \bar{q}\,\partial \bar{p}}\,\pi - \frac{\partial V}{\partial q}\,\sin(\omega t + \delta) \ . \tag{3.381}$$

Let us analyze the coupled equations[15]

$$\dot{\zeta} = A\,\zeta + B\,\pi \tag{3.382}$$

$$\dot{\pi} = -C\,\zeta - A\,\pi + F\,e^{-i\omega t} \ . \tag{3.383}$$

The solution is of the form

$$\begin{pmatrix} \zeta \\ \pi \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} e^{-i\omega t} \ . \tag{3.384}$$

Plugging in, we find

$$\alpha = \frac{BF}{BC - A^2 - \omega^2} = -\frac{BF}{\omega^2} + \mathcal{O}\big(\omega^{-4}\big) \tag{3.385}$$

$$\beta = -\frac{(A + i\omega)F}{BC - A^2 - \omega^2} = \frac{iF}{\omega} + \mathcal{O}\big(\omega^{-3}\big) \ . \tag{3.386}$$

[15]With real coefficients A, B, and C, one can always take the real part to recover the fast variable equations of motion.

Taking the real part, and restoring the phase shift δ, we have

$$\zeta(t) = \frac{-BF}{\omega^2}\,\sin(\omega t + \delta) = \frac{1}{\omega^2}\,\frac{\partial V}{\partial \bar{q}}\,\frac{\partial^2 H_0}{\partial \bar{p}^2}\,\sin(\omega t + \delta) \tag{3.387}$$

$$\pi(t) = -\frac{F}{\omega}\,\cos(\omega t + \delta) = \frac{1}{\omega}\,\frac{\partial V}{\partial \bar{q}}\,\cos(\omega t + \delta)\ . \tag{3.388}$$

The desired averages, to lowest order, are thus

$$\big\langle \zeta^2 \big\rangle = \frac{1}{2\omega^4}\left(\frac{\partial V}{\partial \bar{q}}\right)^2 \left(\frac{\partial^2 H_0}{\partial \bar{p}^2}\right)^2 \tag{3.389}$$

$$\big\langle \pi^2 \big\rangle = \frac{1}{2\omega^2}\left(\frac{\partial V}{\partial \bar{q}}\right)^2 \tag{3.390}$$

$$\big\langle \zeta\,\sin(\omega t + \delta)\big\rangle = \frac{1}{2\omega^2}\,\frac{\partial V}{\partial \bar{q}}\,\frac{\partial^2 H_0}{\partial \bar{p}^2}\ , \tag{3.391}$$

along with $\big\langle \zeta\pi \big\rangle = 0$.

Finally, we substitute the averages into the equations of motion for the slow variables $\bar{q}$ and $\bar{p}$, resulting in the time-independent *effective Hamiltonian*

$$K(\bar{q},\bar{p}) = H_0(\bar{q},\bar{p}) + \frac{1}{4\omega^2}\,\frac{\partial^2 H_0}{\partial \bar{p}^2}\left(\frac{\partial V}{\partial \bar{q}}\right)^2\ , \tag{3.392}$$

and the equations of motion

$$\dot{\bar{q}} = \frac{\partial K}{\partial \bar{p}} \quad , \quad \dot{\bar{p}} = -\frac{\partial K}{\partial \bar{q}}\ . \tag{3.393}$$

3.12.1 Example : Pendulum with Oscillating Support

Consider a pendulum with a vertically oscillating point of support. The coordinates of the pendulum bob are

$$x = \ell\,\sin\theta \quad , \quad y = a(t) - \ell\,\cos\theta\ . \tag{3.394}$$

The Lagrangian is easily obtained:

$$L = \tfrac{1}{2}m\ell^2\,\dot\theta^2 + m\ell\dot a\,\dot\theta\sin\theta + mg\ell\cos\theta + \tfrac{1}{2}m\dot a^2 - mga \tag{3.395}$$

$$= \tfrac{1}{2}m\ell^2\,\dot\theta^2 + m(g+\ddot a)\ell\cos\theta + \overbrace{\tfrac{1}{2}m\dot a^2 - mga - \frac{d}{dt}\big(m\ell\dot a\,\sin\theta\big)}^{\text{these may be dropped}}\ . \tag{3.396}$$

Thus we may take the Lagrangian to be

$$\bar L = \tfrac{1}{2}m\ell^2\,\dot\theta^2 + m(g+\ddot a)\ell\cos\theta\ , \tag{3.397}$$

from which we derive the Hamiltonian

$$H(\theta,p_\theta,t) = \frac{p_\theta^2}{2m\ell^2} - mg\ell\cos\theta - m\ell\ddot a\,\cos\theta \tag{3.398}$$

$$= H_0(\theta,p_\theta,t) + V_1(\theta)\,\sin\omega t\ . \tag{3.399}$$

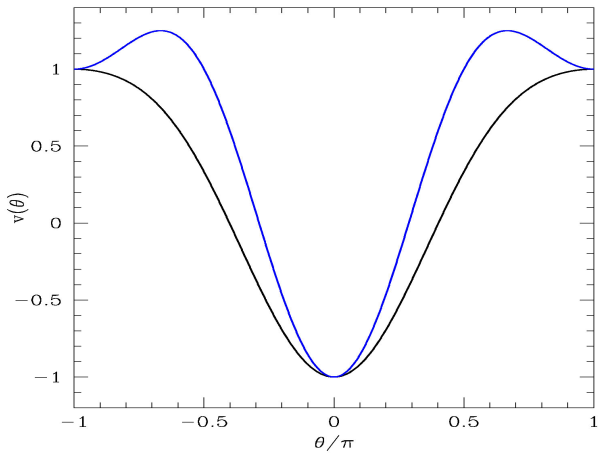

Figure 3.12: Dimensionless potential $v(\theta)$ for $\epsilon = 1.5$ (black curve) and $\epsilon = 0.5$ (blue curve).

We have assumed $a(t) = a_0 \sin \omega t$, so

$$V_1(\theta) = m\ell a_0\, \omega^2 \cos\theta \ . \tag{3.400}$$

The effective Hamiltonian, per eqn. 3.392, is

$$K(\bar{\theta}, \bar{p}_\theta) = \frac{\bar{p}_\theta}{2m\ell^2} - mg\ell \cos\bar{\theta} + \tfrac{1}{4} m\, a_0^2\, \omega^2 \sin^2\bar{\theta} \ . \tag{3.401}$$

Let's define the dimensionless parameter

$$\epsilon \equiv \frac{2g\ell}{\omega^2 a_0^2} \ . \tag{3.402}$$

The slow variable $\bar{\theta}$ executes motion in the *effective potential* $V_{\text{eff}}(\bar{\theta}) = mg\ell\, v(\bar{\theta})$, with

$$v(\bar{\theta}) = -\cos\bar{\theta} + \frac{1}{2\epsilon} \sin^2\bar{\theta} \ . \tag{3.403}$$

Differentiating, and dropping the bar on θ, we find that $V_{\text{eff}}(\theta)$ is stationary when

$$v'(\theta) = 0 \quad \Rightarrow \quad \sin\theta\cos\theta = -\epsilon\sin\theta \ . \tag{3.404}$$

Thus, $\theta = 0$ and $\theta = \pi$, where $\sin\theta = 0$, are equilibria. When $\epsilon < 1$ (note $\epsilon > 0$ always), there are two new solutions, given by the roots of $\cos\theta = -\epsilon$.

To assess stability of these equilibria, we compute the second derivative:

$$v''(\theta) = \cos\theta + \frac{1}{\epsilon} \cos 2\theta \ . \tag{3.405}$$

From this, we see that $\theta = 0$ is stable (*i.e.* $v''(\theta = 0) > 0$) always, but $\theta = \pi$ is stable for $\epsilon < 1$ and unstable for $\epsilon > 1$. When $\epsilon < 1$, two new solutions appear, at $\cos\theta = -\epsilon$, for which

$$v''(\cos^{-1}(-\epsilon)) = \epsilon - \frac{1}{\epsilon} \ , \tag{3.406}$$

which is always negative since $\epsilon < 1$ in order for these equilibria to exist. The situation is sketched in fig. 3.12, showing $v(\theta)$ for two representative values of the parameter ϵ. For $\epsilon > 1$, the equilibrium at $\theta = \pi$ is unstable, but as ϵ decreases, a subcritical pitchfork bifurcation is encountered at $\epsilon = 1$, and $\theta = \pi$ becomes stable, while the outlying $\theta = \cos^{-1}(-\epsilon)$ solutions are unstable.