Contents

List of Figures

List of Tables

Chapter 1

Hamiltonian Mechanics

1.1 References

– R. Z. Sagdeev, D. A. Usikov, and G. M. Zaslavsky, *Nonlinear Physics* (Harwood, 1988)
A thorough treatment of nonlinear Hamiltonian particle and wave mechanics.

– E. Ott, *Chaos in Dynamical Systems* (Cambridge, 2002)
An excellent introductory text appropriate for graduate or advanced undergraduate students.

– W. Dittrich and M. Reuter, *Classical and Quantum Dynamics* (Springer, 2001)
More a handbook than a textbook, but reliably covers a large amount of useful material.

– G. M. Zaslavsky, *Hamiltonian Chaos & Fractional Dynamics* (Oxford, 2005)
An advanced text for graduate students and researchers.

– I. Percival and D. Richards, *Introduction to Dynamics* (Cambridge, 1994)
An excellent advanced undergraduate text.

– A. J. Lichenberg and M. A. Lieberman, *Regular and Stochastic Motion* (Springer, 1983)
An advanced graduate level text. Excellent range of topics, but quite technical and often lacking physical explanations.

1.2 The Hamiltonian

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

$$p_\sigma = \frac{\partial L}{\partial \dot{q}_\sigma} \quad , \tag{1.1}$$

with $n = Nd$ for a system of N particles in d space dimensions. The Hamiltonian, $H(q, p, t)$ is obtained by a Legendre transformation,

$$H(q,p) = \sum_{\sigma=1}^{n} p_\sigma \, \dot{q}_\sigma - L(q, \dot{q}, t) \quad . \tag{1.2}$$

Note that

$$\begin{aligned} dH &= \sum_{\sigma=1}^{n} \left(p_\sigma \, d\dot{q}_\sigma + \dot{q}_\sigma \, dp_\sigma - \frac{\partial L}{\partial q_\sigma} \, dq_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma} \, d\dot{q}_\sigma \right) - \frac{\partial L}{\partial t} \, dt \\ &= \sum_{\sigma=1}^{n} \left(\dot{q}_\sigma \, dp_\sigma - \frac{\partial L}{\partial q_\sigma} \, dq_\sigma \right) - \frac{\partial L}{\partial t} \, dt \quad . \end{aligned} \tag{1.3}$$

Thus, we obtain Hamilton's equations of motion,

$$\frac{\partial H}{\partial p_\sigma} = \dot{q}_\sigma \qquad , \qquad \frac{\partial H}{\partial q_\sigma} = -\frac{\partial L}{\partial q_\sigma} = -\dot{p}_\sigma \tag{1.4}$$

and

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \quad . \tag{1.5}$$

Some remarks:

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

  $$L = \tfrac{1}{2} m \left(\dot{r}^2 + r^2 \, \dot{\theta}^2 + r^2 \sin^2\!\theta \, \dot{\phi}^2 \right) - U(r, \theta, \phi) \quad . \tag{1.6}$$

  We have

  $$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \quad , \quad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2 \, \dot{\theta} \quad , \quad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2 \, \sin^2\!\theta \, \dot{\phi} \quad , \tag{1.7}$$

  and thus

  $$\begin{aligned} H &= p_r \, \dot{r} + p_\theta \, \dot{\theta} + p_\phi \, \dot{\phi} - L \\ &= \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2 \sin^2\!\theta} + U(r, \theta, \phi) \quad . \end{aligned} \tag{1.8}$$

  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,

$$\frac{\partial p_\alpha}{\partial \dot{q}_\beta} = \frac{\partial^2 L}{\partial \dot{q}_\alpha \, \partial \dot{q}_\beta} \tag{1.9}$$

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

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

$$\xi_i = \begin{cases} q_i & \text{if } 1 \leq i \leq n \\ p_{i-n} & \text{if } n < i \leq 2n \end{cases} \quad . \tag{1.10}$$

  Then we may write Hamilton's equations compactly as

$$\dot{\xi}_i = \mathbb{J}_{ij} \, \frac{\partial H}{\partial \xi_j} \quad , \tag{1.11}$$

  where

$$\mathbb{J} = \begin{pmatrix} \mathbb{O}_{n\times n} & \mathbb{I}_{n\times n} \\ -\mathbb{I}_{n\times n} & \mathbb{O}_{n\times n} \end{pmatrix} \tag{1.12}$$

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

1.2.1 Modified Hamilton's principle

Let's vary the action now with respect to *both* $\{q_\sigma\}$ *and* $\{p_\sigma\}$, considering them as *independent variations*. We then have

$$\begin{aligned} 0 = \delta \int\limits_{t_a}^{t_b} \! dt \, L &= \delta \int\limits_{t_a}^{t_b} \! dt \, \big(p_\sigma \, \dot{q}_\sigma - H \big) \\ &= \int\limits_{t_a}^{t_b} \! dt \left\{ p_\sigma \, \delta\dot{q}_\sigma + \dot{q}_\sigma \, \delta p_\sigma - \frac{\partial H}{\partial q_\sigma} \, \delta q_\sigma - \frac{\partial H}{\partial p_\sigma} \, \delta p_\sigma \right\} \\ &= \int\limits_{t_a}^{t_b} \! dt \left\{ -\left(\dot{p}_\sigma + \frac{\partial H}{\partial q_\sigma} \right) \delta q_\sigma + \left(\dot{q}_\sigma - \frac{\partial H}{\partial p_\sigma} \right) \delta p_\sigma \right\} + \big(p_\sigma \, \delta q_\sigma \big) \Big|_{t_a}^{t_b} \quad . \end{aligned} \tag{1.13}$$

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

1.2.2 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:

$$\begin{aligned}\boldsymbol{\nabla}\cdot\dot{\boldsymbol{\xi}}&=\sum_{\sigma=1}^{n}\bigg\{\frac{\partial\dot{q}_\sigma}{\partial q_\sigma}+\frac{\partial\dot{p}_\sigma}{\partial p_\sigma}\bigg\}\\&=\sum_{i=1}^{2n}\frac{\partial\dot{\xi}_i}{\partial\xi_i}=\sum_{i,j}\mathbb{J}_{ij}\,\frac{\partial^2 H}{\partial\xi_i\,\partial\xi_j}=0\quad.\end{aligned}\tag{1.14}$$

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

$$\frac{\partial\rho}{\partial t}+\boldsymbol{\nabla}\cdot(\rho\,\dot{\boldsymbol{\xi}})=0\quad.\tag{1.15}$$

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

$$\frac{D\rho}{Dt}=\frac{\partial\rho}{\partial t}+\dot{\boldsymbol{\xi}}\cdot\boldsymbol{\nabla}\rho=0\quad,\tag{1.16}$$

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.

1.2.3 Poincaré recurrence theorem

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

$$\dot{q}_\sigma=+\frac{\partial H}{\partial p_\sigma}\qquad,\qquad\dot{p}_\sigma=-\frac{\partial H}{\partial q_\sigma}\tag{1.17}$$

for a time interval $\Delta t=\tau$. Consider a region Ω in phase space. Define $g_\tau^n\Omega$ to be the $n^{\rm th}$ image of Ω 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,

$$\frac{\partial\varrho}{\partial t}+\nabla\cdot(\boldsymbol{u}\varrho)=0\tag{1.18}$$

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.* $\nabla\cdot\boldsymbol{u}=0$:

$$\begin{aligned}\nabla\cdot\boldsymbol{u}&=\sum_{\sigma=1}^{n}\bigg\{\frac{\partial\dot{q}_\sigma}{\partial q_\sigma}+\frac{\partial\dot{p}_\sigma}{\partial p_\sigma}\bigg\}\\&=\sum_{\sigma=1}^{n}\bigg\{\frac{\partial}{\partial q_\sigma}\bigg(\frac{\partial H}{\partial p_\sigma}\bigg)+\frac{\partial}{\partial p_\sigma}\bigg(-\frac{\partial H}{\partial q_\sigma}\bigg)\bigg\}=0\quad.\end{aligned}\tag{1.19}$$

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

$$\frac{D\varrho}{Dt} \equiv \frac{\partial\varrho}{\partial t} + \boldsymbol{u}\cdot\nabla\varrho = 0 \quad , \tag{1.20}$$

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.*

$$\int\! d\mu\, \delta\big(E - H(\boldsymbol{q},\boldsymbol{p})\big) < \infty \quad , \tag{1.21}$$

where $d\mu = \prod_i dq_i\, dp_i$ 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 $g^m_\tau\,\Omega$ for all m:

$$\Upsilon = \bigcup_{m=0}^{\infty} g^m_\tau\,\Omega \tag{1.22}$$

We assume that the set $\{g^m_\tau\,\Omega \,|\, m \in \mathbb{Z}_{\ge 0}\}$ is disjoint. The volume of a union of disjoint sets is the sum of the individual volumes. Thus,

$$\mathsf{vol}(\Upsilon) = \sum_{m=0}^{\infty} \mathsf{vol}(g^m_\tau\,\Omega) = \mathsf{vol}(\Omega)\cdot\sum_{m=1}^{\infty} 1 = \infty \quad , \tag{1.23}$$

since $\mathsf{vol}(g^m_\tau\,\Omega) = \mathsf{vol}(\Omega)$ 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 $\{g^m_\tau\,\Omega \,|\, m \in \mathbb{Z} \quad , m \ge 0\}$ is disjoint fails. This means that there exists some pair of integers k and l, with $k \ne l$, such that $g^k_\tau\,\Omega \cap g^l_\tau\,\Omega \ne \emptyset$. Without loss of generality we may assume $k > l$. Apply the inverse g^{-1}_τ to this relation l times to get $g^{k-l}_\tau\,\Omega \cap \Omega \ne \emptyset$. Now choose any point $\boldsymbol{\varphi} \in g^n_\tau\,\Omega \cap \Omega$, where $n = k - l$, and define $\boldsymbol{\varphi}_0 = g^{-n}_\tau \boldsymbol{\varphi}$. Then by construction both $\boldsymbol{\varphi}_0$ and $g^n_\tau\,\boldsymbol{\varphi}_0$ 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 $f\colon \mathbb{R} \to \mathbb{R}$ 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 is equal to $D_t\varrho=-(\nabla\cdot\boldsymbol{u})\varrho=+2\beta\varrho$ 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 $\mathsf{vol}(g_\tau^n\,\Omega)=e^{-2n\beta\tau}\,\mathsf{vol}(\Omega)$. A damped pendulum, released from rest at some small angle θ_0, will not return arbitrarily close to these initial conditions.

1.2.4 Poisson brackets

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

$$\begin{aligned}\frac{d}{dt}\,F\big(\boldsymbol{q}(t),\boldsymbol{p}(t),t\big)&=\frac{\partial F}{\partial t}+\sum_{\sigma=1}^{n}\bigg\{\frac{\partial F}{\partial q_\sigma}\,\dot{q}_\sigma+\frac{\partial F}{\partial p_\sigma}\,\dot{p}_\sigma\bigg\}\\&\equiv\frac{\partial F}{\partial t}+\big\{F,H\big\}\quad,\end{aligned}\tag{1.24}$$

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

$$\begin{aligned}\big\{A,B\big\}&\equiv\sum_{\sigma=1}^{n}\bigg(\frac{\partial A}{\partial q_\sigma}\,\frac{\partial B}{\partial p_\sigma}-\frac{\partial A}{\partial p_\sigma}\,\frac{\partial B}{\partial q_\sigma}\bigg)\\&=\sum_{i,j=1}^{2n}\mathbb{J}_{ij}\,\frac{\partial A}{\partial\xi_i}\,\frac{\partial B}{\partial\xi_j}\quad.\end{aligned}\tag{1.25}$$

Properties of the Poisson bracket:

- Antisymmetry:
$$\big\{f,g\big\}=-\big\{g,f\big\}\quad.\tag{1.26}$$
- Bilinearity: if λ is a constant, and f, g, and h are functions on phase space, then
$$\big\{f+\lambda\,g,h\big\}=\big\{f,h\big\}+\lambda\{g,h\big\}\quad.\tag{1.27}$$
Linearity in the second argument follows from this and the antisymmetry condition.
- Associativity:
$$\big\{fg,h\big\}=f\big\{g,h\big\}+g\big\{f,h\big\}\quad.\tag{1.28}$$
- Jacobi identity:
$$\big\{f,\{g,h\}\big\}+\big\{g,\{h,f\}\big\}+\big\{h,\{f,g\}\big\}=0\quad.\tag{1.29}$$

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

$$\{q_\alpha,q_\beta\}=0 \qquad , \qquad \{p_\alpha,p_\beta\}=0 \qquad , \qquad \{q_\alpha,p_\beta\}=\delta_{\alpha\beta} \quad . \tag{1.30}$$

1.3 Canonical Transformations

1.3.1 Point transformations in Lagrangian mechanics

In Lagrangian mechanics, we are free to redefine our generalized coordinates, *viz.*

$$Q_\sigma = Q_\sigma(q_1,\ldots,q_n,t) \quad . \tag{1.31}$$

This is called a "point transformation." The transformation is invertible if

$$\det\left(\frac{\partial Q_\alpha}{\partial q_\beta}\right) \neq 0 \quad . \tag{1.32}$$

The transformed Lagrangian, $\tilde{L}$, written as a function of the new coordinates $\boldsymbol{Q}$ and velocities $\dot{\boldsymbol{Q}}$, is

$$\tilde{L}(\boldsymbol{Q},\dot{\boldsymbol{Q}},t) = L\big(\boldsymbol{q}(\boldsymbol{Q},t),\dot{\boldsymbol{q}}(\boldsymbol{Q},\dot{\boldsymbol{Q}},t),t\big) + \frac{d}{dt}\,F\big(\boldsymbol{q}(\boldsymbol{Q},t),t\big) \quad , \tag{1.33}$$

where $F(\boldsymbol{q},t)$ is a function only of the coordinates $q_\sigma(\boldsymbol{Q},t)$ and time[1]. Finally, Hamilton's principle,

$$\delta\int\limits_{t_1}^{t_b} dt\,\tilde{L}(\boldsymbol{Q},\dot{\boldsymbol{Q}},t) = 0 \tag{1.34}$$

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

$$\frac{\partial\tilde{L}}{\partial Q_\sigma} - \frac{d}{dt}\left(\frac{\partial\tilde{L}}{\partial\dot{Q}_\sigma}\right) = 0 \quad . \tag{1.35}$$

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

$$\frac{d}{dt}\left(\frac{\partial\tilde{L}}{\partial\dot{Q}_\sigma}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_\alpha}\,\frac{\partial\dot{q}_\alpha}{\partial\dot{Q}_\sigma}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_\alpha}\,\frac{\partial q_\alpha}{\partial Q_\sigma}\right) \quad , \tag{1.36}$$

[1] We must have that the relation $Q_\sigma = Q_\sigma(\boldsymbol{q},t)$ is invertible.

where the relation $\partial\dot{q}_\alpha/\partial\dot{Q}_\sigma = \partial q_\alpha/\partial Q_\sigma$ follows from $\dot{q}_\alpha = \frac{\partial q_\alpha}{\partial Q_\sigma}\,\dot{Q}_\sigma + \frac{\partial q_\alpha}{\partial t}$. We know that adding a total time derivative of a function $\tilde{F}(\boldsymbol{Q},t) = F\big(\boldsymbol{q}(\boldsymbol{Q},t),t\big)$ to the Lagrangian does not alter the equations of motion. Hence we can set $F=0$ and compute

$$
\begin{aligned}
\frac{\partial\tilde{L}}{\partial Q_\sigma} &= \frac{\partial L}{\partial q_\alpha}\,\frac{\partial q_\alpha}{\partial Q_\sigma} + \frac{\partial L}{\partial \dot{q}_\alpha}\,\frac{\partial \dot{q}_\alpha}{\partial Q_\sigma} \\
&= \frac{\partial L}{\partial q_\alpha}\,\frac{\partial q_\alpha}{\partial Q_\sigma} + \frac{\partial L}{\partial \dot{q}_\alpha}\left(\frac{\partial^2 q_\alpha}{\partial Q_\sigma\,\partial Q_{\sigma'}}\,\dot{Q}_{\sigma'} + \frac{\partial^2 q_\alpha}{\partial Q_\sigma\,\partial t}\right) \\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right)\frac{\partial q_\alpha}{\partial Q_\sigma} + \frac{\partial L}{\partial \dot{q}_\alpha}\,\frac{d}{dt}\left(\frac{\partial q_\alpha}{\partial Q_\sigma}\right) \\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\,\frac{\partial q_\alpha}{\partial Q_\sigma}\right) = \frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{Q}_\sigma}\right) \quad ,
\end{aligned}
\tag{1.37}
$$

where the last equality is what we obtained earlier in eqn. 1.36.

1.3.2 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

$$
\begin{aligned}
q_\sigma &= q_\sigma\big(Q_1,\ldots,Q_n\,;P_1,\ldots,P_n;t\big) \\
p_\sigma &= p_\sigma\big(Q_1,\ldots,Q_n\,;P_1,\ldots,P_n;t\big) \quad ,
\end{aligned}
\tag{1.38}
$$

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

$$
\xi_i = \xi_i\big(\Xi_1,\ldots,\Xi_{2n};t\big) \quad ,
\tag{1.39}
$$

with $i \in \{1,\ldots,2n\}$. The transformed Hamiltonian is $\tilde{H}(\boldsymbol{Q},\boldsymbol{P},t)$., where, as we shall see below, $\tilde{H}(\boldsymbol{Q},\boldsymbol{P},t) = H(\boldsymbol{q},\boldsymbol{p},t) + \frac{\partial}{\partial t}F(\boldsymbol{q},\boldsymbol{Q},t)$.

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

$$
\dot{Q}_\sigma = \frac{\partial\widetilde{H}}{\partial P_\sigma} \quad , \quad \dot{P}_\sigma = -\frac{\partial\widetilde{H}}{\partial Q_\sigma} \quad ,
\tag{1.40}
$$

which gives

$$
\frac{\partial\dot{Q}_\sigma}{\partial Q_\sigma} + \frac{\partial\dot{P}_\sigma}{\partial P_\sigma} = 0 = \frac{\partial\dot{\Xi}_i}{\partial\Xi_i} \quad .
\tag{1.41}
$$

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.*

$$
\mathsf{det}\left(\frac{\partial\Xi_i}{\partial\xi_j}\right) = \left|\left|\frac{\partial(\boldsymbol{Q},\boldsymbol{P})}{\partial(\boldsymbol{q},\boldsymbol{p})}\right|\right| = 1 \quad .
\tag{1.42}
$$

Additional conditions will be discussed below.

1.3.3 Hamiltonian evolution

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

$$\xi_i(t + dt) = \xi_i(t) + \mathbb{J}_{ij}\, \frac{\partial H}{\partial \xi_j}\, dt + \mathcal{O}\big(dt^2\big) \quad . \tag{1.43}$$

Thus,

$$\begin{aligned} \frac{\partial \xi'_i}{\partial \xi_j} &= \frac{\partial}{\partial \xi_j} \bigg(\xi_i + \mathbb{J}_{ik}\, \frac{\partial H}{\partial \xi_k}\, dt + \mathcal{O}\big(dt^2\big) \bigg) \\ &= \delta_{ij} + \mathbb{J}_{ik}\, \frac{\partial^2 H}{\partial \xi_j\, \partial \xi_k}\, dt + \mathcal{O}\big(dt^2\big) \quad . \end{aligned} \tag{1.44}$$

Now, using the result $\mathsf{det}\big(1 + \epsilon M\big) = 1 + \epsilon\ \mathsf{Tr}\ M + \mathcal{O}(\epsilon^2)$, we have

$$\left|\left|\frac{\partial \xi'_i}{\partial \xi_j}\right|\right| = 1 + \mathbb{J}_{jk}\, \frac{\partial^2 H}{\partial \xi_j\, \partial \xi_k}\, dt + \mathcal{O}\big(dt^2\big) = 1 + \mathcal{O}\big(dt^2\big) \quad . \tag{1.45}$$

1.3.4 Symplectic structure

We have that

$$\dot{\xi}_i = \mathbb{J}_{ij}\, \frac{\partial H}{\partial \xi_j} \quad . \tag{1.46}$$

Suppose we make a time-independent canonical transformation from $\{\xi_i\}$ to new phase space coordinates $\{\varXi_a\}$, where $\varXi_a = \varXi_a(\xi)$. We then have

$$\dot{\varXi}_a = \frac{\partial \varXi_a}{\partial \xi_j}\, \dot{\xi}_j = \frac{\partial \varXi_a}{\partial \xi_j}\, \mathbb{J}_{jk}\, \frac{\partial H}{\partial \xi_k} \quad . \tag{1.47}$$

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

$$\dot{\varXi}_a = \mathbb{J}_{ab}\, \frac{\partial \widetilde{H}}{\partial \varXi_b} = \mathbb{J}_{ab}\, \frac{\partial H}{\partial \xi_k}\, \frac{\partial \xi_k}{\partial \varXi_b} \quad . \tag{1.48}$$

Equating these two expressions, we have

$$M_{aj}\, \mathbb{J}_{jk}\, \frac{\partial H}{\partial \xi_k} = \mathbb{J}_{ab}\, M^{-1}_{kb}\, \frac{\partial H}{\partial \xi_k} \quad , \tag{1.49}$$

where $M_{aj} \equiv \partial \varXi_a/\partial \xi_j$ is the Jacobian of the transformation. Since the equality must hold for all ξ, we conclude

$$M\mathbb{J} = \mathbb{J}\big(M^{\mathrm{t}}\big)^{-1} \quad \Longrightarrow \quad M\mathbb{J}M^{\mathrm{t}} = \mathbb{J} \quad . \tag{1.50}$$

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

[2]Note that the rank of a symplectic matrix is always even. Note also $M\mathbb{J}M^{\mathrm{t}} = \mathbb{J}$ implies $M^{\mathrm{t}}\mathbb{J}M = \mathbb{J}$.

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

$$\begin{aligned}\big\{A,B\big\}_\xi &= \mathbb{J}_{ij}\,\frac{\partial A}{\partial\xi_i}\,\frac{\partial B}{\partial\xi_j} = \mathbb{J}_{ij}\,\frac{\partial A}{\partial\Xi_a}\,\frac{\partial\Xi_a}{\partial\xi_i}\,\frac{\partial B}{\partial\Xi_b}\,\frac{\partial\Xi_b}{\partial\xi_j}\\ &= \big(M_{ai}\,\mathbb{J}_{ij}\,M^{\rm t}_{jb}\big)\,\frac{\partial A}{\partial\Xi_a}\,\frac{\partial B}{\partial\Xi_b} = \mathbb{J}_{ab}\,\frac{\partial A}{\partial\Xi_a}\,\frac{\partial B}{\partial\Xi_b} = \big\{A,B\big\}_\Xi \quad .\end{aligned} \tag{1.51}$$

1.3.5 Generating functions for canonical transformations

For a transformation to be canonical, we require

$$\delta\int\limits_{t_a}^{t_b}\!dt\,\Big\{p_\sigma\,\dot q_\sigma - H(\boldsymbol{q},\boldsymbol{p},t)\Big\} = 0 = \delta\int\limits_{t_a}^{t_b}\!dt\,\Big\{P_\sigma\,\dot Q_\sigma - \widetilde H(\boldsymbol{Q},\boldsymbol{P},t)\Big\} \quad . \tag{1.52}$$

This is satisfied provided

$$\Big\{p_\sigma\,\dot q_\sigma - H(\boldsymbol{q},\boldsymbol{p},t)\Big\} = \lambda\,\Big\{P_\sigma\,\dot Q_\sigma - \widetilde H(\boldsymbol{Q},\boldsymbol{P},t) + \frac{d}{dt}\,F(\boldsymbol{q},\boldsymbol{Q},t)\Big\} \quad , \tag{1.53}$$

where λ is a constant. For canonical transformations[3], $\lambda=1$. Thus,

$$\widetilde H(\boldsymbol{Q},\boldsymbol{P},t) = H(\boldsymbol{q},\boldsymbol{p},t) + P_\sigma\,\dot Q_\sigma - p_\sigma\,\dot q_\sigma + \frac{\partial F}{\partial q_\sigma}\,\dot q_\sigma + \frac{\partial F}{\partial Q_\sigma}\,\dot Q_\sigma + \frac{\partial F}{\partial t} \quad . \tag{1.54}$$

Thus, we require

$$\frac{\partial F}{\partial q_\sigma} = p_\sigma \quad , \quad \frac{\partial F}{\partial Q_\sigma} = -P_\sigma \quad , \quad \frac{\partial F}{\partial p_\sigma} = 0 \quad , \quad \frac{\partial F}{\partial P_\sigma} = 0 \quad , \tag{1.55}$$

which says that $F = F(\boldsymbol{q},\boldsymbol{Q},t)$ is only a function of $(\boldsymbol{q},\boldsymbol{Q},t)$ and not a function of any of the momentum variables $\boldsymbol{p}$ and $\boldsymbol{P}$. The transformed Hamiltonian is then

$$\widetilde H(\boldsymbol{Q},\boldsymbol{P},t) = H(\boldsymbol{q},\boldsymbol{p},t) + \frac{\partial F(\boldsymbol{q},\boldsymbol{Q},t)}{\partial t} \quad . \tag{1.56}$$

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

$$F(\boldsymbol{q},\boldsymbol{Q},t) = \begin{cases} F_1(\boldsymbol{q},\boldsymbol{Q},t) & ; \quad p_\sigma = +\frac{\partial F_1}{\partial q_\sigma} \quad , \quad P_\sigma = -\frac{\partial F_1}{\partial Q_\sigma} \quad \text{(type I)} \\[2ex] F_2(\boldsymbol{q},\boldsymbol{P},t) - P_\sigma\,Q_\sigma & ; \quad p_\sigma = +\frac{\partial F_2}{\partial q_\sigma} \quad , \quad Q_\sigma = +\frac{\partial F_2}{\partial P_\sigma} \quad \text{(type II)} \\[2ex] F_3(\boldsymbol{p},\boldsymbol{Q},t) + p_\sigma\,q_\sigma & ; \quad q_\sigma = -\frac{\partial F_3}{\partial p_\sigma} \quad , \quad P_\sigma = -\frac{\partial F_3}{\partial Q_\sigma} \quad \text{(type III)} \\[2ex] F_4(\boldsymbol{p},\boldsymbol{P},t) + p_\sigma\,q_\sigma - P_\sigma\,Q_\sigma & ; \quad q_\sigma = -\frac{\partial F_4}{\partial p_\sigma} \quad , \quad Q_\sigma = +\frac{\partial F_4}{\partial P_\sigma} \quad \text{(type IV)} \end{cases} \tag{1.57}$$

[3]Solutions of eqn. 1.53 with $\lambda\neq 1$ are known as *extended* canonical transformations. We can always rescale coordinates and/or momenta to achieve $\lambda=1$.

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

$$\widetilde{H}(\boldsymbol{Q}, \boldsymbol{P}, t) = H(\boldsymbol{q}, \boldsymbol{p}, t) + \frac{\partial F_\gamma}{\partial t} \quad . \tag{1.58}$$

Let's work out some examples:

- Consider the type-II transformation generated by

$$F_2(\boldsymbol{q}, \boldsymbol{P}) = A_\sigma(\boldsymbol{q})\, P_\sigma \quad , \tag{1.59}$$

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

$$Q_\sigma = \frac{\partial F_2}{\partial P_\sigma} = A_\sigma(\boldsymbol{q}) \qquad , \qquad p_\sigma = \frac{\partial F_2}{\partial q_\sigma} = \frac{\partial A_\alpha}{\partial q_\sigma}\, P_\alpha \quad . \tag{1.60}$$

Thus,

$$Q_\sigma = A_\sigma(\boldsymbol{q}) \qquad , \qquad P_\sigma = \frac{\partial q_\alpha}{\partial Q_\sigma}\, p_\alpha \quad . \tag{1.61}$$

This is a general point transformation of the kind discussed in eqn. 1.31. 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.* $\boldsymbol{Q} = M\boldsymbol{q}$, $\boldsymbol{P} = \boldsymbol{p}\, M^{-1}$. 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

$$F_1(\boldsymbol{q}, \boldsymbol{Q}) = A_\sigma(\boldsymbol{q})\, Q_\sigma \quad . \tag{1.62}$$

We then have

$$\begin{aligned} p_\sigma &= \frac{\partial F_1}{\partial q_\sigma} = \frac{\partial A_\alpha}{\partial q_\sigma}\, Q_\alpha \\ P_\sigma &= -\frac{\partial F_1}{\partial Q_\sigma} = -A_\sigma(\boldsymbol{q}) \quad . \end{aligned} \tag{1.63}$$

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

$$\begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} \longrightarrow \begin{pmatrix} -\boldsymbol{P} \\ +\boldsymbol{Q} \end{pmatrix} \quad . \tag{1.64}$$

- A mixed transformation is also permitted. For example,

$$F(\boldsymbol{q}, \boldsymbol{Q}) = q_1\, Q_1 + (q_3 - Q_2)\, P_2 + (q_2 - Q_3)\, P_3 \tag{1.65}$$

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

$$Q_1 = p_1 \quad , \quad Q_2 = q_3 \quad , \quad Q_3 = q_2 \quad , \quad P_1 = -q_1 \quad , \quad P_2 = p_3 \quad , \quad P_3 = p_2 \quad . \tag{1.66}$$

- Consider the $n = 1$ harmonic oscillator,

$$H(q,p) = \frac{p^2}{2m} + \tfrac{1}{2}kq^2 \quad . \tag{1.67}$$

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

$$p = \sqrt{2mf(P)}\,\cos Q \qquad , \qquad q = \sqrt{\frac{2\,f(P)}{k}}\,\sin Q \quad , \tag{1.68}$$

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

$$p = \sqrt{mk}\,q\,\operatorname{ctn} Q \quad , \tag{1.69}$$

which suggests the type-I transformation

$$F_1(q,Q) = \tfrac{1}{2}\sqrt{mk}\,q^2\operatorname{ctn} Q \quad . \tag{1.70}$$

This leads to

$$p = \frac{\partial F_1}{\partial q} = \sqrt{mk}\,q\,\operatorname{ctn} Q \quad , \quad P = -\frac{\partial F_1}{\partial Q} = \frac{\sqrt{mk}\,q^2}{2\,\sin^2\!Q} \quad . \tag{1.71}$$

Thus,

$$q = \frac{\sqrt{2P}}{\sqrt[4]{mk}}\,\sin Q \quad \Longrightarrow \quad f(P) = \sqrt{\frac{k}{m}}\,P = \omega P \quad , \tag{1.72}$$

where $\omega = \sqrt{k/m}$ is the oscillation frequency. We therefore have that $\widetilde{H}(Q,P) = \omega P$, whence $P = E/\omega$. The equations of motion are

$$\dot{P} = -\frac{\partial \widetilde{H}}{\partial Q} = 0 \qquad , \qquad \dot{Q} = \frac{\partial \widetilde{H}}{\partial P} = \omega \quad , \tag{1.73}$$

which yields

$$Q(t) = \omega t + \varphi_0 \qquad , \qquad q(t) = \sqrt{\frac{2E}{m\omega^2}}\,\sin\big(\omega t + \varphi_0\big) \quad . \tag{1.74}$$

1.4 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 $\widetilde{H}(\boldsymbol{Q},\boldsymbol{P},t)$ and the corresponding equations of motion as simple as possible.

Recall the general form of the canonical transformation:

$$\widetilde{H}(\boldsymbol{Q},\boldsymbol{P},t) = H(\boldsymbol{q},\boldsymbol{p},t) + \frac{\partial F(\boldsymbol{q},\boldsymbol{Q},t)}{\partial t} \quad , \tag{1.75}$$

with

$$\frac{\partial F}{\partial q_\sigma} = p_\sigma \quad , \quad \frac{\partial F}{\partial p_\sigma} = 0 \quad , \quad \frac{\partial F}{\partial Q_\sigma} = -P_\sigma \quad , \quad \frac{\partial F}{\partial P_\sigma} = 0 \quad . \tag{1.76}$$

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

$$\frac{\partial F}{\partial t} = -H \quad , \quad \frac{\partial F}{\partial q_\sigma} = p_\sigma \quad . \tag{1.77}$$

The remaining functional dependence may be taken to be either on $\boldsymbol{Q}$ (type I) or on $\boldsymbol{P}$ (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.

1.4.1 The action as a function of coordinates and time

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

$$S(\boldsymbol{q}, t) = \int_{t_a}^{t} d\tau \, L\big(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}, \tau\big) \quad , \tag{1.78}$$

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

$$\frac{\partial L}{\partial \eta_\sigma} - \frac{d}{d\tau}\bigg(\frac{\partial L}{\partial \dot{\eta}_\sigma}\bigg) = 0 \tag{1.79}$$

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

$$\begin{aligned}
dS = S\big[\tilde{\boldsymbol{\eta}}(\tau)\big] - S\big[\boldsymbol{\eta}(\tau)\big] &= \int_{t_a}^{t+dt} d\tau \, L(\tilde{\boldsymbol{\eta}}, \dot{\tilde{\boldsymbol{\eta}}}, \tau) - \int_{t_a}^{t} d\tau \, L\big(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}, \tau\big) \\
&= \int_{t_a}^{t} d\tau \left\{ \frac{\partial L}{\partial \eta_\sigma} \Big[\tilde{\eta}_\sigma(\tau) - \eta_\sigma(\tau)\Big] + \frac{\partial L}{\partial \dot{\eta}_\sigma} \Big[\dot{\tilde{\eta}}_\sigma(\tau) - \dot{\eta}_\sigma(\tau)\Big] \right\} + L\big(\tilde{\boldsymbol{\eta}}(t), \dot{\tilde{\boldsymbol{\eta}}}(t), t\big) \, dt \\
&= \int_{t_a}^{t} d\tau \left\{ \frac{\partial L}{\partial \eta_\sigma} - \frac{d}{d\tau}\bigg(\frac{\partial L}{\partial \dot{\eta}_\sigma}\bigg) \right\} \Big[\tilde{\eta}_\sigma(\tau) - \eta_\sigma(\tau)\Big] + \frac{\partial L}{\partial \dot{\eta}_\sigma}\bigg|_t \Big[\tilde{\eta}_\sigma(t) - \eta_\sigma(t)\Big] + L\big(\tilde{\boldsymbol{\eta}}(t), \dot{\tilde{\boldsymbol{\eta}}}(t), t\big) \, dt \\
&= 0 + \pi_\sigma(t) \, \delta\eta_\sigma(t) + L\big(\boldsymbol{\eta}(t), \dot{\boldsymbol{\eta}}(t), t\big) \, dt + \mathcal{O}(\delta\boldsymbol{q} \, dt) \quad ,
\end{aligned} \tag{1.80}$$

where we have defined $\pi_\sigma = \partial L / \partial \dot{\eta}_\sigma$, and $\delta\eta_\sigma(\tau) \equiv \tilde{\eta}_\sigma(\tau) - \eta_\sigma(\tau)$.

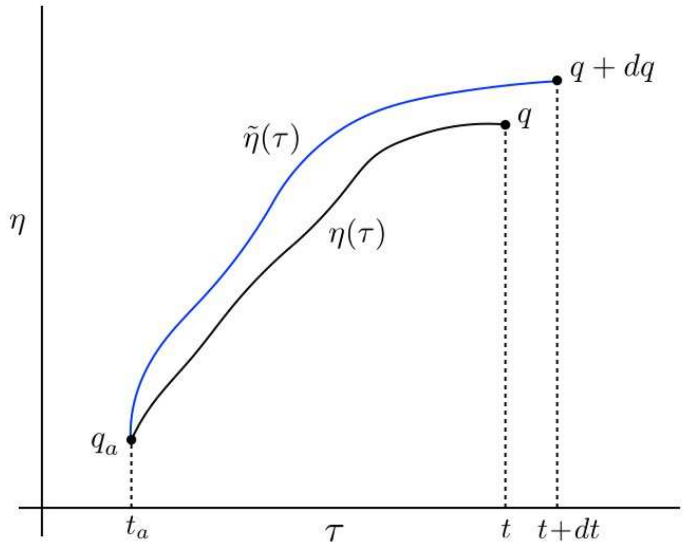

Figure 1.1: The paths $\boldsymbol{\eta}(\tau)$ and $\tilde{\boldsymbol{\eta}}(\tau)$.

Note that the differential dq_σ is given by

$$\begin{aligned} dq_\sigma &= \tilde{\eta}_\sigma(t+dt) - \eta_\sigma(t) \\ &= \tilde{\eta}_\sigma(t+dt) - \tilde{\eta}_\sigma(t) + \tilde{\eta}_\sigma(t) - \eta_\sigma(t) \\ &= \dot{\tilde{\eta}}_\sigma(t)\, dt + \delta\eta_\sigma(t) = \dot{q}_\sigma(t)\, dt + \delta\eta_\sigma(t) + \mathcal{O}(\delta\boldsymbol{q}\, dt) \quad . \end{aligned} \tag{1.81}$$

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

$$\begin{aligned} dS &= p_\sigma\, dq_\sigma + \big(L - p_\sigma\, \dot{q}_\sigma\big)\, dt \\ &= p_\sigma\, dq_\sigma - H\, dt \quad . \end{aligned} \tag{1.82}$$

We therefore obtain

$$\frac{\partial S}{\partial q_\sigma} = p_\sigma \quad , \quad \frac{\partial S}{\partial t} = -H \quad , \quad \frac{dS}{dt} = L \quad . \tag{1.83}$$

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

$$S = S(q_1, \ldots, q_n; \Lambda_1, \ldots, \Lambda_n, t) + \Lambda_{n+1} \quad , \tag{1.84}$$

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

$$\Gamma_\sigma = \frac{\partial S}{\partial \Lambda_\sigma} = \begin{cases} +Q_\sigma & \text{if } \Lambda_\sigma = P_\sigma \\ -P_\sigma & \text{if } \Lambda_\sigma = Q_\sigma \end{cases} \tag{1.85}$$

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

1.4.2 The Hamilton-Jacobi equation

Since the action $S(\boldsymbol{q}, \boldsymbol{\Lambda}, t)$ generates a canonical transformation for which $\tilde{H}(\boldsymbol{Q}, \boldsymbol{P}) = 0$, this requirement may be written as

$$H\Big(q_1, \ldots, q_n, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}, t\Big) + \frac{\partial S}{\partial t} = 0 \quad . \tag{1.86}$$

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 $\widetilde{H}(\boldsymbol{Q}, \boldsymbol{P}, t) = 0$, the equations of motion are trivial, and

$$Q_\sigma(t) = \text{const.} \qquad , \qquad P_\sigma(t) = \text{const.} \tag{1.87}$$

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

$$\det\left(\frac{\partial^2 S}{\partial q_\alpha \, \partial \Lambda_\beta}\right) \neq 0 \quad , \tag{1.88}$$

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 in one dimension, with $H(q,p) = p^2/2m$. The HJE is

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{\partial S}{\partial t} = 0 \quad . \tag{1.89}$$

One solution of the HJE is

$$S(q, \Lambda, t) = \frac{m\,(q-\Lambda)^2}{2t} \quad . \tag{1.90}$$

For this we find

$$\Gamma = \frac{\partial S}{\partial \Lambda} = -\frac{m}{t}\,(q-\Lambda) \quad \Rightarrow \quad q(t) = \Lambda - \frac{\Gamma}{m}\, t \quad . \tag{1.91}$$

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

$$S(q, \Lambda, t) = q\,\sqrt{2m\Lambda} - \Lambda\, t \quad . \tag{1.92}$$

This yields

$$\Gamma = \frac{\partial S}{\partial \Lambda} = q\,\sqrt{\frac{m}{2\Lambda}} - t \quad \Rightarrow \quad q(t) = \sqrt{\frac{2\Lambda}{m}}\,(t+\Gamma) \quad . \tag{1.93}$$

For this solution, $\Lambda = \frac{1}{2}mv^2$ is the energy and Γ may be related to the initial value $q(0) = \Gamma\sqrt{2\Lambda/m}$.

1.4.3 Time-independent Hamiltonians

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

$$S(\boldsymbol{q},\boldsymbol{\Lambda},t) = W(\boldsymbol{q},\boldsymbol{\Lambda}) + T(\boldsymbol{\Lambda},t) \quad . \tag{1.94}$$

The HJE becomes

$$H\bigg(\boldsymbol{q}\,,\frac{\partial W}{\partial \boldsymbol{q}}\bigg) = -\frac{\partial T}{\partial t} \quad . \tag{1.95}$$

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

$$S(\boldsymbol{q},\boldsymbol{\Lambda},t) = W(\boldsymbol{q},\boldsymbol{\Lambda}) - \Lambda_1 t \quad . \tag{1.96}$$

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

$$H\bigg(q_1,\ldots,q_n\,,\frac{\partial W}{\partial q_1}\,,\ldots\,,\frac{\partial W}{\partial q_n}\bigg) = \Lambda_1 \quad . \tag{1.97}$$

Note that adding an arbitrary constant C to S generates the same equation, and simply shifts the last constant $\Lambda_{n+1} \to \Lambda_{n+1} + C$. According to eqn. 1.96, 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.

1.4.4 Example: one-dimensional motion

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

$$H(q,p) = \frac{p^2}{2m} + U(q) \quad . \tag{1.98}$$

The HJE is

$$\frac{1}{2m}\bigg(\frac{\partial W}{\partial q}\bigg)^2 + U(q) = \Lambda \quad . \tag{1.99}$$

Clearly $\Lambda = E$ is the total energy. The HJE may be recast as

$$\frac{\partial W}{\partial q} = \pm\sqrt{2m\big[\Lambda - U(q)\big]} \quad , \tag{1.100}$$

with solution

$$W(q,\Lambda) = \pm\sqrt{2m}\int^q \! dq'\,\sqrt{\Lambda - U(q')} \quad , \tag{1.101}$$

with $S(q,\Lambda,t) = W(q,\Lambda) - \Lambda t$. We now have

$$p = \frac{\partial W}{\partial q} = \pm\sqrt{2m\big[\Lambda - U(q)\big]} \quad , \tag{1.102}$$

as well as

$$\Gamma = \frac{\partial S}{\partial \Lambda} = \frac{\partial W}{\partial \Lambda} - t = \pm\sqrt{\frac{m}{2}} \int^{q(t)} \frac{dq'}{\sqrt{\Lambda - U(q')}} - t \quad . \tag{1.103}$$

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

$$\Gamma + t = \pm\sqrt{\frac{m}{2}} \int^{q(t)} \frac{dq'}{\sqrt{\Lambda - U(q')}} \quad , \tag{1.104}$$

where Λ and Γ are constants. The lower limit on the integral is arbitrary and merely shifts t by another constant. The characteristic function $W(q, \Lambda)$ is actually double-valued in q, corresponding to right-moving and left-moving parts of the motion.

1.4.5 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:

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2\!\theta}\right) + \overbrace{A(r) + \frac{B(\theta)}{r^2} + \frac{C(\phi)}{r^2 \sin^2\!\theta}}^{\text{potential } U(r,\theta,\phi)} \quad . \tag{1.105}$$

We seek a characteristic function of the form $W(r, \theta, \phi) = W_r(r) + W_\theta(\theta) + W_\phi(\phi)$. The HJE is then

$$\begin{aligned} \frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{1}{2mr^2}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{1}{2mr^2 \sin^2\!\theta}\left(\frac{\partial W_\phi}{\partial \phi}\right)^2 \\ + A(r) + \frac{B(\theta)}{r^2} + \frac{C(\phi)}{r^2 \sin^2\!\theta} = \Lambda_1 = E \quad . \end{aligned} \tag{1.106}$$

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

$$\begin{aligned} \frac{1}{2m}\left(\frac{\partial W_\phi}{\partial \phi}\right)^2 + C(\phi) = -\sin^2\!\theta \left\{\frac{1}{2m}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + B(\theta)\right\} \\ - r^2 \sin^2\!\theta \left\{\frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2 + A(r) - \Lambda_1\right\} \quad . \end{aligned} \tag{1.107}$$

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

$$\frac{1}{2m}\left(\frac{\partial W_\phi}{\partial \phi}\right)^2 + C(\phi) = \Lambda_2 \quad . \tag{1.108}$$

Proceeding, we replace the LHS in eqn. 1.107 with Λ_2, arriving at

$$\frac{1}{2m}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + B(\theta) + \frac{\Lambda_2}{\sin^2\!\theta} = -r^2 \left\{\frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2 + A(r) - \Lambda_1\right\} \quad . \tag{1.109}$$

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

$$\frac{1}{2m}\left(\frac{\partial W_\theta}{\partial\theta}\right)^2 + B(\theta) + \frac{\mathit{\Lambda}_2}{\sin^2\theta} = \mathit{\Lambda}_3 \quad . \tag{1.110}$$

We're left with

$$\frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2 + A(r) + \frac{\mathit{\Lambda}_3}{r^2} = \mathit{\Lambda}_1 \quad . \tag{1.111}$$

The full solution is therefore

$$\begin{aligned} S(\boldsymbol{q},\boldsymbol{\mathit{\Lambda}},t) = \sqrt{2m}\int^r dr'\,\sqrt{\mathit{\Lambda}_1 - A(r') - \frac{\mathit{\Lambda}_3}{r'^2}} + \sqrt{2m}\int^\theta d\theta'\,\sqrt{\mathit{\Lambda}_3 - B(\theta') - \frac{\mathit{\Lambda}_2}{\sin^2\theta'}} \\ + \sqrt{2m}\int^\phi d\phi'\,\sqrt{\mathit{\Lambda}_2 - C(\phi')} - \mathit{\Lambda}_1 t \quad . \end{aligned} \tag{1.112}$$

We then have

$$\begin{aligned} \mathit{\Gamma}_1 &= \frac{\partial S}{\partial \mathit{\Lambda}_1} = \sqrt{\tfrac{m}{2}}\int^{r(t)} \frac{dr'}{\sqrt{\mathit{\Lambda}_1 - A(r') - \mathit{\Lambda}_3\, r'^{-2}}} - t \\ \mathit{\Gamma}_2 &= \frac{\partial S}{\partial \mathit{\Lambda}_2} = -\sqrt{\tfrac{m}{2}}\int^{\theta(t)} \frac{d\theta'}{\sin^2\theta'\,\sqrt{\mathit{\Lambda}_3 - B(\theta') - \mathit{\Lambda}_2\csc^2\theta'}} + \sqrt{\tfrac{m}{2}}\int^{\phi(t)} \frac{d\phi'}{\sqrt{\mathit{\Lambda}_2 - C(\phi')}} \\ \mathit{\Gamma}_3 &= \frac{\partial S}{\partial \mathit{\Lambda}_3} = -\sqrt{\tfrac{m}{2}}\int^{r(t)} \frac{dr'}{r'^2\,\sqrt{\mathit{\Lambda}_1 - A(r') - \mathit{\Lambda}_3\, r'^{-2}}} + \sqrt{\tfrac{m}{2}}\int^{\theta(t)} \frac{d\theta'}{\sqrt{\mathit{\Lambda}_3 - B(\theta') - \mathit{\Lambda}_2\csc^2\theta'}} \quad . \end{aligned} \tag{1.113}$$

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 $(\mathit{\Lambda}_1, \mathit{\Lambda}_2, \mathit{\Lambda}_3, \mathit{\Gamma}_1, \mathit{\Gamma}_2, \mathit{\Gamma}_3)$. 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[4]

$$W(\boldsymbol{q},\boldsymbol{\mathit{\Lambda}}) = \sum_{\sigma=1}^{n} W_\sigma(q_\sigma\,,\boldsymbol{\mathit{\Lambda}}) \quad . \tag{1.114}$$

Each $W_\sigma(q_\sigma,\boldsymbol{\mathit{\Lambda}})$ may be regarded as a function of the single variable q_σ, and is obtained by satisfying an ODE of the form[5]

$$H_\sigma\bigg(q_\sigma, \frac{dW_\sigma}{dq_\sigma}\bigg) = \mathit{\Lambda}_\sigma \quad . \tag{1.115}$$

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

[5] Note that $H_\sigma(q_\sigma, p_\sigma)$ may itself depend on several of the constants $\mathit{\Lambda}_\alpha$. For example, eqn. 1.111 is of the form $H_r\big(r, \partial_r W_r, \mathit{\Lambda}_3\big) = \mathit{\Lambda}_1$.

We then have

$$p_\sigma = \frac{\partial W_\sigma}{\partial q_\sigma} \quad , \quad \Gamma_\sigma = \frac{\partial W_\sigma}{\partial \Lambda_\sigma} + \delta_{\sigma,1}\, t \quad . \tag{1.116}$$

Note that while each W_σ depends on only a single q_σ, it may depend on several of the Λ_σ.

1.5 Action-Angle Variables

1.5.1 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:

$$p_\sigma = \frac{\partial W_\sigma}{\partial q_\sigma} = p_\sigma(q_\sigma, \boldsymbol{\Lambda}) \quad . \tag{1.117}$$

The motion satisfies $H_\sigma(q_\sigma, p_\sigma) = \Lambda_\sigma$. The level sets of H_σ are curves $\mathcal{C}_\sigma$. In general, these curves each depend on all of the constants $\boldsymbol{\Lambda}$, so we write $\mathcal{C}_\sigma = \mathcal{C}_\sigma(\boldsymbol{\Lambda})$. 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

$$H(\phi, p_\phi) = \frac{p_\phi^2}{2I} + \tfrac{1}{2} I\omega^2 \left(1 - \cos\phi\right) \quad . \tag{1.118}$$

- 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[6]. 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, $\mathbb{S}^1$. For n freedoms, the motion is therefore confined to an n-torus, $\mathbb{T}^n$:

$$\mathbb{T}^n = \overbrace{\mathbb{S}^1 \times \mathbb{S}^1 \times \cdots \times \mathbb{S}^1}^{n \text{ times}} \quad . \tag{1.119}$$

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.

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

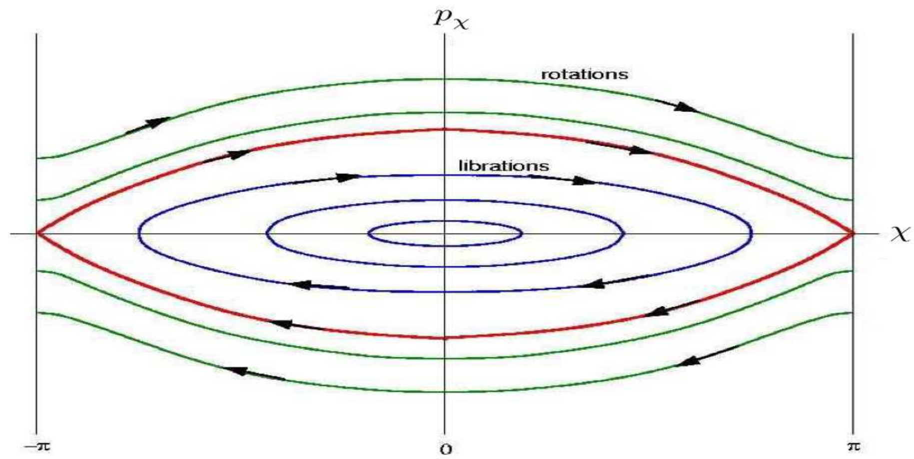

Figure 1.2: Phase curves for the simple pendulum, showing librations (in blue), rotations (in green), and the separatrix (in red). This phase flow is most correctly viewed as taking place on a cylinder, obtained from the above sketch by identifying the lines $\phi = \pi$ and $\phi = -\pi$.

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.

1.5.2 Action-Angle Variables

For a completely integrable system, one can transform canonically from $(\boldsymbol{q}, \boldsymbol{p})$ to new coordinates $(\boldsymbol{\phi}, \boldsymbol{J})$ which specify a particular n-torus $\mathbb{T}^n$ 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

$$\dot{J}_\sigma = -\frac{\partial H}{\partial \phi_\sigma} = 0 \quad \Longrightarrow \quad H = H(\boldsymbol{J}) \quad . \tag{1.120}$$

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

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

$$\oint\limits_{\mathcal{C}_\sigma} d\phi_\sigma = 2\pi \qquad (\text{once around } \mathcal{C}_\sigma) \quad . \tag{1.121}$$

The dynamics of the angle variables are given by

$$\dot{\phi}_\sigma = \frac{\partial H}{\partial J_\sigma} \equiv \nu_\sigma(\boldsymbol{J}) \quad . \tag{1.122}$$

Thus, the motion is given by

$$\phi_\sigma(t) = \phi_\sigma(0) + \nu_\sigma(\boldsymbol{J})\, t \quad . \tag{1.123}$$

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

1.5.3 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 $F_2(\boldsymbol{q}, \boldsymbol{J})$:

$$p_\sigma = \frac{\partial F_2}{\partial q_\sigma} \quad , \quad \phi_\sigma = \frac{\partial F_2}{\partial J_\sigma} \quad . \tag{1.124}$$

Note that[7]

$$2\pi = \oint\limits_{\mathcal{C}_\sigma} d\phi_\sigma = \oint\limits_{\mathcal{C}_\sigma} d\left(\frac{\partial F_2}{\partial J_\sigma}\right) = \oint\limits_{\mathcal{C}_\sigma} \frac{\partial^2 F_2}{\partial J_\sigma\, \partial q_\sigma}\, dq_\sigma = \frac{\partial}{\partial J_\sigma} \oint\limits_{\mathcal{C}_\sigma} p_\sigma\, dq_\sigma \quad , \tag{1.125}$$

which suggests the definition

$$J_\sigma = \frac{1}{2\pi} \oint\limits_{\mathcal{C}_\sigma} p_\sigma\, dq_\sigma \quad . \tag{1.126}$$

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

If, separating variables,

$$W(\boldsymbol{q}, \boldsymbol{\Lambda}) = \sum_\sigma W_\sigma(q_\sigma, \boldsymbol{\Lambda}) \tag{1.127}$$

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

$$J_\sigma = \frac{1}{2\pi} \oint\limits_{\mathcal{C}_\sigma} \frac{\partial W_\sigma}{\partial q_\sigma}\, dq_\sigma = J_\sigma(\boldsymbol{\Lambda}) \tag{1.128}$$

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

$$F_2(\boldsymbol{q}, \boldsymbol{J}) = W\big(\boldsymbol{q}, \boldsymbol{\Lambda}(\boldsymbol{J})\big) = \sum_\sigma W_\sigma\big(q_\sigma, \boldsymbol{\Lambda}(\boldsymbol{J})\big) \quad . \tag{1.129}$$

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

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

(2) Find the orbits $\mathcal{C}_\sigma(\boldsymbol{\Lambda})$, *i.e.* the level sets satisfying $H_\sigma(q_\sigma, p_\sigma) = \Lambda_\sigma$.

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

[7] 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. 1.125 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.

(4) $F_2(\boldsymbol{q},\boldsymbol{J}) = \sum_\sigma W_\sigma\big(q_\sigma, \boldsymbol{\Lambda}(\boldsymbol{J})\big)$ is the desired type-II generator[8].

1.5.4 Example : Harmonic Oscillator

The Hamiltonian is

$$H = \frac{p^2}{2m} + \tfrac{1}{2} m\omega_0^2\, q^2 \quad , \tag{1.130}$$

hence the Hamilton-Jacobi equation is

$$\bigg(\frac{dW}{dq}\bigg)^2 + m^2\omega_0^2\, q^2 = 2m\Lambda \quad . \tag{1.131}$$

Thus,

$$p = \frac{dW}{dq} = \pm\sqrt{2m\Lambda - m^2\omega_0^2 q^2} \quad . \tag{1.132}$$

We now define

$$q \equiv \sqrt{\frac{2\Lambda}{m\omega_0^2}}\,\sin\theta \qquad \Rightarrow \qquad p = \sqrt{2m\Lambda}\,\cos\theta \quad , \tag{1.133}$$

in which case

$$J = \frac{1}{2\pi}\oint p\,dq = \frac{1}{2\pi}\cdot\frac{2\Lambda}{\omega_0}\cdot\int\limits_0^{2\pi}\!d\theta\,\cos^2\!\theta = \frac{\Lambda}{\omega_0} \quad . \tag{1.134}$$

Solving the HJE, we write

$$\frac{dW}{d\theta} = \frac{\partial q}{\partial\theta}\cdot\frac{dW}{dq} = 2J\cos^2\!\theta \quad . \tag{1.135}$$

Integrating, we obtain

$$W = J\theta + \tfrac{1}{2} J\sin 2\theta \quad , \tag{1.136}$$

up to an irrelevant constant. We then have

$$\phi = \frac{\partial W}{\partial J}\bigg|_q = \theta + \tfrac{1}{2}\sin 2\theta + J\big(1+\cos 2\theta\big)\,\frac{\partial\theta}{\partial J}\bigg|_q \quad . \tag{1.137}$$

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

$$dq = \frac{\sin\theta}{\sqrt{2m\omega_0 J}}\,dJ + \sqrt{\frac{2J}{m\omega_0}}\cos\theta\,d\theta \qquad \Rightarrow \qquad \frac{\partial\theta}{\partial J}\bigg|_q = -\frac{1}{2J}\,\tan\theta \quad . \tag{1.138}$$

Plugging this result into eqn. 1.137, we obtain $\phi = \theta$. Thus, the full transformation is

$$q = \sqrt{\frac{2J}{m\omega_0}}\sin\phi \quad , \quad p = \sqrt{2m\omega_0 J}\,\cos\phi \quad . \tag{1.139}$$

The Hamiltonian is

$$H = \omega_0\, J \quad , \tag{1.140}$$

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 $J(t) = J(0)$.

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

1.5.5 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

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} + mgz \quad . \tag{1.141}$$

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

$$\frac{1}{2m}\left(\frac{\partial W_x}{\partial x}\right)^2 + \frac{1}{2m}\left(\frac{\partial W_y}{\partial y}\right)^2 + \frac{1}{2m}\left(\frac{\partial W_z}{\partial z}\right)^2 + mgz = E \equiv \Lambda_z \quad . \tag{1.142}$$

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

$$W_x(x) = \sqrt{2m\Lambda_x}\, x \quad , \quad W_y(y) = \sqrt{2m\Lambda_y}\, y \quad . \tag{1.143}$$

We then have[9]

$$\begin{aligned} W_z'(z) &= -\sqrt{2m\big(\Lambda_z - \Lambda_x - \Lambda_y - mgz\big)} \\ W_z(z) &= \frac{2\sqrt{2}}{3\sqrt{m}\, g}\big(\Lambda_z - \Lambda_x - \Lambda_y - mgz\big)^{3/2} \quad . \end{aligned} \tag{1.144}$$

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

$$p_z(z) = W_z'(z) = \sqrt{2m\big(\Lambda_z - \Lambda_x - \Lambda_y - mgz\big)} \quad , \tag{1.145}$$

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

$$\begin{aligned} J_x &= \frac{1}{2\pi}\oint_{\mathcal{C}_x} p_x\, dx = \frac{1}{\pi}\int_0^{L_x} dx\sqrt{2m\Lambda_x} = \frac{L_x}{\pi}\sqrt{2m\Lambda_x} \\ J_y &= \frac{1}{2\pi}\oint_{\mathcal{C}_y} p_y\, dy = \frac{1}{\pi}\int_0^{L_y} dy\sqrt{2m\Lambda_y} = \frac{L_y}{\pi}\sqrt{2m\Lambda_y} \end{aligned} \tag{1.146}$$

and

$$\begin{aligned} J_z &= \frac{1}{2\pi}\oint_{\mathcal{C}_z} p_z\, dz = \frac{1}{\pi}\int_0^{z_{\rm max}} dz\, \sqrt{2m\big(\Lambda_z - \Lambda_x - \Lambda_y - mgz\big)} \\ &= \frac{2\sqrt{2}}{3\pi\sqrt{m}\, g}\big(\Lambda_z - \Lambda_x - \Lambda_y\big)^{3/2} \quad . \end{aligned} \tag{1.147}$$

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

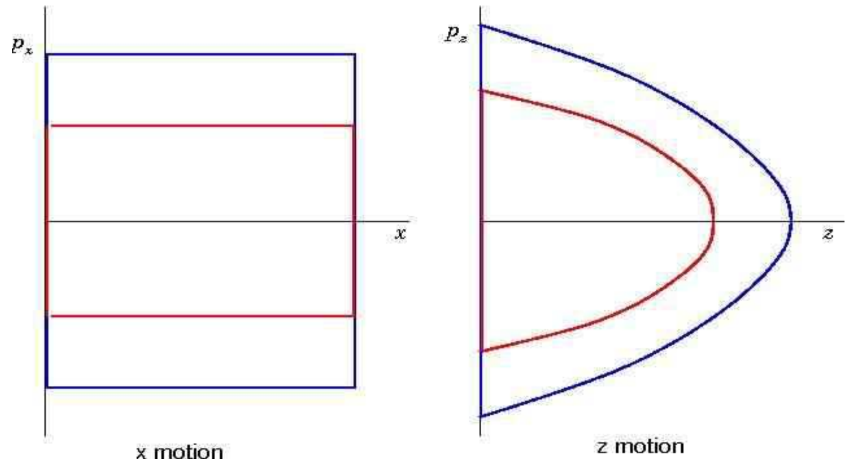

Figure 1.3: The librations $\mathcal{C}_z$ and $\mathcal{C}_x$. Not shown is $\mathcal{C}_y$, which is of the same shape as $\mathcal{C}_x$.

We now invert to obtain

$$\begin{aligned} \Lambda_x &= \frac{\pi^2}{2mL_x^2}\,J_x^2 \quad , \quad \Lambda_y = \frac{\pi^2}{2mL_y^2}\,J_y^2 \\ \Lambda_z &= \bigg(\frac{3\pi\sqrt{m}\,g}{2\sqrt{2}}\bigg)^{2/3} J_z^{2/3} + \frac{\pi^2}{2mL_x^2}\,J_x^2 + \frac{\pi^2}{2mL_y^2}\,J_y^2 \quad . \end{aligned} \tag{1.148}$$

$$F_2\big(x,y,z,J_x,J_y,J_z\big) = \frac{\pi x}{L_x}\,J_x + \frac{\pi y}{L_y}\,J_y + \pi\bigg(J_z^{2/3} - \frac{2m^{2/3}g^{1/3}z}{(3\pi)^{2/3}}\bigg)^{3/2} \quad . \tag{1.149}$$

We now find

$$\phi_x = \frac{\partial F_2}{\partial J_x} = \frac{\pi x}{L_x} \quad , \quad \phi_y = \frac{\partial F_2}{\partial J_y} = \frac{\pi y}{L_y} \tag{1.150}$$

and

$$\phi_z = \frac{\partial F_2}{\partial J_z} = \pi\sqrt{1 - \frac{2m^{2/3}g^{1/3}z}{(3\pi J_z)^{2/3}}} = \pi\sqrt{1-\frac{z}{z_{\rm max}}} \quad , \tag{1.151}$$

where $z_{\rm max}(J_z) = (3\pi J_z/m)^{2/3}/2g^{1/3}$. The momenta are

$$p_x = \frac{\partial F_2}{\partial x} = \frac{\pi J_x}{L_x} \quad , \quad p_y = \frac{\partial F_2}{\partial y} = \frac{\pi J_y}{L_y} \tag{1.152}$$

and

$$p_z = \frac{\partial F_2}{\partial z} = -\sqrt{2m}\,\Bigg(\bigg(\frac{3\pi\sqrt{m}\,g}{2\sqrt{2}}\bigg)^{2/3} J_z^{2/3} - mgz\Bigg)^{1/2} \quad . \tag{1.153}$$

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. 1.125. 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

$$W_x(x) = \begin{cases} \sqrt{2m\Lambda_x}\, x & \text{if } p_x > 0 \\ \sqrt{2m\Lambda_x}\,(2L_x - x) & \text{if } p_x < 0 \end{cases} \quad . \tag{1.154}$$

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

$$W_x(x) = \begin{cases} (\pi x/L_x)\, J_x & \text{if } p_x > 0 \\ 2\pi J_x - (\pi x/L_x)\, J_x & \text{if } p_x < 0 \end{cases} \quad . \tag{1.155}$$

and

$$\phi_x = \begin{cases} \pi x/L_x & \text{if } p_x > 0 \\ \pi(2L_x - x)/L_x & \text{if } p_x < 0 \end{cases} \quad . \tag{1.156}$$

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

1.6 Integrability and Motion on Invariant Tori

1.6.1 Librations and rotations

As discussed above, a completely integrable Hamiltonian system is solvable by separation of variables. The angle variables evolve as

$$\phi_\sigma(t) = \nu_\sigma(\boldsymbol{J})\, t + \phi_\sigma(0) \quad . \tag{1.157}$$

Thus, they wind around the invariant torus, specified by $\{J_\sigma\}$ at constant rates. In general, while each ϕ_σ executes periodic motion around a circle, the motion of the system as a whole is not periodic, since the frequencies $\nu_\sigma(\boldsymbol{J})$ are not, in general, commensurate. Periodic motion requires that there exists a time T such that $\nu_\sigma(\boldsymbol{J})\, T = 2\pi k_\sigma$ with $k_\sigma \in \mathbb{Z}$ for each $\sigma \in \{1, \ldots, n\}$ where each . This means the ratio of any two frequencies $\nu_\sigma/\nu_{\sigma'} = k_\sigma/k_{\sigma'} \in \mathbb{Q}$ must be a rational number. T is the smallest possible such period provided the set $\{k_1, \ldots, k_n\}$ has no common factors. 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 each such invariant torus, on a given torus either all orbits are periodic, or none are.

In terms of the original coordinates $\boldsymbol{q}$, there are two possibilities:

$$q_\sigma(t) = \sum_{\ell_1=-\infty}^{\infty} \cdots \sum_{\ell_n=-\infty}^{\infty} A^{(\sigma)}_{\ell_1 \ell_2 \cdots \ell_n}\, e^{i\ell_1\phi_1(t)} \cdots e^{i\ell_n\phi_n(t)} \equiv \sum_{\boldsymbol{\ell}} A^\sigma_{\boldsymbol{\ell}}\, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\phi}(t)} \qquad \text{(libration)} \tag{1.158}$$

or

$$q_\sigma(t) = \frac{q^\circ_\sigma\, \phi_\sigma(t)}{2\pi} + \sum_{\boldsymbol{\ell}} B^\sigma_{\boldsymbol{\ell}}\, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\phi}(t)} \qquad \text{(rotation)} \quad . \tag{1.159}$$

For rotations, the variable $q_\sigma(t)$ increased by $\Delta q_\sigma = q_\sigma^\circ$.

I want to distinguish two important concepts. *Complete periodicity*, as we have defined, requires that there exists a time $T(\boldsymbol{J})$ such that $\nu_\sigma(\boldsymbol{J})\,T(\boldsymbol{J}) = 2\pi k_\sigma$, with $k_\sigma \in \mathbb{Z}$ for all $\sigma \in \{1,\ldots,n\}$. The period is then defined to be the smallest nonzero such value of $T(\boldsymbol{J})$. The second condition, *resonance* is weaker and only requires that there exists some $\boldsymbol{\ell} \in \mathbb{Z}^n$ such that $\boldsymbol{\ell}\cdot\boldsymbol{\nu}(\boldsymbol{J}) = 0$[10]. *Resonance is thus equivalent to periodicity on a lower-dimensional sub-torus* $\mathbb{T}^k$ *with* $k < n$. In other words, if the *projected dynamics* $\boldsymbol{\phi}(t)$ onto *any* 2-torus $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ spanned by coordinates $(\phi_\sigma, \phi_{\bar\sigma})$ is periodic, where $\sigma, \bar\sigma \in \{1,\ldots,n\}$, then the original n-torus is said to be resonant. Complete periodicity is thus a maximal state of resonance, where the motion projected onto *any* subtorus $\mathbb{T}^2$ is periodic.

1.6.2 Liouville-Arnol'd theorem

Another statement of complete integrability is the content of the *Liouville-Arnol'd theorem*, which says the following. Suppose that a time-independent Hamiltonian $H(\boldsymbol{q},\boldsymbol{p})$ has n *first integrals* $I_k(\boldsymbol{q},\boldsymbol{p})$ with $k \in \{1,\ldots,n\}$. This means that (see eqn. 1.24)

$$0 = \frac{d}{dt} I_k(\boldsymbol{q},\boldsymbol{p}) = \sum_{\sigma=1}^n \left(\frac{\partial I_k}{\partial q_\sigma}\,\dot q_\sigma + \frac{\partial I_k}{\partial p_\sigma}\,\dot p_\sigma \right) = \big\{ I_k \,,\, H \big\} \quad . \tag{1.160}$$

If the $\{I_k\}$ are *independent functions*, meaning that the phase space gradients $\{\boldsymbol{\nabla} I_k\}$ constitute a set of n linearly independent vectors at almost every point $(\boldsymbol{q},\boldsymbol{p}) \in \mathcal{M}$ in phase space, and the different first integrals *commute* with respect to the Poisson bracket, *i.e.* $\{I_k, I_l\} = 0$, then the set of Hamilton's equations of motion is completely solvable[11]. The theorem establishes that[12]

(i) The space $\mathcal{M}_{\boldsymbol{I}} = \big\{ (\boldsymbol{q},\boldsymbol{p}) \in \mathcal{M} \,\big|\, I_k(\boldsymbol{p},\boldsymbol{q}) = C_k \ \forall\ k \in \{1,\ldots,n\} \big\}$ is diffeomorphic to an n-torus $T^n \equiv \mathbb{S}^1 \times \mathbb{S}^1 \times \cdots \mathbb{S}^1$, on which one can introduce action-angle variables $(\boldsymbol{\phi}, \boldsymbol{J})$ on patches, where $\boldsymbol{\phi}$ are coordinates on $\mathcal{M}_I$ and $\boldsymbol{J}$ are the first integrals, *i.e.* $J_k(I_1,\ldots,I_n) = I_k$.

(ii) The equations of motion are $\dot I_k = 0$ and $\dot\phi_k = \omega_k(I_1,\ldots,I_n)$.

Note that the Liouville-Arnol'd theorem does *not* require that H that $\widetilde H(\boldsymbol{I}) = \sum_k \widetilde H^{(k)}(I_k)$, which would be a trivial state of affairs.

1.7 Adiabatic Invariants

1.7.1 Slow perturbations

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

[10] Clearly if $\boldsymbol{\ell}\cdot\boldsymbol{\nu}(\boldsymbol{J}) = 0$, then replacing $\boldsymbol{\ell}$ by $p\boldsymbol{\ell}$ for any $p \in \mathbb{Z}$ also satisfies the resonance condition.

[11] Two first integrals I_k and I_l whose Poisson bracket $\{I_k, I_l\} = 0$ vanishes are said to be *in involution*.

[12] See chapter 1 of http://www.damtp.cam.ac.uk/user/md327/ISlecture_notes_2012.pdf for a proof.

write $H = H\big(\boldsymbol{q}, \boldsymbol{p}\,;\lambda(t)\big)$. All explicit time-dependence to H comes through $\lambda(t)$. Typically, a dimensionless parameter ϵ may be associated with the perturbation:

$$\epsilon = \frac{1}{\omega_0}\left|\frac{d\ln\lambda}{dt}\right| \quad , \tag{1.161}$$

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 harmonic 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,

$$E = \nu J = \tfrac{1}{2} mgl\,\theta_0^2 \quad \Rightarrow \quad \theta_0(l) \approx \frac{2J}{m\sqrt{g}\,l^{3/2}} \quad , \tag{1.162}$$

where $\theta_0(l)$ is the amplitude of the oscillation. Adiabatic invariance of J thus entails $\theta_0(\ell) \propto l^{-3/2}$.

Consider an $n = 1$ system, and suppose that for fixed λ the Hamiltonian is transformed to action-angle variables via the generator $S(q, J; \lambda)$. Now let $\lambda = \lambda(t)$. $S\big(q, J; \lambda(t)\big)$ is still a type-II generating function of a canonical transformation. The resulting transformed Hamiltonian is

$$\widetilde{H}(\phi, J, t) = H(J;\lambda) + \frac{\partial S}{\partial\lambda}\,\frac{d\lambda}{dt} \quad , \tag{1.163}$$

where

$$H(J;\lambda) = H\big(q(\phi, J;\lambda), p(\phi, J;\lambda);\lambda\big) \tag{1.164}$$

is a function only of J and the instantaneous value of λ. Hamilton's equations are now

$$\begin{aligned} \dot{\phi} &= +\frac{\partial\widetilde{H}}{\partial J} = \nu(J;\lambda) + \frac{\partial^2 S}{\partial\lambda\,\partial J}\,\frac{d\lambda}{dt} \\ \dot{J} &= -\frac{\partial\widetilde{H}}{\partial\phi} = -\frac{\partial^2 S}{\partial\lambda\,\partial\phi}\,\frac{d\lambda}{dt} \quad , \end{aligned} \tag{1.165}$$

where $\nu(J;\lambda) \equiv \partial H(J;\lambda)/\partial J$, and where $S(\phi, J;\lambda) = S\big(q(\phi, J;\lambda), J;\lambda\big)$ The second of eqns. 1.165 may then be Fourier decomposed as

$$\dot{J} = -i\dot{\lambda}\sum_{m=-\infty}^{\infty} m\,\frac{\partial S_m(J;\lambda)}{\partial\lambda}\,e^{im\phi} \quad , \tag{1.166}$$

hence

$$\Delta J = J(t = +\infty) - J(t = -\infty) = \sum_{m=-\infty}^{\infty}(-im)\int\limits_{-\infty}^{\infty}\!dt\,\frac{\partial S_m(J;\lambda)}{\partial\lambda}\,\frac{d\lambda}{dt}\,e^{im\phi} \quad . \tag{1.167}$$

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

$$\mathcal{I}_m = \int\limits_{-\infty}^{\infty}\!dt\,\left\{\frac{\partial S_m(J;\lambda)}{\partial\lambda}\,\frac{d\lambda}{dt}\right\}e^{im\nu t} \quad . \tag{1.168}$$

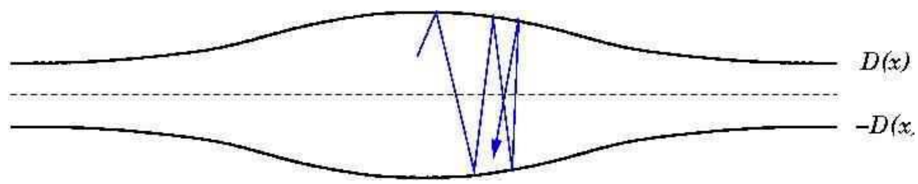

Figure 1.4: A mechanical mirror.

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 singularities in the complex t plane lies at $\mathrm{Im}\, t = \pm\tau$, where $|\nu\tau| \gg 1$. In this case $\mathcal{I}_m$ behaves as $\exp(-|m|\nu\tau)$. Consider, for example, the Lorentzian,

$$f(t) = \frac{1}{\pi}\,\frac{\tau}{t^2+\tau^2} \quad\Rightarrow\quad \int\limits_{-\infty}^{\infty}\!dt\, f(t)\, e^{im\nu t} = e^{-|m\nu\tau|} \quad , \tag{1.169}$$

which is exponentially small in the dimensionless product $|\nu\tau|$. 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 $\lambda(t)$.

1.7.2 Example: mechanical mirror

Consider a two-dimensional version of a mechanical mirror, depicted in fig. 1.4. A particle bounces between two curves, $y = \pm D(x)$, where $|D'(x)| \ll 1$. The bounce time given by $\tau_{\mathrm{b}\perp} = 2D/v_y$. 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,

$$J = \frac{1}{2\pi}\int\limits_{-D}^{D}\!dy\, m\, v_y + \frac{1}{2\pi}\int\limits_{D}^{-D}\!dy\, m\,(-v_y) = \frac{2}{\pi}\, m v_y\, D(x) \quad . \tag{1.170}$$

Thus,

$$E = \tfrac{1}{2}m\big(v_x^2 + v_y^2\big) = \tfrac{1}{2}m v_x^2 + \frac{\pi^2 J^2}{8mD^2(x)} \quad , \tag{1.171}$$

or

$$v_x^2 = \frac{2E}{m} - \left(\frac{\pi J}{2mD(x)}\right)^{\!2} \quad . \tag{1.172}$$

The particle is reflected in the throat of the device at horizontal coordinate x^*, where

$$D(x^*) = \frac{\pi J}{\sqrt{8mE}} \quad . \tag{1.173}$$

1.7.3 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

$$\frac{mv_\perp^2}{\rho} = \frac{e}{c}\,v_\perp B \quad\Rightarrow\quad \rho = \frac{mcv_\perp}{eB} \quad , \tag{1.174}$$

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

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

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = \boldsymbol{\nabla}_\perp\!\cdot\boldsymbol{B}_\perp + \frac{\partial B_z}{\partial z} = 0 \quad . \tag{1.175}$$

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 $T = 2\pi/\omega_\mathrm{c}$, then the system possesses an adiabatic invariant:

$$\begin{aligned} J &= \frac{1}{2\pi}\oint\limits_{\mathcal{C}} \boldsymbol{p}\cdot d\boldsymbol{\ell} = \frac{1}{2\pi}\oint\limits_{\mathcal{C}} \big(m\boldsymbol{v} + \tfrac{e}{c}\boldsymbol{A}\big)\cdot d\boldsymbol{\ell} \\ &= \frac{m}{2\pi}\oint\limits_{\mathcal{C}} \boldsymbol{v}\cdot d\boldsymbol{\ell} + \frac{e}{2\pi c}\oint\limits_{\mathrm{int}(\mathcal{C})} \boldsymbol{B}\cdot\hat{\boldsymbol{n}}\,d\Sigma \quad . \end{aligned} \tag{1.176}$$

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

$$\begin{aligned} J &= -\frac{m}{2\pi}\cdot\frac{\rho e B_z}{mc}\cdot 2\pi\rho + \frac{e}{2\pi c}\cdot B_z\cdot\pi\rho^2 \\ &= -\frac{eB_z\rho^2}{2c} = -\frac{e}{2\pi c}\cdot\Phi_B(\mathcal{C}) = -\frac{m^2v_\perp^2 c}{2eB_z} \quad , \end{aligned} \tag{1.177}$$

where $\Phi_B(\mathcal{C})$ is the magnetic flux enclosed by $\mathcal{C}$. The energy is $E = \frac{1}{2}mv_\perp^2 + \frac{1}{2}mv_z^2$, hence we have

$$v_z = \sqrt{\frac{2}{m}(E - MB)} \quad . \tag{1.178}$$

where

$$M \equiv -\frac{e}{mc}\,J = \frac{e^2}{2\pi mc^2}\,\Phi_B(\mathcal{C}) \tag{1.179}$$

is the *magnetic moment*. Note that v_z vanishes when $B = B_\mathrm{max} = E/M$. When this limit is reached, the particle turns around. This is a *magnetic mirror*. A pair of magnetic mirrors may be used to confine charged particles in a *magnetic bottle*, depicted in fig. 1.5.

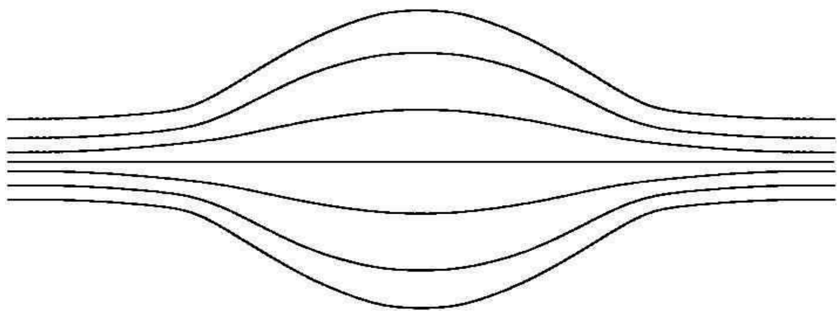

Figure 1.5: $\boldsymbol{B}$ field lines in a magnetic bottle.

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, for J and E, guarantee $v_{\parallel}^2(z) + v_{\perp}^2(z) = v_{\parallel,0}^2 + v_{\perp,0}^2$ and

$$\frac{v_{\perp}^2(z)}{B_{\parallel}(z)} = \frac{v_{\perp,0}^2}{B_{\parallel,0}} \quad . \tag{1.180}$$

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

$$B_{\parallel}(z^*) = B_{\parallel,0}\left(1 + \frac{v_{\parallel,0}^2}{v_{\perp,0}^2}\right)^{1/2} \quad . \tag{1.181}$$

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

1.7.4 Resonances

When $n > 1$, we have

$$\begin{aligned} \dot{J}^\alpha &= -i\dot{\lambda} \sum_{\boldsymbol{m}\in\mathbb{Z}^n} m^\alpha \, \frac{\partial S_{\boldsymbol{m}}(\boldsymbol{J};\lambda)}{\partial\lambda} \, e^{i\boldsymbol{m}\cdot\boldsymbol{\phi}} \\ \Delta J^\alpha &= -i \sum_{\boldsymbol{m}\in\mathbb{Z}^n} m^\alpha \int\limits_{-\infty}^{\infty} dt \, \frac{\partial S_{\boldsymbol{m}}(\boldsymbol{J};\lambda)}{\partial\lambda} \, \frac{d\lambda}{dt} \; e^{i\boldsymbol{m}\cdot\boldsymbol{\nu}t} \, e^{i\boldsymbol{m}\cdot\boldsymbol{\beta}} \quad . \end{aligned} \tag{1.182}$$

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

1.8 Canonical Perturbation Theory

1.8.1 Canonical transformations and perturbation theory

Suppose we have a Hamiltonian

$$H(\boldsymbol{\xi},t) = H_0(\boldsymbol{\xi},t) + \epsilon H_1(\boldsymbol{\xi},t) \quad , \tag{1.183}$$

where ϵ is a small dimensionless parameter. Let's implement a type-II transformation, generated by $S(\boldsymbol{q},\boldsymbol{P},t)$:[13]

$$\widetilde{H}(\boldsymbol{Q},\boldsymbol{P},t) = H(\boldsymbol{q},\boldsymbol{p},t) + \frac{\partial}{\partial t}\, S(\boldsymbol{q},\boldsymbol{P},t) \quad . \tag{1.184}$$

Let's expand everything in powers of ϵ:

$$\begin{aligned} q_\sigma &= Q_\sigma + \epsilon\, q_{1,\sigma} + \epsilon^2\, q_{2,\sigma} + \ldots \\ p_\sigma &= P_\sigma + \epsilon\, p_{1,\sigma} + \epsilon^2\, p_{2,\sigma} + \ldots \\ \widetilde{H} &= \widetilde{H}_0 + \epsilon \widetilde{H}_1 + \epsilon^2 \widetilde{H}_2 + \ldots \\ S &= \underbrace{q_\sigma P_\sigma}_{\substack{\text{identity} \\ \text{transformation}}} + \epsilon S_1 + \epsilon^2 S_2 + \ldots \quad . \end{aligned} \tag{1.185}$$

Then

$$\begin{aligned} Q_\sigma = \frac{\partial S}{\partial P_\sigma} &= q_\sigma + \epsilon\, \frac{\partial S_1}{\partial P_\sigma} + \epsilon^2\, \frac{\partial S_2}{\partial P_\sigma} + \ldots \\ &= Q_\sigma + \bigg(q_{1,\sigma} + \frac{\partial S_1}{\partial P_\sigma}\bigg)\epsilon + \bigg(q_{2,\sigma} + \frac{\partial S_2}{\partial P_\sigma}\bigg)\epsilon^2 + \ldots \end{aligned} \tag{1.186}$$

and

$$\begin{aligned} p_\sigma = \frac{\partial S}{\partial q_\sigma} &= P_\sigma + \epsilon\, \frac{\partial S_1}{\partial q_\sigma} + \epsilon^2\, \frac{\partial S_2}{\partial q_\sigma} + \ldots \\ &= P_\sigma + \epsilon\, p_{1,\sigma} + \epsilon^2\, p_{2,\sigma} + \ldots \quad . \end{aligned} \tag{1.187}$$

We therefore conclude, order by order in ϵ,

$$q_{k,\sigma} = -\frac{\partial S_k}{\partial P_\sigma} \quad , \quad p_{k,\sigma} = +\frac{\partial S_k}{\partial q_\sigma} \quad . \tag{1.188}$$

Now let's expand the Hamiltonian:

$$\begin{aligned} \widetilde{H}(\boldsymbol{Q},\boldsymbol{P},t) &= H_0(\boldsymbol{q},\boldsymbol{p},t) + \epsilon H_1(q,p,t) + \frac{\partial S}{\partial t} \\ &= H_0(\boldsymbol{Q},\boldsymbol{P},t) + \frac{\partial H_0}{\partial Q_\sigma}\,(q_\sigma - Q_\sigma) + \frac{\partial H_0}{\partial P_\sigma}\,(p_\sigma - P_\sigma) + \epsilon H_1(\boldsymbol{Q},\boldsymbol{P},t) + \epsilon\, \frac{\partial}{\partial t}\, S_1(\boldsymbol{Q},\boldsymbol{P},t) + \mathcal{O}(\epsilon^2) \quad . \end{aligned} \tag{1.189}$$

[13] Here $S(\boldsymbol{q},\boldsymbol{P},t)$ is not meant to signify Hamilton's principal function.

Collecting terms, we have

$$\begin{aligned}\widetilde{H}(\boldsymbol{Q},\boldsymbol{P},t) &= H_0(\boldsymbol{Q},\boldsymbol{P},t) + \left(-\frac{\partial H_0}{\partial Q_\sigma}\,\frac{\partial S_1}{\partial P_\sigma} + \frac{\partial H_0}{\partial P_\sigma}\,\frac{\partial S_1}{\partial Q_\sigma} + \frac{\partial S_1}{\partial t} + H_1\right)\epsilon + \mathcal{O}(\epsilon^2) \\ &= H_0(\boldsymbol{Q},\boldsymbol{P},t) + \left(H_1 + \big\{S_1, H_0\big\} + \frac{\partial S_1}{\partial t}\right)\epsilon + \mathcal{O}(\epsilon^2) \quad .\end{aligned} \tag{1.190}$$

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

$$\widetilde{H}(\boldsymbol{Q},\boldsymbol{P},t) = \widetilde{H}_0(\boldsymbol{Q},\boldsymbol{P},t) + \epsilon \widetilde{H}_1(\boldsymbol{Q},\boldsymbol{P},t) + \ldots \tag{1.191}$$

with

$$\widetilde{H}_0(\boldsymbol{Q},\boldsymbol{P},t) = H_0(\boldsymbol{Q},\boldsymbol{P},t) \tag{1.192}$$

$$\widetilde{H}_1(\boldsymbol{Q},\boldsymbol{P},t) = H_1 + \big\{S_1, H_0\big\} + \frac{\partial S_1}{\partial t} \quad . \tag{1.193}$$

The problem, though, is this: we have one equation, eqn, 1.193, for the two unknowns $\widetilde{H}_1$ and S_1. Thus, the problem is underdetermined. Of course, we could choose $\widetilde{H}_1 = 0$, for example. But we might just as well demand that $\widetilde{H}_1$ satisfy some other desideratum, such as that $\widetilde{H}_0 + \epsilon \widetilde{H}_1$ be 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,

$$i\hbar \frac{\partial \psi}{\partial t} = (\mathcal{H}_0 + \epsilon\, \mathcal{H}_1)\, \psi \quad , \tag{1.194}$$

and define χ by $\psi \equiv e^{iS/\hbar}\,\chi$, with

$$S = \epsilon\, S_1 + \epsilon^2\, S_2 + \ldots \quad . \tag{1.195}$$

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

$$i\hbar \frac{\partial \chi}{\partial t} = \mathcal{H}_0\, \chi + \epsilon \left(\mathcal{H}_1 + \frac{1}{i\hbar}\big[S_1, \mathcal{H}_0\big] + \frac{\partial S_1}{\partial t}\right)\chi + \ldots \equiv \widetilde{\mathcal{H}}\, \chi \quad , \tag{1.196}$$

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

$$\{A,B\} \longleftrightarrow \frac{1}{i\hbar}\,[A,B] \quad . \tag{1.197}$$

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

1.8.2 Canonical perturbation theory for $n = 1$ systems

Here and henceforth we shall assume $H(\boldsymbol{q},\boldsymbol{p},t) = H(\boldsymbol{q},\boldsymbol{p})$ is time-independent, and we write the perturbed Hamiltonian as

$$H(\boldsymbol{q},\boldsymbol{p}) = H_0(\boldsymbol{q},\boldsymbol{p}) + \epsilon H_1(\boldsymbol{q},\boldsymbol{p}) \quad . \tag{1.198}$$

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

$$\widetilde{H}_0(\phi_0, J_0) = H_0\big(q(\phi_0,J_0), p(\phi_0,J_0)\big) = \widetilde{H}_0(J_0) \quad . \tag{1.199}$$

We define

$$\widetilde{H}_1(\phi_0, J_0) = H_1\big(q(\phi_0,J_0), p(\phi_0,J_0)\big) \quad . \tag{1.200}$$

We assume that $\widetilde{H} = \widetilde{H}_0 + \epsilon \widetilde{H}_1$ is integrable[14], so it, too, possesses action-angle variables, which we denote by (ϕ, J)[15]. Thus, there must be a canonical transformation taking $(\phi_0, J_0) \to (\phi, J)$, with

$$\widetilde{H}\big(\phi_0(\phi,J), J_0(\phi,J)\big) \equiv E(J) \quad . \tag{1.201}$$

We solve via a type-II canonical transformation:

$$S(\phi_0, J) = \phi_0 J + \epsilon\, S_1(\phi_0, J) + \epsilon^2\, S_2(\phi_0, J) + \ldots \quad , \tag{1.202}$$

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

$$\begin{aligned} J_0 &= \frac{\partial S}{\partial \phi_0} = J + \epsilon\, \frac{\partial S_1}{\partial \phi_0} + \epsilon^2\, \frac{\partial S_2}{\partial \phi_0} + \ldots \\ \phi &= \frac{\partial S}{\partial J} = \phi_0 + \epsilon\, \frac{\partial S_1}{\partial J} + \epsilon^2\, \frac{\partial S_2}{\partial J} + \ldots \quad , \end{aligned} \tag{1.203}$$

and

$$\begin{aligned} E(J) &= E_0(J) + \epsilon\, E_1(J) + \epsilon^2\, E_2(J) + \ldots \\ &= \widetilde{H}_0(\phi_0, J_0) + \epsilon \widetilde{H}_1(\phi_0, J_0) \quad . \end{aligned} \tag{1.204}$$

How is it that the second line terminates after order ϵ while the first line contains terms of order ϵ^k for all $k \geq 0$? The answer is that when we express (ϕ_0, J_0) in terms of (ϕ, J), the canonical transformation itself involve terms to all orders in ϵ, as we see from eqn. 1.203. In general, when a *nonlinear* system is perturbed, the response will include expressions to all orders in the perturbation.

We now expand $\widetilde{H}(\phi_0, J_0)$ in powers of $J_0 - J$, keeping in mind that $\widetilde{H}_0(\phi_0, J_0) = \widetilde{H}_0(J_0)$:

$$\begin{aligned} \widetilde{H}(\phi_0, J_0) &= \widetilde{H}_0(J_0) + \epsilon\, \widetilde{H}_1(\phi_0, J_0) \\ &= \widetilde{H}_0(J) + \frac{\partial \widetilde{H}_0}{\partial J}\,(J_0 - J) + \frac{1}{2}\,\frac{\partial^2 \widetilde{H}_0}{\partial J^2}\,(J_0 - J)^2 + \epsilon\, \widetilde{H}_1(\phi_0, J) + \epsilon\, \frac{\partial \widetilde{H}_1}{\partial J}\bigg|_{\phi_0} (J_0 - J) + \ldots \quad . \end{aligned} \tag{1.205}$$

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

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

Collecting terms,

$$\widetilde{H}(\phi_0,J_0) = \widetilde{H}_0(J) + \left(\widetilde{H}_1 + \frac{\partial \widetilde{H}_0}{\partial J}\,\frac{\partial S_1}{\partial \phi_0}\right)\epsilon + \left(\frac{\partial \widetilde{H}_0}{\partial J}\,\frac{\partial S_2}{\partial \phi_0} + \frac{1}{2}\,\frac{\partial^2 \widetilde{H}_0}{\partial J^2}\left(\frac{\partial S_1}{\partial \phi_0}\right)^2 + \frac{\partial \widetilde{H}_1}{\partial J}\,\frac{\partial S_1}{\partial \phi_0}\right)\epsilon^2 + \ldots \tag{1.206}$$

where all terms on the RHS are expressed as functions of ϕ_0 *and* J. Equating terms, then,

$$\begin{aligned} E_0(J) &= \widetilde{H}_0(J) \\ E_1(J) &= \widetilde{H}_1(\phi_0,J) + \frac{\partial \widetilde{H}_0}{\partial J}\,\frac{\partial S_1}{\partial \phi_0} \\ E_2(J) &= \frac{\partial \widetilde{H}_0}{\partial J}\,\frac{\partial S_2}{\partial \phi_0} + \frac{1}{2}\,\frac{\partial^2 \widetilde{H}_0}{\partial J^2}\left(\frac{\partial S_1}{\partial \phi_0}\right)^2 + \frac{\partial \widetilde{H}_1}{\partial J}\,\frac{\partial S_1}{\partial \phi_0} \quad . \end{aligned} \tag{1.207}$$

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

$$\big\langle f\big(\phi_0\big)\big\rangle = \int\limits_0^{2\pi} \frac{d\phi_0}{2\pi}\, f\big(\phi_0\big) \quad . \tag{1.208}$$

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, we are able to write

$$S_k(\phi_0,J) = \sum_{\ell=-\infty}^{\infty} S_{k,\ell}(J)\, e^{i\ell\phi_0} \tag{1.209}$$

for each $k > 0$, in which case

$$\left\langle \frac{\partial S_k}{\partial \phi_0} \right\rangle = \frac{1}{2\pi}\big[S_k(2\pi,J) - S_k(0,J)\big] = 0 \quad . \tag{1.210}$$

Let's see how this averaging works to the first two orders of the hierarchy. Since $\widetilde{H}_0(J)$ is independent of ϕ_0 and since $\partial S_1/\partial\phi_0$ is periodic, we have

$$E_1(J) = \big\langle \widetilde{H}_1(\phi_0,J)\big\rangle + \frac{\partial \widetilde{H}_0}{\partial J}\,\overbrace{\left\langle \frac{\partial S_1}{\partial \phi_0} \right\rangle}^{\text{this vanishes!}} \tag{1.211}$$

and hence S_1 must satisfy

$$\frac{\partial S_1}{\partial \phi_0} = \frac{\big\langle \widetilde{H}_1\big\rangle - \widetilde{H}_1}{\nu_0(J)} \quad , \tag{1.212}$$

where $\nu_0(J) = \partial \widetilde{H}_0/\partial J$. Clearly the RHS of eqn. 1.212 has zero average, and must be a periodic function of ϕ_0. The solution is $S_1 = S_1(\phi_0,J) + f(J)$, where $f(J)$ is an arbitrary function of J. However, $f(J)$

affects only the difference $\phi - \phi_0$, changing it by a constant value $f'(J)$. So there is no harm in taking $f(J) = 0$.

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

$$E_2(J) = \nu_0(J) \overbrace{\left\langle \frac{\partial S_2}{\partial \phi_0} \right\rangle}^{\text{this vanishes!}} + \frac{1}{2}\,\frac{\partial \nu_0}{\partial J} \left\langle \left(\frac{\partial S_1}{\partial \phi_0} \right)^2 \right\rangle + \left\langle \frac{\partial \widetilde{H}_1}{\partial J}\,\frac{\partial S_1}{\partial \phi_0} \right\rangle \quad . \tag{1.213}$$

The equation for S_2 is then

$$\begin{aligned} \frac{\partial S_2}{\partial \phi_0} = \frac{1}{\nu_0^2(J)} \Bigg\{ \left\langle \frac{\partial \widetilde{H}_1}{\partial J} \right\rangle \langle \widetilde{H}_1 \rangle - \left\langle \frac{\partial \widetilde{H}_1}{\partial J}\,\widetilde{H}_1 \right\rangle - \frac{\partial \widetilde{H}_1}{\partial J}\,\langle \widetilde{H}_1 \rangle + \frac{\partial \widetilde{H}_1}{\partial J}\,\widetilde{H}_1 \\ + \frac{1}{2}\,\frac{\partial \ln \nu_0}{\partial J} \left(\langle \widetilde{H}_1^2 \rangle - 2\,\langle \widetilde{H}_1 \rangle^2 + 2\,\langle \widetilde{H}_1 \rangle\,\widetilde{H}_1 - \widetilde{H}_1^2 \right) \Bigg\} \quad . \end{aligned} \tag{1.214}$$

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

$$E(J) = \widetilde{H}_0(J) + \epsilon\,\langle \widetilde{H}_1 \rangle + \frac{\epsilon^2}{\nu_0(J)} \left\{ \left\langle \frac{\partial \widetilde{H}_1}{\partial J} \right\rangle \langle \widetilde{H}_1 \rangle - \left\langle \frac{\partial \widetilde{H}_1}{\partial J}\,\widetilde{H}_1 \right\rangle + \frac{1}{2}\,\frac{\partial \ln \nu_0}{\partial J} \left(\langle \widetilde{H}_1^2 \rangle - \langle \widetilde{H}_1 \rangle^2 \right) \right\} + \mathcal{O}(\epsilon^3) \quad . \tag{1.215}$$

Note that we don't need S to find $E(J)$! The perturbed frequencies are $\nu(J) = \partial E/\partial J$. 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,

$$(\phi, J) \longrightarrow (\phi_0, J_0) \longrightarrow (q, p) \quad . \tag{1.216}$$

1.8.3 Example : quartic oscillator

Consider a harmonic oscillator with a quartic nonlinearity[16]. The Hamiltonian is

$$H(q,p) = \overbrace{\frac{p^2}{2m} + \tfrac{1}{2} m\nu_0^2\, q^2}^{H_0} + \tfrac{1}{4}\epsilon\alpha q^4 \quad . \tag{1.217}$$

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

$$\phi_0 = \tan^{-1}\!\big(m\nu_0 q/p\big) \qquad , \qquad J_0 = \frac{p^2}{2m\nu_0} + \tfrac{1}{2} m\nu_0 q^2 \tag{1.218}$$

hence

$$q = \sqrt{\frac{2J_0}{m\nu_0}}\,\cos\phi_0 \qquad , \qquad p = \sqrt{2J_0 m\nu_0}\,\sin\phi_0 \quad , \tag{1.219}$$

[16] In §1.11.5 below, we discuss the case of a cubic nonlinearity.

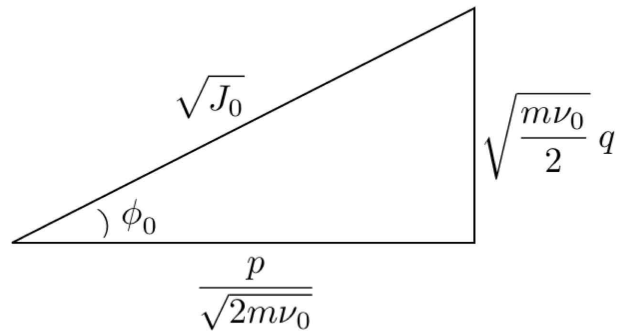

Figure 1.6: Action-angle variables for the harmonic oscillator.

as depicted in fig. 1.6. Note $H_0 = \nu_0\, J_0$. For the full Hamiltonian, we have

$$\begin{aligned}\widetilde{H}(\phi_0, J_0) &= \nu_0 J_0 + \tfrac{1}{4}\epsilon\alpha\bigg(\sqrt{\frac{2J_0}{m\nu_0}}\,\sin\phi_0\bigg)^4 \\ &= \nu_0 J_0 + \frac{\epsilon\alpha}{m^2\nu_0^2}\,J_0^2\,\sin^4\!\phi_0 \equiv H_0(\phi_0, J_0) + \epsilon\widetilde{H}_1(\phi_0, J_0) \quad .\end{aligned} \tag{1.220}$$

We may now evaluate

$$E_1(J) = \big\langle \widetilde{H}_1(\phi_0, J)\big\rangle = \frac{\alpha J^2}{m^2\nu_0^2}\int\limits_0^{2\pi}\frac{d\phi_0}{2\pi}\,\sin^4\!\phi_0 = \frac{3\alpha J^2}{8m^2\nu_0^2} \quad . \tag{1.221}$$

The frequency, to order ϵ, is

$$\nu(J) = \nu_0 + \frac{3\epsilon\alpha J}{4m^2\nu_0^2} \quad . \tag{1.222}$$

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:

$$\nu(A) = \nu_0 + \frac{3\epsilon\alpha A^2}{8m\nu_0} \quad . \tag{1.223}$$

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

$$\begin{aligned}\nu_0\,\frac{\partial S_1}{\partial\phi_0} &= \frac{\alpha J^2}{m^2\nu_0^2}\left(\tfrac{3}{8} - \sin^4\!\phi_0\right) \qquad \Rightarrow \\ S(\phi_0, J) &= \phi_0\,J + \frac{\epsilon\alpha J^2}{8m^2\nu_0^3}\left(3 + 2\sin^2\!\phi_0\right)\sin\phi_0\cos\phi_0 + \mathcal{O}(\epsilon^2) \quad .\end{aligned} \tag{1.224}$$

Thus,

$$\begin{aligned}\phi &= \frac{\partial S}{\partial J} = \phi_0 + \frac{\epsilon\alpha J}{4m^2\nu_0^3}\left(3 + 2\sin^2\!\phi_0\right)\sin\phi_0\cos\phi_0 + \mathcal{O}(\epsilon^2) \\ J_0 &= \frac{\partial S}{\partial\phi_0} = J + \frac{\epsilon\alpha J^2}{8m^2\nu_0^3}\left(4\cos 2\phi_0 - \cos 4\phi_0\right) + \mathcal{O}(\epsilon^2) \quad .\end{aligned} \tag{1.225}$$

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

$$\begin{aligned} J &= J_0 - \frac{\epsilon\alpha J_0^2}{8m^2\nu_0^3}\big(4\cos 2\phi_0 - \cos 4\phi_0\big) + \mathcal{O}(\epsilon^2) \\ \phi &= \phi_0 + \frac{\epsilon\alpha J_0}{8m^2\nu_0^3}\big(3 + 2\sin^2\!\phi_0\big)\,\sin 2\phi_0 + \mathcal{O}(\epsilon^2) \quad . \end{aligned} \tag{1.226}$$

Writing $q = (2J_0/m\nu_0)^{1/2}\sin\phi_0$ and $p = (2m\nu_0 J_0)^{1/2}\cos\phi_0$, one can substitute the above relations, replacing (ϕ_0, J_0) with (ϕ, J) in the $\mathcal{O}(\epsilon)$ terms on the RHS of each equation, to obtain (q,p) in terms of (ϕ, J), valid to $\mathcal{O}(\epsilon)$.

1.8.4 $n > 1$ systems : degeneracies and resonances

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

$$\begin{aligned} J_0^\alpha &= \frac{\partial S}{\partial \phi_0^\alpha} = J^\alpha + \epsilon\,\frac{\partial S_1}{\partial \phi_0^\alpha} + \epsilon^2\,\frac{\partial S_2}{\partial \phi_0^\alpha} + \ldots \\ \phi^\alpha &= \frac{\partial S}{\partial J^\alpha} = \phi_0^\alpha + \epsilon\,\frac{\partial S_1}{\partial J^\alpha} + \epsilon^2\,\frac{\partial S_2}{\partial J^\alpha} + \ldots \end{aligned} \tag{1.227}$$

and

$$\begin{aligned} E_0(\boldsymbol{J}) &= \widetilde{H}_0(\boldsymbol{J}) \\ E_1(\boldsymbol{J}) &= \widetilde{H}_1 + \nu_0^\alpha\,\frac{\partial S_1}{\partial \phi_0^\alpha} \\ E_2(\boldsymbol{J}) &= \nu_0^\alpha\,\frac{\partial S_2}{\partial \phi_0^\alpha} + \frac{1}{2}\,\frac{\partial \nu_0^\alpha}{\partial J^\beta}\,\frac{\partial S_1}{\partial \phi_0^\alpha}\,\frac{\partial S_1}{\partial \phi_0^\beta} + \frac{\partial \widetilde{H}_1}{\partial J^\alpha}\,\frac{\partial S_1}{\partial \phi_0^\alpha} \quad , \end{aligned} \tag{1.228}$$

where $\nu_0^\alpha(\boldsymbol{J}) = \partial\widetilde{H}_0(\boldsymbol{J})/\partial J^\alpha$. We now implement the averaging procedure, with

$$\big\langle f(\phi_0^1,\ldots,\phi_0^n,J^1,\ldots,J^n)\big\rangle = \int\limits_0^{2\pi}\!\frac{d\phi_0^1}{2\pi}\,\cdots\int\limits_0^{2\pi}\!\frac{d\phi_0^n}{2\pi}\,f\big(\phi_0^1,\ldots,\phi_0^n,J^1,\ldots,J^n\big) \quad . \tag{1.229}$$

The equation for S_1 is

$$\nu_0^\alpha\,\frac{\partial S_1}{\partial \phi_0^\alpha} = \big\langle \widetilde{H}_1(\boldsymbol{\phi}_0,\boldsymbol{J})\big\rangle - \widetilde{H}_1(\boldsymbol{\phi}_0,\boldsymbol{J}) \equiv -{\sum_{\boldsymbol{\ell}}}'\,\hat{V}_{\boldsymbol{\ell}}(\boldsymbol{J})\,e^{i\boldsymbol{\ell}\cdot\boldsymbol{\phi}_0} \quad , \tag{1.230}$$

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

$$S_1(\boldsymbol{\phi}_0,\boldsymbol{J}) = i\,{\sum_{l}}'\,\frac{\hat{V}_{\boldsymbol{\ell}}(\boldsymbol{J})}{\boldsymbol{\ell}\cdot\boldsymbol{\nu}_0(\boldsymbol{J})}\,e^{i\boldsymbol{\ell}\cdot\boldsymbol{\phi}_0} \quad . \tag{1.231}$$

where $\boldsymbol{\ell}\cdot\boldsymbol{\nu}_0 = \sum_{\alpha=1}^n l^\alpha \nu_0^\alpha$. When two or more of the frequencies $\nu_0^\alpha(\boldsymbol{J})$ are *commensurate*, there exists a set of integers $\boldsymbol{\ell}$ 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 $|\boldsymbol{\ell}|$ the denominator can become arbitrarily small.

1.8.5 Nonlinear oscillator with two degrees of freedom

As an example of how to implement canonical perturbation theory for $n > 1$, consider the nonlinear oscillator system,

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \tfrac{1}{2}m\omega_1^2\, q_1^2 + \tfrac{1}{2}m\omega_2^2\, q_2^2 + \tfrac{1}{4}\epsilon\, b\,\omega_1^2\,\omega_2^2\, q_1^2\, q_2^2 \quad . \tag{1.232}$$

Writing $H = H_0 + \epsilon H_1$, we have, in terms of the action-angle variables $(\phi_0^{(1,2)}, J_0^{(1,2)})$,

$$\widetilde{H}_0(\boldsymbol{J}_0) = \omega_1 J_0^{(1)} + \omega_2 J_0^{(2)} \tag{1.233}$$

with $q_k = (2J_0^k/m\omega_k)^{1/2}\sin\phi_0^k$ and $p_k = (2m\omega_k J_0^k)^{1/2}\cos\phi_0^k$ with $k \in \{1,2\}$. We then have

$$\widetilde{H}_1(\boldsymbol{\phi}_0, \boldsymbol{J}) = b\,\omega_1\,\omega_2\, J^{(1)} J^{(2)}\, \sin^2\!\phi_0^{(1)}\, \sin^2\!\phi_0^{(2)} \quad . \tag{1.234}$$

We therefore have $E(\boldsymbol{J}) = E_0(\boldsymbol{J}) + \epsilon E_1(\boldsymbol{J})$ with $E_0(\boldsymbol{J}) = H_0(\boldsymbol{J}) = \omega_1 J^{(1)} + \omega_2 J^{(2)}$ and

$$E_1(\boldsymbol{J}) = \big\langle \widetilde{H}_1(\boldsymbol{\phi}_0, \boldsymbol{J})\big\rangle = \tfrac{1}{4}b\,\omega_1\omega_2\, J^{(1)} J^{(2)} \quad . \tag{1.235}$$

Next, we work out the generator $S_1(\boldsymbol{\phi}_0, \boldsymbol{J})$ from eqn. 1.230:

$$\begin{aligned}\big\langle \widetilde{H}_1(\boldsymbol{\phi}_0, \boldsymbol{J})\big\rangle - \widetilde{H}_1(\boldsymbol{\phi}_0, \boldsymbol{J}) &= b\,\omega_1\omega_2\, J^{(1)} J^{(2)} \left\{ \tfrac{1}{4} - \sin^2\!\phi_0^{(1)}\, \sin^2\!\phi_0^{(2)} \right\} \\ &= b\,\omega_1\omega_2\, J^{(1)} J^{(2)} \Big\{ -\tfrac{1}{2}\cos\big(2\phi_0^{(1)} + 2\phi_0^{(2)}\big) - \tfrac{1}{2}\cos\big(2\phi_0^{(1)} - 2\phi_0^{(2)}\big) \\ &\qquad\qquad + \cos 2\phi_0^{(1)} + \cos 2\phi_0^{(2)} \Big\} \quad ,\end{aligned} \tag{1.236}$$

and therefore, from eqn. 1.231,

$$S_1(\boldsymbol{\phi}_0, \boldsymbol{J}) = \tfrac{1}{4}b\,\omega_1\omega_2\, J^{(1)} J^{(2)} \left\{ -\frac{\sin(2\phi_0^{(1)} + 2\phi_0^{(2)})}{\omega_1 + \omega_2} - \frac{\sin(2\phi_0^{(1)} - 2\phi_0^{(2)})}{\omega_1 - \omega_2} + \frac{2\sin 2\phi_0^{(1)}}{\omega_1} + \frac{2\sin 2\phi_0^{(2)}}{\omega_2} \right\} \quad . \tag{1.237}$$

We see that there is a vanishing denominator if $\omega_1 = \omega_2$.

1.8.6 Periodic time-dependent perturbations

Periodic time-dependent perturbations present a similar problem. Consider the system

$$H(\boldsymbol{\phi}, \boldsymbol{J}, t) = H_0(\boldsymbol{J}) + \epsilon\, V(\boldsymbol{\phi}, \boldsymbol{J}, t) \quad , \tag{1.238}$$

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

$$\begin{aligned} V(\boldsymbol{\phi}, \boldsymbol{J}, t) &= \sum_k \hat{V}_k(\boldsymbol{\phi}, \boldsymbol{J})\, e^{-ik\Omega t} \\ &= \sum_{\boldsymbol{k}} \sum_{\boldsymbol{\ell}} \hat{V}_{k,\boldsymbol{\ell}}(\boldsymbol{J})\, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\phi}}\, e^{-ik\Omega t} \quad .\end{aligned} \tag{1.239}$$

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 $V^*_{k,\boldsymbol{\ell}} = \hat{V}_{-\boldsymbol{k},-l}$. The equations of motion are

$$
\begin{aligned}
\dot{J}^\alpha &= -\frac{\partial H}{\partial \phi^\alpha} = -i\epsilon \sum_{k,\boldsymbol{\ell}} l^\alpha \, \hat{V}_{k,\boldsymbol{\ell}}(\boldsymbol{J}) \, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\phi}} \, e^{-ik\Omega t} \\
\dot{\phi}^\alpha &= +\frac{\partial H}{\partial J^\alpha} = \nu_0^\alpha(\boldsymbol{J}) + \epsilon \sum_{k,\boldsymbol{\ell}} \frac{\partial \hat{V}_{k,\boldsymbol{\ell}}(\boldsymbol{J})}{\partial J^\alpha} \, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\phi}} \, e^{-ik\Omega t} \quad .
\end{aligned}
\tag{1.240}
$$

We now expand in ϵ:

$$
\begin{aligned}
\phi^\alpha &= \phi_0^\alpha + \epsilon \, \phi_1^\alpha + \epsilon^2 \, \phi_2^\alpha + \ldots \\
J^\alpha &= J_0^\alpha + \epsilon \, J_1^\alpha + \epsilon^2 \, J_2^\alpha + \ldots \quad .
\end{aligned}
\tag{1.241}
$$

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

$$
\dot{J}_1^\alpha = -i \sum_{k,l} l^\alpha \, \hat{V}_{k,\boldsymbol{\ell}}(\boldsymbol{J}_0) \, e^{i(\boldsymbol{\ell}\cdot\boldsymbol{\nu}_0 - k\Omega)\, t} \, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\beta}_0}
\tag{1.242}
$$

and

$$
\dot{\phi}_1^\alpha = \frac{\partial \nu_0^\alpha}{\partial J^\beta} \, J_1^\beta + \sum_{k,\boldsymbol{\ell}} \frac{\partial \hat{V}_{k,\boldsymbol{\ell}}(\boldsymbol{J})}{\partial J^\alpha} \, e^{i(\boldsymbol{\ell}\cdot\boldsymbol{\nu}_0 - k\Omega)\, t} \, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\beta}_0} \quad ,
\tag{1.243}
$$

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

$$
\begin{aligned}
J_1^\alpha &= \sum_{k,\boldsymbol{\ell}} \frac{l^\alpha \, \hat{V}_{k,\boldsymbol{\ell}}(\boldsymbol{J}_0)}{k\Omega - \boldsymbol{\ell} \cdot \boldsymbol{\nu}_0} \, e^{i(\boldsymbol{\ell}\cdot\boldsymbol{\nu}_0 - k\Omega)\, t} \, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\beta}_0} \\
\phi_1^\alpha &= \left\{ \frac{\partial \nu_0^\alpha}{\partial J_0^\beta} \, \frac{l^\beta \, \hat{V}_{k,\boldsymbol{\ell}}(\boldsymbol{J}_0)}{(k\Omega - \boldsymbol{\ell}\cdot\boldsymbol{\nu}_0)^2} + \frac{\partial \hat{V}_{k,\boldsymbol{\ell}}(\boldsymbol{J}_0)}{\partial J_0^\alpha} \, \frac{1}{k\Omega - \boldsymbol{\ell}\cdot\boldsymbol{\nu}_0} \right\} e^{i(\boldsymbol{\ell}\cdot\boldsymbol{\nu}_0 - k\Omega)\, t} \, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\beta}_0} \quad .
\end{aligned}
\tag{1.244}
$$

When the resonance condition $k\Omega = \boldsymbol{\ell}\cdot\boldsymbol{\nu}_0(\boldsymbol{J}_0)$ is satisfied, the denominators vanish, and the perturbation theory breaks down.

1.8.7 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

$$
H = \frac{1}{2m}\big(\boldsymbol{p} - \tfrac{e}{c}\boldsymbol{A}\big)^2 + \epsilon \, e\hat{V}_0 \, \cos(k_\perp x + k_z z - \omega t) \quad ,
\tag{1.245}
$$

where ϵ is a dimensionless expansion parameter. This is an $n = 3$ system with canonical pairs (x, p_x), (y, p_y), and (z, p_z).

Working in the gauge $\boldsymbol{A} = Bx\hat{\boldsymbol{y}}$, we transform the first two pairs (x, y, p_x, p_y) to convenient variables (Q, P, ϕ, J), explicitly discussed in §1.11.2 below), such that

$$H = \omega_c J + \frac{p_z^2}{2m} + \epsilon\, e\hat{V}_0 \, \cos\!\bigg(k_z z + \frac{k_\perp P}{m\omega_c} + k_\perp \sqrt{\frac{2J}{m\omega_c}}\, \sin\phi - \omega t\bigg) \quad . \tag{1.246}$$

Here,

$$x = \frac{P}{m\omega_c} + \sqrt{\frac{2J}{m\omega_c}}\, \sin\phi \qquad , \qquad y = Q + \sqrt{\frac{2J}{m\omega_c}}\, \cos\phi \quad , \tag{1.247}$$

with $\omega_c = eB/mc$, the cyclotron frequency. Here, (Q, P) describe the *guiding center* degrees of freedom, and (ϕ, J) the *cyclotron* degrees of freedom.

We now make a mixed canonical transformation, generated by

$$F = \phi\tilde{J} + \Big(k_z z + \frac{k_\perp P}{m\omega_c} - \omega t\Big)\tilde{K} - P\tilde{Q} \quad , \tag{1.248}$$

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

$$\tilde{\phi} = \frac{\partial F}{\partial \tilde{J}} = \phi \qquad\qquad J = \frac{\partial F}{\partial \phi} = \tilde{J} \tag{1.249}$$

$$Q = -\frac{\partial F}{\partial P} = -\frac{k_\perp \tilde{K}}{m\omega_c} + \tilde{Q} \qquad\qquad \tilde{P} = -\frac{\partial F}{\partial \tilde{Q}} = P \tag{1.250}$$

$$\tilde{\psi} = \frac{\partial F}{\partial \tilde{K}} = k_z z + \frac{k_\perp P}{m\omega_c} - \omega t \qquad\qquad p_z = \frac{\partial F}{\partial z} = k_z \tilde{K} \quad . \tag{1.251}$$

The transformed Hamiltonian is

$$\begin{aligned} H' &= H + \frac{\partial F}{\partial t} \\ &= \omega_c \tilde{J} + \frac{k_z^2}{2m}\, \tilde{K}^2 - \omega\tilde{K} + \epsilon\, e\hat{V}_0 \, \cos\left(\tilde{\psi} + k_\perp \sqrt{\frac{2\tilde{J}}{m\omega_c}}\, \sin\tilde{\phi}\right) \quad . \end{aligned} \tag{1.252}$$

Note the guiding center pair $(\tilde{Q}, \tilde{P})$ doesn't appear in the transformed Hamiltonian H'.

We now drop the tildes and the prime on H and write $H = H_0 + \epsilon\, H_1$, with

$$\begin{aligned} H_0 &= \omega_c J + \frac{k_z^2}{2m}\, K^2 - \omega K \\ H_1 &= e\hat{V}_0 \, \cos\!\bigg(\psi + k_\perp \sqrt{\frac{2J}{m\omega_c}}\, \sin\phi\bigg) \quad . \end{aligned} \tag{1.253}$$

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

$$\omega_\phi^0 = \frac{\partial H_0}{\partial J} = \omega_c \qquad , \qquad \omega_\psi^0 = \frac{\partial H_0}{\partial K} = \frac{k_z^2 K}{m} - \omega = k_z v_z - \omega \quad , \tag{1.254}$$

where $v_z = p_z/m$ is the z-component of the particle's velocity.

We are now in position to implement the time-independent canonical perturbation theory approach. We invoke a generator

$$S(\phi,\mathcal{J},\psi,\mathcal{K}) = \phi\mathcal{J} + \psi\mathcal{K} + \epsilon\, S_1(\phi,\mathcal{J},\psi,\mathcal{K}) + \epsilon^2 S_2(\phi,\mathcal{J},\psi,\mathcal{K}) + \ldots \tag{1.255}$$

to transform from (ϕ, J, ψ, K) to $(\Phi, \mathcal{J}, \Psi, \mathcal{K})$. We must now solve eqn. 1.230:

$$\omega^0_\phi\, \frac{\partial S_1}{\partial\phi} + \omega^0_\psi\, \frac{\partial S_1}{\partial\psi} = \langle\, H_1 \,\rangle - H_1 \quad . \tag{1.256}$$

That is,

$$\begin{aligned}
\omega_{\rm c}\, \frac{\partial S_1}{\partial\phi} + \bigg(\frac{k_z^2\,\mathcal{K}}{m} - \omega\bigg)\frac{\partial S_1}{\partial\psi} &= -eA_0 \cos\!\bigg(\psi + k_\perp\sqrt{\frac{2\mathcal{J}}{m\omega_{\rm c}}}\,\sin\phi\bigg)\\
&= -eA_0 \sum_{n=-\infty}^{\infty} J_n\bigg(k_\perp\sqrt{\frac{2\mathcal{J}}{m\omega_{\rm c}}}\bigg)\,\cos(\psi + n\phi) \quad ,
\end{aligned}$$

where we have used the result

$$e^{iz\sin\theta} = \sum_{n=-\infty}^{\infty} J_n(z)\, e^{in\theta} \quad . \tag{1.257}$$

The solution for S_1 is then

$$S_1(\phi,\mathcal{J},\psi,\mathcal{K}) = \sum_n \frac{e\hat{V}_0}{\omega - n\omega_{\rm c} - k_z^2\,\mathcal{K}/m}\, J_n\bigg(k_\perp\sqrt{\frac{2\mathcal{J}}{m\omega_{\rm c}}}\bigg)\,\sin(\psi + n\phi) \quad . \tag{1.258}$$

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

$$\begin{aligned}
J &= \mathcal{J} + \epsilon\,\frac{\partial S_1}{\partial\phi} + \mathcal{O}(\epsilon^2)\\
K &= \mathcal{K} + \epsilon\,\frac{\partial S_1}{\partial\psi} + \mathcal{O}(\epsilon^2) \quad .
\end{aligned} \tag{1.259}$$

Defining the dimensionless variable

$$\lambda \equiv k_\perp\sqrt{\frac{2J}{m\omega_{\rm c}}} \quad , \tag{1.260}$$

we obtain the result[17]

$$\bigg(\frac{m\omega_{\rm c}^2}{2e\hat{V}_0 k_\perp^2}\bigg)\varLambda^2 = \bigg(\frac{m\omega_{\rm c}^2}{2e\hat{V}_0 k_\perp^2}\bigg)\lambda^2 - \epsilon\sum_n \frac{nJ_n(\varLambda)\cos(\psi + n\phi)}{\omega/\omega_{\rm c} - n - k_z^2\,\mathcal{K}/m\omega_{\rm c}} + \mathcal{O}(\epsilon^2) \quad , \tag{1.261}$$

where $\varLambda \equiv k_\perp(2\mathcal{J}/m\omega_{\rm c})^{1/2}$.

We see that resonances occur whenever

$$\frac{\omega}{\omega_{\rm c}} - \frac{k_z^2\,\mathcal{K}}{m\omega_{\rm c}} = n \quad , \tag{1.262}$$

[17] Note that the argument of J_n in eqn. 1.261 is λ and not $\varLambda$. This arises because we are computing the new action $\mathcal{J}$ in terms of the old variables (ϕ, J) and (ψ, K).

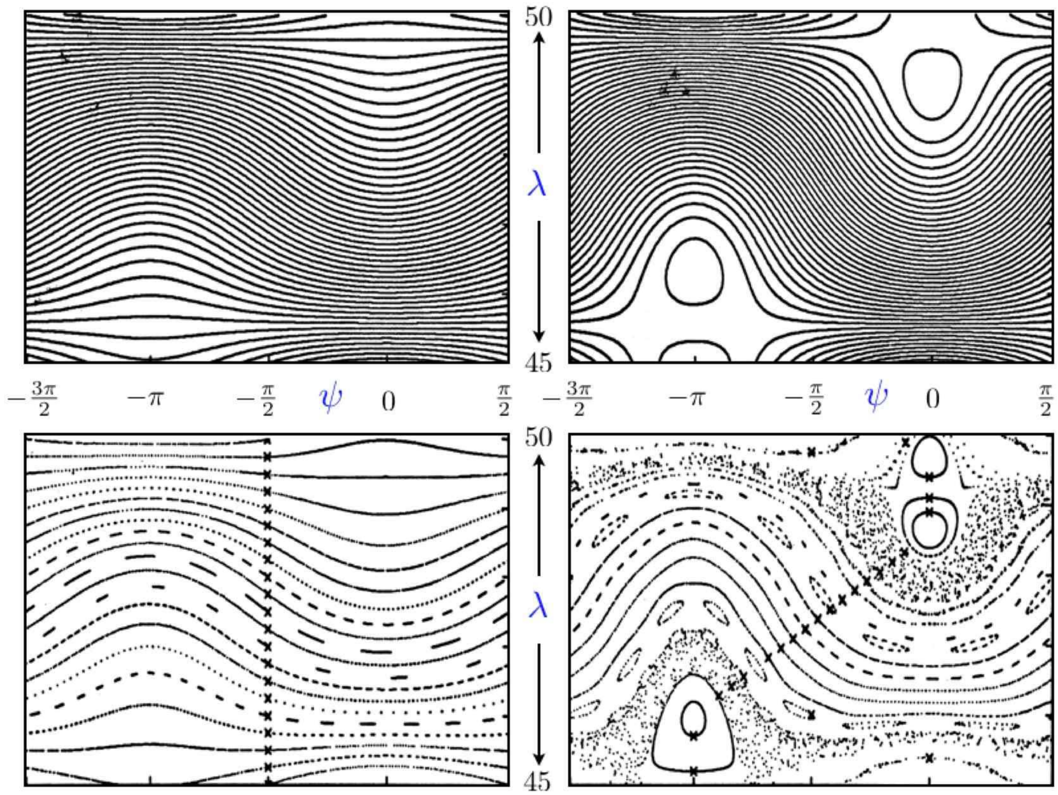

Figure 1.7: Plot of Λ *versus* ψ for $\phi = 0$ (Poincaré section) for $\omega = 30.11\,\omega_c$ Top panels are nonresonant invariant curves calculated to first order. Bottom panels are exact numerical dynamics, with x symbols marking the initial conditions. Left panels: weak amplitude (no trapping). Right panels: stronger amplitude (shows trapping). From Lichtenberg and Lieberman (1983).

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

$$\frac{\Lambda^2}{2\alpha} = \frac{\lambda^2}{2\alpha} - \epsilon \sum_n \frac{n\,J_n(\Lambda)\cos(\psi + n\phi)}{\omega/\omega_c - n} \quad , \tag{1.263}$$

where

$$\alpha = \frac{E_0}{B} \cdot \frac{ck_\perp}{\omega_c} \tag{1.264}$$

is a dimensionless measure of the strength of the perturbation, with $E_0 \equiv k_\perp \hat{V}_0$. In fig. 1.7 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.

1.9 Removal of Resonances in Perturbation Theory

We follow the treatment in chapter 3 of Lichtenberg and Lieberman.

1.9.1 The case of $n = \frac{3}{2}$ degrees of freedom

Consider the time-dependent Hamiltonian,

$$H(\phi, J, t) = H_0(J) + \epsilon\, V(\phi, J, t) \quad , \tag{1.265}$$

where $V(\phi, J, t) = V(\phi + 2\pi, J, t) = V(\phi, J, t + T)$ is periodic in time as well as in the angle variable ϕ. We may express the perturbation as a double Fourier sum,

$$V(\phi, J, t) = \sum_{k,\ell} \hat{V}_{k,\ell}(J)\, e^{ik\phi}\, e^{-i\ell\Omega t} \quad , \tag{1.266}$$

where $\Omega = 2\pi/T$. Hamilton's equations of motion are

$$\begin{aligned} \dot{J} &= -\frac{\partial H}{\partial \phi} = -i\epsilon \sum_{k,\ell} k\, \hat{V}_{k,l}(J)\, e^{ik\phi}\, e^{-i\ell\Omega t} \\ \dot{\phi} &= +\frac{\partial H}{\partial J} = \omega_0(J) + \epsilon \sum_{k,\ell} \frac{\partial \hat{V}_{k,l}(J)}{\partial J}\, e^{ik\phi}\, e^{-i\ell\Omega t} \quad , \end{aligned} \tag{1.267}$$

where $\omega_0(J) \equiv \partial H_0/\partial J$.[18] The resonance condition is obtained by inserting the zeroth order solution $\phi(t) = \omega_0(J)\, t + \beta$ into the perturbation terms. When $k\,\omega_0(J) = l\Omega$, the perturbation results in a secular forcing, leading to a linear time increase and a failure of the solution at sufficiently large values of t.

To resolve this crisis, we focus on a particular resonance, where $(k, \ell) = \pm(k_0, \ell_0)$. The resonance condition $k_0\,\omega(J) = \ell_0\Omega$ fixes the value of J. There may be several solutions, and we focus on a particular one, which we write as $J = J_0$. There is still an infinite set of possible (k, l) values, because if (k_0, ℓ_0) yields a solution for $J = J_0$, so does $(k, \ell) = (pk_0, p\ell_0)$ for $p \in \mathbb{Z}$. However, the amplitude of the Fourier components $\hat{V}_{pk_0,p\ell_0}$ is, in general, a rapidly decreasing function of $|p|$, provided $V(J, \phi, t)$ is smooth in ϕ and t. Furthermore, $p = 0$ always yields a solution. Therefore, we will assume k_0 and ℓ_0 are relatively prime and take $p = 0$ and $p = \pm 1$. This simplifies the system in eqn. 1.267 to

$$\begin{aligned} \dot{J} &= 2\epsilon\, k_0 \hat{V}_1(J)\, \sin(k_0\phi - \ell_0\Omega t + \delta) \\ \dot{\phi} &= \omega_0(J) + \epsilon\, \frac{\partial \hat{V}_0(J)}{\partial J} + 2\epsilon\, \frac{\partial \hat{V}_1(J)}{\partial J}\, \cos(k_0\phi - \ell_0\Omega t + \delta) \quad , \end{aligned} \tag{1.268}$$

where $\hat{V}_{0,0}(J) \equiv \hat{V}_0(J)$ and $\hat{V}_{k_0,\ell_0}(J) = V^*_{-k_0,-\ell_0}(J) \equiv \hat{V}_1(J)\, e^{i\delta}$, where $\hat{V}_0(J)$ and $\hat{V}_1(J)$ are both real. We then expand, writing

$$J = J_0 + \Delta J \quad , \quad \psi = k_0\phi - \ell_0\Omega t + \delta + \pi \quad , \tag{1.269}$$

[18] In this section we write $\partial H_0/\partial J = \omega_0(J)$ rather than $\nu_0(J)$ in order to obviate any confusion between the frequency ν_0 and the potential $\hat{V}_1$ and its various Fourier components.

resulting in the system

$$\begin{aligned}\frac{d\Delta J}{dt} &= -2\epsilon\, k_0 \hat{V}_1(J_0)\sin\psi \\ \frac{d\psi}{dt} &= k_0\, \omega_0'(J_0)\, \Delta J + \epsilon\, k_0 \hat{V}_0'(J_0) - 2\epsilon\, k_0 \hat{V}_1'(J_0)\cos\psi \quad ,\end{aligned} \tag{1.270}$$

which follow from the Hamiltonian

$$K(\Delta J,\psi) = \tfrac{1}{2} k_0\, \omega_0'(J_0)\,(\Delta J)^2 + \epsilon k_0 \hat{V}_0'(J_0)\,\Delta J - 2\epsilon k_0 \hat{V}_1(J_0+\Delta J)\cos\psi \quad , \tag{1.271}$$

with $d\psi/dt = \partial K/\partial(\Delta J)$ and $d(\Delta J)/dt = -\partial K/\partial\psi$. Concerning the last term, we can drop the ΔJ term in the argument of $\hat{V}_1$, leaving $\hat{V}_1(J_0)$, because it will yield a term of second order in smallness in the equation of motion for ψ. The remaining term in K linear in ΔJ can then be removed by a shift of $\Delta J \to \Delta J - \epsilon\, \hat{V}_0'(J_0)/\omega_0'(J_0)$. This is tantamount to shifting the value of J_0, which we could have done at the outset by absorbing the term $\epsilon \hat{V}_0(J)$ into $H_0(J)$, and defining $\omega(J) \equiv \omega_0(J) + \epsilon\, \partial\hat{V}_0/\partial J$. We are left with a simple pendulum, with

$$\ddot{\psi} + \gamma^2 \sin\psi = 0 \tag{1.272}$$

with $\gamma = \sqrt{2\epsilon\, k_0^2\, \omega'(J_0)\, \hat{V}_1(J_0)}$. In Fig. 1.8, we plot the level sets of the function

$$\check{K}(\Delta J,\psi) \equiv \tfrac{1}{2} k_0\, \omega_0'(J_0)\,(\Delta J)^2 + \epsilon k_0 \hat{V}_0'(J_0)\,\Delta J - 2\epsilon k_0 \hat{V}_1(J_0)\cos\psi \tag{1.273}$$

in the *rotating* (q,p) plane, *i.e.* in the $(\check{q},\check{p})$ plane, where

$$\check{q} \propto (J_0+\Delta J)^{1/2}\cos(k_0\,\check{\phi}) \qquad , \qquad \check{p} \propto (J_0+\Delta J)^{1/2}\sin(k_0\,\check{\phi}) \quad , \tag{1.274}$$

where $\check{\phi} = \phi - l_0\, \Omega t/k_0$.

What do we conclude? The original 1-torus (*i.e.* circle) with $J = J_0$ and $\phi(t) = \omega_0(J_0)\, t + \beta$ is destroyed. It and its neighboring tori are replaced, in the c ase $k_0 = 1$, by the separatrix in the left panel of fig. 1.8 and the neighboring librational and rotational phase curves. The structure for $k_0 = 6$ is shown in the right panel. The amplitude of the separatrix is $(\Delta J)_{\rm max} = \sqrt{8\epsilon\hat{V}_1(J_0)/\omega_0'(J_0)}$. In order for the approximations leading to this structure to be justified, we need $(\Delta J)_{\rm max} \ll J_0$ and $\Delta\omega \ll \omega_0$, where $\Delta\omega = \gamma$. These conditions may be written as

$$\epsilon \ll \alpha \ll \frac{1}{\epsilon} \quad , \tag{1.275}$$

where $\alpha = d\ln\omega_0/d\ln J\big|_{J_0} = J_0\, |\omega_0'|/\omega_0$.

1.9.2 $n = 2$ systems

Consider now the time-independent Hamiltonian $H = H_0(\boldsymbol{J}) + \epsilon H_1(\boldsymbol{\phi},\boldsymbol{J})$ with $n = 2$ degrees of freedom, *i.e.* $\boldsymbol{J} = (J_1, J_2)$ and $\boldsymbol{\phi} = (\phi_1,\phi_2)$. We Fourier expand

$$H_1(\boldsymbol{\phi},\boldsymbol{J}) = \sum_{\boldsymbol{\ell}} \hat{V}_{\boldsymbol{\ell}}(\boldsymbol{J})\, e^{i\boldsymbol{\ell}\cdot\boldsymbol{\phi}} \quad , \tag{1.276}$$

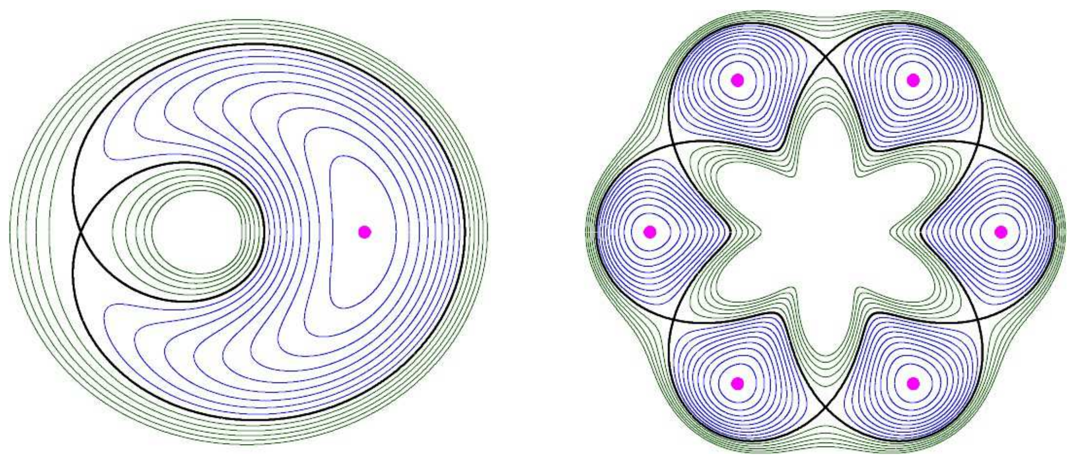

Figure 1.8: Librations, separatrices, and rotations for $k_0 = 1$ (left) and $k_0 = 6$ (right), plotted in the (q, p) phase plane. Elliptical fixed points are shown in magenta. Hyperbolic fixed points are located at the self-intersection of the separatrices (black curves).

with $\boldsymbol{\ell} = (\ell_1, \ell_2)$ and $\hat{V}_{-\boldsymbol{\ell}}(\boldsymbol{J}) = V^*_{\boldsymbol{\ell}}(\boldsymbol{J})$ since $\hat{V}_{\boldsymbol{\ell}}(\boldsymbol{J})$ are the Fourier components of a real function. A resonance exists between the frequencies $\omega_{1,2} = \partial H_0/\partial J_{1,2}$ if there exist nonzero integers r and s such that $r\omega_1 = s\omega_2$. We eliminate the resonance in two steps. First, we employ a canonical transformation $(\boldsymbol{\phi}, \boldsymbol{J}) \to (\boldsymbol{\varphi}, \boldsymbol{\mathcal{J}})$, generated by

$$F_2(\boldsymbol{\phi}, \boldsymbol{\mathcal{J}}) = (r\phi_1 - s\phi_2)\mathcal{J}_1 + \phi_2\mathcal{J}_2 \quad . \tag{1.277}$$

We then have

$$J_1 = \frac{\partial F_2}{\partial \phi_1} = r\mathcal{J}_1 \qquad \varphi_1 = \frac{\partial F_2}{\partial \mathcal{J}_1} = r\phi_1 - s\phi_2 \tag{1.278}$$

$$J_2 = \frac{\partial F_2}{\partial \phi_2} = \mathcal{J}_2 - s\mathcal{J}_1 \qquad \varphi_2 = \frac{\partial F_2}{\partial \mathcal{J}_2} = \phi_2 \quad . \tag{1.279}$$

This transforms us to a rotating frame in which $\dot{\varphi}_1 = r\dot{\phi}_1 - s\dot{\phi}_2$ is slowly varying, while $\dot{\varphi}_2 = \dot{\phi}_2 \approx \omega_2$. Note that we could have chosen $F_2 = \phi_1\mathcal{J}_1 + (r\phi_1 - s\phi_2)\mathcal{J}_2$, in which case we'd have obtained $\varphi_1 = \phi_1$ with an unperturbed natural frequency of ω_1 and $\varphi_2 = r\phi_1 - s\phi_2$ slowly varying, *i.e.* with an unperturbed natural frequency of zero. Which transformation are we to choose? The answer is that we want to end up averaging over the *slower* of $\omega_{1,2}$, so the generator in eqn. 1.277 is appropriate if $\omega_1 > \omega_2$. The reason has to do with what happens when there are higher order resonances to be removed – a state of affairs we shall discuss in the following section.

At this stage, our transformed Hamiltonian is

$$\begin{aligned} \widetilde{H}(\boldsymbol{\varphi}, \boldsymbol{\mathcal{J}}) &= H_0\big(\boldsymbol{J}(\boldsymbol{\mathcal{J}})\big) + \epsilon H_1\big(\boldsymbol{\phi}(\boldsymbol{\varphi}), \boldsymbol{J}(\boldsymbol{\mathcal{J}})\big) \\ &\equiv \widetilde{H}_0(\boldsymbol{\mathcal{J}}) + \epsilon \sum_{\boldsymbol{\ell}} \hat{\widetilde{V}}_{\boldsymbol{\ell}}(\boldsymbol{\mathcal{J}}) \, \exp\left[\frac{i\ell_1}{r}\,\varphi_1 + i\left(\frac{s\,\ell_1}{r} + \ell_2\right)\varphi_2\right] \quad , \end{aligned} \tag{1.280}$$

where $\widetilde{H}(\boldsymbol{\mathcal{J}}) \equiv H_0\big(\boldsymbol{J}(\boldsymbol{\mathcal{J}})\big)$ and $\hat{\tilde{V}}_{\boldsymbol{\ell}}(\boldsymbol{\mathcal{J}}) \equiv \hat{\tilde{V}}_{\boldsymbol{\ell}}\big(\boldsymbol{J}(\boldsymbol{\mathcal{J}})\big)$. Note that $\phi_1 = \varphi_1/r + s\,\varphi_2/r$. We now average over the angle φ_2 , which requires $s\ell_1 + r\ell_2 = 0$. Thus, $\ell_1 = pr$ and $\ell_2 = -ps$ for some $p \in \mathbb{Z}$, and

$$\langle H_1 \rangle = \sum_p \hat{\tilde{V}}_{pr,-ps}(\boldsymbol{\mathcal{J}})\, e^{-ip\varphi_1} \quad . \tag{1.281}$$

The averaging is valid close to the resonance, where $|\dot{\varphi}_2| \gg |\dot{\varphi}_1|$. We are now left with the Hamiltonian

$$\mathcal{H}(\varphi_1, \boldsymbol{\mathcal{J}}) = \widetilde{H}_0(\boldsymbol{\mathcal{J}}) + \epsilon \sum_p \hat{\tilde{V}}_{pr,-ps}(\boldsymbol{\mathcal{J}})\, e^{-ip\varphi_1} \quad . \tag{1.282}$$

Here, $\mathcal{J}_2$ is to be regarded as a parameter which itself has no dynamics: $\dot{\mathcal{J}}_2 = 0$. Note $\mathcal{J}_2 = (s/r)J_1 + J_2$ is the new invariant.

At this point, φ_2 has been averaged out, $\mathcal{J}_2$ is a constant, and only the $(\varphi_1, \mathcal{J}_1)$ variables are dynamical. A stationary point for these dynamics, satisfying $\partial\mathcal{H}/\partial\mathcal{J}_1 = \partial\mathcal{H}/\partial\varphi_1 = 0$ corresponds to a periodic solution to the original perturbed Hamiltonian, since we are now in a rotating frame. Since the Fourier amplitudes $\hat{\tilde{V}}_{-pr,ps}(\boldsymbol{\mathcal{J}})$ generally decrease rapidly with increasing $|p|$, we make the approximation of restricting to $p = 0$ and $p = \pm 1$. Thus,

$$\mathcal{H}(\varphi_1, \boldsymbol{\mathcal{J}}) \approx \widetilde{H}_0(\boldsymbol{\mathcal{J}}) + \epsilon\, \hat{\tilde{V}}_{0,0}(\boldsymbol{\mathcal{J}}) + 2\epsilon\, \hat{\tilde{V}}_{r,-s}(\boldsymbol{\mathcal{J}}) \cos\varphi_1 \quad , \tag{1.283}$$

where we have absorbed any phase in the Fourier amplitude $\hat{\tilde{V}}_{r,-s}(\boldsymbol{\mathcal{J}})$ into a shift of φ_1 , and subsequently take $\hat{\tilde{V}}_{r,-s}(\boldsymbol{\mathcal{J}})$ to be real. The fixed points $\big(\varphi_1^{(0)}, \mathcal{J}_1^{(0)}\big)$ of the $(\varphi_1, \mathcal{J}_1)$ dynamics satisfy

$$\begin{aligned} 0 &= \frac{\partial \widetilde{H}_0}{\partial \mathcal{J}_1} + \epsilon\, \frac{\partial \hat{\tilde{V}}_{0,0}}{\partial \mathcal{J}_1} + 2\epsilon\, \frac{\partial \hat{\tilde{V}}_{r,-s}}{\partial \mathcal{J}_1} \cos\varphi_1 \\ 0 &= 2\epsilon\, \hat{\tilde{V}}_{r,-s} \sin\varphi_1 \quad . \end{aligned} \tag{1.284}$$

Thus, $\varphi_1 = 0$ or π at the fixed points. Note that

$$\frac{\partial \widetilde{H}_0}{\partial \mathcal{J}_1} = \frac{\partial H_0}{\partial J_1}\,\frac{\partial J_1}{\partial \mathcal{J}_1} + \frac{\partial H_0}{\partial J_2}\,\frac{\partial J_2}{\partial \mathcal{J}_1} = r\omega_1 - s\omega_2 = 0 \quad , \tag{1.285}$$

and therefore fixed points occur for solutions $\mathcal{J}_1^{(0)}$ to

$$\frac{\partial \hat{\tilde{V}}_{0,0}}{\partial \mathcal{J}_1} \pm 2\, \frac{\partial \hat{\tilde{V}}_{r,-s}}{\partial \mathcal{J}_1} = 0 \quad , \tag{1.286}$$

where the upper sign corresponds to $\varphi_1^{(0)} = 0$ and the lower sign to $\varphi_1^{(0)} = \pi$. We now consider two cases.

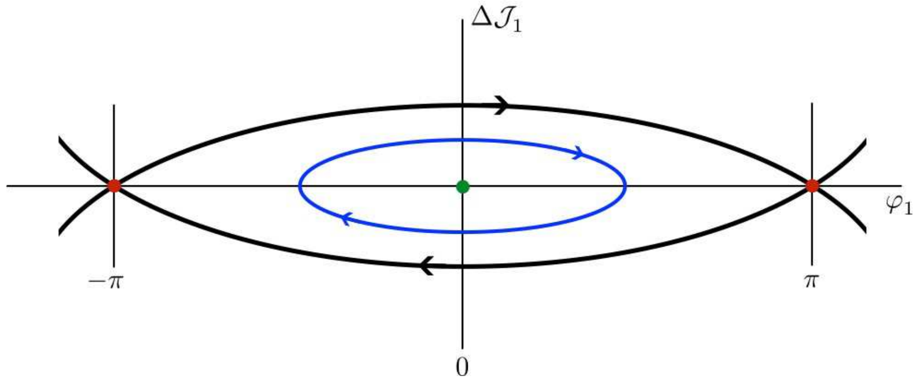

Figure 1.9: Motion in the vicinity of a resonance, showing elliptical fixed point in green, hyperbolic fixed point in red, and separatrix in black.

(i) accidental degeneracy

In the case of accidental degeneracy, the resonance condition $r\omega_1 = s\omega_2$ is satisfied only for particular values of (J_1, J_2), *i.e.* on a set $J_2 = J_2(J_1)$. This corresponds to the case where $H_0(J_1, J_2)$ is a generic function of its two arguments. According to eqn. 1.283, excursions of $\mathcal{J}_1$ relative to its value $\mathcal{J}_1^{(0)}$ at the fixed points are on the order of $\epsilon\hat{\tilde{V}}_{r,-s}$, while excursions of φ_1 are $\mathcal{O}(1)$. We may then expand

$$\widetilde{H}_0(\mathcal{J}_1, \mathcal{J}_2) = \widetilde{H}_0\big(\mathcal{J}_1^{(0)}, \mathcal{J}_2\big) + \frac{\partial \widetilde{H}_0}{\partial \mathcal{J}_1}\,\Delta\mathcal{J}_1 + \frac{1}{2}\frac{\partial^2 \widetilde{H}_0}{\partial \mathcal{J}_1^2}\,(\Delta\mathcal{J}_1)^2 + \ldots \quad , \tag{1.287}$$

where the derivatives are evaluated at $\mathcal{J}_1 = \mathcal{J}_1^{(0)}$. Thus, we arrive at what is often called the *standard Hamiltonian*,

$$\mathcal{H}(\varphi_1, \Delta\mathcal{J}_1) = \tfrac{1}{2}G\,(\Delta\mathcal{J}_1)^2 - F\cos\varphi_1 \quad , \tag{1.288}$$

with

$$G(\mathcal{J}_2) = \frac{\partial^2 \widetilde{H}_0}{\partial \mathcal{J}_1^2}\bigg|_{\mathcal{J}_1^{(0)}} \qquad , \qquad F(\mathcal{J}_2) = -2\epsilon\,\hat{\tilde{V}}_{r,-s}\big(\mathcal{J}_1^{(0)}, \mathcal{J}_2\big) \quad . \tag{1.289}$$

Thus, the motion in the vicinity of every resonance is like that of a pendulum, meaning libration, separatrix, and rotation in the phase plane. F is the amplitude of the first Fourier mode of the perturbation (*i.e.* $|p| = 1$), and G the 'nonlinearity parameter'. For $FG > 0$ the elliptic fixed point (EFP) is at $\varphi_1 = 0$ and the hyperbolic fixed point (HFP) at $\varphi_1 = \pi$. For $FG < 0$, the locations are switched. The frequency of libration about the EFP is given by $\nu_1 = \sqrt{FG} = \mathcal{O}\Big(\sqrt{\epsilon\hat{\tilde{V}}_{r,-s}}\Big)$. The frequency decreases to zero as the separatrix is approached. The maximum excursion along the separatrix is $(\Delta\mathcal{J}_1)_{\rm max} = 2\sqrt{F/G}$ which is also $\mathcal{O}\Big(\sqrt{\epsilon\hat{\tilde{V}}_{r,-s}}\Big)$. The ratio of semiminor to semimajor axis lengths for motion in the vicinity of the EFP is

$$\frac{(\Delta\mathcal{J}_1)_{\rm max}}{(\Delta\varphi_1)_{\rm max}} = \sqrt{\frac{F}{G}} = \mathcal{O}\big(\epsilon^{1/2}\big) \quad . \tag{1.290}$$

(ii) intrinsic degeneracy

In this case, $H_0(J_1, J_2)$ is a function of only the combination $sJ_1 + rJ_2 = r\mathcal{J}_2$, so

$$\mathcal{H}(\varphi_1, \boldsymbol{\mathcal{J}}) = \widetilde{H}_0(\mathcal{J}_2) + \epsilon\, \hat{\tilde{V}}_{0,0}(\boldsymbol{\mathcal{J}}) + 2\epsilon\, \hat{\tilde{V}}_{r,-s}(\boldsymbol{\mathcal{J}}) \cos\varphi_1 \quad . \tag{1.291}$$

In this case excursions of $\mathcal{J}_1$ and φ_1 are both $\mathcal{O}(\epsilon\, \hat{\tilde{V}}_{\bullet,\bullet})$, and we are not licensed to expand in $\Delta\mathcal{J}_1$. However, in the vicinity of an EFP, we may expand, both in $\Delta\mathcal{J}_1$ and $\Delta\varphi_1$, resulting in

$$\mathcal{H} = \tfrac{1}{2} G\,(\Delta\mathcal{J}_1)^2 + \tfrac{1}{2} F\,(\Delta\varphi_1)^2 \quad , \tag{1.292}$$

where

$$G(\mathcal{J}_2) = \left[\frac{\partial^2 \widetilde{H}_0}{\partial \mathcal{J}_1^2} + \epsilon\, \frac{\partial^2 \hat{\tilde{V}}_{0,0}}{\partial \mathcal{J}_1^2} + 2\epsilon\, \frac{\partial^2 \hat{\tilde{V}}_{r,-s}}{\partial \mathcal{J}_1^2}\right]_{(\mathcal{J}_1^{(0)}, \mathcal{J}_2)} \quad , \qquad F(\mathcal{J}_2) = -2\epsilon\, \hat{\tilde{V}}_{r,-s}\big(\mathcal{J}_1^{(0)}, \mathcal{J}_2\big) \quad . \tag{1.293}$$

For the case of intrinsic degeneracy, the first term in brackets on the RHS of the equation for $G(\mathcal{J}_2)$ vanishes, since $\widetilde{H}_0$ is a function only of $\mathcal{J}_2$. Hence F and G are both $\mathcal{O}(\epsilon\, \hat{\tilde{V}}_{\bullet,\bullet})$, hence $\nu_1 = \sqrt{FG} = \mathcal{O}(\epsilon)$ and the ratio of semiminor to semimajor axis lengths of the motion is

$$\frac{(\Delta\mathcal{J}_1)_{\text{max}}}{(\Delta\varphi_1)_{\text{max}}} = \sqrt{\frac{F}{G}} = \mathcal{O}(1) \quad . \tag{1.294}$$

1.9.3 Secondary resonances

By averaging over the φ_2 motion and expanding about the EFP, we obtained the Hamiltonian in Eqns. 1.292 and 1.293. In so doing, we dropped all terms on the RHS of eqn. 1.280 with $s\ell_1 + r\ell_2 \neq 0$. We now restore those terms, and continue to expand about the EFP. The first step is to transform the harmonic oscillator Hamiltonian in eqn. 1.292 to action-angle variables; this was already done in §1.8.3. The canonical transformation from $(\Delta\varphi_1, \Delta\mathcal{J}_1)$ to (χ_1, I_1) is given by

$$\Delta\mathcal{J}_1 = (2RI_1)^{1/2} \cos\chi_1 \quad , \qquad \Delta\varphi_1 = (2R^{-1}I_1)^{1/2} \sin\chi_1 \quad , \tag{1.295}$$

with $R = (F/G)^{1/2}$. We will also define $I_2 \equiv \mathcal{J}_2$ and $\chi_2 \equiv \varphi_2$. Then we may write

$$\mathcal{H}(\varphi_1, \boldsymbol{\mathcal{J}}) \longrightarrow \widetilde{\mathcal{H}}_0(\boldsymbol{I}) = \widetilde{H}_0\big(\mathcal{J}_1^{(0)}, I_2\big) + \nu_1(I_2)\, I_1 - \tfrac{1}{16}\, G(I_2)\, I_1^2 + \ldots \quad , \tag{1.296}$$

where the last term on the RHS before the ellipses is from nonlinear terms in $\Delta\varphi_1$. The missing terms we seek are

$$\widetilde{H}_1' = \sum_{\boldsymbol{\ell}} \hat{\tilde{V}}_{\boldsymbol{\ell}}\big(\mathcal{J}_1^{(0)}, I_2\big)\, \exp\!\big[ir^{-1}\ell_1 (2R^{-1}I_1)^{1/2} \sin\chi_1\big]\, \exp\!\big[i(r^{-1}s\,\ell_1 + \ell_2)\,\chi_2\big] \quad . \tag{1.297}$$

Note that we set $\mathcal{J}_1 = \mathcal{J}_1^{(0)}$ in the argument of $\hat{\tilde{V}}_{\boldsymbol{\ell}}(\boldsymbol{\mathcal{J}})$, because $\Delta\mathcal{J}_1$ is of order $\epsilon^{1/2}$. Next we invoke the Bessel function identity,

$$e^{iu\sin\chi} = \sum_{-\infty}^{\infty} J_n(u)\, e^{in\chi} \quad , \tag{1.298}$$

so we write

$$\widetilde{H}_1' \longrightarrow \mathcal{H}_1(\boldsymbol{\chi},\boldsymbol{I}) = \sum_{\boldsymbol{\ell}} \sum_n W_{\boldsymbol{\ell},n}(\boldsymbol{I})\, e^{in\chi_1}\, e^{i(r^{-1}s\,\ell_1+\ell_2)\,\chi_2} \quad , \tag{1.299}$$

where

$$W_{\boldsymbol{\ell},n}(\boldsymbol{I}) = \hat{\tilde{V}}_{\boldsymbol{\ell}}\big(\mathcal{J}_1^{(0)}, I_2\big)\, J_n\!\left(\frac{\ell_1}{r}\sqrt{\frac{2I_1}{R}}\right) \quad . \tag{1.300}$$

We now write

$$\mathcal{H}(\boldsymbol{\chi},\boldsymbol{I}) = \mathcal{H}_0(\boldsymbol{I}) + \tilde{\epsilon}\,\mathcal{H}_1(\boldsymbol{\chi},\boldsymbol{I}) \quad . \tag{1.301}$$

Here, while $\tilde{\epsilon} = \epsilon$ it is convenient to use a new symbol since ϵ itself appears within $\mathcal{H}_0$.

We now see that a secondary resonance will occur if $r'\nu_1 = s'\nu_2$, with $\nu_j(\boldsymbol{I}) = \partial\mathcal{H}_0/\partial I_j$ and $r', s' \in \mathbb{Z}$. But note that $\nu_1 = \mathcal{O}(\epsilon^{1/2})$ while $\nu_2 = \mathcal{O}(1)$ in the case of an accidental primary resonance. As before, we may eliminate this new resonance by transforming to a moving frame in which the resonance shifts to zero frequency to zeroth order and then averaging over the remaining motion. That is, we canonically transform $(\boldsymbol{\chi},\boldsymbol{I}) \to (\boldsymbol{\psi},\boldsymbol{\mathcal{I}})$ via a type-II generator $F_2' = (r'\chi_1 - s'\chi_2)\,\mathcal{I}_1 + \chi_2\,\mathcal{I}_2$, yielding

$$I_1 = \frac{\partial F_2'}{\partial \chi_1} = r'\mathcal{I}_1 \qquad\qquad \psi_1 = \frac{\partial F_2'}{\partial \mathcal{I}_1} = r'\chi_1 - s'\chi_2 \tag{1.302}$$

$$I_2 = \frac{\partial F_2'}{\partial \chi_2} = \mathcal{I}_2 - s'\mathcal{I}_1 \qquad\qquad \psi_2 = \frac{\partial F_2'}{\partial \mathcal{I}_2} = \chi_2 \quad . \tag{1.303}$$

The phase angle in eqn. 1.299 is then

$$n\chi_1 + \left(\frac{s}{r}\,\ell_1 + \ell_2\right)\chi_2 = \frac{n}{r'}\,\psi_1 + \left(\frac{ns'}{r'} + \frac{s}{r}\,\ell_1 + \ell_2\right)\psi_2 \quad . \tag{1.304}$$

Averaging over $\psi_2(t)$ then requires $nrs' + sr'\ell_1 + rr'\ell_2 = 0$, which is satisfied when

$$n = jr' \quad , \quad \ell_1 = kr \quad , \quad \ell_2 = -js' - ks \tag{1.305}$$

for some $j, k \in \mathbb{Z}$. The result of the averaging is

$$\big\langle \mathcal{H} \big\rangle_{\psi_2} = \mathcal{H}_0\big(\boldsymbol{I}(\boldsymbol{\mathcal{I}})\big) + \tilde{\epsilon} \sum_j \varGamma_{jr',-js'}(\boldsymbol{\mathcal{I}})\, e^{-ij\psi_1} \tag{1.306}$$

where

$$\varGamma_{jr',-js'}(\boldsymbol{\mathcal{I}}) = W_{kr,-js'-ks,jr'}\big(\boldsymbol{I}(\boldsymbol{\mathcal{I}})\big) = \hat{\tilde{V}}_{kr,-js'-ks}\big(\mathcal{J}_1^{(0)}, I_2\big)\, J_{jr'}\!\left(k\sqrt{\frac{2I_1(\boldsymbol{\mathcal{I}})}{R}}\right) \quad . \tag{1.307}$$

Since $\big\langle \mathcal{H} \big\rangle_{\psi_2}$ is independent of ψ_2, the corresponding action $\mathcal{I}_2 = (s'/r')\,I_1 + I_2$ is the adiabatic invariant for the new oscillation.

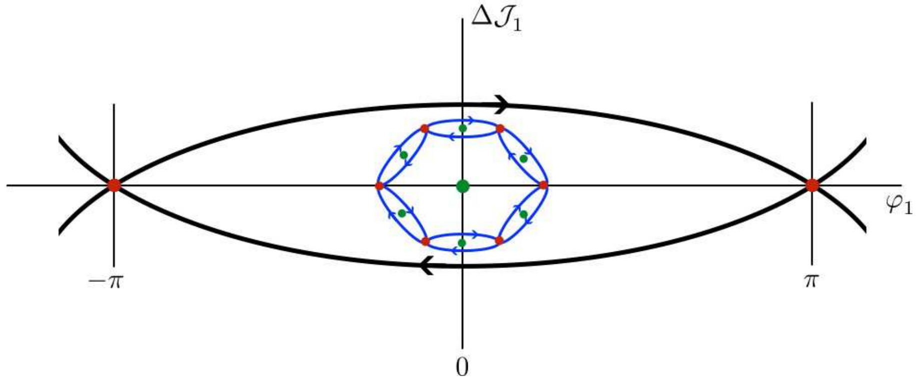

Figure 1.10: Motion in the vicinity of a secondary resonance with $r' = 6$ and $s' = 1$. Elliptical fixed points are in green, hyperbolic fixed points in red, and separatrices in black and blue.

Strength of island resonances

To assess the strength of the secondary resonances, we consider $r = s = j = k = s' = 1$, in which case $r' = \nu_2/\nu_1 = \mathcal{O}(\epsilon^{-1/2})$ is parametrically large. The resulting structure in the phase plane is depicted in fig. 1.10 for $r' = 6$. The amplitude of the $\mathcal{M}_1$ oscillations is proportional to

$$J_r\big((2I_1(\boldsymbol{\mathcal{I}})/2R)^{1/2}\big) \sim \frac{\big(2I_1(\boldsymbol{\mathcal{I}})/2R\big)^{r'/2}}{r'!} = \mathcal{O}\bigg(\frac{1}{(\epsilon^{-1/2})!}\bigg) \quad . \tag{1.308}$$

The frequency of the island oscillations is of the same order of magnitude. Successive higher order resonances result in an increasingly tiny island chain amplitude.

1.10 Whither Integrability?

We are left with the following question: what happens when we perturb an integrable Hamiltonian, $H(\boldsymbol{\phi}, \boldsymbol{J}) = H_0(\boldsymbol{J}) + \epsilon\, H_1(\boldsymbol{\phi}, \boldsymbol{J})$? Two extreme conjectures, and their refutations:

(i) $H(\boldsymbol{\phi}, \boldsymbol{J})$ is always integrable, even though we may not always be able to obtain the corresponding action-angle variables. Tori are deformed but not destroyed. If this were the case, there would be n conserved quantities, *i.e.* the first integrals of motion I_j. This would violate the fundamental tenets of equilibrium statistical physics, as the canonical Gibbs distribution $\varrho = \exp(-\beta H)/Z$ would be replaced with the *pseudo-Gibbs* distribution, $\varrho = \exp(-\lambda_j I_j)/Z$, where $\{\lambda_j\}$ are a set of Lagrange multipliers[19].

(ii) Integrability is destroyed for any $\epsilon > 0$, in which case $E = H(\boldsymbol{\phi}, \boldsymbol{J})$ is the only conserved quantity[20]. If this were the case, the solar system would be unstable, and we wouldn't be here to study Hamiltonian mechanics.

[19]The corresponding microcanonical distribution would be $\prod_{j=1}^{n} \delta\big(I_j - \langle I_j \rangle\big)$, as opposed to $\delta(H - E)$.

[20]Without loss of generality, we may assume $\epsilon \geq 0$.

So the truth lies somewhere in between, and is the focus of the celebrated KAM theorem[21]. We have already encountered the problem of resonances, which arise for tori which satisfy $\boldsymbol{\ell}\cdot\boldsymbol{\omega}_0(\boldsymbol{J})=0$ for some integers $\boldsymbol{\ell}=\{\ell_1,\ldots,\ell_n\}$. Such tori, which are dense in the phase space $\mathcal{M}$ yet still of Lebesgue measure zero, are destroyed by arbitrarily small perturbations, as we have seen. This observation dates back to Poincaré. For a given torus with an $(n-1)$-dimensional family of periodic orbits, $J_n=J_n(J_1,\ldots,J_{n-1})$, it is generally the case that only a finite number of periodic orbits survive the perturbation. Since, in a nondegenerate system, the set of resonant tori is dense, it seems like the situation is hopeless and that arbitrarily small ϵ will induce ergodicity on each energy surface. Until the early 1950s, it was generally believed that this was the case, and the stability of the solar system was regarded as a deep mystery.

Enter Andrey Nikolaevich Kolmogorov, who in 1954 turned conventional wisdom on its head, showing that, in fact, the *majority* of tori survive. Specifically, Kolmogorov proved that *strongly nonresonant* tori survive small perturbations. A strongly nonresonant torus is one for which there exist constants $\alpha>0$ and $\tau>0$ such that $\big|\boldsymbol{\ell}\cdot\boldsymbol{\omega}_0(\boldsymbol{J})\big|\geq\alpha\,|\boldsymbol{\ell}|^{-\tau}$, where $|\boldsymbol{\ell}|\equiv|\ell_1|+\ldots+|\ell_n|$. From a measure theoretic point of view, almost all tori are strongly nonresonant for any $\tau>n-1$, but in order to survive the perturbation, it is necessary that $\epsilon\ll\alpha^2$. For these tori, perturbation theory converges, although not quite in the naïve form we have derived, *i.e.* from the generator $S(\boldsymbol{\phi},\boldsymbol{\mathcal{J}})=S_0+\epsilon S_1+\epsilon^2 S_2+\ldots$, but rather using the 'superconvergent' method pioneered by Kolmogorov.

Since the arithmetic of the strongly nonresonant tori is a bit unusual, let's first convince ourselves that such tori actually exist[22]. Let Δ^τ_α denote the set of all $\boldsymbol{\omega}\in\mathbb{R}^n$ satisfying, for fixed α and τ, the infinitely many conditions $\boldsymbol{\ell}\cdot\boldsymbol{\omega}\geq\alpha|\boldsymbol{\ell}|^{-\tau}$, for all nonzero $\boldsymbol{\ell}\in\mathbb{Z}^n$. Clearly Δ^τ_α is the complement of the open and dense set $R^\tau_\alpha=\bigcup_{0\neq\boldsymbol{\ell}\in\mathbb{Z}^n}R^\tau_{\alpha,\boldsymbol{\ell}}$, where

$$R^\tau_{\alpha,\boldsymbol{\ell}}=\Big\{\boldsymbol{\omega}\in\mathbb{R}^n\ :\ |\boldsymbol{\ell}\cdot\boldsymbol{\omega}|<\alpha\,|\boldsymbol{\ell}|^{-\tau}\Big\}\quad. \tag{1.309}$$

For any bounded region $\Omega\in\mathbb{R}^n$, we can estimate the Lebesgue measure of the set $R^\tau_\alpha\cap\Omega$ from the calculation

$$\mu\big(R^\tau_\alpha\cap\Omega\big)\leq\sum_{\boldsymbol{\ell}\neq0}\mu\big(R^\tau_{\alpha,\boldsymbol{\ell}}\cap\Omega\big)=\mathcal{O}(\alpha)\quad, \tag{1.310}$$

The sum converges provided $\tau>n-1$ since $\mu\big(R^\tau_{\alpha,\boldsymbol{\ell}}\cap\Omega\big)=\mathcal{O}\big(\alpha/|\boldsymbol{\ell}|^{\tau+1}\big)$. Taking the intersection over *all* $\alpha>0$, we conclude $R^\tau=\bigcap_{\alpha>0}R^\tau_\alpha$ is a set of measure zero, and therefore its complement, $\Delta^\tau=\bigcup_{\alpha>0}\Delta^\tau_\alpha$, is a set of full measure in $\mathbb{R}^n$. This means that almost every $\boldsymbol{\omega}\in\mathbb{R}^n$ belongs to the set Δ^τ, which is the set of all $\boldsymbol{\omega}$ satisfying the *Diophantine condition* $|\boldsymbol{\ell}\cdot\boldsymbol{\omega}|\geq\alpha\,|\boldsymbol{\ell}|^{-\tau}$ for *some* value of α, again provided $\tau>n-1$.

We say that a torus *survives* the perturbation if for $\epsilon>0$ there exists a deformed torus in phase space homotopic to that for $\epsilon=0$, and for which the frequencies satisfy $\boldsymbol{\omega}_\epsilon=f(\epsilon)\,\boldsymbol{\omega}_0$, with $\lim_{\epsilon\to0}f(\epsilon)=1$. Note this says $\omega_j/\omega_k=\omega_{0,j}/\omega_{0,k}$. Only tori with frequencies in Δ^τ_α with $\alpha\gg\sqrt{\epsilon}$ survive. The KAM theorem says that the measure of the space of surviving tori approaches unity as $\epsilon\to0$.

[21] KAM = Kolmogorov-Arnol'd-Moser, who developed the theory in a series of papers during the 1950s and 1960s. For a "friendly introduction to the content, history, and significance" of KAM, I highly recommend H. Scott Dumas, *The KAM Story* (World Scientific, 2014).

[22] See J. Pöschel, *A Lesson on the Classical KAM Theorem*, *Proc. Symp. Pure Math.* **69**, 707 (2001), in §1.d.

1.11 Appendix : Examples

1.11.1 Hamilton-Jacobi theory for point charge plus electric field

Consider a potential of the form

$$U(r) = \frac{k}{r} - Fz \quad , \tag{1.311}$$

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):

$$x = \sqrt{\xi\eta}\,\cos\varphi \quad , \quad y = \sqrt{\xi\eta}\,\sin\varphi \quad , \quad z = \tfrac{1}{2}(\xi - \eta) \quad . \tag{1.312}$$

Note that

$$\begin{aligned} \rho &\equiv \sqrt{x^2 + y^2} = \sqrt{\xi\eta} \\ r &= \sqrt{\rho^2 + z^2} = \tfrac{1}{2}(\xi + \eta) \quad . \end{aligned} \tag{1.313}$$

The kinetic energy is

$$\begin{aligned} T &= \tfrac{1}{2}m\big(\dot\rho^2 + \rho^2\,\dot\varphi^2 + \dot z^2\big) \\ &= \tfrac{1}{8}m\,(\xi + \eta)\bigg(\frac{\dot\xi^2}{\xi} + \frac{\dot\eta^2}{\eta}\bigg) + \tfrac{1}{2}m\,\xi\eta\,\dot\varphi^2 \quad , \end{aligned} \tag{1.314}$$

and hence the Lagrangian is

$$L = \tfrac{1}{8}m\,(\xi + \eta)\bigg(\frac{\dot\xi^2}{\xi} + \frac{\dot\eta^2}{\eta}\bigg) + \tfrac{1}{2}m\,\xi\eta\,\dot\varphi^2 - \frac{2k}{\xi + \eta} + \tfrac{1}{2}F\,(\xi - \eta) \quad . \tag{1.315}$$

Thus, the conjugate momenta are

$$\begin{aligned} p_\xi &= \frac{\partial L}{\partial\dot\xi} = \tfrac{1}{4}m\,(\xi + \eta)\,\frac{\dot\xi}{\xi} \\ p_\eta &= \frac{\partial L}{\partial\dot\eta} = \tfrac{1}{4}m\,(\xi + \eta)\,\frac{\dot\eta}{\eta} \\ p_\varphi &= \frac{\partial L}{\partial\dot\varphi} = m\,\xi\eta\,\dot\varphi \quad , \end{aligned} \tag{1.316}$$

and the Hamiltonian is

$$\begin{aligned} H &= p_\xi\,\dot\xi + p_\eta\,\dot\eta + p_\varphi\,\dot\varphi \\ &= \frac{2}{m}\bigg(\frac{\xi\,p_\xi^2 + \eta\,p_\eta^2}{\xi + \eta}\bigg) + \frac{p_\varphi^2}{2m\xi\eta} + \frac{2k}{\xi + \eta} - \tfrac{1}{2}F\,(\xi - \eta) \quad . \end{aligned} \tag{1.317}$$

Notice that $\partial H/\partial t = 0$, which means $dH/dt = 0$, *i.e.* $H = E \equiv \varLambda_1$ 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

$$\begin{aligned} S(q,\Lambda) &= W(q,\Lambda) - Et \\ &= W_\xi(\xi,\Lambda) + W_\eta(\eta,\Lambda) + W_\varphi(\varphi,\Lambda) - Et \quad . \end{aligned} \tag{1.318}$$

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

$$W_\varphi(\varphi,\Lambda) = P_\varphi\,\varphi \quad , \tag{1.319}$$

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

$$\xi\bigg(\frac{dW_\xi}{d\xi}\bigg)^2 + \frac{P_\varphi^2}{4\xi} + mk - \tfrac{1}{4}F\xi^2 - \tfrac{1}{2}mE\xi = -\eta\bigg(\frac{dW_\eta}{d\eta}\bigg)^2 - \frac{P_\varphi^2}{4\eta} - \tfrac{1}{4}F\eta^2 + \tfrac{1}{2}mE\eta \equiv \Upsilon \quad , \tag{1.320}$$

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

$$\begin{aligned} S\big(\overbrace{\xi,\eta,\varphi}^{q};\underbrace{E,P_\varphi,\Upsilon}_{\Lambda}\big) &= \int^\xi\! d\xi' \sqrt{\tfrac{1}{2}mE + \frac{\Upsilon - mk}{\xi'} + \tfrac{1}{4}mF\xi' - \frac{P_\varphi^2}{4\xi'^2}} \\ &\qquad + \int^\eta\! d\eta' \sqrt{\tfrac{1}{2}mE - \frac{\Upsilon}{\eta'} - \tfrac{1}{4}mF\eta' - \frac{P_\varphi^2}{4\eta'^2}} + P_\varphi\,\varphi - Et \quad . \end{aligned} \tag{1.321}$$

1.11.2 Hamilton-Jacobi theory for charged particle in a magnetic field

The Hamiltonian is

$$H = \frac{1}{2m}\Big(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\Big)^2 \quad . \tag{1.322}$$

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

$$S(x,y,P_1,P_2) = W_x(x,P_1,P_2) + W_y(y,P_1,P_2) - P_1\,t \quad . \tag{1.323}$$

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

The Hamilton-Jacobi equation is then

$$\bigg(\frac{\partial W_x}{\partial x}\bigg)^2 + \bigg(\frac{\partial W_y}{\partial y} - \frac{eBx}{c}\bigg)^2 = 2mP_1 \quad . \tag{1.324}$$

We solve by writing

$$W_y = P_2\,y \qquad \Rightarrow \qquad \bigg(\frac{dW_x}{dx}\bigg)^2 + \bigg(P_2 - \frac{eBx}{c}\bigg)^2 = 2mP_1 \quad . \tag{1.325}$$

This equation suggests the substitution

$$x = \frac{cP_2}{eB} + \frac{c}{eB}\sqrt{2mP_1}\,\sin\theta \quad . \tag{1.326}$$

in which case

$$\frac{\partial x}{\partial\theta} = \frac{c}{eB}\sqrt{2mP_1}\,\cos\theta \tag{1.327}$$

and

$$\frac{\partial W_x}{\partial x} = \frac{\partial W_x}{\partial\theta}\cdot\frac{\partial\theta}{\partial x} = \frac{eB}{c\sqrt{2mP_1}}\,\frac{1}{\cos\theta}\,\frac{\partial W_x}{\partial\theta} \quad . \tag{1.328}$$

Substitution into eqn. 1.325, we have $\partial W_x/\partial\theta = (2mcP_1/eB)\cos^2\!\theta$ which integrates to

$$W_x = \frac{mcP_1}{eB}\,\theta + \frac{mcP_1}{2eB}\,\sin(2\theta) \quad . \tag{1.329}$$

We then have

$$p_x = \frac{\partial W_x}{\partial x} = \frac{\partial W_x}{\partial\theta}\bigg/\frac{\partial x}{\partial\theta} = \sqrt{2mP_1}\,\cos\theta \tag{1.330}$$

and $p_y = \partial W_y/\partial y = P_2$. The type-II generator we seek is then

$$S(q,P,t) = \frac{mcP_1}{eB}\,\theta + \frac{mcP_1}{2eB}\,\sin(2\theta) + P_2\,y - P_1\,t \quad , \tag{1.331}$$

where

$$\theta = \frac{eB}{c\sqrt{2mP_1}}\,\sin^{-1}\!\bigg(x - \frac{cP_2}{eB}\bigg) \quad . \tag{1.332}$$

Note that, from eqn. 1.326, we may write

$$dx = \frac{c}{eB}\,dP_2 + \frac{mc}{eB}\,\frac{1}{\sqrt{2mP_1}}\,\sin\theta\,dP_1 + \frac{c}{eB}\sqrt{2mP_1}\,\cos\theta\,d\theta \quad , \tag{1.333}$$

from which we derive

$$\frac{\partial\theta}{\partial P_1} = -\frac{\tan\theta}{2P_1} \quad , \qquad \frac{\partial\theta}{\partial P_2} = -\frac{1}{\sqrt{2mP_1}\,\cos\theta} \quad . \tag{1.334}$$

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

$$\begin{aligned} Q_1 = \frac{\partial S}{\partial P_1} &= \frac{mc}{eB}\,\theta + \frac{mcP_1}{eB}\,\frac{\partial\theta}{\partial P_1} + \frac{mc}{2eB}\,\sin(2\theta) + \frac{mcP_1}{eB}\,\cos(2\theta)\,\frac{\partial\theta}{\partial P_1} - t \\ &= \frac{\theta}{\omega_c} - t \end{aligned} \tag{1.335}$$

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

$$\begin{aligned} Q_2 = \frac{\partial S}{\partial P_2} &= y + \frac{mcP_1}{eB}\big[1+\cos(2\theta)\big]\,\frac{\partial\theta}{\partial P_2} \\ &= y - \frac{c}{eB}\sqrt{2mP_1}\,\cos\theta \quad . \end{aligned} \tag{1.336}$$

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:

$$\begin{aligned} x(t) &= x_0 + A\,\sin(\omega_c t + \delta) \\ y(t) &= y_0 + A\,\cos(\omega_c t + \delta) \quad , \end{aligned} \tag{1.337}$$

and

$$x_0 = \frac{cP_2}{eB} \quad , \quad y_0 = Q_2 \quad , \quad \delta \equiv \omega_c\, Q_1 \quad , \quad A = \frac{c}{eB}\sqrt{2mP_1} \quad . \tag{1.338}$$

1.11.3 Action-angle variables for the Kepler problem

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[23]

$$W(r,\theta,\varphi) = W_r(r) + W_\theta(\theta) + W_\varphi(\varphi) \quad , \tag{1.339}$$

hence

$$\frac{1}{2m}\bigg(\frac{\partial W_r}{\partial r}\bigg)^2 + \frac{1}{2mr^2}\bigg(\frac{\partial W_\theta}{\partial \theta}\bigg)^2 + \frac{1}{2mr^2\sin^2\!\theta}\bigg(\frac{\partial W_\varphi}{\partial \varphi}\bigg)^2 + V(r) = E \equiv \Lambda_r \quad . \tag{1.340}$$

Separating, we have

$$\frac{1}{2m}\bigg(\frac{dW_\varphi}{d\varphi}\bigg)^2 = \Lambda_\varphi \quad \Rightarrow \quad J_\varphi = \oint\limits_{\mathcal{C}_\varphi} d\varphi\,\frac{dW_\varphi}{d\varphi} = 2\pi\sqrt{2m\Lambda_\varphi} \quad . \tag{1.341}$$

Next we deal with the θ coordinate. We have

$$\frac{1}{2m}\bigg(\frac{dW_\theta}{d\theta}\bigg)^2 = \Lambda_\theta - \frac{\Lambda_\varphi}{\sin^2\!\theta} \quad , \tag{1.342}$$

and therefore

$$\begin{aligned} J_\theta &= 4\sqrt{2m\Lambda_\theta}\int\limits_{\theta_0}^{\pi/2} d\theta\,\sqrt{1 - \big(\Lambda_\varphi/\Lambda_\theta\big)\csc^2\!\theta} \\ &= 2\pi\sqrt{2m}\,\Big(\sqrt{\Lambda_\theta} - \sqrt{\Lambda_\varphi}\Big) \quad , \end{aligned} \tag{1.343}$$

where $\theta_0 = \sin^{-1}(\Lambda_\varphi/\Lambda_\theta)$. Finally, we have for the radial coordinate

$$\frac{1}{2m}\bigg(\frac{dW_r}{dr}\bigg)^2 = E + \frac{k}{r} - \frac{\Lambda_\theta}{r^2} \quad , \tag{1.344}$$

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

and so[24]

$$
\begin{aligned}
J_r &= \oint\limits_{\mathcal{C}_r} dr \sqrt{2m\left(E + \frac{k}{r} - \frac{\Lambda_\theta}{r^2}\right)} \\
&= -(J_\theta + J_\varphi) + \pi k \sqrt{\frac{2m}{|E|}} \quad ,
\end{aligned} \tag{1.345}
$$

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

Thus, we find

$$
H = E = -\frac{2\pi^2 m k^2}{\big(J_r + J_\theta + J_\varphi\big)^2} \quad . \tag{1.346}
$$

Note that the frequencies are completely degenerate:

$$
\nu \equiv \nu_{r,\theta,\varphi} = \frac{\partial H}{\partial J_{r,\theta,\varphi}} = \frac{4\pi^2 m k^2}{\big(J_r + J_\theta + J_\varphi\big)^3} = \left(\frac{\pi^2 m k^2}{2\,|E|^3}\right)^{1/2} \quad . \tag{1.347}
$$

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

$$
\Big\{(\phi_r, J_r),\ (\phi_\theta, J_\theta),\ (\phi_\varphi, J_\varphi)\Big\} \longrightarrow \Big\{(\chi_1, \mathcal{J}_1),\ (\chi_2, \mathcal{J}_2),\ (\chi_3, \mathcal{J}_3)\Big\} \quad , \tag{1.348}
$$

using the type-II generator

$$
F_2(\phi_r, \phi_\theta, \phi_\varphi; \mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3) = (\phi_\varphi - \phi_\theta)\,\mathcal{J}_1 + (\phi_\theta - \phi_r)\,\mathcal{J}_2 + \phi_r\,\mathcal{J}_3 \quad , \tag{1.349}
$$

which results in

$$
\chi_1 = \frac{\partial F_2}{\partial \mathcal{J}_1} = \phi_\varphi - \phi_\theta \qquad\qquad J_r = \frac{\partial F_2}{\partial \phi_r} = \mathcal{J}_3 - \mathcal{J}_2 \tag{1.350}
$$

$$
\chi_2 = \frac{\partial F_2}{\partial \mathcal{J}_2} = \phi_\theta - \phi_r \qquad\qquad J_\theta = \frac{\partial F_2}{\partial \phi_\theta} = \mathcal{J}_2 - \mathcal{J}_1 \tag{1.351}
$$

$$
\chi_3 = \frac{\partial F_2}{\partial \mathcal{J}_3} = \phi_r \qquad\qquad J_\varphi = \frac{\partial F_2}{\partial \phi_\varphi} = \mathcal{J}_1 \quad . \tag{1.352}
$$

The new Hamiltonian is

$$
H(\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3) = -\frac{2\pi^2 m k^2}{\mathcal{J}_3^2} \quad , \tag{1.353}
$$

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

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

1.11.4 Action-angle variables for charged particle in a magnetic field

For the case of the charged particle in a magnetic field, studied above in section 1.11.2, we found

$$x = \frac{cP_2}{eB} + \frac{c}{eB}\sqrt{2mP_1}\,\sin\theta \tag{1.354}$$

with $p_x = \sqrt{2mP_1}\,\cos\theta$ and $p_y = P_2$. The action variable J is then

$$J = \oint p_x\,dx = \frac{2mcP_1}{eB}\int_0^{2\pi} d\theta\,\cos^2\theta = \frac{mcP_1}{eB} \quad . \tag{1.355}$$

We then have

$$W = J\theta + \tfrac{1}{2}J\sin(2\theta) + Py \quad , \tag{1.356}$$

where $P \equiv P_2$. Thus,

$$\begin{aligned}\phi = \frac{\partial W}{\partial J} &= \theta + \tfrac{1}{2}\sin(2\theta) + J\big[1+\cos(2\theta)\big]\,\frac{\partial\theta}{\partial J}\\ &= \theta + \tfrac{1}{2}\sin(2\theta) + 2J\cos^2\theta\cdot\left(-\frac{\tan\theta}{2J}\right) = \theta \quad .\end{aligned} \tag{1.357}$$

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

$$Q = \frac{\partial W}{\partial P} = y - \sqrt{\frac{2cJ}{eB}}\,\cos\phi \quad . \tag{1.358}$$

Therefore, we have

$$x = \frac{cP}{eB} + \sqrt{\frac{2cJ}{eB}}\,\sin\phi \qquad , \qquad y = Q + \sqrt{\frac{2cJ}{eB}}\,\cos\phi \tag{1.359}$$

and

$$p_x = \sqrt{\frac{2eBJ}{c}}\,\cos\phi \qquad , \qquad p_y = P \quad . \tag{1.360}$$

The Hamiltonian is

$$\begin{aligned}H &= \frac{P_x^2}{2m} + \frac{1}{2m}\left(p_y - \frac{eBx}{c}\right)^2\\ &= \frac{eBJ}{mc}\,\cos^2\phi + \frac{eBJ}{mc}\,\sin^2\phi = \omega_c\,J \quad ,\end{aligned} \tag{1.361}$$

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

$$\dot\phi = \frac{\partial H}{\partial J} = \omega_c \qquad , \qquad \dot J = -\frac{\partial H}{\partial\phi} = 0 \tag{1.362}$$

and

$$\dot Q = \frac{\partial H}{\partial P} = 0 \qquad , \qquad \dot P = -\frac{\partial H}{\partial Q} = 0 \quad . \tag{1.363}$$

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

1.11.5 Canonical perturbation theory for the cubic oscillator

Consider the Hamiltonian

$$H = \frac{p^2}{2m} + \tfrac{1}{2} m\,\omega_0^2\,q^2 + \tfrac{1}{3}\epsilon\,m\,\omega_0^2\,\frac{q^3}{a} \quad , \tag{1.364}$$

where ϵ is a small dimensionless parameter.

(a) Show that the oscillation frequency satisfies $\nu(J) = \omega_0 + \mathcal{O}(\epsilon^2)$. That is, show that the first order (in ϵ) frequency shift vanishes.

Solution: It is good to recall the basic formulae

$$q = \sqrt{\frac{2J_0}{m\omega_0}}\,\sin\phi_0 \qquad , \qquad p = \sqrt{2m\,\omega_0\,J_0}\,\cos\phi_0 \tag{1.365}$$

as well as the results

$$\begin{aligned} J_0 &= \frac{\partial S}{\partial \phi_0} = J + \epsilon\,\frac{\partial S_1}{\partial \phi_0} + \epsilon^2\,\frac{\partial S_2}{\partial \phi_0} + \ldots \\ \phi &= \frac{\partial S}{\partial J} = \phi_0 + \epsilon\,\frac{\partial S_1}{\partial J} + \epsilon^2\,\frac{\partial S_2}{\partial J} + \ldots \quad , \end{aligned} \tag{1.366}$$

and

$$\begin{aligned} E_0(J) &= \widetilde{H}_0(J) \\ E_1(J) &= \widetilde{H}_1(\phi_0, J) + \frac{\partial \widetilde{H}_0}{\partial J}\,\frac{\partial S_1}{\partial \phi_0} \\ E_2(J) &= \frac{\partial \widetilde{H}_0}{\partial J}\,\frac{\partial S_2}{\partial \phi_0} + \frac{1}{2}\,\frac{\partial^2 \widetilde{H}_0}{\partial J^2}\left(\frac{\partial S_1}{\partial \phi_0}\right)^2 + \frac{\partial \widetilde{H}_1}{\partial J}\,\frac{\partial S_1}{\partial \phi_0} \quad . \end{aligned} \tag{1.367}$$

Expressed in action-angle variables,

$$\begin{aligned} \widetilde{H}_0(\phi_0, J) &= \omega_0\,J \\ \widetilde{H}_1(\phi_0, J) &= \frac{2}{3}\sqrt{\frac{2\omega_0}{ma^2}}\,J^{3/2}\,\sin^3\!\phi_0 \quad . \end{aligned} \tag{1.368}$$

Thus, $\nu_0 = \frac{\partial \widetilde{H}_0}{\partial J} = \omega_0$.

Averaging the equation for $E_1(J)$ yields

$$E_1(J) = \big\langle \widetilde{H}_1(\phi_0, J) \big\rangle = \frac{2}{3}\sqrt{\frac{2\omega_0}{ma^2}}\,J^{3/2}\,\big\langle \sin^3\!\phi_0 \big\rangle = 0 \quad . \tag{1.369}$$

(b) Compute the frequency shift $\nu(J)$ to second order in ϵ.

Solution : From the equation for E_1, we also obtain

$$\frac{\partial S_1}{\partial \phi_0} = \frac{1}{\nu_0}\Big(\langle \widetilde{H}_1 \rangle - \widetilde{H}_1\Big) \quad . \tag{1.370}$$

Inserting this into the equation for $E_2(J)$ and averaging then yields

$$E_2(J) = \frac{1}{\nu_0}\left\langle \frac{\partial \widetilde{H}_1}{\partial J}\Big(\langle \widetilde{H}_1 \rangle - \widetilde{H}_1\Big)\right\rangle = -\frac{1}{\nu_0}\left\langle \widetilde{H}_1\,\frac{\partial \widetilde{H}_1}{\partial J}\right\rangle = -\frac{4\nu_0 J^2}{3ma^2}\,\langle \sin^6\!\phi_0 \rangle \tag{1.371}$$

In computing the average of $\sin^6\!\phi_0$, it is good to recall the binomial theorem, or the Fibonacci tree. The sixth order coefficents are easily found to be $\{1, 6, 15, 20, 15, 6, 1\}$, whence

$$\begin{aligned} \sin^6\!\phi_0 &= \frac{1}{(2i)^6}\big(e^{i\phi_0} - e^{-i\phi_0}\big)^6 \\ &= \tfrac{1}{64}\big(-2\sin 6\phi_0 + 12\sin 4\phi_0 - 30\sin 2\phi_0 + 20\big) \quad . \end{aligned} \tag{1.372}$$

Thus $\langle \sin^6\!\phi_0 \rangle = \frac{5}{16}$, whence

$$E(J) = \omega_0\, J - \tfrac{5}{12}\epsilon^2\,\frac{J^2}{ma^2} \tag{1.373}$$

and

$$\nu(J) = \frac{\partial E}{\partial J} = \omega_0 - \tfrac{5}{6}\epsilon^2\,\frac{J}{ma^2} \quad . \tag{1.374}$$

(c) Find $q(t)$ to order ϵ. Your result should be finite for all times.

Solution : From the equation for $E_1(J)$, we have

$$\frac{\partial S_1}{\partial \phi_0} = -\frac{2}{3}\sqrt{\frac{2J^3}{m\omega_0 a^2}}\,\sin^3\!\phi_0 \quad . \tag{1.375}$$

Integrating, we obtain

$$\begin{aligned} S_1(\phi_0, J) &= \frac{2}{3}\sqrt{\frac{2J^3}{m\omega_0 a^2}}\,\big(\cos\phi_0 - \tfrac{1}{3}\cos^3\!\phi_0\big) \\ &= \frac{J^{3/2}}{\sqrt{2m\omega_0 a^2}}\,\big(\cos\phi_0 - \tfrac{1}{9}\cos 3\phi_0\big) \quad . \end{aligned} \tag{1.376}$$

Thus, with

$$S(\phi_0, J) = \phi_0\, J + \epsilon\, S_1(\phi_0, J) + \ldots \quad , \tag{1.377}$$

we have

$$\begin{aligned} \phi &= \frac{\partial S}{\partial J} = \phi_0 + \frac{3}{2}\,\frac{\epsilon\, J^{1/2}}{\sqrt{2m\omega_0 a^2}}\,\big(\cos\phi_0 - \tfrac{1}{9}\cos 3\phi_0\big) \\ J_0 &= \frac{\partial S}{\partial \phi_0} = J - \frac{\epsilon\, J^{3/2}}{\sqrt{2m\omega_0 a^2}}\,\big(\sin\phi_0 - \tfrac{1}{3}\sin 3\phi_0\big) \quad . \end{aligned} \tag{1.378}$$

Inverting, we may write ϕ_0 and J_0 in terms of ϕ and J:

$$\begin{aligned}\phi_0 &= \phi + \frac{3}{2}\,\frac{\epsilon\, J^{1/2}}{\sqrt{2m\omega_0 a^2}}\,\big(\tfrac{1}{9}\cos 3\phi - \cos\phi\big)\\ J_0 &= J + \frac{\epsilon\, J^{3/2}}{\sqrt{2m\omega_0 a^2}}\,\big(\tfrac{1}{3}\sin 3\phi - \sin\phi\big) \quad .\end{aligned} \tag{1.379}$$

Thus,

$$\begin{aligned}q(t) &= \sqrt{\frac{2J_0}{m\omega_0}}\,\sin\phi_0\\ &= \sqrt{\frac{2J}{m\omega_0}}\,\sin\phi\cdot\Big(1+\frac{\delta J}{2J}+\ldots\Big)\Big(\sin\phi+\delta\phi\,\cos\phi+\ldots\Big)\\ &= \sqrt{\frac{2J}{m\omega_0}}\,\sin\phi - \frac{\epsilon\, J}{m\omega_0 a}\,\big(1+\tfrac{1}{3}\cos 2\phi\big) + \mathcal{O}\big(\epsilon^2\big) \quad ,\end{aligned} \tag{1.380}$$

with

$$\phi(t) = \phi(0) + \nu(J)\,t \quad . \tag{1.381}$$