Prepared for submission to JHEP

Classical and Quantum Mechanics for Mathematicians

Marcos Mariño

Département de Physique Théorique et Section de Mathématiques
Université de Genève, Genève, CH-1211 Switzerland

E-mail: marcos.marino@unige.ch

Abstract: Notes for the course.

Contents

1 Lagrangian mechanics

In this course, we will assume that the configuration space of the system is $M = \mathbb{R}^n$, but will point out the appropriate generalizations to an arbitrary manifold M.

1.1 The principle of least action

In Classical Mechanics (CM), the *state* of a system at any instant of time is specified by the values of its generalized coordinates $\mathbf{q} \in M = \mathbb{R}^n$ and its generalized velocities $\mathbf{v} \in \mathbb{R}^n$. The basic principle of CM is that, given the state of a system at an initial time t, its motion is determined at all times (past and future). The motion is described by the classical trajectory

$$\mathbf{q}(t) = (q_1(t), \cdots, q_n(t)) \in \mathbb{R}^n. \tag{1.1}$$

A *Lagrangian system* is defined by a real-valued, smooth function

$$\begin{aligned} L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} &\to \mathbb{R} \\ (\mathbf{q}, \mathbf{v}, t) &\mapsto L(\mathbf{q}, \mathbf{v}, t) \end{aligned} \tag{1.2}$$

called the *Lagrangian function*. This determines the motion of the system, as follows. We consider the space of smooth parametrized paths in $\mathbb{R}^n$ connecting the points $\mathbf{q}_0$ and $\mathbf{q}_1$,

$$P_{\mathbf{q}_0,t_0}^{\mathbf{q}_1,t_1} = \{\mathbf{q} : [t_0, t_1] \to \mathbb{R}^n : \ \mathbf{q}(t_0) = \mathbf{q}_0, \ \mathbf{q}(t_1) = \mathbf{q}_1\}. \tag{1.3}$$

A variation $\mathbf{Q}$ with fixed ends is a family of paths

$$\mathbf{q}_\epsilon(t) = \mathbf{Q}(t, \epsilon) \tag{1.4}$$

given by a smooth map

$$\mathbf{Q} : [t_0, t_1] \times [-\epsilon_0, \epsilon_0] \to \mathbb{R}^n \tag{1.5}$$

such that

$$\mathbf{Q}(t, 0) = \mathbf{q}(t), \quad t \in [t_0, t_1], \tag{1.6}$$

and

$$\mathbf{Q}(t_0, \epsilon) = \mathbf{q}_0, \quad \mathbf{Q}(t_1, \epsilon) = \mathbf{q}_1, \qquad \epsilon \in [-\epsilon_0, \epsilon_0]. \tag{1.7}$$

The quantity

$$\delta\mathbf{q} = \left.\frac{\partial \mathbf{Q}}{\partial \epsilon}\right|_{\epsilon=0} \tag{1.8}$$

is called an *infinitesimal variation*.

Definition 1.1. The *action functional* $S : P_{\mathbf{q}_0,t_0}^{\mathbf{q}_1,t_1} \to \mathbb{R}$ of a Lagrangian system is

$$S(\gamma) = \int_{t_0}^{t_1} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) \mathrm{d}t. \tag{1.9}$$

The motion of a Lagrangian system is determined by the *principle of least action* or *Hamilton's principle*: at path $\mathbf{q}(t) \in P_{\mathbf{q}_0,t_0}^{\mathbf{q}_1,t_1}$ describes the motion of the system if and only if it is an extremal of the action functional, i.e. if

$$\left.\frac{\mathrm{d}S(\mathbf{q}_\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = 0, \tag{1.10}$$

where $\mathbf{q}_\epsilon$ is an arbitrary variation of $\mathbf{q}(t)$ with fixed ends.

The principle of least action leads to equations of motion (EOM), and we have the following

Theorem 1.2. *A path $\mathbf{q}(t)$ is an extremal of the action functional if and only if it satisfies the EOM*

$$\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0, \qquad i = 1, \cdots, n. \tag{1.11}$$

Proof: this follows by direct computation:

$$\begin{aligned}\left.\frac{\mathrm{d}S(\mathbf{q}_\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0}\int_{t_0}^{t_1} L(\mathbf{q}_\epsilon(t),\dot{\mathbf{q}}_\epsilon(t),t)\mathrm{d}t = \sum_{i=1}^{n}\int_{t_0}^{t_1}\left(\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i\right)\mathrm{d}t \\ &= \sum_{i=1}^{n}\int_{t_0}^{t_1}\left\{\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right\}\delta q_i\mathrm{d}t + \sum_{i=1}^{n}\left.\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right|_{t_0}^{t_1},\end{aligned} \tag{1.12}$$

and the last term vanishes since

$$\delta q_i(t_0) = \delta q_i(t_1) = 0. \tag{1.13}$$

Since the first term must vanish for arbitrary infinitesimal variations δq_i, the EOM (1.11) follows. □

Example 1.3. The dynamics of N interacting particles in $\mathbb{R}^3$ requires $3N$ generalized coordinates described by a configuration space

$$M = \mathbb{R}^{3N}. \tag{1.14}$$

We will denote the position vectors of these particles by

$$\boldsymbol{r} = (\boldsymbol{r}_1, \cdots, \boldsymbol{r}_N), \tag{1.15}$$

where $\boldsymbol{r}_a \in \mathbb{R}^3$ is the position vector of the a-th particle. The Lagrangian is given by

$$L = \frac{1}{2}\sum_{a=1}^{N} m_a\dot{\boldsymbol{r}}_a^2 - V(\boldsymbol{r}) = T - V. \tag{1.16}$$

Here, m_a is the mass of the a-th particle, and

$$T = \frac{1}{2}\sum_{a=1}^{N} m_a\dot{\boldsymbol{r}}_a^2 \tag{1.17}$$

is called the kinetic energy of the particles. $V(\boldsymbol{r})$ is called the potential energy. The Euler–Lagrange equations give Newton's equation of motion,

$$m_a\ddot{\boldsymbol{r}}_a = \boldsymbol{F}_a, \tag{1.18}$$

where

$$\boldsymbol{F}_a = -\frac{\partial V}{\partial \boldsymbol{r}_a} \tag{1.19}$$

is the force on the a-th particle.

1.2 Noether's theorem

Definition 1.4. A smooth function

$$I : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{1.20}$$

is called an *integral of motion* for a Lagrangian system if

$$\frac{\mathrm{d}}{\mathrm{d}t} I\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right) = 0 \tag{1.21}$$

for all extremals $\mathbf{q}(t)$ of the action functional.

Example 1.5. The quantity

$$E(\mathbf{q}(t), \dot{\mathbf{q}}(t)) = \sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L(\mathbf{q}(t), \dot{\mathbf{q}}(t)), \tag{1.22}$$

called the *energy* of the Lagrangian system, is a conserved quantity. This follows immediately from the Euler–Lagrange equations.

Definition 1.6. A Lagrangian L is invariant under a diffeomorphism

$$g : M \to M \tag{1.23}$$

if

$$L(g(\mathbf{q}), g_\star(\mathbf{q})\mathbf{v}, t) = L(\mathbf{q}, \mathbf{v}, t), \tag{1.24}$$

for all $\mathbf{q} \in \mathbb{R}^n$, $\mathbf{v} \in \mathbb{R}^n$. Here,

$$\left(g_\star(\mathbf{q})\mathbf{v}\right)_i = \sum_{j=1}^{n} \frac{\partial g_i}{\partial q_j} v_j. \tag{1.25}$$

Such a diffeomorphism is called a *symmetry* of L.

Theorem 1.7. *(Noether) Let us suppose that L is invariant under a one-parameter group of diffeomorphisms g_s, $s \in \mathbb{R}$. Then, the following is a constant of motion:*

$$I(\mathbf{q}, \dot{\mathbf{q}}) = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} \left(\frac{\partial g_s^i}{\partial s} \right)_{s=0}. \tag{1.26}$$

Proof: Let $\mathbf{q}(t)$ be a solution to the Euler–Lagrange equations. By assumption,

$$L\left(g_s(\mathbf{q}(t)), (g_s)_\star(\mathbf{q})\dot{\mathbf{q}}(t)\right) = L\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right), \qquad s \in \mathbb{R}. \tag{1.27}$$

Note that

$$(g_s)_\star(\mathbf{q})\dot{\mathbf{q}}(t) = \frac{\mathrm{d} g_s(\mathbf{q}(t))}{\mathrm{d}t}. \tag{1.28}$$

It follows from (1.27) that

$$\frac{\partial}{\partial s} \left[L\left(g_s(\mathbf{q}(t)), \dot{g}_s(\mathbf{q}(t))\right) \right]_{s=0} = 0. \tag{1.29}$$

Calculating the derivative, we find

$$\sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} \left(\frac{\partial g_s^i}{\partial s} \right)_{s=0} + \frac{\partial L}{\partial \dot{q}_i} \left(\frac{\partial \dot{g}_s^i}{\partial s} \right)_{s=0} \right) = 0. \tag{1.30}$$

If we now use the Euler–Lagrange equations, we obtain

$$\sum_{i=1}^{n} \left(\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \left(\frac{\partial g_s^i}{\partial s} \right)_{s=0} + \frac{\partial L}{\partial \dot{q}_i} \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial g_s^i}{\partial s} \right)_{s=0} \right) = 0, \tag{1.31}$$

or

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[\sum_{i=1}^{n} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \left(\frac{\partial g_s^i}{\partial s} \right)_{s=0} \right] = 0, \tag{1.32}$$

which is precisely what we wanted to show. □

Example 1.8. Let

$$g_s(\mathbf{q}) = \mathbf{q} + s\boldsymbol{\xi}, \qquad \boldsymbol{\xi} \in \mathbb{R}^n. \tag{1.33}$$

Then, according to Noether's theorem,

$$I = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} \xi_i \tag{1.34}$$

is an integral of motion. A particular example is a system of N interacting particles in three dimensions, where the potential is of the form

$$V(\boldsymbol{r}) = \sum_{1 \le a < b \le N} V_{ab}(|\boldsymbol{r}_a - \boldsymbol{r}_b|). \tag{1.35}$$

In this case, the Lagrangian is clearly invariant if we translate all coordinates $\boldsymbol{r}_a$ by a multiple of the same arbitrary vector $\boldsymbol{c} = (c_1, c_2, c_3) \in \mathbb{R}^3$, so that we can take

$$\boldsymbol{\xi} = (\boldsymbol{c}, \cdots, \boldsymbol{c})\,. \tag{1.36}$$

It follows that the integral of motion is

$$I = \sum_{a=1}^{N} \sum_{i=1}^{3} c_i \frac{\partial L}{\partial \dot{r}_a^i} = \sum_{i=1}^{3} c_i P_i, \tag{1.37}$$

where P_i, $i = 1, 2, 3$ define the vector

$$\boldsymbol{P} = \sum_{a=1}^{n} \frac{\partial L}{\partial \dot{\boldsymbol{r}}_a}, \tag{1.38}$$

which is the (conserved) total momentum of the system.

1.3 Generalization to smooth manifolds

In all previous considerations, we have assumed that the configuration space M was $\mathbb{R}^n$. However, we can generalize everything we have done to an arbitrary differentiable manifold M. The space of positions and velocities $\mathbb{R}^n \times \mathbb{R}^n$ is naturally generalized to the tangent bundle TM. Let (U, ϕ) be a local coordinate patch. Then, we will denote by $\mathbf{q} = (q_1, \cdots, q_n)$ the local coordinates associated to the chart, i.e. $q_i = x_i \circ \phi$. The *standard coordinates* on the chart TU are

$$(\mathbf{q}, \mathbf{v}) = (q_1, \cdots, q_n; v_1, \cdots, v_n), \tag{1.39}$$

where the coordinates v_i correspond to the basis

$$\left(\frac{\partial}{\partial q_1}\right)_p, \cdots, \left(\frac{\partial}{\partial q_n}\right)_p \tag{1.40}$$

of $T_p(M)$, i.e. a vector X_p in $T_p(M)$ will be given by

$$X_p = \sum_{i=1}^{n} v_i \left(\frac{\partial}{\partial q_i}\right)_p. \tag{1.41}$$

Let (U, ϕ), (U', ϕ') be two different coordinate charts of the atlas on M. Under a change of coordinates, we have a diffeomorphism

$$\phi' \circ \phi^{-1} : \phi(U \cap U') \to \phi'(U \cap U'). \tag{1.42}$$

We will write

$$\mathbf{q}' = \boldsymbol{F}(\mathbf{q}), \qquad \mathbf{v}' = \frac{\partial \boldsymbol{F}}{\partial \mathbf{q}} \mathbf{v}, \tag{1.43}$$

where $\boldsymbol{F} = (F_1, \cdots, F_n)$ is a vector of n functions with the coordinates of the diffeomorphism (1.42), and in the r.h.s. of the second equation we have the $n \times n$ matrix,

$$\left(\frac{\partial \mathbf{F}}{\partial \mathbf{q}} \right)_{ij} = \frac{\partial F_i}{\partial q_j}, \qquad i, j = 1, \cdots, n. \tag{1.44}$$

In particular, if we consider a curve in local coordinates $\mathbf{q}(t)$, we find that, under a change of coordinates,

$$\dot{\mathbf{q}}' = \frac{\partial \boldsymbol{F}}{\partial \mathbf{q}} \dot{\mathbf{q}}. \tag{1.45}$$

We also have that

$$\mathrm{d}\mathbf{q}' = \frac{\partial \boldsymbol{F}}{\partial \mathbf{q}} \mathrm{d}\mathbf{q} \tag{1.46}$$

and

$$\mathrm{d}\dot{q}'_i = \sum_{j,k=1}^{n} \frac{\partial^2 F_i}{\partial q_j \partial q_k} \dot{q}_j \mathrm{d}q_k + \sum_{j=1}^{n} \frac{\partial F_i}{\partial q_j} \mathrm{d}\dot{q}_j. \tag{1.47}$$

We can write

$$\mathrm{d}\dot{\mathbf{q}}' = \boldsymbol{G}(\mathbf{q}, \dot{\mathbf{q}}) \mathrm{d}\mathbf{q} + \frac{\partial \boldsymbol{F}}{\partial \mathbf{q}} \mathrm{d}\dot{\mathbf{q}}, \tag{1.48}$$

where $\boldsymbol{G}(\mathbf{q}, \dot{\mathbf{q}})$ is a matrix-valued function.

In this more general context, the Lagrangian is a differentiable function

$$L : TM \times \mathbb{R} \to \mathbb{R}. \tag{1.49}$$

In a coordinate patch TU, this leads to a function in local coordinates $L(\mathbf{q}, \mathbf{v}, t)$.

Let us suppose that we change coordinates from $\mathbf{q}$ to $\mathbf{q}'$, as in (1.43). Let us first consider $L(\mathbf{q}', \dot{\mathbf{q}}, t)$. Under an infinitesimal variation of its arguments, the Lagrangian changes as

$$\mathrm{d}L = \frac{\partial L}{\partial \mathbf{q}'} \cdot \mathrm{d}\mathbf{q}' + \frac{\partial L}{\partial \dot{\mathbf{q}}'} \cdot \mathrm{d}\dot{\mathbf{q}}' + \frac{\partial L}{\partial t} \mathrm{d}t. \tag{1.50}$$

Using the transformation laws above, and the fact that L is a globally well-defined function on TM, we find that

$$\begin{aligned} \mathrm{d}L &= \frac{\partial L}{\partial \mathbf{q}'} \cdot \frac{\partial \boldsymbol{F}}{\partial \mathbf{q}} \mathrm{d}\mathbf{q} + \frac{\partial L}{\partial \dot{\mathbf{q}}'} \cdot \left(\boldsymbol{G}(\mathbf{q}, \dot{\mathbf{q}}) \mathrm{d}\mathbf{q} + \frac{\partial \boldsymbol{F}}{\partial \mathbf{q}} \mathrm{d}\dot{\mathbf{q}} \right) + \frac{\partial L}{\partial t} \mathrm{d}t \\ &= \left(\frac{\partial L}{\partial \mathbf{q}'} \cdot \frac{\partial \boldsymbol{F}}{\partial \mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}'} \cdot \boldsymbol{G}(\mathbf{q}, \dot{\mathbf{q}}) \right) \mathrm{d}\mathbf{q} + \frac{\partial L}{\partial \dot{\mathbf{q}}'} \cdot \frac{\partial \boldsymbol{F}}{\partial \mathbf{q}} \mathrm{d}\dot{\mathbf{q}} + \frac{\partial L}{\partial t} \mathrm{d}t \\ &= \frac{\partial L}{\partial \mathbf{q}} \cdot \mathrm{d}\mathbf{q} + \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \mathrm{d}\dot{\mathbf{q}} + \frac{\partial L}{\partial t} \mathrm{d}t. \end{aligned} \tag{1.51}$$

It follows that

$$\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i'} \frac{\partial F_i}{\partial q_j} = \frac{\partial L}{\partial \dot{q}_j}, \qquad j = 1, \cdots, n. \tag{1.52}$$

If we multiply by $\mathrm{d}q_j$ both sides and we sum over j, we find

$$\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i'} \mathrm{d}q_i' = \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j} \mathrm{d}q_j. \tag{1.53}$$

This shows that the components of

$$\frac{\partial L}{\partial \dot{\mathbf{q}}} = \left(\frac{\partial L}{\partial \dot{q}_1}, \cdots, \frac{\partial L}{\partial \dot{q}_n} \right) \tag{1.54}$$

transform as the components of a one-form on M.

When the configuration space is a general manifold M, the diffeomorphism g in (1.23) induces a map between the tangent bundles,

$$g_\star : TM \to TM \tag{1.55}$$

and the invariance of L is simply stated as

$$L(g_\star(p)) = L(p), \qquad p \in TM. \tag{1.56}$$

We also point out that, given a one-parameter group of diffeomorphisms g_s, we can associate its infinitesimal generator, which is a vector field with local coordinates

$$X = \sum_{i=1}^{n} \left(\frac{\partial g_s^i}{\partial s} \right)_{s=0} \frac{\partial}{\partial q_i}. \tag{1.57}$$

2 Hamiltonian Mechanics

2.1 Hamilton's equations

Let $f(\mathbf{x})$ be a convex function of the vector $\mathbf{x} \in \mathbb{R}^n$. Convexity here means that the Hessian matrix

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \tag{2.1}$$

is positive definite. Let us consider the function

$$F(\mathbf{p}, \mathbf{x}) = \mathbf{p} \cdot \mathbf{x} - f(\mathbf{x}), \tag{2.2}$$

where

$$\mathbf{p} = \frac{\partial f}{\partial \mathbf{x}}. \tag{2.3}$$

Then, the *Legendre transform* of $f(\mathbf{x})$ is given by

$$g(\mathbf{p}) = F\left(\mathbf{p}, \mathbf{x}(\mathbf{p})\right) = \max_{\mathbf{x}} F(\mathbf{p}, \mathbf{x}). \tag{2.4}$$

Example 2.1. Let

$$f(x) = \frac{mx^2}{2}. \tag{2.5}$$

Then,

$$p = \frac{\partial f}{\partial x} = mx \Rightarrow x(p) = \frac{p}{m}. \tag{2.6}$$

We also have that

$$g(p) = px - f(x(p)) = \frac{p^2}{m} - \frac{p^2}{2m} = \frac{p^2}{2m}. \tag{2.7}$$

We will now apply this construction to a Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, which we assume to be convex w.r.t. the second argument $\dot{\mathbf{q}}$. We have the following

Theorem 2.2. *The system of Euler–Lagrange equations is equivalent to the system of* $2n$ *first order equations*

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}, \qquad \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}, \tag{2.8}$$

where

$$H(\mathbf{q}, \mathbf{p}, t) = \mathbf{p} \cdot \dot{\mathbf{q}}(\mathbf{p}) - L\left(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{p}), t\right) \tag{2.9}$$

is the Legendre transform of the Lagrangian w.r.t. the variable $\dot{\mathbf{q}}$.

Proof: first of all, we note that, according to the Legendre transform, the coordinate $\mathbf{p}$ is given by

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}. \tag{2.10}$$

The Hamiltonian is a function of $\mathbf{q}$, $\mathbf{p}$ and t, and its total differential is

$$\mathrm{d}H = \frac{\partial H}{\partial \mathbf{q}}\mathrm{d}\mathbf{q} + \frac{\partial H}{\partial \mathbf{p}}\mathrm{d}\mathbf{p} + \frac{\partial H}{\partial t}\mathrm{d}t. \tag{2.11}$$

On the other hand, from its definition as a Legendre transform, we have

$$\begin{aligned} \mathrm{d}H &= \dot{\mathbf{q}} \cdot \mathrm{d}\mathbf{p} + \mathbf{p} \cdot \mathrm{d}\dot{\mathbf{q}} - \frac{\partial L}{\partial \mathbf{q}}\mathrm{d}\mathbf{q} - \frac{\partial L}{\partial \dot{\mathbf{q}}}\mathrm{d}\dot{\mathbf{q}} - \frac{\partial L}{\partial t}\mathrm{d}t \\ &= \dot{\mathbf{q}} \cdot \mathrm{d}\mathbf{p} - \frac{\partial L}{\partial \mathbf{q}}\mathrm{d}\mathbf{q} - \frac{\partial L}{\partial t}\mathrm{d}t. \end{aligned} \tag{2.12}$$

By comparing the two expressions, and using Euler–Lagrange equations, which say that

$$\frac{\partial L}{\partial \mathbf{q}} = \dot{\mathbf{p}}, \tag{2.13}$$

Hamilton's equations follow, as well as

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{2.14}$$

□

Example 2.3. Let us consider a Lagrangian of the form $L = T - V(\mathbf{q})$, where

$$T = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\dot{q}_i\dot{q}_j. \tag{2.15}$$

The coefficients a_{ij} form a symmetric, positive-definite matrix, and can be themselves functions of $\mathbf{q}$ and t. We then have

$$H = \frac{\partial T}{\partial \dot{\mathbf{q}}}\cdot\dot{\mathbf{q}} - L = 2T - (T - V) = T + V, \tag{2.16}$$

where we used the fact that T is a homogeneous function of degree 2 in the $\dot{q}_i$, and therefore, by Euler's theorem,

$$\frac{\partial T}{\partial \dot{\mathbf{q}}}\cdot\dot{\mathbf{q}} = 2T. \tag{2.17}$$

Note that, in this case,

$$p_i = \frac{\partial T}{\partial \dot{q}_i} = \sum_{j=1}^{n} a_{ij}\dot{q}_j \Rightarrow \dot{q}_i = \sum_{j=1}^{n} a_{ij}^{-1}p_j, \tag{2.18}$$

and

$$H = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}^{-1}p_ip_j + V(\mathbf{q}). \tag{2.19}$$

Hamilton's equations are, in this case,

$$\dot{q}_i = \sum_{j=1}^{n} a_{ij}^{-1}p_j, \qquad \dot{p}_i = -\frac{1}{2}\sum_{k,l=1}^{n}\frac{\partial a_{kl}^{-1}}{\partial q_i}p_kp_l - \frac{\partial V}{\partial q_i}. \tag{2.20}$$

When $a_{ij} = m_i\delta_{ij}$, we recover Newton's equations.

Corollary 2.1. When evaluated on solutions to the Hamilton's equations, we have

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t}. \tag{2.21}$$

Proof:

$$\begin{aligned}\frac{\mathrm{d}H}{\mathrm{d}t} &= \frac{\partial H}{\partial \mathbf{q}}\cdot\dot{\mathbf{q}} + \frac{\partial H}{\partial \mathbf{p}}\cdot\dot{\mathbf{p}} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial \mathbf{q}}\cdot\frac{\partial H}{\partial \mathbf{p}} - \frac{\partial H}{\partial \mathbf{p}}\cdot\frac{\partial H}{\partial \mathbf{q}} + \frac{\partial H}{\partial t}\\ &= \frac{\partial H}{\partial t}.\end{aligned} \tag{2.22}$$

□

We conclude from the corollary above that, if H does not depend explicitly on time, it is constant on solutions of Hamilton's equations. From now on we will assume for simplicity that this is the case.

Definition 2.4. A coordinate q_i is called *cyclic* if it does not enter into the Lagrangian, i.e.

$$\frac{\partial L}{\partial q_i} = 0. \tag{2.23}$$

If the coordinate q_i does not enter the Lagrangian, it does not enter the Hamiltonian either, so one has

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = 0. \tag{2.24}$$

and the momentum associated to a cyclic coordinate

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.25}$$

is conserved.

Hamilton's equations can be again expressed in terms of a variational principle. To do this, we consider trajectories $\sigma : [t_0, t_1] \to \mathbb{R}^n \times \mathbb{R}^n$, $\sigma(t) = (\mathbf{q}(t), \mathbf{p}(t))$, satisfying

$$\sigma(t_0) = (\mathbf{q}_0, \mathbf{p}(t_0)), \qquad \sigma(t_1) = (\mathbf{q}_1, \mathbf{p}(t_1)). \tag{2.26}$$

We now have the following

Theorem 2.5. *(Poincaré). The path σ is an extremal of the functional*

$$S(\sigma) = \int_{t_0}^{t_1} \left(\mathbf{p}(t) \cdot \dot{\mathbf{q}}(t) - H(\mathbf{q}(t), \mathbf{p}(t))\right) \mathrm{d}t \tag{2.27}$$

if and only if $\mathbf{q}(t)$, $\mathbf{p}(t)$ solve Hamilton's equations.

Proof: An infinitesimal variation of the path

$$(\delta q, \delta p) \tag{2.28}$$

leads to the following variation in the functional,

$$\sum_{i=1}^{n} \int_{t_0}^{t_1} \left(\dot{q}_i \delta p_i - \dot{p}_i \delta q_i - \frac{\partial H}{\partial q_i} \delta q_i - \frac{\partial H}{\partial p_i} \delta p_i \right) \mathrm{d}t + \sum_{i=1}^{n} p_i \delta q_i \Big|_{t_0}^{t_1}. \tag{2.29}$$

Since $\delta q_i(t_0) = \delta q_i(t_1) = 0$, we see that the variation vanishes if and only if Hamilton's equations hold. □

The Hamiltonian is a function on *phase space* $\mathcal{M} = \mathbb{R}^n \times \mathbb{R}^n$, whose coordinates are $(\mathbf{q}, \mathbf{p})$. When the configuration space is a smooth manifold M, the phase space is naturally given by the cotangent bundle

$$\mathcal{M} = T^*M. \tag{2.30}$$

If (U, ϕ) is a coordinate chart for M, with $\phi = (q_1, \cdots, q_n)$, and $T^*U \simeq U \times \mathbb{R}^n$ is the corresponding coordinate chart for T^*M, the coordinates $(p_1, \cdots, p_n)$ are the coordinates in the fiber corresponding to the basis $\mathrm{d}q_1, \cdots, \mathrm{d}q_n$ for T^*U. This is consistent with (2.10) and the fact that the r.h.s. in this equations transform as the components of a one-form on M. We will call these coordinates the *standard coordinates*.

In this, more general setting, the Hamiltonian is a function on phase space

$$H : \mathcal{M} \to \mathbb{R}. \tag{2.31}$$

This function defines a *vector field* X_H on T^*M, which locally, on the chart T^*U, is given by

$$X_H = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} \right). \tag{2.32}$$

The one-parameter flow on $\mathcal{M}$ associated to this vector field is called the *Hamiltonian phase flow*, and it is given by

$$g_t(\mathbf{q}, \mathbf{p}) = (\mathbf{q}(t), \mathbf{p}(t)), \tag{2.33}$$

where $\mathbf{q}(t)$, $\mathbf{p}(t)$ are solutions of Hamilton's equations with the initial conditions

$$\mathbf{q}(0) = \mathbf{q}, \qquad \mathbf{p}(0) = \mathbf{p}. \tag{2.34}$$

Definition 2.6. The one-form θ on T^*M, defined in standard coordinates by

$$\theta = \sum_{i=1}^{n} p_i \mathrm{d}q_i \tag{2.35}$$

is called *Liouville's canonical one-form*. Note that, since the p_i transform as the components of a one-form, Liouville's one-form is globally well-defined.

Definition 2.7. The *canonical symplectic form* on the phase space T^*M is given by

$$\omega = \mathrm{d}\theta. \tag{2.36}$$

In local coordinates on T^*U, it is given by

$$\omega = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q_i. \tag{2.37}$$

Note that the canonical symplectic form ω defines a one-to-one correspondence between one forms ξ in T^*M and vector fields X on T^*M, by the following equation

$$\xi\,(Y) = \omega(Y, X). \tag{2.38}$$

In particular, if we denote by $E^1(\mathcal{M})$ the space of smooth one-forms in $\mathcal{M}$, and by $\mathcal{X}(\mathcal{M})$ the space of smooth vector fields, there is an invertible map

$$\begin{aligned} J : E^1(\mathcal{M}) &\to \mathcal{X}(\mathcal{M}) \\ \xi &\mapsto J(\xi), \end{aligned} \tag{2.39}$$

where $J(\xi)$ is defined by

$$\xi(Y) = \omega(Y, J(\xi)). \tag{2.40}$$

In standard coordinates, this means that, if we are given a one-form

$$\xi = \sum_{i=1}^{n} \left(\xi_i^q \mathrm{d}q_i + \xi_i^p \mathrm{d}p_i\right), \tag{2.41}$$

we obtain a vector field

$$J(\xi) = \sum_{i=1}^{n} \left(\xi_i^p \frac{\partial}{\partial q_i} - \xi_i^q \frac{\partial}{\partial p_i}\right). \tag{2.42}$$

In particular, if the one-form is given by $\xi = \mathrm{d}f$, where $f \in C^\infty(T^*M)$, the associated vector field will be denoted by

$$X_f = J(\mathrm{d}f) \tag{2.43}$$

and it is given by

$$X_f = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial}{\partial p_i}\right). \tag{2.44}$$

If $f = H$, this is precisely the Hamiltonian vector field associated to the Hamiltonian function.

Theorem 2.8. *The Hamiltonian phase flow on T^*M preserves the canonical symplectic form.*

Proof: we calculate in standard coordinates:

$$\begin{aligned}\frac{\mathrm{d}\omega}{\mathrm{d}t} &= \sum_{i=1}^{n} (\mathrm{d}\dot{p}_i \wedge \mathrm{d}q_i + \mathrm{d}p_i \wedge \mathrm{d}\dot{q}_i) = \sum_{i=1}^{n} \left\{ \mathrm{d}\left(-\frac{\partial H}{\partial q_i}\right) \wedge \mathrm{d}q_i + \mathrm{d}p_i \wedge \mathrm{d}\left(\frac{\partial H}{\partial p_i}\right) \right\} \\ &= -\sum_{i=1}^{n} \left\{ \mathrm{d}\left(\frac{\partial H}{\partial q_i}\right) \wedge \mathrm{d}q_i + \mathrm{d}\left(\frac{\partial H}{\partial p_i}\right) \wedge \mathrm{d}p_i \right\} = -\mathrm{d}(\mathrm{d}H) = 0.\end{aligned} \tag{2.45}$$

□

Corollary 2.2. (Liouville's theorem) The *Liouville volume form*

$$\frac{\omega^n}{n!} \tag{2.46}$$

is invariant under the Hamiltonian phase flow.

2.2 Hamilton–Jacobi equation

The action functional (1.9) can be regarded as follows. Let us consider a path $\mathbf{q}(s)$, $s \in [t_0, t] \subset \mathbb{R}$, which solves the Euler–Lagrange equations with initial condition $\mathbf{q}(t_0) = \mathbf{q}_0$. At the time t the path is at the point

$$\mathbf{q}(t) = \mathbf{q}. \tag{2.47}$$

Let us now consider the action as a function of the final time t and final position $\mathbf{q}$ of the path:

$$S(\mathbf{q}, t; \mathbf{q}_0, t_0) = \int_{t_0}^{t} L(\mathbf{q}(s), \dot{\mathbf{q}}(s), s)\mathrm{d}s. \tag{2.48}$$

We have the following

Theorem 2.9. *The differential of the function (2.48) of $\mathbf{q}$, t satisfies*

$$\mathrm{d}S = \mathbf{p} \cdot \mathrm{d}\mathbf{q} - H\mathrm{d}t, \tag{2.49}$$

where $\mathbf{p} = \partial L/\partial\dot{\mathbf{q}}$ and $H = \mathbf{p} \cdot \dot{\mathbf{q}} - L$ are determined by $\dot{\mathbf{q}}$, the velocity of the path $\mathbf{q}(t)$ at the time t.

Proof: let us first fix t and consider an infinitesimal variation of $\mathbf{q}$. Since $\mathbf{q}$ is the final position of a solution of the Euler–Lagrange equations, we consider a family of paths $\mathbf{q}_\epsilon(t)$, such that $\mathbf{q}_0(t) = \mathbf{q}(t)$, as in (1.12). These paths satisfy $\mathbf{q}_\epsilon(t_0) = \mathbf{q}_0$, and they are all extremal. It follows that $\delta\mathbf{q}(t_0) = 0$. The variation of their endpoints is

$$\delta\mathbf{q}(t) = \mathrm{d}\mathbf{q}. \tag{2.50}$$

It follows from (1.12) that

$$\mathrm{d}S = \frac{\partial L}{\partial\dot{\mathbf{q}}}(t) \cdot \mathrm{d}\mathbf{q}, \tag{2.51}$$

therefore

$$\frac{\partial S}{\partial\mathbf{q}} = \frac{\partial L}{\partial\dot{\mathbf{q}}}(t) = \mathbf{p}. \tag{2.52}$$

Let us now consider the general variation. From Barrow's rule,

$$\frac{\mathrm{d}S}{\mathrm{d}t} = L, \tag{2.53}$$

and on the other hand

$$\frac{\mathrm{d}}{\mathrm{d}t} S(\mathbf{q}(t), t; \mathbf{q}_0, t_0)) = \frac{\partial S}{\partial \mathbf{q}} \cdot \dot{\mathbf{q}} + \frac{\partial S}{\partial t}. \tag{2.54}$$

We conclude that

$$\frac{\partial S}{\partial t} = L - \mathbf{p} \cdot \dot{\mathbf{q}} = -H. \tag{2.55}$$

□

Corollary 2.3. (Hamilton–Jacobi equation) The classical action, as a function of the final time t and final position $\mathbf{q}$ of the path, satisfies the following PDE

$$\frac{\partial S}{\partial t} + H\left(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}\right) = 0. \tag{2.56}$$

2.3 Classical observables and Poisson brackets

Let us assume that we have a C^∞ differentiable function on the phase space $\mathcal{M} = \mathbb{R}^n \times \mathbb{R}^n$,

$$f : \mathcal{M} \to \mathbb{R}. \tag{2.57}$$

Such a function will be called a *classical observable*. These functions form an algebra w.r.t. the product structure given by the point-wise multiplication of functions. How does the function f evolve with time? Clearly, if $\mathbf{q}(t)$, $\mathbf{p}(t)$ are solutions to Hamilton's equations with initial conditions (2.34), we have

$$f_t(\mathbf{q}, \mathbf{p}) = f(\mathbf{q}(t), \mathbf{p}(t)). \tag{2.58}$$

Therefore, a classical observable satisfies the equation

$$\frac{\mathrm{d}f_t}{\mathrm{d}t} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i \right) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right). \tag{2.59}$$

This suggests the following

Definition 2.10. The *Poisson bracket* of two classical observables f, g is given by

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right). \tag{2.60}$$

Using the Poisson bracket, the EOM for f_t can be written as

$$\frac{\mathrm{d}f_t}{\mathrm{d}t} = \{f, H\}. \tag{2.61}$$

The Poisson bracket can be clearly extended to an arbitrary phase space of the form $\mathcal{M} = T^*M$ if we note that

$$\{f, g\} = -X_f(g) = X_g(f) = \mathrm{d}f(X_g) = \omega\left(X_g, X_f\right), \tag{2.62}$$

where X_f is given in (2.44) and ω is the canonical symplectic form on T^*M.

Proposition 2.11. The Poisson bracket has the following properties:

1. Skew-symmetry: $\{f, g\} = -\{g, f\}$.
2. Leibniz rule: $\{fg, h\} = f\{g, h\} + g\{f, h\}$.
3. Jacobi identity: $\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0$.

The Poisson bracket turns the algebra of classical observables into a *Lie algebra*, where the Lie bracket is the Poisson bracket. Let us define

$$[X_f, X_g] = X_g X_f - X_f X_g. \tag{2.63}$$

We follow here the convention in [1]. Then, by using Jacobi's identity, we find that

$$[X_f, X_g](h) = \{g, \{f, h\}\} - \{f, \{g, h\}\} - = \{h, \{f, g\}\} = X_{\{f,g\}}(h), \tag{2.64}$$

therefore

$$[X_f, X_g] = X_{\{f,g\}}. \tag{2.65}$$

This means that Hamiltonian vector fields form a subalgebra of the algebra of vector fields on $\mathcal{M}$, which is isomorphic to the Lie algebra of classical observables.

Using the Poisson bracket, we can introduce two useful concepts.

Definition 2.12. A classical observable I is an *integral of motion* for Hamilton's equations if it is constant along the Hamiltonian phase flow. Equivalently, and in view of (2.61), we have

$$\{H, I\} = 0. \tag{2.66}$$

Definition 2.13. Two classical observables I_1, I_2 are *in involution*, or *Poisson commute*, if

$$\{I_1, I_2\} = 0. \tag{2.67}$$

2.4 Canonical transformations

Definition 2.14. A diffeomorphism of phase space T^*M is called a *canonical transformation* if it preserves the canonical symplectic form ω on T^*M.

Let us focus on the case in which $T^*M = \mathbb{R}^n \times \mathbb{R}^n$. Let us write the diffeomorphism as

$$\mathbf{Q} = \mathbf{Q}(\mathbf{q}, \mathbf{p}), \qquad \mathbf{P} = \mathbf{P}(\mathbf{q}, \mathbf{p}). \tag{2.68}$$

Since

$$\mathrm{d}\mathbf{P} \wedge \mathrm{d}\mathbf{Q} = \mathrm{d}(\mathbf{P} \cdot \mathrm{d}\mathbf{Q}) = \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{q} = \mathrm{d}(\mathbf{p} \cdot \mathrm{d}\mathbf{q}) \tag{2.69}$$

we conclude that

$$\mathrm{d}\left(\mathbf{p} \cdot \mathrm{d}\mathbf{q} - \mathbf{P} \cdot \mathrm{d}\mathbf{Q}\right) = 0. \tag{2.70}$$

But, on $\mathbb{R}^{2n}$, a closed form is exact, therefore there exists a function $S(\mathbf{q}, \mathbf{p})$ on $\mathbb{R}^{2n}$ such that

$$\mathbf{p} \cdot \mathrm{d}\mathbf{q} - \mathbf{P} \cdot \mathrm{d}\mathbf{Q} = \mathrm{d}S. \tag{2.71}$$

Let us obtain a useful condition for the transformation (2.68) to be canonical. First of all, we put together the coordinates in phase space in a single column vector. For the "old" coordinates we have,

$$\boldsymbol{\eta} = \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}, \tag{2.72}$$

while the vector of "new" coordinates will be denoted as

$$\boldsymbol{\zeta} = \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \boldsymbol{\zeta}(\boldsymbol{\eta}). \tag{2.73}$$

The Jacobian of the change of coordinates is

$$\boldsymbol{M} = \frac{\partial \boldsymbol{\zeta}}{\partial \boldsymbol{\eta}} = \begin{pmatrix} \dfrac{\partial \mathbf{Q}}{\partial \mathbf{q}} & \dfrac{\partial \mathbf{Q}}{\partial \mathbf{p}} \\ \dfrac{\partial \mathbf{P}}{\partial \mathbf{q}} & \dfrac{\partial \mathbf{P}}{\partial \mathbf{p}} \end{pmatrix}. \tag{2.74}$$

We will also introduce the symplectic matrix

$$\boldsymbol{J} = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix} \tag{2.75}$$

where $\mathbf{1}$ is the $n \times n$ identity matrix. Since

$$\mathbf{P} \cdot \mathrm{d}\mathbf{Q} - \mathbf{p} \cdot \mathrm{d}\mathbf{q} \tag{2.76}$$

is an exact form, we deduce that

$$2\left(\mathbf{P} \cdot \mathrm{d}\mathbf{Q} - \mathbf{p} \cdot \mathrm{d}\mathbf{q}\right) + \mathrm{d}\left(\mathbf{p} \cdot \mathbf{q} - \mathbf{P} \cdot \mathbf{Q}\right) \tag{2.77}$$

is also exact. We can write it as

$$\mathbf{P} \cdot \mathrm{d}\mathbf{Q} - \mathbf{Q} \cdot \mathrm{d}\mathbf{P} - \mathbf{p} \cdot \mathrm{d}\mathbf{q} + \mathbf{q} \cdot \mathrm{d}\mathbf{p} = \boldsymbol{\eta}^{\mathrm{T}} \boldsymbol{J} \mathrm{d}\boldsymbol{\eta} - \boldsymbol{\zeta}^{\mathrm{T}} \boldsymbol{J} \mathrm{d}\boldsymbol{\zeta} = \left(\boldsymbol{\eta}^{\mathrm{T}} \boldsymbol{J} - \boldsymbol{\zeta}^{\mathrm{T}} \boldsymbol{J} \boldsymbol{M}\right) \mathrm{d}\boldsymbol{\eta}. \tag{2.78}$$

If this is exact, it is also closed, and by acting with d we conclude that

$$\mathrm{d}\boldsymbol{\eta}^{\mathrm{T}} \left(\boldsymbol{J} - \boldsymbol{M}^{\mathrm{T}} \boldsymbol{J} \boldsymbol{M}\right) \mathrm{d}\boldsymbol{\eta} = 0 \tag{2.79}$$

We then obtain the condition

$$\boldsymbol{J} = \boldsymbol{M}^{\mathrm{T}} \boldsymbol{J} \boldsymbol{M} \tag{2.80}$$

which characterizes the canonical transformation as a condition on its Jacobian $\boldsymbol{M}$. By using that $\boldsymbol{J}^2 = -\mathbf{1}_{2n \times 2n}$, we can write (2.80) as

$$\boldsymbol{M} \boldsymbol{J} \boldsymbol{M}^{\mathrm{T}} = \boldsymbol{J}. \tag{2.81}$$

We will now show that *canonical transformations preserve the form of Hamilton's equations*. In terms of $\boldsymbol{\eta}$, Hamilton's equations in the "old" coordinates read

$$\dot{\boldsymbol{\eta}} = \boldsymbol{J} \frac{\partial H}{\partial \boldsymbol{\eta}}. \tag{2.82}$$

We conclude that the EOM for the new coordinates $\boldsymbol{\zeta}$ are

$$\dot{\boldsymbol{\zeta}} = \boldsymbol{M} \dot{\boldsymbol{\eta}} = \boldsymbol{M} \boldsymbol{J} \frac{\partial H}{\partial \boldsymbol{\eta}}, \tag{2.83}$$

Let us write the Hamiltonian in the new coordinates as

$$K(\mathbf{Q}, \mathbf{P}, t) = H(\mathbf{q}, \mathbf{p}, t). \tag{2.84}$$

We can now change variables in the Hamiltonian function,

$$\frac{\partial H}{\partial \boldsymbol{\eta}} = \boldsymbol{M}^{\mathrm{T}} \frac{\partial K}{\partial \boldsymbol{\zeta}}, \tag{2.85}$$

and write

$$\dot{\boldsymbol{\zeta}} = \boldsymbol{M}\boldsymbol{J}\boldsymbol{M}^{\mathrm{T}} \frac{\partial K}{\partial \boldsymbol{\zeta}}. \tag{2.86}$$

Due to (2.81), this agrees with the EOM for $\boldsymbol{\zeta}$,

$$\dot{\boldsymbol{\zeta}} = \boldsymbol{J} \frac{\partial K}{\partial \boldsymbol{\zeta}}. \tag{2.87}$$

We have then obtained the EOM as Hamilton's equation for the new Hamiltonian K.

Let us now find yet another condition for the transformation to be canonical. The entries of the matrix in the l.h.s. of (2.81) can be computed in terms of the Poisson brackets,

$$\begin{aligned} \boldsymbol{M}\boldsymbol{J}\ {}^{t}\boldsymbol{M} &= \begin{pmatrix} \dfrac{\partial \mathbf{Q}}{\partial \mathbf{q}} \left(\dfrac{\partial \mathbf{Q}}{\partial \mathbf{p}}\right)^{\mathrm{T}} - \dfrac{\partial \mathbf{Q}}{\partial \mathbf{p}} \left(\dfrac{\partial \mathbf{Q}}{\partial \mathbf{q}}\right)^{\mathrm{T}} & \dfrac{\partial \mathbf{Q}}{\partial \mathbf{q}} \left(\dfrac{\partial \mathbf{P}}{\partial \mathbf{p}}\right)^{\mathrm{T}} - \dfrac{\partial \mathbf{Q}}{\partial \mathbf{p}} \left(\dfrac{\partial \mathbf{P}}{\partial \mathbf{q}}\right)^{T} \\ \dfrac{\partial \mathbf{P}}{\partial \mathbf{q}} \left(\dfrac{\partial \mathbf{Q}}{\partial \mathbf{p}}\right)^{\mathrm{T}} - \dfrac{\partial \mathbf{P}}{\partial \mathbf{p}} \left(\dfrac{\partial \mathbf{Q}}{\partial \mathbf{q}}\right)^{\mathrm{T}} & \dfrac{\partial \mathbf{P}}{\partial \mathbf{q}} \left(\dfrac{\partial \mathbf{P}}{\partial \mathbf{p}}\right)^{\mathrm{T}} - \dfrac{\partial \mathbf{P}}{\partial \mathbf{p}} \left(\dfrac{\partial \mathbf{P}}{\partial \mathbf{q}}\right)^{\mathrm{T}} \end{pmatrix} \\ &= \begin{pmatrix} \{Q_i, Q_j\} & \{Q_i, P_j\} \\ \{P_i, Q_j\} & \{P_i, P_j\} \end{pmatrix}. \end{aligned} \tag{2.88}$$

Therefore, for a canonical transformation, we have that

$$\{Q_i, Q_j\} = \{P_i, P_j\} = 0 \tag{2.89}$$

and

$$\{Q_i, P_j\} = \delta_{ij}. \tag{2.90}$$

In addition, it follows from this that the Poisson brackets are invariant: if we denote by $\{f, g\}_{\{\mathbf{q},\mathbf{p}\}}$ the bracket computed in the coordinates $\mathbf{q}$, $\mathbf{p}$, and by $\{f, g\}_{\{\mathbf{Q},\mathbf{P}\}}$ the bracket computed in the coordinates $\mathbf{Q}$, $\mathbf{P}$, we have that

$$\begin{aligned} &\{f, g\}_{\{\mathbf{q},\mathbf{p}\}} \\ &= \sum_{k,l=1}^{n} \left(\frac{\partial f}{\partial Q_k} \frac{\partial g}{\partial Q_l} \{Q_k, Q_l\} + \frac{\partial f}{\partial Q_k} \frac{\partial g}{\partial P_l} \{Q_k, P_l\} + \frac{\partial f}{\partial P_k} \frac{\partial g}{\partial Q_l} \{P_k, Q_l\} + \frac{\partial f}{\partial P_k} \frac{\partial g}{\partial P_l} \{P_k, P_l\} \right). \end{aligned} \tag{2.91}$$

By using (2.89) and (2.90), we conclude that

$$\{f, g\}_{\{\mathbf{q},\mathbf{p}\}} = \sum_{k,l=1}^{n} \left(\frac{\partial f}{\partial Q_k} \frac{\partial g}{\partial P_l} - \frac{\partial f}{\partial P_k} \frac{\partial g}{\partial Q_l} \right) = \{f, g\}_{\{\mathbf{Q},\mathbf{P}\}}. \tag{2.92}$$

Let us now assume that, in some open set in phase space, we can take $(\mathbf{q}, \mathbf{Q})$ as independent coordinates, in other words, we assume that the $n \times n$ matrix

$$\frac{\partial \mathbf{Q}}{\partial \mathbf{p}} \tag{2.93}$$

is non-singular in an open set. We can then express S in (2.71) in terms of $(\mathbf{q}, \mathbf{Q})$, and we will denote the resulting function by $S_1(\mathbf{q}, \mathbf{Q})$, i.e.

$$S(\mathbf{q}, \mathbf{p}) = S_1(\mathbf{q}, \mathbf{Q}). \tag{2.94}$$

In view of (2.71), this function satisfies

$$\mathbf{p} = \frac{\partial S_1}{\partial \mathbf{q}}, \qquad -\mathbf{P} = \frac{\partial S_1}{\partial \mathbf{Q}}. \tag{2.95}$$

Definition 2.15. The function $S_1(\mathbf{q}, \mathbf{Q})$ is called the *generating function* of the canonical transformation.

It is possible to obtain different generating functions depending on the choices of parametrization. For example, if we can parametrize phase space in an open subset by the coordinates $(\mathbf{q}, \mathbf{P})$, we introduce the generating function

$$S_2(\mathbf{q}, \mathbf{P}) = S(\mathbf{q}, \mathbf{p}) + \mathbf{P} \cdot \mathbf{Q} = S_1(\mathbf{q}, \mathbf{Q}) + \mathbf{P} \cdot \mathbf{Q} \tag{2.96}$$

which is a Legendre transform of $S_1(\mathbf{q}, \mathbf{Q})$ w.r.t. $\mathbf{Q}$. This function satisfies, again due to (2.71),

$$\mathbf{p} = \frac{\partial S_2}{\partial \mathbf{q}}, \qquad \mathbf{Q} = \frac{\partial S_2}{\partial \mathbf{P}}. \tag{2.97}$$

Canonical transformations give a method of solving Hamilton's equations. Let us suppose that there is a canonical transformation such that the new Hamiltonian $K(\mathbf{P})$ does not depend explicitly on $\mathbf{Q}$. Then, in the new coordinates, we have

$$\dot{\mathbf{P}} = 0, \qquad \dot{\mathbf{Q}} = \frac{\partial K}{\partial \mathbf{P}}. \tag{2.98}$$

This can be integrated immediately,

$$\mathbf{P}(t) = \mathbf{P}(0), \qquad \mathbf{Q}(t) = \mathbf{Q}(0) + \boldsymbol{\omega} t, \tag{2.99}$$

where

$$\boldsymbol{\omega} = \frac{\partial K}{\partial \mathbf{P}}(\mathbf{P}(0)). \tag{2.100}$$

The generating functional $S_2(\mathbf{q}, \mathbf{P})$ of this canonical transformation satisfies

$$H\left(\mathbf{q}, \frac{\partial S_2}{\partial \mathbf{q}}\right) = K(\mathbf{P}). \tag{2.101}$$

This can be regarded as the Hamilton–Jacobi equation for the action

$$S = S_2 - Et, \tag{2.102}$$

where

$$S_2 = \int_0^{\mathbf{q}} \mathbf{p}(\mathbf{q}', \mathbf{P}) \cdot \mathrm{d}\mathbf{q}', \tag{2.103}$$

and we identify $K(\mathbf{P}) = E$. The $\mathbf{P} = (P_1, \cdots, P_n)$ are integrals of motion in involution, i.e.

$$\{P_i, P_j\} = 0. \tag{2.104}$$

Example 2.16. Let us consider the harmonic oscillator

$$H = \frac{1}{2m}\left(p^2 + m^2\omega^2 q^2\right). \tag{2.105}$$

The conserved P is taken to be the energy E. Let us denote $W(q,E) = S_2(q,E)$. Then, we have

$$\frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + \frac{m^2\omega^2}{2}q^2 = E. \tag{2.106}$$

We can solve this as

$$W(q,E) = \int^q \sqrt{2mE - m^2\omega^2 x^2}\,\mathrm{d}x. \tag{2.107}$$

The equation of motion reads

$$\frac{\partial W}{\partial E} = t + c, \tag{2.108}$$

where c is a constant. Note that, since $K = E$, we have $\omega = \partial K/\partial E = 1$. We find

$$\sqrt{\frac{m}{2E}}\int_{q_0}^{q} \frac{\mathrm{d}x}{\sqrt{1 - \frac{m\omega^2}{2E}x^2}} = t - t_0, \tag{2.109}$$

which immediately leads to the standard solution for the harmonic oscillator.

Although we have developed Hamiltonian mechanics on manifolds of the form T^*M, it is possible to generalize the construction to any *symplectic manifold*, i.e. a pair $(\mathcal{M},\omega)$ consisting of a differentiable manifold $\mathcal{M}$ endowed with a closed, non-degenerated two-form ω. The manifold $\mathcal{M}$ is the phase space, and the symplectic form ω generalizes the canonical symplectic form (2.36). A Hamiltonian system is then obtained when we are given a symplectic manifold $(\mathcal{M},\omega)$ and a real-valued function on $\mathcal{M}$, H, called the Hamiltonian. It can be easily seen that all the constructions we have developed have a generalization the more general, symplectic case. For example, the fact that ω is non degenerated induces an identification between one-forms and vector fields on $\mathcal{M}$, i.e. we have a bijection

$$J : E^1(\mathcal{M}) \to \mathcal{X}(\mathcal{M}) \tag{2.110}$$

and the Hamiltonian vector field defining the motion on $\mathcal{M}$ is given by

$$X_H = J(\mathrm{d}H). \tag{2.111}$$

2.5 Integrable systems

Definition 2.17. An *integrable system* (in the sense of Liouville) is a mechanical system whose phase space $\mathcal{M}$ is of dimension $2n$ and has n independent integrals of motion $F_1 = H, \cdots, F_n$ in involution,

$$\{F_i, F_j\} = 0, \qquad i,j = 1,\cdots,n. \tag{2.112}$$

The condition of independence means that the n one-forms $\mathrm{d}F_i$ are linearly independent at almost all points of $\mathcal{M}$.

In an integrable system, the EOM can be completely solved in terms of quadratures. To see this, we will first study the level set of the functions F_i:

$$M_{\mathbf{f}} = \{\boldsymbol{F}(\mathbf{q},\mathbf{p}) = \mathbf{f}\}. \tag{2.113}$$

Lemma 2.18. *On the n-dimensional manifold $M_{\mathbf{f}}$, there exist n tangent vector fields which commute with another and are linearly independent. In addition, we have that*

$$\omega\big|_{M_{\mathbf{f}}} = 0, \tag{2.114}$$

i.e. $M_{\mathbf{f}}$ is a Lagrangian submanifold.

Proof: Let us denote by X_{F_i} the Hamiltonian vector fields associated to the F_i, which are defined by

$$\mathrm{d}F_i(Y) = \omega(Y, X_{F_i}). \tag{2.115}$$

These vectors are independent, since the $\mathrm{d}F_i$ are independent. They are tangent to $M_{\mathbf{f}}$, since

$$\mathrm{d}F_j(X_{F_i}) = X_{F_i}(F_j) = -\{F_i, F_j\} = 0. \tag{2.116}$$

They also commute, since

$$[X_{F_i}, X_{F_j}] = X_{\{F_i, F_j\}} = 0. \tag{2.117}$$

Finally, we have that

$$\omega(X_{F_i}, X_{F_j}) = \{F_j, F_i\} = 0. \tag{2.118}$$

□

It can be shown that, if $M_{\mathbf{f}}$ is compact and connected, it is diffeomorphic to an n-dimensional torus, i.e.

$$M_{\mathbf{f}} \simeq \mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1. \tag{2.119}$$

This is called the *Liouville torus*. In particular, it has n non-trivial one-cycles $\gamma_1, \cdots, \gamma_n$. We will now find a canonical transformation to a set of new coordinates

$$(\mathbf{q}, \mathbf{p}) \to (\boldsymbol{I}, \boldsymbol{\phi}). \tag{2.120}$$

The new coordinates are called *action-angle variables*. The action variables are defined by

$$I_j(\mathbf{f}) = \frac{1}{2\pi} \oint_{\gamma_j} \mathbf{p} \cdot \mathrm{d}\mathbf{q}. \tag{2.121}$$

Note that, since ω vanishes on $M_{\mathbf{f}}$, these integrals only depend on the homology class of γ_i, by Stokes' theorem. They are functions of the integrals of motion $\mathbf{f}$. Let us now construct the generating function S of this canonical transformation. Let $\mathbf{x}_0$ be a point in $M_{\mathbf{f}}$, with coordinates $\mathbf{q}_0$, and let suppose that on a simply-connected neighborhood of $\mathbf{x}_0$ we can solve for p_i as:

$$\mathbf{p} = \mathbf{p}(\mathbf{q}, \boldsymbol{I}). \tag{2.122}$$

Let us consider the function

$$S_2(\mathbf{q}, \boldsymbol{I}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{p}(\mathbf{q}', \boldsymbol{I}) \cdot \mathrm{d}\mathbf{q}', \tag{2.123}$$

where the integration path is contained in $M_{\mathbf{f}}$. Since ω vanishes on $M_{\mathbf{f}}$, this integral does not depend on the path chosen. Therefore, $S_2(\mathbf{q}, \boldsymbol{I})$ is a well-defined function in a neigborhood of $\mathbf{x}_0$. In addition, we have that

$$\mathbf{p} = \frac{\partial S_2}{\partial \mathbf{q}}, \qquad \boldsymbol{\phi} = \frac{\partial S_2}{\partial \boldsymbol{I}}. \tag{2.124}$$

Note that the variables ϕ_j are multivalued, with period 2π, i.e. they are angle variables. To see this, note that, as we go around one cycle γ_i in the Liouville torus, the function S_2 changes as

$$\Delta_i S_2 = 2\pi I_i, \tag{2.125}$$

and this induces a change in ϕ_j given by

$$\Delta_i \phi_j = \frac{\partial}{\partial I_j} \Delta_i S_2 = 2\pi \delta_{ij}. \tag{2.126}$$

In addition, the EOM can be integrated immediately. The Hamiltonian H only depends on the $\boldsymbol{I}$, therefore Hamilton's equations read

$$\dot{\boldsymbol{I}} = 0, \qquad \dot{\boldsymbol{\phi}} = \frac{\partial H}{\partial \boldsymbol{I}} = \boldsymbol{\omega}, \tag{2.127}$$

which can be immediately integrated to

$$\boldsymbol{\phi}(t) = \boldsymbol{\omega} t + \boldsymbol{\phi}(0). \tag{2.128}$$

We note that, in order to integrate the EOM, it is also possible to consider a canonical transformation in which the generating functional is regarded as a function of the original integrals of motion, $S_2(\mathbf{q}, \mathbf{f})$. In this case, we still have a canonical transformation

$$(\mathbf{q}, \mathbf{p}) \to (\mathbf{f}, \boldsymbol{\psi}), \tag{2.129}$$

where

$$\boldsymbol{\psi} = \frac{\partial S_2}{\partial \mathbf{f}}. \tag{2.130}$$

Like before, the EOM can be integrated in closed form as

$$\boldsymbol{\psi}(t) = \frac{\partial H}{\partial \mathbf{f}} t + \boldsymbol{\psi}(0). \tag{2.131}$$

Example 2.19. *The Kepler problem.* Let us consider a central potential $V(r)$ in three dimensions (Kepler problem). We use spherical coordinates. The conserved quantities are the Hamiltonian,

$$H = \frac{1}{2}\left(p_r^2 + \frac{1}{r^2} p_\theta^2 + \frac{1}{r^2 \sin^2\theta} p_\phi^2\right) + V(r), \tag{2.132}$$

the total angular momentum

$$J^2 = \sum_{i=1}^{3} J_i^2 = p_\theta^2 + \frac{1}{\sin^2\theta} p_\phi^2 \tag{2.133}$$

and the third component of the angular momentum,

$$J_3 = p_\phi. \tag{2.134}$$

Let us consider the surface in phase space where these conserved quantities take fixed values, $M_{\mathbf{f}}$. We can solve for the momenta in terms of H, J^2 and J_3,

$$\begin{aligned} p_r &= \sqrt{2\left(H - V(r)\right) - \frac{J^2}{r^2}}, \\ p_\theta &= \sqrt{J^2 - \frac{J_3^2}{\sin^2\theta}}, \\ p_\phi &= J_3. \end{aligned} \tag{2.135}$$

We can regard the conserved quantities as our new momenta after a canonical transformation, with Hamilton–Jacobi function given as in (2.103),

$$S_2 = \int^r \sqrt{2\left(H - V(r')\right) - \frac{J^2}{(r')^2}}\mathrm{d}r' + \int^\theta \sqrt{J^2 - \frac{J_3^2}{\sin^2\theta'}}\mathrm{d}\theta' + \int^\phi J_3 \mathrm{d}\phi'. \tag{2.136}$$

We have now the new variables

$$\psi_H = \frac{\partial S_2}{\partial H}, \qquad \psi_{J^2} = \frac{\partial S_2}{\partial J^2}, \qquad \psi_{J_3} = \frac{\partial S_2}{\partial J_3}. \tag{2.137}$$

It follows from (2.99) and (2.100) that ψ_{J^2} and ψ_{J_3} are constant, while ψ_H evolves in time according to

$$\psi_H(t) = t + c, \tag{2.138}$$

where c is a constant. We deduce that

$$t - t_0 = \int_{r_0}^r \frac{\mathrm{d}r}{\sqrt{2\left(H - V(r)\right) - \frac{J^2}{r^2}}}, \tag{2.139}$$

which is the standard formula for the Kepler motion. Of course, we can also solve for the motion in terms of action-angle variables (see exercises). □

Example 2.20. A much less trivial example of an integrable model is the *periodic Toda chain*. This consists of $n+1$ particles with positions q_i and momenta p_i. The EOM are given by

$$\dot{q}_i = p_i, \qquad \dot{p}_i = \mathrm{e}^{q_{i-1}-q_i} - \mathrm{e}^{q_i - q_{i+1}}, \qquad i = 1, \cdots, n+1. \tag{2.140}$$

These follow from the Hamiltonian

$$H = \frac{1}{2}\sum_{i=1}^{n+1} p_i^2 + \sum_{i=1}^{n+1} \mathrm{e}^{q_i - q_{i+1}}, \tag{2.141}$$

where we have normalized $m = 1$ for all the particles. We set $(q_{n+1+j}, p_{n+1+j}) = (q_j, p_j)$, which enforces the periodicity condition. It turns out that this system has $n+1$ conserved quantities in involution. The first two of them are of course the total momentum,

$$P = \sum_{i=1}^{n+1} p_i \tag{2.142}$$

and the Hamiltonian itself. Constructing the others can be done by using the technique of Lax matrices. There are various versions of this, but the most useful one was introduced by Sklyanin [8]. We first introduce the Lax matrix, depending on a parameter μ, as follows:

$$L_i(\mu) = \begin{pmatrix} 0 & \mathrm{e}^{q_i} \\ -\mathrm{e}^{-q_i} & \mu - p_i \end{pmatrix}, \qquad i = 1, \cdots, n+1. \tag{2.143}$$

μ is sometimes called the *spectral parameter*. Note that this matrix only involves the coordinates and momenta of the i-th particle. Out of these matrix, we form the *monodromy matrix*

$$T(\mu) = L_{n+1}(\mu)\cdots L_1(\mu) = \begin{pmatrix} A(\mu) & B(\mu) \\ C(\mu) & D(\mu) \end{pmatrix}. \tag{2.144}$$

Let us define $t(\mu)$ by

$$2t(\mu) = \operatorname{Tr} T(\mu) = A(\mu) + D(\mu). \tag{2.145}$$

From its form, it is clear that it is a polynomial of degree $n+1$ in μ, and we will write it as

$$2t(\mu) = \sum_{k=0}^{n+1} (-1)^k \mu^{n+1-k} H_k. \tag{2.146}$$

It is relatively straightforward to show that $H_0 = 1$ and that

$$H_1 = P, \qquad H_2 = \frac{P^2}{2} - H, \tag{2.147}$$

where H is the Hamiltonian. We will now show the $n+1$ functions on phase space $H_1, \cdots, H_{n+1}$ are in involution. To do this, we need some technology which was developed in the modern study of integrable systems. Let $M(\mu)$ be a two times two matrix of the form

$$M(\mu) = \sum_{i,j} M(\mu)_{ij} E_{ij} = \begin{pmatrix} a(\mu) & b(\mu) \\ c(\mu) & d(\mu) \end{pmatrix}, \tag{2.148}$$

which acts on $\mathbb{C}^2$. Here, E_{ij} is the canonical basis for square matrices,

$$(E_{ij})_{kl} = \delta_{ik}\delta_{jl}. \tag{2.149}$$

We now introduce the matrices acting on $\mathbb{C}^2 \otimes \mathbb{C}^2$,

$$\begin{aligned} M^1(\mu) = M(\mu) \otimes I &= \sum_{i,j} M(\mu)_{ij} \left(E_{ij} \otimes I\right) = \begin{pmatrix} a(\mu) & 0 & b(\mu) & 0 \\ 0 & a(\mu) & 0 & b(\mu) \\ c(\mu) & 0 & d(\mu) & 0 \\ 0 & c(\mu) & 0 & d(\mu) \end{pmatrix}, \\ M^2(\mu) = I \otimes M(\mu) &= \sum_{i,j} M(\mu)_{ij} \left(I \otimes E_{ij}\right) = \begin{pmatrix} a(\mu) & b(\mu) & 0 & 0 \\ c(\mu) & d(\mu) & 0 & 0 \\ 0 & 0 & a(\mu) & b(\mu) \\ 0 & 0 & c(\mu) & d(\mu) \end{pmatrix}. \end{aligned} \tag{2.150}$$

Their product is given by

$$\begin{aligned} M^1(\mu) M^2(\mu') &= \sum_{ij,kl} M_{ij}(\mu) M_{kl}(\mu') E_{ij} \otimes E_{kl} \\ &= \begin{pmatrix} a(\mu)a(\mu') & a(\mu)b(\mu') & b(\mu)a(\mu') & b(\mu)b(\mu') \\ a(\mu)c(\mu') & a(\mu)d(\mu') & b(\mu)c(\mu') & b(\mu)d(\mu') \\ c(\mu)a(\mu') & c(\mu)b(\mu') & d(\mu)a(\mu') & d(\mu)b(\mu') \\ c(\mu)c(\mu') & c(\mu)d(\mu') & d(\mu)c(\mu') & d(\mu)d(\mu') \end{pmatrix}. \end{aligned} \tag{2.151}$$

If the entries of $M(\mu)$ are functions on phase space, we also have the matrix of Poisson brackets,

$$\begin{aligned} \{M^1(\mu), M^2(\mu')\} &= \sum_{ij,kl} \{M_{ij}(\mu), M_{kl}(\mu')\} E_{ij} \otimes E_{kl} \\ &= \begin{pmatrix} \{a(\mu), a(\mu')\} & \{a(\mu), b(\mu')\} & \{b(\mu), a(\mu')\} & \{b(\mu), b(\mu')\} \\ \{a(\mu), c(\mu')\} & \{a(\mu), d(\mu')\} & \{b(\mu), c(\mu')\} & \{b(\mu), d(\mu')\} \\ \{c(\mu), a(\mu')\} & \{c(\mu), b(\mu')\} & \{d(\mu), a(\mu')\} & \{d(\mu), b(\mu')\} \\ \{c(\mu), c(\mu')\} & \{c(\mu), d(\mu')\} & \{d(\mu), c(\mu')\} & \{d(\mu), d(\mu')\} \end{pmatrix}. \end{aligned} \tag{2.152}$$

A direct calculation shows that

$$\{L_i^1(\mu), L_i^2(\mu')\} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mathrm{e}^{q_i} \\ 0 & 0 & 0 & \mathrm{e}^{q_i} \\ 0 & -\mathrm{e}^{-q_i} & \mathrm{e}^{-q_i} & 0 \end{pmatrix} \tag{2.153}$$

Let us consider the matrix

$$C_{12} = \sum_{i,j} E_{ij} \otimes E_{ji}, \tag{2.154}$$

which is explicitly given by

$$C_{12} = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,0\,1\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix} \tag{2.155}$$

We have, for a general matrix $M(\mu)$, that

$$\begin{aligned} &\left[C_{12}, M^1(\mu)M^2(\mu')\right] \\ &= \begin{pmatrix} 0 & a(\mu)b\,(\mu') - a\,(\mu')\,b(\mu) & a\,(\mu')\,b(\mu) - a(\mu)b\,(\mu') & 0 \\ a\,(\mu')\,c(\mu) - a(\mu)c\,(\mu') & b\,(\mu')\,c(\mu) - b(\mu)c\,(\mu') & a\,(\mu')\,d(\mu) - a(\mu)d\,(\mu') & b\,(\mu')\,d(\mu) - b(\mu)d\,(\mu') \\ a(\mu)c\,(\mu') - a\,(\mu')\,c(\mu) & a(\mu)d\,(\mu') - a\,(\mu')\,d(\mu) & b(\mu)c\,(\mu') - b\,(\mu')\,c(\mu) & b(\mu)d\,(\mu') - b\,(\mu')\,d(\mu) \\ 0 & c(\mu)d\,(\mu') - c\,(\mu')\,d(\mu) & c\,(\mu')\,d(\mu) - c(\mu)d\,(\mu') & 0 \end{pmatrix} \end{aligned} \tag{2.156}$$

In the case of $L_i(\mu)$, we have

$$\begin{aligned} &\left[C_{12}, L_i^1(\mu)L_i^2(\mu')\right] \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{e}^{q_i}(\mu - \mu') \\ 0 & 0 & 0 & -\mathrm{e}^{q_i}(\mu - \mu') \\ 0 & \mathrm{e}^{-q_i}(\mu - \mu') & -\mathrm{e}^{-q_i}(\mu - \mu') & 0 \end{pmatrix} \end{aligned} \tag{2.157}$$

We conclude that

$$\{L_i^1(\mu), L_i^2(\mu')\} = \left[L_i^1(\mu), L_i^2(\mu'), r_{12}(\mu - \mu')\right], \tag{2.158}$$

where

$$r_{12}(\mu - \mu') = \frac{C_{12}}{\mu - \mu'} \tag{2.159}$$

is called the *classical r-matrix*. The property (2.158) also holds for products of the matrices L_i. Indeed, we have

$$\begin{aligned} \{L_i^1(\mu)L_j^1(\mu), L_i^2(\mu')L_j^2(\mu')\} &= LL_i^1(\mu)L_i^2(\mu')\{L_j^1(\mu), L_j^2(\mu')\} + \{L_i^1(\mu), L_i^2(\mu')\}L_j^1(\mu)L_j^2(\mu') \\ &= L_i^1(\mu)L_i^2(\mu')\left[r_{12}(\mu - \mu'), L_j^1(\mu), L_j^2(\mu')\right] \\ &+ \left[r_{12}(\mu - \mu'), L_i^1(\mu), L_i^2(\mu')\right] L_j^1(\mu)L_j^2(\mu') \\ &= \left[r_{12}(\mu - \mu'), L_i^1(\mu), L_i^2(\mu')L_j^1(\mu)L_j^2(\mu')\right], \end{aligned} \tag{2.160}$$

where we used that

$$A_1[R, A_2] + [R, A_1]A_2 = [R, A_1A_2] \tag{2.161}$$

for the standard matrix commutator. In the first line we used that the Poisson bracket of matrices with different indices is zero, since the matrices $L_i(\mu)$ only depend on the variables q_i, p_i. By induction, we conclude that (2.158) also holds for an arbitrary product of matrices L_i with different indices, in particular for the monodromy matrix $T(\mu)$, and we obtain

$$\{T^1(\mu), T^2(\mu')\} = \left[T^1(\mu), T^2(\mu'), r_{12}(\mu - \mu'),\right]. \tag{2.162}$$

This implies the following properties:

$$\{A(\mu), A(\mu')\} = \{B(\mu), B(\mu')\} = \{C(\mu), C(\mu')\} = \{D(\mu), D(\mu')\} = 0, \tag{2.163}$$

as well as

$$\{A(\mu), D(\mu')\} = \frac{B(\mu)C(\mu') - C(\mu)B(\mu')}{\mu - \mu'}. \tag{2.164}$$

It follows from the above results that

$$\{t(\mu), t(\mu')\} = 0, \tag{2.165}$$

for arbitrary μ, μ'. Therefore, the coefficients of $t(\mu)$ are in involution, i.e.

$$\{H_k, H_l\} = 0, \qquad k = 1, \cdots, n+1. \tag{2.166}$$

In particular, it follows that

$$\{H, H_k\} = 0, \qquad k = 3, \cdots, n+1. \tag{2.167}$$

so all these quantities are integrals of motion. It is not difficult to see that the H_k are independent (they involve different independent functions of the momenta, for example). Therefore, the periodic Toda chain is an integrable system in the sense of Liouville.

3 Basic aspects of Quantum Mechanics

3.1 Introduction: position and momentum in QM

In QM, states are described by elements in a Hilbert space $\mathcal{H}$, while physical quantities, like position, momentum and energy, will be operators acting on $\mathcal{H}$. We will denote the product of two elements in the Hilbert space ϕ, ψ by

$$\langle \phi, \psi \rangle \in \mathbb{C}, \tag{3.1}$$

which is linear in the second factor, i.e.

$$\langle \phi, \lambda\psi \rangle = \lambda \langle \phi, \psi \rangle, \qquad \langle \lambda\phi, \psi \rangle = \overline{\lambda} \langle \phi, \psi \rangle. \tag{3.2}$$

Most of the interesting operators appearing in QM are unbounded. Typically, they are not defined on the whole of $\mathcal{H}$. To understand why, let us consider $\mathcal{H} = L^2(\mathbb{R})$. The function

$$\varphi(q) = \frac{1}{\sqrt{q^2 + 1}} \tag{3.3}$$

belongs to $L^2(\mathbb{R})$. Let us now consider the operator q defined by multiplication by q,

$$\mathsf{q}\psi(q) = q\psi(q). \tag{3.4}$$

As we will see, this is the *position operator* in QM. Clearly, the function

$$\mathsf{q}\varphi(q) \tag{3.5}$$

obtained by acting with q on $\varphi(q)$, is not in $L^2(\mathbb{R})$.

Definition 3.1. An unbounded operator A on the Hilbert space $\mathcal{H}$ is a linear map from a dense subspace $\mathrm{Dom}(A) \subset \mathcal{H}$ to $\mathcal{H}$.

Definition 3.2. A linear map A on $\mathcal{H}$ is *symmetric* if

$$\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle, \tag{3.6}$$

for all ϕ, $\psi \in \mathrm{Dom}(A)$.

For bounded operators, being symmetric is equivalent to being self-adjoint, but this is not the case for unbounded operators. Symmetric operators satisfy some important and elementary properties:

1. For all $\phi \in \mathrm{Dom}(A)$, $\langle \phi, A\phi \rangle \in \mathbb{R}$. The proof is immediate, since

$$\overline{\langle \phi, A\phi \rangle} = \langle A\phi, \phi \rangle = \langle \phi, A\phi \rangle. \tag{3.7}$$

2. Suppose that λ is an eigenvalue for A, meaning that

$$A\psi = \lambda\psi, \qquad \psi \in \mathrm{Dom}(A), \qquad \psi \neq 0. \tag{3.8}$$

   Then, $\lambda \in \mathbb{R}$. This is because

$$\lambda\langle \psi, \psi \rangle = \langle \psi, A\psi \rangle = \langle A\psi, \psi \rangle = \overline{\lambda}\langle \psi, \psi \rangle. \tag{3.9}$$

3. Let us suppose that $\psi_1, \psi_2 \in \mathrm{Dom}(A)$ satisfy

$$A\psi_1 = \lambda_1\psi_1, \qquad A\psi_2 = \lambda_2\psi_2, \tag{3.10}$$

   with $\lambda_1 \neq \lambda_2$. Then, ψ_1 and ψ_2 are orthogonal, i.e.

$$\langle \psi_2, \psi_1 \rangle = 0. \tag{3.11}$$

   To see this, note that

$$\langle \psi_2, A\psi_1 \rangle = \langle \psi_2, \lambda_1\psi_1 \rangle = \lambda_1\langle \psi_2, \psi_1 \rangle. \tag{3.12}$$

   On the other hand,

$$\langle \psi_2, A\psi_1 \rangle = \langle A\psi_2, \psi_1 \rangle = \lambda_2\langle \psi_2, \psi_1 \rangle. \tag{3.13}$$

   We deduce that

$$(\lambda_2 - \lambda_1)\,\langle \psi_2, \psi_1 \rangle = 0, \tag{3.14}$$

   and since $\lambda_2 \neq \lambda_1$ by assumption, the orthogonality property (3.11) follows.

These properties are important in QM. If A is an operator which represents a physical quantity, $\langle \psi, A\psi \rangle$ will be interpreted as the average value of A in the state ψ, while the eigenvalues λ will be the possible values that A can take. Both quantities should be real, since physical quantities lead to real numbers. Therefore, we will require the operators appearing in QM to be symmetric, since in this case these reality requirements are guaranteed by the results above on symmetric operators. In fact, more is needed: we want our operators to be self-adjoint or essentially self-adjoint, but we will not discuss these more technical aspects here. A detailed analysis can be found in [6].

Let us now discuss the most elementary operators appearing in QM: the position and momentum operators. Let us first consider a particle moving on the real line $\mathbb{R}$. In QM, we can not, in general, specify the precise position of such a particle, but we can only provide a *probability distribution* for its position. This means that, if we prepare many physical systems consisting of a quantum particle in the same state, the measurements of its position will lead to random outcomes, distributed according to a precise pattern typical of the state. This state is specified by a *wave function*,

$$\psi : \mathbb{R} \to \mathbb{C}, \tag{3.15}$$

and the probability distribution is given by

$$\rho(q) = |\psi(q)|^2 . \tag{3.16}$$

If we want the probability distribution to be normalized, we need

$$\int_{\mathbb{R}} \rho(q) \mathrm{d}q = \int_{\mathbb{R}} |\psi(q)|^2 \, \mathrm{d}q = 1. \tag{3.17}$$

This means that the wavefunction must be a unit element of $\mathcal{H} = L^2(\mathbb{R})$, which is the relevant Hilbert space. It follows that the average value of the position in such a state is given by

$$E(q) = \int_{\mathbb{R}} q|\psi(q)|^2 \mathrm{d}q, \tag{3.18}$$

provided the integral converges. We want to express this in terms of an operator and the inner product in $\mathcal{H}$. To do this, we recall the position operator introduced in (3.4), and we write (3.18) as

$$E(q) = \langle \mathsf{q} \rangle_\psi = \langle \psi, \mathsf{q}\psi \rangle, \tag{3.19}$$

where the inner product is the standard one in $\mathcal{H} = L^2(\mathbb{R})$:

$$\langle \phi, \psi \rangle = \int_{\mathbb{R}} \overline{\phi(q)} \psi(q) \mathrm{d}q. \tag{3.20}$$

Note that q is an unbounded, symmetric operator. As we discussed above, $\mathsf{q}\psi$ might fail to be in $L^2(\mathbb{R})$, therefore the average position is not necessarily well-defined, unless $\psi \in \mathrm{Dom}(\mathsf{q})$. Similarly, if ψ belongs to the appropriate domain,

$$E(q^m) = \langle \mathsf{q}^m \rangle_\psi = \langle \psi, \mathsf{q}^m \psi \rangle = \int_{\mathbb{R}} q^m |\psi(q)|^2 \mathrm{d}q. \tag{3.21}$$

The wavefunction $\psi(q)$ encodes information about the position of the particle, but what about its momentum? It turns out that, according to de Broglie's hypothesis, the momentum is roughly encoded in the oscillations of $\psi(q)$. We define the *momentum operator* p as

$$\mathsf{p}\psi(q) = -\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}q} \psi(q). \tag{3.22}$$

In this equation, $\hbar$ is called *Planck's constant*. It is a fundamental constant of nature. It sets the scale at which quantum phenomena are important. Note that it has the dimensions of an action, i.e. momentum times distance (or energy times time). As we mentioned above, although $\psi(q) \in L^2(\mathbb{R})$, its derivative is not necessarily well-defined, nor in $L^2(\mathbb{R})$. The first problem can be fixed by considering derivatives in the sense of distributions, the second problem can be fixed by restricting p to an appropriate domain $\mathrm{Dom}(\mathsf{p})$, but we will not explore this in detail. When needed, we will restrict wave functions to belong to a sufficiently "nice" class of functions, like Schwartz functions. We have the following

Proposition 3.3. For all sufficiently nice functions ψ, ϕ in $L^2(\mathbb{R})$, the operator p is symmetric, i.e.

$$\langle \phi, \mathsf{p}\psi \rangle = \langle \mathsf{p}\phi, \psi \rangle. \tag{3.23}$$

Proof: as for sufficiently nice conditions, we will require that ϕ, ψ are continuously differentiable, that they tend to zero as $|q| \to \infty$, and that their derivatives ϕ', ψ' are in $L^2(\mathbb{R})$. Then, if $A > 0$, we have, after integrating by parts,

$$-\mathrm{i}\hbar \int_{-A}^{A} \overline{\phi(q)}\psi'(q)\mathrm{d}q = -\mathrm{i}\hbar\overline{\phi(q)}\psi(q)\Big|_{-A}^{A} + \mathrm{i}\hbar \int_{-A}^{A} \overline{\phi'(q)}\psi(q)\mathrm{d}q \tag{3.24}$$

Let us now take $A \to \infty$, and use the hypothesis on ψ, ϕ. We conclude that

$$\langle \phi, \mathsf{p}\psi \rangle = \int_{-\infty}^{\infty} \overline{\phi(q)} \left(-\mathrm{i}\hbar\psi'(q)\right) \mathrm{d}q = \mathrm{i}\hbar \int_{-A}^{A} \overline{(-\mathrm{i}\hbar\phi'(q))}\psi(q)\mathrm{d}q = \langle \mathsf{p}\phi, \psi \rangle. \tag{3.25}$$

□

It follows from the above description that the average value of the momentum in a state ψ should be given by

$$\langle \mathsf{p} \rangle_\psi = \left\langle \psi, -\mathrm{i}\hbar\psi'(q) \right\rangle, \tag{3.26}$$

provided we restrict the function ψ to an appropriate domain.

A natural question is: if the momentum of a particle described by ψ is also a random variable, what is its probability distribution? It turns out that the answer can be expressed in terms of the Fourier transform of $\psi(q)$:

$$\widehat{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}pq/\hbar}\psi(q)\mathrm{d}q, \tag{3.27}$$

and we have that

$$\rho(p) = \left|\widehat{\psi}(p)\right|^2. \tag{3.28}$$

Note that the standard Fourier transform of $\psi(q)$ is given by

$$\mathcal{F}(\psi)(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kq}\psi(q)\mathrm{d}q = \hbar^{1/2}\widehat{\psi}\left(\hbar k\right). \tag{3.29}$$

It can be easily seen that the distribution $\rho(p)$ in (3.28) is properly normalized. Indeed, as a consequence of Plancherel's theorem, we have

$$\int_{-\infty}^{\infty} |\psi(q)|^2 \,\mathrm{d}q = \int_{-\infty}^{\infty} |\mathcal{F}(\psi)(k)|^2 \,\mathrm{d}k = \int_{-\infty}^{\infty} \left|\widehat{\psi}(p)\right|^2 \mathrm{d}p = 1, \tag{3.30}$$

where we changed variables $p = \hbar k$. In addition, we have the following

Proposition 3.4. If ψ belongs to an appropriate domain in $L^2(\mathbb{R})$, we have

$$E(p^m) = \langle \psi, \mathsf{p}^m\psi \rangle = \int_{-\infty}^{\infty} p^m \left|\widehat{\psi}(p)\right|^2 \mathrm{d}p, \tag{3.31}$$

for all positive integers m.

Proof: We can suppose for example that ψ belongs to the Schwartz space $\mathcal{S}(\mathbb{R})$. In this case,

$$\mathsf{p}^m \psi(q) = (-\mathrm{i}\hbar)^m \, \psi^{(m)}(q). \tag{3.32}$$

The Fourier transform of $\psi^{(m)}(q)$ is given by

$$\mathcal{F}(\psi^{(m)})(k) = (\mathrm{i}k)^m \, \mathcal{F}(\psi)(k), \tag{3.33}$$

We now use the fact that the inner product in $L^2(\mathbb{R})$ is preserved by Fourier transform, therefore

$$\begin{aligned} \langle \psi, \mathsf{p}^m \psi \rangle = \langle \mathcal{F}(\psi), \mathcal{F}(\mathsf{p}^m \psi) \rangle &= \int_{-\infty}^{\infty} \overline{\mathcal{F}(\psi)(k)} \, (-\mathrm{i}\hbar)^m \, (\mathrm{i}k)^m \, \mathcal{F}(\psi)(k) \mathrm{d}k \\ &= \int_{-\infty}^{\infty} p^m \left| \widehat{\psi}(p) \right|^2 \mathrm{d}p \end{aligned} \tag{3.34}$$

□

As in the case of the average of the position operator and its powers, in (3.21), we define the average of the m-th power of the momentum in the state ψ as

$$\langle \mathsf{p}^m \rangle_\psi = \langle \psi, \mathsf{p}^m \psi \rangle. \tag{3.35}$$

The above proposition shows that $\rho(p)$ has indeed the properties expected for the distribution of momentum in the state ψ. We can now derive one of the basic properties of the position and momentum operators in QM.

Proposition 3.5. The operators q and p do not commute, but satisfy the relation

$$[\mathsf{q}, \mathsf{p}] = \mathrm{i}\hbar. \tag{3.36}$$

Proof: this is a simple calculation. If we act on a function belonging to the appropriate domain $\mathrm{Dom}(\mathsf{q}) \cap \mathrm{Dom}(\mathsf{p})$, we find

$$(\mathsf{qp} - \mathsf{pq}) \, \psi(q) = -\mathrm{i}\hbar q \psi'(q) + \mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}q} \left(q \psi(q) \right) = \mathrm{i}\hbar \psi(q). \tag{3.37}$$

□

So far we have focused on one-dimensional systems, with a single position and momentum. If our phase space is $\mathbb{R}^{2n}$, with coordinates and momenta q_i, p_i, $i = 1, \cdots, n$, the corresponding Hilbert space will be $L^2(\mathbb{R}^n)$. A wave function in this space will be denoted by $\psi(\mathbf{q})$. The position and momentum operators q_i, p_i are defined by

$$\mathsf{q}_i \psi(\mathbf{q}) = q_i \psi(\mathbf{q}), \qquad \mathsf{p}_i \psi(\mathbf{q}) = -\mathrm{i}\hbar \frac{\partial}{\partial q_i} \psi(\mathbf{q}). \tag{3.38}$$

They satisfy the commutation relations

$$[\mathsf{q}_i, \mathsf{q}_j] = 0, \qquad [\mathsf{p}_i, \mathsf{p}_j] = 0, \qquad [\mathsf{q}_i, \mathsf{p}_j] = \mathrm{i}\hbar \delta_{ij}, \qquad i, j = 1, \cdots, n. \tag{3.39}$$

The wavefunction in the momentum representation, $\widehat{\psi}(\mathbf{p})$, is defined by a straightforward multi-dimensional generalization of the Fourier transform.

After this discussion, we can formalize in more detail the kinematic and dynamics aspects of QM.

3.2 Axioms of QM

Axiom 1: *The state of a quantum system os represented by a unit vector in an appropriate Hilbert space $\mathcal{H}$. Two unit vectors satisfying $\psi_1 = c\psi_2$ represent the same physical state.*

Remark 3.6. This axiom concerns only the "pure states" of the theory. There is a more general notion of state described by a density matrix, but we will not deal with this in this course.

Axiom 2: *To each real-valued function f on classical phase space, there is associated a self-adjoint operator f on $\mathcal{H}$.*

Remark 3.7. Most operators f appearing in QM are unbounded, therefore being self-adjoint and being symmetric are *not* equivalent.

Given a classical system, defined by a phase space $\mathcal{M}$ and an algebra of classical observables $C^\infty(\mathcal{M})$, its *quantization* should lead to a Hilbert space $\mathcal{H}$ and to a correspondence from $C^\infty(\mathcal{M})$ to the set of self-adjoint operators on $\mathcal{H}$. There are various frameworks to do this. In the case that $\mathcal{M} = \mathbb{R}^n \times \mathbb{R}^n$, the Hilbert space is taken to be $\mathcal{H} = L^2(\mathbb{R}^n)$, and one can use the so-called Weyl–Wigner quantization to define a map from a set of functions in phase space, to the set of self-adjoint operators on $\mathcal{H}$, as we will do later on.

Example 3.8. The standard form of the classical Hamiltonian in one-dimensional motion is of course

$$H = \frac{p^2}{2m} + V(q). \tag{3.40}$$

We can promote H to an operator H as

$$\mathsf{H} = \frac{\mathsf{p}^2}{2m} + V(\mathsf{q}). \tag{3.41}$$

When acting on wave functions $\psi(q) \in L^2(\mathbb{R})$, we have

$$\mathsf{H}\psi(q) = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial q^2} + V(q)\psi(q), \tag{3.42}$$

Axiom 3: *If a quantum system is in a state described by a unit vector $\psi \in \mathcal{H}$, the probability distribution for the measurement of some observable f satisfies*

$$E(f^m) = \langle \mathsf{f}^m \rangle_\psi = \langle \psi, \mathsf{f}^m \psi \rangle. \tag{3.43}$$

Remark 3.9. Note that, since f is self-adjoint, therefore symmetric, these quantities are real. Note also that, if ψ is an eigenvector of f with eigenvalue λ, i.e. if

$$\mathsf{f}\psi = \lambda\psi, \tag{3.44}$$

then

$$E(f^m) = \lambda^m, \tag{3.45}$$

for all positive integers m. This means that the probability distribution for f is the one in which it has the definite value λ, with probability one.

Example 3.10. Let us suppose that f has an orthonormal basis $\{e_j\}$ of eigenvectors with real eigenvalues λ_j. Suppose as well that

$$\psi = \sum_{j=1}^{\infty} a_j e_j \tag{3.46}$$

is a unit vector. Then, if we measure f in the state ψ, the possible outcomes will be the values λ_j, with probability

$$\text{Prob}\,(f = \lambda_j) = |a_j|^2. \tag{3.47}$$

Remark 3.11. In some cases, the appropriate Hilbert space for a quantum system is a finite dimensional space of the form $\mathbb{C}^n$ (this is the case for spin systems, for example). In this case, the quantum observables are the self-adjoint linear operators on $\mathbb{C}^n$, which can be represented by Hermitian matrices, and the possible values of an observable are the eigenvalues of the corresponding matrix.

3.3 Time evolution and the Schrödinger equation

Our next axiom concerns time evolution. Axiom 2 tells us that, given the classical Hamiltonian H, there is a self-adjoint operator on $\mathcal{H}$ associated to it, which we will denote by H.

Axiom 4: The time evolution of the wave function ψ is given by the Schrödinger equation,

$$\mathrm{i}\hbar \frac{\mathrm{d}\psi}{\mathrm{d}t} = \mathsf{H}\psi. \tag{3.48}$$

One of the most beautiful implications of Schrödinger's equation is that the average of operators evolve with time in a way reminiscent of classical mechanics, and we have the following

Proposition 3.12. Let $\psi(t)$ be a solution to the Schrödinger equation, and A a self-adjoint operator. Assuming certain natural domain conditions to hold, one has

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_{\psi(t)} = \left\langle \frac{1}{\mathrm{i}\hbar}[A, \mathsf{H}] \right\rangle_{\psi(t)}, \tag{3.49}$$

where $[\cdot,\cdot]$ denotes the commutator of two operators,

$$[A, B] = AB - BA. \tag{3.50}$$

Proof: we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\langle \psi(t), A\psi(t)\rangle &= \left\langle \frac{\mathrm{d}\psi}{\mathrm{d}t}, A\psi(t) \right\rangle + \left\langle \psi(t), A\frac{\mathrm{d}\psi}{\mathrm{d}t} \right\rangle \\ &= \frac{\mathsf{i}}{\hbar}\langle \mathsf{H}\psi, A\psi\rangle - \frac{\mathsf{i}}{\hbar}\langle \psi, A\mathsf{H}\psi\rangle \\ &= \frac{\mathsf{i}}{\hbar}\langle \psi, \mathsf{H}A\psi\rangle - \frac{\mathsf{i}}{\hbar}\langle \psi, A\mathsf{H}\psi\rangle \\ &= \frac{1}{\mathsf{i}\hbar}\left\langle \psi, [A, \mathsf{H}]\psi \right\rangle. \end{aligned} \tag{3.51}$$

□

By comparing (3.49) to (2.61), it follows that the quantity

$$\frac{1}{\mathsf{i}\hbar}[A, B] \tag{3.52}$$

plays the rôle in Quantum Mechanics of the Poisson bracket in Classical Mechanics. This similarity is strengthened by the following result, to be compared to the properties of the Poisson bracket in Proposition 2.11.

Proposition 3.13. For any complex vector space and linear operators A, B, C, the following relations hold:

1. $[B, A] = -[A, B]$.
2. $[A, BC] = B[A, C] + [A, B], C$
3. $[A, [B, C]] = [[A, B], C] + [B, [A, C]]$

In some cases, one can solve the Schrödinger equation as follows. Let us consider the equation defining eigenvectors and eigenvalues of the Hamiltonian operator,

$$\mathsf{H}\psi = E\psi. \tag{3.53}$$

This is sometimes called the *time-independent Schrödinger equation*. Solving this equation provides very important information for a quantum system, since the eigenvalues of H give the possible values of the energy of the system (indeed, historically, the first success of Quantum Mechanics was to derive the spectrum of energy of the hydrogen atom by solving this equation). Let us now assume that H has a pure point spectrum, i.e. that there exists an orthonormal basis $\{\varphi_n\}_{n=0,1,\dots}$ for the Hilbert space $\mathcal{H}$ consisting of eigenvectors of H, such that

$$\mathsf{H}\varphi_n = E_n\varphi_n, \qquad n = 0, 1, \cdots \tag{3.54}$$

Let $\psi(0)$ be a initial state. By assumption, we can expand it as

$$\psi(0) = \sum_{n=0}^{\infty} c_n\varphi_n, \qquad c_n \in \mathbb{C}. \tag{3.55}$$

Then, it is easy to see that

$$\psi(t) = \sum_{n=0}^{\infty} c_n \mathrm{e}^{-\frac{\mathsf{i}E_n t}{\hbar}}\varphi_n \tag{3.56}$$

solves the Schrödinger equation with the appropriate initial condition. Another way to see this is that (3.48) can be formally solved as

$$\psi(t) = \mathrm{e}^{-\mathsf{i}\mathsf{H}t/\hbar}\psi(0), \tag{3.57}$$

but in order to make sense of this in a precise way we have to give a meaning to the exponential. If A is a bounded operator, its exponential can be defined by the series

$$\mathrm{e}^{\mathsf{A}} = \sum_{n=0}^{\infty} \frac{\mathsf{A}^n}{n!}. \tag{3.58}$$

The Hamiltonian is rarely bounded, but the above exponential can be defined by using more sophisticated tools (in particular, the spectral theorem). The resulting operator has all expected properties. In particular, if φ is an eigenfunction of H with eigenvalue E, then

$$\mathrm{e}^{-\mathsf{i}\mathsf{H}t/\hbar}\varphi = \mathrm{e}^{-\mathsf{i}Et/\hbar}\varphi. \tag{3.59}$$

Example 3.14. In one dimension, with the standard form of the Hamiltonian, the Schrödinger equation takes the form

$$\mathsf{i}\hbar\frac{\partial\psi(q,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(q,t)}{\partial q^2} + V(q)\psi(q,t), \tag{3.60}$$

while the time-independent Schrödinger equation is of the form

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial q^2} + V(q)\psi(q) = E\psi(q). \tag{3.61}$$

Example 3.15. *Particle in a box.* A simple example of quantum system is a "particle in a box", i.e. a free particle moving in a segment $[0, L] \subset \mathbb{R}$. We assume that the walls of the box are impenetrable, which means that the wave function describing the particle vanishes at the endpoints of the segment,

$$\psi(0) = \psi(L) = 0. \tag{3.62}$$

The relevant Hilbert space is therefore $L^2([0, L])$, with the boundary condition above. The time-independent Schrödinger equation is simply

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial q^2} = E\psi(q) \tag{3.63}$$

with the boundary conditions (3.62). It is easy to see that the solutions to this equation are of the form

$$\psi(q) = A\sin\left(\frac{\sqrt{2mE}}{\hbar}q\right) + B\cos\left(\frac{\sqrt{2mE}}{\hbar}q\right), \tag{3.64}$$

but imposing $\psi(0) = 0$ leads to $B = 0$. The resulting function does not verify $\psi(L) = 0$, unless

$$\frac{\sqrt{2mE}}{\hbar}L = n\pi, \qquad n = 1, 2, \cdots. \tag{3.65}$$

Therefore, the boundary condition (3.62) leads to the *quantization* of the energy, which can only take discrete values,

$$E = \frac{n^2\pi^2\hbar^2}{2mL^2}, \qquad n = 1, 2, \cdots. \tag{3.66}$$

The corresponding eigenfunctions,

$$\psi_n(q) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi q}{L}\right), \qquad n = 1, 2, \cdots, \tag{3.67}$$

give an orthonormal basis for the Hilbert space $\mathcal{H}$, as it follows from the theory of Fourier series.

So far we have considered that the operators do not evolve in time, while states do (following Schrödinger's equation). This is the so-called "Schrödinger picture" of time evolution in Quantum Mechanics. There is another picture, called the "Heisenberg picture," in which states do not evolve, but operators do. To see how the two pictures are related, we just note that the average of an operator A at time t is computed, in the Schrödinger picture, by

$$\langle\psi(t), A\psi(t)\rangle = \langle \mathrm{e}^{-\mathsf{i}\mathsf{H}t/\hbar}\psi(0), A\mathrm{e}^{-\mathsf{i}\mathsf{H}t/\hbar}\psi(0)\rangle = \langle\psi(0), \mathrm{e}^{\mathsf{i}\mathsf{H}t/\hbar}A\mathrm{e}^{-\mathsf{i}\mathsf{H}t/\hbar}\psi(0)\rangle. \tag{3.68}$$

This should equal

$$\langle \psi(0), A(t)\psi(0)\rangle \tag{3.69}$$

which tells us that

$$A(t) = \mathrm{e}^{\mathsf{i}\mathsf{H}t/\hbar} A \mathrm{e}^{-\mathsf{i}\mathsf{H}t/\hbar}. \tag{3.70}$$

Equivalently, in the Heisenberg picture, operators satisfy the EOM

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = \frac{1}{\mathsf{i}\hbar}[A(t), \mathsf{H}], \tag{3.71}$$

which should be again compared to (2.61).

3.4 The uncertainty principle

Many fundamental operators in QM do not commute, like for example the position and momentum operators. As in the case of matrices, this means that we can not diagonalize them simultaneously, i.e. that in general we can not find states which are eigenfunctions of the two operators simultaneously. Recall that, if a state is an eigenfunction of the operator A, the value of the physical quantity associated to A is known with total certainty and given by the corresponding eigenvalue (i.e. the probability distribution for this physical quantity is peaked at this eigenvalue). Therefore, if two operators do not commute, we can not find states where the corresponding physical quantities are certain. The uncertainty principle gives a precise mathematical expression to this result.

Let us define the uncertainty of a symmetric operator A in the state ψ by

$$(\Delta_\psi A)^2 = \langle A^2\rangle_\psi - \langle A\rangle_\psi^2. \tag{3.72}$$

In probability theory, this is the mean square deviation, which tells us how "spread out" is the probability distribution of the corresponding physical quantity. Note that we can write this uncertainty as the norm of a certain vector in $\mathcal{H}$,

$$(\Delta_\psi A)^2 = \| \left(A - \langle A\rangle_\psi\right)\psi\|^2, \tag{3.73}$$

therefore the uncertainty vanish if and only if this vector vanishes, i.e. if and only if

$$A\psi = \langle A\rangle_\psi \psi, \tag{3.74}$$

which means that ψ is an eigenvector of A. We now have the following

Theorem 3.16. *(Uncertainty principle) Let A and B be two symmetric operators, and ψ a unit vector in the appropriate domain. Then,*

$$(\Delta_\psi A)^2 (\Delta_\psi B)^2 \geq \frac{1}{4}\left|\langle [A,B]\rangle_\psi\right|^2. \tag{3.75}$$

Proof: Let us denote by

$$A' = A - \langle A\rangle_\psi, \qquad B' = B - \langle B\rangle_\psi \tag{3.76}$$

the operators appearing in the definition of the uncertainties. We can now use the Cauchy–Schwarz inequality to find,

$$\begin{aligned} \|A'\psi\|^2\|B'\psi\|^2 &\geq \left|\langle A'\psi, B'\psi\rangle\right|^2 \\ &\geq \left|\mathrm{Im}\langle A'\psi, B'\psi\rangle\right|^2 \\ &= \frac{1}{4}\left|\langle A'\psi, B'\psi\rangle - \langle B'\psi, A'\psi\rangle\right|^2. \end{aligned} \tag{3.77}$$

Since A and B are symmetric, A' and B' are symmetric, too, and we can write

$$\|A'\psi\|^2\|B'\psi\|^2 \geq \frac{1}{4}\left|\langle\psi, A'B'\psi\rangle - \langle\psi, B'A'\psi\rangle\right|^2 = \frac{1}{4}\left|\langle\psi, [A', B']\psi\rangle\right|^2. \tag{3.78}$$

It is obvious that $[A', B'] = [A, B]$, therefore we conclude that

$$(\Delta_\psi A)^2\,(\Delta_\psi B)^2 \geq \frac{1}{4}\left|\langle\psi, [A', B']\psi\rangle\right|^2. \tag{3.79}$$

which is what we wanted to show. □

It is interesting to see when the inequality in the above theorem is saturated. This means that we have to saturate the Cauchy–Schwarz inequality in (3.77). This only happens if (a) $A'\psi = 0$ or (b) $B'\psi = 0$ or (c) $A'\psi$ is proportional to $B'\psi$. Options (a) and (b) hold if ψ is an eigenvector of A or of B, respectively, and in these cases the inequality is trivially saturated (the l.h.s. and the r.h.s. are both zero). Let us then assume that $A'\psi = cB'\psi$. We still have to saturate the inequality in going from the second to the third line in (3.77), i.e. we need

$$\mathrm{Re}\langle A'\psi, B'\psi\rangle = 0, \tag{3.80}$$

or, equivalently, $\mathrm{Re}(c) = 0$. We conclude that the inequality is saturated when

$$(A' - \mathsf{i}\gamma B')\psi = 0, \qquad \gamma \in \mathbb{R}. \tag{3.81}$$

This happens if ψ is an eigenvector of $A - \mathsf{i}\gamma B$, since

$$A' - \mathsf{i}\gamma B' = A - \mathsf{i}\gamma B - (\langle A\rangle_\psi - \mathsf{i}\gamma\langle B\rangle_\psi), \tag{3.82}$$

and the corresponding eigenvalue is

$$\lambda = \langle A\rangle_\psi - \mathsf{i}\gamma\langle B\rangle_\psi. \tag{3.83}$$

As an application of the above theorem, we note that the uncertainties of q and p satisfy

$$(\Delta_\psi \mathsf{q})\,(\Delta_\psi \mathsf{p}) \geq \frac{\hbar}{2}. \tag{3.84}$$

This is the archi-famous *Heisenberg's uncertainty principle.*

3.5 The harmonic oscillator

The Hamiltonian of the harmonic oscillator is given by

$$\mathsf{H} = \frac{1}{2m}\left(\mathsf{p}^2 + m^2\omega^2\mathsf{q}^2\right). \tag{3.85}$$

We want to find the spectrum of this operator, i.e. we want to solve the time-independent Schrödinger equation. There are at least two ways to do that. The first one is purely algebraic and introduce the important notion of annihilation and creation operators. The second one is based on the analysis of the boundary problem for the Schrödinger equation, regarded as a second order ODE. We will go through both methods in this section.

We first consider the algebraic approach. We introduce the *annihilation (or lowering) operator*

$$a = \frac{m\omega\mathsf{q} + \mathsf{i}\mathsf{p}}{\sqrt{2\hbar m\omega}}. \tag{3.86}$$

Its adjoint is called the *creation (or raising) operator*

$$a^\dagger = \frac{m\omega \mathsf{q} - \mathsf{i}\mathsf{p}}{\sqrt{2\hbar m\omega}}. \tag{3.87}$$

These operators have two important algebraic properties. The first one is

$$\begin{aligned} a^\dagger a &= \frac{1}{2m\hbar\omega}\left(m^2\omega^2\mathsf{q}^2 + \mathsf{p}^2 + \mathsf{i}m\omega[\mathsf{q},\mathsf{p}]\right) = \frac{1}{2\hbar m\omega}\left(\mathsf{p}^2 + m^2\omega^2\mathsf{q}^2\right) - \frac{1}{2} \\ &= \frac{1}{\hbar\omega}\mathsf{H} - \frac{1}{2}. \end{aligned} \tag{3.88}$$

The second one is

$$[a, a^\dagger] = \frac{1}{2m\hbar\omega}\left([m\omega\mathsf{q}, -\mathsf{i}\mathsf{p}] + [\mathsf{i}\mathsf{p}, m\omega\mathsf{q}]\right) = 1. \tag{3.89}$$

Let us now define the *number operator*

$$N = a^\dagger a. \tag{3.90}$$

This operator is clearly symmetric. We now have the following

Proposition 3.17. Suppose ψ is an eigenvector of N with eigenvalue λ. Then,

$$\begin{aligned} N(a\psi) &= (\lambda - 1)a\psi, \\ N(a^\dagger\psi) &= (\lambda + 1)a^\dagger\psi. \end{aligned} \tag{3.91}$$

Proof: By using (3.89) and property (3) in Proposition 3.13, we find

$$[a, N] = a, \qquad [a^\dagger, N] = -a^\dagger. \tag{3.92}$$

It follows that

$$Na\psi = (aN - a)\psi = (\lambda - 1)\psi. \tag{3.93}$$

The other equality is proved in a similar way. □

It is easy to see that the eigenvalue λ is non-negative. Indeed, we have

$$\langle\psi, N\psi\rangle = \lambda\|\psi\|^2 = \|a\psi\|^2 \geq 0. \tag{3.94}$$

Let us now assume that N has at least one eigenvector ψ with eigenvalue λ. If we apply a to ψ repeatedly, we obtain eigenvectors of N with lower and lower eigenvalues. Since these have to be non-negative, there must be a non-negative integer M such that

$$a^{M+1}\psi = 0, \qquad a^M\psi \neq 0. \tag{3.95}$$

Let us denote

$$\psi_0 = a^M\psi, \tag{3.96}$$

so that

$$a\psi_0 = 0. \tag{3.97}$$

Theorem 3.18. *If ψ_0 is a unit vector that satisfies (3.97), the vectors*

$$\psi_n = \frac{\left(a^\dagger\right)^n}{\sqrt{n!}}\psi_0, \qquad n = 0, 1, 2, \cdots \tag{3.98}$$

are orthonormal eigenvectors of H with eigenvalues

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right), \qquad n = 0, 1, 2, \cdots \tag{3.99}$$

Proof: let us first show that the ψ_n are eigenvectors of the number operator N. We have that

$$N(a^\dagger)^n\psi_0 = n(a^\dagger)^n\psi_0, \tag{3.100}$$

where we used

$$[N,(a^\dagger)^n] = a^\dagger[a,(a^\dagger)^n] = a^\dagger n(a^\dagger)^{n-1} = n(a^\dagger)^n. \tag{3.101}$$

It follows from (3.88) that ψ_n are eigenvalues of H with the eigenvalues E_n above. Since N is symmetric, eigenvectors corresponding to different eigenvalues are orthogonal. To see that the ψ_n are properly normalized, we take into account that

$$a^\dagger\psi_n = \frac{(a^\dagger)^{n+1}}{\sqrt{n!}}\psi_0 = \sqrt{n+1}\psi_{n+1}, \tag{3.102}$$

so that

$$\psi_n = \frac{1}{\sqrt{n}}a^\dagger\psi_{n-1}. \tag{3.103}$$

We also have

$$a\psi_n = a\frac{(a^\dagger)^n}{\sqrt{n!}}\psi_0 = n\frac{(a^\dagger)^{n-1}}{\sqrt{n!}}\psi_0 = \sqrt{n}\psi_{n-1}. \tag{3.104}$$

Then,

$$\|\psi_n\|^2 = \frac{1}{\sqrt{n}}\langle a^\dagger\psi_{n-1},\psi_n\rangle = \frac{1}{\sqrt{n}}\langle\psi_{n-1},a\psi_n\rangle = \|\psi_{n-1}\|^2. \tag{3.105}$$

It follows that

$$\|\psi_n\|^2 = \|\psi_0\|^2 = 1. \tag{3.106}$$

□

In order to proceed, we show that the unit vector ψ_0 exists and is unique (up to a phase). We rescale

$$\tilde{q} = \frac{q}{D}, \qquad D = \sqrt{\frac{\hbar}{m\omega}}. \tag{3.107}$$

In terms of this rescaled variable, and in the position representation, we have

$$a = \frac{1}{\sqrt{2}}\left(\tilde{q} + \frac{\mathrm{d}}{\mathrm{d}\tilde{q}}\right). \tag{3.108}$$

We can now solve (3.97) as an ordinary first-order differential equation, to find that

$$\psi_0(\tilde{q}) = C\mathrm{e}^{-\tilde{q}^2/2}. \tag{3.109}$$

The overall constant is fixed by normalization:

$$\int_{\mathbb{R}}|\psi_0(q)|^2\,\mathrm{d}q = D\int_{\mathbb{R}}|\psi_0(\tilde{q})|^2\,\mathrm{d}\tilde{q} = DC^2\int_{\mathbb{R}}\mathrm{e}^{-\tilde{q}^2}\mathrm{d}\tilde{q} = DC^2\pi^{1/2} = 1, \tag{3.110}$$

and we obtain

$$\psi_0(q) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left(-\frac{m\omega}{2\hbar}q^2\right). \tag{3.111}$$

The "excited" states ψ_n can be also described in closed form. They have the structure

$$\psi_n(q) = H_n(q)\psi_0(q), \tag{3.112}$$

where the polynomials H_n are defined recursively as:

$$\begin{aligned} H_0(q) &= 1, \\ H_{n+1}(\tilde{q}) &= \frac{1}{\sqrt{2(n+1)}} \left(2\tilde{q} H_n(\tilde{q}) - \frac{\mathrm{d}H_n(\tilde{q})}{\mathrm{d}\tilde{q}} \right), \qquad n \geq 0. \end{aligned} \tag{3.113}$$

This follows immediately from (3.102). One can check that these polynomials are, up to a normalization, the Hermite polynomials.

Remark 3.19. Since the $\psi_n \in \mathcal{S}(\mathbb{R})$, the space of Schwartz functions, they belong to the domain of all the operators involved in the analysis above, and all of our derivations are in fact justified.

Let us now quickly consider a more analytic derivation of the spectrum of the Hamiltonian of the harmonic oscillator. The time-independent Schrödinger equation, in the position representation, is given by

$$-\psi'' + \left(u^2 - \mathcal{E}\right)\psi = 0, \tag{3.114}$$

where we have changed variables

$$q = \left(\frac{\hbar}{m\omega}\right)^{1/2} u, \qquad \mathcal{E} = \frac{2E}{\hbar\omega}. \tag{3.115}$$

This ODE can be in turn put in the form of the parabolic cylinder equation (following the notations in [2], p. 96 ss):

$$y'' + \left(\nu + \frac{1}{2} - \frac{1}{4}z^2\right) y = 0. \tag{3.116}$$

We find that, after the change of variables

$$z = \sqrt{2}u, \tag{3.117}$$

we recover (3.114) with

$$\mathcal{E} = 2\nu + 1. \tag{3.118}$$

The parabolic cylinder equation has two independent solutions which are analytic in the complex plane. One of these is parabolic cylinder function $D_\nu(z)$, which behaves as

$$D_\nu(z) \sim z^\nu \mathrm{e}^{-z^2/4}, \qquad z \to +\infty. \tag{3.119}$$

The other linearly independent solution behaves as

$$D_{-\nu-1}(-\mathrm{i}z) \sim z^{-\nu-1}\mathrm{e}^{+z^2/4}, \qquad z \to +\infty. \tag{3.120}$$

Clearly, if $\psi(u)$ must belong to $L^2(\mathbb{R})$, it has to decrease at infinity. This means that the wavefunction must be given by the parabolic cylinder function,

$$\psi(u) \propto D_\nu\left(\sqrt{2}u\right). \tag{3.121}$$

What happens when $u \to -\infty$? For this we need to study the asymptotic behavior of the above function for negative u. This can be done with standard asymptotic techniques [2], and one finds

$$D_\nu(z) \sim z^\nu \mathrm{e}^{-z^2/4}\left(1 + \mathcal{O}(z^{-2})\right) - \frac{\sqrt{2\pi}}{\Gamma(-\nu)} \mathrm{e}^{\pi\mathrm{i}\nu} z^{-\nu-1}\mathrm{e}^{z^2/4}\left(1 + \mathcal{O}(z^{-2})\right), \qquad z \to -\infty. \tag{3.122}$$

Therefore, this solution grows at infinity *unless*

$$\frac{1}{\Gamma(-\nu)} = 0. \tag{3.123}$$

This implies

$$\nu = n, \qquad n = 0, 1, 2, \cdots \tag{3.124}$$

The boundary conditions imposed on the wavefunction select a solution and in addition lead to a "quantization condition" on the parameter ν (which a priori could be complex). In particular, by using (3.118) and (3.124), we recover the spectrum of the harmonic oscillator in (3.99).

4 Phase space formulation of Quantum Mechanics

4.1 Weyl–Wigner quantization

Axiom 2 of QM tell us that, given a classical observable, there is a corresponding self-adjoint operator. However, we have not been very explicit on the structure of this correspondence. In this Chapter we will consider a classical system whose phase space is $\mathbb{R}^2$, parametrized by q, p, and we will present the so-called Weyl quantization (or Weyl–Wigner quantization) of certain classical observables in this phase space.

We have seen that the position and momentum become two operators q and p, and that the classical Hamiltonian in one dimension becomes the quantum operator H. Generalizing this procedure for polynomials is already far from obvious, and in fact quantization procedures are not unique. Let us consider for example the classical observable

$$q^2 p^2. \tag{4.1}$$

Clearly, the quantization of this function can not be $\mathsf{q}^2\mathsf{p}^2$, since this is not self-adjoint. We can fix this by considering the operator

$$\frac{1}{2}\left(\mathsf{q}^2\mathsf{p}^2 + \mathsf{p}^2\mathsf{q}^2\right), \tag{4.2}$$

and this is one possible quantization procedure (called "symmetrized pseudodifferential operator quantization"). Another possibility however is to consider the operator

$$\frac{1}{6}(\mathsf{q}^2\mathsf{p}^2 + \mathsf{p}^2\mathsf{q}^2 + \mathsf{pqpq} + \mathsf{pq}^2\mathsf{p} + \mathsf{qp}^2\mathsf{q} + \mathsf{qpqp}) \tag{4.3}$$

which is also self-adjoint. This last scheme, in which we consider all possible orderings of a given monomial in q and p, is called Weyl quantization, and it is singled-out among all quantization schemes of the phase space $\mathbb{R}^{2n}$ by having the best properties.

Definition 4.1. Given a polynomial in q and p, we define the *Weyl quantization* by the following formula acting on monomials,

$$W(q^j p^k) = \frac{1}{(j+k)!} \sum_{\sigma \in S_{j+k}} \sigma(\mathsf{q}, \mathsf{q}, \cdots, \mathsf{q}, \mathsf{p}, \mathsf{p}, \cdots, \mathsf{p}). \tag{4.4}$$

Here, S_n is the symmetric group of n elements, and for any set of n operators $\mathsf{A}_1, \cdots, \mathsf{A}_n$ and any $\sigma \in S_n$, we define

$$\sigma(\mathsf{A}_1, \cdots, \mathsf{A}_1) = \mathsf{A}_{\sigma(1)} \cdots \mathsf{A}_{\sigma(n)}. \tag{4.5}$$

The map W is extended to arbitrary polynomials by linearity, and it is sometimes called *Weyl transform*.

It is easy to see that the map W satisfies

$$W\left((uq+vp)^j\right) = (u\mathsf{q}+u\mathsf{p})^j, \qquad u,v\in\mathbb{R}. \tag{4.6}$$

We would like now to extend this map to more general functions on $\mathbb{R}^2$. Let us introduce the following operators,

$$\mathsf{U}(u) = \mathrm{e}^{\mathrm{i}u\mathsf{q}}, \qquad \mathsf{V}(v) = \mathrm{e}^{\mathrm{i}v\mathsf{p}}, \qquad u,v\in\mathbb{R}. \tag{4.7}$$

A rigorous definition of these operators requires sophisticated tools, like the spectral theorem and Stone's theorem. It can be shown that, when acting on a function $\psi(q)\in L^2(\mathbb{R})$, one has

$$\left(\mathsf{U}(u)\psi\right)(q) = \mathrm{e}^{\mathrm{i}uq}\psi(q), \qquad \left(\mathsf{V}(v)\psi\right)(q) = \psi\left(q+\hbar v\right). \tag{4.8}$$

Heuristically, these relations can be obtained by defining the exponentials through an infinite series (as in (3.58)) and assuming that ψ is real analytic. It follows from (4.8) that these operators satisfy on $\mathcal{H}$ the *Weyl commutation relation*,

$$\mathsf{U}(u)\mathsf{V}(v) = \mathrm{e}^{-\mathrm{i}\hbar uv}\mathsf{V}(v)\mathsf{U}(u). \tag{4.9}$$

Let us now introduce the operator,

$$\mathsf{S}(u,v) = \mathrm{e}^{\mathrm{i}\hbar uv/2}\mathsf{U}(u)\mathsf{V}(v). \tag{4.10}$$

This acts on wavefunctions $\psi(q)$ as

$$\left(\mathsf{S}(u,v)\psi\right)(q) = \mathrm{e}^{\frac{\mathrm{i}\hbar}{2}uv+\mathrm{i}uq}\psi(q+\hbar v). \tag{4.11}$$

The Baker–Campbell–Hausdorff formula says that, given two bounded operators A, B, such that

$$[\mathsf{A},[\mathsf{A},\mathsf{B}]] = [\mathsf{B},[\mathsf{A},\mathsf{B}]] = 0, \tag{4.12}$$

one has

$$\mathrm{e}^{\mathsf{A}}\mathrm{e}^{\mathsf{B}} = \mathrm{e}^{\mathsf{A}+\mathsf{B}+\frac{1}{2}[\mathsf{A},\mathsf{B}]}. \tag{4.13}$$

Applying this (again, heuristically) to the operators $\mathsf{A} = \mathrm{i}u\mathsf{q}$, $\mathsf{B} = \mathrm{i}v\mathsf{p}$, we obtain

$$\mathsf{S}(u,v) = \mathrm{e}^{\mathrm{i}u\mathsf{q}+\mathrm{i}v\mathsf{p}}. \tag{4.14}$$

This can be again justified by using Stone's theorem and the spectral theorem.

In order to motivate the extension of Weyl's transform to more general functions, we note that, if we multiply (4.6) by $\mathrm{i}^j/j!$, we would expect to have

$$W\left(\mathrm{e}^{\mathrm{i}uq+\mathrm{i}vp}\right) = \mathsf{S}(u,v). \tag{4.15}$$

Let us now suppose that we have a function $f(q,p)$, which is sufficiently nice so that we can write it as a Fourier transform

$$f(q,p) = \frac{1}{2\pi}\int_{\mathbb{R}^2}\check{f}(u,v)\mathrm{e}^{\mathrm{i}(uq+vp)}\mathrm{d}u\mathrm{d}v. \tag{4.16}$$

We can now promote this to an operator by using (4.15) as a guiding principle. It is then natural to define the Weyl transform of the function f, $W(f)$, as

$$W(f) = \frac{1}{2\pi}\int_{\mathbb{R}^2}\check{f}(u,v)\mathsf{S}(u,v)\mathrm{d}u\mathrm{d}v. \tag{4.17}$$

Note that $W(f)$ is defined by integrating an operator depending on a set of continuous parameters, $\mathsf{S}(u,v)$. The resulting operator satisfies the following property,

$$\langle \psi_2, W(f)\psi_1\rangle = \frac{1}{2\pi}\int_{\mathbb{R}^2} \check{f}(u,v)\langle \psi_2, \mathsf{S}(u,v)\psi_1\rangle \mathrm{d}u\mathrm{d}v. \tag{4.18}$$

It is easy to see that, if $\check{f} \in L^1(\mathbb{R}^2)\cap L^2(\mathbb{R}^2)$, the Weyl transform defines a bounded operator on $\mathcal{H} = L^2(\mathbb{R})$, satisfying

$$\|W(f)\| \leq \frac{1}{2\pi}\|f\|_{L^1}. \tag{4.19}$$

We can regard the Weyl transform as a sort of *non-commutative version of the Fourier transform.*

It turns out that the operator $W(f)$ can be represented by an integral operator, and its kernel can be computed as follows. By using (4.11), we obtain

$$\begin{aligned}\langle \psi_2, W(f)\psi_1\rangle &= \frac{1}{2\pi}\int_{\mathbb{R}^2}\check{f}(u,v)\left(\int_{\mathbb{R}}\psi_2^*(q)\mathrm{e}^{\frac{\mathrm{i}\hbar}{2}uv+\mathrm{i}uq}\psi_1(q+\hbar v)\mathrm{d}q\right)\mathrm{d}u\mathrm{d}v \\ &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} K_f(q,q')\psi_1(q')\mathrm{d}q'\right)\psi_2^*(q)\mathrm{d}q,\end{aligned} \tag{4.20}$$

where we set

$$q' = q + \hbar v, \tag{4.21}$$

so that

$$v = \frac{q'-q}{\hbar}, \qquad \frac{\hbar}{2}v + q = \frac{q+q'}{2}, \qquad \mathrm{d}q\mathrm{d}v = \frac{1}{\hbar}\mathrm{d}q\mathrm{d}q', \tag{4.22}$$

and we can exchange the order of integration thanks to Fubini's theorem. The function $K_f(q,q')$ is given by

$$K_f(q,q') = \frac{1}{2\pi\hbar}\int \check{f}\left(u, \frac{q'-q}{\hbar}\right)\mathrm{e}^{\frac{\mathrm{i}}{2}(q+q')u}\mathrm{d}u. \tag{4.23}$$

We conclude that $W(f)$ can be represented by an integral operator with kernel $K_f(q,q')$:

$$(W(f)\psi)(q) = \int_{\mathbb{R}} K_f(q,q')\psi(q')\mathrm{d}q', \qquad \psi \in L^2(\mathbb{R}). \tag{4.24}$$

Through a change of variables, the kernel $K_f(q,q')$ can be thought of as depending on the "relative" and "center of mass" coordinates,

$$q_r = q' - q, \qquad q_c = \frac{q+q'}{2}, \tag{4.25}$$

and (4.23) can be thought of as a partial Fourier transform of the function $\check{f}(u, q_r/\hbar)$ from u to the "center of mass" coordinate. Since $\check{f}(u,v)$ is itself a Fourier transform, we obtain

$$K_f(q,q') = \frac{1}{2\pi\hbar}\int_{\mathbb{R}} f\left(\frac{q+q'}{2}, p\right)\mathrm{e}^{\frac{\mathrm{i}}{\hbar}p(q-q')}\mathrm{d}p. \tag{4.26}$$

We also deduced that

$$\begin{aligned}\int |K_f(q,q')|^2\mathrm{d}q\mathrm{d}q' &= \int \left(|K_f(q_r,q_c)|^2\mathrm{d}q_c\right)\mathrm{d}q_r = \frac{1}{\hbar^2}\int\left(|\check{f}(v,q_r/\hbar)|^2\mathrm{d}u\right)\mathrm{d}q_r \\ &= \frac{1}{\hbar}\int |\check{f}(u,v)|^2\mathrm{d}u\mathrm{d}v = \frac{1}{\hbar}\int |f(q,p)|^2\mathrm{d}q\mathrm{d}p,\end{aligned} \tag{4.27}$$

where we have used Plancherel's theorem. Therefore, if $f \in L^2(\mathbb{R}^2)$, the kernel is square integrable, as a function on $\mathbb{R}^2$. Such a kernel defines a *Hilbert–Schmidt* (HS) operator. In particular, it is a well-defined operator on $L^2(\mathbb{R})$, in the sense that it maps functions in $L^2(\mathbb{R})$ to functions in $L^2(\mathbb{R})$. We will denote the space of HS operators on a Hilbert space $\mathcal{H}$ by $\mathcal{J}_2(\mathcal{H})$.

In view of (4.24) and (4.26), we can now use (4.24) and (4.26) to define the Weyl transform as a map

$$W : L^2(\mathbb{R}^2) \to \mathcal{J}_2\left(L^2(\mathbb{R})\right), \tag{4.28}$$

from the set of square-integrable functions on phase space $\mathbb{R}^2$, to the space of HS operators on $L^2(\mathbb{R})$. We can construct explicitly an inverse map

$$W^{-1} : \mathcal{J}_2(\mathcal{H}) \to L^2(\mathbb{R}^2), \tag{4.29}$$

which is given by an inverse Fourier transform

$$W^{-1}(K) = K_{\mathrm{W}}(q,p) = \int_{\mathbb{R}} K\left(q - \frac{q'}{2}, q + \frac{q'}{2}\right) \mathrm{e}^{\frac{\mathrm{i}}{\hbar}pq'} \mathrm{d}q'. \tag{4.30}$$

(Here, we are implicitly using the fact that every HS operator on $L^2(\mathbb{R})$ can be represented by a kernel in $L^2(\mathbb{R}^2)$). The function $K_{\mathrm{W}}(q,p)$ is sometimes called *Wigner transform* of the operator with kernel $K(q,q')$. In physics notation, we have

$$K_{\mathrm{W}}(q,p) = \int \mathrm{d}q' \left\langle q - \frac{q'}{2} \middle| \mathsf{K} \middle| q + \frac{q'}{2} \right\rangle \mathrm{e}^{\mathrm{i}pq'/\hbar}. \tag{4.31}$$

An important example of the inverse Weyl transform occurs when the kernel is of the form,

$$\rho_\psi(q,q') = \psi(q)\psi^*(q'), \tag{4.32}$$

The corresponding operator is the projector on the subspace generated by ψ. In this case, the Weyl transform (with an extra $2\pi\hbar$ factor)

$$f_\psi(p,q) = \frac{1}{2\pi\hbar} W^{-1}(\rho_\psi) = \frac{1}{2\pi\hbar} \int_{\mathbb{R}} \psi^*\left(q + \frac{q'}{2}\right) \psi\left(q - \frac{q'}{2}\right) \mathrm{e}^{\frac{\mathrm{i}}{\hbar}pq'} \mathrm{d}q'. \tag{4.33}$$

is called the *Wigner function* associated to the wavefunction ψ.

4.2 Moyal product

We will now discuss some properties of the Weyl transform.

Lemma 4.2. *Conjugation becomes Hermitian conjugation*

$$(W(f))^\dagger = W(f^*). \tag{4.34}$$

Proof: This is easy to prove, since the kernel of the adjoint operator $(W(f))^\dagger$ is given by

$$K_f^\dagger(q,q') = K_f^*(q',q), \tag{4.35}$$

therefore, from the expression (4.26), we find

$$K_f^\dagger(q,q') = \frac{1}{2\pi\hbar} \int_{\mathbb{R}} f^*\left(\frac{q+q'}{2}, p\right) \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p(q-q')} \mathrm{d}p = K_{f^*}(q,q'). \tag{4.36}$$

□

Lemma 4.3. *Calculation of traces. If $f \in \mathcal{S}(R^2)$, the trace of $W(f)$ is well-defined and given by*

$$\operatorname{Tr} W(f) = \int K_f(q,q)\mathrm{d}q = \frac{1}{2\pi\hbar}\int_{\mathbb{R}^2} f(q,p)\mathrm{d}p\mathrm{d}q. \tag{4.37}$$

The most important property of the Weyl transforms is the $\star$ or Moyal product. Given two functions in phase space, f_1, f_2, a natural question is to obtain the function on phase space corresponding to the product operator $W(f_1)W(f_2)$, which is also HS.

Proposition 4.4. (*Moyal product*). Let f_1, $f_2 \in L^2(\mathbb{R}^2)$. Then,

$$W^{-1}\left(W(f_1)W(f_2)\right) = f_1 \star f_2, \tag{4.38}$$

where

$$(f_1 \star f_2)(q,p) = \frac{1}{(2\pi)^2}\int \mathrm{e}^{-\frac{\mathrm{i}\hbar}{2}(u_1v_2-u_2v_1)+\mathrm{i}q(u_1+u_2)+\mathrm{i}p(v_1+v_2)}\check{f}_1(u_1,v_1)\check{f}_2(u_2,v_2)\mathrm{d}u_1\mathrm{d}v_1\mathrm{d}u_2\mathrm{d}v_2 \tag{4.39}$$

is the $\star$ or *Moyal product* of f_1 and f_2.

Proof: The proof of this is straightforward. We first note that

$$\left[W^{-1}\left(W(f_1)W(f_2)\right)\right](q,p) = \int \left(W(f_1)W(f_2)\right)\left(q-\frac{1}{2}v, q+\frac{1}{2}v\right)\mathrm{e}^{\frac{\mathrm{i}}{\hbar}pv}\mathrm{d}v. \tag{4.40}$$

But

$$\begin{aligned}\left(W(f_1)W(f_2)\right)(q_1,q_2) &= \int K_{f_1}(q_1,q'')K_{f_2}(q'',q_2)\mathrm{d}q'' \\ &= \frac{1}{(2\pi\hbar)^2}\int \check{f}_1\left(u_1,\frac{q''-q_1}{\hbar}\right)\check{f}_2\left(u_1,\frac{q_2-q''}{\hbar}\right)\mathrm{e}^{\frac{\mathrm{i}}{2}(q_1+q'')u_1+\frac{\mathrm{i}}{2}(q''+q_2)u_2}\mathrm{d}q''\mathrm{d}u_1\mathrm{d}u_2.\end{aligned} \tag{4.41}$$

Therefore,

$$\begin{aligned}\left[W^{-1}\left(W(f_1)W(f_2)\right)\right](q,p) = \frac{1}{(2\pi\hbar)^2}\int \check{f}_1\left(u_1,\frac{1}{\hbar}\left(q''-q+\frac{v}{2}\right)\right)\check{f}_2\left(u_1,\frac{1}{\hbar}\left(q+\frac{v}{2}-q''\right)\right) \\ \times\exp\left\{\frac{\mathrm{i}}{\hbar}pv+\frac{\mathrm{i}}{2}\left(q-\frac{v}{2}+q''\right)u_1+\frac{\mathrm{i}}{2}\left(q''+q+\frac{v}{2}\right)u_2\right\}\mathrm{d}q''\mathrm{d}u_1\mathrm{d}u_2\mathrm{d}v.\end{aligned} \tag{4.42}$$

We now change variables

$$v,\, q'' \to v_1,\, v_2 \tag{4.43}$$

with

$$\begin{aligned}\hbar v_1 &= q''-q+\frac{v}{2}, \\ \hbar v_2 &= q+\frac{v}{2}-q''.\end{aligned} \tag{4.44}$$

This is inverted as

$$\begin{aligned}v &= \hbar(v_1+v_2), \\ q'' &= q+\hbar\frac{v_1-v_2}{2}.\end{aligned} \tag{4.45}$$

and the determinant of the Jacobian is $\hbar^2$. We find, for the exponent,

$$\frac{\mathrm{i}}{\hbar}pv+\frac{\mathrm{i}}{2}\left(q-\frac{v}{2}+q''\right)u_1+\frac{\mathrm{i}}{2}\left(q''+q+\frac{v}{2}\right)u_2 = -\frac{\mathrm{i}\hbar}{2}(u_1v_2-u_2v_1)+\mathrm{i}q(u_1+u_2)+\mathrm{i}p(v_1+v_2), \tag{4.46}$$

and we obtain the wished-for result. □

Example 4.5. *Star product of Gaussian functions.* A particularly important example of the $\star$ product are Gaussian functions, i.e. we consider

$$f_1(q,p) = \exp\left(-\frac{a}{\hbar}(q^2+p^2)\right), \qquad f_2(q,p) = \exp\left(-\frac{b}{\hbar}(q^2+p^2)\right). \tag{4.47}$$

Their Fourier transforms are

$$\check{f}_1(u,v) = \frac{1}{2\pi}\int \mathrm{e}^{-a/\hbar(q^2+p^2)}\mathrm{e}^{\mathrm{i}(uq+vp)}\mathrm{d}q\mathrm{d}p = \left(\frac{\hbar}{2a}\right)\mathrm{e}^{-\frac{\hbar}{4a}(u^2+v^2)}, \tag{4.48}$$

and similarly for $\check{f}_2(u,v)$ (with $a \leftrightarrow b$). The integral in (4.39) is given by four inverse Fourier transforms of Gaussians, and after some elementary calculations one obtains,

$$(f_1 \star f_2)(q,p) = \frac{1}{1+ab}\exp\left(-\frac{a+b}{1+ab}\frac{q^2+p^2}{\hbar}\right), \tag{4.49}$$

which is the hyperbolic tangent $\star$-composition law of Gaussians.

A very useful form of the Moyal product is the following one. We can write (4.39) in the more convenient form,

$$(f_1 \star f_2)(q,p) = \frac{1}{(2\pi)^2}\int \check{f}_1(u_1,v_1)\mathrm{e}^{\mathrm{i}(qu_1+pv_1)}\mathrm{e}^{-\frac{\mathrm{i}\hbar}{2}(u_1v_2-u_2v_1)}\mathrm{e}^{\mathrm{i}(qu_2+pv_2)}\check{f}_2(u_2,v_2)\mathrm{d}u_1\mathrm{d}v_1\mathrm{d}u_2\mathrm{d}v_2. \tag{4.50}$$

The exponent

$$\mathrm{e}^{\mathrm{i}(qu_1+pv_1)}\mathrm{e}^{-\frac{\mathrm{i}\hbar}{2}(u_1v_2-u_2v_1)}\mathrm{e}^{\mathrm{i}(qu_2+pv_2)} \tag{4.51}$$

can be obtained as

$$\mathrm{e}^{\mathrm{i}(qu_1+pv_1)}\exp\left[\frac{\mathrm{i}\hbar}{2}\left(\overleftarrow{\partial}_q\overrightarrow{\partial}_p - \overleftarrow{\partial}_p\overrightarrow{\partial}_q\right)\right]\mathrm{e}^{\mathrm{i}(qu_2+pv_2)}, \tag{4.52}$$

where the arrows indicate the direction in which the derivatives act, and we find that the r.h.s. of (4.39) is given by

$$\begin{aligned}\left(\int \mathrm{e}^{\mathrm{i}(qu_1+pv_1)}\check{f}_1(u_1,v_1)\frac{\mathrm{d}u_1\mathrm{d}v_1}{2\pi}\right)\exp\left[\frac{\mathrm{i}\hbar}{2}\left(\overleftarrow{\partial}_q\overrightarrow{\partial}_p - \overleftarrow{\partial}_p\overrightarrow{\partial}_q\right)\right]\\ \left(\int \check{\mathrm{e}}^{\mathrm{i}(qu_2+pv_2)}f_2(u_2,v_2)\frac{\mathrm{d}u_2\mathrm{d}v_2}{2\pi}\right),\end{aligned} \tag{4.53}$$

i.e.

$$(f_1 \star f_2)(q,p) = f_1(q,p)\exp\left[\frac{\mathrm{i}\hbar}{2}\left(\overleftarrow{\partial}_q\overrightarrow{\partial}_p - \overleftarrow{\partial}_p\overrightarrow{\partial}_q\right)\right]f_2(q,p). \tag{4.54}$$

This expression is extremely useful to obtain asymptotic, formal $\hbar$ expansions of this product. At next-to-leading order in an expansion in $\hbar$, we find

$$(f_1 \star f_2)(q,p) = f_1(q,p)f_2(q,p) + \frac{\mathrm{i}\hbar}{2}\{f_1,f_2\} + \cdots. \tag{4.55}$$

An alternative form of the product involves the Bopp shifts:

$$\begin{aligned}f_1(q,p) \star f_2(q,p) &= f_1\left(q, p - \frac{\mathrm{i}\hbar}{2}\overrightarrow{\partial}_q\right)f_2\left(q, p + \frac{\mathrm{i}\hbar}{2}\overleftarrow{\partial}_q\right)\\ &= f_1\left(q + \frac{\mathrm{i}\hbar}{2}\overrightarrow{\partial}_p, p - \frac{\mathrm{i}\hbar}{2}\overrightarrow{\partial}_q\right)f_2\left(q,p\right).\end{aligned} \tag{4.56}$$

Remark 4.6. The formula (4.39) for the Moyal product involves the Fourier transform of the functions $f_{1,2}(q,p)$, which only exists as a function under certain conditions (for example, if $f_{1,2} \in L^2(\mathbb{R}^2)$, as we have assumed). However, this formula can be extended to more general functions, like polynomials, by using distributions. In the case of polynomials, the result agrees with what one would obtain by using (4.54).

An easy consequence of (4.39) is that

$$\int \mathrm{d}q\mathrm{d}p\, f_1 \star f_2 = \int \mathrm{d}q\mathrm{d}p\, f_2 \star f_1 = \int \mathrm{d}q\mathrm{d}p\, f_1 f_2. \tag{4.57}$$

Indeed, when we integrate (4.39) we obtain

$$(2\pi)^2 \int \check{f}_1(u,v)\check{f}_2(-u,-v)\mathrm{d}u\mathrm{d}v = \int \mathrm{d}q\mathrm{d}p\, f_1 f_2, \tag{4.58}$$

and the $\hbar$ dependence drops out. This can be easily found by changing variables to "relative" and "center of mass" variables for u_i, v_i,

$$u_c = \frac{u_1+u_2}{2}, \qquad u_r = u_1 - u_2, \tag{4.59}$$

and similarly for v_i, and noting that, by integrating over q and p, we are computing the function for $u_c = v_c = 0$. Heuristically, we have

$$\int \mathrm{e}^{\mathrm{i}q(u_1+u_2)}\mathrm{d}q = 2\pi\delta(u_1+u_2), \tag{4.60}$$

which after integration sets $u_c = 0$.

If the star product is the analogue of the product of operators for Weyl transforms, the analogue of the commutator is the *Moyal bracket*, defined by

$$[A,B]_\star = A\star B - B\star A \tag{4.61}$$

By using the explicit asymptotic evaluation of the $\star$ product in (4.55), we see that, in the limit $\hbar \to 0$,

$$[A,B]_\star = \mathrm{i}\hbar\{A,B\} + \mathcal{O}(\hbar), \tag{4.62}$$

which makes more transparent the correspondence principle between classical and quantum mechanics.

Another consequence of the results above is that, if ψ is normalized, we have that the corresponding projector satisfies $\rho_\psi^2 = \rho_\psi$, therefore f_ψ is an idempotent, up to a normalization factor:

$$f_\psi \star f_\psi = \frac{1}{2\pi\hbar} f_\psi \tag{4.63}$$

4.3 Quantum Mechanics in phase space

Weyl quantization suggests that, instead of considering wave functions in $L^2(\mathbb{R})$, we should be able to consider functions on phase space, provided that we used the Moyal product instead of the usual product. This leads to the *phase space* or *Wigner formulation* of QM, see [3, 7] for some useful references. In this formulation, a state is described by Wigner's function (4.33), which is defined on the phase space $\mathbb{R}^2$. This function and its time evolution are completely determined

by a set of axioms. First of all, the analogue of time-dependent Schrödinger's equation is *Moyal's evolution equation*,

$$\mathrm{i}\hbar\frac{\partial f}{\partial t} = [H, f]_\star. \tag{4.64}$$

As $\hbar \to 0$, this becomes Liouville's equation for the time evolution of a density.

In the standard formulation of QM, a stationary state is an eigenstate of the Hamiltonian. In the phase space formulation, a stationary Wigner function satisfies the $\star$-genvalue equation,

$$H \star f = f \star H = E_f f. \tag{4.65}$$

$\star$-genfunctions corresponding to different energies are *orthogonal*: if

$$H \star f = f \star H = E_f f, \qquad H \star g = g \star H = E_g g, \tag{4.66}$$

with $E_f \neq E_g$, then

$$f \star g = 0. \tag{4.67}$$

Indeed, by associativity,

$$f \star H \star g = (f \star H) \star g = E_f f \star g = f \star (H \star g) = E_g f \star g, \tag{4.68}$$

and orthogonality follows.

Moyal's evolution equation is solved by the $\star$-unitary evolution, which is given by a $\star$-exponential,

$$U_\star(q, p; t) = \mathrm{e}_\star^{\mathrm{i}tH/\hbar} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\mathrm{i}t}{\hbar}\right)^n H \star \cdots \star H. \tag{4.69}$$

If we are given a Wigner function at $t = 0$, $f(q, p; 0)$, we have

$$f(q, p; t) = U_\star^{-1}(q, p; t) \star f(q, p; 0) \star U_\star(q, p; t). \tag{4.70}$$

We note that this operator satisfies the equation

$$-\mathrm{i}\hbar\frac{\partial U}{\partial t} = H \star U, \tag{4.71}$$

with initial condition

$$U(q, p; 0) = 1. \tag{4.72}$$

Example 4.7. *$\star$-genvalue equations for the harmonic oscillator.* We will now solve the $\star$-genvalue equations for the harmonic oscillator, with Hamiltonian

$$H(q, p) = \frac{1}{2}(q^2 + p^2). \tag{4.73}$$

We have set $m = \omega = 1$. Using (4.56), we find

$$H \star f = \left[\frac{1}{2}\left(p - \frac{\mathrm{i}\hbar}{2}\partial_q\right)^2 + \frac{1}{2}\left(q - \frac{\mathrm{i}\hbar}{2}\partial_p\right)^2\right] f(q, p) = E f(q, p). \tag{4.74}$$

The imaginary part of the equation gives

$$(p\partial_q - q\partial_p)\, f(q, p) = 0, \tag{4.75}$$

which means that $f(q,p)$ is a function of the "radial" coordinate only, $f(q,p) = f(z)$, where

$$z = 4H = 2(p^2 + q^2). \tag{4.76}$$

A simple calculation shows that the real part of the $\star$-genvalue equation becomes,

$$\left[\frac{z}{4} - \hbar^2 z^2 \partial_z^2 - \hbar^2 \partial_z - E\right] f(z) = 0. \tag{4.77}$$

We can rescale $\hbar$ away by setting $z = \hbar y$, $E = \hbar e$, and we find

$$\left[\frac{y}{4} - y^2 \partial_y^2 - \partial_y - e\right] f(y) = 0. \tag{4.78}$$

Set now

$$f(y) = \mathrm{e}^{-y/2}\phi(y). \tag{4.79}$$

Then, the ODE for $f(y)$ becomes,

$$\left(y\partial_y^2 + (1-y)\partial_y + e - \frac{1}{2}\right)\phi(y) = 0. \tag{4.80}$$

When compared to the differential equation for the confluent hypergeometric equation,

$$\left(y\partial_y^2 + (b-y)\partial_y - a\right)\phi(y) = 0, \tag{4.81}$$

we conclude that

$$b = 1, \qquad a = 1/2 - e. \tag{4.82}$$

There are two linearly independent solutions of this equation, usually denoted by $M(a,b,y)$ and $U(a,b,y)$. Therefore, we could have in principle

$$\phi(y) = c_1 M(a,b,y) + c_2 U(a,b,y). \tag{4.83}$$

We want $f(y)$ to decay at infinity, and also to be finite at $y = 0$. Given the behavior at infinity

$$M(a,b,y) \sim \mathrm{e}^y y^{a-b}, \qquad U(a,b,y) \sim y^{-a}, \tag{4.84}$$

we conclude that $c_1 = 0$. Also, since for $b = 1$ we have, for y small,

$$U(a,1,y) = -\frac{1}{\Gamma(a)}\left(\log(y) + \psi(a) + 2\gamma\right) + \mathcal{O}\left(y\log y\right), \tag{4.85}$$

we must have $a = -n$, $n = 0, 1, 2, \cdots$, in which case the confluent hypergeometric function becomes a Laguerre polynomial,

$$U(-n,1,y) = (-1)^n n! L_n(y). \tag{4.86}$$

This already leads to the usual quantization of the energy levels,

$$e = n + 1/2. \tag{4.87}$$

Example 4.8. *⋆-genvalue equations for the linear potential.* We will now solve the ⋆-genvalue equations for the linear potential. The Hamiltonian is

$$H = p^2 + x. \tag{4.88}$$

Let us solve the ⋆-genvalue equation. It reads,

$$\left[\left(q + \frac{\mathrm{i}}{2}\partial_p\right) + \left(p - \frac{\mathrm{i}}{2}\partial_q\right)^2 - E\right] f(q,p) = 0. \tag{4.89}$$

The imaginary part says that

$$\left(\frac{1}{2}\partial_p - p\partial_q\right) f(q,p) = 0, \tag{4.90}$$

i.e.

$$f(q,p) = f(H). \tag{4.91}$$

Set $u = H = q + p^2$. Then, the real part reads

$$\left(u - \frac{1}{4}\partial_u^2 - E\right) f(u) = 0 \Rightarrow f(u) = \frac{2^{2/3}}{2\pi}\mathrm{Ai}\left(2^{2/3}(u-E)\right), \tag{4.92}$$

or

$$f(q,p) = \frac{1}{(2\pi)^2}\int \mathrm{d}z\, \exp\left[\mathrm{i}z\left(E - q - p^2 - \frac{z^2}{12}\right)\right]. \tag{4.93}$$

Example 4.9. *Evolution operator for the harmonic oscillator.* Let us compute U for the harmonic oscillator. First of all, we note that a function of the Hamiltonian (4.73) satisfies, as we saw in (4.77),

$$H \star f(H) = Hf(H) - \frac{\hbar^2}{4}\left(f'(H) + Hf''(H)\right). \tag{4.94}$$

It is clear that U only depends on H, and it satisfies the equation

$$-\mathrm{i}\hbar\frac{\partial U}{\partial t} = HU - \frac{\hbar^2}{4}\left(U'(H) + HU''(H)\right), \tag{4.95}$$

with initial condition

$$U(H;0) = 1. \tag{4.96}$$

To solve this equation, we use the following ansatz,

$$U(H;t) = \exp\left[-a(t)H - b(t)\right]. \tag{4.97}$$

We deduce the following first order ODEs for $a(t)$, $b(t)$,

$$\mathrm{i}\hbar\dot{a}(t) = 1 - \frac{\hbar^2}{4}a(t), \qquad \mathrm{i}\hbar\dot{b}(t) = \frac{\hbar^2}{4}a(t), \tag{4.98}$$

with the initial conditions

$$a(0) = b(0) = 0, \tag{4.99}$$

They can be integrated immediately to give,

$$a(t) = -\frac{2\mathrm{i}}{\hbar}\tan\left(\frac{t}{2} + c_1\right), \qquad b(t) = \log\left[\cos\left(\frac{t}{2} + c_1\right)\right] + c_2. \tag{4.100}$$

and the initial conditions fix $c_1 = c_2 = 0$. We conclude that

$$U(H;t) = \frac{1}{\cos\left(\frac{t}{2}\right)}\exp\left[\frac{2\mathrm{i}}{\hbar}\tan\left(\frac{t}{2}\right)H\right]. \tag{4.101}$$

5 Semiclassical methods

In this final section we will discuss some approximate solutions to problems appearing in QM in the limit in which $\hbar$ is small. We will restrict ourselves to the analysis of bound states in one-dimensional problems.

5.1 The WKB method in one dimension

The time-independent Schrödinger equation (3.61), as any other ODE, is difficult to solve. It is then important then to have approximation methods in situations in which exact solutions are not available. In particular, we would like to have information about the energy spectrum. Let us first write (3.61) in the form

$$\hbar^2\psi''(x) + p^2(x,E)\psi(x) = 0, \qquad p(x,E) = \sqrt{2(E-V(x))}, \tag{5.1}$$

where we set $m = 1$ for simplicity. If we write the wavefunction as

$$\psi(x) = \exp\left[\frac{\mathrm{i}}{\hbar}\int^x Y(x')\mathrm{d}x'\right] \tag{5.2}$$

we transform the Schrödinger equation into a Riccati equation

$$Y^2(x) - \mathrm{i}\hbar\frac{\mathrm{d}Y(x)}{\mathrm{d}x} = p^2(x,E), \tag{5.3}$$

which we solve in power series in $\hbar$:

$$Y(x,E,\hbar) = \sum_{k=0}^{\infty} Y_k(x,E)\hbar^k. \tag{5.4}$$

We will regard this as a *formal* power series, i.e. we will not address issues of convergence. The functions $Y_k(x,E)$ can be computed recursively as

$$\begin{aligned} Y_0(x,E) &= p(x,E), \\ Y_{n+1}(x,E) &= \frac{1}{2Y_0(x,E)}\left(\mathrm{i}\frac{\mathrm{d}Y_n(x,E)}{\mathrm{d}x} - \sum_{k=1}^{n} Y_k(x,E)Y_{n+1-k}(x,E)\right). \end{aligned} \tag{5.5}$$

If we split $Y(x,E,\hbar)$ into even and odd powers of $\hbar$,

$$Y(x,E,\hbar) = Y_{\mathrm{odd}}(x,E,\hbar) + P(x,E,\hbar^2), \tag{5.6}$$

we find that (5.3) splits into two different equations:

$$\begin{aligned} &\text{even:} \quad Y_{\mathrm{odd}}^2 + P^2 - \mathrm{i}\hbar Y_{\mathrm{odd}}' = p^2, \\ &\text{odd:} \quad 2Y_{\mathrm{odd}}P - \mathrm{i}\hbar P' = 0. \end{aligned} \tag{5.7}$$

and the second equation gives

$$Y_{\mathrm{odd}}(x,E,\hbar) = \frac{\mathrm{i}\hbar}{2}\frac{P'(x,E,\hbar^2)}{P(x,E,\hbar^2)} = \frac{\mathrm{i}\hbar}{2}\frac{\mathrm{d}}{\mathrm{d}x}\log P(x,E,\hbar^2). \tag{5.8}$$

Therefore,

$$\frac{\mathrm{i}}{\hbar}\int^x Y(x')\mathrm{d}x' = -\frac{1}{2}\log P(x,E,\hbar^2) + \frac{\mathrm{i}}{\hbar}\int^x P(x',E,\hbar^2)\mathrm{d}x' \tag{5.9}$$

and the wavefunction reads

$$\psi(x,E,\hbar) = \frac{1}{\sqrt{P(x,E,\hbar^2)}}\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\int^x P(x',E,\hbar^2)\mathrm{d}x'}. \tag{5.10}$$

Notice that the very first solutions to the recursive equation are

$$\begin{aligned} Y_0(x,E) &= p(x,E), \\ Y_1(x,E) &= \frac{\mathrm{i}}{2}\frac{Y_0'}{Y_0}, \\ Y_2(x,E) &= \frac{3}{8}\frac{(Y_0')^2}{Y_0^3} - \frac{1}{4}\frac{Y_0''}{Y_0^2}. \end{aligned} \tag{5.11}$$

Let us now consider a situation in which $V(x)$ is analytic when x is extended to the complex domain. By the elementary theory of ODEs in the complex plane, it follows that the function $\psi(x)$ is also analytic. Let us also assume that the potential $V(x)$ is confining, i.e.

$$V(x) \to \infty, \qquad |x| \to \infty. \tag{5.12}$$

In this case, it can be shown that the Schrödinger equation has a pure point spectrum of energies

$$E_0 < E_1 < \cdots. \tag{5.13}$$

In addition, the corresponding eigenfunctions ψ_n have n zeroes on the real line. These zeroes are located along the interval of allowed classical motion with the energy E_n.

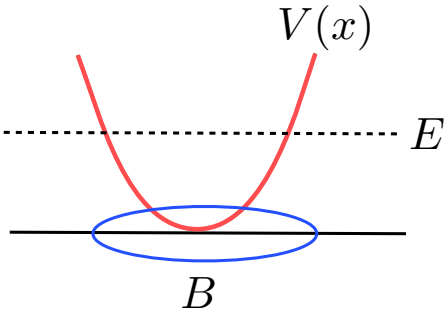

Figure 1. A contour encircling the classical turning points.

Given this information, an approximate formulae for the energies E_n can be obtained as follows. Since ψ is analytic, and due to the argument theorem, we have

$$\frac{1}{2\pi\mathrm{i}}\oint_B \frac{\psi_n'(z)}{\psi_n(z)}\mathrm{d}z = n, \qquad n = 0,1,2,\cdots, \tag{5.14}$$

where B is a contour around the interval of allowed classical motion for the corresponding energy, see Fig. 1. The two points where $E = V(x)$, which we will denote by $x_{1,2}$, are called the *turning points* of the classical motion. Using the ansatz (5.2), we conclude that

$$\oint_B Y(z)\mathrm{d}z = 2\pi n\hbar. \tag{5.15}$$

To understand this condition, let us consider the approximation in which we keep both the first and the second terms, i.e.

$$Y(z) \approx Y_0(z) + \hbar Y_1(z). \tag{5.16}$$

Then, we have

$$\oint_B Y_0(z)\mathrm{d}z + \frac{\mathrm{i}\hbar}{2}\oint_B \frac{Y_0'}{Y_0}\mathrm{d}z + \cdots = 2\pi n\hbar. \tag{5.17}$$

The second term can be evaluated as follows. Note that $Y_0(z)$ is not analytic, but

$$\frac{Y_0'}{Y_0} = \frac{\mathrm{d}}{\mathrm{d}z}\log Y_0 = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}z}\log\left(2(E-V(z))\right) = \frac{1}{2}\frac{\sigma'(z)}{\sigma(z)}, \tag{5.18}$$

where $\sigma(z) = 2(E - V(z))$, an analytic function which has two zeroes inside B, namely, the two turning points $x_{1,2}$. By applying the theorem of the argument to $\sigma(z)$, we so that

$$\frac{\mathrm{i}\hbar}{2}\oint_B \frac{Y_0'(z)}{Y_0(z)}\mathrm{d}z = \frac{\mathrm{i}\hbar}{4}\oint_B \frac{\sigma'(z)}{\sigma(z)}\mathrm{d}z = \frac{\mathrm{i}\hbar}{4}(4\pi\mathrm{i}) = -\pi\hbar, \tag{5.19}$$

and we find, at this order of approximation,

$$\oint_B Y_0(z)\mathrm{d}z \approx 2\pi\hbar\left(n+\frac{1}{2}\right), \tag{5.20}$$

which can be also written as

$$\int_{x_1}^{x_2}\sqrt{2(E-V(x))}\mathrm{d}x \approx \pi\hbar\left(n+\frac{1}{2}\right), \qquad n = 0,1,2,\cdots \tag{5.21}$$

This is the famous *Bohr–Sommerfeld quantization condition*, which gives the leading approximation to the spectrum of energies when the classical action appearing in the l.h.s. is large as compared to $\hbar \to$. This happens when $\hbar$ is small but also when n is large. The Bohr–Sommerfeld quantization condition leads to a constraint relating the energy E to the quantum number, whose solution gives the energy levels $E(n) \equiv E_n$.

The condition (5.15) involves higher order corrections to the Bohr–Sommerfeld quantization condition. Note that Y_{odd} only contributes through its first term, since

$$Y_{\mathrm{odd}}(z) = \frac{\mathrm{i}\hbar}{2}\frac{Y_0'(z)}{Y_0(z)} + \sum_{n=1}^{\infty} f_n'(z), \tag{5.22}$$

where $f_n(z)$ is a rational function, and the integration of the total derivatives along the B-cycle gives a vanishing contribution. We find then an all-orders quantization condition

$$\sum_{n=0}^{\infty}\hbar^{2n}\oint_B Y_{2n}(z)\mathrm{d}z = 2\pi\hbar\left(n+\frac{1}{2}\right), \qquad n = 0,1,2,\cdots \tag{5.23}$$

which is due to Dunham [4].

Example 5.1. The next-to-leading correction is given by

$$\begin{aligned}\oint_B \mathrm{d}x\, Y_2(x,E) &= \oint_B \mathrm{d}x\left(\frac{3}{8}\frac{(Y_0')^2}{Y_0^3} - \frac{1}{4}\frac{Y_0''}{Y_0^2}\right) = -\frac{1}{8}\oint_B \mathrm{d}x\frac{(Y_0')^2}{Y_0^3} \\ &= -\frac{1}{8}\oint_B \mathrm{d}x\frac{(V'(x))^2}{(2(E-V(x)))^{5/2}} = \frac{1}{24}\oint_B \mathrm{d}x\frac{V''(x)}{(2(E-V(x)))^{3/2}} \\ &= -\frac{1}{24}\frac{\mathrm{d}}{\mathrm{d}E}\oint_B \mathrm{d}x\frac{V''(x)}{(2(E-V(x)))^{1/2}}.\end{aligned} \tag{5.24}$$

Example 5.2. *WKB for the harmonic oscillator.* Let us consider $V(x) = \omega^2 x^2/2$. Then, the integral in the l.h.s. of the Bohr–Sommerfeld quantization condition (5.21) is

$$\int_{-x_*}^{x_*} \sqrt{2(E-\omega^2 x^2/2)}\mathrm{d}x = \frac{2E}{\omega}\int_{-1}^{1}\sqrt{1-u^2}\mathrm{d}u = \frac{\pi E}{\omega}, \tag{5.25}$$

where we performed the change of variables

$$u = \frac{\omega x}{\sqrt{2E}}. \tag{5.26}$$

At this order, we obtain

$$\frac{\pi E}{\omega} \approx \hbar\pi\left(n+\frac{1}{2}\right), \tag{5.27}$$

which is already the exact answer. In fact, the first correction to the Bohr–Sommerfeld quantization condition involves,

$$-\frac{\omega^2}{12}\frac{\mathrm{d}}{\mathrm{d}E}\oint_{-x_*}^{x_*}\frac{\mathrm{d}x}{(2(E-\omega^2x^2/2))^{1/2}} = -\frac{\omega}{12}\frac{\mathrm{d}}{\mathrm{d}E}\int_{-1}^{1}\frac{1}{\sqrt{1-u^2}} = 0. \tag{5.28}$$

Therefore, in the case of the harmonic oscillator, the leading WKB approximation is already exact, and further corrections vanish. This is of course quite exceptional.

5.2 The WKB method in the Weyl–Wigner quantization

An alternative, and more general point of view on the all-orders WKB quantization condition of Dunham, has been developed in [5, 10] in the context of Weyl–Wigner quantization.

Given any function $f(\mathsf{H})$ of H, we can expand it around $H_{\mathrm{W}}(q,p)$, which is a c-number, as [5]

$$f(\mathsf{H}) = \sum_{r\geq 0}\frac{1}{r!}f^{(r)}(H_{\mathrm{W}})\left(\mathsf{H}-H_{\mathrm{W}}(q,p)\right)^r. \tag{5.29}$$

The semiclassical expansion of this object is obtained simply by evaluating its Wigner transform, and we obtain

$$f(\mathsf{H})_{\mathrm{W}} = \sum_{r\geq 0}\frac{1}{r!}f^{(r)}\left(H_{\mathrm{W}}\right)\mathcal{G}_r \tag{5.30}$$

where

$$\mathcal{G}_r = [(\mathsf{H}-H_{\mathrm{W}}(q,p))^r]_{\mathrm{W}} \tag{5.31}$$

and the Wigner transform is evaluated at the *same* point q, p. The quantities $\mathcal{G}_r$ can be computed by using (4.39). One finds, for the first orders $\mathcal{G}_0 = 1$, $\mathcal{G}_1 = 0$, and

$$\begin{aligned}
\mathcal{G}_2 &= -\frac{\hbar^2}{4}\left[\frac{\partial^2 H_{\mathrm{W}}}{\partial q^2}\frac{\partial^2 H_{\mathrm{W}}}{\partial p^2} - \left(\frac{\partial^2 H_{\mathrm{W}}}{\partial q\partial p}\right)^2\right] + \mathcal{O}(\hbar^4),\\
\mathcal{G}_3 &= \frac{\hbar^2}{4}\left[\left(\frac{\partial H_{\mathrm{W}}}{\partial q}\right)^2\frac{\partial^2 H_{\mathrm{W}}}{\partial p^2} + \left(\frac{\partial H_{\mathrm{W}}}{\partial p}\right)^2\frac{\partial^2 H_{\mathrm{W}}}{\partial q^2} - 2\frac{\partial H_{\mathrm{W}}}{\partial q}\frac{\partial H_{\mathrm{W}}}{\partial p}\frac{\partial^2 H_{\mathrm{W}}}{\partial q\partial p}\right] + \mathcal{O}(\hbar^4).
\end{aligned} \tag{5.32}$$

One can then apply this method to compute the semiclassical expansion of any function of the Hamiltonian operator. We will now apply this to an operator which counts the number of states up to given energy. This is the operator

$$\mathsf{N}(E) = \theta(E-\mathsf{H}) \tag{5.33}$$

where $\theta(x)$ is the Heaviside step function. The trace of this operator gives the function $n(E)$ counting the number of eigenstates whose energy is less than E:

$$n(E) = \mathrm{Tr}\,\mathsf{N}(E) = \sum_{n=0}^{\infty} \langle \psi_n, \theta(E - \mathsf{H})\psi_n \rangle = \sum_{n=0}^{\infty} \theta(E - E_n) \tag{5.34}$$

We can now apply the Wigner transform formula (5.30) to perform the semiclassical expansion of the operator (5.33). One finds,

$$\mathsf{N}(E)_{\mathrm{W}} = \theta(E - H_{\mathrm{W}}) + \sum_{r=2}^{\infty} \frac{(-1)^r}{r!} \mathcal{G}_r \delta^{(r-1)}(E - H_{\mathrm{W}}). \tag{5.35}$$

We can now compute the trace (5.34) by using the formula

$$\begin{aligned} n(E) &= \int_{H_{\mathrm{W}}(q,p) \le E} \frac{\mathrm{d}q\mathrm{d}p}{2\pi\hbar} + \sum_{r=2}^{\infty} \frac{(-1)^r}{r!} \int \frac{\mathrm{d}q\mathrm{d}p}{2\pi\hbar} \mathcal{G}_r \delta^{(r-1)}(E - H_{\mathrm{W}}) \\ &= \int_{H_{\mathrm{W}}(q,p) \le E} \frac{\mathrm{d}q\mathrm{d}p}{2\pi\hbar} + \sum_{r=2}^{\infty} \frac{(-1)^r}{r!} \frac{\mathrm{d}^r}{\mathrm{d}E^r} \int \frac{\mathrm{d}q\mathrm{d}p}{2\pi\hbar} \mathcal{G}_r \theta(E - H_{\mathrm{W}}). \end{aligned} \tag{5.36}$$

As a formal power series expansion in $\hbar$, it agrees with the l.h.s. of (5.23).

References

[1] V. I. Arnold, *Mathematical methods of classical mechanics*, second edition, Springer-Verlag, 1989.

[2] C. M. Bender and S. Orszag, *Advanced mathematical methods for scientists and engineers*, Springer.

[3] T. Curtright, D. Fairlie and C. K. Zachos, "Features of time independent Wigner functions," Phys. Rev. D **58**, 025002 (1998) [hep-th/9711183].

[4] J. L. Dunham,"The Wentzel-Brillouin-Kramers method of solving the wave equation." *Phys. Rev.* **41**, 713 (1932).

[5] B. Grammaticos and A. Voros, "Semiclassical Approximations For Nuclear Hamiltonians. 1. Spin Independent Potentials," Annals Phys. **123**, 359 (1979).

[6] B. C. Hall, *Quantum theory for mathematicians*, Springer-Verlag, 2013.

[7] M. Hillery, R. F. O'Connell, M. O. Scully and E. P. Wigner, "Distribution functions in physics: Fundamentals," Phys. Rept. **106**, 121-167 (1984).

[8] E. K. Sklyanin, "The Quantum Toda Chain," Lect. Notes Phys. **226**, 196 (1985).

[9] L. Tahktajan, *Quantum Mechanics for mathematicians*, American Mathematical Society, 2008.

[10] A. Voros, "Asymptotic $\hbar$-expansions of stationary quantum states," Ann. Inst. H. Poincaré Sect. A (N.S.) **26** (1977), no. 4, 343-403.