H. Hergert
Facility for Rare Isotope Beams
and Department of Physics & Astronomy

MICHIGAN STATE UNIVERSITY

PHY422/820: Classical Mechanics

FS 2020

Worksheet #15 (Dec 7 – Dec 11)

December 11, 2020

1 Reading

- Lemos, Chapter 9.1-9.6 (9.4 optional), 10.1-10.4 (10.5 optional), 11.1-11.5
- Goldstein, Chapter 10.1-10.5 (10.6-10.8 optional), 12, 13

2 Hamilton-Jacobi Theory and Action-Angle Variables

2.1 The Hamilton-Jacobi Equation

Canonical transformations offer us a great deal of freedom that can be used to simplify the process of solving the equations of motion of a dynamical system. We have used one possible strategy, which is to map a given Hamiltonian onto one we know to solve. Hamilton-Jacobi theory explores another approach, namely to design a canonical transformation that will make all variables *cyclic*. If we can make the new Hamiltonian K independent of (Q, P), then

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} = 0\,, \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0\,. \tag{1}$$

and

$$Q_i = \text{const.}, \quad P_i = \text{const.} \tag{2}$$

We could try to generate a constant K, but the simplest approach is to actually look for $K = 0$. Using a generating function of the type $F = F_2(q, P, t)$, we will have

$$K = H(q,p,t) + \frac{\partial F_2}{\partial t} \stackrel{!}{=} 0 \tag{3}$$

and since

$$p_i = \frac{\partial F_2}{\partial q_i}\,, \tag{4}$$

we can write the condition (3) as

$$H\left(q_1, \ldots, q_n, \frac{\partial F_2}{\partial q_1}, \ldots, \frac{\partial F_2}{\partial q_n}, t\right) + \frac{\partial F_2}{\partial t} = 0\,. \tag{5}$$

which is the **time-dependent Hamilton-Jacobi equation**. We see that it is a first-order differential equation for F_2 in the variables $(q_1, \ldots, q_n, t)$. The solution for F_2 will depend on $n+1$ independent constants of integration $\alpha_1, \ldots, \alpha_{n+1}$. Since Eq. (5) only depends on derivatives of F_2, it is invariant under the shift

$$F_2 \longrightarrow F_2 + \alpha_{n+1} \,, \tag{6}$$

one of the constants — here chosen to be α_{n+1} — can be ignored. Let us assume that the solution is

$$F_2 \equiv S = S\left(q_1, \ldots, q_n, \alpha_1, \ldots \alpha_n, t\right) \tag{7}$$

where S is called **Hamilton's principal function**. Since our goal was to make variables cyclic, we can choose associated conserved momenta as our independent constants:

$$P_i = \alpha_i = \text{const.}\,. \tag{8}$$

This will specify $F_2 = F_2(q, P, t) = F_2(q, \alpha, t)$ as required. For a generating function of this type, the new coordinates are given by (cf. worksheet #14)

$$\beta_i \equiv Q_i = \frac{\partial S(q, \alpha, t)}{\partial \alpha_i} \,, \tag{9}$$

where the notation β_i is meant to emphasize that these are constants as well. To solve for the motion, we invert Eq. (9) to obtain

$$\beta_i = \frac{\partial S}{\partial \alpha_i} \quad \Rightarrow \quad q_i = q_i(\alpha, \beta, t) \,, \tag{10}$$

and plug it into the partial derivative linking p_i to S (cf. worksheet #14):

$$p_i = \frac{\partial S}{\partial q_i} \quad \Rightarrow \quad p_i = p_i(\alpha, \beta, t) \,. \tag{11}$$

Let us now try to interpret the function S. We know that

$$\dot{S} = \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial P_i}\dot{P}_i + \frac{\partial S}{\partial t} \tag{12}$$

as well as

$$\frac{\partial S}{\partial q_i} = p_i \,, \quad \dot{P}_i = 0 \,, \quad \frac{\partial S}{\partial t} = -H \,. \tag{13}$$

Compining these relations, we find

$$\dot{S} = p_i \dot{q}_i - H = L \quad \Rightarrow \quad S = \int dt \; L \,. \tag{14}$$

Thus, S is the action! Whenever it is compatible with the assumptions for a generating function, the action generates a canonical transformation that brings the state of the system at a general time t to its state at some fixed initial time t_0. In that case, we may also with swap our $2n$ constants α_i and β_i for the initial conditions q_{i0} and p_{i0} by solving the equations obtained above at $t = t_0$:

$$q_{i0} = q_i\left(\alpha, \beta, t_0\right) \,, \quad p_{i0} = p_i\left(\alpha, \beta, t_0\right) \,. \tag{15}$$

2.2 Separation of Variables and the Time-Independent Hamilton-Jacobi Equation

The structure of the Hamilton-Jacobi equation makes it amenable to a separation of variables whenever the Hamiltonian does not explicitly depend on one or more of the canonical variables.

Cyclic Variables

Let us assume that q_n is cyclic. Then we will have

$$\dot{p}_n = \frac{\partial H}{\partial q_n} = 0\,, \quad p_n = \alpha_n\,, \tag{16}$$

so we do not have to do any work to turn p_n into a constant of the motion. and we only need to apply an identity transformation to the pair (q_n, p_n). The generating function for the identity transformation of these variables is

$$F_2 = q_n P_n = q_n \alpha_n\,, \tag{17}$$

which is of the same type as S (cf. worksheet #14). This implies that we can make the following ansatz for Hamilton's principal function:

$$S = \alpha_n q_n + \overline{S}(q_1, \ldots, q_{n-1}) \tag{18}$$

where

$$\alpha_n = \frac{\partial S}{\partial q_n}\,, \tag{19}$$

as required, and $\overline{S}$ will satisfy

$$H\left(q_1, \ldots, q_{n-1}, \frac{\partial \overline{S}}{\partial q_1}, \ldots, \frac{\partial \overline{S}}{\partial q_{n-1}}, \alpha_n, t\right) + \frac{\partial \overline{S}}{\partial t} = 0\,. \tag{20}$$

This result is readily generalized: If $q_{k+1}, \ldots, q_n$ are cyclic, S will have the form

$$S = \sum_{i=k+1}^{n} \alpha_i q_i + \overline{S}(q_1, \ldots, q_{k-1})\,, \tag{21}$$

and

$$H\left(q_1, \ldots, q_k, \frac{\partial \overline{S}}{\partial q_1}, \ldots, \frac{\partial \overline{S}}{\partial q_k}, \alpha_{k+1}, \ldots, \alpha_n, t\right) + \frac{\partial \overline{S}}{\partial t} = 0\,. \tag{22}$$

Time-Independent Hamiltonians

If the Hamiltonian does not explicitly depend on time, we have

$$\dot{H} = \frac{\partial H}{\partial t} = 0\,, \quad \Rightarrow \quad H = \text{const.} \tag{23}$$

The Hamilton-Jacobi equation becomes

$$H\left(q, \frac{\partial S}{\partial q}\right) + \frac{\partial S}{\partial t} = 0 \tag{24}$$

which means

$$\alpha_1 \equiv H = -\frac{\partial S}{\partial t}\,. \tag{25}$$

Upon integration, this implies

$$S(q,\alpha,t) = W(q,\alpha) - \alpha_1 t\,, \tag{26}$$

where we have dropped an additional irrelevant constant, as explained in the previous section. The function $W(q,\alpha)$ is known as **Hamilton's characteristic function**. It satisfies the **time-independent Hamilton-Jacobi equation**

$$H\left(q, \frac{\partial W}{\partial q}\right) = \alpha_1\,. \tag{27}$$

If we apply the same strategy as in the time-dependent case, we obtain the equations of motion by using

$$p_i = \frac{\partial W}{\partial q_i}\,, \qquad Q_1 \equiv \beta_1 = \frac{\partial S}{\partial \alpha_1} = \frac{\partial W}{\partial \alpha_1} - t\,, \qquad Q_j \equiv \beta_j = \frac{\partial W}{\partial \alpha_j} \quad \text{for } j > 1\,. \tag{28}$$

Alternatively, we can directly consider $W = F_2(q,P)$ as the generating function and avoid any reference to $S(q,\alpha,t)$ altogether. Just like before, we will have

$$p_i = \frac{\partial W}{\partial q_i}\,, \qquad P_i = \alpha_i\,, \qquad H\left(q, \frac{\partial W}{\partial q}\right) = \alpha_1\,, \tag{29}$$

but we will have a non-zero K and a different equation for Q_1:

$$K = \alpha_1 = H\,, \tag{30}$$

$$Q_i = \frac{\partial W}{\partial \alpha_i}\,. \tag{31}$$

As a consequence, the equation of motion for Q_1 is modified:

$$\dot{Q}_1 = \frac{\partial K}{\partial \alpha_1} = 1 \quad \Rightarrow \quad Q_1 = t + \beta_1 = \frac{\partial W}{\partial \alpha_1}\,. \tag{32}$$

Of course, this is just a re-definition of Eq. (28) that allows an explicit linear time dependence. The equations for the remaining coordinates are the same as before:

$$\dot{Q}_j = \frac{\partial K}{\partial \alpha_j} = 0 \quad \Rightarrow \quad Q_j = \beta_j = \frac{\partial W}{\partial \alpha_j}\,. \tag{33}$$

2.3 Transition to Quantum Mechanics

The time-dependent Hamilton-Jacobi equation can be understood as the leading-order term in an effective theory of quantum mechanics for $\hbar \to 0$ (or $\hbar\omega/E \ll 1$, the so-called **Wenzel-Kramers-Brillouin** or **WKB approximation**). To see this, we consider the Schrödinger equation

$$i\hbar\frac{\partial \psi}{\partial t} = H\psi = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q)\right)\psi\,. \tag{34}$$

We write the wave function as

$$\psi = \exp\left(\frac{i}{\hbar}S\right)\,, \tag{35}$$

where $S(q,t)$ is complex, and plug it into Eq. (34), obtaining

$$-\frac{\partial S}{\partial t} = -\frac{i\hbar}{2m}\frac{\partial^2 S}{\partial q^2} + \frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + V(q)\,. \tag{36}$$

If we now take $\hbar \to 0$, we have

$$0 = \frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + V(q) = \frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}\right)\,, \tag{37}$$

which is the Hamilton-Jacobi equation for S if the Hamiltonian has the usual form

$$H = \frac{p^2}{2m} + V(q)\,. \tag{38}$$

2.4 Examples

2.4.1 Harmonic Oscillator

Let us study our trusty harmonic oscillator using the time-dependent Hamilton-Jacobi equation. The Hamiltonian is given by

$$H = \frac{p^2}{2m}p^2 + \frac{1}{2}m\omega^2 q^2 = E\,. \tag{39}$$

Here we will look for one constant $P = \alpha$ and one constant $Q = \beta$. The Hamilton-Jacobi equation reads

$$\frac{1}{2m}\left(\left(\frac{\partial S}{\partial q}\right)^2 + (m\omega q)^2\right) + \frac{\partial S}{\partial t} = 0\,, \tag{40}$$

and we can make the separation ansatz

$$S(q,\alpha,t) = W(q,\alpha) - \alpha t\,, \tag{41}$$

as discussed in Sec. 2.2. The Hamilton-Jacobi equation now implies

$$H = -\frac{\partial S}{\partial t} = \alpha \overset{!}{=} E\,, \tag{42}$$

hence we have to solve

$$\frac{1}{2m}\left(\left(\frac{\partial W}{\partial q}\right)^2 + (m\omega q)^2\right) = E\,. \tag{43}$$

Rearranging and integrating, we have

$$W = \pm \int dq\,\sqrt{2mE - (m\omega q)^2} \tag{44}$$

and Hamilton's principal function is given by

$$S = -Et \pm \int dq\,\sqrt{2mE - (m\omega q)^2}\,. \tag{45}$$

Now we can derive the equations of motion:

$$\beta = \frac{\partial S}{\partial E} = -t \pm m \int \frac{dq}{\sqrt{2mE - (m\omega q)^2}} \tag{46}$$

This integral starts to look very familiar. Evaluating it, we obtain

$$t + \beta = \pm \frac{1}{\omega} \arcsin\left(\sqrt{\frac{m\omega^2}{2E}}\, q\right), \tag{47}$$

and inverting, we have

$$q(t) = \pm\sqrt{\frac{2E}{m\omega^2}} \sin(\omega(t+\beta)) \tag{48}$$

so β is related to the phase. Using these results, we can evaluate p:

$$p = \frac{\partial S}{\partial q} = \pm\sqrt{2mE - (m\omega q)^2} = \pm\sqrt{2mE}\cos(\omega(t+\beta))\,. \tag{49}$$

These results are exactly what we would expect, of course. We can switch from (α, β) to the initial conditions (q_0, p_0) at $t = 0$, and choose β such that we only need to consider the positive branch of the solutions.

2.4.2 The Kepler Problem

As another example, we solve the Kepler problem using the Hamilton-Jacobi method. The Hamiltonian is given by

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\phi^2}{r^2}\right) + V(r) = \alpha_1 = E\,. \tag{50}$$

Since ϕ is cyclic, we have another constant of the motion:

$$p_\phi \equiv \alpha_2\,. \tag{51}$$

Making the ansatz

$$W = W_1(r) + \alpha_2 \phi\,, \tag{52}$$

the time-independent Hamilton-Jacobi equation becomes

$$\frac{1}{2m}\left(\left(\frac{\partial W_1}{\partial r}\right)^2 + \frac{\alpha_2^2}{r^2}\right) + V(r) = \alpha_1\,. \tag{53}$$

Rearranging, we have

$$\frac{\partial W_1}{\partial r} = \sqrt{2m\left(\alpha_1 - V(r)\right) - \frac{\alpha_2^2}{r^2}} \tag{54}$$

and integration yields

$$W = \alpha_2 \phi + \int dr\, \sqrt{2m\left(\alpha_1 - V(r)\right) - \frac{\alpha_2^2}{r^2}}\,. \tag{55}$$

The transformation equations are:

$$t + \beta_1 = \frac{\partial W}{\partial \alpha_1} = m \int \frac{dr}{\sqrt{2m\left(\alpha_1 - V(r)\right) - \frac{\alpha_2^2}{r^2}}}\,, \tag{56}$$

$$\beta_2 = \frac{\partial W}{\partial \alpha_2} = \phi - \alpha_2 \int \frac{dr}{r^2\sqrt{2m\left(\alpha_1 - V(r)\right) - \frac{\alpha_2^2}{r^2}}}\,. \tag{57}$$

We immediately obtain the radial and orbital equations $t = t(r)$ and $\phi = \phi(r)$, with $\alpha_1 = E$ and $\alpha_2 = l$.

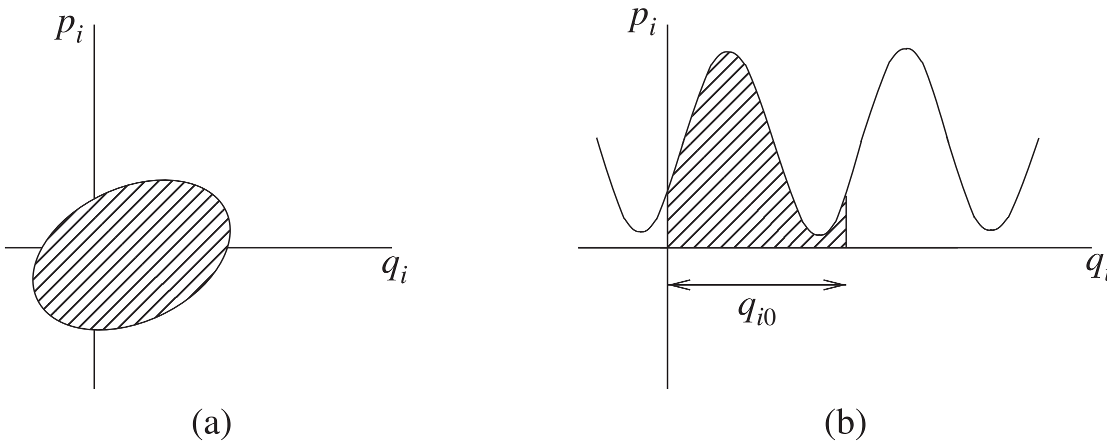

Figure 1: Action-angle variables: Libration (oscillation, (a)) and rotation (b) in phase space. Areas under the curves indicate the action variables J.

2.5 Action-Angle Variables

Periodic systems are of great importance in practically all branches of physics. Hamilton-Jacobi theory provides a powerful method for finding the frequencies of such systems that does not require the detailed solution of the equations of motion. The key is to use canonical transformations to introduce the so-called **action-angle variables**.

2.5.1 Definitions

A system whose Hamiltonian does not explicitly depend on time is called **separable** if for some set of generalised coordinates $q_1, \ldots, q_n$ Hamilton's characteristic function can be written in the form

$$W\left(q_1, \ldots, q_n, \alpha_1, \ldots, \alpha_n\right) = W_1\left(q_1, \alpha_1, \ldots, \alpha_n\right) + \cdots + W_n\left(q_n, \alpha_1, \ldots, \alpha_n\right) \tag{58}$$

For such a system, we have

$$p_i = \frac{\partial S}{\partial q_i} = \frac{\partial W_i}{\partial q_i} = f_i\left(q_i, \alpha_1, \ldots, \alpha_n\right) \tag{59}$$

This last equation is the projection of the system's motion in phase space on the (q_i, p_i) phase plane.

A separable Hamiltonian system is said to be **multiply periodic** if the projection of the system's motion on these phase planes fits into one of the following categories (cf. worksheet #13):

1. The system undergoes **oscillation** or **libration**. The curve $p_i = p_i\left(q_i, \alpha\right)$ is closed – that is, q_i oscillates between two definite limits $q_i = a_i$, and $q_i = b_i$, as in Fig. 2(a).

2. p_i is a periodic function of q_i, with period q_{i0}, although q_i is not a periodic function of time, as in Fig. 2(b). Usually, this is referred to as **rotation**, because it occurs when q_i is an angular coordinate. Upon each complete turn, q changes by 2π and the state of the system repeats itself.

For multiply periodic systems it is possible to make use of action-angle variables to calculate the frequencies associated with the motion without solving Hamilton's equations. If the system has n degrees of freedom, the **action variables** are defined by

$$J_i = \frac{1}{2\pi} \oint p_i dq_i, \quad i = 1, \ldots, n \, , \tag{60}$$

where the integrals are extended over a period of libration or rotation. Geometrically, $2\pi J_i$ represents either of the dashed areas in Fig. 1. According to Eq. (59), the J_i are functions of the α_i. Conversely, we have

$$\alpha_i = \alpha_i \left(J_1, \ldots, J_n \right) \, . \tag{61}$$

Using these relations, we can express Hamilton's characteristic function as $W\left(q, J\right)$, and identify the constants J_i instead of the α_i with the new momenta. For the canonical transformation generated by $W(q, J)$, the transformed Hamiltonian becomes

$$K = H = \alpha_1 = \alpha_1 \left(J_1, \ldots, J_n \right) \equiv H \left(J_1, \ldots, J_n \right) \, , \tag{62}$$

which is simply the original Hamiltonian H expressed as a function of the action variables. The **angle variables** ϕ_i are the canonical conjugates to J_i

$$\phi_i = \frac{\partial W}{\partial J_i} \, , \tag{63}$$

and their equations of motion are

$$\dot{\phi}_i = \frac{\partial H}{\partial J_i} = \omega_i \tag{64}$$

where the frequencies $\omega_i = \omega_i \left(J_1, \ldots, J_n \right)$ are also constants because they can only depend on the constants J_i. Thus, we can immediately solve the equations of motion (64) to obtain

$$\phi_i(t) = \phi_i(0) + \omega_i t \, . \tag{65}$$

2.5.2 Fundamental Frequencies

In order to interpret the physical meaning of the ωs appearing in Eq. (64), we assume that the motion is periodic with period τ. Evidently, the projection of the motion on each (q_i, p_i) phase plane is also periodic and the ratios of the corresponding frequencies are rational numbers. In other words, after a time τ each canonical variable will have performed an integer number of complete cycles. The corresponding change in each angular variable is due to the variation of the coordinates q_i, since the J_i are constants. Therefore, in a period of the motion in phase space, we must have

$$\Delta\phi_i = \oint \sum_k \frac{\partial \phi_i}{\partial q_k} dq_k = \oint \sum_k \frac{\partial^2 W}{\partial q_k \partial J_i} dq_k = \frac{\partial}{\partial J_i} \oint \sum_k \frac{\partial W}{\partial q_k} dq_k \tag{66}$$

where we have used the definition of the ϕ_i and inverted the order of differentiation and integration. Using Eqs. 58 and (59), we can rewrite this as

$$\Delta\phi_i = \frac{\partial}{\partial J_i} \oint \sum_k \frac{\partial W_k}{\partial q_k} dq_k = \frac{\partial}{\partial J_i} \sum_k \oint p_k dq_k \tag{67}$$

If n_k is the number of complete cycles performed by the coordinate q_k in period τ, we have

$$\Delta\phi_i = \frac{\partial}{\partial J_i} \sum_k n_k 2\pi J_k = 2\pi n_i \tag{68}$$

Each variable ϕ_i increases by an integer multiple of 2π in a period of the motion, which justifies considering it an angle. On the other hand,

$$\tau = n_i \tau_i \tag{69}$$

where τ_i is the period associated with the ith degree of freedom. Finally, from Eq. (65), we obtain

$$\Delta\phi_i = \omega_i \tau \,. \tag{70}$$

We can infer

$$\omega_i \tau_i = 2\pi, \quad i = 1, \ldots, n \tag{71}$$

so that the ω_i are the **fundamental frequencies** of the system, i.e., the frequencies of the periodic motion executed by each degree of freedom. The partial derivatives of the Hamiltonian with respect to the action variables yield the fundamental frequencies.

2.5.3 The Harmonic Oscillator

As a first example, we again consider the harmonic oscillator. The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2 \equiv \alpha \,. \tag{72}$$

Rearranging the Hamiltonian, we have

$$p = \frac{\partial W}{\partial q} = \pm\sqrt{2m\omega J - (m\omega q)^2} \,. \tag{73}$$

Now we introduce the action-angle variables (ϕ, J). The action variable is

$$\begin{aligned} 2\pi J &= \oint p\, dq = 2\int_{-q_0}^{q_0} dq\, \sqrt{2m\alpha - (m\omega q)^2} = 2m\omega \int_{-q_0}^{q_0} dq\, \sqrt{\frac{2\alpha}{m\omega^2} - q^2} \\ &= m\omega \left(q\sqrt{\frac{2\alpha}{m\omega^2} - q^2} + \frac{2\alpha}{m\omega^2} \arctan \frac{q}{\sqrt{\frac{2\alpha}{m\omega^2} - q^2}} \right) \Bigg|_{-q_0}^{q_0} \\ &= \frac{2\alpha}{\omega} \left(\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right) = 2\pi \frac{\alpha}{\omega} \,, \end{aligned} \tag{74}$$

where the branches of the solutions have been chosen appropriately for the stages of the oscillation from $-q_0$ to q_0 and back. Thus, Hamilton's characteristic function is

$$W(q, J) = \int dq\, \sqrt{2m\omega J - (m\omega q)^2} \,, \tag{75}$$

and we obtain the angle variable

$$\phi = \frac{\partial W}{\partial J} = \int dq\, \frac{m\omega}{\sqrt{2m\omega J - (m\omega q)^2}} = \arcsin\left(\sqrt{\frac{m\omega}{2J}} q \right) \,, \tag{76}$$

In the action-angle variables, our Hamiltonian reads

$$H = J\omega \,, \tag{77}$$

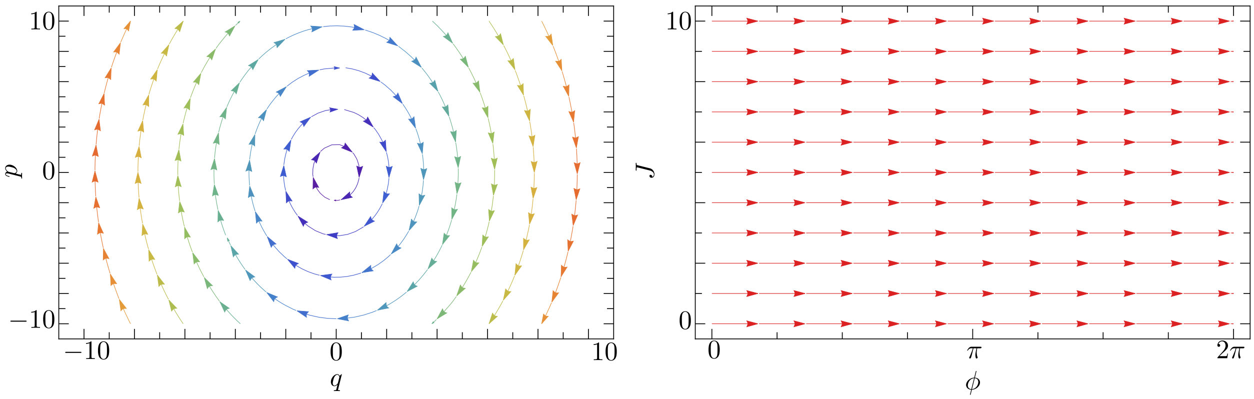

Figure 2: Hamiltonian flow of the harmonic oscillator in the original variables (q, p) (*left panel*), and in action-angle variables (ϕ, J) (*right panel*).

hence ϕ is a cyclic variable. Hamilton's equations then yield

$$\dot{\phi} = \frac{\partial H}{\partial J} = \omega\,, \quad \dot{J} = -\frac{\partial H}{\partial \phi} = 0\,, \tag{78}$$

and we obtain the solutions

$$J = \text{const.}\,, \quad \phi(t) = \omega t + \phi_0\,. \tag{79}$$

In Fig. 2, we show the Hamiltonian flow in phase space in the original variables as well as the action-angle variables. We see that the transformation has "straightened out" the flow from the original ellipses to straight lines.

As a sanity check, we derive the complete transformation and check its canonicity. Combining Eqs. (73) and (76), we obtain the transformation

$$q = \sqrt{\frac{2J}{m\omega}}\sin\phi\,, \quad p = \sqrt{2Jm\omega}\cos\phi\,. \tag{80}$$

We can see that it is canonical by evaluating the Poisson bracket, working backward because the expressions are simpler:

$$\begin{aligned}
\{q, p\}_{(\phi,J)} &= \frac{\partial q}{\partial \phi}\frac{\partial p}{\partial J} - \frac{\partial p}{\partial \phi}\frac{\partial q}{\partial J} \\
&= \sqrt{\frac{2J}{m\omega}}\cos\phi \cdot \frac{m\omega}{\sqrt{2Jm\omega}}\cos\phi - \left(-\sqrt{2Jm\omega}\sin\phi\right)\cdot\frac{1}{m\omega}\frac{1}{\sqrt{\frac{2J}{m\omega}}}\sin\phi \\
&= \cos^2\phi + \sin^2\phi = 1\,.
\end{aligned} \tag{81}$$

2.5.4 Harmonic Oscillator with Two Degrees of Freedom

For a harmonic oscillator with two degrees of freedom, we have

$$H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \frac{p_y^2}{2m} + \frac{1}{2}m\omega^2 y^2 \equiv \alpha_x + \alpha_y\,. \tag{82}$$

The Hamiltonian and Hamilton's characteristic function are separable. Thus,

$$p_x = \frac{\partial W_x}{\partial x} = \pm\sqrt{2m\alpha_x - (m\omega x)^2}\,, \quad p_y = \frac{\partial W_y}{\partial y} = \pm\sqrt{2m\alpha_y - (m\omega y)^2}\,. \tag{83}$$

Taking the appropriate signs as the oscillator moves from x_- to x_+ and back, we have

$$2\pi J_x = 2\int_{x_-}^{x_+} dx\,\sqrt{2m\alpha_x - (m\omega_x x)^2}\,. \tag{84}$$

Using the result for the oscillator with a single degree of freedom, we have

$$J_x = \frac{\alpha}{\omega_x}\,, \tag{85}$$

and analogously,

$$J_y = \frac{\alpha}{\omega_y}\,. \tag{86}$$

We obtain

$$\dot{\phi}_x = \frac{\partial H}{\partial J_x} = \omega_x \tag{87}$$

and

$$\dot{\phi}_y = \frac{\partial H}{\partial J_y} = \omega_y\,. \tag{88}$$

We note that the overall motion is *not* periodic unless the ratio of the frequencies is a rational number:

$$\frac{\omega_y}{\omega_x} = \frac{m}{n}\,, \quad m, n \in \mathbb{Z}\,. \tag{89}$$

2.5.5 Kepler Problem

3 Perturbation Theory

3.1 General Considerations

[...] As our model system, we will consider the nonlinear **Duffing oscillator**[1]

$$H(q,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 q^2 + \frac{1}{4}\alpha q^4 \tag{90}$$

with $\alpha > 0$. Hamilton's equations are

$$\dot{q} = \frac{p}{m}, \quad \dot{p} = -kq - \alpha q^3, \tag{91}$$

which can be combined into the equation of motion

$$\ddot{q} = -\omega_0^2 q - \frac{\alpha}{m} q^3. \tag{92}$$

In general, this equation needs to be solved numerically, but if the quartic term is weak, we can attempt to solve it by performing a perturbation expansion around the solution of a harmonic oscillator.

We introduce the small parameter

$$\epsilon \equiv \frac{\alpha}{m} \ll 1. \tag{93}$$

and make an ansatz of the form

$$q(t) = q_0(t) + \epsilon q_1(t) + \epsilon^2 q_2(t) + \ldots, \tag{94}$$

where $q_0(t)$ is the solution of the unperturbed oscillator. Inserting this into the equation of motion, we obtain

$$\ddot{q}_0 + \epsilon\ddot{q}_1 + \epsilon^2\ddot{q}_2 + \ldots + \omega_0^2\left(q_0 + \epsilon q_1 + \epsilon^2 q_2 + \ldots\right) + \epsilon\left(q_0 + \epsilon q_1 + \epsilon^2 q_2^2 + \ldots\right)^3 = 0. \tag{95}$$

To satisfy this equation, the coefficients of each term in this polynomial must vanish independently, hence we obtain the following system of initial-value problems:

$$\ddot{q}_0 + \omega_0^2 q_0 = 0, \qquad q_0(0) = A, \qquad \dot{q}_0(0) = 0, \tag{96}$$

$$\ddot{q}_1 + \omega_0^2 q_1 = -q_0^3, \qquad q_1(0) = 0, \qquad \dot{q}_1(0) = 0, \tag{97}$$

$$\ddot{q}_2 + \omega_0^2 q_2 = -3q_0^2 q_1, \qquad q_2(0) = 0, \qquad \dot{q}_2(0) = 0, \tag{98}$$

$$\ddot{q}_3 + \omega_0^2 q_3 = -3q_0 q_1^2 - 3q_0^2 q_2, \qquad q_2(0) = 0, \qquad \dot{q}_2(0) = 0, \tag{99}$$

$$\vdots$$

where we have assume that the oscillator is released from rest with some amplitude A. The first of these is easy to solve, and gives us the unperturbed oscillation:

$$q_0(t) = A\cos\omega_0 t. \tag{100}$$

Plugging this into the equation of motion for the next-to-leading term q_1, we have

$$\ddot{q}_1 + q_1 = -A^3\cos^3\omega_0 t, \tag{101}$$

[1] In project #3, we studied the emergence of chaos in the damped, driven version of this oscillator for $\omega_0^2 < 0$.

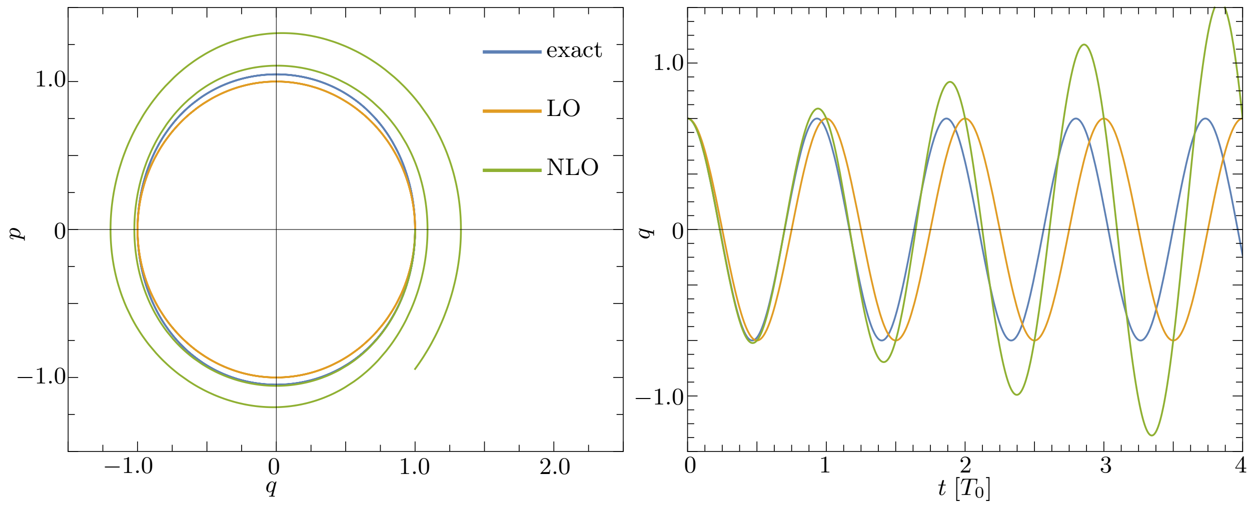

Figure 3: Perturbative solutions for the Duffing oscillator with $A = 1, \omega_0, \epsilon = 0.2$.

To solve this equation, we rewrite the right-hand side as

$$\cos^3 \omega_0 t = \frac{1}{4}\left(3\cos\omega_0 t + \cos 3\omega_0 t\right) , \tag{102}$$

which is nothing but the Fourier series expansion of $\cos^3 \omega_0 t$. Then we obtain

$$\ddot{q}_1 + \omega_0^2 q_1 = -\frac{3A}{4}\left(3\cos\omega_0 t + \cos 3\omega_0 t\right) , \tag{103}$$

which describes an oscillator with two periodic driving forces. Using the initial conditions, the solution is

$$q_1(t) = \frac{A^3}{32\omega_0^2}\left(\cos 3\omega_0 t - \cos\omega_0 t\right) - \frac{3A^3}{8\omega_0} t \sin\omega_0 t \, . \tag{104}$$

We immediately see that this supposed next-to-leading order correction to the oscillator is unphysical: While the first term is purely periodic, the second term grows linearly in t, and is therefore unbounded. The appearance of such **secular terms** is a general phenomenon that leads to the failure of naive perturbative expansions. In the present example, the secular term appears because there is a driving term that oscillates with the same frequency ω_0 as the undamped oscillator, and therefore causes an undamped resonance. In Figure 3, we show the exact phase space trajectories as well as the leading-order and next-to-leading order perturbative solutions

$$q_{\mathrm{LO}}(t) = q_0(t) \, , \quad q_{\mathrm{NLO}}(t) = q_0(t) + \epsilon q_1(t) \, . \tag{105}$$

The resonant behavior of the latter is clearly visible.

3.2 The Poincaré-Lindstedt Method

A solution to the problem of the undamped resonances can be found if we recognize another problem of our perturbative solution, namely that it oscillates at the wrong frequency. Our solution depends on ω_0 and the higher harmonic $3\omega_0$, so the motion is still periodic with the frequency ω_0, which disregards the impact of the quartic perturbation on the oscillation frequency.

Now note that if q_0 were to oscillate at a frequency other than ω_0, we would no longer have have a resonance in Eq. (103). Let us therefore make the following ansatz for the frequency,

$$\omega(\epsilon) = \omega_0 + \epsilon\omega_1 + \epsilon^2\omega_2 + \dots \,. \tag{106}$$

We introduce a new time variable

$$\tau = \omega t \,, \tag{107}$$

which implies

$$d\tau = \omega dt \qquad \Rightarrow \qquad \frac{d}{dt} = \omega^2 \frac{d}{d\tau} \,. \tag{108}$$

Our equation of motion now becomes

$$\omega^2 q'' + \omega_0^2 q + \epsilon q^3 = 0 \,, \tag{109}$$

where the prime indicates the derivative with respect to τ. Now we plug in the expansions for $\omega(\epsilon)$ and q:

$$(\omega_0 + \epsilon\omega_1 + \dots)^2 \left(q_0'' + \epsilon q_1'' + \dots\right) + \omega_0^2 \left(q_0 + \epsilon q_1 + \dots\right) + \epsilon \left(q_0 + \epsilon q_1 + \dots\right)^3 = 0 \,. \tag{110}$$

We obtain the following system of ODEs:

$$\begin{aligned}
\omega_0^2 q_0'' + \omega_0^2 q_0 &= 0 \,, & (111)\\
\omega_0^2 q_1'' + \omega_0^2 q_1 &= -q_0^3 - 2\omega_0\omega_1 q_0'' \,, & (112)\\
\omega_0^2 q_2'' + \omega_0^2 q_2 &= -3q_0^2 q_1 - 2\omega_0\omega_1 q_1'' - \left(\omega_1^2 + 2\omega_0\omega_2\right) q_0'' \,, & (113)\\
&\vdots
\end{aligned}$$

The first equation of motion

$$q_0'' + q_0 = 0 \,, \tag{114}$$

has the general solution

$$q_0(\tau) = A\cos\tau + B\sin\tau = A\cos\omega t + B\sin\omega t \,. \tag{115}$$

or, to leading order in the frequency,

$$q_0(t) = A\cos\omega_0 t + B\sin\omega_0 t \,. \tag{116}$$

Now consider the next-to-leading order equation,

$$q_1'' + q_1 = -2\frac{\omega_1}{\omega_0} q_0'' - \frac{1}{\omega_0^2} y_0^3 \,. \tag{117}$$

Choosing the initial conditions of an oscillator released from rest,

$$q(t=0) = A \,, \quad \dot{q}(t=0) = 0 \,, \tag{118}$$

the chain rule implies that

$$q(\tau=0) = A \,, \quad \omega y'(\tau=0) = 0 \Rightarrow y'(\tau=0) = 0 \,. \tag{119}$$

Our leading-order solution is

$$q_0(\tau) = A\cos(\tau) \tag{120}$$

and the next-to-leading order equation becomes

$$q_1'' + q_1 = 2A\frac{\omega_1}{\omega_0}\cos\tau - \frac{A^3}{\omega_0^2}\cos^3\tau\,. \tag{121}$$

Using the identity (102), we find

$$\begin{aligned} q_1'' + q_1 &= -\frac{A^3}{4\omega_0^2}\left(3\cos\tau + \cos 3\tau\right) + 2A\frac{\omega_1}{\omega_0}\cos\tau \\ &= \frac{2}{\omega_0}\left(A\omega_1 - \frac{3A^3}{8\omega_0}\right)\cos\tau - \frac{A^3}{4\omega_0^2}\cos 3\tau\,. \end{aligned} \tag{122}$$

The first term on the right-hand side could still produce a resonance, but we can cancel it by setting

$$\omega_1 = \frac{3A^2}{8\omega_0}\,. \tag{123}$$

The same approach can be continued through the higher orders in ϵ, so that the cancellation of all divergences will define the coefficients ω_i of the series expansion of $\omega(\epsilon)$. This technique is known as the **Poincaré-Lindstedt method**.

Applying the cancellation, our next-to-leading order equation of motion is

$$q_1'' + q_1 = -\frac{A^3}{4\omega_0^2}\cos 3\tau\,. \tag{124}$$

The solution for the correction is

$$q_1(\tau) = \frac{A^3}{32\omega_0^2}\left(\cos 3\tau - \cos\tau\right) \tag{125}$$

so the full next-to-leading order trajectory is given by

$$q(\tau) = A\cos\tau + \epsilon\frac{A^3}{32\omega^2}\left(\cos 3\tau - \cos\tau\right)\,, \tag{126}$$

or switching back to the original time variable,

$$q(t) = A\cos\omega t + \epsilon\frac{A^3}{32\omega^2}\left(\cos 3\omega t - \cos\omega t\right) \tag{127}$$

with

$$\omega = \omega_0 - \epsilon\frac{3A^2}{8\omega_0}\,. \tag{128}$$

In Fig. 4, we show the perturbative solutions as well as their phase space trajectories. We see that both the LO and NLO solutions are stable for all times, and we note that the NLO solution deforms the shape of the phase space trajectory from that of the unperturbed LO oscillator's ellipse to that of the exact solution.

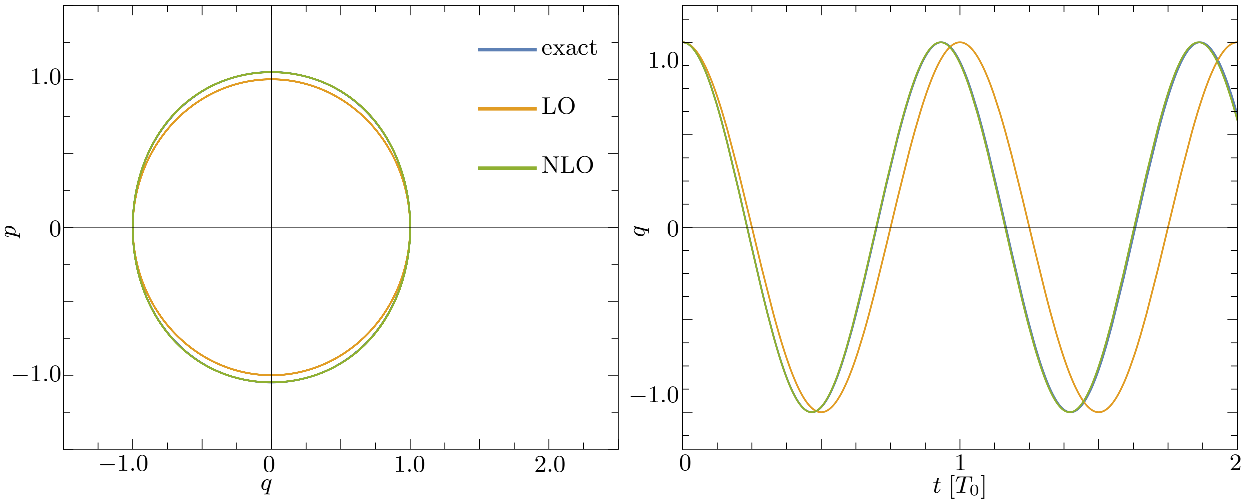

Figure 4: Poincaré-Lindstedt solutions for the Duffing oscillator with $A = 1, \omega_0, \epsilon = 0.2$.

3.3 Canonical Perturbation Theory

3.3.1 Perturbative Canonical Transformations

In the previous sections, we have discussed perturbation theories that are based on expansions of the equations of motion and their solution. Another common strategy is to work in the Hamiltonian-Jacobi formalism instead, and to perform a perturbative construction of Hamilton's principal function S (or the characteristic function W in the time-independent case).

Let us assume we have a Hamiltonian that can be split into unperturbed and perturbed contributions according to

$$H(q,p,t) = H_0(q,p,t) + \epsilon H_1(q,p,t) \tag{129}$$

where ϵ is a small dimensionless parameter. We introduce the usual type-2 canonical transformation, generated by $S(q,P,t)$,

$$\widetilde{H}(Q,P,t) = H(q,p,t) + \frac{\partial}{\partial t}S(q,P,t)\,, \tag{130}$$

and expand all quantities in powers of ϵ:

$$q_\sigma = Q_\sigma + \epsilon q_{1,\sigma} + \epsilon^2 q_{2,\sigma} + \dots \tag{131}$$

$$p_\sigma = P_\sigma + \epsilon p_{1,\sigma} + \epsilon^2 p_{2,\sigma} + \dots \tag{132}$$

$$\widetilde{H} = \widetilde{H}_0 + \epsilon \widetilde{H}_1 + \epsilon^2 \widetilde{H}_2 + \dots \tag{133}$$

$$S = q_\sigma P_\sigma + \epsilon S_1 + \epsilon^2 S_2 + \dots \tag{134}$$

We use Greek letters to distinguish coordinate indices from perturbation orders. Also note that the leading-order term of S is the identity transformation (cf. Sec. 2.2). Plugging these equations into the expressions for the coordinates Q and p, we find

$$Q_\sigma = \frac{\partial S}{\partial P_\sigma} = q_\sigma + \epsilon\frac{\partial S_1}{\partial P_\sigma} + \epsilon^2\frac{\partial S_2}{\partial P_\sigma} + \dots \tag{135}$$

$$= Q_\sigma + \left(q_{1,\sigma} + \frac{\partial S_1}{\partial P_\sigma}\right)\epsilon + \left(q_{2,\sigma} + \frac{\partial S_2}{\partial P_\sigma}\right)\epsilon^2 + \dots\,, \tag{136}$$

and

$$p_\sigma = \frac{\partial S}{\partial q_\sigma} = P_\sigma + \epsilon \frac{\partial S_1}{\partial q_\sigma} + \epsilon^2 \frac{\partial S_2}{\partial q_\sigma} + \dots \tag{137}$$

$$= P_\sigma + \epsilon p_{1,\sigma} + \epsilon^2 p_{2,\sigma} + \dots\,. \tag{138}$$

Comparing coefficients, we see that order-by-order in ϵ, we will have

$$Q_{k,\sigma} = \frac{\partial S_k}{\partial P_\sigma}\,, \quad p_{k,\sigma} = \frac{\partial S_k}{\partial q_\sigma}\,. \tag{139}$$

Next, we need to expand the Hamiltonian. Since our two sets of coordinates are related by the perturbative canonical transformation, we perform a Taylor expansion of $H_k(q,p,t)$ and $S_k(q,P,t)$ around $q = Q$ and $p = P$, and then express the differences $q - Q$ and $p - P$ in our perturbative expansion. Thus,

$$\begin{aligned}\widetilde{H}(Q,P,t) &= H_0(q,p,t) + \epsilon H_1(q,p,t) + \frac{\partial S}{\partial t}\\ &= H_0(Q,P,t) + \sum_\sigma \left(\frac{\partial H_0}{\partial Q_\sigma}(q_\sigma - Q_\sigma) + \frac{\partial H_0}{\partial P_\sigma}(p_\sigma - P_\sigma)\right)\\ &\quad + \epsilon H_1(Q,P,t) + \epsilon \frac{\partial}{\partial t} S_1(Q,P,t) + O\left(\epsilon^2\right)\end{aligned} \tag{140}$$

Collecting terms and using the relation between S and the coordinates, we have

$$\widetilde{H}(Q,P,t) = H_0(Q,P,t) + \sum_\sigma \left(-\frac{\partial H_0}{\partial Q_\sigma}\frac{\partial S_1}{\partial P_\sigma} + \frac{\partial H_0}{\partial P_\sigma}\frac{\partial S_1}{\partial Q_\sigma} + \frac{\partial S_1}{\partial t} + H_1\right)\epsilon + O\left(\epsilon^2\right) \tag{141}$$

$$= H_0(Q,P,t) + \left(H_1 + \left\{S_1, H_0\right\} + \frac{\partial S_1}{\partial t}\right)\epsilon + O\left(\epsilon^2\right)\,. \tag{142}$$

This implies

$$\widetilde{H}(Q,P,t) = \widetilde{H}_0(Q,P,t) + \epsilon \widetilde{H}_1(Q,P,t) + \dots \tag{143}$$

with

$$\widetilde{H}_0(Q,P,t) = H_0(Q,P,t) \tag{144}$$

$$\widetilde{H}_1(Q,P,t) = H_1 + \{S_1, H_0\} + \frac{\partial S_1}{\partial t} \tag{145}$$

$$\vdots$$

Inspecting this tower of equations, we notice a complication: Through next-to-leading order, the system is underdetermined because Eq. (145) only provides a single equation for two unknowns, $\widetilde{H}$ and S_1. Thus, we must specify some additional requirement, e.g., that the transformation eliminates the perturbation $\widetilde{H}_1$ completely through $O(\epsilon)$.

Relation to Quantum Mechanics

Our perturbative construction of a canonical transformation mirrors the so-called canonical or **Van Vleck Perturbation Theory** in (many-body) quantum mechanics, where a unitary transformation is implemented to eliminate a perturbation to lowest order in a small parameter.

In this approach, we consider the Schrdinger equation,

$$i\hbar\frac{\partial}{\partial t}\left|\psi\right\rangle = (H+\epsilon H_1)\left|\psi\right\rangle \tag{146}$$

where $\left|\psi\right\rangle$ is an exact eigenstate, and define a mapping between an unperturbed *reference state* $\left|\phi\right\rangle$ and this exact solution via

$$\left|\psi\right\rangle \equiv e^{iS(\epsilon)/\hbar}\left|\phi\right\rangle \equiv U(\epsilon)\left|\phi\right\rangle \tag{147}$$

The Hermitian generator S of the transformation is expanded perturbatively,

$$S(\epsilon) = \epsilon S_1 + \epsilon^2 S_2 + \dots . \tag{148}$$

By construction, $U(\epsilon)$ is the identity in the limit $\epsilon \to 0$ limit.

The Schrödinger equation for $\left|\phi\right\rangle$ can be written as

$$i\hbar\frac{\partial}{\partial t}\left|\phi\right\rangle = H_0\left|\phi\right\rangle + \epsilon\left(H_1 + \frac{1}{i\hbar}\left[S_1, H_0\right] + \frac{\partial S_1}{\partial t}\right)\left|\phi\right\rangle + \ldots \equiv \widetilde{H}\left|\phi\right\rangle \tag{149}$$

where $\left[A, B\right]$ is the commutator. Note the correspondence with Eq. (145).

3.3.2 Application to Systems with a Single Degree of Freedom

In the following, we will assume that H does not depend explicitly on time, so that

$$H(q,p) = H_0(q,p) + \epsilon H_1(q,p)\,. \tag{150}$$

Let us now assume that H_0 describes a bounded system that is described in the action angle variables (ϕ_0, J_0). Then

$$\widetilde{H}_0\left(\phi_0, J_0\right) = H_0\left(q\left(\phi_0, J_0\right), p\left(\phi_0, J_0\right)\right) = \widetilde{H}_0\left(J_0\right)\,, \tag{151}$$

as discussed in Sec. 2.5 (cf. Eq. (62)). The transformed perturbation is analogously,

$$\widetilde{H}_1\left(\phi_0, J_0\right) \equiv H_1\left(q\left(\phi_0, J_0\right), p\left(\phi_0, J_0\right)\right)\,, \tag{152}$$

but it will in general not be cyclic in ϕ_0. We assume that $\widetilde{H} = \widetilde{H}_0 + \epsilon\widetilde{H}_1$ is integrable so that action-angle variables exists, which we denote by (ϕ, J). Thus, there must be a canonical transformation from (ϕ_0, J_0) to (ϕ, J), such that

$$\widetilde{H}\left(\phi_0(\phi, J), J_0(\phi, J)\right) \equiv E(J)\,. \tag{153}$$

Writing Hamilton's principal function as

$$S\left(\phi_0, J\right) = \phi_0 J + \epsilon S_1\left(\phi_0, J\right) + \epsilon^2 S_2\left(\phi_0, J\right) + \ldots\,, \tag{154}$$

where $\phi_0 J$ is the identity transformation, we have

$$J_0 = \frac{\partial S}{\partial \phi_0} = J + \epsilon\frac{\partial S_1}{\partial \phi_0} + \epsilon^2\frac{\partial S_2}{\partial \phi_0} + \ldots \tag{155}$$

$$\phi = \frac{\partial S}{\partial J} = \phi_0 + \epsilon\frac{\partial S_1}{\partial J} + \epsilon^2\frac{\partial S_2}{\partial J} + \ldots \tag{156}$$

and

$$E(J) = E_0(J) + \epsilon E_1(J) + \epsilon^2 E_2(J) + \ldots = \widetilde{H}_0(\phi_0, J_0) + \epsilon \widetilde{H}_1(\phi_0, J_0)\ . \tag{157}$$

Now we can expand $\widetilde{H}(\phi_0, J_0)$ in powers of $J_0 - J$:

$$\begin{aligned} \widetilde{H}(\phi_0, J_0) &= \widetilde{H}_0(\phi_0, J_0) + \epsilon \widetilde{H}_1(\phi_0, J_0) \\ &= \widetilde{H}_0(J) + \left.\frac{\partial \widetilde{H}_0}{\partial J}\right|_{\phi_0}(J_0 - J) + \frac{1}{2}\left.\frac{\partial^2 \widetilde{H}_0}{\partial J^2}\right|_{\phi_0}(J_0 - J)^2 + \ldots \\ &\quad + \epsilon \widetilde{H}_1(\phi_0, J) + \epsilon \left.\frac{\partial \widetilde{H}_1}{\partial J}\right|_{\phi_0}(J_0 - J) + \ldots\ . \end{aligned} \tag{158}$$

Grouping the terms according to powers of ϵ, we obtain

$$\widetilde{H}(\phi_0, J_0) = \widetilde{H}_0(J) + \left(\widetilde{H}_1 + \frac{\partial \widetilde{H}_0}{\partial J}\frac{\partial S_1}{\partial \phi_0}\right)\epsilon + \left(\frac{\partial \widetilde{H}_0}{\partial J}\frac{\partial S_2}{\partial \phi_0} + \frac{1}{2}\frac{\partial^2 \widetilde{H}_0}{\partial J^2}\left(\frac{\partial S_1}{\partial \phi_0}\right)^2 + \frac{\partial \widetilde{H}_1}{\partial J}\frac{\partial S_1}{\partial \phi_0}\right)\epsilon^2 + \ldots \tag{159}$$

where all terms on the right-hand side are functions of ϕ_0 and J. Comparing Eqs. (157) and (159), we find

$$E_0(J) = \widetilde{H}_0(J) \tag{160}$$

$$E_1(J) = \widetilde{H}_1(\phi_0, J) + \frac{\partial \widetilde{H}_0}{\partial J}\frac{\partial S_1}{\partial \phi_0} \tag{161}$$

$$E_2(J) = \frac{\partial \widetilde{H}_0}{\partial J}\frac{\partial S_2}{\partial \phi_0} + \frac{1}{2}\frac{\partial^2 \widetilde{H}_0}{\partial J^2}\left(\frac{\partial S_1}{\partial \phi_0}\right)^2 + \frac{\partial \widetilde{H}_1}{\partial J}\frac{\partial S_1}{\partial \phi_0}\,. \tag{162}$$

We now choose the S_k such that the dependence of the right-hand sides of these equations on the angle variable ϕ_0 is eliminated. To that end, we average the expressions on the right-hand side over ϕ_0,

$$\langle f(\phi_0)\rangle = \int_0^{2\pi} \frac{d\phi_0}{2\pi} f(\phi_0) \tag{163}$$

These averages are performed at fixed J and *not* at fixed J_0.

Now we note that if we hold J constant and increase ϕ_0 by 2π, we will returns to the same starting point in phase space if the motion is bounded. Therefore, J is a periodic function of ϕ_0, and we can write

$$S_k(\phi_0, J) = \sum_{m=-\infty}^{\infty} S_{k,m}(J) e^{im\phi_0} \tag{164}$$

for each $k > 0$, hence

$$\left\langle \frac{\partial S_k}{\partial \phi_0} \right\rangle = \frac{1}{2\pi}\left[S_k(2\pi, J) - S_k(0, J)\right] = 0\,. \tag{165}$$

Let us now apply the averaging to the first two orders of the hierarchy. Since $\widetilde{H}_0(J)$ is independent of ϕ_0 and $\frac{\partial S_1}{\partial \phi_0}$ is periodic (cf. Eq. (164), we have

$$E_1(J) = \left\langle \widetilde{H}_1(\phi_0, J)\right\rangle + \frac{\partial \widetilde{H}_0}{\partial J}\underbrace{\left\langle \frac{\partial S_1}{\partial \phi_0}\right\rangle}_{=0} \tag{166}$$

and S_1 must satisfy

$$\frac{\partial S_1}{\partial \phi_0} = \frac{\left\langle \widetilde{H}_1 \right\rangle - \widetilde{H}_1}{\omega_0(J)} \tag{167}$$

where $\omega_0(J) = \frac{\partial H_0}{\partial J}$. The right-hand side of this equation averages to zero, and must be a periodic function of ϕ_0, per our previous discussion. Thus, the solution is

$$S_1 = S_1\left(\phi_0, J\right) + f(J)\,, \tag{168}$$

where $f(J)$ is an arbitrary function of J. However, $f(J)$ affects only the difference $\phi - \phi_0$, which is changed by a constant value $f'(J)$, so we can simply take $f(J) = 0$.

Now consider the second order in ϵ. We have

$$E_2(J) = \left\langle \frac{\partial \widetilde{H}_1}{\partial J}\frac{\partial S_1}{\partial \phi_0} \right\rangle + \frac{1}{2}\frac{\partial \omega_0}{\partial J}\left\langle \left(\frac{\partial S_1}{\partial \phi_1}\right)^2 \right\rangle + \nu_0(J)\underbrace{\left\langle \frac{\partial S_2}{\partial \phi_0} \right\rangle}_{=0}\,. \tag{169}$$

Thus, we obtain

$$\begin{aligned}\frac{\partial S_2}{\partial \phi_0} = \frac{1}{\omega_0^2(J)}\Bigg\{ &\left\langle \frac{\partial \widetilde{H}_1}{\partial J} \right\rangle \left\langle \widetilde{H}_0 \right\rangle - \left\langle \frac{\partial \widetilde{H}_1}{\partial J}\widetilde{H}_0 \right\rangle - \frac{\partial \widetilde{H}_1}{\partial J}\left\langle \widetilde{H}_1 \right\rangle + \frac{\partial \widetilde{H}_1}{\partial J}\widetilde{H}_1 \\ &+ \frac{1}{2}\frac{\partial \ln \omega_0}{\partial J}\left(\left\langle \widetilde{H}_1^2 \right\rangle - 2\left\langle \widetilde{H}_1 \right\rangle^2 + 2\left\langle \widetilde{H}_1 \right\rangle - \widetilde{H}_1^2 \right) \Bigg\}\,,\end{aligned} \tag{170}$$

and the expansion for the energy $E(J)$ becomes

$$\begin{aligned}E(J) = \widetilde{H}_0(J) + \epsilon\left\langle \widetilde{H}_1 \right\rangle + \frac{\epsilon^2}{\omega_0(J)}\Bigg\{ &\left\langle \frac{\partial \widetilde{H}_1}{\partial J} \right\rangle \left\langle \widetilde{H}_1 \right\rangle - \left\langle \frac{\partial \widetilde{H}_1}{\partial J}\widetilde{H}_1 \right\rangle \\ &+ \frac{1}{2}\frac{\partial \ln \omega_0}{\partial J}\left(\left\langle \widetilde{H}_1^2 - \left\langle \widetilde{H}_1 \right\rangle^2 \right) \right) \Bigg\} + \mathcal{O}\left(\epsilon^3\right)\end{aligned} \tag{171}$$

Note that we do not need to know S to find $E(J)$. The perturbed fundamental frequencies are

$$\omega(J) = \partial E / \partial J\,. \tag{172}$$

Sometimes these frequencies are all that is desired, but if necessary, we can reconstruct the full motion of the system via the successive canonical transformations

$$(\phi, J) \longrightarrow (\phi_0, J_0) \longrightarrow (q, p)\,. \tag{173}$$

3.3.3 Example: The Duffing Oscillator

Let us now return to the Duffing oscillator, and define

$$H(q,p) = \underbrace{\frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 q^2}_{\equiv H_0} + \frac{1}{4}\epsilon\alpha q^4\,. \tag{174}$$

Note the slight change in definition form previous sections. Here, we will set $\epsilon = 1$ in the end.

The action-angle variables for the harmonic oscillator Hamiltonian H_0 are (cf. Sec. 2.5)

$$\phi_0 = \tan^{-1}\left(\frac{m\omega_0 q}{p}\right), \quad J_0 = \frac{p^2}{2m\omega_0} + \frac{1}{2}m\omega_0 q^2\,, \tag{175}$$

and we have

$$H_0 = \nu_0 J_0\,. \tag{176}$$

For the full Hamiltonian, we have

$$\begin{aligned}\widetilde{H}\left(\phi_0, J_0\right) &= \omega_0 J_0 + \frac{1}{4}\epsilon\alpha\left(\sqrt{\frac{2J_0}{m\omega_0}}\sin\phi_0\right)^4 = \omega_0 J_0 + \frac{\epsilon\alpha}{m^2\omega_0^2}J_0^2\sin^4\phi_0 \\ &\equiv H_0\left(\phi_0, J_0\right) + \epsilon\widetilde{H}_1\left(\phi_0, J_0\right)\,.\end{aligned} \tag{177}$$

We can now evaluate the energy contribution from $O(\epsilon)$, which is given by

$$E_1(J) = \left\langle \widetilde{H}_1\left(\phi_0, J\right)\right\rangle = \frac{\alpha J^2}{m^2\omega_0^2}\int_0^{2\pi}\frac{d\phi_0}{2\pi}\sin^4\phi_0 = \frac{3\alpha J^2}{8m^2\omega_0^2}\,. \tag{178}$$

For the fundamental frequency, we have

$$\omega(J) = \omega_0 + \frac{3\epsilon\alpha J}{4m^2\omega_0^2} \tag{179}$$

To lowest order in ϵ, we may replace J by

$$J_0 = \frac{1}{2}m\omega_0 A^2 \tag{180}$$

where A is the amplitude of the q motion. Thus,

$$\omega(A) = \omega_0 + \frac{3\epsilon\alpha A^2}{8m\omega_0}\,, \tag{181}$$

which matches the result we obtained using the Poincar-Lindstedt method (identifying $(\epsilon\frac{\alpha}{m})_{\text{can}} = (\epsilon)_{\text{PL}}$).

Next, we can construct the canonical transformation $(\phi_0, J_0) \to (\phi, J)$. We have

$$\omega_0\frac{\partial S_1}{\partial\phi_0} = \frac{\alpha J^2}{m^2\omega_0^2}\left(\frac{3}{8} - \sin^4\phi_0\right) \tag{182}$$

and therefore

$$S\left(\phi_0, J\right) = \phi_0 J + \frac{\epsilon\alpha J^2}{8m^2\omega_0^3}\left(3 + 2\sin^2\phi_0\right)\sin\phi_0\cos\phi_0 + O\left(\epsilon^2\right)\,. \tag{183}$$

Using S, we obtain

$$\phi = \frac{\partial S}{\partial J} = \phi_0 + \frac{\epsilon\alpha J}{4m^2\omega_0^3}\left(3 + 2\sin^2\phi_0\right)\sin\phi_0\cos\phi_0 + O\left(\epsilon^2\right) \tag{184}$$

$$J_0 = \frac{\partial S}{\partial\phi_0} = J + \frac{\epsilon\alpha J^2}{8m^2\omega_0^3}\left(4\cos 2\phi_0 - \cos 4\phi_0\right) + O\left(\epsilon^2\right)\,. \tag{185}$$

To lowest order, we may again replace J by J_0 in these expressions, which yields

$$J = J_0 - \frac{\epsilon\alpha J_0^2}{8m^2\omega_0^3}\left(4\cos 2\phi_0 - \cos 4\phi_0\right) + O\left(\epsilon^2\right)\,, \tag{186}$$

$$\phi = \phi_0 + \frac{\epsilon\alpha J_0}{8m^2\omega_0^3}\left(3 + 2\sin^2\phi_0\right)\sin 2\phi_0 + O\left(\epsilon^2\right)\,. \tag{187}$$

These relations implicitly define the coordinates (q, p), but they cannot be inverted analytically.

3.4 Renormalization Group Approach

See Ref. [1].

References

[1] E. Kirkinis, SIAM Review **54**, 374 (2012).

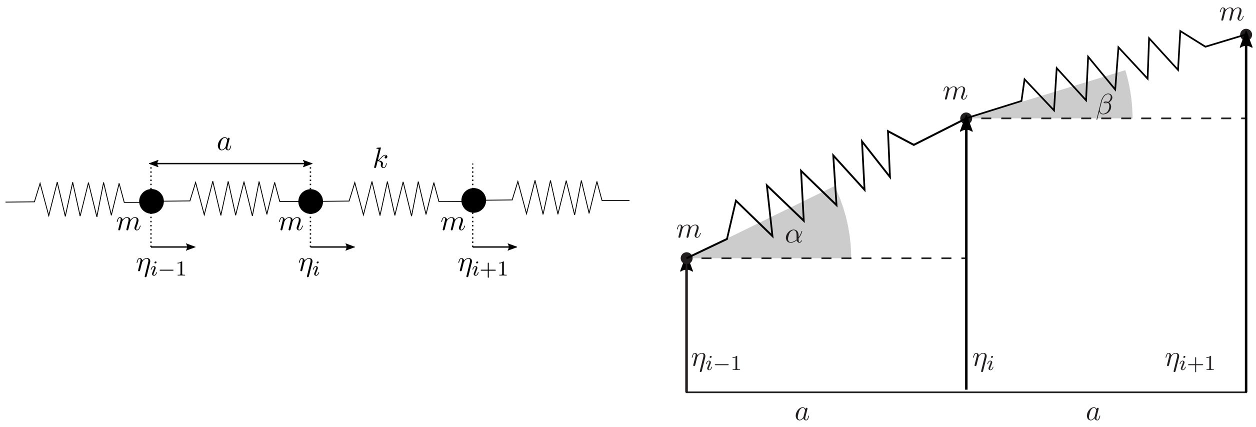

Figure 5: Coupled longitudinal and transversal oscillators.

4 Classical Field Theory

4.1 From Discrete to Continuous Systems

In our discussion of coupled oscillators, we found that the equation of motion for a mass m that is connected via identical springs to equal masses m is given by (cf. worksheet #12, and exercises H22, P16)

$$m\ddot{\eta}_i = k(\eta_{i+1} - \eta_i) - k(\eta_i - \eta_{i-1})\,, \tag{188}$$

where η_i are the displacements of the masses out of equilibrium. It does not matter if the displacement out ouf equilibrium is in *longitudinal* direction, along the chain of coupled oscillators, or *transversal*. The longitudinal case was discussed before, and for the transversal case, we refer to Fig. 5: We have

$$m\ddot{i} = -k\frac{a}{\cos\alpha}\sin\alpha + k\frac{a}{\cos\beta}\sin\beta = -ka\tan\alpha + ka\tan\beta \tag{189}$$

and using

$$\tan\alpha = \frac{\eta_i - \eta_{i-1}}{a}\,, \quad \tan\beta = \frac{\eta_{i+1} - \eta_i}{a}\,, \tag{190}$$

we also obtain

$$m\ddot{\eta}_i = ka\tan\beta - ka\tan\alpha = k(\eta_{i+1} - \eta_i) - k(\eta_i - \eta_{i-1})\,. \tag{191}$$

The Lagrangian leading to these equations of motion is given by

$$L = \frac{1}{2}\sum_i m\dot{\eta}_i^2 - \frac{1}{2}\sum_i k\left(\eta_{i+1} - \eta_i\right)^2\,. \tag{192}$$

Now let us take the continuous limit of this expression, so that the coupled oscillator become an elastic medium. First, we rewrite the Lagrangian as

$$L = \frac{1}{2}\sum_i a\frac{m}{a}\dot{\eta}_i^2 - \frac{1}{2}\sum_i ka^2\left(\frac{\eta_{i+1} - \eta_i}{a}\right)^2 = \frac{1}{2}\sum_i a\mu\dot{\eta}_i^2 - \frac{1}{2}\sum_i a\tau\left(\frac{\eta_{i+1} - \eta_i}{a}\right)^2\,, \tag{193}$$

where $\mu = \frac{m}{a}$ is the mass density and $\tau = ka$ the "string" tension or modulus of elasticity of the medium. We now let $a \to 0$ while keeping μ and τ fixed. Then

$$\eta_i(t) \quad \to \quad \varphi(x,t)\,, \tag{194}$$

$$\frac{\eta_{i+1}(t)-\eta_i(t)}{a} \quad \to \quad \frac{\partial\varphi(x,t)}{\partial x}\,, \tag{195}$$

$$\frac{1}{a}\left(\frac{\eta_{i+1}(t)-\eta_i(t)}{a}-\frac{\eta_i(t)-\eta_{i-1}(t)}{a}\right) \quad \to \quad \frac{\partial^2\varphi(x,t)}{\partial x^2}\,, \tag{196}$$

$$\sum_i a \quad \to \quad \int dx\,. \tag{197}$$

The Lagrangian becomes

$$L = \int dx\,\mathcal{L}\left(\varphi, \frac{\partial\varphi}{\partial t}, \frac{\partial\varphi}{\partial x}\right) \equiv \int dx\,\left(\frac{1}{2}\mu\left(\frac{\partial\varphi}{\partial t}\right)^2 - \frac{1}{2}\tau\left(\frac{\partial\varphi}{\partial x}\right)^2\right)\,, \tag{198}$$

where we have introduced the **Lagrangian density**

$$\mathcal{L}\left(\varphi, \frac{\partial\varphi}{\partial t}, \frac{\partial\varphi}{\partial x}\right) = \frac{1}{2}\mu\left(\frac{\partial\varphi}{\partial t}\right)^2 - \frac{1}{2}\tau\left(\frac{\partial\varphi}{\partial x}\right)^2\,. \tag{199}$$

In continuum mechanics or classical field theory, $\mathcal{L}$ is often also referred to as the Lagrangian, for simplicity.

The continuous limit of the equation of motion is

$$\mu\frac{\partial^2\varphi}{\partial t^2} = \tau\frac{\partial^2\varphi}{\partial x^2}\,. \tag{200}$$

Rearranging, we obtain the **wave equation**

$$\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} - \frac{\partial^2\varphi}{\partial x^2} = 0\,, \tag{201}$$

where we have introduced the **phase velocity** of the medium,

$$c = \sqrt{\frac{\tau}{\mu}}\,. \tag{202}$$

4.2 Lagrange Formalism for Fields in Three Dimensions

The results of the previous section were obtained by taking the continuous limit of a Lagrangian and its Lagrange equations that were defined in discrete variables. Let us now generalize this result to N fields in three spatial dimensions, and derive the general form of the Lagrange equations for these fields.

Just as in the case of discrete variables, we introduce the tuple of fields

$$\varphi(\vec{x},t) \equiv \left(\varphi_1(\vec{x},t), \ldots, \varphi_N(\vec{x},t)\right)\,, \tag{203}$$

where we have used the notation $\vec{x} = (x,y,z)^T$ for the spatial vector instead of $\vec{r}$ to prepare for the extension to relativity later on. The action for these fields can be written as

$$S = \int_{t_1}^{t_2} dt \int_V d^3x\,\mathcal{L}\left(\varphi, \frac{\partial\varphi}{\partial t}, \vec{\nabla}\varphi, \vec{x}, t\right)\,. \tag{204}$$

The principle of least action implies

$$\begin{aligned}\delta S &= \delta \int_{t_1}^{t_2} dt \int_V d^3x \; \mathcal{L}\left(\varphi, \frac{\partial \varphi}{\partial t}, \vec{\nabla}\varphi, \vec{x}, t\right) \\ &= \int_{t_1}^{t_2} dt \int_V d^3x \sum_{\alpha=1}^{N} \left\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha}\delta\varphi_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha}\delta\dot{\varphi}_\alpha + \frac{\partial \mathcal{L}}{\partial\left(\vec{\nabla}\varphi_\alpha\right)} \cdot \delta\left(\vec{\nabla}\varphi_\alpha\right) \right\} \\ &= 0\,, \end{aligned} \tag{205}$$

where the variations of the φ_α are mutually independent and vanish at the endpoints of the time integration as well as the boundary ∂V of our volume. Using

$$\delta\dot{\varphi}_\alpha = \frac{\partial \delta\varphi_\alpha}{\partial t}\,, \tag{206}$$

just as for discrete coordinates, we can integrate by parts and obtain

$$\int_{t_1}^{t_2} dt \int_V d^3x \; \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha}\delta\dot{\varphi}_\alpha = -\int_{t_1}^{t_2} dt \int_V d^3x \; \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha}\right)\delta\varphi_\alpha \tag{207}$$

For the spatial derivatives, we can use

$$\delta\left(\vec{\nabla}\varphi_\alpha\right) = \vec{\nabla}\left(\delta\varphi_\alpha\right)\,, \tag{208}$$

the identity

$$\vec{F}\cdot\vec{\nabla}g = \vec{\nabla}\cdot(g\vec{F}) - g\vec{\nabla}\cdot\vec{F} \tag{209}$$

and Gauss' theorem

$$\int_V d^3x \; \vec{\nabla}\cdot\vec{F} = \oint_{\partial V} d\vec{A}\cdot\vec{F} \tag{210}$$

to perform another integration by parts:

$$\begin{aligned}&\int_{t_1}^{t_2} dt \int_V d^3x \frac{\partial \mathcal{L}}{\partial\left(\vec{\nabla}\varphi_\alpha\right)} \cdot \delta\left(\vec{\nabla}\varphi_\alpha\right) = \int_{t_1}^{t_2} dt \int_V d^3x \frac{\partial \mathcal{L}}{\partial\partial\left(\vec{\nabla}\varphi_\alpha\right)} \cdot \vec{\nabla}\left(\delta\varphi_\alpha\right) \\ &= \int_{t_1}^{t_2} dt \oint_{\partial V} d\vec{A}\cdot \frac{\partial \mathcal{L}}{\partial\left(\vec{\nabla}\varphi_\alpha\right)}\delta\varphi_\alpha - \int_{t_1}^{t_2} dt \int_V d^3x \vec{\nabla}\cdot\left(\frac{\partial \mathcal{L}}{\partial\left(\vec{\nabla}\varphi_\alpha\right)}\right)\delta\varphi_\alpha \\ &= -\int_{t_1}^{t_2} dt \int_V d^3x \; \vec{\nabla}\cdot\left(\frac{\partial \mathcal{L}}{\partial\partial\left(\vec{\nabla}\varphi_\alpha\right)}\right)\delta\varphi_\alpha\,, \end{aligned} \tag{211}$$

where we have used that the variations $\delta\varphi_\alpha$ vanish on ∂V.

Plugging these results into the principle of least action, we have

$$\delta S = \int_{t_1}^{t_2} dt \int_V d^3x \; \sum_{\alpha=1}^{N} \left\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha}\right) - \vec{\nabla}\cdot\left(\frac{\partial \mathcal{L}}{\partial\left(\vec{\nabla}\varphi_\alpha\right)}\right) \right\}\delta\varphi_\alpha = 0\,, \tag{212}$$

and since this relation must hold for arbitrary variations $\delta\varphi_\alpha$, we obtain the **Lagrange equations**

$$\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha}\right) + \vec{\nabla}\cdot\left(\frac{\partial \mathcal{L}}{\partial\left(\vec{\nabla}\varphi_\alpha\right)}\right) - \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} = 0, \qquad \alpha = 1, \ldots, N\,. \tag{213}$$

Box 4.1: Einstein's Summation Convention

Einstein's summation convention offers a compact notation for writing contractions in relativistic calculations. It states that any index that appears **exactly twice** in a given term, once as an upper, contravariant index and once as a lower, covariant index and that is **not otherwise defined** is implicitly summed over its entire range.
For example, the scalar product of two four-vectors a and b can be written as

$$a_\mu b^\mu = a^\mu b_\mu = a^0 b_0 - \vec{a}\cdot\vec{b} = a_0 b^0 - \vec{a}\cdot\vec{b}\,. \tag{I4.1-1}$$

4.3 Relativistic Field Theory

4.3.1 Lagrange Equations

Let us now move on to relativistic fields. We define the *contravariant* four vector as

$$x^\mu \equiv (ct, \vec{x})^T\,, \quad \mu = 0,1,2,3\,, \tag{214}$$

and the *covariant* four-vector gradient as

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \equiv \left(\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla}\right)^T\,. \tag{215}$$

The corresponding covariant vector and contravariant gradient are

$$x_\mu = \eta_{\mu\nu} x^\nu\,, \quad \partial^\mu = \eta^{\mu\nu}\partial_\nu \tag{216}$$

where we have used Einstein's summation convention and introduced the metric for flat **Minkowski spacetime**[2],

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)\,. \tag{217}$$

For future use, we note that

$$\eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu_\rho = \mathrm{diag}(1,1,1,1)\,. \tag{218}$$

The Lagrange equations (213) are invariant under scale transformations of the coordinates $\vec{x}, t$, and we have

$$\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\varphi_\alpha/\partial t)}\right) = \frac{\partial}{\partial x^0}\left(\frac{\partial\mathcal{L}}{\partial(\partial\varphi_\alpha/\partial x^0)}\right)\,. \tag{219}$$

Thus, we can assume that time derivatives are taken with respect to x^0, which allows us to write the Lagrange equations in the form

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_\alpha)}\right) - \frac{\partial\mathcal{L}}{\partial\varphi_\alpha} = 0, \qquad \alpha = 1,\ldots,N\,, \tag{220}$$

which is *manifestly covariant*: If φ_α is a scalar, the first term is a scalar product between covariant and contravariant objects, and therefore invariant under any Lorentz transformation between inertial frames. Manifest covariance of the Lagrange equations (220) is ensured if $\mathcal{L}$ is a

[2] We use the particle-physics convention for the signature of the metric, $\eta^{\mu\nu} = \mathrm{diag}(+1,-1,-1,-1)$. In General Relativity, it is customary to use the opposite signature, $\eta^{\mu\nu} = \mathrm{diag}(-1,+1,+1,+1)$.

scalar. Since the volume element in Minkowski spacetime, d^4x, is a scalar, this requirement also ensures that the action

$$S = \int d^4x \, \mathcal{L} \tag{221}$$

is a scalar as well.

4.3.2 The Klein-Gordon Field

As a first example of a relativistic field theory, we consider the theory of an uncharged scalar particle, which is described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 \, . \tag{222}$$

Here, m is the particle's mass (in units such that $\hbar = c = 1$). As discussed in the previous section, the Lagrangian is a scalar under Lorentz transformations, since ϕ is a scalar field. Comparing this to the usual expression $L = T - V$, we see that

$$T = \frac{1}{2}\int d^3x \; \dot{\phi}^2 \tag{223}$$

and

$$V = \frac{1}{2}\int d^3x \; \left((\vec{\nabla}\phi)^2 + m^2\phi^2\right) \, . \tag{224}$$

Let us now compute the partial derivatives: We have

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} &= \frac{1}{2}\frac{\partial}{\partial(\partial_\mu\phi)}\left(\eta^{\nu\rho}\partial_\rho\phi\;\partial_\nu\phi\right) = \frac{1}{2}\left(\eta^{\nu\rho}\delta^\mu_\rho\delta_\nu\phi + \eta^{\nu\rho}\delta^\mu_\nu\delta_\rho\phi\right) = \frac{1}{2}\left(\eta^{\nu\mu}\delta_\nu\phi + \eta^{\mu\rho}\delta_\rho\phi\right) \\ &= \partial^\mu\phi\end{aligned} \tag{225}$$

and

$$\frac{\partial\mathcal{L}}{\partial\phi} = -m^2\phi \, . \tag{226}$$

Inserting these results into Eq. (220), we obtain the **Klein-Gordon equation**

$$\partial_\mu\partial^\mu\phi + m^2\phi = \left(\Box + m^2\right)\phi = 0 \, , \tag{227}$$

where we have introduced the **D'Alembert operator**

$$\Box \equiv \partial^\mu\partial_\mu = \partial_0^2 - \vec{\nabla}^2 \, . \tag{228}$$

4.3.3 The Electromagnetic Field

An even more important example of relativistic field theory is the electromagnetic field in the presence of charges and currents described by the **four-current**

$$J^\mu = (c\rho, \vec{j})^T \, . \tag{229}$$

Here, our independent fields are the components of the four-potential

$$A^\mu = (\phi/c, \vec{A})^T \, . \tag{230}$$

The Lagrangian can be written as

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu v} - \mu_0 J^\mu A_\mu \tag{231}$$

where we have defined the **electromagnetic field strength tensor**

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \,. \tag{232}$$

The partial derivatives are

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} &= -\frac{1}{4} \frac{\partial}{\partial(\partial_\mu A_\nu)} \eta^{\alpha\rho} \eta^{\beta\sigma} \left((\partial_\rho A_\sigma - \partial_\sigma A_\rho) (\partial_\alpha A_\beta - \partial_\beta A_\alpha) \right) \\
&= -\frac{1}{4} \eta^{\alpha\rho} \eta^{\beta\sigma} \left(\left(\delta^\mu_\rho \delta^\nu_\sigma - \delta^\mu_\sigma \delta^\nu_\rho \right) (\partial_\alpha A_\beta - \partial_\beta A_\alpha) + (\partial_\rho A_\sigma - \partial_\sigma A_\rho) \left(\delta^\mu_\alpha \delta^\nu_\beta - \delta^\mu_\beta \delta^\nu_\alpha \right) \right) \\
&= -\frac{1}{2} \left(\partial^\mu A^\nu - \partial^\nu A^\mu + \partial^\mu A^\nu - \partial^\nu A^\mu \right) \\
&= -F^{\mu\nu}
\end{aligned} \tag{233}$$

and

$$\frac{\partial \mathcal{L}}{\partial A_\nu} = -\mu_0 J^\nu \,, \tag{234}$$

so the Lagrange equations become

$$-\partial_\mu F^{\mu\nu} + \mu_0 J^\nu = 0 \,, \tag{235}$$

or after rearrangement

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu \,. \tag{236}$$

Traditional Form of Maxwell's Equations

In terms of the fields $\vec{E}$ and $\vec{B}$, the field-strength tensor is given by

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} \,, \tag{237}$$

hence

$$\begin{aligned}
\partial_\mu F^{\mu\nu} &= \begin{pmatrix} \frac{1}{c}\frac{\partial}{\partial t} & \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{pmatrix} \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{c}\frac{\partial}{\partial t} & \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{pmatrix} \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}
\end{aligned}$$

$$= \begin{pmatrix} \vec{\nabla} \cdot \vec{E}/c \\ -\dot{E}_x/c^2 + (\vec{\nabla} \times \vec{B})_x \\ -\dot{E}_y/c^2 + (\vec{\nabla} \times \vec{B})_y \\ -\dot{E}_z/c^2 + (\vec{\nabla} \times \vec{B})_z \end{pmatrix} . \tag{238}$$

Thus, the Lagrange equations yield the **inhomogenous Maxwell equations:**

$$\vec{\nabla} \cdot \vec{E} = \mu_0 c^2 \rho = \frac{\rho}{\epsilon_0} , \tag{239}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{j} + \frac{1}{c^2} \dot{\vec{E}} = \mu_0 \left(\vec{j} + \epsilon_0 \dot{\vec{E}} \right) , \tag{240}$$

where we have used $c = 1/\sqrt{\mu_0 \epsilon_0}$.

Maxwell's homogenous equations are automatically satisfied when the electromagnetic field is described by the potentials ϕ and $\vec{A}$. In the manifestly covariant formalism the homogeneous equations do not arise from the Lagrangian, but from the identities (see Exercise 4.1):

$$\partial^\alpha F^{\beta\gamma} + \partial^\beta F^{\gamma\alpha} + \partial^\gamma F^{\alpha\beta} = 0 . \tag{241}$$

Exercise 4.1: The Homogenous Maxwell Equations

Prove that Eqs. (11.43) are automatically satisfied if the electromagnetic field strength tensor is defined as in Eq. (232).

4.4 Hamiltonian Field Theory

While the Lagrangian formalism has the advantage of manifest covariance, the transition to quantum field theory is perhaps easier to achieve in the Hamiltonian formalism, where we can rely on canonical quantization. Thus, it is worth having a look at Hamiltonian field theory.

4.4.1 Canonical Momenta and Hamiltonian Field Equations

Analogous to Hamiltonian mechanics, we can define the canonical momentum associated with a field $\varphi_\alpha(x)$ as

$$\pi^\alpha(x) \equiv \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha(x)} . \tag{242}$$

For simplicity, we are going to assume here that we do not worry about constraints and that we can find a unique solution of these equations for $\dot{\varphi}_\alpha$. Then the **Hamiltonian density** $\mathcal{H}$ or **energy density** can be defined

$$\mathcal{H} \equiv \sum_\alpha \pi^\alpha \dot{\varphi}_\alpha - \mathcal{L} , \tag{243}$$

The Hamiltonian is a functional of $\pi^\alpha, \varphi_\alpha$:

$$H\left[\varphi_\alpha, \pi^\alpha\right] = \int d^3x \, \mathcal{H} \left(\varphi_\alpha(x), \vec{\nabla}\varphi_\alpha(x), \pi^\alpha(x), \vec{\nabla}\pi^\alpha(x) \right) \tag{244}$$

is a functional of the fields and their conjugate momenta.

Using the Hamiltonian, we can write the action as

$$S = \int_\Omega d^4x \left\{ \sum_\alpha \pi^\alpha \dot{\varphi}_\alpha - \mathcal{H}\left(\varphi_\alpha, \vec{\nabla}\varphi_\alpha, \pi^\alpha, \vec{\nabla}\pi^\alpha\right) \right\} \tag{245}$$

and derive Hamilton's equations from the variatonal principle $\delta S = 0$. Varying the fields and their conjugate momenta, we have

$$\begin{aligned}
\delta S = &\int_\Omega d^4x \sum_\alpha \left\{ \pi^\alpha \delta\dot{\varphi}_\alpha + \delta\pi^\alpha \dot{\varphi}_\alpha - \frac{\partial \mathcal{H}}{\partial \varphi_\alpha}\delta\varphi_\alpha - \frac{\partial \mathcal{H}}{\partial\left(\vec{\nabla}\varphi_\alpha\right)} \cdot \delta\left(\vec{\nabla}\varphi_\alpha\right) \right. \\
&\qquad \left. - \frac{\partial \mathcal{H}}{\partial \pi^\alpha}\delta\pi^\alpha - \frac{\partial \mathcal{H}}{\partial\left(\vec{\nabla}\pi^\alpha\right)} \cdot \delta\left(\vec{\nabla}\pi^\alpha\right) \right\} \\
= &\int_\Omega d^4x \sum_\alpha \left\{ \left(-\dot{\pi}^\alpha - \frac{\partial \mathcal{H}}{\partial \varphi_\alpha} + \vec{\nabla} \cdot \frac{\partial \mathcal{H}}{\partial\left(\vec{\nabla}\varphi_\alpha\right)} \right) \delta\varphi_\alpha \right. \\
&\qquad \left. + \left(\dot{\varphi}_\alpha - \frac{\partial \mathcal{H}}{\partial \pi^\alpha} + \vec{\nabla} \cdot \frac{\partial \mathcal{H}}{\partial\left(\vec{\nabla}\pi^\alpha\right)} \right) \delta\pi^\alpha \right\} \\
= &\, 0
\end{aligned} \tag{246}$$

where we performed the usual integration by parts and exploited that the boundary terms vanish. Requiring that the coefficients of $\delta\varphi_\alpha$ and $\delta\pi^\alpha$ vanish, we obtain

$$\dot{\varphi}_\alpha = \frac{\mathcal{H}}{\pi^\alpha} - \nabla \cdot \frac{\mathcal{H}}{(\nabla\pi^\alpha)} \tag{247}$$

$$\dot{\pi}^\alpha = -\frac{\mathcal{H}}{\varphi_\alpha} + \nabla \cdot \frac{\mathcal{H}}{(\nabla\varphi_\alpha)} \tag{248}$$

which are the field equations in Hamiltonian form. They can be written more compactly as **functional derivatives** of H (cf. Box 4.2),

$$\dot{\varphi}_\alpha(x) = \frac{\delta H}{\delta\pi^\alpha(x)}, \quad \dot{\pi}^\alpha(x) = -\frac{\delta H}{\delta\varphi_\alpha(x)}\,, \tag{249}$$

mirroring the structure of Hamilton's equations for discrete systems.

4.4.2 Application to the Klein-Gordon Field

As an example, we construct the Hamiltonian density and field equations for the Klein-Gordon field (cf. Eq. (222)). Making the time derivative explicit, we can write

$$\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\vec{\nabla}\phi \cdot \vec{\nabla}\phi - \frac{m^2}{2}\phi^2 \tag{250}$$

and compute the canonical momentum via

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial\dot{\phi}(x)} = \dot{\phi}(x) \tag{251}$$

Box 4.2: Functional Derivatives

The procedure we used to study the variation of the action functional (or other functionals in the calculus of variations) can be used to define the notion of a **functional derivative**. It generalizes the derivative of a function with respect to its variables. To that of a functional with respect to the functions that are its argument.
The Hamiltonian field equations (249) can be understood as the defining equations for the functional derivatives of H. While we had no nned to use them before, the Lagrange equations for fields can be understood as the functional derivatives of the action S,

$$\frac{\delta S}{\delta \varphi_\alpha} = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)} - \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} = 0\,, \tag{I4.2-1}$$

and the discrete Lagrange equations can be written as

$$\frac{\delta S}{\delta q} = 0\,. \tag{I4.2-2}$$

Consequently,

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}\vec{\nabla}\phi \cdot \vec{\nabla}\phi + \frac{m^2}{2}\phi^2 \tag{252}$$

and Hamilton's equations read

$$\dot{\phi}(x) = \pi(x), \quad \dot{\pi}(x) = \vec{\nabla}^2\phi(x) - m^2\phi(x)\,. \tag{253}$$

Taking another time derivative of the first equation and plugging in $\dot{\pi}(x)$, we obtain the Klein-Gordon equation:

$$\ddot{\phi}(x) - \vec{\nabla}^2\phi(x) + m^2\phi(x) = (\Box + m^2)\phi(x) = 0\,, \tag{254}$$

(recall $\hbar = c = 1$).

4.5 Noether's Theorem

The version of Noether's theorem that we discussed earlier is really just a special case of the more general version, which is one of the most powerful tools of classical and quantum field theory.

4.5.1 Infinitesimal Transformation

Let us consider the infinitesimal transformation

$$x^\mu \longrightarrow x'^\mu = x^\mu + \Delta x^\mu \tag{255}$$

$$\varphi_\alpha(x) \longrightarrow \varphi'_\alpha\left(x'\right) = \varphi_\alpha(x) + \Delta\varphi_\alpha(x) \tag{256}$$

with $\Delta x^\mu = \Delta x^\mu(x)$. The variation Δ differs from the usual variation δ because it takes into account how the field is affected by the change of *both* its functional form and its argument. This is the generalization of the treatment of transformations of the time variable in our earlier discussion of Noether's theorem to all spacetime variables. The variation due to the change of form alone is defined by

$$\delta\varphi_\alpha(x) = \varphi'_\alpha(x) - \varphi_\alpha(x)\,, \tag{257}$$

so we have

$$\Delta\varphi_\alpha(x) = \varphi'_\alpha\left(x'\right) - \varphi_\alpha\left(x'\right) + \varphi_\alpha\left(x'\right) - \varphi_\alpha(x) = \delta\varphi_\alpha\left(x'\right) + \partial_\mu\varphi_\alpha(x)\Delta x^\mu\,. \tag{258}$$

Neglecting terms of second order in the infinitesimal variations, this reduces to

$$\Delta\varphi_\alpha(x) = \delta\varphi_\alpha(x) + (\partial_\mu\varphi_\alpha)\Delta x^\mu\,. \tag{259}$$

We need to be careful here, because the partial derivatives ∂_μ commute with the δ variation, but *not* with the Δs due to the changes in the arguments! Thus,

$$\Delta\left(\partial_\mu\varphi_\alpha\right) \neq \partial_\mu\left(\Delta\varphi_\alpha\right)\,. \tag{260}$$

Applying Eq. (259) to $\partial_\mu\varphi_\alpha$ we find

$$\Delta[\partial_\mu\varphi_\alpha(x)] = \delta[\partial_\mu\varphi_\alpha(x)] + [\partial_\nu\partial_\mu\varphi_\alpha]\Delta x^\nu\,. \tag{261}$$

4.5.2 Invariance of the Action

The variation of the action is now defined by

$$\Delta S = \int_{\Omega'} d^4x'\,\mathcal{L}\left(\varphi'_\alpha\left(x'\right), \partial'_\mu\varphi'_\alpha\left(x'\right), x'\right) - \int_\Omega d^4x\,\mathcal{L}\left(\varphi_\alpha(x), \partial_\mu\varphi_\alpha(x), x\right)\,. \tag{262}$$

The varied Lagrangian is given by

$$\begin{aligned}
\mathcal{L}' &\equiv \mathcal{L}\left(\varphi'_\alpha\left(x'\right), \partial'_\mu\varphi'_\alpha\left(x'\right), x'\right) \\
&= \mathcal{L}\left(\varphi_\alpha(x) + \Delta\varphi_\alpha(x), \partial_\mu\varphi_\alpha(x) + \Delta\partial_\mu\varphi_\alpha(x), x + \Delta x\right) \\
&= \mathcal{L}\left(\varphi_\alpha(x), \partial_\mu\varphi_\alpha(x), x\right) + \frac{\partial\mathcal{L}}{\partial\varphi_\alpha}\Delta\varphi_\alpha + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_\alpha)}\Delta\partial_\mu\varphi_\alpha + \frac{\partial\mathcal{L}}{\partial x^\mu}\Delta x^\mu \\
&= \mathcal{L} + \frac{\partial\mathcal{L}}{\partial\varphi_\alpha}\delta\varphi_\alpha + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_\alpha)}\delta(\partial_\mu\varphi_\alpha) + \frac{d\mathcal{L}}{dx^\mu}\Delta x^\mu \\
&\equiv \mathcal{L} + \delta\mathcal{L} + \frac{d\mathcal{L}}{dx^\mu}\Delta x^\mu\,,
\end{aligned} \tag{263}$$

where we have used Eqs. (259) and (261), and introduced the "total partial derivative" with respect to x^μ:

$$\frac{d\mathcal{L}}{dx^\mu} = \frac{\partial\mathcal{L}}{\partial\varphi_\alpha}\frac{\partial\varphi_\alpha}{\partial x^\mu} + \frac{\partial\mathcal{L}}{\partial(\partial_\nu\varphi_\alpha)}\frac{\partial(\partial_\nu\varphi_\alpha)}{\partial x^\mu} + \frac{\partial\mathcal{L}}{\partial x^\mu}\,. \tag{264}$$

Next, we consider the spacetime volume element. Using Eq. (255) and the following relation for infinitesimal changes in a matrix,

$$\boldsymbol{B} = \mathbb{1} + \epsilon\boldsymbol{A} \qquad \Rightarrow \qquad \det\boldsymbol{B} = \mathbb{1} + \epsilon\,\mathrm{tr}\,\boldsymbol{A}\,, \tag{265}$$

we see that

$$d^4x' = \frac{\partial\left(x'^0, x'^1, x'^2, x'^3\right)}{\partial(x^0, x^1, x^2, x^3)}d^4x = \left(1 + \frac{\partial\Delta x^\mu}{\partial x^\mu}\right)d^4x \tag{266}$$

Plugging our intermediate results into Eq. (262) and keeping terms up to linear order in the variations, we obtain

$$\Delta S = \int_\Omega d^4x\left[\delta\mathcal{L} + \frac{d\mathcal{L}}{dx^\mu}\Delta x^\mu + \mathcal{L}\frac{\partial\Delta x^\mu}{\partial x^\mu}\right] = \int_\Omega d^4x\left[\delta\mathcal{L} + \frac{d}{dx^\mu}\left(\mathcal{L}\Delta x^\mu\right)\right] \tag{267}$$

Using the field equations of motion

$$\frac{\partial \mathcal{L}}{\partial \varphi_\alpha} = \frac{d}{dx^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)}\right), \tag{268}$$

we find

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \varphi_\alpha}\delta\varphi_\alpha + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)}\delta(\partial_\mu\varphi_\alpha) \\ &= \frac{d}{dx^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)}\right)\delta\varphi_\alpha + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)}\partial_\mu\left(\delta\varphi_\alpha\right) \\ &= \frac{d}{dx^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)}\delta\varphi_\alpha\right).\end{aligned} \tag{269}$$

Plugging this into Eq. (267), the invariance condition for the action becomes

$$\Delta S = \int_\Omega d^4x \; \frac{d}{dx^\mu}\left\{\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)}\delta\varphi_\alpha + \mathcal{L}\Delta x^\mu\right\} = 0\,. \tag{270}$$

Since we can choose an arbitray spacetime volume, the integrand needs to vanish, and we obtain a conservation law for the **Noether current**,

$$J^\mu \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)}\delta\varphi_\alpha + \mathcal{L}\Delta x^\mu \tag{271}$$

in the form of the **continuity equation**

$$\partial_\mu J^\mu = 0\,. \tag{272}$$

4.5.3 Conserved Quantities

It is usually more convenient to express the four-current conservation in terms of the infinitesimal parameters of the transformation. Suppose the transformation (255), (256) is specified by R independent infinitesimal parameters $\epsilon_1, \ldots, \epsilon_R$ in the form

$$\Delta x^\mu = \sum_{r=1}^{R} X^{\mu(r)}\epsilon_r \equiv X^{\mu(r)}\epsilon_r, \quad \Delta\varphi_\alpha = \sum_{r=1}^{R}\Psi_\alpha^{(r)}\epsilon_r \equiv \Psi_\alpha^{(r)}\epsilon_r\,. \tag{273}$$

The indices α and r of the fields and of the transformation parameters may or may not have tensor character,but we still adhere to Einstein's summation convention. Substituting the expressions (273) into Eq. (259) we have

$$\Delta x^\mu = X^{\mu(r)}\epsilon_r, \quad \delta\varphi_\alpha = \left(\Psi_\alpha^{(r)} - (\delta_\nu\varphi_\alpha)X^{\nu(r)}\right)\epsilon_r \tag{274}$$

and when we insert this into Eq. (262), we obtain

$$\Delta S = -\int_\Omega d^4x \; \epsilon_r \frac{d}{dx^\mu}\Theta^{\mu(r)} = 0 \tag{275}$$

with

$$\Theta^{\mu(r)} = -\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)} \left(\Psi_\alpha^{(r)} - (\partial_\nu \varphi_\alpha) X^{\nu(r)} \right) - \mathcal{L} X^{\mu(r)} . \tag{276}$$

Since the integration region and the parameters ϵ_r are arbitrary, Eq. (275) implies that we have R local conservation laws

$$\partial_\mu \Theta^{\mu(r)} = 0, \quad r = 1, \ldots, R , \tag{277}$$

and noting $\Theta^{\mu(r)} = \left(\Theta^{0(r)}, \vec{\Theta}^{(r)} \right)$ we can write this in the form of the continuity equations

$$\partial_0 \Theta^{0(r)} + \nabla \cdot \mathbf{\Theta}^{(r)} = 0 , \qquad r = 1, \ldots, R \tag{278}$$

Now

$$\frac{d}{dx^0} \int_V d^3x \; \Theta^{0(r)} = \int_V d^3x \; \partial_0 \Theta^{0(r)} = -\int_V d^3x \; \vec{\nabla} \cdot \Theta^{(r)} = -\oint_{\partial V} d\vec{A} \cdot \Theta^{(r)} \tag{279}$$

where we have again used the divergence theorem. If V is the entire three-dimensional space and the fields vanish sufficiently fast at infinity, the surface integral vanishes and the volume integrals

$$C^{(r)} = \int d^3x \; \Theta^{0(r)} , \qquad r = 1, \ldots, R , \tag{280}$$

are conserved quantities, since they are time independent. Thus, for each R-parametric infinitesimal transformation of coordinates and fields that leaves the action invariant, we obtain R conserved quantites $C^{(r)}$ which are known as **Noether charges**.

4.5.4 Conservation of Four-Momentum

As an example, we consider the spacetime translation

$$x'^\mu = x^\mu + \epsilon^\mu \tag{281}$$

which does not modify the fields:

$$\Delta \varphi_\alpha = 0 . \tag{282}$$

The Jacobian of the transformation is the identity, hence the action is invariant if the Lagrangian does not depend explicitly oon spacetime coordinates, generalizing our previous results for the explicit time dependence of the Lagrangian or Hamiltonian.

In the present example, the index r of the transformation parameters has four-vector character and Eq. (273) implies

$$X^{\mu(\nu)} = \eta^{\mu\nu}, \quad \Psi_\alpha^{(\nu)} = 0 . \tag{283}$$

Substituting these expressions into Eq. (276), we obtain the **(canonical) energy-momentum tensor**

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)} (\partial_\rho \varphi_\alpha) \eta^{\rho\nu} - \mathcal{L} \eta^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_\alpha)} \frac{\partial \varphi_\alpha}{\partial x_\nu} - \mathcal{L} \eta^{\mu\nu} \tag{284}$$

The conserved Noether charges constitute the four-vector

$$P^\nu = \int d^3x \; T^{0\nu} . \tag{285}$$

The zeroth component of this four-vector is the spatial integral of the Hamiltonian density $\mathcal{H}$, i.e., the field energy. Covariance and the fact that the conservation of linear momentum is associated with invariance under translations in space imply that $P^\nu = (P^0\,\vec{p})$, hence it is the field four-momentum. Thus, the name for $T^{\mu\nu}$ is appropriate.

Exercise 4.2: The Homogenous Maxwell Equations

Show that the energy-momentum tensor of the scalar field described by the Klein-Gordon Lagrangian (222) is given by

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \frac{1}{2}\left(\partial_\rho\phi\partial^\rho\phi - m^2\phi^2\right)\eta^{\mu\nu} \qquad \text{(E4.2-1)}$$

and find the expression for the (conserved) field momentum.

4.5.5 Gauge Invariance and Conservation of Electric Charge

Let us now consider a complex scalar field described by the Lagrangian

$$\mathcal{L} = \partial_\mu\phi^*\,\partial^\mu\phi - m^2\phi^*\phi \qquad (286)$$

which is easily seen to be invariant under the one-parameter transformation

$$\phi' = e^{i\lambda}\phi, \quad \phi^{*\prime} = e^{-i\lambda}\phi^*\,, \qquad (287)$$

where λ is an arbitrary real number. This transformation is known as a **global gauge transformation**. Since the phase factor can be understood as a one-dimensional unitary "matrix", we say that the Lagrangian has a global $U(1)$ gauge symmetry.

The infinitesimal version of the transformation with $\lambda = \epsilon$ is

$$\Delta\phi = i\epsilon\phi, \quad \Delta\phi^* = -i\epsilon\phi^*\,, \qquad (288)$$

which implies

$$X^{\mu(1)} = 0, \quad \Psi_1^{(1)} = i, \quad \Psi_2^{(1)} = -i \qquad (289)$$

with $\varphi_1 = \phi, \varphi_2 = \phi^*$ and the index r taking the single value $r = 1$. Since the coordinates remain unchanged, invariance of the Lagrangian implies invariance of the action.

The Noether current (276) has the form

$$J^\mu \equiv \Theta^{\mu(1)} = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_1)}\Psi_1^{(1)} - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_2)}\Psi_2^{(1)} = -i\left(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*\right) \qquad (290)$$

Computing the spatial integral, we see that the conserved Noether charge is

$$Q = i\int d^3x\left(\phi\dot{\phi}^* - \dot{\phi}\phi^*\right)\,. \qquad (291)$$

It is intepreted as the **electric charge** of the scalar particles described by the Lagrangian (286). Equations (290) and (291) imply that a *real* scalar field cannot be used to describe charged particles that couple to the electromagnetic field because J^μ and Q would vanish.