

Chapter 10. A Bit More of Analytical Mechanics

This concluding chapter reviews two alternative approaches to analytical mechanics, whose major value is a closer parallel to quantum mechanics in general and its quasiclassical (WKB) approximation in particular. One of them, the Hamiltonian formalism, is also convenient for the derivation of an important asymptotic result, the adiabatic invariance, for classical systems with slowly changing parameters.

10.1. Hamilton equations

Throughout this course, we have seen how analytical mechanics, in its Lagrangian form, is invaluable for solving various particular problems of classical mechanics. Now let us discuss several alternative formulations¹ that may not be much more useful for this purpose, but shed additional light on possible extensions of classical mechanics, most importantly to quantum mechanics.

As was already discussed in Sec. 2.3, the partial derivative $p_j \equiv \partial L / \partial \dot{q}_j$ participating in the Lagrange equation (2.19),

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0, \quad (10.1)$$

may be considered as the generalized momentum corresponding to the generalized coordinate q_j , and the full set of these momenta may be used to define the Hamiltonian function (2.32):

$$H \equiv \sum_j p_j \dot{q}_j - L. \quad (10.2)$$

Now let us rewrite the full differential of this function² in the following form:

$$\begin{aligned} dH &= d\left(\sum_j p_j \dot{q}_j - L\right) = \sum_j [d(p_j) \dot{q}_j + p_j d(\dot{q}_j)] - dL \\ &= \sum_j [d(p_j) \dot{q}_j + p_j d(\dot{q}_j)] - \left[\frac{\partial L}{\partial t} dt + \sum_j \left(\frac{\partial L}{\partial q_j} dq_j + \frac{\partial L}{\partial \dot{q}_j} d(\dot{q}_j) \right) \right]. \end{aligned} \quad (10.3)$$

According to the definition of the generalized momentum, the second terms of each sum over j in the last expression cancel each other, while according to the Lagrange equation (1), the derivative $\partial L / \partial q_j$ is equal to \dot{p}_j , so

$$dH = -\frac{\partial L}{\partial t} dt + \sum_j (\dot{q}_j dp_j - \dot{p}_j dq_j). \quad (10.4)$$

So far, this is just a universal identity. Now comes the main trick of Hamilton's approach: let us consider H as a function of the following independent arguments: time t , the generalized coordinates q_j ,

¹ Due to not only William Rowan Hamilton (1805-1865), but also Carl Gustav Jacob Jacobi (1804-1851).

² Actually, this differential was already spelled out (but partly and implicitly) in Sec. 2.3 – see Eqs. (2.33)-(2.35).

and the generalized momenta p_j – rather than generalized velocities \dot{q}_j as in the Lagrangian formalism. With this new commitment, the general “chain rule” of differentiation of a function of several arguments gives

$$dH = \frac{\partial H}{\partial t} dt + \sum_j \left(\frac{\partial H}{\partial q_j} dq_j + \frac{\partial H}{\partial p_j} dp_j \right), \quad (10.5)$$

where dt , dq_j , and dp_j are independent differentials. Since Eq. (5) should be valid for any choice of these argument differentials, it should hold in particular if they correspond to the real law of motion, for which Eq. (4) is valid as well. The comparison of Eqs. (4) and (5) gives us three relations:

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (10.6)$$

$$\dot{q}_j = \frac{\partial H}{\partial p_j}, \quad p_j = -\frac{\partial H}{\partial q_j}. \quad (10.7) \quad \text{Hamilton equations}$$

Comparing the first of them with Eq. (2.35), we see that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}, \quad (10.8)$$

meaning that the function $H(t, q_j, p_j)$ can change in time only via its explicit dependence on t . Two Eqs. (7) are even more substantial: provided that such function $H(t, q_j, p_j)$ has been calculated, they give us two first-order differential equations (called the *Hamilton equations*) for the time evolution of the generalized coordinate and generalized momentum of each degree of freedom of the system.³

Let us have a look at these equations for the simplest case of a system with one degree of freedom, with the Lagrangian function (3.3):

$$L = \frac{m_{\text{ef}}}{2} \dot{q}^2 - U_{\text{ef}}(q, t). \quad (10.9)$$

In this case, $p \equiv \partial L / \partial \dot{q} = m_{\text{ef}} \dot{q}$, and $H \equiv p\dot{q} - L = m_{\text{ef}} \dot{q}^2 / 2 + U_{\text{ef}}(q, t)$. To honor our new commitment, we need to express the Hamiltonian function explicitly via t , q , and p (rather than \dot{q}). From the above expression for p , we immediately have $\dot{q} = p / m_{\text{ef}}$; plugging this expression back to Eq. (9), we get

$$H = \frac{p^2}{2m_{\text{ef}}} + U_{\text{ef}}(q, t). \quad (10.10)$$

Now we can spell out Eqs. (7) for this particular case:

$$\dot{q} \equiv \frac{\partial H}{\partial p} = \frac{p}{m_{\text{ef}}}, \quad (10.11)$$

$$\dot{p} \equiv -\frac{\partial H}{\partial q} = -\frac{\partial U_{\text{ef}}}{\partial q}. \quad (10.12)$$

³ Of course, the right-hand side of each equation (7) may include coordinates and momenta of other degrees of freedom as well, so the equations of motion for different j are generally coupled.

While the first of these equations just repeats the definition of the generalized momentum corresponding to the coordinate q , the second one gives the equation of momentum's change. Differentiating Eq. (11) over time, and plugging Eq. (12) into the result, we get:

$$\ddot{q} = \frac{\dot{p}}{m_{\text{ef}}} = -\frac{1}{m_{\text{ef}}} \frac{\partial U_{\text{ef}}}{\partial q}. \quad (10.13)$$

So, we have returned to the same equation (3.4) that had been derived from the Lagrangian approach.⁴

Thus, Hamiltonian formalism does not give much help for the solution of this problem – and indeed most problems of classical mechanics. (This is why its discussion had been postponed until the very end of this course.) Moreover, since the Hamiltonian function $H(t, q_j, p_j)$ does not include generalized velocities explicitly, the phenomenological introduction of dissipation in this approach is less straightforward than that in the Lagrangian equations, whose precursor form (2.17) is valid for dissipative forces as well. However, the Hamilton equations (7), which treat the generalized coordinates and momenta in a manifestly symmetric way, are heuristically fruitful – besides being very appealing aesthetically. This is especially true in the cases where these arguments participate in H in a similar way. For example, in the very important case of a dissipation-free linear (“harmonic”) oscillator, for which $U_{\text{ef}} = \kappa_{\text{ef}} q^2/2$, Eq. (10) gives the symmetric form

$$H = \frac{p^2}{2m_{\text{ef}}} + \frac{\kappa_{\text{ef}} x^2}{2} \equiv \frac{p^2}{2m_{\text{ef}}} + \frac{m_{\text{ef}} \omega_0^2 x^2}{2}, \quad \text{where } \omega_0^2 \equiv \frac{\kappa_{\text{ef}}}{m_{\text{ef}}}. \quad (10.14)$$

The Hamilton equations (7) for this system preserve that symmetry, especially evident if we introduce the normalized momentum $\rho \equiv p/m_{\text{ef}}\omega_0$ (already used in Secs. 5.6 and 9.2):

$$\frac{dq}{dt} = \omega_0 \rho, \quad \frac{d\rho}{dt} = -\omega_0 q. \quad (10.15)$$

More practically, the Hamilton approach gives additional tools for the search for the integrals of motion. To see that, let us consider the full time derivative of an arbitrary function $f(t, q_j, p_j)$:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_j \left(\frac{\partial f}{\partial q_j} \dot{q}_j + \frac{\partial f}{\partial p_j} \dot{p}_j \right). \quad (10.16)$$

Plugging in \dot{q}_j and \dot{p}_j from the Hamilton equations (7), we get

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_j \left(\frac{\partial H}{\partial p_j} \frac{\partial f}{\partial q_j} - \frac{\partial H}{\partial q_j} \frac{\partial f}{\partial p_j} \right) \equiv \frac{\partial f}{\partial t} + \{H, f\}. \quad (10.17)$$

The last term on the right-hand side of this expression is the so-called *Poisson bracket*,⁵ and is defined, for two arbitrary functions $f(t, q_j, p_j)$ and $g(t, q_j, p_j)$, as

Dynamics
of arbitrary
variable

⁴ The reader is strongly encouraged to perform a similar check for a few more problems, for example those listed at the end of the chapter, to get a better feeling of how the Hamiltonian formalism works.

⁵ Named after Siméon Denis Poisson (1781-1840), of the Poisson equation and the Poisson statistical distribution fame.

$$\{g, f\} \equiv \sum_j \left(\frac{\partial g}{\partial p_j} \frac{\partial f}{\partial q_j} - \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_j} \right). \quad (10.18) \quad \text{Poisson bracket}$$

From this definition, one can readily verify that besides evident relations $\{f, f\} = 0$ and $\{f, g\} = -\{g, f\}$, the Poisson brackets obey the following important *Jacobi identity*:

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \quad (10.19)$$

Now let us use these relations for a search for integrals of motion. First, Eq. (17) shows that if a function f does not depend on time explicitly, and

$$\{H, f\} = 0, \quad (10.20)$$

then $df/dt = 0$, i.e. that function is an integral of motion. Moreover, it turns out that if we already know two integrals of motion, say f and g , then the following function,

$$F \equiv \{f, g\}, \quad (10.21)$$

is also an integral of motion – the so-called *Poisson theorem*. In order to prove it, we may use the Jacobi identity (19) with $h = H$. Next, using Eq. (17) to express the Poisson brackets $\{g, H\}$, $\{H, g\}$, and $\{H, \{f, g\}\} = \{H, F\}$ via the full and partial time derivatives of the functions f , g , and F , we get

$$\left\{ f, \frac{\partial g}{\partial t} - \frac{dg}{dt} \right\} + \left\{ g, \frac{df}{dt} - \frac{\partial f}{\partial t} \right\} + \frac{dF}{dt} - \frac{\partial F}{\partial t} = 0, \quad (10.22)$$

so if f and g are indeed integrals of motion, i.e., $df/dt = dg/dt = 0$, then

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \left\{ g, \frac{\partial f}{\partial t} \right\} - \left\{ f, \frac{\partial g}{\partial t} \right\} = \frac{\partial F}{\partial t} - \left[\left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} \right]. \quad (10.23)$$

Plugging Eq. (21) into the first term of the right-hand side of this equation, and differentiating it by parts, we get $dF/dt = 0$, i.e. F is indeed an integral of motion as well.

Finally, one more important role of the Hamilton formalism is that it allows one to trace the close formal connection between classical and quantum mechanics. Indeed, using Eq. (18) to calculate the Poisson brackets of the generalized coordinates and momenta, we readily get

$$\{q_j, q_{j'}\} = 0, \quad \{p_j, p_{j'}\} = 0, \quad \{q_j, p_{j'}\} = -\delta_{jj'}. \quad (10.24)$$

In quantum mechanics, the operators of these variables (“observables”) obey commutation relations⁶

$$[\hat{q}_j, \hat{q}_{j'}] = 0, \quad [\hat{p}_j, \hat{p}_{j'}] = 0, \quad [\hat{q}_j, \hat{p}_{j'}] = i\hbar \delta_{jj'}, \quad (10.25)$$

where the definition of the commutator, $[\hat{g}, \hat{f}] \equiv \hat{g}\hat{f} - \hat{f}\hat{g}$, is to a certain extent⁷ similar to that (18) of the Poisson bracket. We see that the classical relations (24) are similar to the quantum-mechanical relations (25) if the following parallel has been made:

⁶ See, e.g., QM Sec. 2.1.

$$\{g, f\} \leftrightarrow \frac{i}{\hbar} [\hat{g}, \hat{f}]. \quad (10.26)$$

This analogy extends well beyond Eqs. (24)-(25). For example, by making the replacement (26) in Eq. (17), we get

$$\frac{d\hat{f}}{dt} = \frac{\partial \hat{f}}{\partial t} + \frac{i}{\hbar} [\hat{H}, \hat{f}], \quad \text{i.e. } i\hbar \frac{d\hat{f}}{dt} = i\hbar \frac{\partial \hat{f}}{\partial t} + [\hat{f}, \hat{H}], \quad (10.27)$$

which is the correct equation of operator evolution in the Heisenberg picture of quantum mechanics.⁸ The parallel (26) may give important clues in the search for the proper quantum-mechanical operator of a given observable – which is not always elementary.

10.2. Adiabatic invariance

One more application of the Hamiltonian formalism in classical mechanics is the solution of the following problem.⁹ Earlier in the course, we already studied some effects of time variation of parameters of a single oscillator (Sec. 5.5) and coupled oscillators (Sec. 6.5). However, those discussions were focused on the case when the parameter variation speed is comparable with the own oscillation frequency (or frequencies) of the system. Another practically important case is when some system's parameter (let us call it λ) is changed much more slowly (*adiabatically*¹⁰),

$$\left| \frac{\dot{\lambda}}{\lambda} \right| \ll \frac{1}{\mathcal{T}}, \quad (10.28)$$

where \mathcal{T} is a typical period of oscillations in the system. Let us consider a 1D system whose Hamiltonian $H(q, p, \lambda)$ depends on time only via such a slow evolution of such parameter $\lambda = \lambda(t)$, and whose initial energy restricts the system's motion to a finite coordinate interval – see, e.g., Fig. 3.2c.

Then, as we know from Sec. 3.3, if the parameter λ is constant, the system performs a periodic (though not necessarily sinusoidal) motion back and forth the q -axis, or, in a different language, along a closed trajectory on the phase plane $[q, p]$ – see Fig. 1.¹¹ According to Eq. (8), in this case, H is constant along the trajectory. (To distinguish this particular *value* of H from the Hamiltonian *function* as such, I will call it E , implying that this constant coincides with the full mechanical energy E – as does for the Hamiltonian (10), though this assumption is not necessary for the calculation made below.)

The oscillation period \mathcal{T} may be calculated as a contour integral along this closed trajectory:

⁷ There is, of course, a conceptual difference between the “usual” products of the function derivatives participating in the Poisson brackets, and the operator “products” (meaning their sequential action on a state vector) forming the commutator.

⁸ See, e.g., QM Sec. 4.6.

⁹ Various aspects of this problem and its quantum-mechanical extensions were first discussed by L. Le Cornu (1895), Lord Rayleigh (1902), H. Lorentz (1911), P. Ehrenfest (1916), and M. Born and V. Fock (1928).

¹⁰ This term is also used in thermodynamics and statistical mechanics, where it implies not only a slow parameter variation (if any) but also thermal insulation of the system – see, e.g., SM Sec. 1.3. Evidently, the latter condition is irrelevant in our current context.

¹¹ As a reminder, we discussed such phase-plane representations in Chapter 5 – see, e.g., Figs. 5.5, 5.9, and 5.16.

$$\tau \equiv \int_0^\tau dt = \oint \frac{dt}{dq} dq \equiv \oint \frac{1}{\dot{q}} dq. \quad (10.29)$$

Using the first of the Hamilton equations (7), we may represent this integral as

$$\tau = \oint \frac{1}{\partial H / \partial p} dq. \quad (10.30)$$

At each given point q , $H = E$ is a function of p alone, so we may flip the partial derivative in the denominator just as the full derivative, and rewrite Eq. (30) as

$$\tau = \oint \frac{\partial p}{\partial E} dq. \quad (10.31)$$

For the particular Hamiltonian (10), this relation is immediately reduced to Eq. (3.27), now in the form of a contour integral:

$$\tau = \left(\frac{m_{\text{ef}}}{2} \right)^{1/2} \oint \frac{1}{[E - U_{\text{ef}}(q)]^{1/2}} dq. \quad (10.32)$$

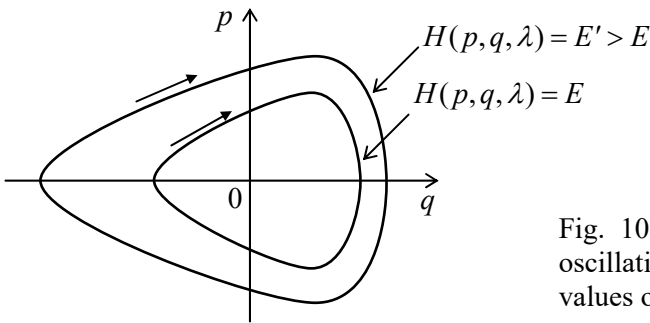


Fig. 10.1. Phase-plane representation of periodic oscillations of a 1D Hamiltonian system, for two values of energy (schematically).

Naively, it may look that these formulas may be also used to find the motion period's change when the parameter λ is being changed adiabatically, for example, by plugging the given functions $m_{\text{ef}}(\lambda)$ and $U_{\text{ef}}(q, \lambda)$ into Eq. (32). However, there is no guarantee that the energy E in that integral would stay constant as the parameter changes, and indeed we will see below that this is not necessarily the case. Even more interestingly, in the most important case of the harmonic oscillator ($U_{\text{ef}} = \kappa_{\text{ef}} q^2 / 2$), whose oscillation period τ does not depend on E (see Eq. (3.29) and its discussion), its variation in the adiabatic limit (28) may be readily predicted: $\mathcal{T}(\lambda) = 2\pi / \omega_0(\lambda) = 2\pi [m_{\text{ef}}(\lambda) / \kappa_{\text{ef}}(\lambda)]^{1/2}$, but the dependence of the oscillation energy E (and hence of the oscillation amplitude) on λ is not immediately obvious.

In order to address this issue, let us use Eq. (8) (with $E = H$) to represent the rate of the energy change with $\lambda(t)$, i.e. in time, as

$$\frac{dE}{dt} = \frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda} \frac{d\lambda}{dt}. \quad (10.33)$$

Since we are interested in a very slow (adiabatic) time evolution of energy, we can average Eq. (33) over fast oscillations in the system, for example over one oscillation period \mathcal{T} , treating $d\lambda/dt$ as a constant during this averaging. (This is the most critical point of this argumentation, because at any non-

vanishing rate of parameter change the oscillations are, strictly speaking, non-periodic.¹²⁾ The averaging yields

$$\frac{d\overline{E}}{dt} = \frac{d\lambda}{dt} \frac{\partial \overline{H}}{\partial \lambda} \equiv \frac{d\lambda}{dt} \frac{1}{\tau} \int_0^{\tau} \frac{\partial H}{\partial \lambda} dt. \quad (10.34)$$

Transforming this time integral to the contour one, just as we did at the transition from Eq. (29) to Eq. (30), and then using Eq. (31) for τ , we get

$$\frac{d\overline{E}}{dt} = \frac{d\lambda}{dt} \frac{\oint \frac{\partial H / \partial \lambda}{\partial H / \partial p} dq}{\oint \frac{\partial p}{\partial E} dq}. \quad (10.35)$$

At each point q of the contour, H is a function of not only λ , but also of p , which may be also λ -dependent, so if E is fixed, the partial differentiation of the relation $E = H$ over λ yields

$$\frac{\partial H}{\partial \lambda} + \frac{\partial H}{\partial p} \frac{\partial p}{\partial \lambda} = 0, \quad \text{i.e.} \quad \frac{\partial H / \partial \lambda}{\partial H / \partial p} = -\frac{\partial p}{\partial \lambda}. \quad (10.36)$$

Plugging the last relation to Eq.(35), we get

$$\frac{d\overline{E}}{dt} = -\frac{d\lambda}{dt} \frac{\oint \frac{\partial p}{\partial \lambda} dq}{\oint \frac{\partial p}{\partial E} dq}. \quad (10.37)$$

Since the left-hand side of Eq. (37) and the derivative $d\lambda/dt$ do not depend on q , we may move them into the integrals over q as constants, and rewrite Eq. (37) as

$$\oint \left(\frac{\partial p}{\partial E} \frac{d\overline{E}}{dt} + \frac{\partial p}{\partial \lambda} \frac{d\lambda}{dt} \right) dq = 0. \quad (10.38)$$

Now let us consider the following integral over the same phase-plane contour,

Action
variable

$$J \equiv \frac{1}{2\pi} \oint pdq, \quad (10.39)$$

called the *action variable*. Just to understand its physical sense, let us calculate J for a harmonic oscillator (14). As we know very well from Chapter 5, for such an oscillator, $q = A \cos \Psi$, $p = -m_{\text{ef}} \omega_0 A \sin \Psi$ (with $\Psi = \omega_0 t + \text{const}$), so J may be easily expressed either via the oscillations' amplitude A , or via their energy $E = H = m_{\text{ef}} \omega_0^2 A^2 / 2$:

$$J = \frac{1}{2\pi} \oint pdq = \frac{1}{2\pi} \int_{\Psi=0}^{\Psi=2\pi} (-m_{\text{ef}} \omega_0 A \sin \Psi) d(A \cos \Psi) = \frac{m_{\text{ef}} \omega_0}{2} A^2 = \frac{E}{\omega_0}. \quad (10.40)$$

¹² Because of the implied nature of this conjecture (which is very close to the assumptions made at the derivation of the reduced equations in Sec. 5.3), new, more strict (but also much more cumbersome) proofs of the final Eq. (42) are still being offered in literature – see, e.g., C. Wells and S. Siklos, *Eur. J. Phys.* **28**, 105 (2007) and/or A. Lobo *et al.*, *Eur. J. Phys.* **33**, 1063 (2012).

Returning to a general system with adiabatically changed parameter λ , let us use the definition of J , Eq. (39), to calculate its time derivative, again taking into account that at each point q of the trajectory, p is a function of E and λ :

$$\frac{dJ}{dt} = \frac{1}{2\pi} \oint \frac{dp}{dt} dq = \frac{1}{2\pi} \oint \left(\frac{\partial p}{\partial E} \frac{dE}{dt} + \frac{\partial p}{\partial \lambda} \frac{d\lambda}{dt} \right) dq. \quad (10.41)$$

Within the accuracy of our approximation, in which the contour integrals (38) and (41) are calculated along a closed trajectory, the factor dE/dt is indistinguishable from its time average, and these integrals coincide, so the result (38) is applicable to Eq. (41) as well. Hence, we have finally arrived at a very important result: at a slow parameter variation, $dJ/dt = 0$, i.e. the action variable remains constant:

$$J = \text{const.} \quad (10.42) \quad \text{Adiabatic invariance}$$

This is the famous *adiabatic invariance*.¹³ In particular, according to Eq. (40), in a harmonic oscillator, the energy of oscillations changes proportionately to its own (slowly changed) frequency.

Before moving on, let me briefly note that the adiabatic invariance is not the only application of the action variable J . Since the initial choice of generalized coordinates and velocities (and hence the generalized momenta) in analytical mechanics is arbitrary (see Sec. 2.1), it is almost evident that J may be taken for a new generalized momentum corresponding to a certain new generalized coordinate Θ ,¹⁴ and that the pair $\{J, \Theta\}$ should satisfy the Hamilton equations (7), in particular,

$$\frac{d\Theta}{dt} = \frac{\partial H}{\partial J}. \quad (10.43)$$

Following the commitment of Sec. 1 (made there for the “old” arguments q_j, p_j), before the differentiation on the right-hand side of Eq. (43), H should be expressed as a function (besides t) of the “new” arguments J and Θ . For time-independent Hamiltonian systems, H is uniquely defined by J – see, e.g., Eq. (40). Hence in this case the right-hand side of Eq. (43) does not depend on either t or Θ , so according to that equation, Θ (called the *angle variable*) is a linear function of time:

$$\Theta = \frac{\partial H}{\partial J} t + \text{const.} \quad (10.44)$$

For a harmonic oscillator, according to Eq. (40), the derivative $\partial H/\partial J = \partial E/\partial J$ is just $\omega_0 \equiv 2\pi/\tau$, so $\Theta = \omega_0 t + \text{const.}$, i.e. it is just the full phase Ψ that was repeatedly used in this course – especially in Chapter 5. It may be shown that a more general form of this relation,

$$\frac{\partial H}{\partial J} = \frac{2\pi}{\tau}, \quad (10.45)$$

¹³ For certain particular oscillators, e.g., a point pendulum, Eq. (42) may be also proved directly – an exercise highly recommended to the reader.

¹⁴ This, again, is a plausible argument but not a strict proof. Indeed: though, according to its definition (39), J is nothing more than a sum of several (formally, the infinite number of) values of the momentum p , they are not independent, but have to be selected on the same closed trajectory on the phase plane. For more mathematical vigor, the reader is referred to Sec. 45 of *Mechanics* by Landau and Lifshitz (which was repeatedly cited above), which discusses the general rules of the so-called *canonical transformations* from one set of Hamiltonian arguments to another one – say from $\{p, q\}$ to $\{J, \Theta\}$.

is valid for an arbitrary system described by Eq. (10). Thus, Eq. (44) becomes

$$\Theta = 2\pi \frac{t}{T} + \text{const} . \quad (10.46)$$

This means that for an arbitrary (nonlinear) 1D oscillator, the angle variable Θ is a convenient generalization of the full phase Ψ . Due to this reason, the variables J and Θ present a convenient tool for discussion of certain fine points of the dynamics of strongly nonlinear oscillators – for whose discussion I, unfortunately, do not have time/space.¹⁵

10.3. The Hamilton principle

Now let me show that the Lagrange equations of motion, which were derived in Sec. 2.1 from the Newton laws, may be also obtained from the so-called *Hamilton principle*,¹⁶ namely the condition of a minimum (or rather an extremum) of the following integral called *action*:

Action

$$S \equiv \int_{t_{\text{ini}}}^{t_{\text{fin}}} L dt , \quad (10.47)$$

where t_{ini} and t_{fin} are, respectively, the initial and final moments of time, at which all generalized coordinates and velocities are considered fixed (not varied) – see Fig. 2.

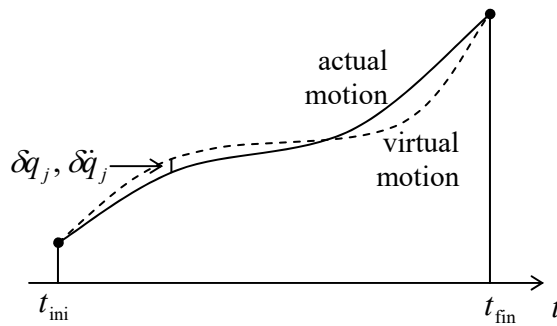


Fig. 10.2. Deriving the Hamilton principle.

The proof of that statement is rather simple. Considering, similarly to Sec. 2.1, a possible virtual variation of the motion, described by infinitesimal deviations $\{ \delta q_j(t), \delta \dot{q}_j(t) \}$ from the real motion, the necessary condition for S to be minimal is

Hamilton principle

$$\delta S \equiv \int_{t_{\text{ini}}}^{t_{\text{fin}}} \delta L dt = 0 , \quad (10.48)$$

where δS and δL are the variations of the action and the Lagrange function, corresponding to the set $\{ \delta q_j(t), \delta \dot{q}_j(t) \}$. As has been already discussed in Sec. 2.1, we can use the operation of variation just

¹⁵ An interested reader may be referred, for example, to Chapter 6 in J. Jose and E. Saletan, *Classical Dynamics*, Cambridge U. Press, 1998.

¹⁶ It is also called the “principle of least action”. (This name may be fairer in the context of a long history of the development of the principle, starting from its simpler particular forms, which includes the names of P. de Fermat, P. Maupertuis, L. Euler, and J.-L. Lagrange.)

as the usual differentiation (but at a fixed time, see Fig. 2), swapping these two operations if needed – see Fig. 2.3 and its discussion. Thus, we may write

$$\delta L = \sum_j \left(\frac{\partial L}{\partial q_j} \delta q_j + \frac{\partial L}{\partial \dot{q}_j} \delta \dot{q}_j \right) = \sum_j \frac{\partial L}{\partial q_j} \delta q_j + \sum_j \frac{\partial L}{\partial \dot{q}_j} \frac{d}{dt} \delta q_j. \quad (10.49)$$

After plugging the last expression into Eq. (48), we can integrate the second term by parts:

$$\begin{aligned} \delta S &= \int_{t_{\text{ini}}}^{t_{\text{fin}}} \sum_j \frac{\partial L}{\partial q_j} \delta q_j dt + \sum_j \int_{t_{\text{ini}}}^{t_{\text{fin}}} \frac{\partial L}{\partial \dot{q}_j} \frac{d}{dt} \delta q_j dt \\ &= \int_{t_{\text{ini}}}^{t_{\text{fin}}} \sum_j \frac{\partial L}{\partial q_j} \delta q_j dt + \sum_j \left[\frac{\partial L}{\partial \dot{q}_j} \delta q_j \right]_{t_{\text{ini}}}^{t_{\text{fin}}} - \sum_j \int_{t_{\text{ini}}}^{t_{\text{fin}}} \delta q_j d \left(\frac{\partial L}{\partial \dot{q}_j} \right) = 0. \end{aligned} \quad (10.50)$$

Since the generalized coordinates in the initial and final points are considered fixed (not affected by the variation), all $\delta q_j(t_{\text{ini}})$ and $\delta q_j(t_{\text{fin}})$ vanish, so the second term in the last form of Eq. (50) vanishes as well. Now multiplying and dividing the last term of that expression by dt , we finally get

$$\delta S = \int_{t_{\text{ini}}}^{t_{\text{fin}}} \sum_j \frac{\partial L}{\partial q_j} \delta q_j dt - \sum_j \int_{t_{\text{ini}}}^{t_{\text{fin}}} \delta q_j \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) dt = - \int_{t_{\text{ini}}}^{t_{\text{fin}}} \sum_j \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} \right] \delta q_j dt = 0. \quad (10.51)$$

This relation should hold for an arbitrary set of functions $\delta q_j(t)$, and for any time interval, and this is only possible if the expressions in the square brackets equal zero for all j , giving us the set of the Lagrange equations (2.19). So, the Hamilton principle indeed gives the Lagrange equations of motion.

It is fascinating to see how the Hamilton principle works for particular cases. As a very simple example, let us consider the usual 1D linear oscillator, with the Lagrangian function used so many times before in this course:

$$L = \frac{m}{2} \dot{q}^2 - \frac{m\omega_0^2}{2} q^2. \quad (10.52)$$

As we know very well, the Lagrange equations of motion for this L are exactly satisfied by any sinusoidal function with the frequency ω_0 , in particular by a symmetric function of time

$$q_e(t) = A \cos \omega_0 t, \quad \text{so that } \dot{q}_e(t) = -A \omega_0 \sin \omega_0 t. \quad (10.53)$$

On a limited time interval, say $0 \leq \omega_0 t \leq \pi/2$, this function is rather smooth and may be well approximated by another simple, reasonably selected functions of time, for example

$$q_a(t) = A(1 - \lambda t^2), \quad \text{so that } \dot{q}_a(t) = -2A\lambda t, \quad (10.54)$$

provided that the parameter λ is also selected reasonably. Let us take $\lambda = (\pi/2\omega_0)^2$, so the approximate function $q_a(t)$ coincides with the exact function $q_e(t)$ at both ends of our time interval (Fig.3):

$$q_a(t_{\text{ini}}) = q_e(t_{\text{ini}}) = A, \quad q_a(t_{\text{fin}}) = q_e(t_{\text{fin}}) = 0, \quad \text{where } t_{\text{ini}} \equiv 0, \quad t_{\text{fin}} \equiv \frac{\pi}{2\omega_0}, \quad (10.55)$$

and check which of them the Hamilton principle “prefers”, i.e. which function gives the least action.

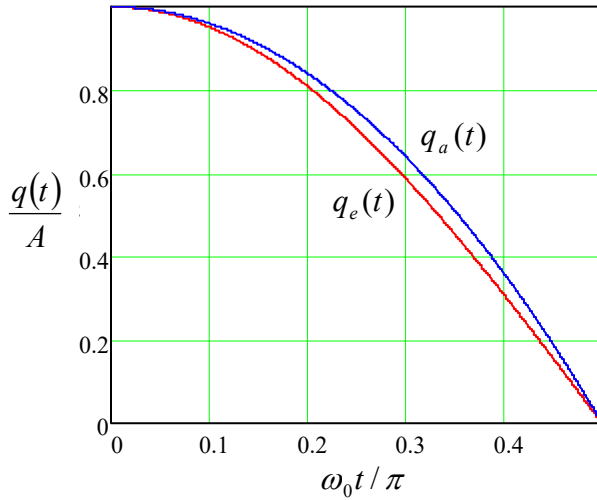


Fig. 10.3. Plots of the functions $q(t)$ given by Eqs. (53) and (54).

An elementary calculation of the action (47) corresponding to these two functions, yields

$$S_e = \left(\frac{\pi}{8} - \frac{\pi}{8} \right) m \omega_0 A^2 = 0, \quad S_a = \left(\frac{4}{3\pi} - \frac{2\pi}{15} \right) m \omega_0 A^2 \approx (0.4244 - 0.4189) m \omega_0 A^2 > 0, \quad (10.56)$$

with the first terms in all the parentheses coming from the time integrals of the kinetic energy, and the second terms, from those of the potential energy.

This result shows, first, that the exact function of time, for which these two contributions *exactly* cancel,¹⁷ is indeed “preferable” for minimizing the action. Second, for the approximate function, the two contributions to the action are rather close to the exact ones, and hence *almost* cancel each other, signaling that this approximation is very reasonable. It is evident that in some cases when the exact analytical solution of the equations of motion cannot be found, the minimization of S by adjusting one or more free parameters, incorporated into a guessed “trial” function, may be used to find a reasonable approximation for the actual law of motion.¹⁸

It is also very useful to make the notion of action S , defined by Eq. (47), more transparent by calculating it for the simple case of a single particle moving in a potential field that conserves its energy $E = T + U$. In this case, the Lagrangian function $L = T - U$ may be represented as

$$L = T - U = 2T - (T + U) = 2T - E = mv^2 - E, \quad (10.57)$$

with a time-independent E , so

$$S = \int L dt = \int mv^2 dt - Et + \text{const.} \quad (10.58)$$

Recasting the expression under the remaining integral as $m\mathbf{v} \cdot \mathbf{v} dt = \mathbf{p} \cdot (d\mathbf{r}/dt) dt = \mathbf{p} \cdot d\mathbf{r}$, we finally get

$$S = \int \mathbf{p} \cdot d\mathbf{r} - Et + \text{const} = S_0 - Et + \text{const}, \quad (10.59)$$

¹⁷ Such cancellation, i.e. the equality $S = 0$, is of course not the general requirement; it is specific only for this particular example, with a specific choice of the arbitrary constant in the potential energy of the system.

¹⁸ This is essentially a classical analog of the variational method of quantum mechanics – see, e.g., QM Sec. 2.9.

where the time-independent integral

$$S_0 \equiv \int \mathbf{p} \cdot d\mathbf{r} \quad (10.60)$$

is frequently called the *abbreviated action*.¹⁹

This expression may be used to establish one more important connection between classical and quantum mechanics – now in its Schrödinger picture. Indeed, in the quasiclassical (WKB) approximation of that picture²⁰ a particle of fixed energy E is described by a de Broglie wave

$$\Psi(\mathbf{r}, t) \propto \exp\left\{i\left(\int \mathbf{k} \cdot d\mathbf{r} - \omega t + \text{const}\right)\right\}, \quad (10.61)$$

where the wave vector \mathbf{k} is proportional to the particle's momentum (which is possibly a slow function of \mathbf{r}) and the frequency ω , to its energy:

$$\mathbf{k} = \frac{\mathbf{p}}{\hbar}, \quad \omega = \frac{E}{\hbar}. \quad (10.62)$$

Plugging these expressions into Eq. (61) and comparing the result with Eq. (59), we see that the WKB wavefunction may be represented as

$$\Psi \propto \exp\{iS/\hbar\}. \quad (10.63)$$

Hence the Hamilton principle (48) means that the total phase of the quasiclassical wavefunction should be minimal along the particle's real trajectory. But this is exactly the so-called *eikonal minimum principle* well known from the optics (though it is valid for any other waves as well), where it serves to define the ray paths in the geometric optics limit – similar to the WKB approximation. Thus, the ratio S/\hbar may be considered just as the eikonal, i.e. the total phase accumulation, of the de Broglie waves.²¹

Now, comparing Eq. (60) with Eq. (39), we see that the action variable J is just the change of the abbreviated action S_0 along a single phase-plane contour, divided by 2π . This means, in particular, that in the WKB approximation, J is the number of de Broglie waves along the classical trajectory of a particle, i.e. an integer value of the corresponding quantum number. If the system's parameters are changed slowly, the quantum number has to stay integer, and hence J cannot change, giving a quantum-mechanical interpretation of the adiabatic invariance. The reader should agree that this is really fascinating: a fact of classical mechanics may be “derived” (or at least understood) more easily from the quantum mechanics' standpoint. (As a reminder, we have run into a similarly pleasant surprise at our discussion of the non-degenerate parametric excitation in Sec. 6.7.)

¹⁹ Comparing Eq. (59) with the Hamilton principle (48), we see that if the variational trajectories are limited to those of only one (actual) energy E , the real motion corresponds to the minimum of not only S but S_0 as well. This fact is called the *Maupertuis principle*. (Historically, this result rather than Eq. (48), was called the “principle of least action”, and some authors still use this terminology, so the reader's caution is advised.)

²⁰ See, e.g., QM Sec. 3.1.

²¹ Indeed, Eq. (63) was the starting point for R. Feynman's development of his path-integral formulation of quantum mechanics – see, e.g., QM Sec. 5.3.

10.4. The Hamilton-Jacobi equation

The action S , defined by Eq. (47), may be used for one more analytical formulation of classical mechanics. For that, we need to make one more, different commitment: S has to be considered as a function of the following independent arguments: the final time point t_{fin} (which I will, for brevity, denote as t in this section), and the set of generalized coordinates (but not of the generalized velocities!) at that point:

$$S \equiv \int_{t_{\text{ini}}}^t L dt = S[t, q_j(t)]. \quad (10.64)$$

Hamilton-
Jacobi
action

Let us calculate the variation of this (from the variational point of view, new!) function, resulting from an arbitrary combination of variations of the final values $q_j(t)$ of the coordinates while keeping t fixed. Formally this may be done by repeating the variational calculations described by Eqs. (49)-(51), besides that now the variations δq_j at the finite point (t) do not necessarily equal zero. As a result, we get

$$\delta S = \sum_j \frac{\partial L}{\partial \dot{q}_j} \delta q_j \Big|_t - \int_{t_{\text{ini}}}^t dt \sum_j \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} \right] \delta q_j. \quad (10.65)$$

For the motion along the real trajectory, i.e. satisfying the Lagrange equations (2.19), the second term of this expression equals zero. Hence Eq. (65) shows that, for (any) fixed time t ,

$$\frac{\partial S}{\partial q_j} = \frac{\partial L}{\partial \dot{q}_j}. \quad (10.66)$$

But the last derivative is nothing else than the generalized momentum p_j , so

$$\frac{\partial S}{\partial q_j} = p_j. \quad (10.67)$$

(As a reminder, both parts of this relation refer to the final moment t of the trajectory.) As a result, the full derivative of the action $S[t, q_j(t)]$ over time takes the form

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \sum_j \frac{\partial S}{\partial q_j} \dot{q}_j = \frac{\partial S}{\partial t} + \sum_j p_j \dot{q}_j. \quad (10.68)$$

Now, by the definition of S , the full derivative dS/dt is nothing more than the Lagrangian function L , so Eq. (67) yields

$$\frac{\partial S}{\partial t} = L - \sum_j p_j \dot{q}_j. \quad (10.69)$$

However, according to the definition (2) of the Hamiltonian function H , the right-hand side of Eq. (69) is just $(-H)$, and we get an extremely simply-looking *Hamilton-Jacobi equation*

$$\frac{\partial S}{\partial t} = -H. \quad (10.70)$$

Hamilton-
Jacobi
equation

This simplicity is, however, rather deceiving, because to use this equation for the calculation of the function $S(t, q_j)$ for any particular problem, the Hamiltonian function has to be first expressed as a function of time t , generalized coordinates q_j , and the generalized momenta p_j (which may be, according

to Eq. (67), represented just as the derivatives $\partial S/\partial q_j$). Let us see how this procedure works for the simplest case of a 1D system with the Hamiltonian function given by Eq. (10). In this case, the only generalized momentum is $p = \partial S/\partial q$, so

$$H = \frac{p^2}{2m_{\text{ef}}} + U_{\text{ef}}(q, t) = \frac{1}{2m_{\text{ef}}} \left(\frac{\partial S}{\partial q} \right)^2 + U_{\text{ef}}(q, t), \quad (10.71)$$

and Eq. (70) is reduced to the following partial differential equation,

$$\frac{\partial S}{\partial t} + \frac{1}{2m_{\text{ef}}} \left(\frac{\partial S}{\partial q} \right)^2 + U_{\text{ef}}(q, t) = 0. \quad (10.72)$$

Its solution may be readily found in the easiest case of time-independent potential energy $U_{\text{ef}} = U_{\text{ef}}(q)$. In this case, Eq. (72) is evidently satisfied by the following variable-separated solution:

$$S(t, q) = S_0(q) + \text{const} \times t. \quad (10.73)$$

Plugging this solution into Eq. (72), we see that since the sum of the two last terms on the left-hand side of that equation represents the full mechanical energy E , the constant in Eq. (73) is nothing but $(-E)$. Thus for the function $S_0(q)$ we get an ordinary differential equation

$$-E + \frac{1}{2m_{\text{ef}}} \left(\frac{dS_0}{dq} \right)^2 + U_{\text{ef}}(q) = 0. \quad (10.74)$$

Integrating it, we get

$$S_0 = \int \{2m_{\text{ef}} [E - U_{\text{ef}}(q)]\}^{1/2} dq + \text{const}, \quad (10.75)$$

so, finally, the action is equal to

$$S = \int \{2m_{\text{ef}} [E - U_{\text{ef}}(q)]\}^{1/2} dq - Et + \text{const}. \quad (10.76)$$

For the case of 1D motion of a single 1D particle, i.e. for $q = x$, $m_{\text{ef}} = m$, $U_{\text{ef}}(q) = U(x)$, this solution is just the 1D case of the more general Eqs. (59)-(60), which were obtained above in a much more simple way. (In particular, S_0 is just the abbreviated action.)

This particular example illustrates that the Hamilton-Jacobi equation is not the most efficient way for the solution of most practical problems of classical mechanics. However, it may be rather useful for studies of certain mathematical aspects of dynamics.²² Moreover, in the early 1950s this approach was extended to a completely different field – the optimal control theory, in which the role of the action S is played by the so-called *cost function* – a certain functional of a system (understood in a very general sense of this term), that should be minimized by an optimal choice of a *control signal* – a function of time that affects the system's evolution in time. From the point of view of this theory, Eq. (70) is a particular case of a more general *Hamilton-Jacobi-Bellman* equation.²³

²² See, e.g., Chapters 6-9 in I. C. Percival and D. Richards, *Introduction to Dynamics*, Cambridge U. Press, 1983.

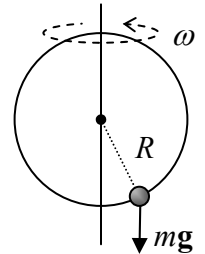
²³ See, e.g., T. Bertsekas, *Dynamic Programming and Optimal Control*, vols. 1 and 2, Aetna Scientific, 2005 and 2007. The reader should not be intimidated by the very unnatural term “dynamic programming”, which was invented by the founding father of this field, Richard Bellman, to lure government bureaucrats into funding his research, deemed too theoretical at that time. (Presently, it has a broad range of important applications.)

10.5. Exercise problems

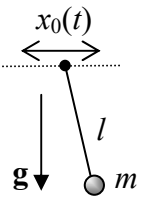
In each of Problems 1-3, for the given system:

- (i) derive the Hamilton equations of motion, and
 (ii) check whether these equations are equivalent to those derived from the Lagrangian formalism.

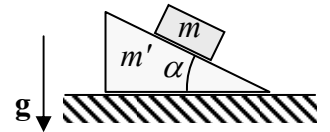
10.1. Our “testbed” system: a bead on a ring rotated, with a fixed angular velocity ω , about its vertical diameter – see Fig. 2.1, partly reproduced on the right.



10.2. The system considered in Problem 2.3: a pendulum hanging from a point whose motion $x_0(t)$ in the horizontal direction is fixed – see the figure on the right. (No vertical-plane constraint.)



10.3. The system considered in Problem 2.8: a block of mass m that can slide, without friction, along the inclined surface of a heavy wedge of mass m' . The wedge is free to move, also without friction, along a horizontal surface – see the figure on the right. (Both motions are within the vertical plane containing the steepest slope line.)



10.4. Derive and solve the equations of motion of a particle with the following Hamiltonian function:

$$H = \frac{1}{2m} (\mathbf{p} + a\mathbf{r})^2,$$

where a is a constant scalar.

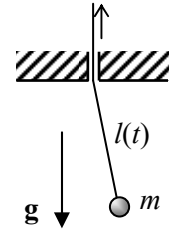
10.5. Let L be the Lagrangian function, and H the Hamiltonian function, of the same system. What three of the following four statements,

$$(i) \frac{dL}{dt} = 0, \quad (ii) \frac{\partial L}{\partial t} = 0, \quad (iii) \frac{dH}{dt} = 0, \quad (iv) \frac{\partial H}{\partial t} = 0,$$

are equivalent? Give an example of when those three equalities hold, but the fourth one does not.

10.6. Calculate the Poisson brackets of a Cartesian component of the angular momentum \mathbf{L} of a particle moving in a central force field and its Hamiltonian function H , and discuss the most evident implication of the result.

10.7. After small oscillations had been initiated in the point pendulum shown in Fig. on the right, the supporting string is being pulled up slowly, so that the pendulum's length l is being reduced. Neglecting dissipation,



(i) prove by a direct calculation that the oscillation energy is indeed changing proportionately to the oscillation frequency, as it follows from the constancy of the corresponding adiabatic invariant (40); and

(ii) find the l -dependence of the amplitudes of the angular and linear deviations from the equilibrium.

10.8. The mass m of a small body that performs 1D oscillations in the potential well $U(x) = ax^{2n}$, with $n > 0$, is being changed slowly, without exerting any additional direct force. Calculate the oscillation energy E as a function of m .

10.9. A stiff ball is bouncing vertically from the floor of an elevator whose upward acceleration changes very slowly. Neglecting the energy dissipation, calculate how much the bounce height h changes during the acceleration's increase from 0 to g . Is your result valid for an equal but abrupt increase of the elevator's acceleration?

10.10.* A 1D particle of a constant mass m moves in a time-dependent potential $U(q, t) = m\omega^2(t)q^2/2$, where $\omega(t)$ is a slow function of time, with $|\dot{\omega}| \ll \omega^2$. Develop the approximate method for the solution of the corresponding equation of motion, similar to the WKB approximation used in quantum mechanics.²⁴ Use the approximation to confirm the conservation of the action variable (40) for this system.

Hint: You may like to look for the solution to the equation of motion in the form

$$q(t) = \exp\{\Lambda(t) + i\Psi(t)\},$$

where Λ and Ψ are some real functions of time, and then make proper approximations in the resulting equations for these functions.

²⁴ See, e.g., QM Sec. 2.4.