Chapter 10. A Bit More of Analytical Mechanics

This concluding chapter reviews two alternative approaches to analytical mechanics, whose main advantage is a closer parallel to quantum mechanics in general and to its quasiclassical (WKB) approximation in particular. One of them, the Hamiltonian formalism, is also used to derive 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 useful the analytical mechanics, in its Lagrangian form, may be invaluable for solving various particular problems of classical mechanics. Now let us discuss several alternative formulations[1] that may not be much more useful for this purpose, but shed 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 equations (2.19)

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

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

Hamiltonian function

$$H \equiv \sum_j p_j \dot{q}_j - L\,. \tag{10.2}$$

Now let us rewrite the full differential of this function[2] in the following form:

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

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

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

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

[1] Due mostly to W. Hamilton (1805-1865) and C. Jacobi (1804-1851).
[2] Actually, this differential has already been used in Sec. 2.3 to derive Eq. (2.35).

and the generalized momenta p_j (rather than generalized velocities). With this commitment, the general rule of differentiation of a function of several arguments gives

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

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 the differentials 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:

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{10.6}$$

$$\boxed{\begin{aligned} \dot{q}_j &= \frac{\partial H}{\partial p_j}, \\ \dot{p}_j &= -\frac{\partial H}{\partial q_j}. \end{aligned}} \tag{10.7}$$

Hamilton equations

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

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}, \tag{10.8}$$

meaning that function $H(t, q_j, p_j)$ can change in time only via its explicit dependence on t. 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 simple Lagrangian function (3.3):

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

In this case, $p \equiv \partial L/\partial \dot{q} = m_{\text{ef}}\dot{q}$, and $H \equiv p\dot{q} - L = m_{\text{ef}}\dot{q}^2/2 + U_{\text{ef}}(q,t)$. In order to honor our new commitment, we need to express the Hamiltonian function explicitly via t, q and p (rather than $\dot{q}$):

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

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

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m_{\text{ef}}}, \tag{10.11}$$

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

$$\dot{p} = -\frac{\partial H}{\partial q} = -\frac{\partial U_{\text{ef}}}{\partial q}. \tag{10.12}$$

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

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

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

Thus, the Hamiltonian formalism does not give much new for the solution of most problems of classical mechanics. (This is why I have postponed its discussion until the very end of this course.) Moreover, since the Hamiltonian function $H(t, q_j, p_j)$ does nor 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 aesthetically appealing and heuristically fruitful. This is especially true in the cases where these arguments participate in H in a similar way. For example, for the very important case of a dissipation-free harmonic oscillator, for which $U_{\text{ef}} = \kappa_{\text{ef}}q^2/2$, Eq. (10) gives the famous symmetric form

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

The Hamilton equations (7) for this system preserve the symmetry, especially evident if we introduce the normalized momentum $\not{p} \equiv p/m_{\text{ef}}\omega_0$ (already used in Secs. 4.3 and 9.2):

$$\frac{dq}{dt} = \omega_0 \not{p}, \quad \frac{d\not{p}}{dt} = -\omega_0 q. \tag{10.15}$$

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

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

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

Dynamics of arbitrary variable

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

where the last term in the right-hand part is the so-called *Poisson bracket*[4] that is defined, for two arbitrary functions $f(t, q_j, p_j)$ and $g(t, q_j, p_j)$, as

Poisson bracket

$$\boxed{\{g, f\} \equiv \sum_j \left(\frac{\partial g}{\partial p_j}\frac{\partial f}{\partial q_j} - \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j} \right).} \tag{10.18}$$

[4] Named after S. Poisson - of the Poisson equation and the Poisson statistical distribution fame.

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*:

$$\{f,\{g,h\}\}+\{g,\{h,f\}\}+\{h,\{f,g\}\}=0. \tag{10.19}$$

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

$$\{H, f\}=0, \tag{10.20}$$

then $df/dt = 0$, i.e. function f is an integral of motion.

Moreover, if we already know two integrals of motion, say f and g, then function

$$F \equiv \{f, g\} \tag{10.21}$$

is also an integral of motion – the so-called *Poisson theorem*. In order to prove it, we may use the Jacobi identity (19) with $h = H$. Now 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 functions f, g, and F, we get

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

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

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

Plugging Eq. (21) into the first term of the right-hand part 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 connection between the classical and quantum mechanics. Indeed, using Eq. (18) to calculate the Poisson brackets of the generalized coordinates and momenta, we readily get

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

In quantum mechanics,[5] operators of these quantities ("observables") obey commutation relations

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

where the definition of the commutator, $[\hat{g}, \hat{f}] \equiv \hat{g}\,\hat{f} - \hat{f}\,\hat{g}$, is *to a certain extent* [6] similar to that (18) of the Poisson bracket. We see that the classical relations (24) are similar to quantum-mechanical relations (25) if we following parallel has been made:

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

CM ↔ QM relation

[5] See, e.g., QM Sec. 2.1.

[6] There is of course a conceptual difference between the "usual" products of function derivatives participating in the Poisson brackets, and the operator "products" (meaning their sequential action on a state vector – see, e.g., QM Sec. 4.1) forming the commutator.

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

$$\frac{d\hat{f}}{dt} = \frac{\partial \hat{f}}{\partial t} + \frac{i}{\hbar}\left[\hat{H}, \hat{f}\right], \qquad \text{i.e. } i\hbar\frac{d\hat{f}}{dt} = i\hbar\frac{\partial \hat{f}}{\partial t} + \left[\hat{f}, \hat{H}\right], \tag{10.27}$$

which is the correct equation of operator evolution in the Heisenberg picture of quantum mechanics.[7]

This analogy implies, in particular, that the quantum-mechanical operators (and the matrices used for their representation in a particular basis) should satisfy the same identities including Eq. (17).

10.2. Adiabatic invariance

One more application of the Hamiltonian formalism in classical mechanics is the solution of the following problem.[8] Earlier in the course, we already studied some effects of time variation of parameters of a single oscillator (Sec. 4.5) and coupled oscillators (Sec. 5.5). However, those discussions were focused on the case when the parameter variation frequency is comparable with the initial 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*[9]),

$$\left|\frac{\dot{\lambda}}{\lambda}\right| << \frac{1}{\mathcal{T}}, \tag{10.28}$$

where $\mathcal{T}$ is a typical time period of oscillations in the system. Let us consider a 1D system whose Hamiltonian $H(q, p, \lambda)$ depends on time only via the slow (28) evolution of parameter $\lambda = \lambda(t)$, and whose initial energy restricts system's motion to a finite coordinate interval – see Fig. 3.2c.

Then, as we know from Sec. 3.3, if parameter λ is constant, the system performs a periodic (though not necessarily sinusoidal) motion back and forth axis q, or, in a different language, along a closed trajectory on the phase plane $[q, p]$ – see Fig. 1.[10] According to Eq. (8), in this case H is constant on the trajectory. (In order to distinguish this particular *value* from the Hamiltonian *function* as such, I will assume that this constant coincides with the full mechanical energy E, like is does for 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:

$$\mathcal{T} \equiv \int_0^{\mathcal{T}} dt = \oint \frac{dt}{dq} dq = \oint \frac{1}{\dot{q}} dq\,. \tag{10.29}$$

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

[7] See, e.g., QM Sec. 4.6.

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

[9] This term has come from thermodynamics and statistical mechanics, where it implies not only a slow parameter variation, but also the thermal insulation of the system - see, e.g., SM Sec. 1.3. Evidently, the latter condition is irrelevant in our current context.

[10] In Sec. 4.6, we discussed this plane for the particular case of sinusoidal oscillations – see Fig. 9

$$\mathcal{T} = \oint \frac{1}{\partial H / \partial p} dq \,. \tag{10.30}$$

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

$$\mathcal{T} = \oint \frac{\partial p}{\partial E} dq \,. \tag{10.31}$$

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

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

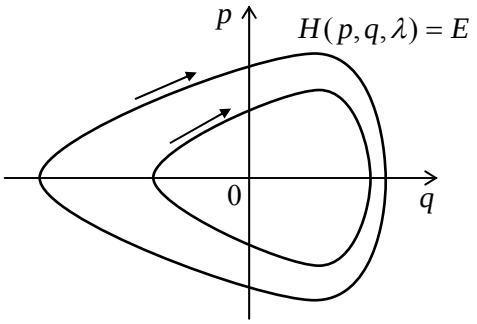


Fig. 10.1. Phase-plane representation of periodic oscillations of a 1D Hamiltonian system, for two values of energy (schematically).

Superficially, it looks that these formulas may be also used to find the motion period change when parameter λ is being changed adiabatically, for example, by plugging known functions $m_{\text{ef}}(\lambda)$ and $U_{\text{ef}}(q, \lambda)$ into Eq. (32). However, there is no guarantee that energy E in that integral would stay constant as the parameter change, 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 $\mathcal{T}$ 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 the oscillation amplitude) on λ is not immediately obvious.

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

$$\frac{dE}{dt} = \frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda} \frac{d\lambda}{dt} \,. \tag{10.33}$$

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.[11] The averaging yields

[11] This is the most critical point of this proof, because at any finite rate of parameter change the oscillations are, strictly speaking, non-periodic. Because of the approximate nature of this conjecture (which is very close to the assumptions made at the derivation of the RWA equations in Sec. 4.3), new, more strict (but also much more

$$\overline{\frac{dE}{dt}} \approx \frac{d\lambda}{dt}\overline{\frac{\partial H}{\partial \lambda}} = \frac{d\lambda}{dt}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \frac{\partial H}{\partial \lambda}\,dt\,. \tag{10.34}$$

Transforming the time integral to the contour one, just as we did at the transition from Eq. (29) to Eq. (30), and using Eq. (31) for $\mathcal{T}$, we get

$$\overline{\frac{dE}{dt}} = \frac{d\lambda}{dt}\frac{\oint \frac{\partial H/\partial \lambda}{\partial H/\partial p}\,dq}{\oint \frac{\partial p}{\partial E}\,dq}\,. \tag{10.35}$$

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 that if E is fixed, the partial differentiation of relation $E = H$ over λ yields

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

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

$$\overline{\frac{dE}{dt}} = -\frac{d\lambda}{dt}\frac{\oint \frac{\partial p}{\partial \lambda}\,dq}{\oint \frac{\partial p}{\partial E}\,dq}\,. \tag{10.37}$$

Since the left-hand part 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 that relation as

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

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

Action variable

$$J \equiv \frac{1}{2\pi}\oint p\,dq\,, \tag{10.39}$$

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 4, for such oscillator, $q = A\cos\Psi$, $p = -m_{\text{ef}}\omega_0 A\sin\Psi$ (with $\Psi = \omega_0 t$ + const), so that J may be easily expressed either via oscillations' amplitude A, or their energy $E = H = m_{\text{ef}}\omega_0^2 A^2/2$:

$$J = \frac{1}{2\pi}\oint p\,dq = \frac{1}{2\pi}\int_{\Psi=0}^{\Psi=2\pi}\left(-m_{\text{ef}}\omega_0 A\sin\Psi\right)d\left(A\cos\Psi\right) = \frac{1}{2\pi}\frac{m_{\text{ef}}\omega_0}{2}A^2 = \frac{E}{\omega_0}. \tag{10.40}$$

Returning to the general oscillator 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:

---

cumbersome) proofs of 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).

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

Within the accuracy of our approximation, in which the contour integrals (38) and (41) are calculated along a closed trajectory, factor dE/dt is indistinguishable from its time average, and these integrals coincide, so that 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:

$$\boxed{J = \text{const}}\,. \tag{10.42}$$

Adiabatic invariance

This is the famous *adiabatic invariance*.[12] In particular, according to Eq. (40), in a harmonic oscillator, energy of oscillation changes proportionately to the (slowly changed) eigenfrequency.

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,[13] and that pair $\{J, \Theta\}$ should satisfy the Hamilton equations (7), in particular,

$$\frac{d\Theta}{dt} = \frac{\partial H}{\partial J}. \tag{10.43}$$

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

$$\Theta = \frac{\partial H}{\partial J}t + \text{const}\,. \tag{10.44}$$

For a harmonic oscillator, according to Eq. (40), derivative $\partial H/\partial J = \partial E/\partial J = \omega_0 = 2\pi/\mathcal{T}$, so that $\Theta = \omega_0 t + \text{const}$. It may be shown that a more general form of this relation,

$$\frac{\partial H}{\partial J} = \frac{2\pi}{\mathcal{T}}, \tag{10.45}$$

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

$$\Theta = 2\pi\frac{t}{\mathcal{T}} + \text{const}\,. \tag{10.46}$$

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

[13] 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, infinite number of) values of 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 $\{J, \Theta\}$.

To summarize, for a harmonic oscillator, the angle variable Θ is just the full phase Ψ that we used so much in Ch. 4, while for an arbitrary (nonlinear) 1D oscillator, this is a convenient generalization of that notion. Due to this reason, variables J and Θ present a convenient tool for discussion of certain fine points of dynamics strongly nonlinear oscillators – for whose discussion I, unfortunately, do not have time.[14]

10.3. The Hamilton principle

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

Action

$$S \equiv \int_{t_{ini}}^{t_{fin}} L dt , \tag{10.47}$$

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

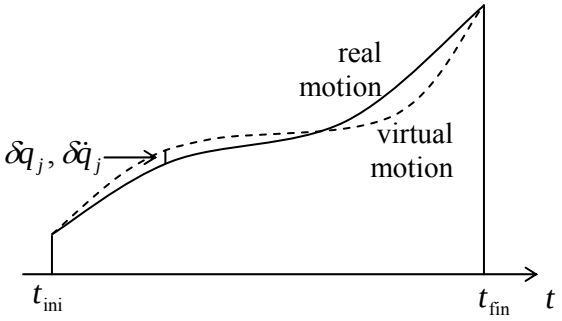


Fig. 10.2. Deriving the Hamilton principle.

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

$$\delta S \equiv \int_{t_{ini}}^{t_{fin}} \delta L \, dt = 0 , \tag{10.48}$$

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 as the usual differentiation (but at fixed time, see Fig. 2.1), swapping these two operations if needed – see Fig. 2.3 and its discussion. Thus, we may write

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

[14] See, e.g., Chapter 6 in J. Jose and E. Saletan, *Classical Dynamics*, Cambridge U. Press, 1998.

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

$$\delta S = \int_{t_{\rm ini}}^{t_{\rm fin}} \sum_j \frac{\partial L}{\partial q_j}\delta q_j dt + \sum_j \int_{t_{\rm ini}}^{t_{\rm fin}} \frac{\partial L}{\partial \dot{q}_j}\frac{d}{dt}\delta q_j dt$$

$$= \int_{t_{\rm ini}}^{t_{\rm fin}} \sum_j \frac{\partial L}{\partial q_j}\delta q_j dt + \sum_j \left[\frac{\partial L}{\partial \dot{q}_j}\delta q_j\right]_{t_{\rm ini}}^{t_{\rm fin}} - \sum_j \int_{t_{\rm ini}}^{t_{\rm fin}} \delta q_j d\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0. \qquad (10.50)$$

Since the generalized coordinates in the initial and final points are considered fixed (not affected by the variation), all $\delta q_j(t_{\rm ini}) = \delta q_j(t_{\rm fin}) = 0$, the second term in the right-hand part of Eq. (50) vanishes. Multiplying and dividing the last term of that part by dt, we finally get

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

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

It is 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 presented as

$$L = T - U = 2T - (T + U) = 2T - E = mv^2 - E, \qquad (10.52)$$

with $E = \text{const}$, so that

$$S = \int L dt = \int mv^2 dt - Et + \text{const}. \qquad (10.53)$$

Presenting 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

$$S = \int \mathbf{p}\cdot d\mathbf{r} - Et + \text{const} = S_0 - Et + \text{const}, \qquad (10.54)$$

where the time-independent integral

$$S_0 \equiv \int \mathbf{p}\cdot d\mathbf{r} \qquad (10.55)$$

is frequently called the *abbreviated action*.[15]

This expression may be used to establish one more connection between the classical and quantum mechanics, now in its Schrödinger picture. Indeed, in the quasiclassical (WKB) approximation of that picture[16] a particle of fixed energy is described by a De Broglie wave

$$\Psi(\mathbf{r},t) \propto \exp\left\{i\left(\int \mathbf{k}\cdot d\mathbf{r} - \omega t + \text{const}\right)\right\}, \qquad (10.56)$$

[15] Please note that despite a close relation between the abbreviated action S_0 and the action variable J defined by Eq. (39), these notions are not identical. Most importantly, J is an integral over a *closed* trajectory, while S_0 in defined for an arbitrary point of a trajectory.

[16] See, e.g., QM Sec. 2.3.

where wavevector **k** is proportional to the particle's momentum, while frequency ω, to its energy:

$$\mathbf{k} = \frac{\mathbf{p}}{\hbar}, \quad \omega = \frac{E}{\hbar}. \tag{10.57}$$

Plugging these expressions into Eq. (56) and comparing the result with Eq. (54), we see that the WKB wavefunction may be presented as

$$\Psi \propto \exp\{iS/\hbar\}. \tag{10.58}$$

Hence the Hamilton's principle (48) means that the total phase of the quasiclassical wavefunction should be minimal along particle's real trajectory. But this is exactly the so-called *eikonal minimum principle* well known from the optics (though 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 condition. Thus, the ratio $S/\hbar$ may be considered just as the eikonal, i.e. the total phase accumulation, of the de Broigle waves.[17]

Now, comparing Eq. (55) with Eq. (33), 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 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 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. It is really fascinating that a fact of classical mechanics may be "derived" (or at least understood) more easily from the quantum mechanics' standpoint.[18]

10.4. The Hamilton-Jacobi equation

Action S, defined by Eq. (47), may be used for one more formulation of classical mechanics. For that, we need one more, different commitment: S to be considered 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:

Hamilton-Jacobi action

$$S \equiv \int_{t_{\text{ini}}}^{t} L\,dt = S\left[t, q_j(t)\right]. \tag{10.59}$$

Let us calculate a variation of this (essentially, new!) function, resulting from an arbitrary combination of variations of final values $q_j(t)$ of the coordinates, while keeping t fixed. Formally this may be done by repeating the variation calculations described by Eqs. (49)-(52), besides that now variations δq_j at the finite point (t) do not necessarily equal zero. As a result, we get

$$\delta S = \sum_j \frac{\partial L}{\partial \dot{q}_j}\delta q_j\Big|_t - \int_{t_{\text{ini}}}^{t} dt \sum_j \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j}\right]\delta q_j\,. \tag{10.60}$$

[17] Eq. (58) was the starting point for R. Feynman's development of his path-integral formulation of quantum mechanics – see, e.g., QM Sec. 5.3.

[18] As a reminder, we have run into a similar situation at our discussion of the non-degenerate parametric excitation in Sec. 5.5.

For the motion along the real trajectory, i.e. satisfying the Lagrange equations of motion, the second term of this expression equals zero. Hence Eq. (60) shows that, for (any) fixed time t,

$$\frac{\partial S}{\partial q_j} = \frac{\partial L}{\partial \dot{q}_j}. \tag{10.61}$$

But the last derivative is nothing else than the generalized momentum p_j – see Eq. (2.31), so that

$$\frac{\partial S}{\partial q_j} = p_j. \tag{10.62}$$

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

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \sum_j \frac{\partial S}{\partial q_j}\dot{q}_j = \frac{\partial S}{\partial t} + \sum_j p_j \dot{q}_j. \tag{10.63}$$

Now, by the very definition (59), the full derivative dS/dt is nothing more that the Lagrange function L, so that Eq. (63) yields

$$\frac{\partial S}{\partial t} = L - \sum_j p_j \dot{q}_j. \tag{10.64}$$

However, according to the definition (2) of the Hamiltonian function H, the right-hand part of Eq. (63) is just $(-H)$, so that we get an extremely simply-looking *Hamilton-Jacobi equation*

$$\boxed{\frac{\partial S}{\partial t} = -H.} \tag{10.65}$$

Hamilton-Jacobi equation

This simplicity is, however, rather deceiving, because in order to use this equation for the calculation of 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. (62), presented just as derivatives $\partial S/\partial q_j$). Let us see how does this procedure work 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 that

$$H = \frac{p^2}{2m_{\text{ef}}} + U_{\text{ef}}(q,t) = \frac{1}{2m_{\text{ef}}}\left(\frac{\partial S}{\partial q}\right)^2 + U_{\text{ef}}(q,t), \tag{10.66}$$

and the Hamilton-Jacobi equation (65) is reduced to a partial differential equation,

$$\frac{\partial S}{\partial t} + \frac{1}{2m_{\text{ef}}}\left(\frac{\partial S}{\partial q}\right)^2 + U_{\text{ef}}(q,t) = 0. \tag{10.67}$$

Its solution may be readily found in the particular case of time-independent potential energy $U_{\text{ef}} = U_{\text{ef}}(q)$. In this case, Eq. (67) is evidently satisfied by a variable-separated solution

$$S(t,q) = S_0(q) + \text{const} \times t. \tag{10.68}$$

Plugging this solution into Eq. (67), we see that since the sum of two last terms in the left-hand part of that equation presents the full mechanical energy E, the constant in Eq. (68) is nothing but ($-E$). Thus for function S_0 we get an ordinary differential equation

$$-E+\frac{1}{2m_{\text{ef}}}\left(\frac{dS_0}{dq}\right)^2+U_{\text{ef}}(q)=0. \tag{10.69}$$

Integrating it, we get

$$S_0=\int\left\{2m_{\text{ef}}\left[E-U_{\text{ef}}(q)\right]\right\}^{1/2}dq+\text{const}, \tag{10.70}$$

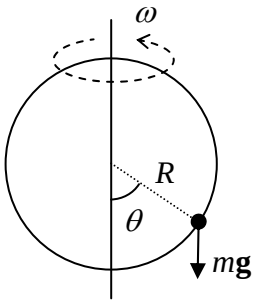
so that, finally, the action is equal to

$$S=\int\left\{2m_{\text{ef}}\left[E-U_{\text{ef}}(q)\right]\right\}^{1/2}dq-Et+\text{const}. \tag{10.71}$$

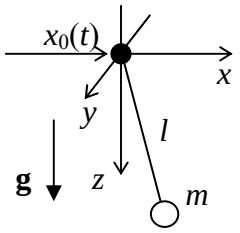
For the case of 1D motion of a single 1D particle, i.e. for $q = x$, $m_{\text{ef}} = m$, $U_{\text{ef}}(q) = U(x)$, this solution is just the 1D case of the more general Eqs. (54)-(55), which were obtained by a much more simple way. (In particular, S_0 is just the abbreviated action.)

This particular case illustrates that the Hamilton-Jacobi equation is not the most efficient way for solution of most practical problems. However, it may be rather useful for studies of certain mathematical aspects of dynamics.[19] Moreover, in the 1940s this approach was extended to a completely different field – the optimal control theory, in which the role of action S is played by the so-called *cost function* – a certain functional of a dynamic system, that should be minimized by an optical choice of a *control signal* – a function of time that affects system's dynamics. From the point of view of this mathematical theory, Eq. (65) is a particular case of a more general *Hamilton-Jacobi-Bellman* equation.[20]

10.5. Exercise problems



10.1. Derive the Hamilton equations of motion for our testbed problem (a bead on a ring rotating about its vertical diameter – see Fig. 2.1, partly reproduced on the right). Check that the equations are equivalent to those derived from the Lagrangian formalism.

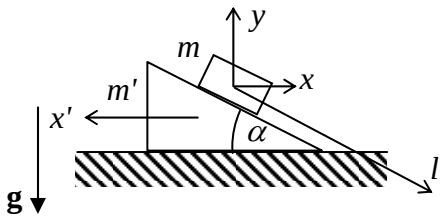


10.2. Perform the same tasks as in Problem 10.1 for the system already considered in Problem 2.3, a fixed-length pendulum hanging from a horizontal support whose motion law $x_0(t)$ is fixed – see Fig. on the right. (No vertical plane constraint.)

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

[20] See, e.g., T. P. Bertsekas, *Dynamic Programming and Optimal Control*, vols. 1 and 2, Aetna Scientific, 2005 and 2007. The reader should not be deceived by the unnatural term "dynamic programming" that was invented by the founding father of this field, R. Bellman, to lure government bureaucrats into funding his research, which had been deemed too theoretical at that time, but now has a broad range of important applications.

10.3. Perform the same tasks as in Problems 1 and 2, for the system already considered in Problem 2.5 - a block of mass m that can slide, without friction, along the inclined surface of a heavy wedge (mass m'). The wedge is free to move, also without friction, along a horizontal surface - see Fig. on the right. (Both motions are within the vertical plane containing the steepest slope line.)



10.4. Find and solve 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.

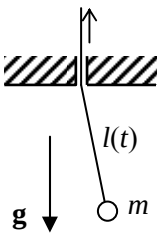
10.5. Let L be the Lagrange function, and H the Hamilton function, of the same system. What three of the following four statements,

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

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

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

10.7. After small oscillations had been initiated in a simple pendulum (Fig. on the right), the thread is being pulled up slowly, so that the pendulum 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 amplitudes of the angular and linear deviations from the equilibrium.

10.8. The mass m of a small body that performs 1D oscillations in potential $U(x) = ax^{2n}$, with $n > 0$, is being changed slowly. 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 energy dissipation, calculate how much does the bounce height h change during acceleration's increase from 0 to g.