PHYS V1100: Analytical Dynamics

Lecture 15: Hamiltonian Mechanics I

Conjugate variables, endpoint structure, and Hamilton's equations

1 Why change the description of a state?

A position is not generally enough to predict a particle's future: two particles can pass through the same point with different velocities. For a regular Lagrangian system with N independent generalized coordinates, the pair $(\boldsymbol{q}, \dot{\boldsymbol{q}})$ supplies the local state data for its second-order equations. If the law depends explicitly on time, the initial time must also be specified. This returns us to the opening question of the course: which variables make the evolution close?

The variational formulation has already distinguished another quantity,

$$\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{q}}}.$$

For a Cartesian particle with $L = \frac{1}{2} m \dot{\boldsymbol{q}}^T \dot{\boldsymbol{q}} - V$, this is just $m\dot{\boldsymbol{q}}$. It would then be easy to regard the change from velocity to momentum as a change of scale. But an angle has a conjugate angular momentum, not a linear momentum, and a configuration-dependent mass matrix relates velocities and momenta in more general coordinates. There is a structural reason to choose $\boldsymbol{p}$, not simply an algebraic convenience.

That reason was visible in the boundary term of the action variation. In Lecture 5, integration by parts produced a momentum multiplying a configuration variation at each endpoint. Fixed endpoint data made that term vanish, but did not make the momentum vanish. We now keep the boundary information and ask what it tells us. The discrete action of Lecture 4 makes the endpoint signs and the matching of neighboring pieces especially transparent.

The chapter develops one chain of ideas:

$$\text{action edge} \longrightarrow \text{endpoint momentum pairing} \longrightarrow (\boldsymbol{q}, \boldsymbol{p}) \text{ as state},$$

$$\text{change from velocity to momentum} \longrightarrow H(\boldsymbol{q}, \boldsymbol{p}, t) \longrightarrow \text{first-order equations}.$$

We will also identify the coefficient paired with an endpoint time change. The oscillator and central-force problem will then test the new description against mechanics we already know. General Poisson-bracket theory remains the next lecture's subject.

2 An action edge has two endpoint responses

2.1 Varying one node changes two edges

Use one coordinate first, so that endpoint signs are not hidden by vector notation. Put $h = \Delta t > 0$, hold the time grid fixed, and write the left-endpoint discrete action from Lecture 4 as

$$\mathcal{S}_d = \sum_{n=0}^{K-1} \mathcal{A}_n(q_n, q_{n+1}), \qquad \mathcal{A}_n = hL(q_n, v_n), \qquad v_n = \frac{q_{n+1} - q_n}{h}. \tag{2.1}$$

Here K is the number of time intervals; N will continue to count configuration degrees of freedom. The symbol v_n denotes a discrete slope, not a second independent node variable. Explicit time dependence could be included as a prescribed argument t_n; we suppress it in this first calculation.

Changing an interior q_n changes the end of the preceding edge and the start of the following edge. No other action term depends on that node. This is why a stationary sum produces a local equation even when the sum contains many terms. Instead of immediately differentiating the two slopes, let us first give names to the endpoint responses of a general smooth edge function $\mathcal{A}_n$.

2.2 The signs belong to the endpoints

At fixed endpoint times, ordinary differentiation gives

$$\delta\mathcal{A}_n = \frac{\partial\mathcal{A}_n}{\partial q_n}\delta q_n + \frac{\partial\mathcal{A}_n}{\partial q_{n+1}}\delta q_{n+1}.$$

Define

$$\boxed{p_n^- = -\frac{\partial\mathcal{A}_n}{\partial q_n}, \qquad p_{n+1}^+ = \frac{\partial\mathcal{A}_n}{\partial q_{n+1}}.} \tag{2.2}$$

Then the edge variation has the oriented form

$$\boxed{\delta\mathcal{A}_n = -p_n^-\delta q_n + p_{n+1}^+\delta q_{n+1}.} \tag{2.3}$$

The minus superscript labels the initial-end response of an edge and the plus superscript its final-end response. They do not label the sign or direction of a physical momentum. At an interior node q_n, p_n^+ comes from the edge arriving there, while p_n^- comes from the edge leaving it.

The sign choice makes addition of edges behave like the familiar endpoint difference in integration by parts. Summing (2.3) over the path yields the full identity

$$\delta\mathcal{S}_d = -p_0^-\delta q_0 + p_K^+\delta q_K + \sum_{n=1}^{K-1}(p_n^+ - p_n^-)\delta q_n. \tag{2.4}$$

If the endpoint configurations are fixed, the first two terms vanish. Each interior variation remains independent, so stationarity requires

$$\boxed{p_n^+ = p_n^- \equiv p_n \qquad (1 \le n \le K-1).} \tag{2.5}$$

A stationary path therefore carries one matched momentum at each interior node. For a trial path not satisfying the equations, the two responses generally differ. Their difference is precisely the coefficient that detects failure of stationarity at that node.

There is no assumption here that the two endpoints of one edge have equal momentum. A force can change momentum from one node to the next. Equation (2.5) instead compares two descriptions of the *same* interior endpoint.

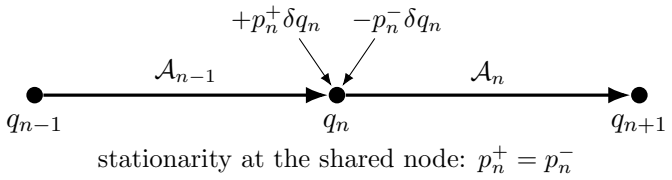

Figure 1: The same configuration variation receives two endpoint contributions. The incoming edge supplies a final-end term, and the outgoing edge an initial-end term. Their cancellation is a dynamical condition, not an identity for arbitrary trial paths.

3 Checking the endpoint momenta against Lecture 4

3.1 Sampled velocity momentum and node momentum

The endpoint definitions become more informative when we calculate them for (2.1). Write $L_n = L(q_n, v_n)$ and introduce a temporary symbol

$$\pi_n \equiv \frac{\partial L}{\partial v}(q_n, v_n).$$

This is the sampled quantity that Lecture 4 denoted by p_n. We use π_n in this comparison to avoid identifying it prematurely with the matched *node* momentum of (2.5).

There are two ways for the starting endpoint to affect the edge action: directly through the position argument of L, and through the slope. Since

$$\frac{\partial v_n}{\partial q_n} = -\frac{1}{h}, \qquad \frac{\partial v_n}{\partial q_{n+1}} = \frac{1}{h},$$

the chain rule gives

$$\frac{\partial \mathcal{A}_n}{\partial q_n} = h\frac{\partial L}{\partial q}(q_n, v_n) - \pi_n,$$
$$\frac{\partial \mathcal{A}_n}{\partial q_{n+1}} = \pi_n.$$

The derivatives of L on the right are its partial derivatives with its other argument held fixed. Therefore

$$\boxed{p_n^- = \pi_n - h\frac{\partial L}{\partial q}(q_n, v_n), \qquad p_{n+1}^+ = \pi_n.} \tag{3.1}$$

From the preceding edge, $p_n^+ = \pi_{n-1}$. Momentum matching now gives

$$\boxed{\frac{\pi_n - \pi_{n-1}}{h} = \frac{\partial L}{\partial q}(q_n, v_n),} \tag{3.2}$$

which is exactly the earlier discrete Euler–Lagrange equation.

Notice what has and has not become equal. On a stationary path,

$$p_n = \pi_{n-1} = \pi_n - h\frac{\partial L}{\partial q}(q_n, v_n).$$

It is not generally true that $p_n = \pi_n$. Different endpoint conventions attach the momentum to different parts of the discretized path. The abstract endpoint pairing is the reliable starting point; a familiar continuum formula must be matched to the actual discretization before it is used.

3.2 A mechanical edge and its update rule

For $L(q,v) = \frac{1}{2}mv^2 - V(q)$, with constant $m > 0$,

$$\mathcal{A}_n = \frac{m}{2h}(q_{n+1} - q_n)^2 - hV(q_n), \qquad \pi_n = mv_n.$$

Thus

$$p_n^- = mv_n + h\frac{dV}{dq}(q_n), \qquad p_{n+1}^+ = mv_n. \tag{3.3}$$

On the resulting matched phase-space path, these equations determine

$$\boxed{\begin{aligned} p_{n+1} &= p_n - h\frac{dV}{dq}(q_n), \\ q_{n+1} &= q_n + \frac{h}{m}p_{n+1}. \end{aligned}} \tag{3.4}$$

We first change the momentum using the force at q_n, then move using the new momentum. The second equation is just the definition of v_n, together with $p_{n+1} = mv_n$.

This does not silently replace Lecture 4's move-then-turn convention. That convention used the forward step $s_n = q_{n+1} - q_n$, or equivalently $\pi_n = ms_n/h$, as part of the state. In those variables the same configuration sequence satisfies

$$\boxed{\begin{aligned} q_{n+1} &= q_n + \frac{h}{m}\pi_n, \\ \pi_{n+1} &= \pi_n - h\frac{dV}{dq}(q_{n+1}). \end{aligned}} \tag{3.5}$$

Along the matched path, $p_{n+1} = \pi_n$. Both descriptions give

$$m\frac{q_{n+1} - 2q_n + q_{n-1}}{h^2} = -\frac{dV}{dq}(q_n).$$

The two state labels differ; the node equation does not. Keeping this attachment explicit prevents an apparent contradiction about the ordering of the update.

For $V = 0$, the two endpoint momenta of each edge agree and the particle moves with constant slope. For nonzero force, their difference is the discrete impulse in (3.4). These equations are exact statements about the chosen discrete action. For a finite h, they are generally an approximation to the continuum trajectory, not the exact finite-time flow of Newton's differential equation.

4 What makes momentum conjugate to configuration?

For many coordinates the endpoint terms take the form

$$\delta\mathcal{A} = -\boldsymbol{p}_i^T\delta\boldsymbol{q}_i + \boldsymbol{p}_f^T\delta\boldsymbol{q}_f.$$

The scalar pairing $\boldsymbol{p}^T\delta\boldsymbol{q}$ is the key. At an endpoint, it measures the first-order action response to a configuration displacement. A momentum component is naturally paired with the corresponding coordinate displacement because that is how it enters the variation of the action.

A generalized velocity describes a possible direction and rate of motion through configuration space. Momentum assigns a scalar to such a displacement. This second role is called *dual* or *covector-like*. Writing both objects as columns is our component convention; it does not erase their different roles. The transpose in $\boldsymbol{p}^T\delta\boldsymbol{q}$ displays the scalar pairing.

4.1 A rescaling exposes the difference

Take one coordinate and replace its label by $Q = aq$, where $a \neq 0$ is constant. A velocity changes as $\dot{Q} = a\dot{q}$. But the endpoint action response must be unchanged:

$$p\,\delta q = P\,\delta Q = Pa\,\delta q.$$

Consequently $P = p/a$. Momentum scales oppositely to velocity under this relabeling because it must pair with the new displacement to give the same scalar. No new dynamics has been introduced; we have only tested what kind of quantity each variable represents.

The units make the same point. Action has units of energy times time, so $p_j \delta q_j$ must have those units. If q_j is a length, p_j has units of linear momentum. If q_j is a dimensionless angle, p_j has units of angular momentum. Generalized momenta need not all have the same units.

4.2 The mass matrix supplies the relation

For a natural mechanical Lagrangian,

$$L = \frac{1}{2}\dot{\boldsymbol{q}}^T M(\boldsymbol{q})\dot{\boldsymbol{q}} - V(\boldsymbol{q}), \qquad M^T = M,$$

differentiation with respect to velocity gives

$$\boxed{\boldsymbol{p} = M(\boldsymbol{q})\dot{\boldsymbol{q}}.} \tag{4.1}$$

The mass matrix converts a velocity into the dual quantity appropriate to the action pairing. It contains physical masses and coordinate geometry, just as in the small-oscillation lectures. In independent nonsingular mechanical coordinates it is positive definite, hence invertible. For a more general Lagrangian, the relation between $\boldsymbol{p}$ and $\dot{\boldsymbol{q}}$ need not be linear.

For planar polar motion,

$$T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2), \qquad p_r = m\dot{r}, \qquad p_\theta = mr^2\dot{\theta}.$$

The factor r^2 comes from the geometry of an angular displacement: $r\delta\theta$ is the corresponding tangential length. The coefficient conjugate to θ is the angular momentum, not the angular velocity. Cartesian coordinates can conceal this distinction because their mass map is particularly simple.

Conjugacy is a pairing, not a conservation statement. Momentum is conjugate to configuration because it multiplies the endpoint configuration displacement in the action response. Whether that momentum is conserved depends on the dynamics and its symmetry.

5 From matched endpoints to a phase-space map

The endpoint equations can be read as an evolution rule. Given (q_n, p_n), first solve

$$p_n = -\frac{\partial \mathcal{A}_n(q_n, q_{n+1})}{\partial q_n}$$

for q_{n+1}, then evaluate

$$p_{n+1} = \frac{\partial \mathcal{A}_n(q_n, q_{n+1})}{\partial q_{n+1}}.$$

This is locally possible when the mixed derivative

$$\frac{\partial^2 \mathcal{A}_n}{\partial q_n \, \partial q_{n+1}}$$

is nonzero. In several coordinates the corresponding mixed matrix must be invertible. The implicit function theorem is the mathematical reason that the first equation can then determine the next configuration. In the mechanical edge above, the scalar mixed derivative is $-m/h$, so the update is regular.

Stationarity has thus lifted a sequence of configurations into a sequence of states,

$$(q_n, p_n) \longmapsto (q_{n+1}, p_{n+1}).$$

Momentum is no longer just an auxiliary label on a formula: it is the endpoint datum needed to select the next edge. In the continuum we will make the corresponding change from $(\boldsymbol{q}, \dot{\boldsymbol{q}})$ to $(\boldsymbol{q}, \boldsymbol{p})$. For N independent configuration coordinates, this regular phase space has local dimension $2N$. If the equations depend explicitly on time, the time remains an additional known argument of the evolution rule.

5.1 An area-preservation preview

There is a geometric consequence of the endpoint identity that is worth seeing before introducing the continuum Hamiltonian. Keep the edge's time arguments fixed. On the graph of the regular update, its ordinary differential is

$$d\mathcal{A}_n = -p_n \, dq_n + p_{n+1} \, dq_{n+1}.$$

Here the endpoint momenta have been identified with the phase-space coordinates of the update. One more exterior differential gives

$$0 = d^2\mathcal{A}_n = -dp_n \wedge dq_n + dp_{n+1} \wedge dq_{n+1},$$

and hence

$$\boxed{dq_{n+1} \wedge dp_{n+1} = dq_n \wedge dp_n.} \tag{5.1}$$

Only two elementary facts are being used: mixed derivatives of a smooth scalar commute, giving $d^2\mathcal{A}_n = 0$, and the alternating product changes sign when its factors are interchanged.

To read the area notation concretely, consider two small independent perturbations (a_q, a_p) and (b_q, b_p) of an initial state. Their oriented parallelogram area is

$$a_q b_p - b_q a_p.$$

The expression $dq \wedge dp$ assigns exactly this value to the pair. Equation (5.1) says that evolving both perturbations by the linearized update preserves that oriented area. A patch may shear or stretch in different directions; it need not keep its shape. Smoothness and regularity allow this local statement to describe the transport of small phase-space regions.

5.2 The same check with an ordinary Jacobian

For (3.4), write the state column as $\boldsymbol{z}_n = (q_n, p_n)^T$. Its update Jacobian is

$$\frac{\partial \boldsymbol{z}_{n+1}}{\partial \boldsymbol{z}_n^T} = \begin{pmatrix} 1 - \dfrac{h^2}{m}\dfrac{d^2V}{dq^2}(q_n) & h/m \\ -h\dfrac{d^2V}{dq^2}(q_n) & 1 \end{pmatrix}.$$

Its determinant is 1. Thus the ordinary change-of-variables area factor confirms (5.1) without requiring the differential-form notation.

For N degrees of freedom the endpoint argument similarly preserves

$$\omega = \sum_{j=1}^{N} dq_j \wedge dp_j.$$

This is the symplectic structure whose full meaning will be developed later. Here it is a preview arising directly from the action, not a substitute for the later theory.

Area preservation is *not* a proof of conservation of a chosen energy function. In particular, a fixed-step discrete variational update can preserve this structure without preserving the continuum Hamiltonian exactly. Energy conservation and phase-space volume preservation answer different questions. We will return to that distinction in Lecture 16; no modified discrete-energy construction is needed here.

6 Moving endpoint events: why time is paired with minus energy

The discrete calculation held the time grid fixed. To uncover the quantity paired with time, return to the continuum action

$$\mathcal{S}[\boldsymbol{q}] = \int_{t_i}^{t_f} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)\,\mathrm{d}t, \qquad \boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{q}}}.$$

Unlike the fixed-event problem of Lecture 5, we now compare solutions whose endpoint *events* may move. An event specifies both a time and a configuration. We must therefore distinguish two changes that were identical when the endpoint time was fixed.

6.1 A path deformation and an endpoint displacement are different

Let $\delta_0\boldsymbol{q}(t)$ denote the infinitesimal change of the history evaluated at the *same* time t. It is the fixed-time path variation used earlier. Let dt_a and $d\boldsymbol{q}_a$ denote the ordinary infinitesimal changes of endpoint event a, where $a = i$ or f. If the time of that endpoint changes, the endpoint samples the neighboring history at a different time as well. The chain rule therefore gives

$$\boxed{d\boldsymbol{q}_a = \delta_0\boldsymbol{q}(t_a) + \dot{\boldsymbol{q}}(t_a)\,dt_a.} \tag{6.1}$$

All terms here are first-order changes. For a family $\boldsymbol{q}_\epsilon(t)$ with endpoint time $t_a(\epsilon)$, the distinction follows by differentiating $\boldsymbol{q}_\epsilon(t_a(\epsilon))$: one derivative changes the history at fixed time, while the other changes the evaluation time.

Two cases fix the interpretation. If $dt_a = 0$, then $d\boldsymbol{q}_a = \delta_0\boldsymbol{q}(t_a)$. If instead the endpoint is extended along the same unaltered history, $\delta_0\boldsymbol{q} = 0$ and $d\boldsymbol{q}_a = \dot{\boldsymbol{q}}_a dt_a$. Thus setting an endpoint's position change to zero while varying its time does *not* set its fixed-time path variation to zero; the latter must then be $-\dot{\boldsymbol{q}}_a dt_a$.

6.2 Keep both sources of action change

The action changes because its integrand changes and because its integration limits change. The moving-limit rule gives

$$\begin{aligned} d\mathcal{S} &= [L\,dt]_i^f \\ &\quad + \int_{t_i}^{t_f} \left[\left(\frac{\partial L}{\partial \boldsymbol{q}} \right)^T \delta_0\boldsymbol{q} + \boldsymbol{p}^T \frac{d}{dt}(\delta_0\boldsymbol{q}) \right] \mathrm{d}t. \end{aligned} \tag{6.2}$$

The explicit time argument inside the integral is held fixed during the path comparison; changing the limits is already accounted for by the first term. Integrating the second part by parts yields

$$d\mathcal{S} = [\boldsymbol{p}^T \delta_0\boldsymbol{q} + L\,dt]_i^f + \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial \boldsymbol{q}} - \dot{\boldsymbol{p}} \right)^T \delta_0\boldsymbol{q}\,\mathrm{d}t. \tag{6.3}$$

On a solution of the Euler–Lagrange equations, the interior coefficient vanishes. This is the meaning of *on shell*: the history satisfies the equations of motion. It does not mean its action is zero.

Use (6.1) to replace $\delta_0\boldsymbol{q}$ by the total endpoint displacement. Then

$$\begin{aligned} d\mathcal{S}_{\text{on shell}} &= [\boldsymbol{p}^T(d\boldsymbol{q} - \dot{\boldsymbol{q}}\,dt) + L\,dt]_i^f \\ &= [\boldsymbol{p}^T d\boldsymbol{q} - (\boldsymbol{p}^T\dot{\boldsymbol{q}} - L)\,dt]_i^f. \end{aligned}$$

The combination

$$\mathcal{E}_L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \equiv \boldsymbol{p}^T\dot{\boldsymbol{q}} - L \tag{6.4}$$

is called the Lagrangian energy function. Once the velocity is expressed in terms of $(\boldsymbol{q}, \boldsymbol{p}, t)$, this same function becomes the Hamiltonian H. We have therefore exposed the endpoint identity

$$\boxed{d\mathcal{S}_{\text{on shell}} = [\boldsymbol{p}^T d\boldsymbol{q} - H\,dt]_i^f.} \tag{6.5}$$

The bracket means final minus initial. In particular, the final endpoint coefficients are $(\boldsymbol{p}_f, -H_f)$; the initial ones have the opposite signs.

The minus sign in front of H was not guessed. The moving integration limit contributes $+L\,dt$, while the conversion from fixed-time variation to endpoint displacement contributes $-\boldsymbol{p}^T\dot{\boldsymbol{q}}\,dt$. Their sum gives $-H\,dt$.

We have identified the endpoint pairings

$$\boldsymbol{q} \longleftrightarrow \boldsymbol{p}, \qquad t \longleftrightarrow -H.$$

This does not add a new independent time coordinate to our ordinary $2N$-dimensional phase space, nor does it assert that both members of a pair are conserved. It identifies how the on-shell action responds to changes of its endpoint data. A local smooth family of solutions is assumed when viewing that action as a function of the endpoint events; arbitrary endpoint data need not determine one global solution.

6.3 A free-particle check of the signs

For a one-dimensional free particle joining (t_i, q_i) to (t_f, q_f), put $\tau = t_f - t_i > 0$. The stationary history is a straight line, so

$$p = \frac{m(q_f - q_i)}{\tau}, \qquad \mathcal{S}_{\text{on shell}} = \frac{m(q_f - q_i)^2}{2\tau}, \qquad E = \frac{p^2}{2m}.$$

Differentiate this explicit action with respect to all four endpoint data:

$$d\mathcal{S}_{\text{on shell}} = p(dq_f - dq_i) - E(dt_f - dt_i).$$

It has exactly the structure of (6.5). In particular, increasing t_f at fixed endpoint positions decreases the free-particle action: the particle has more time to cover the same distance and needs less kinetic energy. The derivative is $-E$, not $+E$.

By contrast, extending the final endpoint along the same straight history gives $dq_f = (p/m)dt_f$. Then $p\,dq_f - E\,dt_f = E\,dt_f = L\,dt_f$, the action of the short added segment. These are different endpoint comparisons, so they need not give the same response.

Stationarity does not erase boundary information. The interior action response vanishes on a solution. If endpoint data change, the remaining response is generally nonzero, and its coefficients identify conjugate momentum and the Hamiltonian.

7 The Legendre transform changes which variable is independent

7.1 When can momentum replace velocity?

To construct a Hamiltonian as a function on phase space, write a temporary independent velocity argument $\boldsymbol{v}$:

$$L = L(\boldsymbol{q}, \boldsymbol{v}, t), \qquad \boldsymbol{p} = \frac{\partial L}{\partial \boldsymbol{v}}(\boldsymbol{q}, \boldsymbol{v}, t).$$

Fix $\boldsymbol{q}$ and t. The question is whether this relation determines $\boldsymbol{v}$ from $\boldsymbol{p}$. Its Jacobian with respect to $\boldsymbol{v}$ is the velocity Hessian

$$\boxed{W(\boldsymbol{q}, \boldsymbol{v}, t) = \frac{\partial^2 L}{\partial \boldsymbol{v}\, \partial \boldsymbol{v}^T}, \qquad \det W \neq 0.} \tag{7.1}$$

Under this condition, the inverse function theorem gives a local inverse $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{q}, \boldsymbol{p}, t)$. This is the regularity assumed in the chapter. It is a local statement, not a claim that an arbitrary Lagrangian has one globally invertible momentum map. Positive-definite mechanical kinetic energy supplies a particularly simple regular case; positive definiteness is sufficient but not required for a local inverse.

The regularity in (7.1) concerns a continuum velocity-to-momentum map. The mixed endpoint derivative in Section 5 instead concerns a chosen discrete action edge. They serve analogous purposes but are not the same derivative test. If the continuum Hessian is singular, momenta can satisfy additional constraints, and the present unconstrained Hamiltonian construction is insufficient. That theory is outside this lecture's scope.

7.2 Why subtract the Lagrangian?

Define

$$\boxed{H(\boldsymbol{q},\boldsymbol{p},t) = \boldsymbol{p}^T\boldsymbol{v}(\boldsymbol{q},\boldsymbol{p},t) - L\big(\boldsymbol{q},\boldsymbol{v}(\boldsymbol{q},\boldsymbol{p},t),t\big).} \tag{7.2}$$

This is the Legendre transform with respect to the velocity arguments only. The position and time remain independent variables; the velocity has been eliminated in favor of its conjugate momentum. Leaving an unevaluated $\dot{\boldsymbol{q}}$ on the right is acceptable shorthand only when this elimination is understood.

There is a simple geometric interpretation in one velocity dimension. With q, t fixed, let $\ell(v) = L(q, v, t)$. At the selected velocity v, the slope is $p = d\ell/dv$. The tangent line, evaluated at a dummy horizontal coordinate u, is

$$y = \ell(v) + p(u - v) = pu - H.$$

Thus the transform records the negative tangent-line intercept while using the slope as its independent label. We do not need a global maximization principle: the local slope-to-velocity inverse is enough for the mechanics here.

The subtraction in (7.2) is also what makes the differential work. Changing $\boldsymbol{p}$ generally changes $\boldsymbol{v}$, but the resulting velocity-dependent terms cancel because $\boldsymbol{p} = \partial L/\partial \boldsymbol{v}$. That cancellation, not an arbitrary definition of energy, is the algebraic mechanism of the transform.

7.3 A useful mechanical formula

For $L = \frac{1}{2}\boldsymbol{v}^T M(\boldsymbol{q})\boldsymbol{v} - V(\boldsymbol{q})$ with symmetric invertible M,

$$\boldsymbol{p} = M\boldsymbol{v}, \qquad \boldsymbol{v} = M^{-1}\boldsymbol{p}.$$

Substituting gives

$$\boxed{H(\boldsymbol{q},\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^T M(\boldsymbol{q})^{-1}\boldsymbol{p} + V(\boldsymbol{q}).} \tag{7.3}$$

In velocity variables the kinetic quadratic form uses M; in momentum variables it uses M^{-1}. The same kinetic energy is being expressed through different independent data.

This replacement retains the local state information: given $(\boldsymbol{q}, \dot{\boldsymbol{q}})$, we can compute $\boldsymbol{p}$, and given $(\boldsymbol{q}, \boldsymbol{p})$, regularity lets us reconstruct $\dot{\boldsymbol{q}}$. It does not add degrees of freedom or turn momenta into extra physical particles.

8 Hamilton's equations from a differential identity

8.1 First change variables; only then use dynamics

Begin with $H = \boldsymbol{p}^T\boldsymbol{v} - L$ and differentiate. No equations of motion are needed for this step:

$$\begin{aligned} dH &= \boldsymbol{v}^T d\boldsymbol{p} + \boldsymbol{p}^T d\boldsymbol{v} - dL \\ &= \boldsymbol{v}^T d\boldsymbol{p} + \boldsymbol{p}^T d\boldsymbol{v} - \left[\left(\frac{\partial L}{\partial \boldsymbol{q}}\right)^T d\boldsymbol{q} + \left(\frac{\partial L}{\partial \boldsymbol{v}}\right)^T d\boldsymbol{v} + \frac{\partial L}{\partial t}\,dt\right]. \end{aligned}$$

Since $\boldsymbol{p} = \partial L/\partial \boldsymbol{v}$, the two coefficients of $d\boldsymbol{v}$ cancel. Thus

$$\boxed{dH = \boldsymbol{v}^T d\boldsymbol{p} - \left(\frac{\partial L}{\partial \boldsymbol{q}}\right)^T d\boldsymbol{q} - \frac{\partial L}{\partial t}\, dt.} \tag{8.1}$$

We have *not* set $d\boldsymbol{v} = 0$. In fact, on the new variable space,

$$d\boldsymbol{v} = \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{q}^T} d\boldsymbol{q} + \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}^T} d\boldsymbol{p} + \frac{\partial \boldsymbol{v}}{\partial t}\, dt.$$

Its entire coefficient vanishes, so no derivatives of the inverse map need to be carried through the calculation. This is the practical strength of the Legendre transform.

On the other hand, the ordinary differential of $H(\boldsymbol{q}, \boldsymbol{p}, t)$ is

$$dH = \left(\frac{\partial H}{\partial \boldsymbol{q}}\right)^T d\boldsymbol{q} + \left(\frac{\partial H}{\partial \boldsymbol{p}}\right)^T d\boldsymbol{p} + \frac{\partial H}{\partial t}\, dt.$$

Because the new arguments may be varied independently, comparison with (8.1) gives

$$\boxed{\frac{\partial H}{\partial \boldsymbol{p}} = \boldsymbol{v}, \qquad \frac{\partial H}{\partial \boldsymbol{q}} = -\frac{\partial L}{\partial \boldsymbol{q}}, \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.} \tag{8.2}$$

The arguments held fixed differ across these identities. In $\partial H/\partial \boldsymbol{q}$, hold $\boldsymbol{p}, t$ fixed. In $\partial L/\partial \boldsymbol{q}$, hold $\boldsymbol{v}, t$ fixed, then evaluate at $\boldsymbol{v}(\boldsymbol{q}, \boldsymbol{p}, t)$. Likewise, the time derivative of H holds $\boldsymbol{q}, \boldsymbol{p}$ fixed, whereas that of L holds $\boldsymbol{q}, \boldsymbol{v}$ fixed. The identity is not permission to differentiate a substituted expression while ignoring its chain rule.

8.2 Put the trajectory back into the calculation

Along a physical Lagrangian history, $\boldsymbol{v} = \dot{\boldsymbol{q}}$, and the Euler–Lagrange equations supply

$$\dot{\boldsymbol{p}} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) = \frac{\partial L}{\partial \boldsymbol{q}}.$$

Only here has the dynamical equation been used. Together with (8.2), it gives

$$\boxed{\dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}}, \qquad \dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}}.} \tag{8.3}$$

The first equation reconstructs velocity from phase-space data. The second evolves the momentum. Conversely, a solution of (8.3) in a regular coordinate neighborhood satisfies $\dot{\boldsymbol{q}} = \boldsymbol{v}(\boldsymbol{q}, \boldsymbol{p}, t)$, hence $\boldsymbol{p} = \partial L/\partial \dot{\boldsymbol{q}}$, and the second equation becomes Euler–Lagrange. The formulations describe the same local mechanical trajectories.

They organize those trajectories differently. The N second-order equations become $2N$ first-order equations for the state column

$$\boldsymbol{z} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix}, \qquad \dot{\boldsymbol{z}} = \begin{pmatrix} \partial H/\partial \boldsymbol{p} \\ -\partial H/\partial \boldsymbol{q} \end{pmatrix}.$$

Both descriptions require $2N$ initial data, together with the initial time when relevant. The Hamiltonian formulation exposes the endpoint-conjugate structure without changing the number of data needed to predict a motion.

9 Two familiar systems in canonical variables

9.1 Harmonic oscillator: a state follows an ellipse

Take $m, k > 0$ and

$$L(q, v) = \frac{1}{2}mv^2 - \frac{1}{2}kq^2.$$

The construction can be performed without any guess about the energy:

$$p = mv, \qquad v = \frac{p}{m}, \qquad H = p\frac{p}{m} - \left(\frac{p^2}{2m} - \frac{1}{2}kq^2\right) = \frac{p^2}{2m} + \frac{1}{2}kq^2.$$

Hamilton's equations are therefore

$$\dot{q} = \frac{p}{m}, \qquad \dot{p} = -kq.$$

Differentiating the first equation and using the second gives $\ddot{q}+(k/m)q = 0$. Conversely, a Newtonian oscillator solution with $p = m\dot{q}$ satisfies both first-order equations.

Put $\omega = \sqrt{k/m}$. For initial data (q_0, p_0) at $t = 0$, the solution is

$$\begin{aligned} q(t) &= q_0 \cos(\omega t) + \frac{p_0}{m\omega} \sin(\omega t), \\ p(t) &= p_0 \cos(\omega t) - m\omega q_0 \sin(\omega t). \end{aligned} \tag{9.1}$$

Both initial components matter. Two states at the same q_0 but different p_0 lie on different trajectories or at different stages of a trajectory.

Direct differentiation checks that

$$\frac{dH}{dt} = kq\dot{q} + \frac{p}{m}\dot{p} = 0.$$

For $E > 0$, the level curve $H = E$ is the ellipse

$$\frac{q^2}{2E/k} + \frac{p^2}{2mE} = 1.$$

With q horizontal and p vertical, the motion at the right-hand turning point points downward because $\dot{q} = 0$ and $\dot{p} = -kq < 0$. The flow is clockwise. The ellipse specifies the possible states at that energy; the vector field or (9.1) also specifies how fast and in which direction the state moves. At $E = 0$ the ellipse reduces to the equilibrium point.

A phase portrait with velocity on its vertical axis can look similar, but the canonical momentum is the distinguished coordinate. In this example $p = m\dot{q}$, so the two portraits differ by a constant scale. That simplification does not hold for general coordinates or general Lagrangians.

9.2 Central force: differentiate at fixed angular momentum

Return to the reduced planar two-body Lagrangian, with reduced mass μ, radial coordinate $r > 0$, and polar angle θ:

$$L = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2) - V(r).$$

The momenta and their inverses are

$$p_r = \mu\dot{r}, \qquad p_\theta = \mu r^2\dot{\theta}, \qquad \dot{r} = \frac{p_r}{\mu}, \qquad \dot{\theta} = \frac{p_\theta}{\mu r^2}.$$

Thus p_θ is the angular momentum about the perpendicular axis. The inverse is local in these polar coordinates and is not a coordinate chart at $r = 0$.

Performing the transform with *both* velocities gives

$$\begin{aligned} H &= p_r\dot{r} + p_\theta\dot{\theta} - L \\ &= \frac{p_r^2}{2\mu} + \frac{p_\theta^2}{2\mu r^2} + V(r). \end{aligned}$$

Now differentiate with the canonical arguments held fixed:

$$\boxed{\begin{aligned} \dot{r} &= \frac{p_r}{\mu}, & \dot{\theta} &= \frac{p_\theta}{\mu r^2}, \\ \dot{p}_r &= \frac{p_\theta^2}{\mu r^3} - \frac{dV}{dr}, & \dot{p}_\theta &= 0. \end{aligned}} \tag{9.2}$$

The radial derivative is a useful test of the variable change. At fixed p_θ, increasing r decreases the angular kinetic-energy term $p_\theta^2/(2\mu r^2)$. At fixed $\dot{\theta}$, increasing r instead increases $\frac{1}{2}\mu r^2\dot{\theta}^2$. These are different derivatives, just as (8.2) requires. Using the wrong quantity as fixed would give the wrong radial force.

To check the mechanics, substitute the inverse momentum relations into the radial equation:

$$\mu\ddot{r} = \mu r\dot{\theta}^2 - \frac{dV}{dr}, \qquad \mu(\ddot{r} - r\dot{\theta}^2) = -\frac{dV}{dr}.$$

This is the usual radial Newton equation. Because θ is absent from H, its conjugate momentum is constant; write $p_\theta = J$. After deriving the full equations, restrict to that conserved value:

$$H_{\text{radial}}(r, p_r; J) = \frac{p_r^2}{2\mu} + V_{\text{eff}}(r), \qquad V_{\text{eff}}(r) = V(r) + \frac{J^2}{2\mu r^2}. \tag{9.3}$$

This recovers the earlier one-dimensional reduction. Once $r(t)$ is found, the angle is reconstructed from

$$\theta(t) = \theta_0 + \int_0^t \frac{J}{\mu r(\tau)^2}\,\mathrm{d}\tau.$$

The angular coordinate is cyclic, not frozen: its momentum can be constant while the angle changes. Hamiltonian mechanics makes the conserved quantity one of the state coordinates and separates it cleanly from the reconstructed motion.

10 Definition, physical energy, and conservation are three questions

10.1 When the Hamiltonian equals kinetic plus potential energy

The definition is the regular Legendre transform (7.2). For a natural Lagrangian $L = T - V$, with V independent of velocity and T homogeneous of degree two in velocity,

$$\boldsymbol{p}^T\dot{\boldsymbol{q}} = 2T.$$

For $T = \frac{1}{2}\dot{\boldsymbol{q}}^T M \dot{\boldsymbol{q}}$, this follows immediately from $\boldsymbol{p} = M\dot{\boldsymbol{q}}$. Therefore

$$\boxed{H = 2T - (T - V) = T + V,} \tag{10.1}$$

with the velocities on the right expressed in phase-space variables. If T, V are the ordinary kinetic and potential energies in a time-independent coordinate description, this is the familiar mechanical energy. The crucial algebraic condition is *homogeneous* quadratic dependence, not merely the presence of a squared velocity somewhere in T.

10.2 Conservation requires a further condition

Along Hamiltonian motion,

$$\begin{aligned}\frac{dH}{dt} &= \left(\frac{\partial H}{\partial \boldsymbol{q}}\right)^T \dot{\boldsymbol{q}} + \left(\frac{\partial H}{\partial \boldsymbol{p}}\right)^T \dot{\boldsymbol{p}} + \frac{\partial H}{\partial t} \\ &= \left(\frac{\partial H}{\partial \boldsymbol{q}}\right)^T \frac{\partial H}{\partial \boldsymbol{p}} - \left(\frac{\partial H}{\partial \boldsymbol{p}}\right)^T \frac{\partial H}{\partial \boldsymbol{q}} + \frac{\partial H}{\partial t}.\end{aligned}$$

The first two terms are equal scalars with opposite signs, so

$$\boxed{\frac{dH}{dt} = \frac{\partial H}{\partial t}.} \tag{10.2}$$

If H has no explicit time dependence, it is conserved. This conclusion does not require guessing that H is a mechanical energy. Similarly, absence of a coordinate q_j gives $\dot{p}_j = -\partial H/\partial q_j = 0$, not $\dot{q}_j = 0$.

A time-dependent spring illustrates why identification and conservation are separate. For a prescribed stiffness $k(t)$,

$$H(q, p, t) = \frac{p^2}{2m} + \frac{1}{2}k(t)q^2 = T + V, \qquad \frac{dH}{dt} = \frac{1}{2}\dot{k}(t)q^2.$$

The Hamiltonian still equals the instantaneous kinetic plus potential energy, but an external change of the stiffness can supply or remove energy. This is a reduced, externally controlled description, not a violation of energy bookkeeping for a larger closed system.

10.3 A moving-coordinate check

Optional clarification. Even a free particle shows why the defining transform should not be replaced by a blanket statement that H always equals a particular physical energy. Let the laboratory coordinate be $x = q + ut$, where u is a prescribed constant speed. Express the laboratory free-particle Lagrangian in q, without subtracting any total derivative:

$$L(q, \dot{q}) = \frac{1}{2}m(\dot{q} + u)^2.$$

Then

$$p = m(\dot{q} + u), \qquad \dot{q} = \frac{p}{m} - u, \qquad H = p\dot{q} - L = \frac{p^2}{2m} - up.$$

The laboratory energy is $E_{\text{lab}} = \frac{1}{2}m\dot{x}^2 = p^2/(2m)$, so this Hamiltonian is $E_{\text{lab}} - up$, not E_{lab}. The kinetic energy written in $\dot{q}$ includes linear and constant terms and is not homogeneous quadratic in that velocity. Hamilton's equations correctly give $\dot{q} = p/m - u$ and $\dot{p} = 0$.

This also tests the endpoint pairing. Since $dx = dq + u\,dt$,

$$p\,dx - E_{\text{lab}}\,dt = p\,dq - (E_{\text{lab}} - up)\,dt.$$

The coefficient paired with endpoint time depends on which spatial coordinate is held fixed, exactly as the Hamiltonian change shows. No new mechanics is needed to explain the difference.

11 What has been gained, and what comes next?

We did not obtain a different oscillator or a different central-force orbit. We obtained a description in which the state uses the quantities naturally paired by the action. The discrete argument revealed the two endpoint responses of one edge, and stationarity matched them where edges meet. The continuum argument kept the endpoint information that fixed-boundary calculations usually set aside.

The main relations are

$$\begin{aligned}
\delta\mathcal{A}_n &= -p_n^-\delta q_n + p_{n+1}^+\delta q_{n+1}, & p_n^+ &= p_n^- \quad \text{at a stationary interior node,}\\
d\mathcal{S}_{\text{on shell}} &= [\boldsymbol{p}^T d\boldsymbol{q} - H\,dt]_i^f, & H &= \boldsymbol{p}^T\dot{\boldsymbol{q}} - L,\\
\dot{\boldsymbol{q}} &= \frac{\partial H}{\partial \boldsymbol{p}}, & \dot{\boldsymbol{p}} &= -\frac{\partial H}{\partial \boldsymbol{q}}.
\end{aligned}$$

The Legendre transform relates the last two lines: it expresses the endpoint-time coefficient in $(\boldsymbol{q}, \boldsymbol{p}, t)$ variables, and its differential converts Euler–Lagrange into a first-order state evolution. Its local invertibility assumption is part of the result, not an optional afterthought.

There are three distinctions to retain. A sampled discrete slope momentum need not be the matched momentum at the same node. A fixed-time path variation need not be the displacement of a moving endpoint event. And a momentum being conjugate to a coordinate does not by itself make that momentum conserved. Each distinction prevents a tempting but incorrect identification.

Finally, let $F(\boldsymbol{q}, \boldsymbol{p}, t)$ be any differentiable scalar observable. Its total change along a trajectory follows from the ordinary chain rule:

$$\dot{F} = \left(\frac{\partial F}{\partial \boldsymbol{q}}\right)^T \dot{\boldsymbol{q}} + \left(\frac{\partial F}{\partial \boldsymbol{p}}\right)^T \dot{\boldsymbol{p}} + \frac{\partial F}{\partial t}.$$

Substituting Hamilton's equations gives

$$\boxed{\dot{F} = \left(\frac{\partial F}{\partial \boldsymbol{q}}\right)^T \frac{\partial H}{\partial \boldsymbol{p}} - \left(\frac{\partial F}{\partial \boldsymbol{p}}\right)^T \frac{\partial H}{\partial \boldsymbol{q}} + \frac{\partial F}{\partial t}.} \tag{11.1}$$

For $F = q_j$ or $F = p_j$, this reproduces the two canonical equations. For $F = H$, the antisymmetric pair cancels and gives (10.2). The first two terms are one operation involving *two* phase-space functions, not a property of either one in isolation. Lecture 16 will name that operation and use it to study conservation, transformations, and phase-space flow.