Graduate Classical Mechanics
Lecture 17: Canonical Transformations

Symplectic structure and the geometry preserved by Hamiltonian mechanics

1 The organizing question

Lecture 16 described Hamiltonian evolution as a flow on phase space and proved that this flow preserves canonical volume. We now ask a different question: which changes of phase-space coordinates preserve the *form of Hamiltonian mechanics*? A useful coordinate change may simplify a Hamiltonian, but we do not want to lose the bracket rules that tell us how to obtain motion from it.

Write the old and new coordinate columns as

$$\boldsymbol{z} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix}, \qquad \boldsymbol{Z} = \begin{pmatrix} \boldsymbol{Q} \\ \boldsymbol{P} \end{pmatrix} = \boldsymbol{Z}(\boldsymbol{z}).$$

For most of this chapter, the coordinate map is smooth, locally invertible, and independent of time. Local invertibility lets us describe the same state using either set of variables. We will state explicitly when we instead discuss the physical flow between two times.

The candidate new coordinates are canonical if their brackets, *computed in the old canonical variables*, satisfy

$$\boxed{\{Q_i, Q_j\}_{\boldsymbol{z}} = 0, \qquad \{P_i, P_j\}_{\boldsymbol{z}} = 0, \qquad \{Q_i, P_j\}_{\boldsymbol{z}} = \delta_{ij}.} \tag{1}$$

The subscript matters. If we merely declared a new bracket formula in $(\boldsymbol{Q}, \boldsymbol{P})$, their coordinate brackets would have this form by definition. The actual test is whether differentiating their expressions in $(\boldsymbol{q}, \boldsymbol{p})$ gives the same operation. We will show that this coordinate test preserves the bracket of *every* pair of scalar observables, and then the canonical form of Hamilton's equations.

This is initially a change of description of the same state, not the active change of state generated by G in Lecture 16. Later, a physical time-evolution map will satisfy the same mathematical test. Keeping the two interpretations distinct prevents a second confusion: a canonical map is not necessarily a symmetry of a particular Hamiltonian. It can change the formula for H without changing the mechanics it describes.

The geometry to be preserved is not an ordinary Euclidean distance. Positions and momenta generally have different units, and phase space comes with no preferred rule for adding their squared coordinate differences. The relevant structure is their *conjugate pairing*. We start with one pair, where that structure can be seen as oriented area.

2 One conjugate pair: what the area form measures

2.1 A two-form takes two directions, not one trajectory

At a phase-space point (q, p), consider two infinitesimal displacement directions

$$\boldsymbol{u} = \begin{pmatrix} u_q \\ u_p \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} v_q \\ v_p \end{pmatrix}.$$

These can be thought of as two sides of a small parallelogram of neighboring states. They are not two successive instants on one particle trajectory. The differentials dq and dp extract components: $dq(\boldsymbol{u}) = u_q$, $dp(\boldsymbol{u}) = u_p$. Their wedge product is the antisymmetric measurement

$$\boxed{(dq \wedge dp)(\boldsymbol{u}, \boldsymbol{v}) = u_q v_p - v_q u_p.} \tag{2}$$

This determinant is the oriented area of the parallelogram in the (q, p) coordinate plane. Interchanging the two directions changes the sign; parallel directions give zero. The units are those of qp, namely action for a mechanical conjugate pair. This is not a metric area obtained by treating both axes as lengths in ordinary space.

The name *two-form* records the need for two direction inputs and their antisymmetric pairing. The notation summarizes the rules

$$dq \wedge dp = -dp \wedge dq, \qquad dq \wedge dq = dp \wedge dp = 0.$$

A single direction paired with itself gives zero, even when its ordinary length is nonzero. Thus this object cannot be a squared-length formula.

2.2 Transform the measurement

Let $Q = Q(q, p)$ and $P = P(q, p)$. Their ordinary differentials are

$$dQ = \frac{\partial Q}{\partial q}\, dq + \frac{\partial Q}{\partial p}\, dp, \qquad dP = \frac{\partial P}{\partial q}\, dq + \frac{\partial P}{\partial p}\, dp.$$

Expanding with the antisymmetry rules gives

$$\begin{aligned} dQ \wedge dP &= \left(\frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} \right) dq \wedge dp \\ &= \{Q, P\}_{\boldsymbol{z}}\, dq \wedge dp. \end{aligned} \tag{3}$$

The coefficient is both the two-dimensional Jacobian determinant and the Poisson bracket of the new coordinate functions. Consequently,

$$\boxed{\{Q, P\}_{\boldsymbol{z}} = 1 \quad \Longleftrightarrow \quad dQ \wedge dP = dq \wedge dp.} \tag{4}$$

The equality means that the transformed directions have the same pairing as the original directions. For a nonlinear map it is a pointwise statement about infinitesimal patches; the change-of-variables theorem then gives preservation of oriented areas of regions on which the map is one-to-one.

Area preservation does not require shape preservation. Reciprocal stretching or a shear may change angles and side lengths while leaving the determinant unchanged. We will use both examples below. Orientation is also essential: $Q = q,\ P = -p$ preserves *unsigned* area but gives $\{Q, P\}_{\boldsymbol{z}} = -1$, so it is not canonical under (1).

Lecture 15 already supplied a particular instance,

$$dq_{n+1} \wedge dp_{n+1} = dq_n \wedge dp_n,$$

for a regular discrete variational update. We are now identifying the structure behind that example rather than identifying an arbitrary numerical update with exact Hamiltonian flow.

3 Many degrees of freedom: the full pairing

3.1 Why there is a sum over conjugate pairs

For N degrees of freedom, define

$$\boxed{\omega = \sum_{i=1}^{N} dq_i \wedge dp_i.} \tag{5}$$

Here ω denotes a two-form, not an oscillator frequency. On two phase-space displacement columns it gives

$$\omega(\boldsymbol{u}, \boldsymbol{v}) = \sum_i (u_{q_i} v_{p_i} - v_{q_i} u_{p_i}) = \boldsymbol{u}_q^T \boldsymbol{v}_p - \boldsymbol{u}_p^T \boldsymbol{v}_q. \tag{6}$$

A pure q_1 displacement pairs with a pure p_1 displacement, but not with a pure p_2 displacement. The two-form records those relations, not just a total $2N$-dimensional volume. At each point the pairing is nondegenerate: if it vanishes for every $\boldsymbol{v}$, then $\boldsymbol{u} = 0$. To check this, choose $\boldsymbol{v}$ in one momentum direction to test each u_{q_i}, and in one coordinate direction to test each u_{p_i}.

Preservation means

$$\boxed{\sum_i dQ_i \wedge dP_i = \sum_i dq_i \wedge dp_i.} \tag{7}$$

It does *not* say that each old conjugate plane is individually fixed. For two coordinates with compatible units, take the same mixing of coordinates and momenta:

$$Q_1 = \frac{q_1 + q_2}{\sqrt{2}}, \quad Q_2 = \frac{q_1 - q_2}{\sqrt{2}}, \qquad P_1 = \frac{p_1 + p_2}{\sqrt{2}}, \quad P_2 = \frac{p_1 - p_2}{\sqrt{2}}.$$

For example,

$$dQ_1 \wedge dP_1 = \frac{1}{2}\big(dq_1 \wedge dp_1 + dq_1 \wedge dp_2 + dq_2 \wedge dp_1 + dq_2 \wedge dp_2\big).$$

The second pair has the same diagonal terms and opposite cross terms. Adding them gives (7). Neither new pair is simply one old pair. The *sum*, evaluated on every pair of directions, is what survives.

3.2 Compress the component conditions into a matrix

Use the ordering and Jacobian convention

$$\boldsymbol{z} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix}, \quad \boldsymbol{Z} = \begin{pmatrix} \boldsymbol{Q} \\ \boldsymbol{P} \end{pmatrix}, \quad \mathsf{C} = \frac{\partial \boldsymbol{Z}}{\partial \boldsymbol{z}^T}, \quad \mathsf{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \tag{8}$$

The rows of C correspond to new components and its columns to old ones, so $d\boldsymbol{Z} = \mathsf{C}\, d\boldsymbol{z}$. For a nonlinear transformation, C depends on the point; this is not a single global matrix approximation to the whole map.

The (a, b) entry of $\mathsf{C}\mathsf{J}\mathsf{C}^T$ is

$$\sum_i \left(\frac{\partial Z_a}{\partial q_i} \frac{\partial Z_b}{\partial p_i} - \frac{\partial Z_a}{\partial p_i} \frac{\partial Z_b}{\partial q_i} \right) = \{Z_a, Z_b\}_{\boldsymbol{z}}.$$

Thus all the canonical-coordinate brackets, including the vanishing cross-pair brackets, are precisely

$$\boxed{\mathsf{C}\mathsf{J}\mathsf{C}^T = \mathsf{J}.} \tag{9}$$

A smooth locally invertible map obeying this condition is called *symplectic*; in the present setting it is a canonical transformation.

There is an equivalent matrix test which acts directly on the two displacement directions. Equation (6) reads $\omega(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^T \mathsf{J} \boldsymbol{v}$. In new components the pairing is

$$(\mathsf{C}\boldsymbol{u})^T \mathsf{J} (\mathsf{C}\boldsymbol{v}) = \boldsymbol{u}^T \mathsf{C}^T \mathsf{J} \mathsf{C} \boldsymbol{v}.$$

Preserving this for every $\boldsymbol{u}, \boldsymbol{v}$ requires

$$\boxed{\mathsf{C}^T \mathsf{J} \mathsf{C} = \mathsf{J}.} \tag{10}$$

The transpose order differs from (9) because this calculation transforms displacement columns rather than scalar derivatives. The conditions are equivalent, not competing conventions. Since $\mathsf{J}^{-1} = -\mathsf{J}$, (9) gives

$$\mathsf{C}^{-1} = -\mathsf{J}\mathsf{C}^T\mathsf{J}.$$

Multiplying by J on the left and by C on the right yields (10); the reverse argument gives (9). These identities also show that the Jacobian is nonsingular. They guarantee local, not necessarily global, invertibility of the coordinate map.

An orthogonal matrix preserves a symmetric Euclidean inner product. A symplectic Jacobian preserves this antisymmetric pairing. There is no requirement that $\mathsf{C}^T\mathsf{C} = I$.

4 Why the criterion preserves every Poisson bracket

It would not be enough to preserve a few convenient trajectories. The coordinate-bracket criterion is stronger: it preserves the operation used to compute the change of *any* observable. Here is the chain-rule proof, with the coordinate representations distinguished explicitly.

Let $F(\boldsymbol{z}) = \widetilde{F}(\boldsymbol{Z}(\boldsymbol{z}))$ be the same scalar observable expressed in old and new coordinates. At fixed time,

$$\begin{aligned} dF &= \left(\frac{\partial \widetilde{F}}{\partial \boldsymbol{Z}}\right)^T d\boldsymbol{Z} \\ &= \left(\frac{\partial \widetilde{F}}{\partial \boldsymbol{Z}}\right)^T \mathsf{C}\, d\boldsymbol{z} \\ &= \left(\mathsf{C}^T \frac{\partial \widetilde{F}}{\partial \boldsymbol{Z}}\right)^T d\boldsymbol{z}. \end{aligned}$$

Comparing with $dF = (\partial F/\partial \boldsymbol{z})^T d\boldsymbol{z}$ gives the column-gradient chain rule

$$\boxed{\frac{\partial F}{\partial \boldsymbol{z}} = \mathsf{C}^T \frac{\partial \widetilde{F}}{\partial \boldsymbol{Z}}.} \tag{11}$$

Likewise $\partial G/\partial \boldsymbol{z} = \mathsf{C}^T \partial \widetilde{G}/\partial \boldsymbol{Z}$. Therefore

$$\begin{aligned}\{F, G\}_{\boldsymbol{z}} &= \left(\frac{\partial F}{\partial \boldsymbol{z}}\right)^T \mathsf{J} \frac{\partial G}{\partial \boldsymbol{z}} \\ &= \left(\frac{\partial \widetilde{F}}{\partial \boldsymbol{Z}}\right)^T \mathsf{C}\mathsf{J}\mathsf{C}^T \frac{\partial \widetilde{G}}{\partial \boldsymbol{Z}} \\ &= \left(\frac{\partial \widetilde{F}}{\partial \boldsymbol{Z}}\right)^T \mathsf{J} \frac{\partial \widetilde{G}}{\partial \boldsymbol{Z}} = \{\widetilde{F}, \widetilde{G}\}_{\boldsymbol{Z}}. \end{aligned} \tag{12}$$

The equality is evaluated at corresponding states. No constancy of C was used; only first derivatives enter the bracket, so the proof also applies to nonlinear canonical transformations. Explicit time dependence of the observables is harmless in this fixed-time calculation: time is held as a parameter, not included as an extra bracket coordinate.

The result explains why checking (1) is sufficient. Those coordinate functions test all entries of the bracket structure, and the chain rule extends their result to arbitrary functions.

5 Hamilton's equations in the new coordinates

For a time-independent canonical coordinate change, define the new Hamiltonian by substituting the inverse map:

$$\boxed{K(\boldsymbol{Z}, t) = H(\boldsymbol{z}(\boldsymbol{Z}), t).} \tag{13}$$

The functions can have different formulas, but their values at a given physical state are the same. The transformation is time independent; the Hamiltonian itself need not be. Using (11) for H,

$$\frac{\partial H}{\partial \boldsymbol{z}} = \mathsf{C}^T \frac{\partial K}{\partial \boldsymbol{Z}}.$$

Along a trajectory $\dot{\boldsymbol{Z}} = \mathsf{C}\dot{\boldsymbol{z}}$, so Hamilton's equations give

$$\begin{aligned}\dot{\boldsymbol{Z}} &= \mathsf{C}\mathsf{J} \frac{\partial H}{\partial \boldsymbol{z}} = \mathsf{C}\mathsf{J}\mathsf{C}^T \frac{\partial K}{\partial \boldsymbol{Z}} \\ &= \mathsf{J} \frac{\partial K}{\partial \boldsymbol{Z}}. \end{aligned} \tag{14}$$

In components,

$$\boxed{\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}.} \tag{15}$$

This is why the word *canonical* is useful: the structural form survives, although the Hamiltonian's expression changes.

Any invertible coordinate change can rewrite an ODE. A noncanonical one instead produces the matrix $\mathsf{C}\mathsf{J}\mathsf{C}^T$ in place of J; we cannot simply pretend its new variables obey the usual canonical brackets. A canonical change retains the standard differentiation rules without this extra coordinate-dependent structure.

The qualification on the map matters. If $\boldsymbol{Z} = \boldsymbol{Z}(\boldsymbol{z}, t)$, the chain rule has the additional term $\partial \boldsymbol{Z}/\partial t$. Substituting the inverse coordinates in H is then not, by itself, the full transformed-Hamiltonian rule. That theory belongs to Lecture 18; throughout the coordinate-change examples below, all transformation parameters are constant.

6 Operational examples: from a map to the mechanics

6.1 Reciprocal scaling

Take

$$\boxed{Q = aq, \qquad P = \frac{p}{a}}, \qquad a \neq 0.$$

Direct calculation gives $\{Q, P\}_{\boldsymbol{z}} = a(1/a) = 1$. A displacement in one direction is stretched while its conjugate direction shrinks. For negative a, both directions reverse, so the oriented area still has positive sign. With a dimensionless this is a rescaling of coordinate magnitudes, not a demand to give positions and momenta the same units.

For an oscillator with mass $m > 0$ and frequency $\nu > 0$,

$$H(q, p) = \frac{p^2}{2m} + \frac{1}{2}m\nu^2 q^2, \qquad K(Q, P) = \frac{a^2 P^2}{2m} + \frac{m\nu^2 Q^2}{2a^2}.$$

Then $\dot{Q} = a^2 P/m$, $\dot{P} = -m\nu^2 Q/a^2$, and $\ddot{Q} = -\nu^2 Q$. The appearance of the phase portrait changes, but its physical frequency does not. Here ν is a frequency; the symbol ω remains reserved for the two-form in this chapter.

6.2 Momentum shear: new canonical momentum is not necessarily mechanical momentum

Now use the second canonical map,

$$\boxed{Q = q, \qquad P = p + cq,} \tag{16}$$

where c is constant and has units of momentum per coordinate. Its derivatives give

$$\{Q, P\}_{\boldsymbol{z}} = 1 \cdot 1 - 0 \cdot c = 1, \qquad \mathsf{C} = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}.$$

The coordinate is unchanged while the momentum axis is sheared. Canonicality says nothing yet about which Hamiltonian is being studied. To see what the map does to dynamics, use the same oscillator and substitute $q = Q$, $p = P - cQ$:

$$\boxed{K(Q, P) = \frac{(P - cQ)^2}{2m} + \frac{1}{2}m\nu^2 Q^2.} \tag{17}$$

The two partial derivatives are taken holding the other *new* variable fixed. In particular,

$$\frac{\partial K}{\partial P} = \frac{P - cQ}{m}, \qquad \frac{\partial K}{\partial Q} = -\frac{c}{m}(P - cQ) + m\nu^2 Q.$$

Thus

$$\boxed{\dot{Q} = \frac{P - cQ}{m}, \qquad \dot{P} = \frac{c}{m}(P - cQ) - m\nu^2 Q.} \tag{18}$$

The extra term in $\dot{P}$ is not a new physical force. Since $P = p + cq$, its derivative must include $c\dot{q}$. Indeed,

$$\ddot{Q} = \frac{\dot{P} - c\dot{Q}}{m} = -\nu^2 Q, \qquad \dot{p} = \dot{P} - c\dot{Q} = -m\nu^2 Q.$$

Both recover the original oscillator. Equally importantly,

$$\boxed{p = m\dot{Q} = P - cQ, \qquad P = m\dot{Q} + cQ.} \tag{19}$$

The new momentum is canonical but is not simply mass times the new coordinate velocity. Canonical conjugacy is a structural relationship, not a universal mechanical-momentum formula.

For completeness, specify initial data (Q_0, P_0) at t_0, and set $\tau = t - t_0$. The solution is

$$\begin{aligned} Q(t) &= Q_0 \cos(\nu\tau) + \frac{P_0 - cQ_0}{m\nu} \sin(\nu\tau), \\ P(t) &= cQ(t) - m\nu Q_0 \sin(\nu\tau) + (P_0 - cQ_0)\cos(\nu\tau). \end{aligned} \tag{20}$$

Reconstruction is just $q(t) = Q(t)$, $p(t) = P(t) - cQ(t)$. The energy curve becomes

$$\frac{(P - cQ)^2}{2m} + \frac{1}{2} m\nu^2 Q^2 = E,$$

a sheared ellipse representing exactly the same motion. Setting $c = 0$ recovers the familiar formulas. This calculation follows the complete sequence: test the map, transform the Hamiltonian, derive the new equations, then reconstruct the original physical variables.

6.3 A noncanonical common scaling

If instead $Q = aq$, $P = ap$, then

$$\boxed{\{Q, P\}_{\boldsymbol{z}} = a^2.}$$

For nonzero a, this is invertible but is not canonical except when $a = \pm 1$. It scales the coordinate and its conjugate momentum in the same direction, changing area by a^2. For $a = 0$ it is not a coordinate change at all.

There is a dynamical consequence. Substitution gives

$$K = \frac{P^2}{2ma^2} + \frac{m\nu^2 Q^2}{2a^2}.$$

Treating (Q, P) as canonical would incorrectly predict $\dot{Q} = P/(ma^2)$. The actual transformed motion has $\dot{Q} = a\dot{q} = P/m$. The missing factor is exactly the noncanonical bracket $\{Q, P\}_{\boldsymbol{z}} = a^2$. An invertible replotting of a trajectory is not automatically a canonical description of its dynamics.

7 Symplectic preservation is stronger than volume preservation

7.1 Build the volume form from the two-form

The two-form ω takes two directions; a top-dimensional volume form takes $2N$ directions. They are different objects when $N > 1$. The connection is

$$\boxed{\frac{\omega^N}{N!} = dq_1 \wedge dp_1 \wedge \cdots \wedge dq_N \wedge dp_N.} \tag{21}$$

Here the power means repeated wedge products. To see both the result and the factorial, let $\alpha_i = dq_i \wedge dp_i$. Any product with a repeated α_i vanishes. Distinct two-forms α_i and α_j commute in the wedge product: moving two one-forms past two others produces four minus signs. In $(\alpha_1 + \cdots + \alpha_N)^N$, the surviving terms are the $N!$ orders of the same product.

For example, with two degrees of freedom,

$$(\alpha_1 + \alpha_2) \wedge (\alpha_1 + \alpha_2) = \alpha_1 \wedge \alpha_2 + \alpha_2 \wedge \alpha_1 = 2\alpha_1 \wedge \alpha_2.$$

Thus preserving ω preserves the canonical oriented volume form, and hence the positive Liouville volume measure. In particular a symplectic Jacobian has $\det \mathsf{C} = +1$, not just absolute determinant one. Taking determinants of (9) alone gives $(\det \mathsf{C})^2 = 1$; the preserved volume form fixes the sign. The interleaved order in (21) and the block order in (8) differ by a fixed orientation convention, not by a physical volume change.

For $N = 1$, the two-form is already the full volume form. Oriented area preservation and canonicality therefore coincide. For several degrees of freedom, knowing only the product of all local stretch factors cannot tell us how every conjugate pairing changes.

7.2 A four-dimensional counterexample

Consider the explicit map

$$\boxed{Q_1 = 2q_1, \qquad Q_2 = \frac{1}{2}q_2, \qquad P_1 = p_1, \qquad P_2 = p_2.}$$

In the order $\boldsymbol{z} = (q_1, q_2, p_1, p_2)^T$,

$$\mathsf{C} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \boxed{\det \mathsf{C} = 1.}$$

But

$$\mathsf{J} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \qquad \mathsf{C}\mathsf{J}\mathsf{C}^T = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ -2 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \end{pmatrix} \neq \mathsf{J}.$$

Equivalently,

$$\boxed{\{Q_1, P_1\}_{\boldsymbol{z}} = 2, \qquad \{Q_2, P_2\}_{\boldsymbol{z}} = \frac{1}{2}.} \tag{22}$$

The first conjugate-plane area doubles while the second halves. Their product, the four-dimensional volume, is unchanged, but their sum as a two-form is

$$dQ_1 \wedge dP_1 + dQ_2 \wedge dP_2 = 2\, dq_1 \wedge dp_1 + \frac{1}{2}\, dq_2 \wedge dp_2 \neq \omega.$$

A pair of displacement directions lying only in the first conjugate plane detects the failure immediately. The compensation in a different plane cannot restore that pair's value. This is why a determinant test is insufficient for canonicality in more than one degree of freedom:

$$\boxed{\text{symplectic} \Longrightarrow \text{volume preserving,}} \qquad \boxed{N > 1 : \text{ the converse fails.}}$$

8 Hamiltonian time evolution is itself canonical

We now change the interpretation of the map. Instead of assigning new labels to the same state, evolve *each* initial state by the Hamiltonian equations. The result is a map between states. Canonicality is a property of that whole map and its nearby-state derivatives, not a property one can test from a single trajectory alone.

8.1 Initial labels and instantaneous coordinates

Let

$$\boldsymbol{z}(t) = \Phi_{t,t_0}(\boldsymbol{z}_0), \qquad \mathsf{M}(t) = \frac{\partial \Phi_{t,t_0}(\boldsymbol{z}_0)}{\partial \boldsymbol{z}_0^T}. \tag{23}$$

The initial time t_0 is fixed. At each observation time t, the components of Φ_{t,t_0} are functions of the initial coordinate column $\boldsymbol{z}_0$. Write them as $Q_i(t;\boldsymbol{z}_0)$ and $P_i(t;\boldsymbol{z}_0)$ when testing canonical brackets. Thus the bracket to prove is

$$\{Q_i(t;\boldsymbol{z}_0), P_j(t;\boldsymbol{z}_0)\}_{\boldsymbol{z}_0} = \delta_{ij}.$$

The derivatives in this bracket vary the initial state while holding both times fixed. They are not derivatives with respect to the instantaneous components $\boldsymbol{z}(t)$, in which the coordinate brackets would already be canonical by definition.

Assume H has the commuting second phase-space derivatives used below, and work on a common interval where its solutions and differentiable flow exist uniquely. The argument permits explicit time dependence of H; conservation of energy is not assumed.

8.2 Transport the pairing of nearby-state displacements

Hamilton's equations have velocity field

$$\boldsymbol{u}(\boldsymbol{z},t) = \mathsf{J}\frac{\partial H}{\partial \boldsymbol{z}}.$$

Differentiating the flow equation with respect to its initial labels gives

$$\boxed{\dot{\mathsf{M}} = \mathsf{JDM}, \qquad \mathsf{M}(t_0) = I,} \tag{24}$$

where the matrix

$$\mathsf{D}(t) = \left.\frac{\partial^2 H}{\partial \boldsymbol{z}\,\partial \boldsymbol{z}^T}\right|_{\boldsymbol{z}=\Phi_{t,t_0}(\boldsymbol{z}_0)}$$

is the instantaneous Hessian, with $D_{ab} = \partial^2 H/(\partial z_a \partial z_b)$. It is symmetric when the mixed derivatives commute. Equation (24) is the chain rule: first differentiate the velocity field at the current state, then multiply by the change of current state caused by changing the initial data. It is the same nearby-state Jacobian idea used in Lecture 16's volume argument.

Now calculate how the entire antisymmetric pairing changes:

$$\begin{aligned}\frac{d}{dt}(\mathsf{M}^T\mathsf{JM}) &= \mathsf{M}^T\big[(\mathsf{JD})^T\mathsf{J} + \mathsf{J}(\mathsf{JD})\big]\mathsf{M} \\ &= \mathsf{M}^T(\mathsf{D}-\mathsf{D})\mathsf{M} = 0.\end{aligned} \tag{25}$$

The cancellation uses $\mathsf{D}^T = \mathsf{D}$, $\mathsf{J}^T = -\mathsf{J}$, and $\mathsf{J}^2 = -I$. Unlike the trace cancellation in Liouville's theorem, this is a full matrix identity. Together with the initial value $\mathsf{M}(t_0) = I$, it gives

$$\boxed{\mathsf{M}(t)^T \mathsf{J} \mathsf{M}(t) = \mathsf{J}, \qquad \mathsf{M}(t) \mathsf{J} \mathsf{M}(t)^T = \mathsf{J}.} \tag{26}$$

The equivalence of the two forms was established in Section 3. The second form reads, component by component,

$$\boxed{\{Q_i, Q_j\}_{\boldsymbol{z}_0} = 0, \qquad \{P_i, P_j\}_{\boldsymbol{z}_0} = 0, \qquad \{Q_i, P_j\}_{\boldsymbol{z}_0} = \delta_{ij}.} \tag{27}$$

The initial labels and the time-dependent arguments in these functions remain those of (23). Thus exact Hamiltonian evolution preserves the conjugate pairing of any two neighboring-state displacements. Its preservation of their full phase-space volume is a consequence of this stronger result.

For an autonomous Hamiltonian, the maps depend only on $\tau = t - t_0$ and compose as $\Phi_{\tau+s} = \Phi_\tau \circ \Phi_s$ wherever the solutions exist. This is the one-parameter family usually described as Hamiltonian evolution. For a driven Hamiltonian, Φ_{t,t_0} generally depends on both endpoint times; each such map is still canonical. The explicit-time energy changes discussed in Lecture 16 do not spoil (25).

As a direct check, the free-particle flow is

$$Q = q_0 + \frac{\tau}{m} p_0, \qquad P = p_0, \qquad \mathsf{M} = \begin{pmatrix} 1 & \tau/m \\ 0 & 1 \end{pmatrix}.$$

Computing the bracket in (q_0, p_0) gives $\{Q, P\}_{\boldsymbol{z}_0} = 1$. This recovers the same transported shear used to illustrate Liouville's theorem, now as an example of canonicality.

8.3 How the Jacobi argument is interpreted

A second view, for an autonomous Hamiltonian. The bracket proof can also be stated using Jacobi, but its variables need to be fixed carefully. Let

$$A_\tau(\boldsymbol{z}_0) = A(\Phi_\tau(\boldsymbol{z}_0))$$

for a time-independent observable A. Then

$$\frac{\partial A_\tau}{\partial \tau} = \{A_\tau, H(\boldsymbol{z}_0)\}_{\boldsymbol{z}_0}. \tag{28}$$

This is not an arbitrary substitution of initial for instantaneous coordinates. To justify it without already assuming canonicality, differentiate the autonomous composition rule $\Phi_\tau(\Phi_s(\boldsymbol{z}_0)) = \Phi_{\tau+s}(\boldsymbol{z}_0)$ with respect to s at zero. It gives

$$\mathsf{M}(\tau) \boldsymbol{u}(\boldsymbol{z}_0) = \boldsymbol{u}(\Phi_\tau(\boldsymbol{z}_0)).$$

The chain rule for A_τ then equates its directional derivative along the initial velocity field with its derivative along the current motion, giving (28).

Apply this to the component observables $Q_i(\tau; \boldsymbol{z}_0)$ and $P_j(\tau; \boldsymbol{z}_0)$, and denote their initial-label bracket by $B_{ij}(\tau, \boldsymbol{z}_0)$. Differentiating at fixed $\boldsymbol{z}_0$ and using Jacobi yields

$$\begin{aligned} \frac{\partial B_{ij}}{\partial \tau} &= \{\{Q_i, H\}, P_j\}_{\boldsymbol{z}_0} + \{Q_i, \{P_j, H\}\}_{\boldsymbol{z}_0} \\ &= \{B_{ij}, H\}_{\boldsymbol{z}_0}. \end{aligned} \tag{29}$$

Every bracket in this equation is taken in $\boldsymbol{z}_0$, with $H = H(\boldsymbol{z}_0)$. The constant $B_{ij} = \delta_{ij}$ solves this equation and its initial condition. It is the unique local solution: for the difference $E = B_{ij} - \delta_{ij}$, the equation is $\partial E/\partial\tau = \boldsymbol{u}^T \partial E/\partial \boldsymbol{z}_0$; along backward characteristics $d\boldsymbol{z}_0/d\tau = -\boldsymbol{u}$, the total derivative of E vanishes. Its zero initial values remain zero. The same reasoning preserves the initially vanishing QQ and PP brackets.

Merely observing that the first time derivative vanishes initially would not prove preservation for later times. The transport equation and its initial data supply that missing step. The Hessian proof above avoids this autonomous-flow qualification and applies directly to $H(\boldsymbol{z}, t)$.

Finally, a flow map indexed by t, t_0 is not the same use of a time-dependent formula as adopting a moving coordinate system during a calculation. Here we compare entire states at two fixed times and differentiate with respect to initial labels. The transformed-Hamiltonian correction for a time-dependent *coordinate description* remains a separate subject for Lecture 18.

9 The canonical one-form: what is integrated along a path?

The two-form compares two directions. To reconnect with action and endpoints we also need a quantity evaluated on *one* direction:

$$\boxed{\vartheta = \boldsymbol{p}^T d\boldsymbol{q} = \sum_i p_i \, dq_i.} \tag{30}$$

At a point $(\boldsymbol{q}, \boldsymbol{p})$, its value on $\boldsymbol{u} = (\boldsymbol{u}_q, \boldsymbol{u}_p)$ is

$$\vartheta(\boldsymbol{u}) = \boldsymbol{p}^T \boldsymbol{u}_q.$$

A pure momentum displacement at fixed configuration gives zero; a configuration displacement is weighted by the momentum at that point. This is the endpoint pairing encountered in Lecture 15. The symbol ϑ is a *one-form*, a rule for this linear measurement at every point, not a scalar function whose value is just $\boldsymbol{p}^T \boldsymbol{q}$.

9.1 From a local measurement to a line integral

Parameterize any phase-space curve by λ:

$$\gamma(\lambda) = (\boldsymbol{q}(\lambda), \boldsymbol{p}(\lambda)), \qquad \lambda_a \le \lambda \le \lambda_b.$$

Then its one-form integral means

$$\boxed{\int_\gamma \vartheta = \int_{\lambda_a}^{\lambda_b} \boldsymbol{p}(\lambda)^T \frac{d\boldsymbol{q}}{d\lambda} \, d\lambda.} \tag{31}$$

The parameter need not be physical time and the curve need not be a mechanical solution. An orientation-preserving reparameterization changes the velocity and parameter interval inversely, leaving the integral unchanged. If $\lambda = t$ on a physical path, this becomes $\int \boldsymbol{p}^T \dot{\boldsymbol{q}} \, dt$, the momentum part of the phase-space action from Lecture 15. It has units of action.

It is important not to replace ϑ by $d(\boldsymbol{p}^T\boldsymbol{q})$. The ordinary product rule gives

$$d(\boldsymbol{p}^T\boldsymbol{q}) = \boldsymbol{p}^T d\boldsymbol{q} + \boldsymbol{q}^T d\boldsymbol{p}.$$

The missing second term prevents ϑ from being an exact differential in general. For a concrete one-pair check, follow the rectangular loop

$$(0,0) \to (a,0) \to (a,b) \to (0,b) \to (0,0), \qquad a, b > 0.$$

The lower horizontal segment has $p = 0$, and the vertical segments have $dq = 0$. The upper segment contributes $\int_a^0 b\, dq = -ab$. Thus

$$\oint p\, dq = -ab \neq 0,$$

whereas the integral of an exact differential around this closed loop would be zero. The one-form's path dependence is real; canonicality will constrain the *difference* between the old and new one-forms, not make each one separately path independent.

9.2 Taking an exterior derivative

The exterior derivative extends the familiar differential from scalar functions to forms. The rules needed here are the product rule, antisymmetry of the wedge product, and $d^2 = 0$. In particular,

$$d(p_i\, dq_i) = dp_i \wedge dq_i + p_i\, d(dq_i) = dp_i \wedge dq_i.$$

The identity $d(dq_i) = 0$ is not a statement about the acceleration of a trajectory. An ordinary second derivative $d^2q_i/d\lambda^2$ may certainly be nonzero. Exterior differentiation antisymmetrizes changes in two independent directions; commuting scalar mixed derivatives cancel in that operation.

Consequently,

$$\begin{aligned} -d\vartheta &= -\sum_i d(p_i\, dq_i) = -\sum_i dp_i \wedge dq_i \\ &= \sum_i dq_i \wedge dp_i = \omega. \end{aligned} \tag{32}$$

The sign is fixed by the chosen order $dq_i \wedge dp_i$. The negative integral around the positively oriented rectangular loop above is consistent with this sign. The one-form and the two-form have related but distinct jobs: one is integrated along a path, while the other pairs two local directions.

10 An exact difference: the doorway to generating functions

For a canonical map, express the new one-form $\Theta = \boldsymbol{P}^T d\boldsymbol{Q}$ in the old variables by substituting $\boldsymbol{Q}(\boldsymbol{z})$, $\boldsymbol{P}(\boldsymbol{z})$. We must compare both forms on the same phase-space patch, not subtract unrelated symbols living in unrelated coordinate systems. Preservation of ω gives

$$d\vartheta = -\omega = d\Theta, \qquad \boxed{d(\vartheta - \Theta) = 0.}$$

A one-form with zero exterior derivative is called *closed*. On a sufficiently small simply connected coordinate neighborhood, a closed one-form is the differential of a scalar, hence is *exact*. Using that local calculus fact here,

$$\boxed{\boldsymbol{p}^T d\boldsymbol{q} - \boldsymbol{P}^T d\boldsymbol{Q} = dF \quad \text{locally.}} \tag{33}$$

We do not assert a single globally defined F across holes, singularities, or multiple coordinate branches. Nor have we specified which mixture of old and new variables can serve as independent arguments of F. Those are additional questions, not consequences of writing (33).

For any curve from state A to state B within the patch,

$$\int_\gamma \boldsymbol{p}^T d\boldsymbol{q} - \int_\gamma \boldsymbol{P}^T d\boldsymbol{Q} = F(B) - F(A). \tag{34}$$

Each integral can depend on the path, but their difference depends only on endpoints. This is the same type of boundary information that repeatedly appeared in the variational formulation. It explains why preserving the two-form need not preserve the one-form exactly.

10.1 Check the distinction on the momentum shear

Return to $Q = q$, $P = p + cq$. Since $dQ = dq$,

$$p\,dq - P\,dQ = -cq\,dq = d\left(-\frac{1}{2}cq^2\right). \tag{35}$$

Here the local scalar is $F = -cq^2/2$. It has units of action and yields the endpoint difference

$$\int p\,dq - \int P\,dQ = -\frac{c}{2}(q_B^2 - q_A^2)$$

for any path joining the two endpoints. The one-forms are not equal when $c \neq 0$, but their exterior derivatives are equal. By contrast, reciprocal scaling has $P\,dQ = (p/a)a\,dq = p\,dq$, so its difference vanishes outright.

The shear check does not yet claim a generating function with independent arguments (q, Q): here $Q = q$, so that pair is not independent. It only verifies the exact-differential relation on the original phase-space patch. Learning how to choose independent arguments and use F to *construct* canonical maps is the next step. Lecture 18 will develop that construction through ordinary multivariable calculus and Green/Stokes before organizing the generating-function representations. We stop at the structural doorway here.

11 Summary and handoff

A canonical transformation preserves more than the picture of a trajectory and, in several degrees of freedom, more than total volume. Its defining structure can be tested in equivalent ways:

$$\boxed{\begin{array}{c} \{Q_i, Q_j\}_{\boldsymbol{z}} = 0, \quad \{P_i, P_j\}_{\boldsymbol{z}} = 0, \quad \{Q_i, P_j\}_{\boldsymbol{z}} = \delta_{ij} \\ \Updownarrow \\ \mathsf{C}\mathsf{J}\mathsf{C}^T = \mathsf{J} \quad \Longleftrightarrow \quad \mathsf{C}^T\mathsf{J}\mathsf{C} = \mathsf{J} \\ \Updownarrow \\ \displaystyle\sum_i dQ_i \wedge dP_i = \sum_i dq_i \wedge dp_i. \end{array}}$$

The chain rule then gives bracket invariance for arbitrary observables. For a time-independent coordinate map, the transformed Hamiltonian is obtained by inverse substitution and Hamilton's equations retain their canonical form. The momentum-shear example shows why canonical momentum need not equal mechanical momentum, and why a canonical coordinate change need not leave the Hamiltonian's formula unchanged.

Preserving ω preserves $\omega^N/N!$, the volume form. The reverse implication holds for one conjugate pair, but fails in higher dimensions: a volume-preserving map can change individual pairing measurements. Exact Hamiltonian flow satisfies the stronger symplectic condition, even for a time-dependent Hamiltonian, on the interval where the differentiable flow exists. Brackets of evolved components in this assertion are computed with respect to the *initial* state labels.

Finally, $\vartheta = \boldsymbol{p}^T d\boldsymbol{q}$ is evaluated on one direction and integrated along a curve, while $\omega = -d\vartheta$ pairs two directions. Canonicality allows the old and new one-forms to differ by a locally exact term:

$$\boxed{\boldsymbol{p}^T d\boldsymbol{q} - \boldsymbol{P}^T d\boldsymbol{Q} = dF.}$$

This is an endpoint-only change, not a claim that either one-form is itself path independent. The next lecture asks how to turn this relation into a practical method for choosing canonical coordinates. Generating-function types, time-dependent coordinate corrections, and action–angle variables have not been developed here.