Graduate Classical Mechanics
Lecture 18: Generating Functions

Constructing canonical coordinates that simplify the motion

1 From testing a coordinate change to constructing one

Lecture 17 gave several ways to test whether a proposed phase-space coordinate change is canonical. For a smooth, locally invertible map

$$\boldsymbol{Q} = \boldsymbol{Q}(\boldsymbol{q}, \boldsymbol{p}), \qquad \boldsymbol{P} = \boldsymbol{P}(\boldsymbol{q}, \boldsymbol{p}),$$

we can calculate the new coordinate brackets using the old variables, test the Jacobian against the symplectic matrix, or compare the old and new conjugate pairings. Those tests answer: *does this map preserve the structure of Hamilton's equations?* They do not yet tell us how to find a useful map.

The present question is constructive. Can we choose coordinates in which the Hamiltonian, and therefore the motion, becomes simpler? A useful coordinate need not resemble a position in ordinary space. For the oscillator, for example, an orbit label and a phase are more natural than two quantities that both oscillate.

The starting point is the difference of the canonical one-forms,

$$\boxed{\alpha = \boldsymbol{p}^T d\boldsymbol{q} - \boldsymbol{P}^T d\boldsymbol{Q}.} \tag{1}$$

Both terms must be expressed on the same phase-space patch: substitute the functions $\boldsymbol{Q}(\boldsymbol{q}, \boldsymbol{p})$, $\boldsymbol{P}(\boldsymbol{q}, \boldsymbol{p})$ before comparing them. The one-form $\boldsymbol{p}^T d\boldsymbol{q}$ is generally path dependent. Lecture 17's rectangular-loop calculation showed that it is not, by itself, the differential of a scalar. What canonicality controls is the *difference* in (1).

We will first calculate why this difference is locally an exact differential,

$$\alpha = dF.$$

The scalar F is a generating function. Its derivatives encode the coordinate change. We will then restore time dependence, construct the oscillator's simplifying coordinates, and only afterward organize the different possible choices of independent arguments of F. The four standard types will be four representations of one differential relation, not four unrelated recipes.

2 One phase plane: canonicality removes circulation

2.1 Write the line element in ordinary coordinates

Begin with one degree of freedom and suppress time for the moment:

$$Q = Q(q, p), \qquad P = P(q, p).$$

The ordinary differential of Q is

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

Consequently,

$$\alpha = A\,dq + B\,dp, \qquad A = p - P\frac{\partial Q}{\partial q}, \qquad B = -P\frac{\partial Q}{\partial p}. \tag{2}$$

Here A and B are just the two coefficients of the line element. A path in this calculation is a comparison path through phase space, not necessarily a solution of any Hamiltonian system. The parameter used to traverse it need not be physical time.

Take two paths γ_1 and γ_2 from the same state z_a to the same state z_b. Their integrals agree precisely when the integral around the resulting closed loop vanishes. Figure 1 fixes the orientation: traverse the lower path forward and the upper path backward, so that the enclosed region lies on the left.

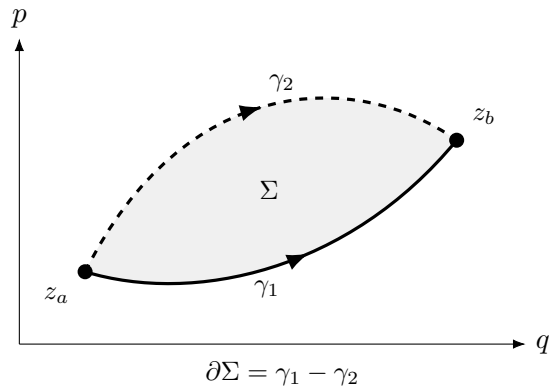

Figure 1: Both path arrows run from z_a to z_b. For the closed boundary, return along γ_2 in the reverse direction. This boundary is positively oriented for the (q, p) plane. The region and both paths must lie in one smooth coordinate patch.

2.2 Calculate the circulation coefficient

Green's theorem gives

$$\begin{aligned}\int_{\gamma_1} \alpha - \int_{\gamma_2} \alpha &= \oint_{\partial\Sigma} \alpha \\ &= \iint_{\Sigma} \left(\frac{\partial B}{\partial q} - \frac{\partial A}{\partial p}\right) dq\,dp.\end{aligned} \tag{3}$$

The condition for path independence is therefore visible in ordinary partial derivatives. Using (2), calculate each term before simplifying:

$$\begin{aligned}\frac{\partial B}{\partial q} &= -\frac{\partial P}{\partial q}\frac{\partial Q}{\partial p} - P\frac{\partial^2 Q}{\partial q\,\partial p}, \\ \frac{\partial A}{\partial p} &= 1 - \frac{\partial P}{\partial p}\frac{\partial Q}{\partial q} - P\frac{\partial^2 Q}{\partial p\,\partial q}.\end{aligned}$$

Assume sufficient smoothness for the mixed derivatives of Q to commute. Subtracting cancels them and leaves

$$\begin{aligned}\frac{\partial B}{\partial q} - \frac{\partial A}{\partial p} &= -1 + \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} \\ &= \boxed{\{Q, P\}_{q,p} - 1.}\end{aligned} \tag{4}$$

For a canonical map, $\{Q, P\}_{q,p} = 1$, so the circulation vanishes. The bracket is computed in the old variables; declaring a new bracket formula in (Q, P) would not establish this equality.

The subtraction of 1 has a structural meaning. It is the old conjugate pairing, while the derivative products measure the new one expressed in the old coordinates. Canonicality makes them agree. As a check, a common scaling $Q = aq$, $P = ap$ gives the coefficient $a^2 - 1$. Unless $a^2 = 1$, the difference of one-forms has nonzero circulation and cannot be dF on an open patch. Invertibility alone is not enough.

2.3 Define the scalar instead of assuming it

Choose a reference state z_0 inside a sufficiently small simply connected neighborhood. The integral

$$F(z) - F(z_0) = \int_{z_0}^{z} \alpha \tag{5}$$

is independent of the path chosen within that neighborhood. It therefore defines an ordinary scalar function of the endpoint. If we move that endpoint a short distance (dq, dp), append a short final segment to the path. To first order, the integral changes by $A\,dq + B\,dp$. Thus

$$\frac{\partial F}{\partial q} = A, \qquad \frac{\partial F}{\partial p} = B, \qquad \boxed{dF = p\,dq - P\,dQ.}$$

This is the reason the generating function exists locally. It is not an additional mechanical law.

The locality matters. A closed loop surrounding a hole or crossing a branch cut need not bound a region inside the chosen patch. Zero local circulation does not by itself supply one globally single-valued function on an arbitrary domain. Likewise, F is determined only up to an additive constant at this stage. Neither qualification changes the local construction.

3 Several degrees of freedom: all conjugate pairs contribute

A map in four-dimensional phase space may mix all four old coordinates. To see what changes, take two degrees of freedom and restrict to the (q_1, p_1) plane by holding q_2, p_2 fixed. Although only one old pair is varied, both new coordinates Q_1 and Q_2 may change. On this plane,

$$\alpha = A_1\,dq_1 + B_1\,dp_1,$$

where

$$A_1 = p_1 - \sum_{i=1}^{2} P_i \frac{\partial Q_i}{\partial q_1}, \qquad B_1 = -\sum_{i=1}^{2} P_i \frac{\partial Q_i}{\partial p_1}.$$

Applying the same product rule and cancellation as before gives

$$\frac{\partial B_1}{\partial q_1} - \frac{\partial A_1}{\partial p_1} = -1 + \sum_{i=1}^{2} \left(\frac{\partial Q_i}{\partial q_1}\frac{\partial P_i}{\partial p_1} - \frac{\partial Q_i}{\partial p_1}\frac{\partial P_i}{\partial q_1} \right). \tag{6}$$

The sum is essential. A canonical map need not preserve each old coordinate plane separately; it preserves the total conjugate pairing of two directions, as the mixing example in Lecture 17 illustrated.

Write $\boldsymbol{z} = (\boldsymbol{q}^T, \boldsymbol{p}^T)^T$, $\boldsymbol{Z} = (\boldsymbol{Q}^T, \boldsymbol{P}^T)^T$, and use the course's Jacobian convention

$$\mathsf{C} = \frac{\partial \boldsymbol{Z}}{\partial \boldsymbol{z}^T}, \qquad \mathsf{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Lecture 17 established the equivalent canonical tests

$$\mathsf{C}\mathsf{J}\mathsf{C}^T = \mathsf{J}, \qquad \mathsf{C}^T\mathsf{J}\mathsf{C} = \mathsf{J}.$$

The second form compares transformed displacement directions. Its (q_1, p_1) entry says that the sum in (6) is 1, so the circulation again vanishes.

For N degrees of freedom there are also planes involving nonmatching old indices and planes with two positions or two momenta. The required identities are

$$\sum_{i=1}^{N} \left(\frac{\partial Q_i}{\partial q_j}\frac{\partial P_i}{\partial p_k} - \frac{\partial Q_i}{\partial p_k}\frac{\partial P_i}{\partial q_j} \right) = \delta_{jk}, \tag{7}$$

$$\sum_{i=1}^{N} \left(\frac{\partial Q_i}{\partial q_j}\frac{\partial P_i}{\partial q_k} - \frac{\partial Q_i}{\partial q_k}\frac{\partial P_i}{\partial q_j} \right) = 0, \tag{8}$$

$$\sum_{i=1}^{N} \left(\frac{\partial Q_i}{\partial p_j}\frac{\partial P_i}{\partial p_k} - \frac{\partial Q_i}{\partial p_k}\frac{\partial P_i}{\partial p_j} \right) = 0. \tag{9}$$

These are the component entries of the same matrix condition, not extra restrictions on a canonical map. Expanding the coefficients of α in all the old coordinates, these identities make every antisymmetric difference of coefficient derivatives vanish. They are precisely the local integrability conditions for those coefficients to be the partial derivatives of one scalar.

On a sufficiently small contractible coordinate patch, paths can be continuously deformed into one another without meeting a singularity or leaving the patch. The zero-circulation condition then gives

$$\boxed{\boldsymbol{p}^T d\boldsymbol{q} - \boldsymbol{P}^T d\boldsymbol{Q} = dF.} \tag{10}$$

Optional notation check. In the exterior-form notation introduced in Lecture 17, all the plane-by-plane calculations compress to

$$d\alpha = 0 \quad \Longrightarrow \quad \alpha = dF \quad \text{locally.}$$

The compact statement now summarizes a calculation we can carry out in components. Notice that the scalar has initially been constructed as a function on the *old phase-space patch*. We have not yet shown that any particular mixture, such as $(\boldsymbol{q}, \boldsymbol{Q})$, can serve as its independent arguments.

4 Time dependence and the Hamiltonian correction

4.1 A fixed-time primitive is not yet the full differential

Suppose the canonical map depends explicitly on time. At each fixed time we may repeat the preceding construction, obtaining a primitive $\widehat{F}(\boldsymbol{q},\boldsymbol{p},t)$. But the full differential of $\boldsymbol{Q}(\boldsymbol{q},\boldsymbol{p},t)$ now contains the additional contribution

$$\left(\frac{\partial \boldsymbol{Q}}{\partial t}\right)_{\boldsymbol{q},\boldsymbol{p}} dt.$$

The fixed-time construction therefore implies

$$\begin{aligned} d\widehat{F} = \boldsymbol{p}^T d\boldsymbol{q} - \boldsymbol{P}^T d\boldsymbol{Q} \\ + \left[\left(\frac{\partial \widehat{F}}{\partial t}\right)_{\boldsymbol{q},\boldsymbol{p}} + \boldsymbol{P}^T \left(\frac{\partial \boldsymbol{Q}}{\partial t}\right)_{\boldsymbol{q},\boldsymbol{p}}\right] dt. \end{aligned} \tag{11}$$

Here the subscripts outside parentheses specify what is held fixed; the derivatives themselves remain explicit. The second term inside the brackets accounts for the motion of the coordinate grid. Ignoring it would confuse a time-dependent change of coordinates with a fixed one.

Define the transformed Hamiltonian K, expressed finally in the new variables, so that

$$\boxed{\boldsymbol{p}^T d\boldsymbol{q} - H\, dt = \boldsymbol{P}^T d\boldsymbol{Q} - K\, dt + d\widehat{F}.} \tag{12}$$

Equation (11) shows exactly which correction is required:

$$K = H + \left(\frac{\partial \widehat{F}}{\partial t}\right)_{\boldsymbol{q},\boldsymbol{p}} + \boldsymbol{P}^T \left(\frac{\partial \boldsymbol{Q}}{\partial t}\right)_{\boldsymbol{q},\boldsymbol{p}},$$

with the old variables on the right subsequently replaced by their expressions in $(\boldsymbol{Q},\boldsymbol{P},t)$. This formula is an intermediate explanation. Convenient independent arguments for the generating function will make it much simpler.

4.2 Why a total derivative preserves the interior dynamics

Evaluate (12) along any path related by the coordinate map. Dividing by dt gives

$$\boldsymbol{p}^T \dot{\boldsymbol{q}} - H = \boldsymbol{P}^T \dot{\boldsymbol{Q}} - K + \frac{d\widehat{F}}{dt}.$$

Thus the two phase-space actions satisfy

$$\mathcal{S} = \mathcal{S}' + \widehat{F}\big|_{t_i}^{t_f}. \tag{13}$$

This identity holds for the related trial paths, not just after solving the equations. Under variations supported away from the endpoints, the last term has zero variation. Since the map is locally invertible, the interior stationarity conditions in the two descriptions are equivalent. Varying the new action gives Hamilton's equations with K.

This does not license imposing unrelated endpoint conditions in the two descriptions. For a boundary-value problem, the endpoint data and the boundary term must be transformed consistently. The safe conclusion is that the *interior* dynamics is unchanged, while the endpoint bookkeeping may change. This is the same distinction encountered when adding a total derivative to a Lagrangian in Lecture 5.

4.3 The first convenient representation

Suppose $(\boldsymbol{q}, \boldsymbol{Q}, t)$ are locally independent coordinates on the transformation. Express the primitive in those arguments and call it $F_1(\boldsymbol{q}, \boldsymbol{Q}, t)$. Rearranging (12) gives

$$dF_1 = \boldsymbol{p}^T d\boldsymbol{q} - \boldsymbol{P}^T d\boldsymbol{Q} + (K - H)dt.$$

Coefficient matching, with the displayed arguments independent, now yields

$$\boxed{\boldsymbol{p} = \frac{\partial F_1}{\partial \boldsymbol{q}}, \qquad \boldsymbol{P} = -\frac{\partial F_1}{\partial \boldsymbol{Q}}, \qquad K = H + \frac{\partial F_1}{\partial t}.} \tag{14}$$

The last derivative holds $\boldsymbol{q}, \boldsymbol{Q}$ fixed, not $\boldsymbol{q}, \boldsymbol{p}$. Its simplicity is consistent with (11): changing the independent arguments incorporates the coordinate-grid term. For a time-independent map we can choose a time-independent primitive, and then K is just H rewritten in the new variables.

Rather than catalog more representations immediately, we now use this one to simplify a familiar motion.

5 The oscillator: choose coordinates adapted to the orbits

5.1 Recover the Hamiltonian and identify the target

Use normalized oscillator variables with unit mass and unit frequency:

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

The normalization fixes the time and phase-space scales used in the following formulas. It is not a claim that position and momentum have identical physical units in an arbitrary system.

The Legendre transformation is explicit:

$$p = \frac{\partial L}{\partial \dot{q}} = \dot{q}, \qquad H = p\dot{q} - L = \frac{1}{2}(p^2 + q^2).$$

Hamilton's equations are $\dot{q} = p$, $\dot{p} = -q$, and the trajectories satisfy

$$p^2 + q^2 = 2E.$$

The circles suggest a new momentum P that labels the orbit and a new coordinate Q that measures phase:

$$\boxed{P = \frac{1}{2}(p^2 + q^2), \qquad q = \sqrt{2P}\sin Q, \qquad p = \sqrt{2P}\cos Q.} \tag{15}$$

The phase is measured from the positive p axis toward positive q, consistent with the direction of Hamiltonian motion. Locally one may write $Q = \tan^{-1}(q/p)$, with the branch and quadrant chosen continuously. The inverse sine and cosine relations in (15) are the more reliable definition of the phase branch. We exclude the origin, where $P = 0$ and phase is undefined.

Choosing a radius and angle is not sufficient to establish canonicality. The conjugate momentum must have the right dependence on radius. Direct differentiation gives

$$\frac{\partial Q}{\partial q} = \frac{p}{p^2+q^2}, \qquad \frac{\partial Q}{\partial p} = -\frac{q}{p^2+q^2}, \qquad \frac{\partial P}{\partial q} = q, \qquad \frac{\partial P}{\partial p} = p.$$

Therefore

$$\{Q, P\}_{q,p} = \frac{p^2+q^2}{p^2+q^2} = 1.$$

By contrast, the amplitude $R = \sqrt{p^2+q^2}$ would give $\{Q, R\} = 1/R$, not 1. The energy-like variable $P = R^2/2$, not the amplitude itself, is conjugate to this phase in the chosen units.

The map has no explicit time dependence, so

$$\boxed{K(Q,P) = P, \qquad \dot{P} = 0, \qquad \dot{Q} = 1.} \tag{16}$$

Two oscillating components have become a constant orbit label and a uniformly advancing phase. We have not changed the physical trajectory; we have changed how it is described.

5.2 Construct the same map from a scalar

On a patch where (q, Q) are independent, (15) gives

$$p = q\cot Q, \qquad P = \frac{1}{2}q^2\csc^2 Q.$$

The first relation tells us how to integrate the first equation in (14):

$$\frac{\partial F_1}{\partial q} = q\cot Q \quad \Longrightarrow \quad F_1(q,Q) = \frac{1}{2}q^2\cot Q + f(Q).$$

The second equation must also hold. It gives

$$-\frac{\partial F_1}{\partial Q} = \frac{1}{2}q^2\csc^2 Q - \frac{df}{dQ} = P,$$

so $df/dQ = 0$. Up to an additive constant,

$$\boxed{F_1(q,Q) = \frac{1}{2}q^2\cot Q.} \tag{17}$$

Integrating one derivative was not the whole construction: checking the other derivative fixed the remaining function. This is a useful habit whenever a generating function is obtained by partial integration.

Because F_1 is time independent, it produces $K = H = P$, just as the direct bracket calculation did. Conversely, beginning with (17) and differentiating constructs the transformation without separately guessing two unrelated coordinate formulas.

There is a representation limitation here beyond the usual angular branch choice. At $q = 0$, Q is 0 or π, regardless of the nonzero oscillation amplitude. The pair (q, Q) cannot then distinguish nearby states with different amplitudes. Accordingly $\cot Q$ becomes singular in this representation. This does not mean that the oscillator flow or the canonical phase coordinate is singular at every crossing of $q = 0$. It means that this particular choice of independent arguments for F_1 has failed. Other patches or representations are needed there.

6 A moving phase coordinate makes the new Hamiltonian vanish

The first transformation leaves a simple motion: $Q(t) = Q_0 + t$. Remove it by defining

$$\bar{Q} = Q - t, \qquad \bar{P} = P.$$

To construct this second map, the convenient independent arguments are $(Q, \bar{P}, t)$, not $(Q, \bar{Q}, t)$. Indeed, $\bar{Q} = Q - t$ makes the latter pair dependent at fixed time.

Let f denote the primitive for this second transformation. The extended relation is

$$df = P\,dQ - \bar{P}\,d\bar{Q} + (\bar{K} - K)dt.$$

Adding the product differential $d(\bar{Q}\bar{P})$ trades $d\bar{Q}$ for $d\bar{P}$. Define $G_2 = f + \bar{Q}\bar{P}$, so that

$$dG_2 = P\,dQ + \bar{Q}\,d\bar{P} + (\bar{K} - K)dt. \tag{18}$$

This is called a type-2 representation. The product rule, rather than a memorized sign convention, explains its coefficients.

Choose

$$\boxed{G_2(Q, \bar{P}, t) = (Q - t)\bar{P}.} \tag{19}$$

Its derivatives give

$$P = \frac{\partial G_2}{\partial Q} = \bar{P}, \qquad \bar{Q} = \frac{\partial G_2}{\partial \bar{P}} = Q - t, \qquad \frac{\partial G_2}{\partial t} = -\bar{P}.$$

Consequently,

$$\boxed{\bar{K} = K + \frac{\partial G_2}{\partial t} = \bar{P} - \bar{P} = 0.} \tag{20}$$

Hamilton's equations now give $\dot{\bar{Q}} = \dot{\bar{P}} = 0$.

6.1 Where did the motion go?

The inverse map still contains time:

$$\boxed{q(t) = \sqrt{2\bar{P}}\sin(\bar{Q} + t), \qquad p(t) = \sqrt{2\bar{P}}\cos(\bar{Q} + t).} \tag{21}$$

Differentiating reproduces $\dot{q} = p$ and $\dot{p} = -q$. Given (q_0, p_0) at time t_0, the constants are

$$\bar{P} = \frac{1}{2}(p_0^2 + q_0^2), \qquad \bar{Q} = Q_0 - t_0,$$

where $\sin Q_0 = q_0/\sqrt{2\bar{P}}$ and $\cos Q_0 = p_0/\sqrt{2\bar{P}}$ on the selected branch. Thus the two constants retain the full initial state for a nonzero orbit.

The equation $\bar{K} = 0$ does *not* say that the oscillator has zero physical energy. Its original energy is still $H = \bar{P}$. A Hamiltonian generates motion in the coordinates being used; explicitly time-dependent coordinates add their own motion to that description. Here the coordinates follow the advancing phase exactly, so the represented state is constant. The time dependence has moved into the reconstruction formula.

The sequence is therefore

$$L \;\longrightarrow\; H = \tfrac{1}{2}(p^2 + q^2) \;\longrightarrow\; K = P \;\longrightarrow\; \bar{K} = 0.$$

The Legendre transformation and the two canonical transformations play different roles, but each uses a differential relation to choose useful independent variables. We have seen the destination before organizing the general representations.

7 Four representations from the product rule

7.1 Exchange conjugate arguments, not the underlying principle

Return to N degrees of freedom and

$$dF_1 = \boldsymbol{p}^T d\boldsymbol{q} - \boldsymbol{P}^T d\boldsymbol{Q} + (K - H)dt.$$

To replace $\boldsymbol{Q}$ by $\boldsymbol{P}$ as an independent argument, use

$$d(\boldsymbol{Q}^T\boldsymbol{P}) = \boldsymbol{P}^T d\boldsymbol{Q} + \boldsymbol{Q}^T d\boldsymbol{P}.$$

Thus

$$\begin{aligned} F_2 &= F_1 + \boldsymbol{Q}^T\boldsymbol{P}, \\ dF_2 &= \boldsymbol{p}^T d\boldsymbol{q} + \boldsymbol{Q}^T d\boldsymbol{P} + (K - H)dt. \end{aligned} \tag{22}$$

The derivatives of $F_2(\boldsymbol{q}, \boldsymbol{P}, t)$ give

$$\boxed{\boldsymbol{p} = \frac{\partial F_2}{\partial \boldsymbol{q}}, \qquad \boldsymbol{Q} = \frac{\partial F_2}{\partial \boldsymbol{P}}, \qquad K = H + \frac{\partial F_2}{\partial t}.} \tag{23}$$

This is a partial Legendre transformation. Wherever the corresponding variable change is possible, the $-\boldsymbol{P}^T d\boldsymbol{Q}$ coefficient has become $+\boldsymbol{Q}^T d\boldsymbol{P}$. The second oscillator transformation used exactly this exchange.

Similarly, replacing the old configuration by its momentum uses subtraction of $\boldsymbol{q}^T\boldsymbol{p}$:

$$\begin{aligned} F_3 &= F_1 - \boldsymbol{q}^T\boldsymbol{p}, \\ dF_3 &= -\boldsymbol{q}^T d\boldsymbol{p} - \boldsymbol{P}^T d\boldsymbol{Q} + (K - H)dt, \end{aligned} \tag{24}$$

so for $F_3(\boldsymbol{p}, \boldsymbol{Q}, t)$,

$$\boxed{\boldsymbol{q} = -\frac{\partial F_3}{\partial \boldsymbol{p}}, \qquad \boldsymbol{P} = -\frac{\partial F_3}{\partial \boldsymbol{Q}}, \qquad K = H + \frac{\partial F_3}{\partial t}.}$$

Doing both exchanges gives

$$\begin{aligned} F_4 &= F_1 + \boldsymbol{Q}^T\boldsymbol{P} - \boldsymbol{q}^T\boldsymbol{p}, \\ dF_4 &= -\boldsymbol{q}^T d\boldsymbol{p} + \boldsymbol{Q}^T d\boldsymbol{P} + (K-H)dt. \end{aligned} \tag{25}$$

For $F_4(\boldsymbol{p}, \boldsymbol{P}, t)$, the rules are

$$\boxed{\boldsymbol{q} = -\frac{\partial F_4}{\partial \boldsymbol{p}}, \qquad \boldsymbol{Q} = \frac{\partial F_4}{\partial \boldsymbol{P}}, \qquad K = H + \frac{\partial F_4}{\partial t}.}$$

These equations are meant on overlapping patches where the chosen arguments are independent. They do not imply that every map admits every representation everywhere.

7.2 What the time derivative and boundary term mean

For each type, $\partial F_i/\partial t$ holds *that type's independent arguments* fixed. After evaluating $H + \partial F_i/\partial t$, eliminate any old variables to obtain an actual function $K(\boldsymbol{Q}, \boldsymbol{P}, t)$. It is not enough to write the formula for K while leaving its arguments unresolved.

The generating functions of different types are not interchangeable as endpoint terms. For example, (22) implies

$$\boldsymbol{p}^T d\boldsymbol{q} - H dt = \boldsymbol{P}^T d\boldsymbol{Q} - K dt + d(F_2 - \boldsymbol{Q}^T\boldsymbol{P}). \tag{26}$$

Thus the primitive in the action difference is $F_2 - \boldsymbol{Q}^T\boldsymbol{P}$, not F_2 alone. This is the product term that was added to change independent arguments. The differential coefficients and the physical interior equations are consistent across representations, while their endpoint expressions look different.

Adding a function of time alone to a generating function changes K by its time derivative. Since that addition has no phase-space derivatives, it does not change the coordinate equations of motion. This freedom is separate from the substantive time correction of a moving coordinate map.

7.3 A generating function must define a map, not just equations

Consider a proposed $F_2(\boldsymbol{q}, \boldsymbol{P}, t)$. Given an old state $(\boldsymbol{q}, \boldsymbol{p})$, the first equation in (23) must be solved for $\boldsymbol{P}$. A sufficient local condition is

$$\boxed{\det\left(\frac{\partial^2 F_2}{\partial \boldsymbol{q}\, \partial \boldsymbol{P}^T}\right) \neq 0.} \tag{27}$$

The implicit-function theorem then supplies $\boldsymbol{P}(\boldsymbol{q}, \boldsymbol{p}, t)$, and the second equation supplies $\boldsymbol{Q}$. Applying an exterior derivative to (22) at fixed time, or using commuting mixed partial derivatives directly, gives preservation of the conjugate two-form. Thus a regular generating function constructs a canonical map, rather than merely parametrizing a set of possibly ambiguous relations.

An analogous condition for F_1 uses its mixed $(\boldsymbol{q}, \boldsymbol{Q})$ derivatives. The oscillator's (q, Q) patch failure is an example of why these conditions belong to the construction. The existence of a local primitive on phase space does not guarantee the usability of every proposed argument list.

8 Checks that make the rules operational

8.1 Identity and reciprocal scaling

The identity transformation is generated by

$$F_2(\boldsymbol{q}, \boldsymbol{P}) = \boldsymbol{q}^T \boldsymbol{P}.$$

Differentiation gives $\boldsymbol{p} = \boldsymbol{P}$, $\boldsymbol{Q} = \boldsymbol{q}$; the mixed derivative matrix is the identity. This example is especially instructive because an identity map cannot use $(\boldsymbol{q}, \boldsymbol{Q})$ as independent arguments. A perfectly good map may require a different type.

In one degree of freedom take a nonzero constant a and

$$F_2(q, P) = aqP.$$

Then $p = aP$, $Q = aq$, hence

$$\boxed{Q = aq, \qquad P = p/a.} \tag{28}$$

The reciprocal momentum scaling is forced by the scalar derivatives. For the normalized oscillator, inverse substitution gives

$$K(Q, P) = \frac{1}{2}\left(a^2 P^2 + \frac{Q^2}{a^2}\right).$$

The new equations are $\dot{Q} = a^2 P$, $\dot{P} = -Q/a^2$. Substituting $Q = aq$, $P = p/a$ recovers the original oscillator equations. Canonicality preserves the Hamiltonian *form* of the equations, not necessarily the formula of the Hamiltonian.

8.2 Recover the earlier momentum shear

Lecture 17 used $Q = q$, $P = p + cq$, with constant c. Its type-2 generating function is

$$F_2(q, P) = qP - \frac{1}{2}cq^2.$$

Indeed, $p = P - cq$ and $Q = q$. The boundary primitive from (26) becomes

$$F_2 - QP = -\frac{1}{2}cq^2,$$

exactly the scalar whose differential was found in Lecture 17. This reconciles the primitive obtained by comparing one-forms with a generating function whose arguments actually define the map.

8.3 An optional sign check with a translating coordinate

A free particle has $H = p^2/(2m)$. Let v be a constant and choose

$$F_2(q, P, t) = (q - vt)P.$$

Then $Q = q - vt$, $P = p$, and

$$\boxed{K(Q, P, t) = \frac{P^2}{2m} - vP.}$$

Its equations give $\dot{Q} = P/m - v$, $\dot{P} = 0$. Restoring $q = Q + vt$ gives $\dot{q} = p/m$, as required. Omitting the time derivative would incorrectly assign the laboratory velocity to the translating coordinate.

Here P is still the laboratory momentum; it is not $m\dot{Q}$. We changed the position coordinate without adding a momentum shift. This check illustrates the time correction and does not require a new mechanical force or a change in the physical motion.

9 The infinitesimal limit recovers Poisson-bracket generators

Choose the near-identity type-2 function

$$F_2(\boldsymbol{q}, \boldsymbol{P}) = \boldsymbol{q}^T \boldsymbol{P} + \epsilon G(\boldsymbol{q}, \boldsymbol{P}), \tag{29}$$

with ϵG having units of action. For sufficiently small ϵ, suppose its mixed derivative matrix remains invertible. The exact implicit relations are

$$\boldsymbol{p} = \boldsymbol{P} + \epsilon \frac{\partial G}{\partial \boldsymbol{q}}, \qquad \boldsymbol{Q} = \boldsymbol{q} + \epsilon \frac{\partial G}{\partial \boldsymbol{P}}.$$

Because $\boldsymbol{P} = \boldsymbol{p} + O(\epsilon)$, replacing $\boldsymbol{P}$ by $\boldsymbol{p}$ inside a term already multiplied by ϵ changes that term only at second order. Therefore

$$\begin{aligned} \boldsymbol{Q} - \boldsymbol{q} &= \epsilon \frac{\partial G(\boldsymbol{q}, \boldsymbol{p})}{\partial \boldsymbol{p}} + O(\epsilon^2), \\ \boldsymbol{P} - \boldsymbol{p} &= -\epsilon \frac{\partial G(\boldsymbol{q}, \boldsymbol{p})}{\partial \boldsymbol{q}} + O(\epsilon^2). \end{aligned} \tag{30}$$

Taylor-expand a scalar phase-space function A between a point and its mapped point, using the same function for this comparison:

$$\begin{aligned} \delta_G A &= \left(\frac{\partial A}{\partial \boldsymbol{q}}\right)^T (\boldsymbol{Q} - \boldsymbol{q}) + \left(\frac{\partial A}{\partial \boldsymbol{p}}\right)^T (\boldsymbol{P} - \boldsymbol{p}) + O(\epsilon^2) \\ &= \boxed{\epsilon\{A, G\} + O(\epsilon^2).} \end{aligned} \tag{31}$$

This is the specified infinitesimal change generated by G, not an arbitrary virtual path variation. If the map is instead used as a passive relabeling of the same state, the observable's coordinate formula is transformed by inverse substitution; the point comparison above fixes the convention behind the displayed sign.

For example, $G = P$ in one degree of freedom gives $F_2 = qP + \epsilon P$, hence $Q = q + \epsilon$, $P = p$. Momentum generates translation with the expected sign. In general, the finite transformation defined implicitly by (29) need not be the exact finite-parameter flow of G. Their first-order generators agree; higher-order terms require separate attention.

10 The next question: choose zero Hamiltonian directly

For the oscillator we recognized the phase-plane geometry, found coordinates with $K = P$, and then removed the uniform phase advance. Could we work backward from the desired endpoint $K = 0$ for a general Hamiltonian?

Let a type-2 generating function be denoted $S(\boldsymbol{q}, \boldsymbol{P}, t)$. Its rules give

$$\boldsymbol{p} = \frac{\partial S}{\partial \boldsymbol{q}}, \qquad K = H\left(\boldsymbol{q}, \frac{\partial S}{\partial \boldsymbol{q}}, t\right) + \frac{\partial S}{\partial t}.$$

Requiring $K = 0$ therefore asks us to solve

$$\boxed{H\left(\boldsymbol{q}, \frac{\partial S}{\partial \boldsymbol{q}}, t\right) + \frac{\partial S}{\partial t} = 0.} \tag{32}$$

This is the Hamilton–Jacobi equation. If a suitable parameterized solution defines a regular canonical map, the new equations say

$$\dot{\boldsymbol{Q}} = \boldsymbol{0}, \qquad \dot{\boldsymbol{P}} = \boldsymbol{0}.$$

As in the oscillator, constants in the new coordinates would reconstruct nontrivial motion through a time-dependent inverse map. Finding that scalar may be difficult; merely writing the PDE has not solved the mechanics.

11 Summary and handoff

The construction began with a comparison of line elements, not with a table of derivative rules. Canonicality made their difference have zero local circulation. Green's theorem and path independence then supplied a scalar primitive. Restoring time gave the extended identity

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

Different independent arguments for that scalar, related by product-rule exchanges of conjugate variables, lead to the familiar generating-function types. Each requires a valid local coordinate patch, and each time derivative must hold its own independent arguments fixed.

The oscillator showed why the machinery is useful: F_1 constructs an orbit label and phase, and a time-dependent type-2 function turns the remaining uniform phase advance into constants. Zero transformed Hamiltonian means that the chosen coordinates absorb the motion, not that physical energy vanishes. The infinitesimal limit reconnects this finite construction with the Poisson-bracket generators of Lecture 16.

Lecture 19 develops the Hamilton–Jacobi equation as a method: the endpoint meaning of S, the parameters needed for a complete integral, and the recovery of trajectories by differentiation with respect to those parameters. We have not yet solved a Hamilton–Jacobi problem here. The systematic treatment of canonical variables adapted to recurrent motion remains a later step.