Classical Mechanics

LECTURE 11

Hamiltonian vector fields

Given any function $f = f(\mathbf{y})$ on phase space, we define the following differential operator

$$\mathcal{D}_f \equiv \sum_{\alpha,\beta=1}^{2n} \frac{\partial f}{\partial y_\alpha} \Omega_{\alpha\beta} \frac{\partial}{\partial y_\beta} . \tag{5.77}$$

This maps functions on phase space to functions on phase space in the obvious way. In fact from (5.46) notice that if we apply $\mathcal{D}_f$ to a function $g = g(\mathbf{y})$ we simply obtain the Poisson bracket between f and g:

$$\mathcal{D}_f\, g = \{f, g\} . \tag{5.78}$$

We shall refer to $\mathcal{D}_f$ as the *Hamiltonian vector field associated to* f.[15] One can deduce some immediate properties:

- $\mathcal{D}_f$ is linear in f: $\mathcal{D}_{f_1+f_2} = \mathcal{D}_{f_1} + \mathcal{D}_{f_2}$,
- $\mathcal{D}_f\, g$ is linear in g: $\mathcal{D}_f\,(g_1 + g_2) = \mathcal{D}_f\, g_1 + \mathcal{D}_f\, g_2$,
- anti-symmetry of the Poisson bracket and (5.78) implies $\mathcal{D}_f\, g = -\mathcal{D}_g\, f$.

Another property of $\mathcal{D}_f$ is that it is invariant under canonical transformations; that is, for all f and g we have

$$(\mathcal{D}_f)_y\, g \equiv \{f, g\}_y = \{f, g\}_Y \equiv (\mathcal{D}_f)_Y\, g . \tag{5.79}$$

Said more simply, in (5.77) it doesn't matter whether we evaluate the partial derivatives with respect to $\mathbf{y}$ or $\mathbf{Y}$, provided they are related by a canonical transformation. Given two Hamiltonian vector fields $\mathcal{D}_f$, $\mathcal{D}_g$, we may consider their *commutator*

$$[\mathcal{D}_f, \mathcal{D}_g] \equiv \mathcal{D}_f\mathcal{D}_g - \mathcal{D}_g\mathcal{D}_f . \tag{5.80}$$

The composition of operators on the right hand side works in the obvious way, *i.e.* applied to another function $h = h(\mathbf{y})$ we have

$$(\mathcal{D}_f\mathcal{D}_g)(h) = \mathcal{D}_f(\mathcal{D}_g\, h) . \tag{5.81}$$

[15] This might seem like an odd name, because so far this seems to have nothing to do with either the Hamiltonian H or vector fields. The relation to the Hamiltonian and Hamilton's equations will become clear below. The term *vector field* is often used for (or conflated with) a *directional derivative*, since one can write $\mathcal{D}_f = \sum_{\alpha=1}^{2n} V_\alpha \partial_{y_\alpha}$, where V_α is a vector field on phase space.

Then a calculation shows that (acting on twice differentiable functions on phase space)

$$[\mathcal{D}_f, \mathcal{D}_g] \;\; = \;\; \mathcal{D}_{\{f,g\}} \; . \tag{5.82}$$

That is, the commutator of Hamiltonian vector fields equals the Hamiltonian vector field of the Poisson bracket. To prove (5.82) one simply applies both sides to an arbitrary function $h = h(\mathbf{y})$: on the left hand side there are *a priori* terms involving second derivatives of h, but one checks that these cancel, and the remaining terms on both sides can be rerranged and shown to be equal – we leave this as an exercise on Problem Sheet 4. The relation (5.82) implies that if f and g Poisson commute, then their Hamiltonian vector fields commute as differential operators.

Armed with this new technology, let us fix any function $f = f(\mathbf{y})$ and consider

$$\mathbf{Y}(\mathbf{y}, s) \;\; \equiv \;\; \mathrm{e}^{-s\mathcal{D}_f}\, \mathbf{y} \;\; \equiv \;\; \sum_{n=0}^{\infty} \frac{(-s)^n}{n!} (\mathcal{D}_f)^n \, \mathbf{y} \; . \tag{5.83}$$

The operator $(\mathcal{D}_f)^n$ is understood to act as in (5.81), and we shall assume that everything is suitably differentiable/convergent for the right hand side of (5.83) to make sense. We claim that $\mathbf{Y}(\mathbf{y}, s)$ defines a one-parameter family of canonical transformations. To see this, note first that clearly $\mathbf{Y}(\mathbf{y}, 0) = \mathbf{y}$, giving the identity transformation on the coordinates $\mathbf{y}$. We next compute

$$\frac{\partial}{\partial s}\mathbf{Y}(\mathbf{y}, s) \;\; = \;\; -\sum_{n=1}^{\infty} \frac{(-s)^{n-1}}{(n-1)!} (\mathcal{D}_f)^n \, \mathbf{y} \;\; = \;\; -\mathcal{D}_f \, \mathbf{Y}(\mathbf{y}, s) \;\; = \;\; \{\mathbf{Y}(\mathbf{y}, s), f(\mathbf{y})\} \; . \tag{5.84}$$

The last equality here is simply (5.78). Writing $\mathbf{Y} = \mathbf{Y}(\mathbf{y}, s)$ we may hence compute

$$\begin{aligned} \frac{\partial}{\partial s}\{Y_\alpha, Y_\beta\} \;\; &= \;\; \left\{\frac{\partial}{\partial s} Y_\alpha, Y_\beta\right\} + \left\{Y_\alpha, \frac{\partial}{\partial s} Y_\beta\right\} \\ &= \;\; \{\{Y_\alpha, f\}, Y_\beta\} + \{Y_\alpha, \{Y_\beta, f\}\} \\ &= \;\; \{\{Y_\alpha, Y_\beta\}, f\} \; . \end{aligned} \tag{5.85}$$

Here we have used (5.84) in the second equality, and the Jacobi identity in the last equality. Notice that this equation takes the same form as (5.84). One can at this point turn things around, and regard (5.84) as a first order differential equation in s, whose solution is (5.83). We thus deduce from (5.85) that

$$\{Y_\alpha(\mathbf{y}, s), Y_\beta(\mathbf{y}, s)\} \;\; = \;\; \mathrm{e}^{-s\mathcal{D}_f}\{Y_\alpha(\mathbf{y}, 0), Y_\beta(\mathbf{y}, 0)\} \;\; = \;\; \mathrm{e}^{-s\mathcal{D}_f}\, \Omega_{\alpha\beta} \;\; = \;\; \Omega_{\alpha\beta} \; . \tag{5.86}$$

Here the second equality follows from $\mathbf{Y}(\mathbf{y}, 0) = \mathbf{y}$ and the canonical Poisson brackets, while the last equality follows since $\Omega_{\alpha\beta}$ is *constant* – hence $\mathcal{D}_f \, \Omega_{\alpha\beta} = 0$. Of course (5.86) is precisely what we wanted to prove: $\mathbf{Y}(\mathbf{y}, s)$ defines a canonical transformation for all $s \in \mathbb{R}$.

It follows that any function $f = f(\mathbf{y})$ on phase space generates an associated one-parameter family of canonical transformations. The simplest example would be to choose f constant, but clearly then $\mathcal{D}_f = 0$ and $\mathbf{Y}(\mathbf{y}, s) = \mathbf{y}$ for all s (a trivial one-parameter family). The next simplest example is to consider f linear. Let's focus on $n = 1$ degree of freedom, and rewrite the Hamiltonian vector field in terms of the canonically conjugate q and p:

$$\mathcal{D}_f \;=\; \frac{\partial f}{\partial q}\frac{\partial}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial}{\partial q} \; . \tag{5.87}$$

Example: Let $f(q,p) = \lambda p - \mu q$, with λ, μ constant. Then we compute $\mathcal{D}_f = -\lambda\partial_q - \mu\partial_p$, and hence

$$\begin{aligned} Q(q,p,s) &\equiv \mathrm{e}^{-s\mathcal{D}_f}\, q \;=\; q + \lambda s \; , \\ P(q,p,s) &\equiv \mathrm{e}^{-s\mathcal{D}_f}\, p \;=\; p + \mu s \; . \end{aligned} \tag{5.88}$$

Notice here that the power series terminate after the first two terms. The resulting linear shifts in coordinates are clearly canonical transformations, albeit not very interesting ones.

Another interesting application of this formalism is to Hamilton's equations (5.29). These may be written

$$\dot{\mathbf{y}} \;=\; \{\mathbf{y}, H\} \;=\; -\mathcal{D}_H\, \mathbf{y} \; . \tag{5.89}$$

Here the function $f = H = H(\mathbf{y})$ is the Hamiltonian, which we assume here doesn't depend on time t. But (5.89) is just equation (5.84), with the variable s replaced by time t. We thus deduce that

$$\mathbf{y}(t) \;=\; \mathrm{e}^{-t\mathcal{D}_H}\, \mathbf{y} \tag{5.90}$$

solves Hamilton's equations, with initial condition $\mathbf{y}(0) = \mathbf{y}$. In this sense, the Hamiltonian vector field $\mathcal{D}_H$ generates the solution to Hamilton's equations. An immediate corollary of our results is that Hamiltonian flow, where we regard $\mathbf{y} = \mathbf{y}(0) \to \mathbf{y}(t)$ as a one-parameter family of coordinate transformations on phase space parametrized by time t, is in fact a family of *canonical transformations*!

Example: As a more interesting example, consider the one-dimensional harmonic oscillator with Hamiltonian $H(q,p) = \frac{1}{2}(p^2 + q^2)$. The Hamiltonian vector field is $\mathcal{D}_H = q\partial_p - p\partial_q$, and one computes

$$\begin{aligned} Q(q,p,t) &\equiv \mathrm{e}^{-t\mathcal{D}_H}\, q \;=\; q\cos t + p\sin t \; , \\ P(q,p,t) &\equiv \mathrm{e}^{-t\mathcal{D}_H}\, p \;=\; p\cos t - q\sin t \; . \end{aligned} \tag{5.91}$$

Of course these indeed solve Hamilton's equations $\partial_t Q = P$, $\partial_t P = -Q$, with initial condition $Q(q,p,0) = q$, $P(q,p,0) = p$.

The principle of least action

A different approach, that also involves a notion of generating functions, arises naturally if we think about the principle of least action in the Hamiltonian formalism. Recall that the action (2.6) is

$$S = \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t)\mathrm{d}t = \int_{t_1}^{t_2} \left(\sum_{a=1}^{n} p_a \dot{q}_a - H \right) \mathrm{d}t = \int_{t_1}^{t_2} \left(\sum_{a=1}^{n} p_a \mathrm{d}q_a - H\mathrm{d}t \right) , \quad (5.92)$$

where we have written the Lagrangian in terms of the Hamiltonian. In fact Hamilton's equations follow from extremizing (5.92) with respect to both $\mathbf{q}$ *and* $\mathbf{p}$, *i.e.* we treat them as independent variables, with $S = S[\mathbf{q}(t), \mathbf{p}(t)]$. Indeed, writing $\delta\mathbf{q}(t) = \epsilon\,\mathbf{u}(t)$, $\delta\mathbf{p}(t) = \epsilon\,\mathbf{w}(t)$, as in our derivation of Lagrange's equations we compute the first order variation

$$\begin{aligned} \delta S &= \epsilon \sum_{a=1}^{n} \int_{t_1}^{t_2} \left(w_a \mathrm{d}q_a + p_a \mathrm{d}u_a - \frac{\partial H}{\partial q_a} u_a \mathrm{d}t - \frac{\partial H}{\partial p_a} w_a \mathrm{d}t \right) \\ &= \epsilon \sum_{a=1}^{n} \int_{t_1}^{t_2} \left[w_a \left(\mathrm{d}q_a - \frac{\partial H}{\partial p_a}\mathrm{d}t \right) - u_a \left(\mathrm{d}p_a + \frac{\partial H}{\partial q_a}\mathrm{d}t \right) \right] . \quad (5.93) \end{aligned}$$

Here we have integrated by parts the second term on the right hand side of the first line, and used the boundary condition $\mathbf{u}(t_1) = \mathbf{u}(t_2) = \mathbf{0}$. Hamilton's equations hence follow, much as we deduced Lagrange's equations.

Also as in our discussion of Lagrange's equations (see around equation (2.17)), adding a total time derivative to the integrand in (5.92) does not affect the equations of motion, as such a term depends only on the boundary data and so has zero variation. If we have two coordinate systems $(\mathbf{q}, \mathbf{p})$ and $(\mathbf{Q}, \mathbf{P})$ on phase space, with corresponding Hamiltonian's H and K, respectively, then our above discussion means that these will lead to the *same* equations of motion if

$$\sum_{a=1}^{n} P_a \mathrm{d}Q_a - K\mathrm{d}t = \sum_{a=1}^{n} p_a \mathrm{d}q_a - H\mathrm{d}t - \mathrm{d}F_1(\mathbf{q}, \mathbf{Q}, t) . \quad (5.94)$$

The function $F_1(\mathbf{q}, \mathbf{Q}, t)$ is called a *generating function of the first kind*. It generates the change of coordinates in the sense that (5.94) implies the relations

$$\mathbf{p} = \frac{\partial F_1}{\partial \mathbf{q}} , \qquad \mathbf{P} = -\frac{\partial F_1}{\partial \mathbf{Q}} , \qquad K = H + \frac{\partial F_1}{\partial t} . \quad (5.95)$$

One must be very careful to read the dependences correctly in these equations. The first equation in (5.95) determines $\mathbf{p}$ as a function of $(\mathbf{q}, \mathbf{Q}, t)$. One then inverts this (assuming it is invertible) to find $\mathbf{Q}$ as a function of $(\mathbf{q}, \mathbf{p}, t)$. Similarly, the second equation in (5.95) determines $\mathbf{P}$ as a function of $(\mathbf{q}, \mathbf{Q}, t)$, which we may then in turn convert to a function of $(\mathbf{q}, \mathbf{p}, t)$. Notice that the final equation in (5.95) relates the two Hamiltonians, and may be compared to equation (5.71) – although in the latter equation everything was a function of the new variables $(\mathbf{Q}, \mathbf{P})$, and we were careful to distinguish $H(\mathbf{q}, \mathbf{p}, t)$ from $\tilde{H}(\mathbf{Q}, \mathbf{P}, t)$. In this section we are being less precise, using the same name for a function and the function composed with a change of variable. This makes the notation cleaner and easier to read, but one must then be particularly careful when applying the chain rule. Notice that if the generating function is independent of time, so $\partial F_1/\partial t = 0$, then $K = H$ and the new Hamiltonian is obtained by simply substituting for $\mathbf{q}, \mathbf{p}$ in H their values in terms of the new variables $\mathbf{Q}, \mathbf{P}$ (so that more precisely $K = \tilde{H}$, as in (5.69)).

By construction, any two Hamiltonian systems related by the change of variables (5.95) will lead to the same Hamilton equations of motion, and be a canonical transformation. One can verify this directly by a brute force computation of Poisson brackets, being *very* careful with the chain rule, although this is rather long and not very enlightening. A more conceptually interesting way to think about the invariance of the Poisson brackets $\{f, g\}_{p,q} = \{f, g\}_{P,Q}$ is as follows. Since time t simply appears as a parameter in the Poisson bracket computation, we might as well consider only time-independent functions. We may then regard the arbitrary function g appearing in $\{f, g\}$ as the Hamiltonian of a ficticious dynamical system, so that via Hamilton's equations for this system $\{f, g\} = \frac{\mathrm{d}f}{\mathrm{d}t}$. Since Hamilton's equations in the two coordinate systems are the same, the time-evolution of the function f must be coordinate-independent, and we deduce that $\{f, g\}_{p,q} = \{f, g\}_{P,Q}$.

Example: The generating function $F_1 = \sum_{a=1}^{n} q_a Q_a$ generates the position and momentum swap $\mathbf{Q} = \mathbf{p}$, $\mathbf{P} = -\mathbf{q}$.

In (5.94) the generating function $F_1 = F_1(\mathbf{q}, \mathbf{Q}, t)$ is a function of the old and new generalized coordinates. However, it might be more convenient to express the generating function in terms of the old coordinate $\mathbf{q}$ and new *momenta* $\mathbf{P}$. The appropriate formulae are obtained via a Legendre transform. That is, we rewrite (5.94) as

$$\mathrm{d}F_1 \;=\; \sum_{a=1}^{n} p_a \mathrm{d}q_a - \sum_{a=1}^{n} P_a \mathrm{d}Q_a + (K - H)\mathrm{d}t \;, \tag{5.96}$$

which in turn may be written

$$\mathrm{d}(F_1 + \sum_{a=1}^{n} P_a Q_a) \;=\; \sum_{a=1}^{n} p_a \mathrm{d}q_a + \sum_{a=1}^{n} Q_a \mathrm{d}P_a + (K - H)\mathrm{d}t \;, \tag{5.97}$$

The argument of the differential on the left hand side may then be regarded as $F_2 = F_2(\mathbf{q}, \mathbf{P}, t) = F_1 + \sum_{a=1}^{n} P_a Q_a$. Up to an overall sign, this is a Legendre transform of $F_1 = F_1(\mathbf{q}, \mathbf{Q}, t)$ with respect to $\mathbf{Q}$, since $\partial F_1/\partial \mathbf{Q} = -\mathbf{P}$. From (5.97) we obtain the Legendre-transformed relations

$$\mathbf{p} \;=\; \frac{\partial F_2}{\partial \mathbf{q}} \;, \qquad \mathbf{Q} \;=\; \frac{\partial F_2}{\partial \mathbf{P}} \;, \qquad K \;=\; H + \frac{\partial F_2}{\partial t}. \tag{5.98}$$

We may similarly obtain formulae for generating functions which depend on $\mathbf{p}$ and $\mathbf{Q}$, or $\mathbf{p}$ and $\mathbf{P}$, respectively, which are generating functions of the third and fourth kind, respectively.

Example: As a less trivial example, consider the generating function

$$F_2(q,P) \;=\; \int^q \sqrt{2P-x^2}\,\mathrm{d}x\ . \tag{5.99}$$

It follows from (5.98) that

$$p \;=\; \frac{\partial F_2}{\partial q} \;=\; \sqrt{2P-q^2}\ , \qquad Q \;=\; \frac{\partial F_2}{\partial P} \;=\; \arctan\frac{q}{\sqrt{2P-q^2}}\ . \tag{5.100}$$

The coordinate transformation is obtained by inverting these appropriately. For example, we may write q,p in terms of Q,P as

$$q \;=\; \sqrt{2P}\sin Q\ , \qquad p \;=\; \sqrt{2P}\cos Q\ . \tag{5.101}$$

The new variables Q,P are essentially polar coordinates on the phase space $\mathscr{P}=\mathbb{R}^2$. Since F_2 is time-independent, the new Hamiltonian is simply $K=H$, written in terms of the new variables. For example, the harmonic oscillator Hamiltonian $H(q,p)=\frac{1}{2}(p^2+q^2)$ becomes simply $H(Q,P)=P$ in the new variables! In particular Q is an ignorable coordinate, leading to $\dot{P}=0$, while $\dot{Q}=\partial H/\partial P=1$. Of course this leads to the usual trigonometric solutions $q=\sqrt{2P}\sin(t-t_0)$ in the original variables, but the point is that the transformation essentially trivialises the Hamiltonian system.