MATHEMATICS 0054 (Analytical Dynamics)
YEAR 2021–2022, TERM 2

HANDOUT #12: THE HAMILTONIAN APPROACH TO MECHANICS

These notes are intended to be read as a supplement to the handout from Gregory, *Classical Mechanics*, Chapter 14.

1 The basic set-up

I assume that you have already studied Gregory, Sections 14.1–14.4. The following is intended only as a succinct summary.

We are considering a system whose equations of motion are written in **Hamiltonian form**. This means that:

1. The **phase space** of the system is parametrized by **canonical coordinates** $\boldsymbol{q} = (q_1, \ldots, q_n)$ and $\boldsymbol{p} = (p_1, \ldots, p_n)$.
2. We are given a **Hamiltonian function** $H(\boldsymbol{q}, \boldsymbol{p}, t)$.
3. The dynamics of the system is given by **Hamilton's equations of motion**

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{1a}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{1b}$$

for $i = 1, \ldots, n$.

In these notes we will consider some deeper aspects of Hamiltonian dynamics.

2 Poisson brackets

Let us start by considering an arbitrary function $f(\boldsymbol{q}, \boldsymbol{p}, t)$. Then its time evolution is given by

$$\frac{df}{dt} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i \right) + \frac{\partial f}{\partial t} \tag{2a}$$

$$= \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right) + \frac{\partial f}{\partial t} \tag{2b}$$

where the first equality used the definition of total time derivative together with the chain rule, and the second equality used Hamilton's equations of motion.

The formula (2b) suggests that we make a more general definition. Let $f(\boldsymbol{q}, \boldsymbol{p}, t)$ and $g(\boldsymbol{q}, \boldsymbol{p}, t)$ be *any* two functions; we then define their **Poisson bracket** $\{f, g\}$ to be

$$\{f, g\} \stackrel{\text{def}}{=} \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right) . \tag{3}$$

The time-evolution equation (2) can then be rewritten in compact form as

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} . \tag{4}$$

But the importance of the Poisson bracket in Hamiltonian mechanics goes far beyond this.

> **Remark.** In the definition (3), the explicit time-dependence, if any, simply goes for the ride; the important thing is how f and g depend on $\boldsymbol{q}$ and $\boldsymbol{p}$. So, in discussing Poisson brackets, we shall often just consider functions $f(\boldsymbol{q}, \boldsymbol{p})$ and $g(\boldsymbol{q}, \boldsymbol{p})$ and not bother to discuss explicit time-dependence.

Let us begin by recording some fundamental properties of the Poisson bracket:

1. **Bilinearity.** We have

$$\{\alpha_1 f_1 + \alpha_2 f_2, g\} = \alpha_1 \{f_1, g\} + \alpha_2 \{f_2, g\} \tag{5}$$

   and likewise for g.

2. **Anticommutativity.** We have

$$\{f, g\} = -\{g, f\} . \tag{6}$$

   In particular it follows that $\{f, f\} = 0$ (why?).

3. **Jacobi identity.** For any three functions f, g, h we have

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0 \tag{7}$$

   or equivalently (using anticommutativity)

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0 \tag{8}$$

   We will prove the Jacobi identity in the next section.

4. **Product identity.** For any three functions f, g, h we have

$$\{fg, h\} = f\{g, h\} + g\{f, h\} . \tag{9}$$

   This is an easy consequence of the product law for partial differentation; you will be asked to prove it in the next problem set. It basically expresses the fact that the Poisson bracket $\{f, g\}$ involves *first* derivatives of f and of g.

5. **Fundamental Poisson brackets.** The Poisson brackets among the canonical coordinates $\boldsymbol{q} = (q_1, \ldots, q_n)$ and $\boldsymbol{p} = (p_1, \ldots, p_n)$ are

$$\{q_i, q_j\} \;=\; 0 \tag{10a}$$
$$\{p_i, p_j\} \;=\; 0 \tag{10b}$$
$$\{q_i, p_j\} \;=\; \delta_{ij} \tag{10c}$$

where δ_{ij} is the Kronecker delta, i.e.

$$\delta_{ij} \;=\; \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{11}$$

The three properties of bilinearity, anticommutativity and the Jacobi identity play such a fundamental role in many areas of mathematics that they have been given a name: an algebraic structure involving a "product" that is bilinear, anticommutative and satisfies the Jacobi identity is called a **Lie algebra**.[1] You already know two other examples of Lie algebras:

- Vectors in $\mathbb{R}^3$, equipped with the cross product $\mathbf{a} \times \mathbf{b}$.
- $n \times n$ matrices, equipped with the **commutator** $[A, B] = AB - BA$.

In both cases the bilinearity and anticommutativity are obvious; I leave it to you to check the Jacobi identity.

We can now prove an important result in Hamiltonian dynamics:

Total time derivative of a Poisson bracket. For any two functions $f(\boldsymbol{q}, \boldsymbol{p}, t)$ and $g(\boldsymbol{q}, \boldsymbol{p}, t)$, we have

$$\frac{d}{dt}\{f, g\} \;=\; \left\{\frac{df}{dt}, g\right\} \,+\, \left\{f, \frac{dg}{dt}\right\} . \tag{12}$$

Despite its fairly obvious-looking form, this formula is *not* obvious; it requires a bit of calculation.

PROOF OF (12). From the fundamental time-evolution equation (4) applied to $\{f, g\}$, we have

$$\frac{d}{dt}\{f, g\} \;=\; \{\{f, g\}, H\} \,+\, \frac{\partial}{\partial t}\{f, g\} . \tag{13}$$

[1] After the Norwegian mathematician Sophus Lie (1842–1899), who created the theory of continuous symmetry — what is now known as the theory of **Lie groups** and **Lie algebras** — and applied it to differential geometry and differential equations. These theories now play a central role in many areas of mathematics and theoretical physics.

The first term on the right-hand side can be transformed using the Jacobi identity and anticommutativity:

$$\{\{f,g\},H\} \;=\; -\{\{g,H\},f\} - \{\{H,f\},g\} \tag{14a}$$

$$\;=\; \{f,\{g,H\}\} + \{\{f,H\},g\}\;. \tag{14b}$$

And for the second term on the right-hand side, we use the fact that $\partial/\partial t$ commutes with the partial derivatives $\partial/\partial q_j$ and $\partial/\partial p_j$ occurring in the definition of the Poisson bracket; it therefore follows that

$$\frac{\partial}{\partial t}\{f,g\} \;=\; \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\} \tag{15}$$

(you should check the details!). Adding (14) and (15) and using the fundamental time-evolution equation (4) for f and for g, we obtain (12). $\square$

In particular, if f and g are constants of motion, then so is $\{f,g\}$. So this provides a method for obtaining new constants of motion, given old ones! Of course, these new constants of motion are not guaranteed to be nontrivial. (For instance, we might have $\{f,g\} = 0$.) But here is one nontrivial example:

Example: Angular momentum. Consider a single particle in Cartesian coordinates, so that $\boldsymbol{q} = (q_1, q_2, q_3)$ is the position and that $\boldsymbol{p} = (p_1, p_2, p_3)$ is the ordinary linear momentum. In the next problem set I will ask you to show that three components of the angular momentum $\mathbf{L} = \boldsymbol{q} \times \boldsymbol{p}$ have the Poisson brackets

$$\{L_1, L_2\} \;=\; L_3 \tag{16a}$$

$$\{L_2, L_3\} \;=\; L_1 \tag{16b}$$

$$\{L_3, L_1\} \;=\; L_2 \tag{16c}$$

It follows that if *two* components of the angular momentum happen to be constants of motion, then the third component of the angular momentum *must also* be a constant of motion.

Note, by contrast, that nothing of the kind follows if *only one* component of the angular momentum is a constant of motion. Indeed, we have seen lots of examples of systems where one component of angular momentum (e.g. the z component) is conserved but the other two are not.

3 A unified notation for phase space

The key idea of the Hamiltonian formulation of mechanics is the extremely symmetric role played by the coordinates $\boldsymbol{q}$ and the conjugate momenta $\boldsymbol{p}$ — in contrast to the Lagrangian formulation, where the coordinates $\boldsymbol{q}$ and the velocities $\dot{\boldsymbol{q}}$ play very different roles. So it would be nice to introduce a notation that makes this symmetry between the $\boldsymbol{q}$ and $\boldsymbol{p}$ more explicit.

This unified notation is defined by the obvious approach of assembling the coordinates $\boldsymbol{q} = (q_1, \ldots, q_n)$ and the conjugate momenta $\boldsymbol{p} = (p_1, \ldots, p_n)$ into a single vector $\boldsymbol{X} = (q_1, \ldots, q_n, p_1, \ldots, p_n)$ of length $2n$. That is, we define **phase-space coordinates** $\boldsymbol{X} = (X_1, \ldots, X_{2n})$ by

$$X_i \;=\; \begin{cases} q_i & \text{for } 1 \le i \le n \\ p_{i-n} & \text{for } n+1 \le i \le 2n \end{cases} \tag{17}$$

We then introduce a $2n \times 2n$ matrix Ω whose $n \times n$ blocks look like

$$\Omega \;=\; \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix} \tag{18}$$

where I_n denotes the $n \times n$ identity matrix and 0_n denotes the $n \times n$ zero matrix; or in more detail,

$$\Omega_{ij} \;=\; \begin{cases} 1 & \text{if } j = i+n \\ -1 & \text{if } i = j+n \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

Note that the matrix Ω is *antisymmetric*, and that $\Omega^2 = -I$ (why?). This matrix is just the trick we need to get the appropriate minus sign into Hamilton's equations: namely, Hamilton's equations

$$\dot{q}_i \;=\; \frac{\partial H}{\partial p_i} \tag{20a}$$

$$\dot{p}_i \;=\; -\frac{\partial H}{\partial q_i} \tag{20b}$$

for $i = 1, \ldots, n$ can trivially be rewritten as

$$\dot{X}_i \;=\; \sum_{j=1}^{2n} \Omega_{ij} \frac{\partial H}{\partial X_j} \tag{21}$$

for $i = 1, \ldots, 2n$. (I leave it to you to check that this works: you will simply need to check separately the cases $1 \le i \le n$ and $n+1 \le i \le 2n$.)

Likewise, in this notation the Poisson bracket of two functions $f(\boldsymbol{X}, t)$ and $g(\boldsymbol{X}, t)$ takes the very simple form

$$\{f, g\} \;=\; \sum_{i,j=1}^{2n} \frac{\partial f}{\partial X_i} \Omega_{ij} \frac{\partial g}{\partial X_j} \,. \tag{22}$$

(Again, you should check this!) And the fundamental Poisson brackets among the canonical coordinates are simply

$$\{X_i, X_j\} \;=\; \Omega_{ij} \,. \tag{23}$$

(You should check this too!)

We see here the fundamental role played by the matrix Ω in defining the "geometry" of Hamiltonian phase space; it is analogous to the fundamental role played by the identity matrix I in defining the geometry of Euclidean space.

In what follows it will also be extremely convenient to use **Einstein's summation convention**: namely, whenever an index appears exactly twice in a product — as does the index j on the right-hand side of (21), and as do both of the indices i and j on the right-hand side of (22) — then it is automatically considered to be summed from 1 to $2n$ unless explicitly stated otherwise. So, for example, we abbreviate (21) by

$$\dot{X}_i \;=\; \Omega_{ij} \frac{\partial H}{\partial X_j} \,, \tag{24}$$

and we abbreviate (22) by

$$\{f, g\} \;=\; \frac{\partial f}{\partial X_i} \, \Omega_{ij} \, \frac{\partial g}{\partial X_j} \,. \tag{25}$$

This convention saves a lot of writing of summation signs, because in practice repeated indices are nearly always intended to be summed. (The prototype for this is matrix multiplication.)

Let me now use this unified notation to give the promised proof of the Jacobi identity for Poisson brackets. Gregory says (p. 416) that Jacobi's identity

> is quite important, but there seems to be no way of proving it apart from crashing it out, which is very tedious. Unless you can invent a smart method, leave this one alone.

I would like to show you that with the unified notation this proof is a fairly straightforward calculation. So let us consider three functions f, g, h of the phase-space coordinate $\boldsymbol{X}$; and let us prove the Jacobi identity in the form

$$\{f, \{g, h\}\} \,+\, \{g, \{h, f\}\} \,+\, \{h, \{f, g\}\} \;=\; 0 \,. \tag{26}$$

By (22) [and using the summation convention] we have

$$\{f, \{g, h\}\} \;=\; \frac{\partial f}{\partial X_i} \, \Omega_{ij} \, \frac{\partial \{g, h\}}{\partial X_j} \tag{27a}$$

$$\;=\; \frac{\partial f}{\partial X_i} \, \Omega_{ij} \, \frac{\partial}{\partial X_j} \left(\frac{\partial g}{\partial X_k} \, \Omega_{kl} \, \frac{\partial h}{\partial X_l} \right) \tag{27b}$$

$$\;=\; \Omega_{ij} \, \Omega_{kl} \left[\underbrace{\frac{\partial f}{\partial X_i} \frac{\partial^2 g}{\partial X_j \partial X_k} \frac{\partial h}{\partial X_l}}_{\text{“term 1”}} \,+\, \underbrace{\frac{\partial f}{\partial X_i} \frac{\partial^2 h}{\partial X_j \partial X_l} \frac{\partial g}{\partial X_k}}_{\text{“term 2”}} \right] . \tag{27c}$$

Similarly, the other two triple brackets $\{g, \{h, f\}\}$ and $\{h, \{f, g\}\}$ will contains “terms 3–6” obtained by replacing (f, g, h) by the cyclic permutation (g, h, f) for terms 3–4 and by (h, f, g) for terms 5–6. Now I claim that term 1 will cancel term 6, term 3 will cancel term 2, and term 5 will cancel term 4. Let me show the proof in detail for $1 \leftrightarrow 6$; the other cases will obviously follow by cyclic permutation of (f, g, h). We have

$$\text{term 1} \,+\, \text{term 6} \;=\; \Omega_{ij} \, \Omega_{kl} \left[\frac{\partial f}{\partial X_i} \frac{\partial^2 g}{\partial X_j \partial X_k} \frac{\partial h}{\partial X_l} \,+\, \frac{\partial h}{\partial X_i} \frac{\partial^2 g}{\partial X_j \partial X_l} \frac{\partial f}{\partial X_k} \right] \tag{28}$$

In the "term 6" part of this equation, let us interchange the summation indices k and l (both of them are being summed from 1 to $2n$, so we have a right to interchange their names): since $\Omega_{lk} = -\Omega_{kl}$, we have

$$\begin{aligned}
\text{term 1 + term 6} &= \Omega_{ij}\,\Omega_{kl}\left[\frac{\partial f}{\partial X_i}\frac{\partial^2 g}{\partial X_j \partial X_k}\frac{\partial h}{\partial X_l} - \frac{\partial h}{\partial X_i}\frac{\partial^2 g}{\partial X_j \partial X_k}\frac{\partial f}{\partial X_l}\right] && \text{(29a)}\\
&= \Omega_{ij}\,\Omega_{kl}\left(\frac{\partial f}{\partial X_i}\frac{\partial h}{\partial X_l} - \frac{\partial h}{\partial X_i}\frac{\partial f}{\partial X_l}\right)\frac{\partial^2 g}{\partial X_j \partial X_k} && \text{(29b)}\\
&\stackrel{\text{def}}{=} \Omega_{ij}\,\Omega_{kl}\,F_{ijkl} && \text{(29c)}
\end{aligned}$$

where F_{ijkl} is symmetric in j, k and antisymmetric in i, l. But we then have

$$\begin{aligned}
\Omega_{ij}\,\Omega_{kl}\,F_{ijkl} &= \Omega_{lk}\,\Omega_{ji}\,F_{lkji} && \text{by renaming dummy indices} && \text{(30a)}\\
&= \Omega_{kl}\,\Omega_{ij}\,F_{lkji} && \text{by antisymmetry of } \Omega \text{ (used twice)} && \text{(30b)}\\
&= -\Omega_{kl}\,\Omega_{ij}\,F_{ijkl} && \text{since } F_{ijkl} \text{ is symmetric in } j,k \text{ and antisymmetric in } i,l && \text{(30c)}
\end{aligned}$$

But a quantity equal to its own negative must be zero: that is, $\Omega_{ij}\,\Omega_{kl}\,F_{ijkl} = 0$ as claimed. □

4 Canonical transformations

When we were studying Lagrangian mechanics, we saw that one of its advantages over the Newtonian formulation is that it is covariant under arbitrary changes of coordinates: that is, instead of the original coordinates $\boldsymbol{q} = (q_1, \ldots, q_n)$ we could use some new coordinates $\boldsymbol{q}' = (q'_1, \ldots, q'_n)$, defined in an arbitrary way as a function of the old coordinates:

$$\begin{aligned}
q'_1 &= q'_1(q_1, \ldots, q_n, t)\\
&\;\;\vdots\\
q'_n &= q'_n(q_1, \ldots, q_n, t)
\end{aligned} \tag{31}$$

If we then define the Lagrangian $L'(\boldsymbol{q}', \dot{\boldsymbol{q}}', t)$ to have the same *values* as L, i.e.

$$L'(\boldsymbol{q}', \dot{\boldsymbol{q}}', t) \;=\; L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)\;, \tag{32}$$

it turns out that the Lagrange equations of motion for L' are equivalent to those for L. (To prove this directly from the differential equations is a nontrivial calculation; but as pointed out by Gregory, Chapter 13, pp. 387–388, this is an immediate consequence of the variational principle.)

One of the outstanding features of the Hamiltonian formalism is that it possesses an even wider flexibility: not only can we reparametrize coordinate space $\boldsymbol{q}$ to $\boldsymbol{q}'$ as in the Lagrangian formalism (with a corresponding "dual" change from $\boldsymbol{p}$ to $\boldsymbol{p}'$); we can even choose new coordinates that *mix* $\boldsymbol{q}$ and $\boldsymbol{p}$! These transformations turn out to be of immense importance, both theoretical and practical. Here we will only have time to scratch the surface of the theory of **canonical transformations**.

Let us consider, as usual, a $2n$-dimensional Hamiltonian phase space parametrized by canonical coordinates $\boldsymbol{q} = (q_1, \ldots, q_n)$ and $\boldsymbol{p} = (p_1, \ldots, p_n)$. These canonical coordinates have the fundamental Poisson brackets

$$\{q_i, q_j\} = 0 \tag{33a}$$

$$\{p_i, p_j\} = 0 \tag{33b}$$

$$\{q_i, p_j\} = \delta_{ij} \tag{33c}$$

Now let $Q_1, \ldots, Q_n$ and $P_1, \ldots, P_n$ be arbitrary functions of $\boldsymbol{q}$ and $\boldsymbol{p}$ (and also t if we wish). We say that $\boldsymbol{Q} = (Q_1, \ldots, Q_n)$ and $\boldsymbol{P} = (P_1, \ldots, P_n)$ form new canonical coordinates for phase space if they have the correct Poisson brackets, i.e.

$$\{Q_i, Q_j\} = 0 \tag{34a}$$

$$\{P_i, P_j\} = 0 \tag{34b}$$

$$\{Q_i, P_j\} = \delta_{ij} \tag{34c}$$

Such a transformation $(\boldsymbol{q}, \boldsymbol{p}) \mapsto (\boldsymbol{Q}, \boldsymbol{P})$ is called a **canonical transformation**.

In terms of the unified notation this can be stated even more simply. We start from canonical coordinates $\boldsymbol{X} = (X_1, \ldots, X_{2n})$ satisfying the fundamental Poisson brackets

$$\{X_i, X_j\} = \Omega_{ij} \,. \tag{35}$$

We then consider new coordinates $\boldsymbol{Y} = (Y_1, \ldots, Y_{2n})$ that depend in a completely arbitrary way on $\boldsymbol{X}$ (and on t if we wish). The coordinates $\boldsymbol{Y}$ form new canonical coordinates if their Poisson brackets are

$$\{Y_i, Y_j\} = \Omega_{ij} \,. \tag{36}$$

In this case the transformation $\boldsymbol{X} \mapsto \boldsymbol{Y}$ is called a canonical transformation.

But we can easily work out what this means concretely, by using the definition (22) of Poisson brackets in the unified notation. We have

$$\{Y_i, Y_j\} = \sum_{k,l=1}^{2n} \frac{\partial Y_i}{\partial X_k} \Omega_{kl} \frac{\partial Y_j}{\partial X_l} \tag{37a}$$

$$= (J \Omega J^{\mathrm{T}})_{ij} \tag{37b}$$

if we define the Jacobian matrix of the transformation from $\boldsymbol{X}$ to $\boldsymbol{Y}$,

$$J_{ij} \stackrel{\text{def}}{=} \frac{\partial Y_i}{\partial X_j} \,. \tag{38}$$

So the transformation $\boldsymbol{X} \mapsto \boldsymbol{Y}$ is a canonical transformation if and only if

$$J \Omega J^{\mathrm{T}} = \Omega \,. \tag{39}$$

Note that J is actually a *function* of $\boldsymbol{X}$; so what we mean by (39) is that this equation should hold *for all* $\boldsymbol{X}$, i.e. *everywhere in phase space*.

A $2n \times 2n$ matrix J satisfying (39) is called a **symplectic matrix**. So a transformation $\boldsymbol{X} \mapsto \boldsymbol{Y}$ is a canonical transformation if and only if its Jacobian at every point of phase space is a symplectic matrix.

Example 1: Linear transformations. A *linear* transformation $\boldsymbol{Y} = J\boldsymbol{X}$ is a canonical transformation if and only if the matrix J (which is indeed the Jacobian matrix of the transformation) is a symplectic matrix.

One example is the transformation

$$Q_i = p_i \tag{40a}$$

$$P_i = -q_i \tag{40b}$$

that interchanges the $\boldsymbol{q}$ and $\boldsymbol{p}$ (and makes one compensating sign change). You should figure out what the matrix J is in this case, and convince yourself that it is indeed symplectic. (*Hint*: Use $\Omega^{\mathrm{T}} = -\Omega$ and $\Omega^2 = -I$.)

Example 2: Time evolution. Consider a system started at time zero at a phase-space point $\boldsymbol{X}_0 = (\boldsymbol{q}_0, \boldsymbol{p}_0)$, and let it evolve for a time t. Let us write $\boldsymbol{X}(t)$ for the phase-space point where the system arrives at time t: this is a function of the initial condition $\boldsymbol{X}_0$ and t. So, for each value of t, we can consider $\boldsymbol{X}(t)$ as a function of $\boldsymbol{X}_0$. Let us show that, for each t, the transformation $\boldsymbol{X}_0 \mapsto \boldsymbol{X}(t)$ is a canonical transformation.

By hypothesis, the function $\boldsymbol{X}(\boldsymbol{X}_0, t)$ satisfies the partial differential equation[2]

$$\frac{\partial X_i(t)}{\partial t} = \{X_i(t), H\} \tag{41}$$

(where Poisson brackets are taken with respect to the canonical coordinates $\boldsymbol{X}_0$) together with the initial conditions

$$\boldsymbol{X}(0) = \boldsymbol{X}_0 \,. \tag{42}$$

[2]The meaning of $\partial/\partial t$ in equation (41) may be a bit confusing, since it is *not* the same as the meaning of $\partial/\partial t$ in equation (4). So let me try to explain.

In equation (4), f is a function of $\boldsymbol{q}$, $\boldsymbol{p}$ and t: $\partial f/\partial t$ denotes the explicit time-dependence (if any) in f, while $\{f, H\}$ captures the time-dependence of f that comes through its dependence on $\boldsymbol{q}$ and $\boldsymbol{p}$, which themselves evolve in time according to Hamilton's equations (1).

In the present situation, by contrast, there is *no* explicit time-dependence, so $\boldsymbol{X}(t)$ evolves in time *only* by the Poisson bracket $\{\boldsymbol{X}, H\}$. So the left-hand side of (41) would normally be written with d, not ∂. The reason I write it here with ∂ is that we are now considering $\boldsymbol{X}$ to be a function not only of t but also of the initial conditions $\boldsymbol{X}_0$; so I write $\partial/\partial t$ to mean the derivative with respect to t when the initial conditions $\boldsymbol{X}_0$ are held fixed. (Usually we just consider the initial conditions to be fixed, so we don't make the dependence on $\boldsymbol{X}_0$ explicit; but here the whole point is to consider simultaneously *all possible* initial conditions $\boldsymbol{X}_0$, in order to study the properties of the *map* $\boldsymbol{X}_0 \mapsto \boldsymbol{X}(t)$.)

So let us now compute how the Poisson bracket $\{X_i(t), X_j(t)\}$ varies with time:

$$\frac{\partial}{\partial t}\{X_i(t), X_j(t)\} = \left\{\frac{\partial X_i}{\partial t}, X_j\right\} + \{X_i, \frac{\partial X_j}{\partial t}\} \tag{43a}$$

$$= \{\{X_i, H\}, X_j\} + \{X_i, \{X_j, H\}\} \qquad \text{by the PDE (41)} \tag{43b}$$

$$= -\{X_j, \{X_i, H\}\} + \{X_i, \{X_j, H\}\} \qquad \text{by antisymmetry} \tag{43c}$$

$$= \{X_j, \{H, X_i\}\} + \{X_i, \{X_j, H\}\} \qquad \text{again by antisymmetry} \tag{43d}$$

$$= -\{H, \{X_i, X_j\}\} \qquad \text{by the Jacobi identity} \tag{43e}$$

$$= \{\{X_i, X_j\}, H\} \qquad \text{again by antisymmetry .} \tag{43f}$$

So $F_{ij}(t) \stackrel{\text{def}}{=} \{X_i(t), X_j(t)\}$ is a function of $\boldsymbol{X}_0$ and t that satisfies the partial differential equation

$$\frac{\partial F_{ij}(t)}{\partial t} = \{F_{ij}(t), H\} \tag{44}$$

with initial condition

$$F_{ij}(t) = \{(\boldsymbol{X}_0)_i, (\boldsymbol{X}_0)_j\} = \Omega_{ij} \ . \tag{45}$$

But the solution of this partial differential equation is simply the constant function Ω_{ij}!

> More precisely, the constant function Ω_{ij} *does solve* this partial differential equation, since $\{\Omega_{ij}, H\} = 0$ [the Poisson bracket of a constant function with any other function is zero]. And I am taking for granted that, by general theory, we can prove that the solution is unique. Therefore, the solution can *only* be the constant function Ω_{ij}.

One property of canonical transformations is that they **preserve phase-space volume**, i.e. $|\det J| = 1$. This is, in fact, an immediate consequence of (39): taking determinants of both sides, we get

$$(\det J)^2 \, (\det \Omega) = \det \Omega \tag{46}$$

and hence (since $\det \Omega \neq 0$) $\det J = +1$ or -1. (In fact, with more work one can prove that $\det J = +1$.)

Applying this in particular to the canonical transformation associated with time evolution (Example 2 above), we obtain **Liouville's theorem**: time evolution under Hamilton's equations preserves phase-space volumes.

Let us now look more closely at **infinitesimal canonical transformations**. That is, we consider a transformation $\boldsymbol{X} \mapsto \boldsymbol{Y}$ that is very close to the identity map, i.e.

$$\boldsymbol{Y} = \boldsymbol{X} + \epsilon \boldsymbol{\Psi}(\boldsymbol{X}) \tag{47}$$

for some vector function $\boldsymbol{\Psi}(\boldsymbol{X})$ — or writing it in components,

$$Y_i = X_i + \epsilon \psi_i(\boldsymbol{X}) \tag{48}$$

for some functions $\psi_i(\boldsymbol{X})$ $(1 \le i \le 2n)$. We now attempt to determine conditions on the $\{\psi_i\}$ such that this transformation is canonical through first order in ϵ; we do this by testing the conditions (39) on the Jacobian matrix

$$J_{ij} \stackrel{\text{def}}{=} \frac{\partial Y_i}{\partial X_j} \tag{49}$$

of the transformation $\boldsymbol{X} \mapsto \boldsymbol{Y}$. We see from (48) that the Jacobian matrix J is

$$J \;=\; I \,+\, \epsilon K \tag{50}$$

where

$$K_{ij} \stackrel{\text{def}}{=} \frac{\partial \psi_i}{\partial X_j} \tag{51}$$

is the Jacobian matrix of the transformation $\boldsymbol{X} \mapsto \boldsymbol{\Psi}$. Substituting (50) into (39) and keeping only terms through first order in ϵ, we see that the infinitesimal transformation (47) is canonical if and only if

$$K\Omega \,+\, \Omega K^{\mathrm{T}} \;=\; 0\,. \tag{52}$$

Since Ω is antisymmetric, we can also write this as

$$K\Omega \,-\, \Omega^{\mathrm{T}} K^{\mathrm{T}} \;=\; 0 \tag{53}$$

or in other words

$$K\Omega \,-\, (K\Omega)^{\mathrm{T}} \;=\; 0\,. \tag{54}$$

So this says that the matrix $K\Omega$ is symmetric, or equivalently that the matrix

$$\Omega(K\Omega)\Omega^{\mathrm{T}} \;=\; \Omega K \tag{55}$$

is symmetric. This suggests that we should define a new vector function $\boldsymbol{\Phi}(\boldsymbol{X})$ by

$$\boldsymbol{\Phi}(\boldsymbol{X}) \;=\; \Omega\, \boldsymbol{\Psi}(\boldsymbol{X}) \tag{56}$$

— or in components,

$$\varphi_i(\boldsymbol{X}) \;=\; \sum_{j=1}^{2n} \Omega_{ij}\, \psi_j(\boldsymbol{X}) \tag{57}$$

— so that its Jacobian matrix will be ΩK, i.e.

$$\frac{\partial \varphi_i}{\partial X_j} \;=\; (\Omega K)_{ij}\,. \tag{58}$$

Then the symmetry of the matrix ΩK says that

$$\frac{\partial \varphi_i}{\partial X_j} \;=\; \frac{\partial \varphi_j}{\partial X_i} \tag{59}$$

for all pairs i, j. But this is precisely the necessary and sufficient condition for the vector function $\boldsymbol{\Phi}(\boldsymbol{X})$ to be (locally at least) the gradient of a scalar function $F(\boldsymbol{X})$. [You know this in 3 dimensions: a vector field is (locally at least) the gradient of a scalar field if and only if its curl is zero. But the principle holds true in any number of dimensions.] Thus, the infinitesimal transformation (47) is canonical if and only if there exists a scalar function $F(\boldsymbol{X})$ such that

$$\varphi_i(\boldsymbol{X}) \;=\; \frac{\partial F}{\partial X_i} \,. \tag{60}$$

Left-multiplying this by Ω and using the fact that $\boldsymbol{\Psi}(\boldsymbol{X}) = -\Omega\boldsymbol{\Phi}(X)$ since $\Omega^2 = -I$, we get

$$\psi_i(\boldsymbol{X}) \;=\; -\sum_{j=1}^{2n} \Omega_{ij}\, \frac{\partial F}{\partial X_j} \,. \tag{61}$$

This is the necessary and sufficient condition for the infinitesimal transformation (47) to be canonical. It is convenient to get rid of the minus sign by defining $G = -F$; we thus conclude that the infinitesimal transformation (47) is canonical if and only if there exists a scalar function $G(\boldsymbol{X})$ such that

$$\psi_i(\boldsymbol{X}) \;=\; \Omega_{ij}\, \frac{\partial G}{\partial X_j} \tag{62}$$

(where we are now using the summation convention to lighten the notation). That is, every infinitesimal canonical transformation is of the form

$$Y_i \;=\; X_i \,+\, \epsilon\,\Omega_{ij}\, \frac{\partial G}{\partial X_j} \,, \tag{63}$$

and conversely every infinitesimal transformation of this form is canonical. We call G the **generator** of this infinitesimal canonical transformation. We also write (63) in the shorthand form

$$\delta X_i \;=\; \epsilon\,\Omega_{ij}\, \frac{\partial G}{\partial X_j} \,. \tag{64}$$

But by (22) this also has an elegant expression in terms of Poisson brackets, namely

$$\delta X_i \;=\; \epsilon\,\{X_i, G\} \,. \tag{65}$$

So this is one reason why Poisson brackets play such a central role in Hamiltonian mechanics: they show how to generate infinitesimal canonical transformations.

One important special case of (65) is when the generator G is simply the Hamiltonian H: then [by (4)] the transformation (65) is simply time evolution (forward by a time ϵ). And we have already seen in Example 2 above that time evolution is a canonical transformation. But this second proof, using infinitesimal transformations, is arguably simpler than the first proof I gave you.

Another important special case is when the generator G is one of the components of angular momentum $\mathbf{L}$, say L_z. You will show in the next problem set that the (infinitesimal) canonical transformation generated by L_z is a (infinitesimal) rotation around the z axis.