Steppingstones in Hamiltonian dynamics

Thomas F. Jordan
Physics Department
University of Minnesota
Duluth, Minnesota 55812
tjordan@d.umn.edu

Easy steps through essential Hamiltonian dynamics are outlined, from necessary definitions of Poisson brackets and canonical transformations, to a quick proof that Hamiltonian evolution is made by canonical transformations, the quickest proof of Liouville’s theorem, and on to Poincare-Cartan integral invariants and completely integrable dynamics, making room, providing tools, and setting the stage for more recent developments.
I. Introduction

At a few key points, my teaching of classical mechanics relies on calculations that I do differently from standard texts. Strung together with bits of familiar material, they form an alternative route through Hamiltonian dynamics that I find advantageous. It fits more easily into a course that moves from traditional to chaotic dynamics. This requires that I provide notes for students. Here I describe points that may be helpful for teachers as well. They supplement what I have found in the standard texts. I take care to refer to the texts I have found most helpful and those that are closest to what I do.

New developments in classical mechanics have given Hamiltonian dynamics a new role in our teaching. It helps prepare us to work with the methods and understand and appreciate the results of chaotic dynamics.\textsuperscript{1–4} To make room for new developments, we have to delete or reduce some traditional topics. Hamiltonian dynamics involves an accumulated wealth of mathematical development that can be shortened without sacrificing either the basic concepts or the results that illuminate contemporary work. We do not need to know much about canonical transformations, for example, or anything about their generating functions, to work with the canonical transformations of Hamiltonian evolution. Using the way they depend on time can be enough.

We can edit Hamiltonian dynamics for economy of expression and aim it in directions of current interest. We can, if we choose, step lightly over some traditional theory and save time for solving problems, or have time to introduce modern geometrical views.\textsuperscript{1} Minimal steps through essential Hamiltonian dynamics are outlined here, from necessary definitions of Poisson brackets and canonical transformations, to a quick proof that Hamiltonian evolution is made by canonical transformations, the quickest proof of Liouville’s theorem, and on toward contemporary goals, to Poincare-Cartan integral invariants, providing tools that can be used with surface-of-section maps,\textsuperscript{5} and to completely integrable dynamics, setting the stage for use of action-angle variables and discussion of the effect of perturbations on resonant and non-resonant tori and the KAM theorem.\textsuperscript{6} Generating functions are particularly useful for proving that the transformation to action-angle variables is canonical, but that is proved here simply by calculating Poisson brackets. We do not need a power tool for this one job; we can do it by hand.
II. Poisson Brackets

We work with canonical coordinates \( q_1, q_2, \ldots q_N \) and momenta \( p_1, p_2, \ldots p_N \). The Poisson bracket is defined by

\[
[F, G] = \sum_{n=1}^{N} \left( \frac{\partial F}{\partial q_n} \frac{\partial G}{\partial p_n} - \frac{\partial F}{\partial p_n} \frac{\partial G}{\partial q_n} \right)
\]

for any two functions \( F \) and \( G \) of the canonical coordinates and momenta. It is linear for both \( F \) and \( G \). It is antisymmetric; \([G, F] = -[F, G]\). These properties are obvious. It is an easy exercise to prove that the Poisson bracket also satisfies the Jacobi identity

\[
[[F, G], H] + [[G, H], F] + [[H, F], G] = 0.
\]

The canonical coordinates and momenta themselves have Poisson brackets

\[
[q_m, q_n] = 0, \quad [p_m, p_n] = 0
\]

\[
[q_m, p_n] = \delta_{mn}
\]

for \( m, n = 1, 2, \ldots \) or \( N \). Hamilton’s equations

\[
\frac{dq_n}{dt} = \frac{\partial H}{\partial p_n}, \quad \frac{dp_n}{dt} = -\frac{\partial H}{\partial q_n}
\]

can be written as

\[
\frac{dq_n}{dt} = [q_n, H], \quad \frac{dp_n}{dt} = [p_n, H].
\]

Indeed, Hamilton’s equations and the definition of the Poisson bracket imply that for any function \( F \) of the canonical coordinates and momenta

\[
\frac{dF}{dt} = \sum_{n=1}^{N} \left( \frac{\partial F}{\partial q_n} \frac{\partial H}{\partial p_n} - \frac{\partial F}{\partial p_n} \frac{\partial H}{\partial q_n} \right) = [F, H].
\]

The Hamiltonian \( H \) is a function of the canonical coordinates and momenta that may be different for different \( t \); it may be a function of the \( q_n, p_n \) and \( t \).
III. Canonical Transformations

The 2N-dimensional space of points specified by canonical coordinates and momenta is called phase space. We consider changes of coordinates in phase space. Let $Q_n, P_n$, for $n = 1,2,...,N$, be functions of the $q_k,p_k$ that determine the $q_k,p_k$ as functions of the $Q_n,P_n$, so the $Q_n,P_n$ label the points of phase space as well as the $q_k,p_k$. We say that the $Q_n,P_n$ are canonical coordinates and momenta, and the transformation from the $q_k,p_k$ to the $Q_n,P_n$ is a canonical transformation, if

$$[Q_m, Q_n] = 0, \quad [P_m, P_n] = 0$$

$$[Q_m, P_n] = \delta_{mn}$$

for $m,n = 1,2,...,N$. Then the $Q_n,P_n$ can be used in place of the $q_n,p_n$ to calculate Poisson brackets; for any functions $F$ and $G$ of the phase-space variables,

$$[F, G] = \sum_{n,m} \left( \frac{\partial F}{\partial Q_m} [Q_m, Q_n] \frac{\partial G}{\partial Q_n} + \frac{\partial F}{\partial Q_m} [Q_m, P_n] \frac{\partial G}{\partial P_n} \right)$$

$$+ \frac{\partial F}{\partial P_m} [P_m, Q_n] \frac{\partial G}{\partial Q_n} + \frac{\partial F}{\partial P_m} [P_m, P_n] \frac{\partial G}{\partial P_n} \right)$$

$$= \sum_n \left( \frac{\partial F}{\partial Q_n} \frac{\partial G}{\partial P_n} - \frac{\partial F}{\partial P_n} \frac{\partial G}{\partial Q_n} \right).$$

From (2.6) and (3.2) we see that

$$\frac{dQ_n}{dt} = [Q_n, H] = \frac{\partial H}{\partial P_n}$$

$$\frac{dP_n}{dt} = [P_n, H] = -\frac{\partial H}{\partial Q_n}.$$

Hamilton’s equations are the same for the $Q_n,P_n$ as for the $q_n,p_n$. The Hamiltonian is the same thing written differently as a function of different variables.
Conversely, if this is true for any Hamiltonian, the transformation from the \( q_k, p_k \) to the \( Q_n, P_n \) must be canonical; with (2.6) used to write the time derivatives as Poisson brackets, Hamilton’s equations for the \( Q_n, P_n \) give

\[
[Q_n, H] = \frac{\partial H}{\partial P_n}, \quad [P_n, H] = -\frac{\partial H}{\partial Q_n},
\]

which imply (3.1) in the particular cases where \( H \) is \( Q_m \) or \( P_m \).
IV. Hamiltonian Evolution

When the dynamics is described by Hamilton’s equations, the evolution in time is made by canonical transformations. Let \( Q_n \) and \( P_n \) be the functions of the \( q_k, p_k \) and \( t \) that are the solutions of the equations of motion

\[
\frac{dQ_n}{dt} = [Q_n, H], \quad \frac{dP_n}{dt} = [P_n, H]
\]

specified by the boundary conditions

\[
Q_n(t = 0) = q_n, \quad P_n(t = 0) = p_n.
\]

The canonical coordinates and momenta \( q_n \) and \( p_n \) at time zero evolve to \( Q_n, P_n \) at time \( t \). If their sets of values are \( q, p \) at time zero, their values at time \( t \) are \( Q_n(q, p, t) \) and \( P_n(q, p, t) \).

Explicitly, when the Hamiltonian does not depend on time and the series converges,

\[
Q_n = q_n + t [q_n, H] + \frac{1}{2} t^2 [[q_n, H], H]
\]

\[
\ldots + \frac{1}{k!} t^k \ldots [q_n, H] \ldots H + \ldots
\]

the bracket with \( H \) being taken \( k \) times in the term with \( t^k \); this and the corresponding formula for \( P_n \) do satisfy the equations of motion (4.1) and the boundary conditions (4.2).

The \( Q_n, P_n \) are canonical coordinates and momenta. The transformation from the \( q_n, p_n \) to the \( Q_n, P_n \) is a canonical transformation. We prove this by showing that the \( Q_n, P_n \) satisfy the bracket relations (3.1). Using the Jacobi identity (2.2), we get

\[
\frac{d}{dt} [Q_m, P_n] = [[Q_m, H], P_n] + [Q_m, [P_n, H]]
\]

\[
= [[Q_m, P_n], H].
\]

The Poisson bracket \([Q_m, P_n]\) must be the function of the \( q_k, p_k \) and \( t \) that is the solution of this equation of motion specified by the boundary condition

\[
[Q_m, P_n] (t = 0) = [q_m, q_n] = \delta_{mn},
\]
which is \([Q_m, P_n] = \delta_{mn}\). Explicitly, for a time-independent Hamiltonian,

\[
[Q_m, P_n] = [q_m, p_n] + t [[q_m, p_n], H] + \ldots = \delta_{mn}.
\]

The other bracket relations (3.1) can be proved similarly.\(^7\)

The notation used here distinguishes the coordinates and momenta \(Q_n, P_n\) at time \(t\) from the \(q_n, p_n\) at time zero. We will use this notation again when it is helpful, to consider the coordinates and momenta at time \(t\) as functions of the initial values. More often, since now we know that the coordinates and momenta are equally canonical at each time, we will follow the usual practice and work with canonical coordinates and momenta that depend on time, as we did in writing Hamilton’s equations in Section II; this will be indicated by absence of \(Q_n, P_n\).

V. Integral Invariants

A point in phase space marks a set of values for the canonical coordinates and momenta. It moves through phase space as they change in time. The character of this motion reflects the structure of the dynamics. Consider a set of points that occupy a volume in phase space. If the dynamics is described by Hamilton’s equations, the size of the volume the points occupy does not change in time as the points move, even though its shape generally does. This property of Hamiltonian dynamics is called Liouville’s theorem.

The proof is very simple.\(^8\) At any time, the time derivative of the volume is the integral over the surface surrounding the volume of the normal component of the velocity of the motion of the points on the surface. That is the integral over the volume of the divergence of the velocity, which is zero, because for Hamiltonian dynamics the divergence is

\[
\sum_{n=1}^{N} \left( \frac{\partial}{\partial q_n} \frac{dq_n}{dt} + \frac{\partial}{\partial p_n} \frac{dp_n}{dt} \right) = \sum_{n=1}^{N} \left( \frac{\partial^2 H}{\partial q_n \partial p_n} - \frac{\partial^2 H}{\partial p_n \partial q_n} \right) = 0.
\]

We see that this holds for a time-dependent Hamiltonian; the Hamiltonian can be a different function of the \(q_k, p_k\) at each time.

Finding quantities that do not change in time helps build a picture of the dynamics. The volume of Liouville’s theorem is one. Others can be constructed as integrals along curves.
Consider a curve of points in phase space. Let $\theta$ be a parameter that measures location along the curve. The phase-space coordinates $q_n, p_n$ of points on the curve are functions of $\theta$. As a point on the curve moves in time, its phase-space coordinates change from $q_n(\theta), p_n(\theta)$ at time zero to

$$Q_n(q(\theta), p(\theta), t), \quad P_n(q(\theta), p(\theta), t)$$

and time $t$, where $Q_n$ and $P_n$ are the solutions of the equations of motion (4.1) that satisfy the boundary conditions (4.2). We write $q$ for the set of $q_k$ and $p$ for the set of $p_k$. The phase-space coordinates of the moving point are functions of $\theta$ and $t$. We can let points at different $\theta$ move for different intervals of time $t$. If we make the time interval zero at the ends of a curve segment, the moved segment and the original segment will form a closed loop, as shown in Figure 1. We can replace the original segment with another moved curve to make a closed loop on which the phase-space coordinates change as functions of both $\theta$ and $t$ all the way around. In this way we can generate any closed loop in phase space that surrounds an area of surface of tracks of motion in time, an area of surface swept by a curve segment moving in time.

For any such loop,

$$\oint \left( \sum_{n=1}^{N} P_n dQ_n - H dt \right) = 0,$$

the integral being around the closed loop. We can see this from

$$\oint \left( \sum_{n=1}^{N} P_n \left( \frac{\partial Q_n}{\partial \theta} d\theta + \frac{\partial Q_n}{\partial t} dt \right) - H dt \right)$$

$$= \oint (A_\theta d\theta + A_t dt) = \oint \vec{A} \cdot d\vec{l} = \oint \nabla \times \vec{A} \cdot d\vec{S}$$

with

$$A_\theta = \sum_{n=1}^{N} P_n \frac{\partial Q_n}{\partial \theta}, \quad A_t = \sum_{n=1}^{N} P_n \frac{\partial Q_n}{\partial t} - H,$$
the integrals involving $\vec{A}$ being around the closed loop in the $\theta, t$ plane and over the area it encloses; because for Hamiltonian dynamics

$$\frac{\partial A_t}{\partial t} - \frac{\partial A_{\theta}}{\partial \theta} = \sum_{n=1}^{N} \frac{\partial P_n}{\partial t} \frac{\partial Q_n}{\partial \theta} - \sum_{n=1}^{N} \frac{\partial P_n}{\partial \theta} \frac{\partial Q_n}{\partial t} + \frac{\partial H}{\partial \theta}$$

$$= \sum_{n=1}^{N} \left( -\frac{\partial H}{\partial Q_n} \frac{\partial Q_n}{\partial \theta} \right) - \sum_{n=1}^{N} \frac{\partial P_n}{\partial \theta} \frac{\partial H}{\partial P_n} + \frac{\partial H}{\partial \theta} = 0.$$  

This implies that when a closed loop in phase space is changed by letting the points on it move for different intervals of time, the integral around the loop

$$\oint \left( \sum_{n=1}^{N} P_n dQ_n - H dt \right)$$

is not changed. We can see this by letting our original curve be a closed loop and letting the points on it move for different intervals of time to form a moved loop. Connecting the two loops along two closely adjacent tracks of motion, as in Figure 2, yields a closed loop to which (5.3) applies. It surrounds a surface of tracks of motion, a surface swept by a curve moving in time. Since the integrals in opposite directions along the two tracks of motion cancel, the integrals in opposite directions around the two loops must cancel, which means the integrals in the same direction around the two loops are the same.

These results are important for surface-of-section maps. The quantities (5.7) that do not change in time are called Poincare-Cartan integral invariants. When the $H dt$ term is absent, they are called Poincare integral invariants. The $H dt$ term will be absent if there are no time differentials or if $H$ can be taken outside the integral. We will use (5.3) in the next section.
VI. Complete Integrability

A canonical transformation can make the equations of motion so simple that their solutions become trivial. This is possible if the dynamics are completely integrable. That means there are \( N \) constants of the motion, which we call \( I_n \) with \( n \) running from 1 to \( N \), that in addition to

\[
[I_n, H] = 0
\]

which says they are constants, satisfy

\[
[I_m, I_n] = 0
\]

for \( m, n = 1, 2, \ldots N \). It is assumed also that these \( I_n \) are functions of the \( q_k, p_k \) that determine the \( p_k \) as functions of the \( q_n, I_n \) in a sufficiently relevant region of phase space, so the \( q_n, I_n \) label those points of phase space as well as the \( q_k, p_k \).

These assumptions of complete integrability imply, as we will show, that there are \( Q_n \) for which the \( Q_n, I_n \) are canonical coordinates and momenta.\(^\text{11}\) The \( P_n \) are the \( I_n \). They are all constants of the motion. That means Hamilton’s equations give

\[
\frac{\partial H}{\partial Q_n} = -\frac{dP_n}{dt} = 0.
\]

Then \( H \) is a function only of the \( P_n \) and

\[
\frac{dQ_n}{dt} = \frac{\partial H}{\partial P_n}
\]

is a constant of the motion. The equations of motion are solved. The \( P_n \) are constant in time. The \( Q_n \) change at constant rates; they are linear functions of time. Action-angle variables are a particular case of these canonical coordinates and momenta \( Q_n \) and \( I_n \) for completely integrable dynamics.

We will take two steps to prove that there are canonical coordinate partners \( Q_n \) for canonical momenta \( I_n \). The first step, needed to define the \( Q_n \), is to show that there is a function \( S \) of the \( q_n, I_n \) such that

\[
\frac{\partial S}{\partial q_n} = p_n.
\]
It is an action integral. Let

$$S(\mathbf{q}, \mathbf{I}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \sum_{k=1}^{N} p_k(q', \mathbf{I}) dq'_k.$$  

We write \( \mathbf{q} \) for the set of \( q_n \) and \( \mathbf{I} \) for the set of \( I_n \). For each fixed set of values of the \( I_n \), the integral gives \( S \) as a function of the \( q_n \). The integral is along a path on the surface in phase space where the \( I_n \) are constant, starting at the point where the \( q \) have values \( \mathbf{q}_0 \). To establish that \( S \) is well defined, we need to show that the integral is not changed by variations in the path.

We have described the motion in phase space generated by the Hamiltonian \( H \) for evolution in time. Now we consider also the motion that is generated the same way when the Hamiltonian is replaced by one of the \( I_n \) or a linear combination of the \( I_n \). From the bracket relations (6.2) we see that this motion does not change any of the \( I_m \). It stays on the surface in phase space where the \( I_m \) stay constant. The velocity vector for the motion generated by \( I_n \), the vector that points along the direction of the motion in phase space and measures the rate of the motion with respect to a parameter \( s \) along the curve of the motion, has components

$$\frac{dq_1}{ds} = \frac{\partial I_n}{\partial p_1}, \quad \ldots \quad \frac{dq_n}{ds} = \frac{\partial I_n}{\partial p_n}. \quad \frac{dp_1}{ds} = -\frac{\partial I_n}{\partial q_1}, \quad \ldots \quad \frac{dp_n}{ds} = -\frac{\partial I_n}{\partial q_n}.$$ 

We are assuming that the \( q_n, I_n \) are coordinates for a region of phase space. For each fixed set of values of the \( q_n \), the \( I_n \) must be coordinates for an \( N \)-dimensional space. They must vary in \( N \) different directions. Their gradients must be \( N \) linearly independent vectors. From (6.7), which relates the components of the gradient of \( I_n \) to the components of the velocity vector for the motion generated by \( I_n \), we see that the \( N \) velocity vectors for the motions generated by the \( I_n \) must be linearly independent. The velocity vectors for the motions generated by the \( I_n \) and linear combinations of the \( I_n \) point in all directions on the surface of constant \( I_m \).

Now we can prove that the integral (6.6) is not changed by variations in the path on the surface of constant \( I_m \). If an infinitesimal segment of the path is moved on the surface of constant \( I_m \), the moved segment and the original
segment form a closed loop, as in Figure 1. Now the motion is generated by a linear combination of the $I_n$ instead of $H$. We can use (5.3) with this replacement. The $Hdt$ term becomes zero, because the linear combination of the $I_n$ is a constant that can be taken out of the integral, and the change in the parameter $t$ or $s$ around the closed loop is zero. Therefore

$$\oint P_k dQ_k = 0$$

for the closed loop, which means the integral (6.6) is the same for the new path as for the old. We conclude that $S$ is well defined. It clearly satisfies (6.5).

That allows us to define the $Q_n$. Let

$$Q_n = \frac{\partial S}{\partial I_n}.$$ 

The final step is to show that the $Q_n$ and $I_n$ are canonical coordinates and momenta, that the bracket relations (3.1) are satisfied with $I_n$ for $P_n$ and these $Q_n$.

Using (6.2), (6.5) and (6.9) and taking partial derivatives of $S$, the $Q_n$ and the $p_n$ with respect to the $q_k, I_k$ and partial derivatives of the $I_n$ with respect to the $q_n, p_n$, we get

$$[Q_m, I_n] = \sum_{j=1}^{N} \frac{\partial Q_m}{\partial q_j} [q_j, I_n] + \sum_{j=1}^{N} \frac{\partial Q_m}{\partial I_j} [I_j, I_n]$$

$$= \sum_{j=1}^{N} \frac{\partial^2 S}{\partial q_j \partial I_m \partial p_j} \frac{\partial I_n}{\partial p_j} + \sum_{j=1}^{N} \frac{\partial p_j}{\partial I_m} \frac{\partial I_n}{\partial p_j}$$

$$= \sum_{j=1}^{N} \frac{\partial I_n}{\partial p_j} \frac{\partial p_j}{\partial I_m} + \sum_{j=1}^{N} \frac{\partial I_n}{\partial q_j} \frac{\partial q_j}{\partial I_m}$$

$$= \frac{\partial I_n}{\partial I_m} = \delta_{mn};$$

in this context $\partial q_j/\partial I_m$ is zero because the derivatives of $I_n$ are with respect to the $q_j$ and $p_j$ and the derivatives of the $p_j$ and $q_j$ are with respect to the $q_n$ and $I_m$. Similarly, we get
\[
\{Q_m, Q_n\} = \sum_{i,j=1}^{N} \frac{\partial Q_m}{\partial q_i} [q_i, q_j] \frac{\partial Q_n}{\partial q_j} + \sum_{i,j=1}^{N} \frac{\partial Q_m}{\partial I_i} [I_i, q_j] \frac{\partial Q_n}{\partial q_j} + \sum_{j=1}^{N} [Q_m, I_j] \frac{\partial Q_n}{\partial I_j}
\]

\[
= \sum_{i,j=1}^{N} \frac{\partial Q_m}{\partial I_i} \left( -\frac{\partial I_i}{\partial p_j} \right) \frac{\partial Q_n}{\partial q_j} + \frac{\partial Q_n}{\partial I_m}
\]

\[
= -\sum_{i,j=1}^{N} \frac{\partial^2 S}{\partial I_i \partial I_m} \frac{\partial I_i}{\partial p_j} \frac{\partial^2 S}{\partial q_j \partial I_m} + \frac{\partial^2 S}{\partial I_m \partial I_n}
\]

\[
= -\sum_{i,j=1}^{N} \frac{\partial^2 S}{\partial I_i \partial I_m} \frac{\partial I_i}{\partial p_j} \frac{\partial p_j}{\partial I_m} + \frac{\partial^2 S}{\partial I_m \partial I_n}
\]

\[
= -\frac{\partial^2 S}{\partial I_m \partial I_n} + \frac{\partial^2 S}{\partial I_m \partial I_n} = 0
\]

by using again the last steps of (6.10) which show that in this context

\[
\sum_{j} \frac{\partial I_i}{\partial p_j} \frac{\partial p_j}{\partial I_m} = \delta_{im}.
\]


5 Reference 2, pp. 123-6 and 183-4.

6 Reference 1, Chapter 9; reference 2, Chapters 3 and 4; reference 3, Chapter 7.


9 Poincare-Cartan integral invariants are usually discussed in the context of an extended phase space that includes the time dimension. For example, see reference 3, pp. 212-3.


11 The method of proof used here is outlined in reference 1, pp. 156-8. For the usual method using generating functions see, for example, reference 4;
Figure 1. A closed loop made by points on a curve segment moving for different intervals of time.
Figure 2. A closed loop made by an initial loop and a moved loop connected along closely adjacent tracks of motion.