PHY411/AST233 Lecture notes – Canonical Transformations

Alice Quillen

November 4, 2024

Contents

1 Canonical Transformations

It is straightforward to transfer coordinate systems using the Lagrangian formulation as minimization of the action can be done in any coordinate system. However, in the Hamiltonian formulation, only some coordinate transformations preserve Hamilton's equations. **Canonical transformations**, defined here as those that preserve the Poisson brackets or equivalently the symplectic 2-form, also preserve Hamilton's equations. A search for conserved quantities and symmetries is equivalent to a search for a nice coordinate system that preserves Hamilton's equations.

1.1 Poisson Brackets

Consider a function $f(q,p,t)$ and a Hamiltonian $H(p,q)$ where p,q coordinates and momenta. The time dependence of f

$$\frac{df}{dt} = \frac{\partial f}{\partial q}\dot{q} + \frac{\partial f}{\partial p}\dot{p} + \frac{\partial f}{\partial t}$$

Using Hamilton's equations we can write this as

$$\frac{df}{dt} = \frac{\partial f}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial f}{\partial t}$$

We can write this short hand with a commutation relation known as the *Poisson bracket*

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} \quad (1)$$

with the Poisson bracket for two functions f, g

$$\{f, g\} \equiv \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q}$$

For more than one dimension

$$\{f, g\} = \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \right)$$

and using summation notation we neglect the $\sum$ symbol.

What are the Poisson brackets of functions equal to the coordinates and momenta, $f(\mathbf{q},\mathbf{p}) = q_i$ and $g(\mathbf{q},\mathbf{p}) = p_i$? We calculate

$$\{p_i, p_j\} = 0 \qquad \{q_i, q_j\} = 0 \qquad \{q_i, p_j\} = \delta_{ij}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{for} \quad i = j \\ 0 & \text{for} \quad i \neq j \end{cases}$$

Coordinates and momenta resemble an orthogonal basis.

Consider the functions $f(q,p) = q$ and $g(q,p) = p$. Inserting these functions into equation 1 we recover Hamilton's equations in terms of Poisson brackets

$$\begin{aligned} \dot{q} &= \frac{\partial H}{\partial p} = \{q, H\} \\ \dot{p} &= -\frac{\partial H}{\partial q} = \{p, H\} \end{aligned}$$

The Poisson bracket satisfies the conditions for a Lie algebra. For functions f, g, h,

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

$$\{f,g\} + \{g,f\} = 0$$

The first of these is called a Jacobi identity, the second is antisymmetry. In addition they satisfy another condition known as a Leibnitz type of product rule

$$\{f, gh\} = g\{f,h\} + \{f,g\}h$$

A Lie algebra with this extra rule is called a Poisson algebra.

Remark In what contexts are these extra rules important? These relations give the Poisson bracket Lie bracket-like constraints. The Leibnitz rule makes the Poisson bracket behave like a derivative. Infinite dimensional continuous systems with equations of motion corresponding to partial differential equations can be described with a a Hamiltonian and a Poisson bracket. However, we might lack canonical coordinates. For example the KdV equation can be described with a Hamiltonian and Poisson bracket but there isn't a pair of canonical coordinates.

1.2 Canonical transformations

A canonical transformation is a transformation from one set of coordinates $\mathbf{q},\mathbf{p}$ to a new one $Q(\mathbf{q},\mathbf{p}), P(\mathbf{q},\mathbf{p})$ that satisfies the Poisson brackets

$$\{P_i, P_j\} = 0 \qquad \{Q_i, Q_j\} = 0 \qquad \{Q_i, P_j\} = \delta_{ij} \tag{2}$$

The above Poisson brackets are computed using derivatives of p, q.

Using $\mathbf{x} = (q_1, q_2 ...., p_1, p_2, ...p_N)$

$$\{x_i, x_j\} = \omega_{ij}$$

with

$$\omega = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

and I the identity matrix.

Given a Hamiltonian, $H(p, q, t)$, we will show that we can find a new Hamiltonian $K(Q, P, t)$ such that Hamilton's equations are obeyed in the new coordinate system.

For a function $f(\mathbf{Q}(\mathbf{p}, \mathbf{q}), \mathbf{P}(\mathbf{p}, \mathbf{q}))$ using the chain rule and using summation notation

$$\begin{aligned}
\frac{\partial f}{\partial q_i} &= \frac{\partial f}{\partial Q_j}\frac{\partial Q_j}{\partial q_i} + \frac{\partial f}{\partial P_j}\frac{\partial P_j}{\partial q_i} \\
\frac{\partial f}{\partial p_i} &= \frac{\partial f}{\partial Q_j}\frac{\partial Q_j}{\partial p_i} + \frac{\partial f}{\partial P_j}\frac{\partial P_j}{\partial p_i}
\end{aligned}$$

The Poisson bracket of $f(\mathbf{Q}(\mathbf{p}, \mathbf{q}), \mathbf{P}(\mathbf{p}, \mathbf{q}))$ and $g(\mathbf{Q}(\mathbf{p}, \mathbf{q}), \mathbf{P}(\mathbf{p}, \mathbf{q}))$

$$\begin{aligned}
\{f, g\} &= \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \\
&= \left(\frac{\partial f}{\partial Q_j}\frac{\partial Q_j}{\partial q_i} + \frac{\partial f}{\partial P_j}\frac{\partial P_j}{\partial q_i}\right)\left(\frac{\partial g}{\partial Q_k}\frac{\partial Q_k}{\partial p_i} + \frac{\partial g}{\partial P_k}\frac{\partial P_k}{\partial p_i}\right) \\
&\quad - \left(\frac{\partial f}{\partial Q_j}\frac{\partial Q_j}{\partial p_i} + \frac{\partial f}{\partial P_j}\frac{\partial P_j}{\partial p_i}\right)\left(\frac{\partial g}{\partial Q_k}\frac{\partial Q_k}{\partial q_i} + \frac{\partial g}{\partial P_k}\frac{\partial P_k}{\partial q_i}\right) \\
&= \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial P_k}\left(\frac{\partial Q_j}{\partial q_i}\frac{\partial P_k}{\partial p_i} - \frac{\partial Q_j}{\partial p_i}\frac{\partial P_k}{\partial q_i}\right) - \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial Q_k}\left(\frac{\partial Q_k}{\partial q_i}\frac{\partial P_j}{\partial p_i} - \frac{\partial Q_k}{\partial p_i}\frac{\partial P_j}{\partial q_i}\right) \\
&\quad + \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial Q_k}\left(\frac{\partial Q_j}{\partial q_i}\frac{\partial Q_k}{\partial p_i} - \frac{\partial Q_j}{\partial p_i}\frac{\partial Q_k}{\partial q_i}\right) + \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial P_k}\left(\frac{\partial P_j}{\partial q_i}\frac{\partial P_k}{\partial p_i} - \frac{\partial P_j}{\partial p_i}\frac{\partial P_k}{\partial q_i}\right) \\
&= \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial P_k}\{Q_j, P_k\} - \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial Q_k}\{Q_k, P_j\} \\
&\quad + \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial Q_k}\{Q_j, Q_k\} + \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial P_k}\{P_j, P_k\}
\end{aligned}$$

If the new coordinates obey the Poisson brackets in equation 2 (so that the transformation is canonical) then we can insert these relations into the above equation.

$$\begin{aligned}
\{f, g\} &= \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial P_k}\delta_{jk} - \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial Q_k}\delta_{jk} \\
&= \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial P_j} - \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial Q_j}
\end{aligned}$$

This is just the definition of the Poisson bracket but with respect to our new coordinates, P, Q rather than p, q. If the transformation is canonical then we can compute Poisson brackets using the new coordinates and momenta. If the coordinate transformation is canonical (the Poisson brackets of equation 2 are obeyed in the new coordinate system) then the Poisson bracket can be computed in the new coordinate system

$$\{f, g\}|_{pq} = \{f, g\}|_{PQ}$$

Above we defined canonical transformations without even specifying a Hamiltonian function. Given $H(\mathbf{q}, \mathbf{p})$, Hamilton's equations give

$$\dot{\mathbf{q}} = \{\mathbf{q}, H\}|_{pq} \qquad \dot{\mathbf{p}} = \{\mathbf{p}, H\}|_{pq}$$

and using the q, p coordinates. But this is true for any time independent function including $\mathbf{Q}(\mathbf{q}, \mathbf{p})$ and $\mathbf{P}(\mathbf{q}, \mathbf{p})$ so

$$\dot{\mathbf{Q}} = \{\mathbf{Q}, H\}|_{pq} \qquad \dot{\mathbf{P}} = \{\mathbf{P}, H\}|_{pq}$$

and the Poisson bracket is computed using the p, q coordinate system. However if the transformation is canonical then the Poisson brackets can be computed using either coordinate system. So

$$\dot{\mathbf{Q}} = \{\mathbf{Q}, H\}|_{PQ} \qquad \dot{\mathbf{P}} = \{\mathbf{P}, H\}|_{PQ}$$

but now we compute the Poisson bracket with the new coordinates P, Q. Thus the new Hamiltonian is equivalent to the old Hamiltonian but using the new variables;

$$K(\mathbf{Q}, \mathbf{P}) = H\left(\mathbf{q}(\mathbf{Q}, \mathbf{P}), \mathbf{p}(\mathbf{Q}, \mathbf{P})\right)$$

You may notice that there is a term missing from this expression. We will discuss time dependent transformations below.

1.3 Canonical Transformations are Symplectic

A symplectic transformation S, obeys

$$J = S^t J S$$

where

$$J \equiv \begin{pmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{pmatrix}$$

and S^t is the transpose of S and $\mathbf{I}$ the identity matrix. Consider a canonical transformation $P(p, q), Q(p, q)$ and the Jacobian matrix

$$S = \begin{pmatrix} \dfrac{\partial Q(q,p)}{\partial q} & \dfrac{\partial Q(q,p)}{\partial p} \\ \dfrac{\partial P(q,p)}{\partial q} & \dfrac{\partial P(q,p)}{\partial p} \end{pmatrix} \tag{3}$$

Let us compute $S^t J S$

$$\begin{aligned}
S^t J S &= \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p} & \frac{\partial P}{\partial p} \end{pmatrix} \begin{pmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p} & \frac{\partial P}{\partial p} \end{pmatrix} \begin{pmatrix} \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \\ -\frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial Q}{\partial q}\frac{\partial P}{\partial q} - \frac{\partial Q}{\partial q}\frac{\partial P}{\partial q} & \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} - \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} & \frac{\partial Q}{\partial p}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial p} \end{pmatrix} \\
&= \begin{pmatrix} 0 & \{Q, P\} \\ -\{Q, P\} & 0 \end{pmatrix}
\end{aligned}$$

If the coordinate transformation is canonical and the Poisson brackets are satisfied, then the transformation is symplectic.

Take a look again at equation 3 for the Jacobian matrix of the coordinate transformation which we repeat here:

$$S = \begin{pmatrix} \frac{\partial Q(q,p)}{\partial q} & \frac{\partial Q(q,p)}{\partial p} \\ \frac{\partial P(q,p)}{\partial q} & \frac{\partial P(q,p)}{\partial p} \end{pmatrix}$$

If the Jacobian has determinant of 1 then the transformation is volume preserving. We compute the determinant

$$\begin{aligned}
\det S &= \frac{\partial Q(q,p)}{\partial q}\frac{\partial P(q,p)}{\partial p} - \frac{\partial Q(q,p)}{\partial p}\frac{\partial P(q,p)}{\partial q} \\
&= \{Q, P\}_{qp}, \qquad (4)
\end{aligned}$$

When the transformation is canonical, Poisson bracket is 1 and the determinant of the Jacobian is 1. This means that the transformation is volume preserving in phase space.

Remark It can be useful in numerical integrations to use discrete transformations that are symplectic. A system that varies continuously with time can be advanced with a discrete time step. A symplectic transformation can be used to transform the system, across a time interval, approximating the time dependence of the real system.

1.4 Generating Functions for Canonical Transformations

Not every coordinate transformation is canonical. Furthermore, the requirement that Poisson brackets are satisfied does not strongly restrict the transformation.

In many classical mechanics texts, canonical transformations are introduced with generating functions. Suppose we start with p, q and refer to these as 'old'. We want to find a better set of momenta and coordinates, P, Q. We refer to these as 'new'. Generating functions are functions of both new and old coordinates and momenta.

Suppose we take a generating function $F_2(q, P)$ of old coordinates and new momenta and define

$$\begin{aligned} p &= \frac{\partial F_2(q,P)}{\partial q} \\ Q &= \frac{\partial F_2(q,P)}{\partial P}. \end{aligned} \tag{5}$$

Really we describe the generating function in terms of p, q or P, Q so what we mean by $F_2(q, P)$ is $F_2(q(Q, P), P)$. The function F_2 has two arguments so we rewrite equations 5 like this:

$$\begin{aligned} p &= \partial_1 F_2(q,P) \\ Q &= \partial_2 F_2(q,P) \end{aligned} \tag{6}$$

where ∂_1 is the derivative with respect to the first argument of the function $F_2(,)$, and ∂_2 is the derivative with respect to the second argument. Taking the second equation (for Q) we can write

$$\begin{aligned} Q &= \partial_2 F_2(q(Q,P),P) \\ \frac{\partial Q}{\partial Q} &= 1 = [\partial_1\partial_2 F_2]\frac{\partial q}{\partial Q} \end{aligned} \tag{7}$$

Taking the equation for p (in equations 6) we can write

$$\begin{aligned} p(Q,P) &= \partial_1 F_2(q(Q,P),P) \\ \frac{\partial p}{\partial Q} &= [\partial_1\partial_1 F_2]\frac{\partial q}{\partial Q} \\ \frac{\partial p}{\partial P} &= [\partial_1\partial_1 F_2]\frac{\partial q}{\partial P} + \partial_1\partial_2 F_2 \end{aligned}$$

We have computed relations for $\frac{\partial p}{\partial Q}, \frac{\partial p}{\partial P}, \frac{\partial q}{\partial Q}$. We don't need to compute a relation for $\frac{\partial q}{\partial P}$ because it will cancel out from our next computation.

Take the Poisson bracket and insert the relations for $\frac{\partial p}{\partial Q}$ and $\frac{\partial p}{\partial P}$

$$\begin{aligned} \{q,p\} &= \frac{\partial q}{\partial Q}\frac{\partial p}{\partial P} - \frac{\partial q}{\partial P}\frac{\partial p}{\partial Q} \\ &= \frac{\partial q}{\partial Q}\left(\partial_1\partial_1 F_2\frac{\partial q}{\partial P} + \partial_1\partial_2 F_2\right) - \frac{\partial q}{\partial P}\partial_1\partial_1 F_2\frac{\partial q}{\partial Q} \\ &= \frac{\partial q}{\partial Q}\partial_1\partial_2 F_2 \\ &= 1 \end{aligned}$$

where the last step uses equation 7.

The coordinate transformation is canonical as long as we define the new coordinate and momenta using equations 5. Similar choices can be made for generating functions that depend on old and new coordinates, old and new momenta or old momenta and new coordinates. Traditionally these are denoted F_1, F_2, F_3, F_4.

1.4.1 Example canonical transformation - action angle coordinates for the harmonic oscillator

Given coordinates ϕ, I we consider new coordinates

$$q(I,\phi) = \sqrt{2I}\sin\phi \qquad p(I,\phi) = \sqrt{2I}\cos\phi \tag{8}$$

We check the Poisson bracket

$$\begin{aligned} \{q,p\} &= \frac{\partial q}{\partial \phi}\frac{\partial p}{\partial I} - \frac{\partial q}{\partial I}\frac{\partial p}{\partial \phi} \\ &= \cos^2\phi + \sin^2\phi = 1 \end{aligned}$$

verifying that this is a canonical transformation. Note that we need a factor of two within the square root in equation 8 so that the Poisson bracket gives 1 instead of 1/2.

This is a handy canonical transformation for the harmonic oscillator with Hamiltonian

$$H(p,q) = \frac{1}{2}\left(p^2 + q^2\right)$$

In the coordinates I, ϕ the Hamiltonian is particularly simple

$$K(I,\phi) = I$$

This system is said to be in action angle variables as I (the action) is conserved and $\dot{\phi}$ is constant.

Can we find a generating function that gives this canonical transformation from p, q to I, ϕ? Consider the generating function of old momenta, p, and new (ϕ) coordinates,

$$F_3(p,\phi) = \frac{p^2}{2}\tan\phi$$

with

$$\begin{aligned} \frac{\partial F_3}{\partial p} &= p\tan\phi = q \\ \frac{\partial F_3}{\partial \phi} &= \frac{p^2}{2}\sec^2\phi = I \end{aligned}$$

From this we find that

$$\tan^2\phi = (q/p)^2$$

or

$$\sec^2\phi = \tan^2\phi + 1 = (q/p)^2 + 1$$

and

$$\begin{aligned} q/p &= \mathrm{atan}\phi \\ \frac{1}{2}(q^2+p^2) &= I \end{aligned}$$

consistent with equation 8.

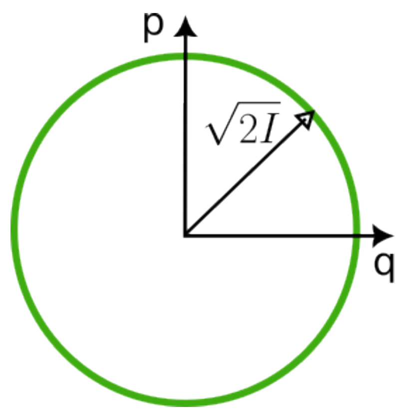

Figure 1: The radius in phase space for an orbit of the harmonic oscillator is equal to $\sqrt{2I}$ where I is the action variable.

1.4.2 The action variable

Let's look again at our new coordinate and momentum

$$q(I,\phi) = \sqrt{2I}\sin\phi \qquad p(I,\phi) = \sqrt{2I}\cos\phi \tag{9}$$

and the associated Hamiltonian in the old and new coordinates $H = p^2/2 + q^2/2 = I$. Notice that $q^2+p^2 = 2I$. In phase space the radius from the origin is $\sqrt{2I}$. Can we instead use a coordinate that is actually the radius in phase space of the orbit? We could, but it would not be a coordinate that is conjugate to the angle ϕ. The canonical momentum that is conjugate to ϕ is I as defined above.

We restore the units using a frequency ω_0.

$$H(p,q) = \frac{p^2}{2} + \omega_0^2\frac{q^2}{2} \tag{10}$$

(taking energy per unit mass). The new momenta and coordinates would satisfy

$$q(I,\phi) = \sqrt{2I/\omega_0}\sin\phi \qquad p(I,\phi) = \sqrt{2I\omega_0}\cos\phi. \tag{11}$$

We insert this into the Hamiltonian to find

$$H(I,\phi) = I\omega_0$$

This makes it clearer that we have an action I and an angle ϕ and that $\frac{\partial H}{\partial I} = \dot{\phi}$ is a frequency ω_0.

Notice that $H = I\omega_0$ is now independent of the angle ϕ. The Hamiltonian is said to be action angle coordinates. Recall that because volume is conserved for Hamiltonian systems, the area in phase space is conserved. The area of an orbit in phase space for a closed orbit is

$$A = \oint pdq$$

For the harmonic oscillator with $H(I,\phi) = I\omega_0$, the area of an orbit is set by I alone. As the radius of the orbit in phase space is $R = \sqrt{2I}$, the area of an orbit that has action variable I is

$$A = \oint pdq = \pi R^2 = 2\pi I.$$

Equivalently

$$A = \oint I d\phi = 2\pi I.$$

This inspires a definition for an action variable for a general system (not just the harmonic oscillator). If you have a closed orbit, the action variable can be defined as

$$I \equiv \frac{1}{2\pi}\oint pdq. \tag{12}$$

Remark We defined canonical transformations as transformations that preserve the Poisson brackets and below we will show that an equivalent statement is that they preserve the symplectic two-form. We did not even mention a Hamiltonian in our definition. Often you see symplectic transformations defined as those that preserved Hamilton's equations. However, Arnold gives an example of a transformation that preserves Hamilton's equations and does not preserve the Poisson bracket.

2 Some Geometry

Why introduce mathematical jargon? It is perhaps insightful to reformulate dynamics in a coordinate-free manor and coordinates can be considered arbitrary or chosen for convenience, rather than related to the underlying dynamics. The geometric view aids in thinking about dynamical problems in a more abstract and general way.

Consider coordinates $\mathbf{p}$ as positions or points on a manifold, $\mathbb{M}$. A manifold is a topological space where each point has a neighborhood that resembles a Euclidean space of dimension n. A map to the Euclidean space is called a *chart*. A collection of charts is called an *atlas*. Here we assume that our manifold is *differentiable* which implies that nearby charts can be smoothly related to one another. Namely, there are transition maps from one chart to another that are differentiable.

2.1 The Tangent Bundle

A curve is a map from $\mathbb{R}$ (parametrized by time t) to the manifold. On each chart (and near a point p) this gives at each time a vector in $\mathbb{R}^N$. If we have a curve c in the manifold we can consider its image in the local Euclidean space at point $\mathbf{x}$. The tangents to the curve at the point $\mathbf{p}$ are in the tangent space.

The tangent space at position $\mathbf{p}$ we can call $\mathbf{T}_p\mathbb{M}$. The *tangent bundle* $\mathbf{T}\mathbb{M}$ consists of the manifold $\mathbb{M}$ with the collection of all its tangent spaces.

It is convenient to use a coordinate basis to describe vectors in the tangent space, for example $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ that is sometimes written $\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$.

A trajectory is a curve on the manifold. Given a curve on the manifold (described by t), the tangent to the curve is $\dot{\mathbf{q}}$. In the Lagrangian formalism the manifold coordinates are specified by $\mathbf{q}$. The Lagrangian is a function of $\mathbf{q}$ (coordinate on the manifold) and $\dot{\mathbf{q}}$ (in the tangent space) and time t (along the curve).

In the Hamiltonian formalism the manifold coordinates are specified by phase space $\mathbf{q}, \mathbf{p}$ and curves given as a function of the phase space coordinates. Here $\dot{\mathbf{q}}$ and $\dot{\mathbf{p}}$ are both vectors in the tangent space.

A *flow* on the manifold is described by velocities which lie in the tangent space at each point $\mathbf{p}$ on the manifold. A vector field on the manifold generates a flow on the manifold. We put a vector on each point of the manifold.

It is sometimes useful to keep track of the map (based on a chart) from the manifold to the Euclidean space $\phi(p)$ which gives a point $\mathbf{x}$ in Euclidean space $\mathbb{R}^N$. The inverse of this map $\phi^{-1}(\mathbf{x}) \to p$ gives points on the manifold as a function of positions in $\mathbb{R}^N$. A function on the manifold $f(p)$ then can act effectively in the Euclidean space with $f \circ \phi^{-1}(x)$.

2.1.1 Differential forms and the wedge product

We define **vectors** as lying in the tangent space $\mathbf{T}_p\mathbb{M}$. We can write vectors as

$$\mathbf{V} = V_i \frac{\partial}{\partial x_i} + V_j \frac{\partial}{\partial x_j} + V_k \frac{\partial}{\partial x_k}.$$

in three dimensions in terms of coordinates in a chart at a particular position in the manifold. This is an example of a vector in $\mathbb{R}^3$.

Tangents to a curve give vectors. Curves are parametrized by a distance along a curve or time.

Another example of a vector is given by a trajectory $\dot{p}, \dot{q}$ with

$$\mathbf{V} = \dot{q}\frac{\partial}{\partial q} + \dot{p}\frac{\partial}{\partial p}.$$

This is an example of a vector in phase space.

We can define a cotangent space $\mathbf{T}_p^*\mathbb{M}$ dual to the tangent space. An element in the cotangent space gives a map from the tangent space to $\mathbf{T}_p\mathbb{M} \to \mathbb{R}$. An element in the cotangent space acts on a vector in the tangent space via a dot product.

For example, a differential of a function f is an example of an element of the cotangent space. The gradient $df \in \mathbf{T}_p^*\mathbb{M}$ and acts on a vector $V \in \mathbf{T}_p\mathbb{M}$ in the tangent space giving a real number

$$\langle df, \mathbf{V} \rangle = V^i \frac{\partial f}{\partial x_i} \in \mathbb{R}.$$

df is called a *one form*. In three dimensions we can write

$$df = \frac{\partial f}{\partial x_i} dx^i + \frac{\partial f}{\partial x_j} dx^j + \frac{\partial f}{\partial x_k} dx^k$$

In a tangent space we have the tangent of a curve. Trajectories on the manifold give tangent vectors. In the cotangent space we have the gradient of a function.

The directional derivative in the direction of vector $\mathbf{A} = a_i \frac{\partial}{\partial x^i}$ of a function f

$$a_i \frac{\partial f}{\partial x^i} = \langle A, df \rangle$$

is the gradient of a function f in the direction of vector $\mathbf{A}$. This can be described as the Lie derivative of a function

$$\mathcal{L}_A f = a_i \frac{\partial f}{\partial x^i}. \tag{13}$$

One forms are members of the *cotangent space* at a point p or $\mathbf{T}_p^*\mathbb{M}$. As the tangent bundle $\mathbf{T}\mathbb{M}$ is formed of $\mathbb{M}$ and its tangent spaces $\mathbf{T}_q\mathbb{M}$, the cotangent bundle $\mathbf{T}^*\mathbb{M}$ is formed of $\mathbb{M}$ and its cotangent spaces $\mathbf{T}_p^*\mathbb{M}$.

Differential forms can be integrated. A one-form can be integrated along a path with a sum of the values of the form along a bunch of tangent vectors to the path.

A q form maps q vectors in $T_pM \to \mathbb{R}$ and can be described as sums of products of one forms, for example

$$a_{ijk} dx^i \otimes dx^j \otimes dx^k$$

is 3 form.

Consider two vectors

$$\mathbf{V} = v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y}$$
$$\mathbf{W} = w_x \frac{\partial}{\partial x} + w_y \frac{\partial}{\partial y}$$

and the two form $\omega = dx \otimes dy$. We operate on $\mathbf{V}, \mathbf{W}$ with the two form,

$$\omega(\mathbf{V}, \mathbf{W}) = v_x w_y.$$

Another example. Let $\omega = dx \otimes dy - 3x dy \otimes dx$.

$$\omega(\mathbf{V}, \mathbf{W}) = v_x w_y - 3x v_y w_x.$$

The wedge product, $\wedge$, of one forms is an antisymmetric sum

$$dx^i \wedge dx^j = dx^i \otimes dx^j - dx^j \otimes dx^i$$

$$\begin{aligned} dx^i \wedge dx^j \wedge dx^k &= dx^i \otimes dx^j \otimes dx^k + dx^j \otimes dx^k \otimes dx^i + dx^k \otimes dx^i \otimes dx^j \\ &\quad - dx^i \otimes dx^k \otimes dx^j - dx^j \otimes dx^i \otimes dx^k - dx^k \otimes dx^j \otimes dx^i \end{aligned}$$

Another example. Antisymmetry implies $dx \wedge dy = -dy \wedge dx$ and $dx \wedge dy \wedge dz = -dx \wedge dz \wedge dy$.

With two form

$$\omega = dx^i \wedge dx^j = dx^i \otimes dx^j - dx^j \otimes dx^i$$

we compute

$$\omega(\mathbf{V}, \mathbf{W}) = v_i w_j - v_j w_i.$$

A *differential form* is an antisymmetric q form

$$\omega = \frac{1}{r!} \omega_{\mu_1 \ldots \mu_r} dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r}$$

Differential forms can be thought of as volume elements (think volume of a parallelpiped calculated from vectors). The *exterior derivative*

$$d\omega = \frac{1}{r!} \left(\frac{\partial}{\partial x_\nu} \omega_{\mu_1 \ldots \mu_r} \right) dx^\nu \wedge dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r}$$

The exterior derivate takes an r differential form and gives back an $r+1$ differential form. The exterior derivative gives the boundary of a volume element. As the exterior derivative is antisymmetric

$$d^2 \omega = 0.$$

This is related to the fact that in three dimensions $\nabla \times \nabla f = 0$ and $\nabla \cdot \nabla \times f = 0$.

If $\omega = w_x dx + w_y dy$ then

$$d\omega = \left(\frac{\partial w_y}{\partial x} - \frac{\partial w_x}{\partial y} \right) dx \wedge dy.$$

This is reminiscent of the cross product.

The geometric formulation is independent of the coordinate system used in the charts. The coefficients of a vector or a differential-form transform as a tensor.

A form ω is *exact* if there is a form θ such that $d\theta = \omega$.

A form is *closed* if $d\omega = 0$. Every exact form is closed, as $d^2\theta = 0$, but not every closed form is exact.

The generalized version of Stokes' theorem relates the integral of a form ω over the boundary of a region ∂C to the exterior derivative $d\omega$ and the region C.

$$\int_C d\omega = \int_{\partial C} \omega$$

Compare the above to Stokes' theorem in three dimensions

$$\int_A \boldsymbol{\nabla} \times \mathbf{F} \cdot d\mathbf{A} = \int_S \mathbf{F} \cdot d\mathbf{S}$$

where A is a surface bounded by a loop $S = \partial A$. Stokes' theorem is also equivalent to Gaus' law

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{F} dV = \int_A \mathbf{F} \cdot d\mathbf{A}$$

where V is a volume with boundary $A = \partial V$. In three dimensions the exterior derivative gives $\boldsymbol{\nabla}\times$ or $\boldsymbol{\nabla}\cdot$ depending on the dimension of the object that is being integrated. In three dimensions dS is a one form, dA is related to a two form and dV is related to a three form.

2.2 The Symplectic form

The one form

$$\theta = q_i dp^i$$

has exterior derivative

$$\omega = d\theta = dq^i \wedge dp^i$$

which is a two form. A manifold with such a two-form (that is not degenerate) is known as a *symplectic* manifold. Since ω is a derivative of θ

$$d\omega = d^2\theta = 0.$$

The symplectic form is non-degenerate and exact. (A form ω is *exact* if there is a form θ such that $d\theta = \omega$).

What does it mean to be non-degenerate? A two-form maps two vectors, $\boldsymbol{\eta}, \boldsymbol{\xi}$ to a real number. For every $\boldsymbol{\eta} \neq 0$ there exists a $\boldsymbol{\xi}$ such that $\omega(\boldsymbol{\eta}, \boldsymbol{\xi}) \neq 0$.

The symplectic form is connected with areas. Consider two vector fields $\mathbf{V}, \mathbf{W}$ at a point $\mathbf{q}, \mathbf{p}$.

$$
\begin{aligned}
\mathbf{V} &= v_{qi}\frac{\partial}{\partial q_i} + v_{pi}\frac{\partial}{\partial p_i} \\
\mathbf{W} &= w_{qi}\frac{\partial}{\partial q_i} + w_{pi}\frac{\partial}{\partial p_i}
\end{aligned}
$$

$$
\begin{aligned}
\omega(\mathbf{V}, \mathbf{W}) &= \left(v_{qi}\frac{\partial}{\partial q_i} + v_{pi}\frac{\partial}{\partial p_i}\right) dq^i \left(w_{qi}\frac{\partial}{\partial q_i} + w_{pi}\frac{\partial}{\partial p_i}\right) dp^i \\
&\quad - \left(v_{qi}\frac{\partial}{\partial q_i} + v_{pi}\frac{\partial}{\partial p_i}\right) dp^i \left(w_{qi}\frac{\partial}{\partial q_i} + w_{pi}\frac{\partial}{\partial p_i}\right) dq^i \\
&= v_{qi}w_{pi} - v_{pi}w_{qi}
\end{aligned}
$$

For each (q_i, p_i) pair we have the area of the parallelogram defined by (v_{qi}, w_{pi}) and (w_{qi}, w_{pi}).[1] The total is the sum of the areas of the n parallelograms. Is this is related to Liouville's volume theorem? Yes, as we will show the symplectic form is preserved by Hamiltonian flows.

Consider the two form $\omega = dq \wedge dp$ in a new coordinate system Q, P so that $q(P,Q), p(P,Q)$. We compute

$$
\begin{aligned}
dq &= \frac{\partial q}{\partial Q}dQ + \frac{\partial q}{\partial P}dP \\
dp &= \frac{\partial p}{\partial Q}dQ + \frac{\partial p}{\partial P}dP.
\end{aligned}
$$

Inserting these into $\omega = dq \wedge dp$

$$
\begin{aligned}
\omega &= \left(\frac{\partial q}{\partial Q}dQ + \frac{\partial q}{\partial P}dP\right) \wedge \left(\frac{\partial p}{\partial Q}dQ + \frac{\partial p}{\partial P}dP\right) \\
&= \left(\frac{\partial q}{\partial Q}\frac{\partial p}{\partial P} - \frac{\partial q}{\partial P}\frac{\partial p}{\partial Q}\right) dQ \wedge dP \\
&= \{q, p\}|_{P,Q}\, dQ \wedge dP,
\end{aligned}
$$

where I have written $\{q,p\}|_{P,Q}$ as the Poisson brackets computed with P, Q. If the coordinate transformation is canonical then the two form can be written

$$
\omega = dQ \wedge dP
$$

Canonical transformations preserve the two form ω.

[1] If $\mathbf{A}$, $\mathbf{B}$ are vectors in 3 dimensions, the area of a parallelogram spanned by these two vectors is $|\mathbf{A}\times\mathbf{B}|$. If $\mathbf{A}, \mathbf{B}$ are on the x, y plane, the z component of the cross product gives the area.

2.3 Generating Functions Geometrically

Consider p, q and P, Q both canonical sets of coordinates. We can look at the one forms

$$\theta_1 = p_i dq_i \qquad \theta_2 = P_i dQ_i$$

Because the two sets of coordinates are canonical, the two form

$$\omega = d\theta_1 = d\theta_2 = dp_i \wedge dq_i = dP_i \wedge dQ_i$$

Consider

$$\theta_1 - \theta_2 = p_i dq_i - P_i dQ_i$$

Because $dp_i \wedge dq_i = dP_i \wedge dQ_i$ we know that

$$d(\theta_1 - \theta_2) = 0$$

If we can find a function F (the generating function) such that

$$p_i dq_i - P_i dQ_i = dF$$

then we ensure that the coordinate transformation is canonical. By specifying that the difference $\theta_1 - \theta_2$ is the differential of a function we guarantee that $dp_i \wedge dq_i = dP_i \wedge dQ_i$ and so that the transformation is canonical. F serves as our generating function. If we start with the generating function and then use it to give us the coordinate transformation then we ensure that the coordinate transformation is canonical.

Let our difference in forms be

$$pdq - PdQ = dF_1(q, Q) = \frac{\partial F_1(q, Q)}{\partial q} dq + \frac{\partial F_1(q, Q)}{\partial Q} dQ. \tag{14}$$

We associate

$$p = \frac{\partial F_1(q, Q)}{\partial q} \qquad P = -\frac{\partial F_1(q, Q)}{\partial Q}$$

so equation 14 cancels to zero and the transformation is canonical.

Let

$$qdp - QdP = dF_4(p, P) = \frac{\partial F_4(p, P)}{\partial p} dp + \frac{\partial F_4(p, P)}{\partial P} dP.$$

We associate

$$q = \frac{\partial F_4(p, P)}{\partial p} \qquad Q = -\frac{\partial F_4(p, P)}{\partial P}$$

Let

$$pdq + QdP = dF_2(q, P)$$

The plus sign here arises because

$$d(pdq + QdP) = dp \wedge dq + dQ \wedge dP = dp \wedge dq - dP \wedge dQ$$

and this we would require is zero for the transformation to be canonical.

$$dF_2(q, P) = \frac{\partial F_2(q, P)}{\partial q} dq + \frac{\partial F_2(q, P)}{\partial P} dP$$

giving

$$p = \frac{\partial F_2(q, P)}{\partial p} \qquad Q = \frac{\partial F_2(q, P)}{\partial P}$$

for a canonical transformation.

Remark The sign of the generating functions can be flipped and the transformation is still canonical.

Given a transformation $Q(q, p)$ and $P(q, p)$ the transformation is canonical (symplectic) if one of the following forms is *exact*

$$\begin{aligned} \sigma_1 &= pdq - PdQ \\ \sigma_2 &= pdq + QdP \\ \sigma_3 &= qdp + PdQ \\ \sigma_4 &= qdp - QdP. \end{aligned}$$

For the different related generating functions

$$\begin{aligned} \frac{\partial F_1(q, Q)}{\partial q} &= p \qquad \frac{\partial F_1(q, Q)}{\partial Q} = -P \\ \frac{\partial F_2(q, P)}{\partial q} &= p \qquad \frac{\partial F_2(q, P)}{\partial P} = Q \\ \frac{\partial F_3(p, Q)}{\partial p} &= q \qquad \frac{\partial F_3(p, Q)}{\partial Q} = P \\ \frac{\partial F_4(p, P)}{\partial p} &= q \qquad \frac{\partial F_4(p, P)}{\partial P} = -Q. \end{aligned}$$

2.4 Vectors generate Flows and trajectories

Consider a vector field, $\mathbf{X}$, and a curve on the manifold, $\sigma(t, x)$, that is a map $\mathbb{R} \times \mathbb{M} \to \mathbb{M}$ such that the tangent vector at each point, x, on the curve is $\mathbf{X}$. An integral curve $\sigma(t, x_0)$ that goes through a point x_0 on the manifold at $t = 0$ has tangent

$$\left. \frac{d}{dt} \sigma^\mu(t, x_0) \right|_{t=0} = X^\mu$$

I have used local coordinates. The initial condition $\sigma^\mu(t=0,x_0)=x_0^\mu$. Now instead of requiring σ to only contain a single curve going through x_0 we extend σ so that it includes curves that go through all the points on the manifold. The map σ is known as a *flow* generated by the vector field $\mathbf{X}$.

We can define an exponential function $\exp(Xt)$ from $\mathbb{M}$ to $\mathbb{M}$ so that we can move along the flow from one point to another, $\sigma^\mu(t,x)=\exp(Xt)x^\mu$.

A vector field generates a flow all over the manifold.

2.5 The Hamiltonian flow and the symplectic two-form

The symplectic two-form $\omega = dq\wedge dp$ is a map from two vectors to a real number. Consider two vectors $\mathbf{V},\mathbf{W}$,

$$\begin{aligned}\mathbf{V} &= v_q\frac{\partial}{\partial q}+v_p\frac{\partial}{\partial p}\\ \mathbf{W} &= w_q\frac{\partial}{\partial q}+w_p\frac{\partial}{\partial p}\end{aligned}$$

recall that

$$\omega(V,W)=v_qw_p-w_qv_p$$

Now what if we consider the map only using vector V

$$\omega(V,?)=v_qdp-v_pdq \tag{15}$$

This gives a one-form. In this way the symplectic two-form gives us a way to take a vector (in the tangent space) and generate a one-form (in the cotangent space) from it. Conversely using the symplectic form ω, we can take a one-form and construct a vector from it, using equation 15.

Remark By providing an invertible map between vectors and one-forms, ω serves like the metric tensor in general relativity or Riemannian geometry. However the map is antisymmetric rather than symmetric and so it is not a metric.

Recall that by specifying a direction, a vector *field* generates a flow, or given an initial condition, a trajectory. We can produces trajectories $(\dot q,\dot p)$ from a function (H) using this map.

Consider the one form associated with a Hamiltonian, $H(q,p)$ and use ω to generate a vector from it.

$$\begin{aligned}dH &= \frac{\partial H}{\partial q}dq+\frac{\partial H}{\partial p}dp\\ &= \omega(V,?)\\ &= v_qdp-v_pdq\end{aligned}$$

On the second line we have used vector V (as written out in equation 15). The V vector has components

$$v_q = \frac{\partial H}{\partial p} \qquad v_p = -\frac{\partial H}{\partial q}$$

These are associated with $\dot{q}$ and $\dot{p}$ for the flow. Let us define (as Arnold) the vector generated by dH as IdH. The equations of motion (Hamilton's equations) are equivalent to $\dot{\mathbf{x}} = IdH$ In this way, a Hamiltonian *flow* is generated from the Hamiltonian *function*.

2.6 Extended Phase space

The statements relating the symplectic two-form to Hamiltonian flows can be illustrated with the one-form (known as the Poincaré-Cartan integral invariant)

$$\sigma^1 = pdq - Hdt$$

in what is known as *extended phase space* or a $\mathbb{R}^{2n+1}$ space that consists of phase space ($\mathbb{R}^{2n}$) with addition of time or $(\mathbf{p}, \mathbf{q}, t)$. The exterior derivative $d\sigma^1$ is a two form. The odd dimension implies that there always exists some direction (a vector V)

$$\exists V \text{ such that} \qquad d\sigma^1(V, \eta) = 0 \qquad \text{for all vectors } \eta.$$

The direction, V, we associate with the equations of motion. Flows along this direction are also called **vortex lines**. The exterior derivative

$$d\sigma^1 = dp \wedge dq - \frac{\partial H}{\partial q} dq \wedge dt - \frac{\partial H}{\partial p} dp \wedge dt \tag{16}$$

Consider the vector

$$V = -\frac{\partial H}{\partial q}\frac{\partial}{\partial p} + \frac{\partial H}{\partial p}\frac{\partial}{\partial q} + \frac{\partial}{\partial t}$$

Let us insert this into $d\sigma^1$

$$\begin{aligned} d\sigma^1(V, *) &= -\frac{\partial H}{\partial q} dq - \frac{\partial H}{\partial p} dp \qquad (\text{from } \; dp \wedge dq) \\ &\quad + \frac{\partial H}{\partial p} dp + \frac{\partial H}{\partial q} dq \qquad (\text{from } \frac{\partial}{\partial t}) \\ &\quad - \frac{\partial H}{\partial q}\frac{\partial H}{\partial p} dt + \frac{\partial H}{\partial q}\frac{\partial H}{\partial p} dt \\ &= 0 \end{aligned}$$

So our vector V is in fact the direction defining the vortex lines. But look again at the vector V

$$\begin{aligned} V &= -\frac{\partial H}{\partial q}\frac{\partial}{\partial p} + \frac{\partial H}{\partial p}\frac{\partial}{\partial q} + \frac{\partial}{\partial t} \\ &\to \dot{p}\frac{\partial}{\partial p} + \dot{q}\frac{\partial}{\partial q} + \frac{\partial}{\partial t} \end{aligned}$$

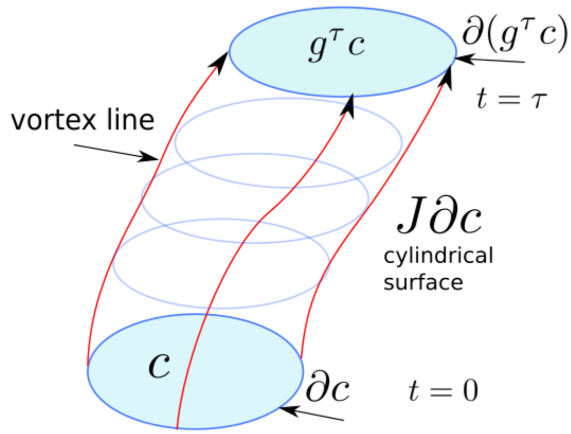

Figure 2: Vortex lines are shown in red and form the cylinder boundary. The map J takes base surface c at time $t = 0$ to top surface $g^\tau c$ at $t = \tau$.

The flow generated by vector V satisfies Hamilton's equations.

2.7 Hamiltonian flows preserve the symplectic two-form

Here we are going to use the generalized Stokes' theorem in extended phase space to prove that the symplectic two-form is conserved by a Hamiltonian flow.

Consider a surface region in (p, q) (call this c) and its Hamiltonian flow along t, from $t = 0$ to a time $t = \tau$ causing a transformation $g^\tau c$ (see Figure 2). The boundary of the surface at $t = 0$ would be ∂c. Boundary at $t = \tau$ would be $\partial g^\tau c$. Flow in time from $t = 0$ to τ is created with an operator J. Every point on c is given a trajectory from $t = 0 \to \tau$ using J. These trajectories are vortex lines and so have tangent equal to the vector V such that $d\sigma(V) = 0$. The total volume of the cylinder is Jc. The sides of the cylinder are covered by $J\partial c$. The boundary of Jc or $\partial(Jc)$ is the sum of the top and lower faces, c and $g^\tau c$, and the sides, $J\partial c$.

The generalized version of Stokes' theorem gives

$$\int_c d\omega = \int_{\partial c} \omega$$

The integral of a differential form of a boundary of a region in an orientable manifold ∂c is equivalent to the integral of the exterior derivative of the form $d\omega$ in the region c.

One consequence of the generalized version of Stokes' theorem is that the integral of an exact form over the boundary of a region is zero. In other words we apply Stokes' theorem

$$\int_{\partial c} d\sigma = \int_c d^2\sigma = 0$$

where the last step is zero because $d^2 = 0$.

Let us integrate the exterior derivative of $d\sigma^1$ over the surface of the cylindrical volume $\partial(Jc)$ or

$$\int_{\partial(Jc)} d\sigma^1$$

Because of Stokes' theorem

$$\int_{\partial(Jc)} d\sigma^1 = \int_{Jc} d^2\sigma^1 = 0$$

Now we divide $\partial(Jc)$ into three pieces, the bottom of the cylinder c, the top of the cylinder $g^t c$ and the cylindrical surface $J\partial c$. The sum of these three integrals of $d\sigma$ must be zero.

Now we perform this sum for the exact two-form $d\sigma^1$ with $\sigma^1 = pdq - H$. Integrating $d\sigma^1$ on the surface $J\partial c$ gives zero because the surface is comprised of vortex lines or null vectors. Consequently integrating $d\sigma^1$ on the top and both of the cylinder c and $g^\tau c$ must give the same result. On these surfaces the vectors are perpendicular to $\frac{\partial}{\partial t}$ so $d\sigma^1 = \omega = dp \wedge dq$ the symplectic two-form (see equation 16) and we find that g^t preserves the integral of the symplectic two-form. Since the symplectic two-form gives the volume in phase space, this is equivalent to Liouville's theorem and also implies that a transformations generated by a Hamiltonian flow are canonical transformations.

Previously we showed that a Hamiltonian flow preserved phase space volume (Liouville's theorem). Phase space volume conservation can also be written in terms of ω, hence Liouville's theorem is equivalent to saying that Hamiltonian flows preserve the two-form ω. Because ω is preserved by a Hamiltonian flow, the flow also generates canonical transformations between coordinates at any two times.

2.8 The symplectic two-form and the Poisson bracket

We ask, if canonical transformations preserve the Poisson bracket and they preserve the two-form ω what is the relation between the Poisson bracket and the symplectic two form?

The Poisson bracket takes derivatives of functions.

$$\begin{aligned}
\{g, h\} &= \frac{\partial g}{\partial q_i}\frac{\partial h}{\partial p_i} - \frac{\partial g}{\partial p_i}\frac{\partial h}{\partial q_i} \\
&= \boldsymbol{\nabla}_x g^t \mathbf{J} \boldsymbol{\nabla}_x h \\
&= \omega(\boldsymbol{\nabla}_x g, \boldsymbol{\nabla}_x h)
\end{aligned}$$

where $\mathbf{x} = (\mathbf{q}, \mathbf{p})$ and we consider $\boldsymbol{\nabla}_x g, \boldsymbol{\nabla}_x h$ as vectors. Here

$$\mathbf{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

and $\omega = dq \wedge dp$ is the symplectic two form. The above is true only in a canonical basis where the two-form provides a direct and trivial way to convert between one-forms and vectors. Really the gradient of a function should be considered a one-form, not a vector.

What if we write the two-form in a non-canonical basis with

$$\omega = f_{ij} dx^i \otimes dx^j$$

and x variables are functions of p, q a canonical set. Here f_{ij} an antisymmetric matrix that effectively gives the wedge product. Now we use the two-form ω to convert between the differential forms of g, h and vectors $\mathbf{V}, \mathbf{W}$ with $\mathbf{V} = v^i \frac{\partial}{\partial x^i}$ and $\mathbf{W} = w^i \frac{\partial}{\partial x^i}$. We assert that ω operating on $\mathbf{V}$ gives us dg,

$$\begin{aligned} \omega(V, ?) &= dg = \frac{\partial g}{\partial x^j} dx^j \\ v^i f_{ij} dx^j &= \frac{\partial g}{\partial x^j} dx^j \\ v^i f_{ij} &= \frac{\partial g}{\partial x^j}. \end{aligned} \tag{17}$$

Likewise we do the same for h but in the second location in the two-form

$$\begin{aligned} \omega(?, W) &= dh = \frac{\partial h}{\partial x^j} dx^j \\ w^j f_{ji} dx^i &= \frac{\partial h}{\partial x^i} dx^i \\ w^j f_{ji} &= \frac{\partial h}{\partial x^i}. \end{aligned} \tag{18}$$

Since the symplectic two-form is not degenerate we can invert the matrix f_{ij}. We call the inverse of the matrix F. It satisfies

$$F^{ij} f_{jk} = \delta^i_k.$$

Both f and F are antisymmetric matrices. We invert the above relations

$$\begin{aligned} \frac{\partial g}{\partial x^j} F^{jk} &= v^i f_{ij} F^{jk} \qquad \text{(using eqn 17)} \\ &= v^i \delta_i^k = v^k. \end{aligned} \tag{19}$$

Likewise

$$w^i = F^{ji} \frac{\partial h}{\partial x^j} \qquad \text{(using eqn 18).}$$

Now let us compute the two-form on these two vectors

$$\begin{aligned}\omega(V,W) &= \frac{\partial g}{\partial x^j}F^{jk}f_{ki}F^{il}\frac{\partial h}{\partial x^l}\\ &= \frac{\partial g}{\partial x^j}\delta^j_i F^{il}\frac{\partial h}{\partial x^l}\\ &= \frac{\partial g}{\partial x^i}F^{il}\frac{\partial h}{\partial x^l}\end{aligned}$$

The expression on the right we recognize as similar to a Poisson bracket. So it makes sense to define

$$\{g,h\} = \frac{\partial g}{\partial x^j}F^{jl}\frac{\partial h}{\partial x^l}$$

in the non-canonical basis. In this sense we can consider the two-form like an inverse of the Poisson bracket.

This form of the Poisson bracket is sometimes seen in classical field (continuum) models.

Remark It is possible to construct a Poisson bracket that cannot be converted into a two-form. For example in 3 dimensions, the Poisson bracket

$$\{f,g\} = \epsilon_{ijk}x_k\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j} \tag{20}$$

cannot be inverted as a three dimensional anti-symmetric matrix has no inverse.

2.8.1 On connection to the Lagrangian

Integrating the one form $\sigma^1 = pdq - Hdt$ is like integrating the Lagrangian

$$\mathcal{L}dt = (p\dot{q} - H)dt$$

So σ^1 gives us a way to generate the action for the associated Lagrangian.

2.8.2 Discretized Systems and Symplectic integrators

Symplectic integrators provide maps from phase space to phase space separated by a time Δt. These are symplectic transformations, preserving phase space volume and the symplectic two-form.

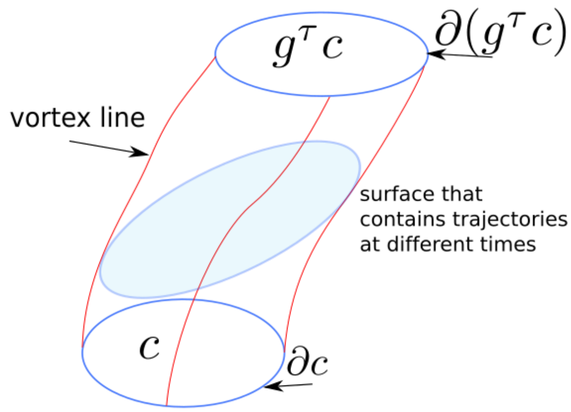

Figure 3: Vortex lines going through a tilted plane. A two-form can be constructed with degrees of freedom lying in the tilted plane. Because the vortex lines give no area, this two-form is equivalent to one at a single time.

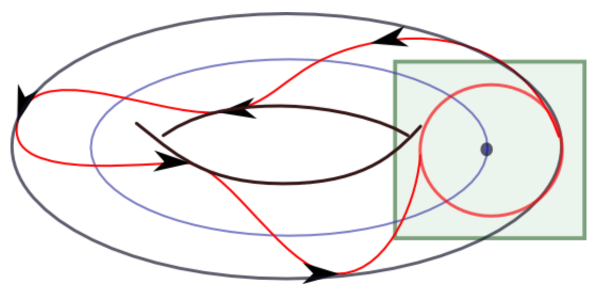

Figure 4: Orbits covering a torus. A map can be constructed from the position of the orbits each time they pass through a plane. This is known as a surface of section and is another way to create an area preserving map from a Hamiltonian flow in 4-dimensional phase space. The flow lines are vortex lines and so are null vectors with respect to the symplectic two-form in extended phase space. On a plane that slices the torus, one piece of the two form can be set to zero. The remain degrees of freedom (in that plane) give a two form that is preserved by the map. If the Hamiltonian is time independent, the time dependent part of the two form is zero. The map is symplectic. Note that the time it takes the red orbit to start from the green plane and return to it can differ from the time it takes the blue orbit to start from this plane and return to it.

2.8.3 Surfaces of section

We can also consider surfaces at an angle in the flow of vortex lines (see Figure 3) and a two form computed at the p, q on this surface but at different times. We can consider maps generated from between the times it takes to cross a planar subspace (see Figure 4). There are a number of ways to generate area or volume preserving maps from Hamiltonian flows. The key point here is that the vortex lines are flow lines and these are null vectors with respect to the symplectic form.

The illustration in Figure 3 is useful for constructing surfaces of section for time dependent Hamiltonians. The illustration in Figure 4 is used for constructing surfaces of section in the restricted 3 body problem.

2.9 The Hamiltonian following a canonical transformation

We defined a canonical transformation as one that satisfied the Poisson brackets or equivalently preserved the symplectic 2-form. We did not require that Hamilton's equations were preserved, or even required a Hamiltonian to determine whether a coordinate transformation was canonical. However we showed that Hamilton's equations were satisfied using a new Hamiltonian in the old coordinates. This is true as long as coordinate transformation is time independent.

What happens if the canonical transformation is time-dependent?

For any function $f(p, q, t)$ recall that

$$\dot{f} = \{f, H\} + \frac{\partial f}{\partial t}$$

Following a time dependent canonical transformation $Q(p, q, t), P(q, p, t)$, (satisfying Poisson brackets) we can insert Q or P finding

$$\dot{Q} = \{Q, H\} + \frac{\partial Q}{\partial t} \qquad \dot{P} = \{P, H\} + \frac{\partial P}{\partial t}$$

A new Hamiltonian is required so that the new coordinates satisfy Hamilton's equations. We need to find a function K such that

$$\dot{Q} = \{Q, K\} \qquad \dot{P} = \{P, K\}$$

Consider an extended phase space defined by q, p, t. We are adding time as an extra dimension. We can construct a form

$$\theta = pdq - Hdt + QdP + Kdt. \tag{21}$$

We previously defined a one form

$$\sigma_a = pdq - Hdt$$

such that $d\sigma_a(V,\eta) = 0$ for V giving the Hamiltonian flow from H for all vectors η. We can similarly define a one form

$$\sigma_b = -QdP - Kdt$$

such that $d\sigma_b(V,\eta) = 0$ for V giving Hamiltonian flow from K for all vectors η. This follows from the same argument given previously in section 2.6 above as $d^2\sigma_b = 0$. The one form from equation 21

$$\theta = \sigma_a - \sigma_b.$$

So if $d\theta = 0$ then if Hamiltonian flow for H implies that Hamiltonian flow for K is obeyed.

Let $\theta = dF_2$ where $F_2(q, P, t)$ is a generating function that now depends on t. Remember we are now working in extended phase space so

$$dF_2(q,P,t) = \frac{\partial F_2}{\partial q}dq + \frac{\partial F_2}{\partial P}dP + \frac{\partial F_2}{\partial t}$$

We match

$$\begin{aligned} p &= \frac{\partial F_2}{\partial q} \\ Q &= \frac{\partial F_2}{\partial P} \\ K - H &= \frac{\partial F_2}{\partial t} \end{aligned}$$

With these values our $\theta = dF_2$ is exact and the two Poincaré-Cartan invariants $d\sigma_a$, $d\sigma_b$ are the same and so have the same vortex lines and the *same* Hamiltonian flows.

Thus our new Hamiltonian (and one that by definition satisfies Hamilton's equations) is

$$K = H + \frac{\partial F_2}{\partial t}.$$

Following the same procedure for the other classes of generating functions we find

$$K = H + \frac{\partial F_1}{\partial t}$$

and similarly for F_3 and F_4 for the other classes of generating functions.

2.10 The Lie derivative

The Lie derivative evaluates the change of a scalar function, vector field or one-form, along the flow defined by another vector field. This change is coordinate invariant and so the Lie derivative is defined on any differentiable manifold. Once the Lie derivative is defined on a function and vector field, the derivative can be extended so is can be evaluated on any tensor field.

We express the action of the flow defined by vector field $\mathbf{X}$ in terms of coordinates $\mathbf{X} = X^\mu(\mathbf{x})\frac{\partial}{\partial x^\mu}$ for a local coordinate map $\mathbf{x} \to p$ where p is a point on the manifold. The Lie derivative of a function f is the same as the directional derivative

$$\mathcal{L}_X f(p) = \nabla_X f(p(x)) = X^\mu \frac{\partial f(p(x))}{\partial x^\mu}$$

How do we compute the Lie derivative of a vector? It may be helpful to compute some Lie derivatives of functions

$$\mathcal{L}_X(\mathcal{L}_Y f) = X^\mu \frac{\partial}{\partial x^\mu}\left(Y^\nu \frac{\partial f}{\partial x^\nu}\right) \tag{22}$$

$$\mathcal{L}_Y(\mathcal{L}_X f) = Y^\mu \frac{\partial}{\partial x^\mu}\left(X^\nu \frac{\partial f}{\partial x^\nu}\right) \tag{23}$$

$$\begin{aligned}(\mathcal{L}_X\mathcal{L}_Y - \mathcal{L}_Y\mathcal{L}_X)f &= \left(X^\mu \frac{\partial Y^\nu}{\partial x^\mu}\frac{\partial}{\partial x^\nu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}\frac{\partial}{\partial x^\nu}\right) f \\ &= \left(X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}\right)\frac{\partial}{\partial x^\nu} f \end{aligned} \tag{24}$$

We recognize the last thing as a vector which means we can associate

$$\mathcal{L}_X Y = (\mathcal{L}_X\mathcal{L}_Y - \mathcal{L}_Y\mathcal{L}_X) = [X, Y] = XY - YX \tag{25}$$

giving a vector. Here the vectors operate on each other, for example

$$XY = X^i \frac{\partial}{\partial x^i}\left(Y^j \frac{\partial}{\partial x^j}\right) \tag{26}$$

$$= X^i \frac{\partial Y^j}{\partial x^i}\frac{\partial}{\partial x^j} + X^i Y^j \frac{\partial^2}{\partial x^i \partial x^j}. \tag{27}$$

The Lie derivative can also be defined axiomatically, asserting that it commutes with the exterior derivative and obeys versions of the Leibniz rule.

Another definition of $\mathcal{L}_X Y$ is

$$\mathcal{L}_X Y = \frac{d}{dt}\phi^t_X Y\Big|_{t=0}$$

where ϕ^t_X is the flow caused by vector X.

The Lie derivative $\mathcal{L}_X Y$ of the Y vector field along the direction of the X vector field is a function of two vector fields. It computes the change of one vector along the direction of the other. The Lie derivative is not the same as a covariant derivative. Using a covariant derivative you can compute the change of a basis vector in a particle direction. It is not necessary to have two vector fields, however you do need to know how to transport the basis vectors of the coordinate system between neighboring regions of the manifold.

Once we have a derivative of a vector we can take the Lie derivative of a one form $\omega = \omega^i dx^i$ and a vector $\mathbf{V}$

$$\begin{aligned}
\langle \omega, \mathbf{V} \rangle &= V^i \omega^i \\
\mathcal{L}_X \langle \omega, \mathbf{V} \rangle &= \langle \mathcal{L}_X(\omega), \mathbf{V} \rangle + \langle \omega, \mathcal{L}_X \mathbf{V} \rangle \\
X^j \frac{\partial}{\partial x^j} \left(V^i \omega^i \right) &= \langle \mathcal{L}_X(\omega), \mathbf{V} \rangle + \langle \omega, \left(X^\mu \frac{\partial V^\nu}{\partial x^\mu} - V^\mu \frac{\partial X^\nu}{\partial x^\mu} \right) \frac{\partial}{\partial x^\nu} \rangle \\
&= \langle \mathcal{L}_X(\omega), \mathbf{V} \rangle + \omega^\nu \left(X^\mu \frac{\partial V^\nu}{\partial x^\mu} - V^\mu \frac{\partial X^\nu}{\partial x^\mu} \right) \qquad (28)
\end{aligned}$$

where we have used equations 25 and 27. Continuing,

$$\begin{aligned}
\langle \mathcal{L}_X(\omega), \mathbf{V} \rangle &= -\omega^\nu X^\mu \frac{\partial V^\nu}{\partial x^\mu} + \omega^\nu V^\mu \frac{\partial X^\nu}{\partial x^\mu} + X^j V^i \frac{\partial \omega^i}{\partial x^j} + X^j \omega^i \frac{\partial V^i}{\partial x^j} \\
&= \omega^\nu V^\mu \frac{\partial X^\nu}{\partial x^\mu} + X^j V^i \frac{\partial \omega^i}{\partial x^j}. \qquad (29)
\end{aligned}$$

We compute

$$d\omega = \frac{\partial \omega^i}{\partial x^j} dx^j \wedge dx^i \qquad (30)$$

We can contract a vector with a two form to get a one form

$$i_X d\omega = \langle X, d\omega \rangle = \left(X^i \frac{\partial \omega^j}{\partial x^i} - X^j \frac{\partial \omega^i}{\partial x^j} \right) dx^j \qquad (31)$$

Note the contraction notation $i_V \omega = \langle V, \omega \rangle = V^i \omega^i$. We can also compute

$$d(i_X \omega) = d\langle X^i \omega^i \rangle = \frac{\partial X^i}{\partial x^j} \omega^i dx^j + \frac{\partial \omega^i}{\partial x^j} X^i dx^j \qquad (32)$$

With some manipulation we can show that the Lie derivative of a one form is consistent with Cartan's magic formula or

$$\mathcal{L}_X(\omega) = i_X d\omega + d(i_X \omega) \qquad (33)$$

and that

$$\mathcal{L}_X(\omega) = \left(X^b \frac{\partial \omega^a}{\partial x^b} + \omega^b \frac{\partial X^b}{\partial x^a} \right) dx^a$$

is consistent with equation 29.

3 Examples of Canonical transformations

3.1 Orbits in the plane of a galaxy or around a massive body

The Keplerian problem of a massless particle in orbit about a massive object of mass M can be written in polar coordinates and restricted to a plane

$$\mathcal{L}(r,\theta;\dot{r},\dot{\theta}) = \frac{1}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - V(r)$$

where $V(r) = -GM/r$. The associated momentum are

$$\begin{aligned} p_r &= \frac{\partial \mathcal{L}}{\partial \dot{r}} = \dot{r} \\ p_\theta &= \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = r^2\dot{\theta} = L \end{aligned}$$

where L is the angular momentum. We find a Hamiltonian

$$H(r,\theta;p_r,L) = \frac{p_r^2}{2} + \frac{L^2}{2r^2} + V(r) \tag{34}$$

Because the Hamiltonian is independent of θ, the angular momentum is conserved.

Let us expand this Hamiltonian using $y = r - R_0$ and expand assuming that $y \ll R_0$.

$$\begin{aligned} H(y,\theta,p_r,L) &= \frac{p_r^2}{2} + \frac{L^2}{2(y+R_0)^2} + V(y+R_0) \\ &= \frac{p_r^2}{2} + y\left[V'(R_0) - \frac{L^2}{R_0^3}\right] + \frac{y^2}{2}\left[V''(R_0) + \frac{3L^2}{R_0^4}\right] \end{aligned}$$

Near a circular orbit the term proportional to y must be zero, otherwise $\dot{r} = \frac{\partial H}{\partial y} =$ is a constant and r will continue to increases or decrease. The angular momentum that sets the y term to zero is $L = R_0^2\Omega$ with $\Omega = \dot{\theta}$ and angular rotation rate

$$\Omega(R_0) = \sqrt{\frac{V'(R_0)}{R_0}}.$$

We identify

$$\kappa^2(R_0) = V''(R_0) + \frac{3L^2}{R_0^4} = V''(R_0) + 3\Omega(R_0)^2$$

as the frequency of radial oscillations or the **epicyclic frequency**.

The above Hamiltonian (equation 34) is useful to study dynamics of stars in the mid-plane of a galaxy but with a different potential $V(r)$. A circular orbit has velocity $v_c(r) = \sqrt{rV'(r)}$. Many galaxies have nearly flat rotation curves with $v_c(r) \sim v_c$ with v_c a constant, corresponding to a logarithmic potential $V(r) = v_c^2 \ln r$.

Above we used a Lagrangian in cylindrical coordinates to find the Hamiltonian system. However we could have started with a Hamiltonian in cartesian coordinates

$$H(x, p_x; y, p_y) = \frac{1}{2}(p_x^2 + p_y^2) + V(\sqrt{x^2 + y^2})$$

The transformation is not obviously canonical

$$\begin{aligned}
p_r &= \frac{p_x x}{r} + \frac{p_y y}{r} \\
L &= x p_y - y p_x \\
r &= \sqrt{x^2 + y^2} \\
\theta &= \mathrm{atan}(y/x)
\end{aligned}$$

however we can check that it is using Poisson brackets. Computing derivatives

$$\begin{aligned}
\frac{\partial p_r}{\partial x} = (p_x y - p_y x)\frac{y}{r^{3/2}} \\
\frac{\partial p_r}{\partial y} = (p_y x - p_x y)\frac{x}{r^{3/2}}
\end{aligned}$$

We evaluate the Poisson bracket

$$\begin{aligned}
\{p_r, L\} &= \frac{\partial p_r}{\partial x}\frac{\partial L}{\partial p_x} - \frac{\partial p_r}{\partial p_x}\frac{\partial L}{\partial x} + \frac{\partial p_r}{\partial y}\frac{\partial L}{\partial p_y} - \frac{\partial p_r}{\partial p_y}\frac{\partial L}{\partial y} \\
&= (p_x y - p_y x)\frac{y}{r^{3/2}}(-y) - \frac{x}{r}p_y + (p_y x - p_x y)\frac{x}{r^{3/2}}x + \frac{y}{r}p_x \\
&= \frac{L}{r} - \frac{L}{r} \\
&= 0
\end{aligned}$$

and likewise for the other brackets.

3.2 Epicyclic motion

Orbits in a disk galaxy are nearly circular. The epicylic approximation assumes that the orbit can be described by a radial oscillation around a circular orbit. In the Keplerian setting the radial oscillation period is the same as the orbital period. But in the galactic setting radial oscillations are faster that the rotation and the orbit does not close (see Figure 5). By setting $p_r = 0$, we can define a function $E(L)$ giving the energy of a circular orbit and this can be inverted $L(E)$ to give the angular momentum of a circular orbit with energy E. We can also define a function $r_c(L)$ that gives the radius of a circular orbit with angular momentum L. These are related by

$$E(L) = \frac{L^2}{2r_c(L)^2} + V(r_c(L)) \tag{35}$$

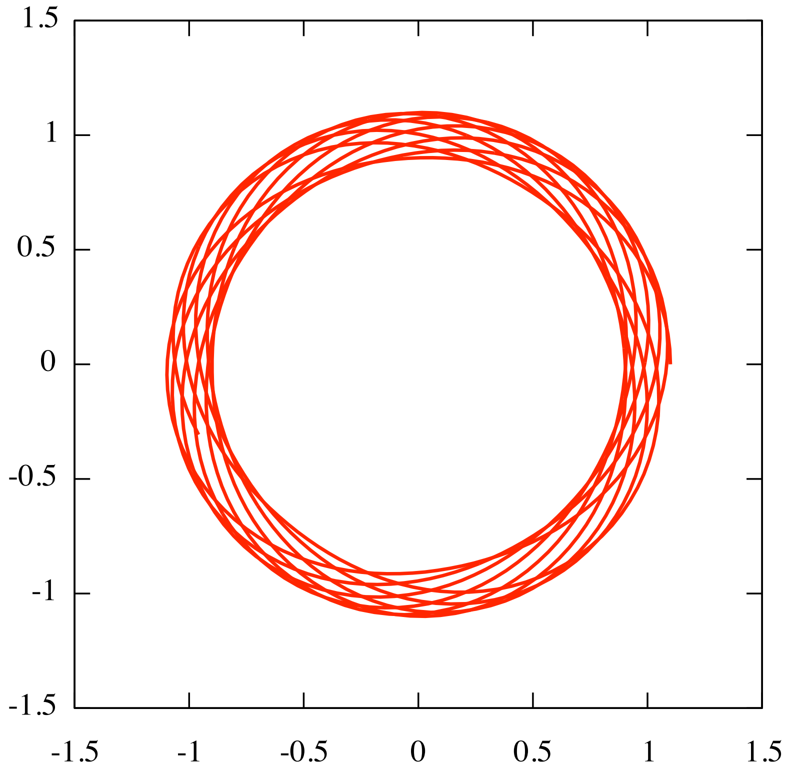

Figure 5: A rosette orbit. The orbit can be described in terms of radial oscillations or epicycles about a circular orbit.

$$
\begin{aligned}
L &= r_c^2(L)\Omega(L) \\
\Omega(L) &= \sqrt{\frac{V'(r_c(L))}{r_c(L)}}
\end{aligned} \tag{36}
$$

where $\Omega(L)$ is the angular rotation rate $\dot{\theta}$ for the circular orbit with angular momentum L.

Example: For a flat rotation curve what is $\Omega(L)$ and $r_c(L)$? For the flat rotation curve $v_c = r_c(L)\Omega(L)$ is constant. The angular momentum $L = r_c(L)v_c$, consequently

$$
r_c(L) = \frac{L}{v_c}
$$

and

$$
\Omega(L) = \frac{v_c}{r_c(L)} = \frac{v_c^2}{L}
$$

The energy

$$
E(L) = \frac{L^2}{2r_c(L)^2} + V(r_c) = v_c^2 \ln L + \text{constant}
$$

We can consider orbits with angular momentum and energy near those of the circular orbit. Consider the following generating function that is a function of old momenta (p_r, L) and new coordinates (θ_r, θ_{new})

$$
F_3(p_r, L; \theta_r, \theta_{new}) = \frac{p_r^2}{2\kappa(L)} \cot\theta_r - r_c(L)p_r - L\theta_{new}
$$

The canonical transformation gives

$$
\begin{aligned}
r &= -\frac{\partial F_3}{\partial p_r} = r_c(L) + \frac{p_r}{\kappa(L)} \cot\theta_r \\
J_r &= \frac{\partial F_3}{\partial \theta_r} = -\frac{p_r^2}{2\kappa(L)} \sin^{-2}\theta_r \\
\theta &= -\frac{\partial F_3}{\partial L} = \theta_{new} - \frac{dr_c(L)}{dr} p_r + \frac{\frac{\kappa(L)}{dL} p_r^2}{2\kappa^2(L)} \cot\theta_r \\
L_{new} &= -\frac{\partial F_3}{\partial \theta_{new}} = L
\end{aligned}
$$

We can rewrite this so that old coordinates are written in terms of new ones that can be directly inserted into the Hamiltonian

$$\begin{aligned}
r &= r_c(L) + \sqrt{\frac{2J_r}{\kappa(L)}} \cos\theta_r \\
p_r &= -\sqrt{2J_r\kappa(L)} \sin\theta_r \\
L &= L_{new} \\
\theta &= \theta_{new} - \frac{dr_c(L)}{dL}\sqrt{2J_r\kappa(L)} \sin\theta_r + \frac{\frac{d\kappa(L)}{dL} J_r}{2\kappa(L)} \cos(2\theta_r)
\end{aligned}$$

We insert these new coordinates into the Hamiltonian

$$H(r,\theta; p_r, L) = \frac{p_r^2}{2} + \frac{L^2}{2r^2} + V(r) \tag{37}$$

We temporarily define $r = r_c(L) + \delta r$ with $\delta r = \sqrt{\frac{2J_r}{\kappa(L)}} \cos\theta_r$ and expand to second order in δr;

$$\begin{aligned}
H(\theta_r, \theta_{new}, J_r, L) = V(r_c) + \frac{L^2}{2r_c(L)^2} + \left(\frac{dV(r_c)}{dr} - \frac{L^2}{r_c(L)^3}\right)\delta r \\
+ \left(\frac{d^2V(r_c)}{dr^2} + \frac{3L^2}{r_c(L)^4}\right)\frac{\delta r^2}{2} + J_r\kappa(L)\sin^2\theta_r
\end{aligned} \tag{38}$$

The term dependent on δr cancels due to the definition of $r_c(L)$. The term independent of δr

$$\begin{aligned}
g_0(L) &\equiv V(r_c(L)) + \frac{L^2}{2r_c(L)^2} \\
\frac{dg_0(L)}{dL} &= \Omega(L)
\end{aligned} \tag{39}$$

consistent with equation 36 and the definition of the angular rotation rate $\Omega(L)$. Using our definition for $g_0(L)$ and δr the Hamiltonian becomes

$$H(\theta_r, \theta_{new}; J_r, L) = g_0(L) + J_r\kappa(L)\sin^2\theta_r + \frac{J_r}{\kappa(L)}\left(\frac{d^2V(r_c)}{dr^2} + \frac{3L^2}{r_c(L)^4}\right)\cos^2\theta_r \tag{40}$$

We retroactively chose $\kappa(L)$ so as to make the Hamiltonian independent of the epicyclic angle θ_r

$$\kappa(L)^2 = \frac{d^2V(r_c)}{dr^2} + \frac{3L^2}{r_c(L)^4} = \frac{d^2V(r_c(L))}{dr^2} + 3\Omega^2(L) \tag{41}$$

and the Hamiltonian becomes

$$H(\theta_r, \theta_{new}; J_r, L) = g_0(L) + \kappa(L)J_r \tag{42}$$

We did this expansion to second order in $J_r^{\frac{1}{2}}$. This is equivalent to or can be described as the *epicyclic approximation*. The action variable J_r sets the amplitude of radial oscillations and the frequency $\kappa(L)$ is the epicyclic frequency and governs the frequency of radial oscillations. It is possible to carry out a higher order expansion in J_r (see George Contopoulos's paper: Contopoulos G., 1975, ApJ, 201, 566).

It may be useful to manipulate the definitions for $L, r_c(L), \Omega(L)$ and $\kappa(L)$ to show that

$$\frac{dr_c(L)}{dL} = \frac{2\Omega(L)}{\kappa^2(L) r_c(L)} \tag{43}$$

$$\frac{d\Omega(L)}{dL} = \frac{1}{r_c^2(L)}\left(1 - \frac{4\Omega^2(L)}{\kappa^2(L)}\right) \tag{44}$$

And expansion of $g_0(L)$ about a particular L value, with $L = L_0 + l$ gives

$$g_0(L_0 + l) = g_0(L_0) + \Omega(L_0) l + \left[\frac{1}{r_c^2}\left(1 - \frac{4\Omega^2}{\kappa^2}\right)\right]_{L_0} \frac{l^2}{2} \tag{45}$$

The coefficient can also be written

$$\frac{1}{2r_c^2}\left(1 - \frac{4\Omega^2}{\kappa^2}\right)\Bigg|_{L_0} = \frac{\Omega}{r_c \kappa^2}\frac{d\Omega}{dr}\Bigg|_{L_0} \tag{46}$$

so as to compare to coefficients in the appendix by Contopoulos 1975.

3.3 The Jacobi integral

Consider the Hamiltonian

$$H(r, \theta, p_r, L) = \frac{p_r^2}{2} + \frac{L^2}{2r^2} + V(r) + \epsilon g(\theta - \Omega_b t)$$

where ϵ is small and $g(\theta - \Omega_b t)$ is a perturbation to the potential that is fixed in a frame rotating with angular rotation rate Ω_b. The perturbation is what one would expect for an oval or bar perturbation, as is found in barred galaxies, which is moving through the galaxy with a pattern speed or angular rotation rate Ω_b. Similarly a planet in a circular orbit about the Sun can cause a periodic perturbation (with a constant angular rotation rate) on an asteroid, also in orbit about the Sun. We take a time dependent generating function of old coordinate θ and new momenta L'

$$F_2(\theta, L', t) = (\theta - \Omega_b t) L'$$

giving new coordinates

$$\begin{aligned} L &= \frac{\partial F_2}{\partial \theta} = L' \\ \theta_{new} &= \frac{\partial F_2}{\partial L'} = \theta - \Omega_b t \end{aligned}$$

The transformation only involves θ, L so we neglect p_r, r in the transformation. Because the generating function is time dependent we must add

$$\frac{\partial F_2}{\partial t} = -\Omega_b L$$

to the new Hamiltonian. The new Hamiltonian in the new coordinates

$$\begin{aligned} K(r.\theta_{new}, p_r, L') &= H - \frac{\partial F_2}{\partial t} \\ &= \frac{p_r^2}{2} + \frac{L^2}{2r^2} - L\Omega_b + V(r) + \epsilon g(\theta_{new}) \end{aligned}$$

We note that the new Hamiltonian is time independent and so is conserved. This conserved energy, computed in the rotating frame is called the Jacobi integral. It is equivalent to the Tisserand relation when written in terms of orbital elements in the context of celestial mechanics (orbital dynamics in a frame rotating with a planet that is in a circular orbit around the Sun). The Tisserand relation is used to classify comets and estimate the range of orbital changes that can be caused by gravitational assists.

3.4 The Shearing Sheet

A two-dimensional system in the plane

$$H(r, \theta; p_r, L) = \frac{p_r^2}{2} + \frac{L^2}{2r^2} + V(r)$$

For a Keplerian system the potential $V(r) = -GM/r$. In a disk galaxy with a flat rotation curve $V(r) = v_c^2 \ln r$.

We want to consider motion near a particle that is in a circular orbit with radius R_0 and has angular rotation rate Ω_0 with

$$\Omega_0^2 = \left(\frac{1}{r} \frac{dV}{dr} \right) \bigg|_{R_0}$$

First let us rescale units so that units of time are in Ω_0^{-1} and units of distance are R_0. We want to go into a rotating frame moving with the particle at $r = R_0 = 1$ and has $\theta = \Omega_0 t = t$. In our rescaled units we can do this using a generating function using old coordinates and new momenta

$$F_2(\theta, r; p_x, p_y; t) = (\theta - t)(p_x + 1) + (r - 1)p_y$$

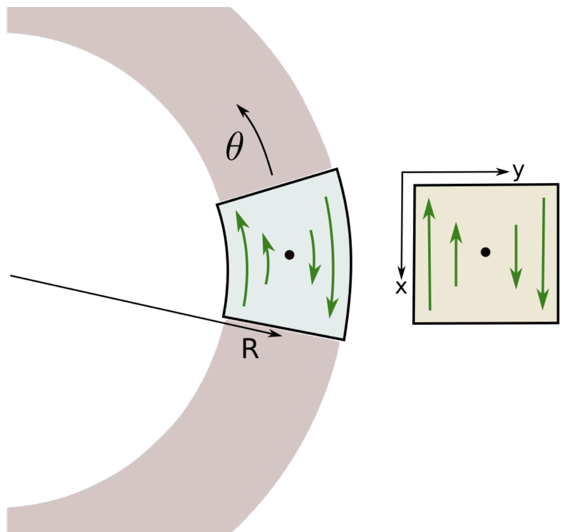

Figure 6: An illustration of a patch of a rotating disc (on left) and how the shearing box (on right) approximates it. Arrows are shown with respect to motion in the centre of the disc patch (on left). In this rotating frame a circular orbit would remain fixed at the black dot. An orbit with zero epicyclic amplitude and located at the centre of the shearing box (on right) would also remain fixed. The orientation of our coordinate system is shown on the right. The shearing sheet is described with Hamiltonian in equation 47 and equations of motion (equations 49). A central particle remains fixed. Particles with no epicyclic oscillations (in circular orbits) exhibit shear in their horizontal (x) velocities as a function of y. Often the shearing sheet is simulated with periodic boundary conditions in both x and y.

giving

$$\begin{aligned}
x &= \frac{\partial F_2}{\partial p_x} = \theta - t \\
y &= \frac{\partial F_2}{\partial p_y} = r - 1 \\
L &= \frac{\partial F_2}{\partial \theta} = p_x + 1 \\
p_r &= \frac{\partial F_2}{\partial r} = p_y \\
\frac{\partial F_2}{\partial t} &= -p_x - 1
\end{aligned}$$

Our new Hamiltonian (neglecting constants)

$$H(x, y; p_x, p_y) = \frac{p_y^2}{2} + \frac{(p_x + 1)^2}{2(1+y)^2} + V(1+y) - p_x$$

The Hamiltonian is independent of x so p_x is a conserved quantity. The Hamiltonian is time independent so H is constant. There is a fixed point at $x = 0, y = 0, p_x = 0, p_y = 0$. For small y we can expand

$$V(1+y) = V(1) + V'(1)y + V''(1)y^2/2$$

Recall that $\Omega_0^2 = R_0^{-1} V'(R_0)$ so with our choice of units $V'(1) = 1$. Below it will be useful to write $\kappa^2 = V''(1) + 3$ (usually one writes $\kappa_0^2 = V''(R_0) + 3\Omega_0^2$ but with $\Omega_0 = R_0 = 1$ this simplifies). Expanding the Hamiltonian for small y,

$$H(x, y; p_x, p_y) \approx \frac{p_y^2}{2} + \frac{(p_x + 1)^2}{2}(1 - 2y + 3y^2) + y + V''(1)\frac{y^2}{2} - p_x$$

I have have dropped $V(1)$ as it is a constant. We can also drop terms that are higher than second order in all coordinates and momenta, (dropping terms $\propto p_x^2 y$ and $\propto p_x y^2$) giving us

$$H(x, y; p_x, p_y) \approx \frac{p_y^2}{2} + \frac{p_x^2}{2} - 2p_x y + \frac{\kappa^2 y^2}{2} \tag{47}$$

and I have used $\kappa^2 \equiv V''(1) + 3$ for the epicyclic frequency.

Hamilton's equations gives

$$\begin{aligned}
\frac{\partial H}{\partial p_x} &= -2y + p_x = \dot{x} \\
\frac{\partial H}{\partial p_y} &= p_y = \dot{y} \\
\frac{\partial H}{\partial x} &= 0 = -\dot{p}_x \\
\frac{\partial H}{\partial y} &= -2p_x + \kappa^2 y = -\dot{p}_y
\end{aligned}$$

We can solve for
$$p_x = \dot{x} + 2y$$
and because p_x is conserved
$$\ddot{x} = -2\dot{y}$$
Inserting $p_x = \dot{x} + 2y$ into the expression for $\dot{p}_y$
$$\ddot{y} = 2\dot{x} + (4 - \kappa^2)y$$
For the Keplerian system $\kappa^2 = \Omega^2 = 1$
$$\begin{aligned} \ddot{x} &= -2\dot{y} \\ \ddot{y} &= 2\dot{x} + 3y \end{aligned} \tag{48}$$
With the addition of an additional local potential these are known as **Hill's equations** (and originally derived for orbits near the Moon and usually x and y are interchanged). But in a system with a different rotation curve (like a disk galaxy)
$$\begin{aligned} \ddot{x} &= -2\dot{y} \\ \ddot{y} &= 2\dot{x} + (4 - \kappa^2)y \end{aligned} \tag{49}$$

If we added in a local potential $W()$ (that is a function of distance from the origin) that could be due to a local mass at the origin then the Hamiltonian would look like this
$$H(x, y; p_x, p_y) \approx \frac{p_y^2}{2} + \frac{p_x^2}{2} - 2p_x y + \frac{\kappa^2 y^2}{2} + W(\sqrt{x^2 + y^2}) \tag{50}$$

Going back to the shearing sheet without any extra perturbations
$$H(x, y; p_x, p_y) \approx \frac{p_y^2}{2} + \frac{p_x^2}{2} - 2p_x y + \frac{\kappa^2 y^2}{2} \tag{51}$$
The momentum p_x sets the mean y value about which y oscillates. The oscillation frequency for y (radial) oscillations is the epicylic frequency κ and it's independent of the mean value of y. The angular rotation rate (here $\dot{x}$) is set by p_x and this sets the mean y value. We can look for solutions that look like $y = y_0 + A\cos(\kappa t + \phi_0)$ with phase ϕ_0. Conservation of p_x implies that we can take
$$\dot{x} = \dot{x}_0 - 2A\cos(\kappa t + \phi_0) \tag{52}$$
with $\dot{x}_0$ a constant. By subbing into equation 49 this gives
$$\dot{x}_0 = \frac{1}{2}(\kappa^2 - 4)y_0. \tag{53}$$

This is the velocity shear. For $\kappa = 1$ (Keplerian rotation) we recover the expected factor of 3/2.

It may be useful to write the equations of motion

$$y = y_0 + A\cos(\kappa t + \phi_0) \tag{54}$$
$$x = x_0 + \frac{1}{2}(\kappa^2 - 4)y_0 t - \frac{2A}{\kappa}\sin(\kappa t + \phi_0) \tag{55}$$
$$\dot{x} = \frac{1}{2}(\kappa^2 - 4)y_0 - 2A\cos(\kappa t + \phi_0) \tag{56}$$
$$\dot{y} = -A\kappa\sin(\kappa t + \phi_0). \tag{57}$$

We can write the above equations in terms of a guiding center x_g, y_g for the epicyclic motion

$$y = y_g + A\cos(\kappa t + \phi_0) \tag{58}$$
$$x = x_g - \frac{2A}{\kappa}\sin(\kappa t + \phi_0) \tag{59}$$

with

$$y_g = y_0 \tag{60}$$
$$x_g = x_0 + \frac{1}{2}(\kappa^2 - 4)y_0 t. \tag{61}$$

We can solve for the guiding centers in terms of current positions and velocities

$$y_g = 4y + \frac{2v_x}{\kappa^2} \tag{62}$$
$$x_g = x - \frac{2v_y}{\kappa^2} \tag{63}$$

with $v_x = \dot{x}$ and $v_y = \dot{y}$. We can solve for the epicyclic amplitude

$$A\cos(\kappa t + \phi_0) = (y - y_g) \tag{64}$$
$$A\sin(\kappa t + \phi_0) = -\frac{1}{2}\kappa(x - x_g) \tag{65}$$

As there is a fixed point at $x = 0, y = 0, p_x = 0, p_y = 0$ the Hamiltonian (equation 51) can be written in the form

$$H = \frac{1}{2}\mathbf{x}M\mathbf{x}$$

with M the Hessian matrix and $\dot{\mathbf{x}} = \boldsymbol{\omega}M\mathbf{x}$ and $\ddot{\mathbf{x}} = (\boldsymbol{\omega}M)^2\mathbf{x}$. As H is second order in all coordinates, M contains no variables (is just constants), conserved quantities can be identified from zero value eigenvalues of $\boldsymbol{\omega}M$ and frequencies of oscillation from eigenvalues of $(\boldsymbol{\omega}M)^2$.

The Hamiltonian is independent of x (giving us a conserved quantity p_x). An associated Lagrangian would be independent of x. That means in a simulation we can change x of a particle and the equations of motion will not change. That means in a simulation, periodic boundary conditions in x would not change the equation of motion (equations 49).

There is another symmetry in the equations of motion (is this because the Jacobi integral is conserved in the rotating frame?)

$$\begin{aligned}\dot{x} &\rightarrow \dot{x} + \epsilon\frac{(\kappa^2-4)}{2} \\ y &\rightarrow y + \epsilon\end{aligned}$$

This symmetry is exploited in simulations so that the y boundary condition can be periodic also.

If we are simulating with the Hamiltonian system then we are change momenta and coordinates rather than velocities and coordinates. Equations of motion are

$$\begin{aligned}\dot{x} &= p_x - 2y \\ \dot{y} &= p_y \\ \dot{p}_x &= 0 \\ \dot{p}_y &= 2p_x - \kappa^2 y\end{aligned}$$

Consider changing y by δy but not affecting any accelerations. To maintain $\dot{p}_y$ we require that

$$\delta p_x = \frac{\kappa^2}{2}\delta y$$

The acceleration in x is $\ddot{x} = \dot{p}_x - 2p_y$ is zero as long as we don't change p_y. How does this transformation affect $\dot{x}$?

$$\delta\dot{x} = \delta p_x - 2\delta y = \frac{1}{2}(\kappa^2-4)\delta y$$

and this is equivalent to the transformation given above. In terms of momenta and coordinates the transformation is

$$\begin{aligned}y &\rightarrow y + \epsilon \\ p_x &\rightarrow p_x + \epsilon\frac{\kappa^2}{2}\end{aligned}$$

As the system has two independent conserved quantities (H, p_x), without any additional perturbations the system is integrable, and this is expected as this system is equivalent to the epicyclic approximation. The shearing sheet, because it is non-trivial, integrable and can be simulated with periodic boundary conditions is a nice place for particle integrations (see recent work by Hanno Rein and collaborators, and including short range interactions between particles).

4 Symmetries and Conserved Quantities

4.1 Functions that commute with the Hamiltonian

Recall that for a function $f(p, q, t)$

$$\dot{f} = \{f, H\} + \frac{\partial f}{\partial t}$$

If the Poisson bracket

$$\{f, H\} = 0 \tag{66}$$

and $\frac{\partial f}{\partial t} = 0$ then f is a conserved quantity.

Consider the gradient of a function on p, q or with one-form $df = \frac{\partial f}{\partial q}dq + \frac{\partial f}{\partial p}dq$. We can use the symplectic form to generate a flow or vectors V such that $\omega(V, ?) = df$. In this way f generates a direction in phase space or a tangent vector. The Hamiltonian generates a flow. The function f also generates a flow. The function f corresponding to a conserved quantity is also one that commutes with H and generates a flow that commutes with the flow generated by H.

Using $\mathbf{x} = (\mathbf{q}, \mathbf{p})$ and $\boldsymbol{\nabla} = (\partial_{\mathbf{q}}, \partial_p)$ recall that we could write the equations of motion as

$$\dot{\mathbf{x}} = \boldsymbol{\omega}\boldsymbol{\nabla} H$$

where $\boldsymbol{\omega}$ is the matrix made up of a positive and negative identity matrix

$$\boldsymbol{\omega} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

Using gradients it is possible to rewrite equation 66 as

$$\{f, H\} = \boldsymbol{\nabla} f^t \boldsymbol{\omega} \boldsymbol{\nabla} H = 0$$

Using the equations of motion

$$\boldsymbol{\nabla} f \cdot \dot{\mathbf{x}} = 0$$

A conserved quantity is one that has a gradient (in phase space) perpendicular to the direction of motion. If there are many possible such directions (there are many conserved quantities) then the orbits must be of low dimension.

The Hamiltonian formalism relates flows that commute with the Hamiltonian flow with conserved quantities.

4.2 Noether's theorem

Noether's theorem relates coordinate symmetries of the Lagrangian with a conserved quantity. Consider a Lagrangian $\mathcal{L}(q, \dot{q}, t)$ and a transformation of the coordinates

$$q_s = h(s, q)$$

with

$$q_{s=0} = h(0, q)$$

and s a continuous parameter. We require that the transformed time derivative

$$\dot{q}_s = \frac{d}{dt} q_s$$

so that h transforms both q and $\dot{q}$ consistently.

The map h is a symmetry of the Lagrangian when

$$\mathcal{L}(q_s, \dot{q}_s, t) = \mathcal{L}(q, \dot{q}, t)$$

The Lagrangian cannot depend on s so

$$\left. \frac{\partial \mathcal{L}(q_s, \dot{q}_s, t)}{\partial s} \right|_{s=0} = 0$$

I don't think it matters what particular s value where you evaluate the derivative but if you choose another s value than that's equivalent to choosing a different set of coordinates and evaluating the derivative at $s = 0$.

Using the chain rule

$$\frac{\partial \mathcal{L}}{\partial s} = \frac{\partial \mathcal{L}}{\partial q} \frac{\partial q_s}{\partial s} + \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{\partial \dot{q}_s}{\partial s} \tag{67}$$

We can write

$$\frac{\partial \dot{q}_s}{\partial s} = \frac{d}{dt} \frac{\partial q_s}{\partial s} \tag{68}$$

and using Lagrange's equation we can replace $\frac{\partial \mathcal{L}}{\partial q}$ with $\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}}$. These inserted into equation 67 give

$$\frac{\partial \mathcal{L}}{\partial s} = \left(\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \frac{dq_s}{ds} + \frac{\partial \mathcal{L}}{\partial \dot{q}} \left(\frac{d}{dt} \frac{\partial q_s}{\partial s} \right)$$

and

$$\left. \frac{\partial \mathcal{L}}{\partial s} \right|_{s=0} = \frac{d}{dt} \left. \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{\partial q_s}{\partial s} \right] \right|_{s=0} = 0.$$

Hence we have a conserved quantity

$$I = \left. \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{\partial q_s}{\partial s} \right|_{s=0}$$

or

$$I = \left. \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\partial q_{i,s}}{\partial s} \right|_{s=0}$$

using summation notation and in more than one dimension.

Note: Rather than conserving $\mathcal{L}$ we should really be thinking about conserving the action. In this case we should be considering infinitesimals in time and end points.

The Lagrangian formalism relates a symmetry of the Lagrangian to a conserved quantity.

In the Hamiltonian context, we use the symplectic two form and a conserved quantity (here a function) to generate a vector and so a flow in phase space. This flow commutes with the Hamiltonian flow.

4.3 Integrability

A time independent Hamiltonian with N degrees of freedom (N coordinates, so phase space is $2N$ dimensional) is said to be *integrable* if N smooth independent functions I_i can be found that

$$\{I_i, H\} = 0$$

that are conserved quantities and

$$\{I_i, I_j\} = 0$$

are in *involution*. The reason that the constants should be smooth and independent is that the equations $I_i(\mathbf{q}, \mathbf{p}) = c_i$, where the c_i's are constants, must define N different surfaces of dimension $2N - 1$ in the 2N-dimensional phase space. If the conserved quantities are in involution then they can be used as *canonical momenta*. Then Hamiltonian can then be written as a function of the N different momenta.

Suppose a Hamiltonian with N coordinates is integrable. The Hamiltonian can be written as $H(\mathbf{p})$ with N momentum, and all momenta are conserved quantities. The vector $\nabla_p H$ gives N frequencies that specify the motion of the N coordinates. The dimension of an orbit is N if none of the frequencies are zero.

For example, consider a 4 dimensional phase space with a time independent Hamiltonian. An orbit is a trajectory in 4 dimensional phase space. Once we specify the energy (Hamiltonian) then there is a constraint on the orbit and the orbit must wander in a 3-dimensional subspace. If we specify an additional conserved quantity then the orbit must wander in a 2 dimensional subspace. At this point we say the system is integrable.

The orbit could be a lower dimensional object if there are additional conserved quantities. A circular orbit in the plane of an axisymmetric galaxy is a one dimensional object (points only a function of angle). Epicyclic motion near the circular orbit covers a 2-dimensional surface (is a torus). I can describe it in terms of two angles given angular momentum and the epicyclic action variable. The two associated frequencies are the angular rotation rate and the epicyclic frequency. This system has two conserved quantities (energy and angular momentum) so is integrable.

If the potential is proportional to $1/r$ (Keplerian setting) then there is an additional conserved quantity (Runge-Lunz vector) and every orbit is one dimensional (an ellipse) instead of a 2-dimensional (torus). Systems with extra conserved quantities (past what is needed to make them integrable) are known as *superintegrable*. The Keplerian system is *maximally superintegrable* as there are 5 conserved quantities for every orbit in 6 dimensional phase space and the orbit is one dimensional.

A superintegrable system can be written in terms of a Hamiltonian that has some frequencies that are zero. In other words $H(p)$ does not depend upon all the momenta. For example the Keplerian system can be written as $H(\Lambda)$ in terms of a single momentum, giving only one angle that increases in time (the mean motion, which specifies where in the ellipse a particle is at each time).

Liouville integrability means there exists a maximal set of Poisson commuting invariants (i.e., functions on the phase space whose Poisson brackets with the Hamiltonian of the system, and with each other, vanish).

Locally there could be a complete set of conserved quantities and it might be possible to construct canonical transformations that give you a Hamiltonian purely in terms of conserved quantities. But it might not be possible to find a complete set the covers the entire manifold (in the Liouville sense of integrability).

Often we call a system integrable if the dimension covered by orbits is low. For example in a 4 dimensional phase space if there are 2 (non-degenerate) conserved quantities (one of them could be the Hamiltonian itself). Orbits are in a 4 dimensional space, and two conserved quantities drops the dimension by 2 so orbits cover a surface. They are like tori as two frequencies describe motion. If the system had only a single conserved quantity (like energy) we could make a surface of section and find area filling orbits that corresponded to orbits filling a 3 dimensional volume. Periodic orbits would be fixed points in a surface of section and these would be 1 dimensional.