Visualizing the Inverse Noether Theorem and Symplectic Geometry

Noah Miller

February 14, 2019 (updated August 1, 2019)

Abstract

A collection of visualizations pertaining to symmetries and conservation laws in Hamiltonian mechanics. This includes visual interpretations of the Poisson bracket and symplectic form. Visual proofs are given for Liouville’s theorem, the inverse Noether Theorem (which states that all conserved quantities generate a symmetry), and the Jacobi identity.

Contents

1 Hamiltonian evolution, Liouville’s theorem 2
2 Hamiltonian mechanics 5
3 Poisson Bracket 6
4 Higher dimensions 10
5 Lie Bracket, Commutativity, and Symmetry 10
6 Symplectic Form 13
7 Visual Proof of the Inverse Noether Theorem 15
8 Visual Proof of the Jacobi Identity 17
9 Appendix 20
   9.1 Lie Bracket .......................................................... 20
   9.2 Visualizing $[X_f, X_g] = X_{\{f,g\}}$ .................................. 21
1 Hamiltonian evolution, Liouville’s theorem

Let’s say you have two variables called \( q \) and \( p \). “Phase space” is the 2D plane comprised of all \( (q, p) \in \mathbb{R}^2 \). \( q \) is called “position” and \( p \) is called “momentum.”

![Phase space diagram](image)

Figure 1: On the left we have phase space. On the right we have lines of constant \( H \) for some arbitrary function \( H(q, p) \).

Now say you have some function on phase space, \( H : \mathbb{R}^2 \to \mathbb{R} \). Let’s say you want to take one point \((q, p)\) and evolve it along a path such that \( H \) remains constant. Parameterizing the path by \( t \), the condition that \( H \) is kept constant can be written

\[
0 = \dot{H} = \dot{q} \frac{\partial H}{\partial q} + \dot{p} \frac{\partial H}{\partial p}.
\]

This implies that

\[
\dot{q} = \lambda(t) \frac{\partial H}{\partial p} \quad \dot{p} = -\lambda(t) \frac{\partial H}{\partial q}
\]

for some arbitrary function \( \lambda(t) \). We might as well take this function to be 1. Note that the components of the gradient of \( H \) are

\[
\nabla H = \left( \frac{\partial H}{\partial q}, \frac{\partial H}{\partial p} \right).
\]

This means that the vector

\[
\left( \frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q} \right)
\]

is just \( \nabla H \) rotated by 90° clockwise.

Imagine drawing lines of constant \( H \), sometimes called “equipotential lines,” as in the right portion of Fig. 1. If you were to plot \( H \) using a 3D graph, where the height of the plot represented the value of \( H(q, p) \), the equipotential lines would be lines of constant height at evenly spaced intervals. They make maps like this for hiking trails called topographic maps.
Continuing with this analogy, $\nabla H$ points in the direction where the height is increasing the fastest. Furthermore, the length of $\nabla H$ is inversely proportional to the distance between adjacent equipotential lines (assuming that the difference in height between two nearby equipotential lines is sufficiently small). When $\nabla H$ is small, the lines are spread out. When $\nabla H$ is large, the lines are close together.

Next, the vector field given by

$$X_H = \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q}\right)$$

is called the “Hamiltonian vector field generated by $H$.” We call it $X_H$ to follow the notation commonly used in mathematics.
Figure 4: Lines of equipotential and gradient vectors $\nabla H$. Notice that $\nabla H$ will always point at a right angle to the equipotential lines.

I should now introduce a piece of terminology. Say you have a point $(q_0, p_0)$ and any vector field $X$. When we say we “flow $(q_0, p_0)$ along $X$ for a time $t$,” it means we evolve $(q_0, p_0)$ along a path whose velocity vector is equal to $X$ for a length of time $t$. Note that $(\dot{q}, \dot{p}) = X$ is a first order differential equation for which $q(0) = q_0$ and $p(0) = p_0$ are its initial conditions. The point $(q(t), p(t))$ would be the final result of this flow. Thinking of $t$ as an adjustable parameter, we can see that every vector field gives a one-parameter family of transformations on phase space.

Figure 5: A random vector field and its “flow lines.” The flow lines are blue where they have larger velocity and red where they have lower velocity.

Let’s return to thinking about the vector field $X_H$ specifically. Recall that $X_H$ is just the gradient vector $\nabla H$ rotated clockwise by $90^\circ$. Therefore, the value of $H$ will remain constant on flow lines. Furthermore, because $\nabla H$ is larger when the equipotential lines are close together, the path will speed up when the equipotential lines are close together and slow down when the equipotential lines are far apart.

This gives a good way to picture Liouville’s theorem. Liouville’s theorem states that time evolution generated by a Hamiltonian vector field preserves “phase space volume.” That means that if we allow any arbitrary “blotch” in phase space to flow
along $X_H$, the area of that blotch will remain constant in time (even though the overall shape may change in complex ways).

This fact is apparent using the visualization tools described above. Imagine a tiny rectangular area flowing along $X_H$. The height of the rectangle will be the distance between equipotential lines and the length will be proportional to the velocity at which the rectangle is moving. This is drawn in Fig. 6 below. However, the velocity of the rectangle is inversely proportional to its height. Therefore, no matter how the equipotential lines wiggle around, the rectangle will always have the same area as it evolves in time. We can therefore see that Liouville’s theorem is just like continuity equation in fluid mechanics: in steady state, water must travel faster down a narrow part of a pipe and slower in a wider part of the pipe.

![Figure 6: A small rectangle of points evolving according to $X_H$. The rectangle’s area will remain constant under this evolution because the velocity of the rectangle’s length will be inversely proportional to the height.](image)

## 2 Hamiltonian mechanics

A point in phase space should be thought of as a physical state. $H$ can in principle be any function on phase space, but in physics, the letter “$H$” specifically is reserved for a distinguished function called the “Hamiltonian” which is equal to the energy of a physical state. Furthermore, it can be taken as an axiom, if you wish, that physical time evolution is given by flowing along $X_H$. Physicists therefore like to say the words “$H$ generates time translations.”

It’s instructive to consider what the Hamiltonian vector fields looks like for simple Hamiltonians. For a free particle in one dimension,

$$H = \frac{p^2}{2m}.$$ 

$H$ increases quadratically in $p$, meaning that the lines of equipotential squeeze closer the further away you get from $p = 0$. Therefore, a particle with a large momentum travels faster. The fact that a particle with positive momentum travels in a different direction from a particle with negative momentum comes from the fact that $X_H$ is a 90° clockwise rotation of $\nabla H$. Because $\nabla H$ always points away from the $p = 0$ axis, that 90° turn makes $X_H$ point to the right for $p > 0$ and to the left for $p < 0$. This is depicted in Fig. [7](image).
Figure 7: Equipotential lines, $\nabla H$, and $X_H$ for the free particle Hamiltonian $H = \frac{1}{2m}p^2$ where $m = 1$.

Let us now look at the “harmonic oscillator” (a particle attached to an idealized spring). In some units, the Hamiltonian for this system is

$$H = \frac{1}{2}(p^2 + q^2).$$

The lines of equipotential are circles. The circles get closer and closer together the further you get from the origin. Furthermore, $\nabla H$ always points outwards from the origin and gets bigger as you go away from the origin. Once you rotate these vectors by $90^\circ$, they all go in a circle. In particular, time evolution by $X_H$ just gives us rotation at a constant angular velocity.

Figure 8: Equipotential lines, $\nabla H$, and $X_H$ for the rescaled harmonic oscillator Hamiltonian $H = \frac{1}{2}(p^2 + q^2)$.

3 Poisson Bracket

Let’s say that we now introduce a test function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ and we want to know how $f$ changes infinitesimally as you flow along $X_H$.

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

$$= \frac{\partial H}{\partial p} \frac{\partial f}{\partial q} - \frac{\partial H}{\partial q} \frac{\partial f}{\partial p}$$

$$= \{f, H\}$$
Above, \( \{f, H\} \) is called the “Poisson bracket” of \( f \) and \( H \).

Peculiarly, mathematicians usually like define vector fields to be first order differential operators which act on test functions. So a mathematician would write

\[
X_H = \frac{\partial H}{\partial p} \partial_q - \frac{\partial H}{\partial q} \partial_p. \tag{4}
\]

Contrast this way of writing \( X_H \) with Eq. 2. Both notions are useful and we will find ourselves flipping between them at multiple points. The vector field \( X_H \) acts on the test function as

\[
X_H(f) = \frac{\partial H}{\partial p} \frac{\partial f}{\partial q} - \frac{\partial H}{\partial q} \frac{\partial f}{\partial p} \tag{5}
= \{f, H\}.
\]

Thinking of \( X_H \) in terms of its components (Eq. 2) we can see that

\[
X_H(f) = X_H \cdot \nabla f. \tag{6}
\]

\( X_H(f) \) can therefore be thought of as the directional derivative of \( f \) in the \( X_H \) direction. This means that \( X_H(f) \) is that it is the rate of change of \( f \) along flow lines, i.e.

\[
\dot{f} = X_H(f). \tag{7}
\]

There is something a little surprising about the equation \( \dot{f} = X_H(f) \). Because \( \{f, H\} = -\{H, f\} \), we can see that

\[
X_H(f) = -X_f(H). \tag{8}
\]

While mathematically trivial, the pictorial interpretation of this fact is at first puzzling. \( X_H \) is the Hamiltonian vector field which keeps \( H \) constant along flow lines. \( X_f \) is the Hamiltonian vector field which keeps \( f \) constant along flow lines. Why does \( f \) change by (negative) the same amount along \( X_H \) that \( H \) changes along \( X_f \)? The answer comes from keeping our whole 90° rotation trick in mind.

Pick any particular point in phase space and look at the equipotential lines of \( H \) and \( f \) that pass through that point. The vectors \( \nabla H \) and \( \nabla f \) point perpendicularly to their respective equipotential lines. Furthermore, say that the angle between the two equipotential lines is \( \theta \).

![Figure 9: Equipotential lines for \( H \) and \( f \), along with the gradient vectors \( \nabla H \) and \( \nabla f \).](image)

If we want to compute \( X_H(f) \), we turn \( \nabla H \) by 90° clockwise and then take its dot product with \( \nabla f \). If we want to compute \( X_f(H) \), we turn \( \nabla f \) by 90° clockwise and then take the dot product with \( \nabla H \).
Recall that the dot product of two vectors is equal to the product of the lengths of both vectors and the cosine of the angle between them. Therefore,

\[ X_H(f) = |\nabla H||\nabla f| \cos(90° - \theta) \quad \text{and} \quad X_f(H) = |\nabla H||\nabla f| \cos(90° + \theta) \]

and

\[ \cos(90° - \theta) = -\cos(90° + \theta) \implies X_H(f) = -X_f(H) \]

thus proving the claim.

I should mention that the statement that \( X_f(H) = -X_H(f) \) is sometimes seen as the avatar of Noether’s theorem in Hamiltonian mechanics. Say you have some function on phase space called \( Q \). \( Q \) generates a one parameter family of transformations given by flow lines of \( X_Q \). (For example, it turns out that if \( J \) is the angular momentum along \( z \)-direction, then \( X_J \) generates rotations around the \( z \) axis. Here, our “parameter” is the angle of rotation.) One could argue that a symmetry is a transformation that does not change the energy of a physical state, i.e. \( X_Q(H) = 0 \). However, we can now see this immediately implies \( X_H(Q) = 0 \), which means \( Q = 0 \), i.e. \( Q \) is conserved. This is why people sometimes cheekily claim Noether’s theorem is just

\[ \{H, Q\} = 0 = \{Q, H\}. \quad (9) \]

I actually think there is a much better way to understand how Noether’s theorem fits into Hamiltonian mechanics, but I will comment on that later.

Let me now speak more generally about the Poisson bracket. Consider a function \( h \) that is the Poisson bracket of two other functions, \( f \) and \( g \):

\[ h = \{f, g\}. \]

How can we visualize the properties of \( h \) based on \( f \) and \( g \)?

We now have built up intuition for how to understand \( \{f, g\} \) as the directional derivative \( X_g(f) \). We know that \( X_g(f) \) should be large where the lines of equipotential for \( f \) and \( g \) are packed close together and nearly perpendicular to each other. Equivalently, \( X_g(f) \) will be small where the lines of equipotential for \( f \) and \( g \) are spaced far apart or when they are nearly parallel. (There is also a sign based on the orientation of the gradients.)
Figure 11: \( \{f, g\} \) is large when lines of equipotential for \( f \) and \( g \) are packed tight and nearly perpendicular. \( \{f, g\} \) is close to zero when the lines of equipotential are spaced far apart and not perpendicular.

The Poisson bracket \( \{f, g\} \) has a direct visual interpretation which I will now describe. Note that the equipotential lines of \( f \) and \( g \) form little parallelograms. It turns out that \( \{f, g\} \) is inversely proportional to the area of these parallelograms. I will prove this for a simple example. Say that \( f = q \) and \( g = rq \sin \theta + rp \cos \theta \). Note that
\[
\{f, g\} = \{q, rq \sin \theta + rp \cos \theta\} = r \cos \theta.
\]

Next, draw the lines of equipotential for these two functions. Let’s take the “height differential” between equipotential lines to be \( \varepsilon \). Then the space between equipotential lines for \( f = q \) will be \( \varepsilon \) and the space between equipotential lines for \( g = rq \sin \theta + rp \cos \theta \) will be \( \varepsilon/r \).

Figure 12: The green lines are the equipotential lines for \( f = q \) and the blue lines are the equipotential lines for \( g = rq \sin \theta + rp \cos \theta \). The area of the shaded parallelogram is \( \frac{\varepsilon^2}{r \cos \theta} \).

We can see that the area of the shaded parallelogram is
\[
\text{area} = \frac{\varepsilon^2}{r \cos \theta} = \frac{\varepsilon^2}{\{f, g\}}.
\]
In fact, this should hold for any $f$ and $g$. If we take $\varepsilon$ to be small, then locally any functions will look linear and the intersection of their equipotential lines will also form small parallelograms.

We will use this visual interpretation of the Poisson bracket in section when proving the Jacobi identity.

(As an aside, you can also think about $\{f, g\}$ as the area of the parallelogram spanned by $\nabla f$ and $\nabla g$.)

4 Higher dimensions

I should comment briefly on what happens in higher dimensions. In general, phase space can be a space of even dimension, i.e. $\mathbb{R}^{2N}$. For example, the phase space for a single particle moving in three spatial dimensions is $\mathbb{R}^6$. Three of the dimensions in phase space parameterize the three components of its position vector $(q_1, q_2, q_3)$, while the other three parameterize the three components of its momentum vector $(p_1, p_2, p_3)$. The Poisson bracket for this phase space is just

$$
\{f, g\} \equiv \sum_{i=1}^{3} \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i}.
$$

(10)

As claimed earlier, the angular momentum in the $z$-direction

$$
J = q_1 p_2 - q_2 p_1
$$

(11)
generates rotations around the $z$-axis:

$$
\dot{q}_1 = -q_2 \quad \dot{q}_2 = q_1 \quad \dot{p}_3 = 0
$$

$$
\dot{p}_1 = -p_2 \quad \dot{p}_2 = p_1 \quad \dot{p}_3 = 0.
$$

Anyway, the higher dimensional situation is not radically different from the two dimensional phase space we have discussed thus far. In particular, much of the intuition we’ve built for $\mathbb{R}^2$ can be used in $\mathbb{R}^{2N}$ by “projecting down” onto the $N$ different $\mathbb{R}^2$’s spanned by $q_i$ and $p_i$. For example, not only is Liouville’s theorem still true in $\mathbb{R}^{2N}$ (that phase space hyper-volume is preserved by Hamiltonain vector flow) but something stronger is also true: if you have a tiny $2N$ dimensional hyper-parallelepiped in your phase space, the areas of the “shadows” of this hyper volume in each of the $N$ different $\mathbb{R}^2$’s will remain constant under Hamiltonian vector flow.

Because the $\mathbb{R}^{2N}$ case does not introduce anything fundamentally new, we will mostly ignore it from here on out.

5 Lie Bracket, Commutativity, and Symmetry

To reiterate, mathematicians like to define vector fields as first order differential operators which act on test functions. The Lie bracket of two vector fields $X$ and $Y$ is defined to be the differential operator which acts on test functions as

$$
[X, Y](f) \equiv X(Y(f)) - Y(X(f)).
$$

(12)
There is a simple way to interpret $[X, Y]$ pictorially. Imagine flowing a point along $X$ for a tiny time $\varepsilon_1$. Then flow the resulting point along $Y$ for $\varepsilon_2$. The final point will, in general, end up at a different place if you were to flow along $Y$ first and $X$ second. When this happens, we say that the vectors fields fail to “close.” The vector pointing between these two possible end points will be given by $\varepsilon_1\varepsilon_2[X, Y]$. (I prove this in the appendix, section 9.1.)

Figure 13: The Lie Bracket of two vector fields $X$ and $Y$ can be understood as measuring the failure of their respective flows to commute infinitesimally.

With this piece of intuition, we can see that when $[X, Y] = 0$, the flows given by $X$ and $Y$ will commute for finite times as well as infinitesimal times. We can visualize this using the following abstract diagram:

Figure 14: Commuting vector flows.

Note that commutativity captures our usual notion of symmetry. For example, the laws of physics have a rotational symmetry. That means that if you rotate a system by some angle $\theta$ and then wait some time $t$, it will end up in the same final state as if you were to wait first and rotate it second.

In quantum mechanics, we capture the above statement mathematically (suppressing $\hbar$) as

$$[e^{-i\hat{H}}, e^{-i\hat{J}}] = 0.$$  

The above equation can actually be understood as four closely related equations.
Figure 15: Rotations commute with time evolution.

1. \[ [e^{-it\hat{H}}, e^{-i\theta \hat{J}}] = 0 \]: Rotating and then time evolving a state is the same as time evolving and then rotating. (We have a symmetry.)

2. \[ [e^{-it\hat{H}}, \hat{J}] = 0 \]: The angular momentum of a state does not change after time evolution. (Angular momentum is conserved.)

3. \[ [\hat{H}, e^{-i\theta \hat{J}}] = 0 \]: The energy of a state does not change if the state is rotated.

4. \[ [\hat{H}, \hat{J}] = 0 \]: If you measure the angular momentum of a state, the probability that the state will have any particular energy afterwards will be unchanged, and vice versa. (\( \hat{H} \) and \( \hat{J} \) can be simultaneously diagonalized.)

We can see that symmetries and conservation laws are interrelated in many ways far beyond the simple statement “symmetries give conservation laws.”

Rather amazingly, three of these four statements about quantum mechanics also have direct analogs in classical mechanics!

1. \[ [X_H, X_J] = 0 \]: Rotating and then time evolving a state is the same as time evolving and then rotating. (We have a symmetry.)

2. \[ X_H(J) = 0 \]: The angular momentum of a state does not change after time evolution. (Angular momentum is conserved.)

3. \[ X_J(H) = 0 \]: The energy of a state does not change if the state is rotated.

4. \[ \{H, J\} = 0 \]: No classical meaning I can think of. (Can you think of one?)

In classical mechanics, the fact which links the first of the above statements to the other three is the so-called “Jacobi identity” of the Poisson bracket.

\[ \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0 \] (13)

The above identity can be readily derived algebraically by using the definition of the Poisson bracket and writing out all 24 terms that arise via differentiation by \( q \) and \( p \). They will all cancel. A visual proof will be given in section 8.

Using the anti-symmetry of the Poisson bracket and rearranging Eq. (13) we find

\[ \{f, \{g, h\}\} = \{(f, g), h\} - \{(f, h), g\} \]

\[ X_{\{g,h\}}(f) = X_h(X_g(f)) - X_g(X_h(f)). \]
Therefore, once again using the letters $H$ and $Q$ as our functions on phase space,

$$X_{\{Q,H\}} = [X_H, X_Q].$$  \hspace{1cm} (14)

I prefer thinking about Noether’s theorem in Hamiltonian mechanics using the above equation. If $Q$ is conserved under time evolution, then $\dot{Q} = \{Q, H\} = 0$ and $X_{\{Q,H\}} = X_0 = 0$. Therefore, we can see $X_Q$ generates a symmetry because $[X_H, X_Q] = X_{\{Q,H\}} = 0$. Eq. 14 lets us explicitly see how symmetries are related to conservation laws.

\section{Symplectic Form}

We are now going to define a fancy mathematical object with a simple geometrical interpretation. The “symplectic form” $\omega$ takes in two vectors as inputs and outputs a number. This number is equal to the area of the parallelogram spanned by those two vectors.

$$\omega(a, b) = \text{area of parallelogram spanned by } a \text{ and } b \hspace{1cm} (15)$$

It turns out that there is a very simple formula for the area of a parallelogram in terms of the components of these vectors.

$$\omega(a, b) = a_1b_2 - a_2b_1$$

Therefore, if $a$ and $b$ are explicitly

$$a = a_1\partial_q + a_2\partial_p$$
$$b = b_1\partial_q + b_2\partial_p$$

then the symplectic form can be written as

$$\omega(a, b) = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \hspace{1cm} (16)$$

I should note that in higher dimensions where phase space is $\mathbb{R}^{2N}$, that $2 \times 2$ matrix
becomes the $2N \times 2N$ matrix
\[
\begin{pmatrix}
0 & 1 & 0 & 0 & \ldots & 0 & 0 \\
-1 & 0 & 0 & 0 & \ldots & 0 & 0 \\
0 & 0 & 0 & 1 & \ldots & 0 & 0 \\
0 & 0 & -1 & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \ldots & 0 & 1 \\
0 & 0 & 0 & 0 & \ldots & -1 & 0 \\
\end{pmatrix}
\]

The symplectic form should then be interpreted as the sum of areas of the “shadows” of the vectors in the $N$ different $\mathbb{R}^2$ subspaces.

Anyway, let’s stay focused on $\mathbb{R}^2$. Notice that the matrix
\[
\begin{pmatrix}
0 & 1 \\
-1 & 0 \\
\end{pmatrix}
\]
represents a clockwise 90° rotation on column vectors and a counter-clockwise 90° rotation on row vectors. We can therefore see that the symplectic form simply rotates one vector by 90° and then takes the dot product of the two resulting vectors. We can see why this specific operation gives us the area of the parallelogram by remembering two facts: 1) the area of a parallelogram is just base $\times$ height, and 2) the dot product of two vectors is equal to the length of the first vector multiplied by the length of the other vector’s projection onto the first vector.

Figure 17: This picture should convince you that the symplectic form given in Eq. [16], which rotates the green vector 90° before taking the dot product with the blue vector, really does compute the area of the parallelogram spanned by the two vectors.

I claim that Fig. [17] provides a visual proof of the equation
\[
\omega(X_f, v) = v(f)
\]
in the sense that $v$ is the blue arrow, $X_f$ is the green arrow on the left and $\nabla f$ is the green arrow on the right.

However, if you are having a tough time seeing this, it is also easy to prove algebraically. First, write $v$ in terms of its components:
\[
v = v_1 \partial_q + v_2 \partial_p \\
v(f) = v_1 \partial_q f + v_2 \partial_p f \\
= (\partial_q f \quad \partial_p f) \begin{pmatrix} v_1 \\
v_2 \end{pmatrix}
\]
Second, we write $\omega(X_f, v)$ in terms of its components, giving us Eq. 17:

$$
\omega(X_f, v) = \left( \frac{\partial f}{\partial p} - \frac{\partial f}{\partial q} \right) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \left( \frac{\partial f}{\partial q} \frac{\partial f}{\partial p} \right) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = v(f) \checkmark
$$

There’s one last fact I must mention. I hope you can agree that if there is some mystery vector “?” that satisfies

$$
\omega(?, v) = v(f)
$$

for all $v$, then $? = X_f$. Keep this in mind.

### 7 Visual Proof of the Inverse Noether Theorem

Say $H$ is the Hamiltonian and $Q$ is some conserved quantity.

$$
\dot{Q} = X_H(Q) = 0. \quad (18)
$$

The inverse Noether theorem states that $X_Q$ must be a symmetry:

$$
[X_H, X_Q] = 0.
$$

We now begin the proof. Let us use the intuition of vectors as “pointing” between two infinitesimally nearby points in phase space.

First, choose any point in phase space and draw two vectors based at that point: $X_Q$ and a completely arbitrary vector $v$. Both of these vectors can be thought of as spanning a tiny parallelogram in phase space.

Next, flow this parallelogram along $X_H$ for any amount of time. This tiny parallelogram will evolve into another tiny parallelogram, as drawn in Fig. 18. The side spanned by $v$ will now be spanned by a new vector which we will call $u$. $X_Q$ will sent to a new vector as well, which we will call “?” for the time being.

![Figure 18: A tiny parallelogram changes after flowing along $X_H$.](image)

The first thing we note is that

$$
\omega(X_Q, v) = v(Q) \quad (19)
$$
which is just Eq. 17 rewritten.

Next, notice that $v(Q)$ can be thought of as the change of $Q$ between the tail of $v$ and the tip of $v$. However, because of Eq. 18, the value of $Q$ of any point in phase space remains constant after flowing along $X_H$, as drawn in Fig. 19.

\[ v(Q) = u(Q). \] (20)

However, from Liouville’s theorem, we also know that the area of the parallelogram must be preserved under time evolution. Therefore

\[ \omega(X_Q, v) = \omega(?, u). \] (21)

Using Eq. 19, Eq. 20, and Eq. 21, we can see that

\[ \omega(?, u) = u(Q). \] (22)

This is true for any $u$. Therefore, we must have

\[ ? = X_Q. \]

Figure 19: Because $Q$ is conserved along the flow of $X_H$, $v(Q) = u(Q)$.

Therefore, the change of $Q$ between the tail of $v$ and the tip of $v$ must be equal to the change of $Q$ between the tail of $u$ and the tip of $u$, implying

\[ v(Q) = u(Q). \]

This is true for any $u$. Therefore, we must have

\[ ? = X_Q. \]

Remember that $X_Q$ is really a field of vectors, with one vector based at every point. We have just shown that when we flow $X_Q$ at one point along $X_H$, the resulting vector will be equal to $X_Q$ at the end point.

This implies that $[X_H, X_Q] = 0$: flowing from the tail of $X_Q$ to the tip, and then flowing along $X_H$, will take you to the same point in phase space as flowing along $X_H$ first, and then flowing from the tail of $X_Q$ to the tip. This concludes the proof.

If your phase space has more dimensions than two, you can picture the flows of $X_H$ and $X_Q$ as living in two-dimensional sub-surfaces in the larger phase space, as drawn in Fig. 20. $H$ and $Q$ will be constant on these surfaces.
8 Visual Proof of the Jacobi Identity

I will now give a visual proof of the Jacobi identity, which you may recall is

\[ \{ f, \{ g, h \} \} + \{ g, \{ h, f \} \} + \{ h, \{ f, g \} \} = 0. \]  
\[ (23) \]

This can be rewritten as

\[ X_h(\{ f, g \}) = \{ X_h(f), g \} + \{ f, X_h(g) \}. \]  
\[ (24) \]

To start, imagine that \( f \) and \( g \) are completely arbitrary functions with completely arbitrary lines of equipotential. As discussed in section 3, these lines of equipotential form tiny parallelograms. The area of these parallelograms is inversely proportional to \( \{ f, g \} \).

Figure 21: On the left we have a “drawing” on phase space, the equipotential lines for arbitrary functions \( f \) and \( g \). On the right we include the function \( \{ f, g \} \).
The equipotential lines of \( f \) and \( g \) can be regarded as a kind of “drawing” on phase space, as in Fig. 21. Next, choose a third arbitrary function, \( h \), which generates a vector field \( X_h \), as shown in Fig. 22.

\[ q \]
\[ p \]

**Figure 22:** A third arbitrary function \( h \) and the vector field it generates \( X_h \).

Imagine subjecting this “drawing” to the infinitesimal vector flow given by \( X_h \) for some tiny time \( \varepsilon \). This will distort the drawing slightly, as shown in Fig. 23.

\[ q \]
\[ p \]

**Figure 23:** On the left we see how the equipotential lines of \( f \) distort under infinitesimal flow along \( X_h \). On the right we see the end result of this distortion on both \( f \) and \( g \).

The values of \( f \) and \( g \) will change slightly under this distortion. For example, \( f \) will change to

\[
f(q, p) \mapsto f(q - \varepsilon \frac{\partial h}{\partial p}, p + \varepsilon \frac{\partial h}{\partial q}) = f - \varepsilon X_h(f).
\]  

(25)

(The minus sign is there because we want the value of \( f \) to flow along \( X_h \).) Therefore, the changes in \( f \) and \( g \), denoted as \( \delta f \) and \( \delta g \), are

\[
\delta f = -\varepsilon X_h(f) \quad \delta g = -\varepsilon X_h(g).
\]  

(26)
The change in the Poisson bracket \( \{ f, g \} \) is therefore given by

\[
\delta(\{ f, g \}) = -\varepsilon \{ X_h(f), g \} - \varepsilon \{ f, X_h(g) \}.
\]  

(27)

However, there is another way to compute \( \delta(\{ f, g \}) \). Note that under our tiny distortion, the locations of our tiny parallelograms will be translated along the vector \( \varepsilon X_h \). Crucially, by Liouville’s theorem, the area of this parallelogram will not change as our drawing is distorted by the vector field \( X_h \). This is shown in Fig. 24.

![Figure 24: Under the distortion caused by flow along \( X_h \), the area of a tiny parallelogram between the equipotential lines of \( f \) and \( g \) will not change. The location of the parallelogram translates along \( \varepsilon X_h \).](image)

Therefore, the value of \( \{ f, g \} \) will “come along for the ride” unchanged when our drawing is distorted, and

\[
\delta(\{ f, g \}) = -\varepsilon X_h(\{ f, g \}).
\]  

(28)

Combining Eq. (27) and Eq. (28) proves Eq. (24)

Incidentally, the argument given here provides a good way to see why time evolution is a “canonical transformation.” The equation \( \{ q, p \} = 1 \) means that the equipotential lines of \( q \) and \( p \) form a set of equal area parallelograms. Distorting the functions \( q \) and \( p \) by flowing along \( X_h \) will not change their Poisson bracket, which will remain 1.
9 Appendix

9.1 Lie Bracket

Earlier I claimed that if you had two vector fields $X$ and $Y$, the Lie bracket $[X,Y]$ measures the failure of the flows given by $X$ and $Y$ to commute. More specifically, if you take a point and first evolve it along $X$ for some time $\varepsilon_1$ and afterwards evolve it along $Y$ for $\varepsilon_2$, your point will end up in a different place than if you first evolved along $Y$ before $X$. The vector pointing between these two possible endpoints is exactly $\varepsilon_1 \varepsilon_2 [X,Y]$ (to leading order in $\varepsilon_2$ and $\varepsilon_2$). In this section I will prove that claim.

Figure 25: The Lie Bracket of two vector fields $X$ and $Y$ can be understood as measuring the failure of their respective flows to commute infinitesimally.

This is a general fact about vector fields so I will use more general coordinates than $q$ and $p$. We will call our coordinate functions $x^i$, where $i = 1 \ldots n$. The function $x^i : \mathbb{R}^n \to \mathbb{R}$ gives the $i^{th}$ coordinate of a point. We then have the differential operators $\partial_i$ which satisfy

$$\partial_i x^j = \delta^j_i \quad (29)$$

where $\delta^j_i$ is the Kronecker delta.

We begin the proof by writing out the Lie bracket explicitly using coordinates. In these coordinates, we may decompose our vector fields into components:

$$X = X^i \partial_i \quad Y = Y^i \partial_i. \quad (30)$$

Here, $X^i$ and $Y^i$ are the components of our vector fields $X$ and $Y$. They are just functions $X^i, Y^i : \mathbb{R}^n \to \mathbb{R}$. Furthermore, in the notation I am using, there is an implied summation over repeated indices. Therefore, $X^i \partial_i$ is shorthand for $\sum_{i=1}^{n} X^i \partial_i$.

Let us now calculate the components of $[X,Y]$. We accomplish this by acting it on a test function $f$.

$$[X,Y](f) = X(Y(f)) - Y(X(f))$$

$$= X^i \partial_i (Y^j \partial_j f) - Y^j \partial_j (X^i \partial_i f)$$

$$= X^i (\partial_i Y^j)(\partial_j f) + X^i \partial_i Y^j \partial_j f - Y^j (\partial_j X^i)(\partial_i f) - Y^j X^i \partial_j \partial_i f$$

$$= X^i (\partial_i Y^j)(\partial_j f) - Y^j (\partial_j X^i)(\partial_i f)$$
so (interchanging the dummy variables $i \leftrightarrow j$ for half the terms)

$$[X, Y] = X^j(\partial_j Y^i)\partial_i - Y^j(\partial_j X^i)\partial_i. \quad (31)$$

Next, we calculate the coordinate of a point that flows along the vector field given by $X$ for an infinitesimal time $\varepsilon_1$. Say the coordinates of the initial point $x_* \in \mathbb{R}^n$ are the constants $x_i^*$. The coordinates of the point after flow is then just

$$x_i^* + \varepsilon_1 X^i.$$

Let us then flow this resulting point along $Y$, now for the infinitesimal time $\varepsilon_2$. Because the point at which we’re evaluating the vector field has changed, the components of $Y$ are now $Y^i + \varepsilon_1 X^j \partial_j Y^i$ instead of just $Y^i$.

$$x_i^* + \varepsilon_1 X^i + \varepsilon_2 (Y^i + \varepsilon_1 X^j \partial_j Y^i).$$

In total, we see that we have

$$x_i^* + \varepsilon_1 X^i + \varepsilon_2 (Y^i + \varepsilon_1 X^j \partial_j Y^i).$$

We can also see what would happen if we flow along $Y$ first and $X$ second.

$$x_i^* + \varepsilon_2 Y^i + \varepsilon_1 (X^i + \varepsilon_2 Y^j \partial_j X^i).$$

Subtracting the above two equations, we can therefore see that the difference between the flows (the green arrow in Fig. 25) is just

$$\varepsilon_1 \varepsilon_2 (X^j \partial_j Y^i - Y^j \partial_j X^i).$$

This expression matches Eq. (31) thus proving our claim.

9.2 Visualizing $[X_f, X_g] = X_{\{g,f\}}$

In this part I provide a visualization of the identity

$$[X_f, X_g] = X_{\{g,f\}}. \quad (32)$$

This section has been relegated to the appendix because it does not present a “visual proof” in the strict sense, but merely demonstrates that Eq. (32) is plausibly true.

Let’s start by considering two functions $f$ and $g$ whose equipotential lines on phase space are drawn on the left side of Fig. 26.

Recall that $X_f$ points parallel to lines of equipotential and gets larger as the lines of equipotential are closer. The same goes for $X_g$. This gives us a way to directly visualize $[X_f, X_g]$, drawn on the right side of Fig. 26.

Let us now qualitatively understand the properties of the function $\{g, f\}$. Based on Fig. 11, we can see that $\{g, f\}$ is close to zero in the upper right hand corner and large in the lower left hand corner. However, it turns out that in this particular case, “larger” actually means “more negative.” Based on this, we can see that the vector $X_{\{g,f\}}$ should point from the upper left hand to the lower right, as drawn in Fig. 10.

Finally, we can see that based on Fig 26 and 27, the vector fields $[X_f, X_g]$ and $X_{\{g,f\}}$ have the same characteristics, showing that the equation $[X_f, X_g] = X_{\{g,f\}}$ is plausible. I will conclude this section with a set of high quality numerical pictures.
Figure 26: We can infer some properties of $X_f$, $X_g$, and $[X_f, X_g]$ by looking at the equipotential lines of $f$ and $g$.

Figure 27: We can infer some properties of $X_{\{g,f\}}$ by looking at the density of the equipotential lines of $f$ and $g$ in different regions. $\{g, f\}$ has a larger magnitude in the lower-left corner and a smaller magnitude in the upper-right corner. Therefore $X_{\{g,f\}}$ points from the upper-left to the lower-right.
Figure 28: The first graph shows the function \(\{g, f\}\). The second graph shows the vector fields \(X_f\) and \(X_g\). Using these two graphs, you should be able to see that \([X_f, X_g]\) and \(X_{\{g,f\}}\) are the same vector field.