

# From Two Assumptions to Hamilton's Mechanics

by

Ryan Coper

A Thesis Presented in Partial Fulfillment  
of the Requirements for the Degree  
Bachelor of Science in Honors Physics

University of Michigan, Ann Arbor

April 2022

### **Abstract**

This work is a re-derivation of Hamiltonian Mechanics from two assumptions about a classical system. Through this derivation we work to better understand the nature of classical systems and eliminate some of the misconceptions surrounding the relationship between classical theories.

### **Acknowledgments**

I would like to thank Gabriele Carcassi for his guidance and seemingly endless patience and Christine Aidala for her support and encouragement. This work is a part of the larger Assumptions of Physics project at the University of Michigan. This work was funded in part by the University of Michigan Department of Physics.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Infinitesimal Reducibility</b>	<b>5</b>
2.1	Constraint on Coordinate Transformations . . . . .	6
2.2	Unit Variables, Conjugate Pairs, and $2n$ Dimensional State Space . . . . .	9
2.3	Degrees of Freedom, Areas in State Space, and Poisson Brackets . . . . .	11
<b>3</b>	<b>Deterministic and Reversible Evolution</b>	<b>12</b>
<b>4</b>	<b>Conclusion</b>	<b>14</b>
<b>5</b>	<b>References</b>	<b>14</b>

# 1 Introduction

The main purpose of this work is to re-derive Hamiltonian Mechanics from base assumptions at an approachable level. Through this derivation it also seeks to eliminate the misconceptions that plague current instructional literature and motivate a more thoughtful understanding of the subject.

Historically the basis for the laws of physics were generalized experiences. These laws were rooted in the physical world rather than a mathematical one. As it is treated now, Hamiltonian Mechanics lacks this physicality. It is taught as a mathematical reformulation of Newtonian Mechanics rather than its own physical theory. Many popular mechanics textbooks contain passages like this from Thorton and Marion’s *Classical Mechanics of Particles and Systems*: “... we need not formulate a *new* theory of mechanics—the Newtonian theory is quite correct—but only devise an alternate method of dealing with complicated problems.” [1] This assumption of an equivalence between Newtonian and Hamiltonian mechanics is a misconception.

Another more egregious example of this kind of hand waving can be found in Douglas A. Davis’s *Classical Mechanics* “Therefore, it is quite worthwhile to spend some considerable effort in reformulating the ideas held in Newtonian mechanics so we can solve otherwise intractable problems. Remember, this is only a reformulation so, as we have done before, we shall check the results by applying them to already familiar examples. There is no new information or new areas of validity. We will simply restate Newtonian mechanics in another form.” [2] This not only equivocates these two theories, but also asserts that Hamiltonian mechanics is only valid for systems in which Newtonian Mechanics is valid. This is not the case.

Understanding when a theory applies is a critical part of Physics. Hamiltonian Mechanics describes the dynamics of a system by assuming that energy is conserved; we will call this assumption Deterministic and Reversible Evolution later on. Newtonian Mechanics relates the dynamics of a system to its kinematics; this assumption of a bijection between the kinematics and the dynamics is called Kinematic Equivalence.<sup>1</sup> These two different assumptions mean different physical theories are applicable to different systems.

For example consider two systems suspended in a viscous fluid. In the first system the observer is at rest and a particle is moving toward them as it decelerates. In the other system the particle is at rest while the observer is moving toward it while decelerating. Kinematically these systems are identical to the observer, but their dynamics are fundamentally different.

Hamiltonian Mechanics cannot accurately treat the first system as energy is not conserved. Take this as a fact for now; we will see why Hamiltonian Mechanics cannot describe dissipative systems later. Meanwhile, Newtonian Mechanics cannot describe our second system as we are working in a non-inertial frame and as such will observe fictitious forces; there is not an invertible map between the dynamics and the kinematics. Evidently these two formalisms are not equivalent; in fact, the relationship between our three favorite classical theories is best described by Figure 1.

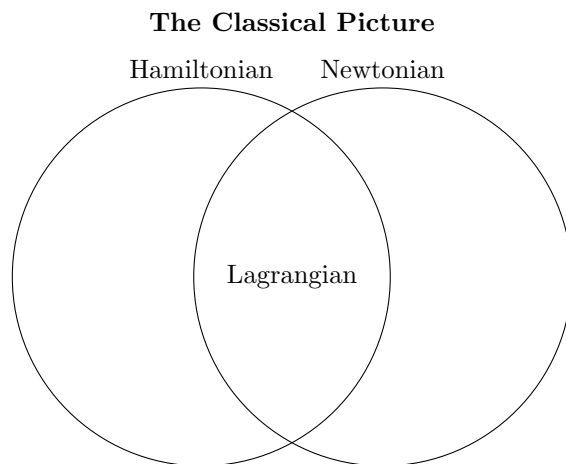


Figure 1: Venn diagram showing the relationship between our three favorite classical theories.

Now that we have motivated our approach, we will begin the derivation with basic physical assumptions about a system in order to show it obeys Hamiltonian Mechanics with a more physical justification. We will

<sup>1</sup>For clarity a system’s dynamics describe how it moves in phase space while its kinematics describe how it moves in physical spacetime.

start with these physical assumptions, then translate those assumptions into precise mathematical definitions. This will lead us to our results.

## 2 Infinitesimal Reducibility

**Assumption I** (Infinitesimal reducibility). *The state of the system is reducible to the state of its infinitesimal parts. That is, giving the state of the whole system is equivalent to giving the state of its parts, which in turn is equivalent to giving the state of its subparts and so on. This relationship holds the other way as well. Giving the states of the smallest subdivisions of the system is equivalent to giving the state of the whole system.*

What does it mean physically for a system to be reducible? This first assumption that underlies classical mechanics prescribes how a system can be broken down into its component parts. This assumption will provide us with a mathematical framework to describe the state of the system we are investigating.

Let's consider a ball that we throw through the air. This ball will follow some path. We assume we can fully describe the state of the ball by its motion through the air. Now, suppose we draw a red dot on this ball. Then if we fully describe the motion of the ball, we have also equivalently described the motion of the red dot. The ball is a reducible system meaning knowledge of the state of the whole is equivalent to knowledge of the states of the parts. The internal dynamics of the system are accessible meaning we can know the states of the parts of the system given knowledge of the state of the whole. Additionally if we give the motion of all possible red dots on the surface of the ball, we know the motion of the whole ball. The state of the whole system can be known from the states of its parts.

Each state of the whole system is uniquely determined by the states of its parts, meaning that given the state of all parts, we can determine exactly the state of the whole system. Two identical systems are in the same state if and only if all of their parts are in the same states. The same is true the other way around. The states of the parts of the system are uniquely determined by the state of the whole. Thus we have an invertible, bijective map between the state of our whole system and the states of its parts. In our example, this means that there is exactly one path traveled by the red dot per distinct state of the whole ball. So if we throw the ball with no angular velocity, the red dot will travel a parabola through the air. We could then throw the ball with the same linear velocity so that its center of mass travels the same path as before, but if the ball is spinning we have a different path traveled by the red dot. So from two distinct states of the overall system, one with angular velocity one without, we find two distinct sets of states of the parts of the system. One spiraling trajectory and one parabola.

We have described reducibility in general, but what does it mean to be infinitesimally reducible? If a system is infinitesimally reducible we can subdivide it as many times as we want and the smallest subdivisions will still be themselves reducible. If we imagine that we divide our system until we approach the limit of this subdivision we will arrive at a collection of infinitesimally small parts of the system; we call these smallest subdivisions *classical particles*.

What does this mean physically? Continuing our previous example, imagine we cover the surface of the ball in red dots. We then remove these dots and replace each of them with a number of smaller dots. We repeat this process until the radius of the dots is approaching zero. Because the ball is assumed to be a reducible system, given the state of the whole ball, we know the states of all of the red dots drawn on it. It is important to note that the dots do not ever become points. They always have radii that are greater than zero. No dot is a point particle, rather they are infinitesimally small subdivisions. *We can always divide them again.* Assuming a system to be infinitesimally reducible means that given the state of the whole system, we can describe the states of all the infinitesimally small parts of the system. Now we must codify this idea formally.

Let's start with a discrete system as it is conceptually simpler. Consider a box with a fixed number of balls of varying colors as our system.<sup>2</sup> Suppose the state of the whole system is determined by the number of balls of each color. We will define the *state space* of the balls as  $\mathcal{S}$  which spans all possible colors a ball can be. Each ball is at a point  $s \in \mathcal{S}$  corresponding to its color. We call the state space of our whole system  $\mathcal{C}$  which spans all possible color combinations of the balls. One state of the whole system is a point  $c \in \mathcal{C}$ . For a state  $c \in \mathcal{C}$  we can define the number of balls in the box of each color as a *distribution*,  $\rho$ , over the state space  $\mathcal{S}$  of the balls. This distribution is unique to each state of the whole system. Two identical systems are in the same state if and only if their distributions are equivalent. For each distinct state  $c \in \mathcal{C}$  of the whole system, we have exactly one  $\rho$ .

---

<sup>2</sup>This is not an example of an infinitesimally reducible system as the balls are the limit of its reducibility, but it serves as an illustrative example.

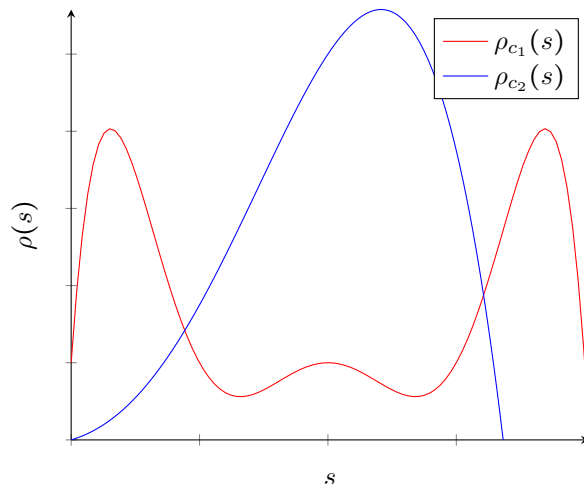


Figure 2: Two different states of the whole system map to two different distributions of the parts of the system over  $\mathcal{S}$ . Functions are not normalized.

For each subset  $U$  of  $\mathcal{S}$ , we can count the number of balls that are those colors; call this counting function  $\mu : \mathcal{S} \rightarrow \mathbb{R}$ . That is, for each  $U \subseteq \mathcal{S}$  we have

$$\mu(U) = \frac{\sum_{s \in U} \rho(s)}{\sum_{s \in \mathcal{S}} \rho(s)}.$$

This function is normalized by definition because  $\mu(\mathcal{S}) = 1$ .

If our system is continuous the main ideas above hold, but we must make two changes to the formalization. First, our distribution  $\rho : \mathcal{S} \rightarrow \mathbb{R}$  becomes a continuous function that describes of states over the space and second,  $\mu$  becomes an integral:

$$\mu(U) = \frac{\int_U \rho(s) d\mathcal{S}}{\int_{\mathcal{S}} \rho(s) d\mathcal{S}}$$

where  $d\mathcal{S}$  gives the number of states in an infinitesimal area. We will drop the denominator from now on unless it is necessary for notation's sake but keep in mind this normalization is implicit. See Figures 2 and 3 for a visual representation of this phase space and counting function.

For a physical example let's consider a classical gas distributed in a volume. The state of the whole is the distribution. Now we can see that our counting function will count the amount of gas in a region of this volume by integrating the density over the region in question.

**Definition 1.** Let  $\mathcal{C}$  be the state space of a system. The system is *infinitesimally reducible* to the infinitesimal parts (i.e. particles) identified by the state space  $\mathcal{S}$  if there exists a measure  $d\mathcal{S}$  such that for every state of the whole system  $c \in \mathcal{C}$  there exists  $\rho : \mathcal{S} \rightarrow \mathbb{R}$  such that for every  $U \subseteq \mathcal{S}$  the counting function  $\mu(U)$  corresponds to the fraction of the system found within those states; as usual this definition means  $\mu(\mathcal{S}) = 1$ .

## 2.1 Constraint on Coordinate Transformations

If we want to further describe the state of the system we need to identify our states in  $\mathcal{S}$  with numbers so that all states are uniquely labeled. This system of labeling will give us our *state variables*. For example imagine an ideal gas in a box. We know that we can describe the state of this gas by giving its pressure, temperature, or volume. By giving two of these quantities we have completely described the state of the gas as the third can be ascertained from knowledge of the other two through the ideal gas law. The state of the gas is uniquely determined by the values of two quantities. Thus we have a one to one map from the state space to the values of a set of numerical quantities. There is a precise number of state variables for every system. From our ideal gas example if we were to also include the third quantity we would have introduced a redundant label. For each pressure-volume pair there is exactly one possible temperature so if we include temperature in our labeling we no longer have a one to one map between the state space and sets of values as

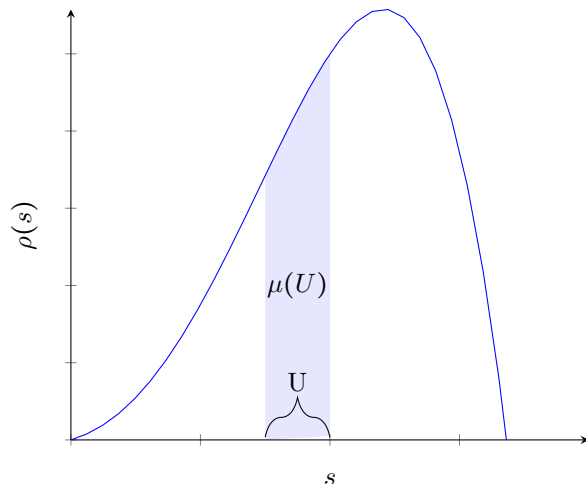


Figure 3: One dimensional continuous example of our counting function  $\mu(U)$ .  $U$  is a region of phase space and  $\mu(U)$  is the fraction of the system that is found in this region.

some triplets do not correspond to possible states. For each state there is one well defined pressure-volume pair.

**Definition 2.** A state variable assigns one numerical quantity to each state. Formally, it is a map  $\xi : \mathcal{S} \rightarrow \mathbb{R}$ .<sup>3</sup> A complete set of state variables fully identifies a state. Formally, it is an invertible map  $\xi^a : \mathcal{S} \rightarrow \mathbb{R}^n$ . This gives us a state space that is locally isomorphic to  $\mathbb{R}^n$  thus our state space is an  $n$ -dimensional **manifold**.

Expressing  $\rho$  in terms of our new state variables is the next natural step. Consider a cannon ball that is shot through the air. If the ball is spinning the position and momentum of each infinitesimal region of the ball will vary. Let the position and momentum be our state variables. The parts of the system will be distributed across these positions and momenta. Writing our distribution in terms of these variables will give us a density. We want  $\rho$  to be a map  $\rho(\xi^a) : \mathbb{R}^n \rightarrow \mathbb{R}$ . When we defined our state variables we noted they created an invertible map between states and numbers labeling those states. Using this we see  $\rho(\xi^a)$  is simply our distribution function  $\rho(s)$  composed with the inverse of  $\xi^a : \mathcal{S} \rightarrow \mathbb{R}^n$ . That is:

$$\rho(\xi^a) = \rho(s(\xi^a)).$$

This density is a map from our state variables to a numerical value that describes how much of the system occupies a region on this manifold.

Note that the density and the distribution are distinct functions that describe the same thing; specifically the distribution is  $\rho(s) : \mathcal{S} \rightarrow \mathbb{R}$  and the density is a composition of the distribution function with the state variables i.e.  $\rho(\xi^a) : \mathbb{R}^n \rightarrow \mathbb{R}$ . The density requires labeling the space with state variables to be defined while the distribution does not.<sup>4</sup>

Under a change of state variables  $\rho(\xi^a)$  must have a well defined transformation rule. For example say we have the density as a function of  $x$  and  $y$ . Then if we change to polar coordinates we now want  $\rho(r, \theta)$ . How do we find  $\rho(r, \theta)$  in terms of  $\rho(x, y)$ ? With a little thought we can write

$$\rho(x, y) = \rho(r, \theta) \begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix}.$$

Evidently these partial derivatives need to be defined to have a well defined transformation of  $\rho$ . More generally our transformation rule is of the form

$$\rho(\xi^a) = \rho(\hat{\xi}^b) \begin{vmatrix} \frac{\partial \hat{\xi}^b}{\partial \xi^a} \end{vmatrix}.$$

<sup>3</sup>Note this  $\mathcal{S}$  indicates a subspace of the particle state space  $\mathcal{S}$ . Each  $\xi$  maps a subspace  $S \leq \mathcal{S}$  to  $\mathbb{R}$ .

<sup>4</sup>Note that we will use  $\xi^a, \xi^b, \xi^c$  for state variables;  $q^i, q^j, q^k$  for unit variables;  $\xi^\alpha, \xi^\beta, \xi^\gamma$  for state variables that include time (e.g. four momentum). For distributions/densities we use  $\rho$ ; where we need to distinguish between the two we will use  $\rho(s)$  for the distribution and  $\rho(\xi^a)$  for the density. States are indicated with  $s$  and state variables with  $\xi$ . Note the duplicity of  $\rho, s$ , and  $\xi$ . We use  $\xi^a(s)$  as  $\mathcal{S} \rightarrow \mathbb{R}^n$ ;  $s(\xi^a)$  as  $\mathbb{R}^n \rightarrow \mathcal{S}$ . Thus we have the following identities:  $\rho(s(\xi^a)) = \rho(\xi^a)$ ;  $\rho(s) = \rho(s(\xi^a(s)))$ .

In our example above we do have one point at which the transformation is not well defined: the origin. We do not have well defined partial derivatives at the origin as theta is not well defined at that point. Thus our density is not well defined at the origin in polar coordinates.

We will require that any state variables labeling our state space be *differentiable* over all of  $\mathcal{S}$ . This ensures we can define a change of variables without introducing points in  $\mathcal{S}$  where our density is not well defined. So our state space will be a manifold labeled by differentiable state variables.

**Proposition 3.** *The state space of our classical particles  $\mathcal{S}$  is an  $n$ -dimensional **differentiable manifold**.*

As we saw above, the density  $\rho(\xi^a)$  will be affected by the labeling of our state space. We also know that our distribution,  $\rho(s)$  should be independent of our choice of state variables; the way the system is distributed over the state space does not change even if we choose new state variables to label the space. This will give us an additional consideration when thinking about how  $\rho$  transforms.

Consider an example using temperature. We can give the temperature at a point in space without a coordinate system. Take a point  $A$  in three-dimensional space. There is no need for a conception of any other unit to define the temperature. Formally this means we can write  $T(A) : \mathcal{M} \rightarrow \mathbb{R}$ . This is a direct map from spacetime to a number; we do not need to label spacetime using a coordinate system. This independence also means logically that changing our coordinate system does not change the numerical value of temperature at a point nor its units. We can write  $T(A) = T(x, y) = T(r, \theta)$  meaning that as long as the point we are discussing remains the same, the value of  $T$  is the same. So temperature is *coordinate independent* and invariant under change of coordinates.<sup>5</sup>

We know that the amount of a system occupying a region of state space does not change with the variables we use to label that space. Our distribution  $\rho(s)$  is coordinate independent thus it is unaffected by the choice of state variables  $\xi^a$ . We want this to translate to our density  $\rho(\xi^a)$  because the distribution and the density represent the same thing: how the system is distributed over state space. This means that under a change of state variables density will be invariant i.e.  $\rho(s(\xi^a)) = \rho(s(\hat{\xi}^b))$ .

We have found two seemingly contradictory transformation rules for our density  $\rho(\xi^a)$ . On one hand,  $\rho$  must transform as a density i.e. its value will vary under a change of variables; on the other hand  $\rho$  must be invariant under a change of variables because it should match the distribution. We solve this issue by restricting which state variables we use to label our state space. We will throw out any choice of variables that do not produce a  $\rho$  that obeys both of these restrictions.

**Proposition 4.** *Because it is a density and is coordinate invariant,  $\rho$  will transform as a scalar and a density under a change of state variables  $\hat{\xi}^b = \hat{\xi}^b(\xi^a)$ . Thus  $\rho(\xi^a) = \rho(\hat{\xi}^b) \left| \frac{\partial \hat{\xi}^b}{\partial \xi^a} \right|$  and  $\rho(s(\xi^a)) = \rho(s(\hat{\xi}^b))$ . The state variables that allow  $\rho$  to be expressed in a way that satisfies these conditions are called **canonical variables**.*

Using these two conditions together we have that  $\left| \frac{\partial \hat{\xi}^b}{\partial \xi^a} \right| = 1$ . Using our counting function let us see if we can make a little more sense of this physically.

We can write our counting function  $\mu(U)$  in terms of our state variables as  $\mu(U) = \int_{\xi(U)} \rho(\xi^a) \prod_a d\xi^a$ . If we change our state variables we know that that fraction of the system in a region  $U \subseteq \mathcal{S}$  will not change. Thus  $\mu(U)$  must be invariant under such a change. Let  $\hat{\xi}^b = \hat{\xi}^b(\xi^a)$  be our change of variables. This gives us

$$\mu(U) = \int_{\xi(U)} \rho(\xi^a) \prod_a d\xi^a = \int_{\hat{\xi}(U)} \rho(\hat{\xi}^b) \prod_b d\hat{\xi}^b.$$

In more concrete terms we know that to find the total mass in a region  $U$  from a density we write the integral

$$\int_U dm = \int_U \rho_m dV.$$

If we want to conserve the value of this integral under a change of variables we know that if  $\rho_m$  changes, the value of  $dV$  must change accordingly. Proposition 4 however, tells us that the value of  $\rho$  cannot change, thus we know the value of our 'volume' element  $d\xi^a$  cannot change either. Great! This is exactly what proposition 4 tells us.

**Proposition 5.** *The state space  $\mathcal{S}$  must admit canonical state variables. A change of state variables must be differentiable and must have a unitary Jacobian i.e.  $\left| \frac{\partial \hat{\xi}^b}{\partial \xi^a} \right| = 1$ . Such a transformation is a **canonical transformation**.*

<sup>5</sup>We say a quantity is *coordinate independent* if it can be defined without a coordinate system. To be *coordinate invariant* means it will not change under a change of coordinates.



Symbol	Name	Description
$\mathcal{C}$	State Space of Whole System	Contains all possible states of the whole system. Each state $c$ maps to a unique distribution $\rho(s)$
$\mathcal{S}$	State Space of Classical Particles	Contains all possible states of the parts of the system. Particles are distributed according to $\rho(s)$ .
$\mathcal{M}$	Space-Time	Physical spacetime.
$\rho(s)$	Distribution of States	$\mathcal{S} \rightarrow \mathbb{R}$
$\rho(\xi^a)$	Density of States	$\mathbb{R}^n \rightarrow \mathbb{R}$
$\mu(U)$	Counting function	$\mathbb{P}(\mathcal{S}) \rightarrow \mathbb{R}$
$\xi^a$	State variables	$\mathcal{S} \rightarrow \mathbb{R}^n$
$\rho(\xi^a)$	Density	$\mathbb{R}^n \rightarrow \mathbb{R}$
$\hat{\xi}^b(\xi^a)$	Change of Variables	$\mathbb{R}^n \rightarrow \mathbb{R}^n$

## 2.2 Unit Variables, Conjugate Pairs, and 2n Dimensional State Space

We know that to define a density we need some kind of units of area or volume. We must remember that we are working with physical systems here. This means that the units of our state variables are important. Some of our state variables will define a unit and others will have their units derived from the units of other state variables. For example if distance in a direction is defined in terms of meters and time is defined in terms of seconds, velocity in that direction must be defined in meters per second and acceleration must be of units of meters per second squared. The units of velocity and acceleration are derived from the definition of the units of distance and time. Changing the definition of a fundamental unit will also change the definitions of all the units derived from it.

Another example of this relationship between variables and their units is the set energy, entropy, and temperature. If we define the units of energy and entropy, we have no choice in the units of temperature because the relation  $TdS = dU \rightarrow T = \frac{dU}{dS}$  ensures that the units of temperature are equivalent to the units of energy divided by the units of entropy. We could just as easily define the units of energy and temperature which would give us the units of entropy through the same relation.

We will define a *unit system* using a subset of our state variables which we will call our *unit variables*. These unit variables will define the fundamental units underlying all other state variables. Changing these unit variables will induce a change in all of the other state variables that rely on the definition of that unit. Take our distance, velocity, and acceleration example; if we change the units of distance from meters to light years, our definitions of velocity and acceleration must change as well, but our units of time remain unchanged. Conversely if we change our units of time, the units of velocity and acceleration will still change, but our unit of distance is unaffected. Unit variables are independent of one another. Ostensibly, the definitions of temperature and time have nothing to do with one another. Generalizing this fact we see that if we change one unit variable we do not induce a change in any other unit variables. This change of units must be unique given our knowledge that state variables are invertible bijections.

**Definition 6.** We define a unit variable  $q \in \xi^a$  as a state variable that is the definition of a unit. The subset of unit variables  $q^i \subset \xi^a$  defines a unit system upon which the other state variables depend meaning that a transformation  $\hat{q}^j = \hat{q}^j(q^i)$  induces a unique change of state variables  $\hat{\xi}^b = \hat{\xi}^b(\xi^a)$ . Unit variables are independent of one another.

Now we must ask how a change of unit variables should affect the units of our density  $\rho(\xi^a)$ . Let's consider a typical example of a density: mass distributed over three dimensional space  $\mathcal{M}$ . If I have a mass distribution, and I ask you how much mass is there in a region  $U \subseteq \mathcal{M}$ , your answer would be a number followed by a unit. This response would be coordinate independent as the definition of the region  $U$  does not require a coordinate system; your answer depends only on the defined region of interest and the units of mass; there is no dependence on the choice of coordinates used to describe the space. We can talk about points in this space without a defined coordinate system. Take a point  $A \in \mathcal{M}$ ; this is the same point in all references, and we do not need to have a coordinate system defined to talk about point  $A$ . If I ask for the mass density at that point, however, I need to specify a unit like  $m^3$  to describe volume. Formally, I have

to specify the unit system of the space. In math terms this means that the density function is defined as  $\rho_m : \mathbb{R}^3 \rightarrow \mathbb{R}$  not  $\rho_m : \mathcal{M} \rightarrow \mathbb{R}$  just as we saw in the previous section. We cannot define a map directly from  $\mathcal{M}$  to  $\mathbb{R}$  because without a well defined notion of distance, density makes no sense. If we were to change our unit system to  $cm^3$  the actual numerical value of the density at each point changes. The change of units labeling the space also means that the units of the density change as well. The unit system of the state space will determine the units used to describe density. See Figure 4 for an explicit example.

As we saw in section 2.1 the Jacobian of a change of variables will give the transformation rule of the density; the transformation rule for the units of the density will similarly be given by the units of the Jacobian. But we also know that like our counting function  $\mu(U)$ , we want our density to be unaffected by changes in our state variables in both its value and units. Because we will only use canonical state variables, we have ensured that the Jacobian of a change of state variables will be *unitary*. This means that not only will the value be equal to one, but also that it will be unitless. This ensures that like our counting function,  $\rho$  will have the same units after a change of variables.

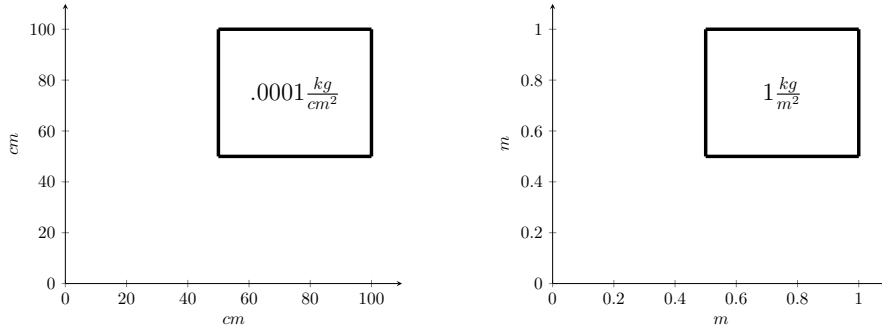


Figure 4: Visualization of the how the units with which we label our space affect the density. Both areas contain exactly  $.25kg$  uniformly distributed, but their density functions are different because their unit systems are different.

Now we will examine formally this change of units. Let's start with the simplest case: one unit variable suffices to describe our unit system. We then can write our set of state variables as  $\xi^a = \{q, k^\mu\}$  where  $k^\mu$  could represent any number of other state variables. Let's suppose we perform an arbitrary change of units. We then have  $\hat{q} = \hat{q}(q)$ . How must our set  $\{k^\mu\}$  change while upholding our restriction to a unitary Jacobian? Our Jacobian block matrix will be of the form

$$1 = \begin{vmatrix} \partial \hat{\xi}^b \\ \partial \xi^a \end{vmatrix} = \begin{vmatrix} \frac{\partial \hat{q}}{\partial q} & \frac{\partial \hat{q}}{\partial k^\mu} \\ \frac{\partial \hat{k}^\nu}{\partial q} & \frac{\partial \hat{k}^\nu}{\partial k^\mu} \end{vmatrix}.$$

If there are no  $k^\mu$  variables then our constraint equation becomes  $\left| \frac{\partial \hat{\xi}^b}{\partial \xi^a} \right| = \left| \frac{\partial \hat{q}}{\partial q} \right| = 1$ . This does not allow us to make an arbitrary change in units meaning this case is invalid. We must have one or more  $k^\mu$  variable(s).

Because  $\hat{q}$  has no dependence on  $k^\mu$  we know that  $\left| \frac{\partial \hat{q}}{\partial k^\mu} \right| = 0$ . This means that our constraint becomes  $\left| \frac{\partial \hat{q}}{\partial q} \right| \left| \frac{\partial \hat{k}^\nu}{\partial k^\mu} \right| = 1$ . We know that  $\hat{q}$  only depends on  $q$  so we can write  $\left( \frac{d\hat{q}}{dq} \right)^{-1} = \left| \frac{dq}{d\hat{q}} \right|$ . We then get  $\left| \frac{\partial \hat{k}^\nu}{\partial k^\mu} \right| = \left| \frac{dq}{d\hat{q}} \right|$ . This provides only one constraint on our transformation. We know that we must have a unique transformation. So if we have two or more  $k^\mu$  variables we cannot say that  $q$  fully defines the unit system for all other state variables as a change in  $q$  does not induce a unique change in the other state variables as it should. Thus we must have exactly one  $k$  variable and our manifold will be two dimensional. This  $k$  variable is called a *conjugate variable*.

What are the units of the conjugate variable  $k$ ? Recall the earlier equality  $\left| \frac{d\hat{q}}{dq} \right| \left| \frac{\partial \hat{k}}{\partial k} \right| = 1$ . It is important that we know this equality gives us a unitless quantity. This means we can relate the units of  $k$  to the units of  $q$ . We determined earlier that  $k$  is a function only of  $q$  and as such the units of  $k$  written as  $[k]$  must only be a function of the units of  $q$ . Rearranging the earlier equality gives us

$$\left| \frac{d\hat{q}}{dq} \right| = \left| \frac{\partial k}{\partial \hat{k}} \right|.$$

This means we will write the units of  $k$  as  $[k] = \frac{1}{[q]}$ .

How does this generalize to higher dimensions? Let's consider a state space whose unit system is described by  $\{q^i\}_{i=1}^n$  where each  $q^i$  is an independent unit variable. This means that a change in one of the unit variables will not affect the other unit variables and that this change will be a function only of the original unit variable. Now we will perform individual changes of unit variables. We will write the first of these changes as  $\hat{q}^1(q^1) = \hat{q}^1$ . Writing out the resulting change of units explicitly:

$$\{\hat{q}^j\}_{j=1}^n = \{\hat{q}^1(q^1), q^2, \dots, q^n\}.$$

Because we are fixing all of the  $\{q^i\}_{i=2}^n$  variables, this case is the same as the one unit variable system. This means that we will find precisely one conjugate  $k_1$  variable that corresponds to  $q^1$  whose units are the inverse of those of  $q^1$ . Continuing this process we see that the set of state variables must be given by  $\{q^i, k_i\}_{i=1}^n$  and that we must have a  $2n$  dimension manifold.

**Proposition 7.** *The state space of the particles is  $2n$  dimensional. The state variables are organized in pairs  $\{q^i, k_i\}$  where the set  $\{q^n\}$  defines the unit system. These state variables have a well defined transformation rule with a unitary Jacobian.*

**Proposition 8.** *The units of conjugate variables  $\{k_i\}$  are the inverse of the units of the corresponding unit variables  $\{q^i\}$ . The volume element  $d\xi^a = d\xi^{2n} = dq^n \wedge dk_n$  is invariant in value and units under a canonical change of variables.*

We have now successfully described the structure of phase space by making one assumption about our system. The next section will elaborate on this structure and discuss the units we use to count states before we move on to our second assumption.

## 2.3 Degrees of Freedom, Areas in State Space, and Poisson Brackets

We showed in section 2.1 that the volume element expressed in canonical variables  $\prod_a d\xi^a = d\xi^1 d\xi^2 \dots d\xi^{2n}$  is invariant under a canonical transformation. We want to relate this differential to our measure  $dS$ . How do the variables with which we label our state space relate to the number of possible states in that space? This requires a few considerations. What units will we use to count these states? The convention in statistical and quantum mechanics is to use  $\hbar$  as the units of number of states; we will continue this convention. We define the units of our measure  $dS$  to be  $\hbar$ . We write  $p = \hbar k$  in order to define a new conjugate variable  $p$  that depends not only on the unit of  $q$ , but also on the units used to count states. Here we are exercising our freedom to choose the units of the conjugate variables. We are giving  $p$  the units of number of states over the units of  $q$ . This choice allows us to relate the state variables labeling the space to the number of states contained in that space in a precise way.

In defining  $p$  this way we arrive at the relation  $dS = \prod_{i=1}^n dq^i dp_i$ . From section 2.2 we know that each conjugate variable is a function of one and only one unit variable. This means that each individual pair of  $\{q^i, p_i\}$  variables will chart a 2D surface in phase space that can be described independently of the other state variables. We will call these pairings and the surfaces associated with them *degrees of freedom* (d.o.f.).

If we want to count the number of possible states over one d.o.f. we can integrate using the measure  $dS^i$ . The integral that gives the number of possible states in a region  $U_i$  of our surface is  $\int_{U_i} dq^i dp_i = \int_{U_i} dS^i$ .

**Definition 9.** *A degree of freedom (d.o.f.) is a pair  $\{q^i, p_i\}$  such that  $dS^i = dq^i dp_i$ . The measure  $dS^i$  quantifies the number of possible configurations within an infinitesimal surface identified by that degree of freedom.*

How do we count states over multiple degrees of freedom? For example imagine we have a state space of four dimensions where  $q^1, q^2$  are positions along the  $x$  and  $y$  axes and  $p_1, p_2$  are momenta along those axes. The two pairings  $\{q^1, p_1\}$  and  $\{q^2, p_2\}$  are our degrees of freedom. We can count the total number of possible states by first counting the possible configurations of position and momentum in each dimension then multiplying these two counts together because the degrees of freedom are independent: the choice of a configuration in one has no effect on the choice of a configuration in the other. Thus the surfaces are orthogonal and the total number of possible states is given by the product of the possibilities in each degree of freedom.

**Definition 10.** *Two degrees of freedom are **independent** if the number of configurations identified by them is the product of the configurations identified by the individual degrees of freedom. That is,  $dS = dS^1 dS^2$ . Since the volume is the product of the areas, independent degrees of freedom are orthogonal.*

How do we determine if a pair of variables forms a degree of freedom? Not all variables are compatible with each other. Consider our earlier example of 4D state space. We could not define the  $x$ -position and  $y$ -momentum as a degree of freedom because a change of units of the  $x$ -position will not induce a change in the units of  $y$ -momentum. To formalize this we will define an operator that tells us if two state variables are compatible.<sup>6</sup> This operator will take two state variables  $f, g$  and return 0 if the variables are not compatible and do not form a d.o.f. or the unit change of the area  $dfdg/dS$  if they are. This operator is called a *Poisson bracket*.

**Definition 11.** *The Poisson bracket is defined as  $\{f, g\} = \sum_i \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q^i} \frac{\partial f}{\partial p_i}$ .*

Let's suppose we have canonical variables such that each  $\{q^i, p_i\}$  form a degree of freedom. If  $f$  and  $g$  are canonical variables that form the same d.o.f. as one of these pairs then the Poisson bracket of  $f$  and  $g$  is  $\{f, g\} = dfdg/dS^i$ . If not  $\{f, g\} = 0$ . We know that because our degrees of freedom are independent and because we've limited ourselves to the use of differentiable state variables only, there will exist an invertible transformation between  $f, g$  and  $q^i, p_i$ . If we recall that the Poisson bracket returns the unit change of area  $dfdg/dS^i$ , we see that the Jacobian of this transformation is precisely the Poisson bracket of  $f$  and  $g$ .

**Proposition 12.** *The Poisson bracket  $\{f, g\}$  translates the densities of states into densities per unit area of  $f, g$ . It is the Jacobian of the transformation  $dfdg \rightarrow dq^i dp_i$  i.e.  $dfdg = \{f, g\} dq^i dp_i$ .*

### 3 Deterministic and Reversible Evolution

**Assumption II** (Deterministic and Reversible Evolution). *The system undergoes deterministic and reversible evolution meaning given the state of the system at any time, its state at all past and future times is known.*

We will now make an assumption about the time evolution of the system. The resulting mathematical structure will describe this evolution in terms of our state variables.

For a system to undergo deterministic and reversible evolution there must be a bijection between all past and future states of the classical particles that make up the composite system. We must also have that the density of states,  $\rho$ , is conserved over time. Now this does not mean that our classical particles cannot change state. It means that if we have a certain number of classical particles in a initial state, all of those particles and no others will evolve together and end up in the same final state. For example, if half of the system begins in state A and one of these particles is later found in state B then we know that exactly half of the system is found in state B at that time.

Why must we have the additional requirement of the conservation of the density when it seems the existence of the bijection would be sufficient for the evolution of our system to be deterministic and reversible? Here is an example of a system for which we can write such a bijection, but for which the density of states is not conserved. We will see that the condition of the bijection alone is not strong enough to provide us with the desired mathematics.

Imagine we have a one dimensional damped harmonic oscillator. We have a one to one map from past states to future states. But we also know that no matter the initial state, as the system evolves it will approach its rest state because the oscillator is damped. Our state variables will be position and momentum. We know that in the real world all measurements have finite precision. This informational granularity tells us that at some point in the above example we will not be able to tell the difference between the state in which the oscillator is barely moving and its rest state; a measurement will not be able to distinguish between the two states, see Figure 5 for a visualization of this. Thinking of this in terms of our knowledge of the system we see that as time progresses and our states collapse towards the rest state, we lose information about the system. This means in reality our system is not reversible because once it passes a certain point we cannot deduce the past state of the system from its current state; when the system is at rest we have no idea of its initial condition. Thus the addition of the requirement that density be conserved over time is necessary to ensure we can always deduce the past states of the system from its current state in a physically meaningful way.

**Definition 13.** *A deterministic and reversible system is one for which all particles in a state are mapped to one and only one past or future state. That is, given a set of initial values  $\xi_0^\alpha$  at a time  $t_0$ , there is only one possible evolution  $\xi^\alpha(t)$ , such that  $\xi^\alpha(t_0) = \xi_0^\alpha$ . Moreover,  $\rho(\xi^\alpha(t))$  is a constant.*

<sup>6</sup>This idea may feel familiar. That's because the operator we will define is the classical equivalent of the quantum commutator.

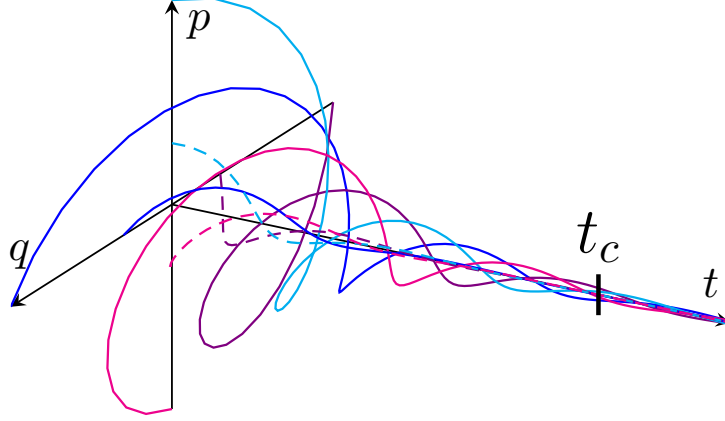


Figure 5: The evolution of a damped harmonic oscillator in phase space from 8 different initial conditions. At some critical time  $t_c$  we can no longer distinguish this system from the oscillator at rest given finite precision of measurement.

We will continue to consider a system with a single degree of freedom and state variables  $\{q, p\}$ . If we transport our 2D state space in time  $t$ , we can map the evolution of each classical particle in  $\mathbb{R}^3$ . The evolution of the state of each particle can be written as  $\xi^\alpha(t)$  and the fraction of the system that is found in that state is  $\rho(\xi^\alpha(t))$ . Note we are using  $\xi^\alpha$  to denote the state variables in the extended phase space that includes time i.e.  $\xi^\alpha = \{\xi^\alpha, t\}$ .

If we change infinitesimally in time we can write  $\xi^\alpha(t+dt) = \xi^\alpha + \frac{d\xi^\alpha}{dt} dt$ . We know that  $\xi^\alpha(t)$  is differentiable because  $\rho$  is always defined. We then will write a vector field  $S^\alpha = \frac{d\xi^\alpha}{dt} \implies S = \left(\frac{dq}{dt}, \frac{dp}{dt}, \frac{dt}{dt}\right)$ . This vector field describes the time evolution of our state variables.

Thinking about our damped oscillator example, we see that the vector field  $S$  for this system will have a nonzero divergence because the states of the classical particles starting from different initial conditions get closer together as the system evolves towards its rest state as illustrated in Figure 5.

Instead applying the requirement that  $\rho(\xi^\alpha(t))$  be conserved in time, we see that the divergence of the vector field must be zero:

$$\nabla \cdot S = \partial_\alpha S^\alpha = \frac{\partial S^q}{\partial q} + \frac{\partial S^p}{\partial p} + \frac{\partial S^t}{\partial t} = 0.$$

Note that this means that we can write our vector field  $S$  as the curl of a vector potential by the fundamental theorem:

$$S = \nabla \times (-\theta + \nabla f)$$

where  $\theta$  is our vector potential and  $\nabla f(q, p, t)$  is a gauge term. In component notation this equation is

$$S^\alpha = \epsilon^{\alpha\beta\gamma} (\partial_\beta (\theta_\gamma + \partial_\gamma f)).$$

We know that  $\theta$  will not be unique because we can choose what gauge we work in.<sup>7</sup> We will write theta as  $\theta = (\theta_q, \theta_p, \theta_t)$  and choose our gauge such that  $\theta_p = 0$  by setting  $\frac{\partial f}{\partial p} = -\theta_p$ . This means we can write

$$\theta = (\theta_q, 0, \theta_t)$$

without loss of generality. Plugging this back into  $\nabla \times \theta = S = \left(\frac{dq}{dt}, \frac{dp}{dt}, \frac{dt}{dt}\right)$  and looking at the third component we get

$$\frac{dt}{dt} = \frac{d\theta_q}{dp}.$$

Here we see that  $\frac{dt}{dt} = 1 = \frac{d\theta_q}{dp}$ .<sup>8</sup> Integrating with respect to  $p$  gives us

$$\theta_q = p + g(q, t)$$

<sup>7</sup>Remember the curl of the gradient of a scalar field is always zero so our choice of  $f$  has no effect on  $S$ . That is to say  $S$  is gauge invariant.

<sup>8</sup>Note here that the  $t$  in the numerator of this differential is technically different from the  $t$  in the denominator. In the numerator  $t$  is an affine parameter while in the denominator it is a state variable.

where  $g$  is an arbitrary function and thus can be set to zero. We have arrived then at  $\theta = (p, 0, \theta_t)$ . Let  $\theta_t \equiv -H$ . Plugging back in again to the cross product equation we arrive at a familiar system of equations

$$\begin{aligned} S^q &= d_t q = \partial_p H \\ S^p &= d_t p = -\partial_q H \\ S^t &= d_t t = 1. \end{aligned}$$

This is the standard form of Hamilton's equations for a single degree of freedom where  $H$  is the Hamiltonian.

Generalizing this to a system with  $n$  degrees of freedom we recall that each infinitesimal surface in phase space is mapped in time to a surface of the same area; this means that the number possible states remains the same and that independent degrees of freedom remain independent and thus their surfaces remain orthogonal. We can break down a higher dimensional system into its d.o.f.s in order to arrive at our final equations:

$$\begin{aligned} d_t q^i &= \partial_{p_i} H \\ d_t p_i &= -\partial_{q^i} H. \end{aligned}$$

This is the generalized form of Hamilton's equations.<sup>9</sup> A system that evolves according to these equations is a Hamiltonian system.

**Proposition 14.** *The deterministic and reversible evolution of an infinitesimally reducible system follows Hamilton's equations. So the evolution of the system can be written in the familiar way:*

$$\begin{aligned} d_t q^i &= \partial_{p_i} H \\ d_t p_i &= -\partial_{q^i} H \end{aligned}$$

## 4 Conclusion

We have now successfully derived all of Hamiltonian Mechanics from two basic assumptions about a system. By assuming that our system is Infinitesimally Reducible we derived the structure of phase space. With the assumption that the system undergoes Deterministic and Reversible Evolution we found a set of equations that describe the dynamics of the system. This is fundamentally different from Newtonian Mechanics, which assumes a connection between a system's kinematics and its dynamics.

This derivation of Hamiltonian Mechanics, while seemingly miraculous, is simply a product of creating a one-to-one mapping of physical principles to mathematical structures. This precise way of understanding the physical world provides a much more clear picture of the underlying physics within the abstract mathematics.

## References

- [1] S. T. Thornton and J. B. Marion, *Hamilton's Principle - Lagrangian and Hamiltonian Dynamics*, p. 228–230. Cengage Learning, 5 ed., 2014.
- [2] A. D. Davis, *Classical mechanics*. Elsevier Science, 2012.
- [3] G. Carcassi, C. A. Aidala, D. J. Baker, and L. Bieri, "From physical assumptions to classical and quantum hamiltonian and lagrangian particle mechanics," *Journal of Physics Communications*, vol. 2, p. 045026, apr 2018.

---

<sup>9</sup>For a much more in depth treatment of this generalization see [3].