THE LAGRANGIAN AND HAMILTONIAN MECHANICAL SYSTEMS

ALEXANDER TOLISH

Abstract. Newton's Laws of Motion, which equate forces with the time-rates of change of momenta, are a convenient way to describe mechanical systems in Euclidean spaces with cartesian coordinates. Unfortunately, the physical world is rarely so cooperative—physicists often explore systems that are neither Euclidean nor cartesian. Different mechanical formalisms, like the Lagrangian and Hamiltonian systems, may be more effective at describing such phenomena, as they are geometric rather than analytic processes. In this paper, I shall construct Lagrangian and Hamiltonian mechanics, prove their equivalence to Newtonian mechanics, and provide examples of both non-Newtonian systems in action.

Contents

1. The Calculus of Variations

Lagrangian mechanics applies physics not only to particles, but to the trajectories of particles. We must therefore study how curves behave under small disturbances or variations.

Definition 1.1. Let V be a Banach space. A curve is a continuous map $\gamma : [t_0, t_1] \to V$. A *variation on the curve* γ is some function h of t that creates a new curve $\gamma + h$. A *functional* is a function from the space of curves to the real numbers.

Example 1.2. $\Phi(\gamma) = \int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2} dt$, where $\dot{x} = \frac{d\gamma}{dt}$, is a functional. It expresses the length of curve γ between t_0 and t_1.

Definition 1.3. A functional Φ is *differentiable* if $\Phi(\gamma + h) - \Phi(\gamma) = F(h) + R$ where F is linear in h and R is $O(h^2)$. $F(h)$ is called the *differential* of Φ.

Definition 1.4. An *extremal* of a differentiable functional $\Phi(\gamma)$ is a curve γ such that $F(h,\gamma)$=0 for all h. It is conceptually akin to a critical point of a traditional function.

Remark 1.5. Lagrangian mechanics is primarily concerned with the variations and extremals of one particular form of functional, $\Phi(\gamma)=\int_{t_0}^{t_1} L(x(t),\dot{x}(t),t)dt$, where L is a function differentiable in x, $\dot{x}$, and t. We would therefore be well advised to pay close attention to this functional.

Theorem 1.6. *Functionals of the form $\Phi(\gamma)=\int_{t_0}^{t_1} L(x(t),\dot{x}(t),t)dt$, where L is differentiable in x, $\dot{x}$, and t, are themselves differentiable.*

Proof.

$$\Phi(\gamma+h)-\Phi(\gamma)=\int_{t_0}^{t_1} L(x+h,\dot{x}+\dot{h},t)dt-\int_{t_0}^{t_1} L(x,\dot{x},t)dt \tag{1.7}$$

$$=\int_{t_0}^{t_1} \Big[L(x+h,\dot{x}+\dot{h},t)-L(x,\dot{x},t)\Big]\,dt \tag{1.8}$$

$$=\int_{t_0}^{t_1} \left[\frac{\partial L}{\partial x}h+\frac{\partial L}{\partial \dot{x}}\dot{h}\right]dt+O(h^2) \tag{1.9}$$

$$=\int_{t_0}^{t_1} \left[\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right]hdt+\left(h\frac{\partial L}{\partial \dot{x}}\right)\Bigg|_{t_0}^{t_1}+O(h^2) \tag{1.10}$$

$$=F(h)+O(h^2)=F+R. \tag{1.11}$$

□

So the differential of Φ is $\int_{t_0}^{t_1}[\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}]hdt+(h\frac{\partial L}{\partial \dot{x}})|_{t_0}^{t_1}$. We now wish to determine the extremals of this functional, but before we do, we must establish an intermediate lemma.

Lemma 1.12. *If a continuous function $f(t)$ satisfies $\int_{t_0}^{t_1} f(t)h(t)dt=0$ for any continuous function h(t) provided $h(t_0)=h(t_1)=0$, then $f(t)\equiv 0$.*

Proof. Assume that $\int_{t_0}^{t_1} f(t)h(t)dt=0$ and $0<c<f(t^*)$ for some $t^*\in(t_0,t_1)$. f is continuous, so there exists an interval Δ around t^* such that f(t)≥c for all $t\in\Delta$. Pick some d such that $[t^*-d,t^*+d]\subset\Delta$. Define h to be 0 outside of Δ, greater than 0 within Δ, and 1 within $[t^*-d,t^*+d]$. $\int_{t_0}^{t_1} f(t)h(t)dt\geq cd>0$, which contradicts our hypothesis; there cannot exist any t^* such that $f(t^*)$ is greater than zero. A similar proof can be constructed for $f(t^*)<0$, so $f(t)$ must be exactly zero for all $t\in(t_0,t_1)$. □

Theorem 1.13. *The curve γ is an extremal of the functional $\Phi(\gamma)=\int_{t_0}^{t_1} L(x,\dot{x},t)dt$ on the space of curves passing through $x(t_0)=x_0$ and $x(t_1)=x_1$ if and only if L satisfies $\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}-\frac{\partial L}{\partial x}=0$ along the curve γ.*

Proof. We have established that the differential of Φ is $\int_{t_0}^{t_1}[\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}]hdt+\frac{\partial L}{\partial \dot{x}}h|_{t_0}^{t_1}$. Because we are considering the extremals of Φ, this expression equals zero. Because we are only considering the curves that pass through x_0 and x_1, $h(t_0)=h(t_1)=0$ and the term outside of the integral equals zero, so $\int_{t_0}^{t_1}[\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}]hdt=0$. This is comparable to the integral we encountered in the lemma, where $\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}=f(t)$

and $h = h(t)$. By lemma 2.6, $f(t) \equiv 0$. On the other hand, if $f(t) \equiv 0$ then the integral as a whole equals zero. □

Remark 1.14. This equation lies at the heart of Lagrangian mechanics. We call

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0 \tag{1.15}$$

the *Lagrange-Euler equation.* However, before we begin to explore the physics of the Lagrange-Euler equation, we must specify what space our system exists in-so far, we have considered only real Euclidean space. We can expand our settings with the geometry of manifolds.

2. Manifold Geometry

Even though a space may not be perfectly Euclidean, it may be 'Euclidean enough' for our purposes. We will consider manifolds, structures which are locally, even if not globally, Euclidean.

Definition 2.1. A *manifold* is a Hausdorff space M covered by a countable number of charts so that every point in M is represented in at least one chart. A *chart* is a homeomorphism from an open subset of $\mathbb{R}^n$ to $U \subset M$. A union of charts which cover an entire manifold is an *atlas.*

Remark 2.2. If two charts (U, ϕ) and (V, ψ) map onto intersecting neighborhoods on M, then we can pass between the charts by inverting and composing the bijections. In other words, if there exist $U' \subset U$ and $V' \subset V$ such that $\text{image}_\phi(U')$=$\text{image}_\psi(V')$ then $\psi^{-1} \circ \phi : U' \to V'$. If the functions between charts are infinitely differentiable, then we say that the charts are *compatible.* If a manifold can be covered by a union of compatible charts, then we call it a *differentiable manifold.* We will be considering differentiable manifolds exclusively in this paper.

Example 2.3. Euclidean $\mathbb{R}^n$ is a manifold covered by the chart $\mathbb{R}^n$. The circle, sphere, and hyperspheres are each manifolds covered by two charts of the same dimension as the manifold itself 'wrapped' around the surface from opposite poles.

Definition 2.4. A *tangent vector* to M at x is the velocity vector of some curve embedded on $M : \dot{x} = lim_{t\to 0}\frac{\gamma(t)-\gamma(0)}{t}$, where $\gamma(0) = x$ and $\gamma(t) \in M$. The space of all vectors tangent to M at x is the *tanget space* at x, T_xM. The union of T_xM for all $x \in M$ is the *tangent bundle* TM on manifold M. If we impose a local coordinate system x_i in a neighborhood of x, then we can break $\dot{x}$ into components: $\dot{x}_i = \frac{d\gamma_i}{dt}|_{t=0}$.

Definition 2.5. A *Riemannian manifold* is a differentiable manifold M with a positive definite quadratic form $(\dot{x}, \dot{x})$:$TM \to \mathbb{R}$. The form $(\dot{x}, \dot{x})$ is called the *Riemannian metric.*

Remark 2.6. This metric allows us to measure tangent vectors; no longer abstract derivatives, they can now be used in calculations. This allows us to carry the Lagrange-Euler equation over to manifolds. Consider a Riemannian manifold M. We can define a differentiable function $L : TM \to \mathbb{R}$ that depends on $x, \dot{x}$, and t and construct the old functional $\Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t)dt$ for any γ embedded on M. We can carry this functional by bijection onto the Euclidean charts, among which we can move freely because this is a differentiable manifold. We established the Lagrange-Euler equation for Euclidean space, so we can now apply it to curves on manifolds as well. To do so, we will define a specific function $L(x, \dot{x}, t)$, the *Lagrangian.*

3. Lagrangian Mechanics

Definition 3.1. $x \in M$ is position along curve γ on Riemannian manifold M. $\dot{x}$ is the velocity of x along γ—i.e., a tangent vector. The body at x has mass m. Define kinetic energy $T : TM \to \mathbb{R}$ so $T = \frac{1}{2}m\dot{x}^2 = \frac{1}{2}m(\dot{x}, \dot{x})$ and system-dependent potential energy $U : M \to \mathbb{R}$. We call the function $L(x, \dot{x}, t) = T - U$ the *Lagrangian* or the *Lagrange function.* $\Phi = \int_{t_0}^{t_1}(T - U)dt$ is called the *action* of a system.

Theorem 3.2. ***Hamilton's Principle of Least Action.*** *The Lagrange-Euler equation over the Lagrangian function is equivalent to Newton's second law. That is, any mechanical system is an extremal of the functional* $\Phi = \int_{t_0}^{t_1} Ldt$ *on that system's configuration space, where L is the Lagrangian.*

Proof. At the extremals of the system's action, the Lagrangian satisfies the Lagrange-Euler equation, or $\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x}$. U is a function of position alone and T is a function of velocity alone, so $\frac{\partial L}{\partial x} = -\frac{dU}{dx}$ and $\frac{\partial L}{\partial \dot{x}} = \frac{dT}{d\dot{x}}$. Therefore $\frac{d}{dt}\frac{d\frac{1}{2}m\dot{x}^2}{d\dot{x}} = -\frac{dU}{dx}$, or $\frac{d}{dt}m\dot{x} = -\frac{dU}{dx}$. By definition, $p = m\dot{x}$ and $F = -\frac{dU}{dx}$, so in the end, $\frac{d}{dt}p = F$, which is Newton's second law. □

Remark 3.3. Nowhere in the proof did we use cartesian coordinates—the physics which we used in our definitions is generalized to any basis. Therefore, we can use Hamilton's principle of least action over a generalized coordinate system q_i with generalized velocities $\dot{q}_i$ and generalized momenta $p_i = \frac{\partial L}{\partial \dot{q}_i}$. The evolution of $\vec{q}$ along γ is subject to the general form of the Lagrange-Euler equation,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial L}{\partial \mathbf{q}} = 0. \tag{3.4}$$

Example 3.5. A mass is in free Euclidean $\mathbb{R}^3$ space so that $T = \frac{m}{2}(q_1^2 + q_2^2 + q_3^2), U = 0$, and $L = T = \frac{m}{2}(q_1^2 + q_2^2 + q_3^2)$. Putting this into the Lagrange-Euler equation, we find that $\frac{d}{dt}p_1 + \frac{d}{dt}p_2 + \frac{d}{dt}p_3 = 0$. In the absence of energy fields (i.e., in the absence of any external forces), linear momentum is conserved, in agreement with Newton's first law.

4. Two Electric Pendula

Let us attempt to use Hamilton's principle of least action in a physical problem. Consider two simple planar pendula, both of length l and mass m, suspended a distance a apart on a horizontal line so that they swing in the same plane. We wish to study only small oscillations, so we can approximate a swing by its linear displacement in the horizontal direction; the first pendulum will have displacement q_1 and the second q_2. In the exact problem, q_i represents an angle so the configuration space is $S^1 \times S^1$, the torus; in the approximation, it is $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, the plane. Both spaces are manifolds, so we are entitled to use Lagrangian mechanics in either case.

Put a charge $+e$ on both pendula and insulate them. We wish to find the motion of the pendula that arises from slight disturbances from their equilibrium position.

Let us begin by finding the Lagrangian in terms of our coordinates q_i. We know that the kinetic energy is given by

$$T = \frac{1}{2}m(\dot{q}_1^2 + \dot{q}_2^2). \tag{4.1}$$

Potential energy is the sum of two terms; one gravitational, the other electrostatic. Using various small-angle trigonometric approximations, we hypothesize the gravitational term to be

$$U_G = \frac{1}{2}\frac{mg}{l}(q_1^2 + q_2^2) \tag{4.2}$$

and the electrostatic term to be

$$U_E = \frac{e^2}{a + q_2 - q_1}. \tag{4.3}$$

Therefore, the entire Lagrangian function is

$$L = T - U = \frac{1}{2}m(\dot{q}_1^2 + \dot{q}_2^2) - \frac{1}{2}\frac{mg}{l}(q_1^2 + q_2^2) - \frac{e^2}{a + q_2 - q_1}. \tag{4.4}$$

Using q_1 and q_2 as the generalized coordinates, we can apply the Lagrange-Euler equation to the system. We get the two equations of motion

$$\frac{d}{dt}(m\dot{q}_1) + \frac{mg}{l}q_1 + \frac{e^2}{(a + q_2 - q_1)^2} = 0 \tag{4.5}$$

$$\frac{d}{dt}(m\dot{q}_2) + \frac{mg}{l}q_2 - \frac{e^2}{(a + q_2 - q_1)^2} = 0 \tag{4.6}$$

If we add these two equations, we get

$$\frac{d}{dt}m(\dot{q}_1 + \dot{q}_2) + \frac{mg}{l}(q_1 + q_2) = 0 \Rightarrow \frac{d^2}{dt^2}(q_1 + q_2) = -\frac{g}{l}(q_1 + q_2). \tag{4.7}$$

If we subtract them, we find

$$\frac{d^2}{dt^2}(q_2 - q_1) + \frac{g}{l}(q_2 - q_1) - \frac{2e^2}{m(a + q_2 - q_1)^2} = 0. \tag{4.8}$$

Because we assume that both q_1 and q_2 are very small and far less than a, we can expand the electrostatic term of (5.8), yielding

$$\frac{d^2}{dt^2}(q_2 - q_1) + (q_2 - q_1)\left(\frac{g}{l} + \frac{4e^2}{ma^3}\right) - \frac{2e^2}{ma^2} = 0. \tag{4.9}$$

Both (4.7) and (4.9) are harmonic oscillations in $(q_1 + q_2)$ and $(q_2 - q_1)$, respectively. They therefore have relatively simple solutions.

$$(q_1 + q_2)(t) = A_1 \cos\left(\sqrt{\frac{g}{l}}t + \phi_1\right) \tag{4.10}$$

$$(q_2 - q_1)(t) - \frac{2e^2}{ma^2\left(\frac{g}{l} + \frac{4e^2}{ma^3}\right)} = A_2 \cos\left(\sqrt{\frac{g}{l} + \frac{4e^2}{ma^3}}t + \phi_2\right) \tag{4.11}$$

A_i and ϕ_i are amplitudes and phases, respectively. They depend on the initial conditions of the system and not on the system itself, so we will leave them unspecified.

Now we can add and subtract these two expressions to find the individual motions of q_1 and q_2.

$$q_1 = \frac{-e^2}{ma^2\left(\frac{g}{l}+\frac{4e^2}{ma^3}\right)} + \frac{A_1}{2}\cos\left(\sqrt{\frac{g}{l}}t+\phi_1\right) - \frac{A_2}{2}\cos\left(\sqrt{\frac{g}{l}+\frac{4e^2}{ma^3}}t+\phi_2\right) \tag{4.12}$$

$$q_2 = \frac{e^2}{ma^2(\frac{g}{l}+\frac{4e^2}{ma^3})} + \frac{A_1}{2}\cos\left(\sqrt{\frac{g}{l}}t+\phi_1\right) + \frac{A_2}{2}\cos\left(\sqrt{\frac{g}{l}+\frac{4e^2}{ma^3}}t+\phi_2\right) \tag{4.13}$$

We can translate the origins for both q_1 and q_2 to eliminate the constant $\frac{e^2}{ma^2(\frac{g}{l}+\frac{4e^2}{ma^3})}$. Now both motions are the composition of two simple sinusoidal motions; depending on the initial conditions, they might simplify further.

5. Differential Forms and Symplectic Geometry

The Hamiltonian formalism provides yet another way to look at classical mechanical systems. But before we develop the physics, we much establish some mathematical groundwork.

Definition 5.1. A *1-form* is a linear vector function $\omega^1 : \mathbb{R}^n \to \mathbb{R}$. $\omega^1(\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2) = \lambda_1\omega(\vec{v_1}) + \lambda_2\omega(\vec{v_2})$. A *2-form* is a bilinear skew-symmetric function of 2 vectors $\omega^2 : \mathbb{R}^n \text{ x } \mathbb{R}^n \to \mathbb{R}$. $\omega^2(\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2, \vec{v}_3) = \lambda_1\omega^2(\vec{v}_1, \vec{v}_3) + \lambda_2\omega^2(\vec{v}_2, \vec{v}_3)$ and $\omega^2(\vec{v}_2, \vec{v}_1) = -\omega^2(\vec{v}_1, \vec{v}_2)$. Similarly, a *k-form* is a multilinear skew-symmetric function from k vectors to the real numbers. If we define form addition and scalar multiplication so that $(\omega_1^k + \omega_2^k)(\ldots\vec{v}_i\ldots) = \omega_1^k(\ldots\vec{v}_i\ldots) + \omega_2^k(\ldots\vec{v}_i\ldots)$ and $(\lambda\omega^k)(\ldots\vec{v}_i\ldots) = \lambda(\omega^k(\ldots\vec{v}_i\ldots))$ then the set of k-forms forms a vector space itself.

Definition 5.2. The *exterior product* $\omega_1^1 \wedge \omega_2^1$ of two 1-forms ω_1 and ω_2 is the 2-form $\omega^2(\vec{v}_1, \vec{v}_2)$ that gives the oriented area of the parallelogram formed by vectors $(\omega_1^1(\vec{v}_1), \omega_2^1(\vec{v}_1)), (\omega_1^1(\vec{v}_2), \omega_2^1(\vec{v}_2))$. Similarly, the exterior product of k 1-forms $\omega_1^1 \wedge \omega_2^2 \wedge \ldots \wedge \omega_k^1$ is the k-form that gives the oriented volume of the parallelotope formed by the vectors $(\omega_1^1(\vec{v}_1), \ldots, \omega_k^1(\vec{v}_1)), (\omega_1^1(\vec{v}_2), \ldots, \omega_k^1(\vec{v}_2)), \ldots, (\omega_1^1(\vec{v}_k), \ldots, \omega_k^1(\vec{v}_k))$.

In matrix notation, $(\omega_1^1 \wedge \ldots \wedge \omega_k^1)(\vec{v}_1 \ldots \vec{v}_k) = \det \begin{bmatrix} \omega_1^1(\vec{v}_1) & \ldots & \omega_k^1(\vec{v}_1) \\ \vdots & & \vdots \\ \omega_1^1(\vec{v}_k) & \ldots & \omega_k^1(\vec{v}_k) \end{bmatrix}$.

Definition 5.3. A *differentiable 1-form* on manifold M is a smooth map $\omega^1 : TM \to \mathbb{R}$ which is linear on each fibre to the tangent bundle. A *differentiable k-form at x* is a k-form at x $\omega^k : T_xM \to \mathbb{R}$. A differentiable k-form may be differentiable for all x on M.

Definition 5.4. We can construct *exterior differentiation*, an operation relating functions and forms. The exterior derivative df of function f is uniquely defined so that df acts on vector field V by $(df)(V) = V(f)$. Additionally, $ddf = dd\omega = 0$ for any form ω or function f and $d(\omega^k \wedge \omega^l) = d\omega^k \wedge \omega^l + (-1)^k\omega^k \wedge d\omega^l$. Exterior differentiation sends $k-$forms to $k+1-$forms; standard functions are regarded as $0-$forms, so the exterior derivative of a function is a $1-$form. If $d\omega = 0$, then ω is said to be a *closed form.*

Definition 5.5. A *symplectic structure* on $2n$-dimensional manifold M^{2n} is a closed, skew-symmetric, nondegenerate differential 2-form $\omega^2 : T_xM^{2n} \to \mathbb{R}$. This makes M^{2n} a *symplectic manifold.* To say that ω^2 is nondegenerate means that for every $\vec{v} \in T_xM^{2n}$, there exists some $\vec{w} \in T_xM^{2n}$ such that $\omega^2(\vec{v}, \vec{w}) \neq 0$. From the linear algebra of the symplectic form, we find that it is impossible for an odd-dimensional manifold to support a symplectic structure.

Definition 5.6. Suppose M^n is an $n-$dimensional differentiable manifold. A 1-form acting on T_xM^n is called a *cotangent vector* to M^n at x. The set of all 1-forms acting on T_xM^n is called the *cotangent space* of M^n at x, or $T_x^*M^n$. The union of all cotangent spaces for all x on M^n forms the manifold's *cotangent bundle*, T^*M^n. If we assign T^*M^n a 2-form $\omega^2 = d\vec{p} \wedge d\vec{q} = dp_1 \wedge dq_1 + dp_2 \wedge dq_2 + \ldots + dp_n \wedge dq_n$, then T^*M^n has a symplectic structure. It can be shown that this symplectic form does not depend upon the coordinate systems of the charts.

Definition 5.7. Let M^{2n} be a symplectic manifold. Because the symplectic form ω^2 is nondegenerate, for every $\vec{v} \in TM^{2n}$ there exists some $\vec{w} \in TM^{2n}$ such that $\omega^2(\vec{w}, \vec{v}) \neq 0$. Therefore we can create a 1-form $\omega^1_{\vec{v}}(\vec{w}) = \omega^2(\vec{w}, \vec{v})$ which serves as an isomorphism $I : T^*M^{2n} \to TM^{2n}$. If H is a differentiable function on M^{2n}, then its exterior derivative is the $1-$form $dH \in T^*M^{2n}$. By applying I to dH, we find a vector field $IdH \subset TM^{2n}$ that provides a tangent vector everywhere on M^{2n}. If we call H the *Hamiltonian function*, this field is the *Hamiltonian vector field* associated with H.

Remark 5.8. Let $\vec{q}$ of dimension n specify some point on M^n, and let $\vec{p}_q$ of dimension n specify some tangent vector at $\vec{q}$. Then the $2n-$dimensional ordered pair $(\vec{q}, \vec{p})$ serves as a coordinate in the cotangent bundle. Furthermore, if M^n is the configuration space of a dynamic system with $\vec{q}$ the generalized position and $\vec{p}$ a generalized momentum, then the cotangent bundle is the system's *phase space*, which exhibits symplectic geometry. Because motion depends exclusively on position and velocity or momentum, every possible condition of the particle is a point in phase space, and every possible trajectory is an embedded curve. With the proper Hamiltonian function, we could produce a Hamiltonian vector field where every mechanical trajectory is an integral curve dependent only on the initial conditions.

6. Hamiltonian Mechanics

We have seen that a Hamiltonian function produces a Hamiltonian vector field. We now wish to find a Hamiltonian function that produces a field worth studying; one corresponding to a vector field that describes how a body's position and momentum (i.e., its coordinates in phase space) evolve in time. Physical motion is then an integral curve that satisfies the vector field, given the system's boundary conditions.

Definition 6.1. $\frac{\partial L}{\partial \dot{q}_i} = p_i$, where L is the Lagrangian and p_i is the i^{th} component of the generalized momentum. Then the *Hamiltonian function* is $H(q, p, t) = \langle \frac{d\vec{q}}{dt}, \vec{p} \rangle - L(q, \dot{q}, t)$. Note that for classical mechanical systems the Hamiltonian is just the total mechanical energy: $H = \langle \frac{d\vec{q}}{dt}, \vec{p} \rangle - L = \langle \vec{v}, m\vec{v} \rangle - (\frac{1}{2}mv^2 - U(q)) = \frac{1}{2}mv^2 + U(q) = T + U$.

Theorem 6.2. ***Hamilton's Canonical Equations.*** *For a mechanical system, $\vec{q} = (\ldots, q_i, \ldots)$ is the generalized position, $\vec{p} = (\ldots, p_i, \ldots) = (\ldots, \frac{\partial L}{\partial \dot{q}_i}, \ldots)$ is the* *generalized momentum, and H is the Hamiltonian. Then*

$$\frac{\partial H}{\partial \vec{q}} = -\frac{d\vec{p}}{dt} \tag{6.3}$$

$$\frac{\partial H}{\partial \vec{p}} = \frac{d\vec{q}}{dt}. \tag{6.4}$$

Properly,

$$\frac{\partial H}{\partial q_i} = -\frac{dp_i}{dt} \tag{6.5}$$

$$\frac{\partial H}{\partial p_i} = \frac{dq_i}{dt}. \tag{6.6}$$

Proof. (*Note: for typographical reasons, I will denote vector status by boldface font within this proof.*) We know, by definition, that

$$H(q, p, t) = \langle \frac{d\mathbf{q}}{dt}, \mathbf{p} \rangle - L(q, \dot{q}, t). \tag{6.7}$$

Taking the differential of both sides, we find that

$$\left(\frac{\partial H}{\partial \mathbf{q}}\right) d\mathbf{q} + \left(\frac{\partial H}{\partial \mathbf{p}}\right) d\mathbf{p} + \left(\frac{\partial H}{\partial t}\right) dt = dH = \left(\frac{d\mathbf{q}}{dt}\right) d\mathbf{p} - \left(\frac{\partial L}{\partial \mathbf{q}}\right) d\mathbf{q} - \left(\frac{\partial L}{\partial t}\right) dt. \tag{6.8}$$

Matching the differentials, we immediately find that

$$\frac{\partial H}{\partial \mathbf{p}} = \frac{d\mathbf{q}}{dt} \tag{6.9}$$

$$\frac{\partial H}{\partial t} = -\frac{dL}{dt}.$$

We also see that

$$\frac{\partial H}{\partial \mathbf{q}} = -\frac{\partial L}{\partial \mathbf{q}}. \tag{6.11}$$

By the chain rule, we can expand this to

$$\frac{\partial H}{\partial \mathbf{q}} = -\frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}}. \tag{6.12}$$

By drawing out the time derivative and identifying $\frac{\partial L}{\partial \dot{\mathbf{q}}} = \mathbf{p}$, we can reduce this to the last of Hamilton's canonical equations,

$$\frac{\partial H}{\partial \mathbf{q}} = -\frac{d}{dt} \frac{\partial \mathbf{q}}{\partial \mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} = -\frac{d\mathbf{p}}{dt}. \tag{6.13}$$

□

Definition 6.14. A solution satisfying Hamilton's canonical equations is called a system's *Hamiltonian phase flow* because it describes movement in phase space.

Definition 6.15. In the canonical coordinates $(\vec{q}, \vec{p})$, we define the *Poisson bracket* of two functions f and H on a symplectic manifold as $(f, H) = \sum \left[\frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i}\right]$. We see that it is skew-symmetric (so $(H, f) = -(f, H)$) and, by the properties of the derivative, bilinear. In terms of the symplectic form, $(f, H) = \omega^{-1}(df, dH)$: although we have introduced and will use the bracket in coordinates, it is important to note that it is actually independent of the coordinate system.

Remark 6.16. The Poisson bracket has a number of interesting properties. For a fixed $f, (f, \cdot)$ satisfies the Leibnitz Rule. Like the Lie bracket, it also satisfies the Jacobi identity; if A, B, and C are functions on a manifold, then $((A, B), C) + ((B, C), A) + ((C, A), B) = 0$. Furthermore, if f is a vector function and $[X, Y]$ is the Lie bracket acting on vector fields X and Y, then $(f_X, f_Y) = -f_{[X,Y]}$, the full implications of which go beyond the scope of this paper. However, the Poisson bracket does have a relevant application at our level.

Theorem 6.17. *$f(\vec{q}, \vec{p})$ is a function in a system's phase space and H is the Hamiltonian function. f is conserved through motion (i.e., it is constant along the phase flow) if and only if $(f, H) = 0$.*

Proof.

$$\frac{d}{dt}f = \frac{\partial f}{\partial \vec{q}}\frac{d\vec{q}}{dt} + \frac{\partial f}{\partial \vec{p}}\frac{d\vec{p}}{dt} \tag{6.18}$$

$$\frac{d}{dt}f = \frac{\partial f}{\partial \vec{q}}\frac{\partial H}{\partial \vec{p}} - \frac{\partial f}{\partial \vec{p}}\frac{\partial H}{\partial \vec{q}} \tag{6.19}$$

$$\frac{d}{dt}f = (f, H) \tag{6.20}$$

t is the parameter of phase flow, so if f is conserved, then $\frac{df}{dt} = 0$; $(f, H) = 0$ by consequence. □

Example 6.21. Consider a body of mass m in a uniform gravitational field $\vec{g}$. $\vec{y}$ is a coordinate with $\hat{y}$ antiparallel to $\vec{g}$. $T = \frac{1}{2}m\dot{y}^2$ and $U = mgy$ so $H = T + U = \frac{1}{2}m\dot{y}^2 + mgy = \frac{p_y^2}{2m} + mgy$. By Hamilton's equations,

$$F_y = \dot{p_y} = -\frac{\partial H}{\partial y} = -mg \tag{6.22}$$

$$v_y = \dot{y} = \frac{\partial H}{\partial p_y} = \frac{p_y}{m} = \frac{mv_y}{m} = v_y \tag{6.23}$$

as we would expect from Newtonian mechanics. $y(t)$ can be found by simple integration.

Remark 6.24. Although our examples are all systems that can be easily described in coordinates, Hamilton's canonical equations work independently of any coordinate system because they were developed on manifolds. They are fundamentally geometric concepts, as are phase space and phase flow; we artificially impose analytic coordinates on the system in order to quantify and describe the phenomena.

7. The Double Planar Pendulum

Consider a standard planar pendulum of length l and mass m in a uniform gravitational field $\vec{g}$. We fix a second identical pendulum onto the bob of the first so that both swing in the same plane. Given initial conditions, we wish to find the pendula's coordinates as a function of time.

We will take as our generalized coordinates (θ_1, θ_2), the angle each pendulum's rod makes with the vertical. One pendulum has configuration space S^1 (the circle, corresponding to a full rotation about the axis), so the double pendulum has configuration space $S^1 \times S^1$, the torus. The torus is a manifold, so we can construct a symplectic phase space from its cotangent bundle and use Hamilton's equations to find a phase flow. First, we must find the energies and Lagrangian and Hamiltonian in terms of the generalized coordinates. To do so, let us assign a temporary cartesian frame with origin at the first pendulum's axis.

$$x_1 = l \sin\theta_1 \tag{7.1}$$

$$y_1 = -l \cos\theta_1 \tag{7.2}$$

$$x_2 = l(\sin\theta_1 + \sin\theta_2) \tag{7.3}$$

$$y_2 = -l(\cos\theta_1 + \cos\theta_2) \tag{7.4}$$

$$T = \frac{1}{2}mv_1^2 + \frac{1}{2}mv_2^2 = \frac{1}{2}m\left(\dot{x}_1^2 + \dot{y}_1^2 + \dot{x}_2^2 + \dot{y}_2^2\right) \tag{7.5}$$

$$U = mgy_1 + mgy_2 \tag{7.6}$$

$$L = T - U = \frac{1}{2}m\left(\dot{x}_1^2 + \dot{y}_1^2 + \dot{x}_2^2 + \dot{y}_2^2\right) - mgy_1 - mgy_2 \tag{7.7}$$

$$H = T + U = \frac{1}{2}m\left(\dot{x}_1^2 + \dot{y}_1^2 + \dot{x}_2^2 + \dot{y}_2^2\right) + mgy_1 + mgy_2 \tag{7.8}$$

Now we must revert back into our general angular coordinates.

$$T = \frac{1}{2}ml^2\left(2\dot{\theta}_1^2 + \dot{\theta}_2^2 + \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2\right) \tag{7.9}$$

$$U = -mgl\left(2\cos(\theta_1) + \cos(\theta_2)\right) \tag{7.10}$$

$$L = \frac{1}{2}ml^2\left(2\dot{\theta}_1^2 + \dot{\theta}_2^2 + \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2\right) + mgl\left(2\cos(\theta_1) + \cos(\theta_2)\right) \tag{7.11}$$

$$H = \frac{1}{2}ml^2\left(2\dot{\theta}_1^2 + \dot{\theta}_2^2 + \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2\right) - mgl\left(2\cos(\theta_1) + \cos(\theta_2)\right) \tag{7.12}$$

We can find our generalized angular momenta by differentiating the Lagrangian:

$$p_1 = \frac{\partial L}{\partial \dot{\theta}_1} = 2ml^2\dot{\theta}_1 + \frac{1}{2}ml^2\cos(\theta_1 - \theta_2)\dot{\theta}_2 \tag{7.13}$$

$$p_2 = \frac{\partial L}{\partial \dot{\theta}_2} = ml^2\dot{\theta}_2 + \frac{1}{2}ml^2\cos(\theta_1 - \theta_2)\dot{\theta}_1 \tag{7.14}$$

Solving these for $\dot{\theta}_1$ and $\dot{\theta}_2$ in terms of p_1 and p_2 and substituting into the Hamiltonian, we find

$$H = \left(\frac{p_1^2 + 2p_2^2 - 2p_1p_2cos(\theta_1 - \theta_2)}{2ml^2(1 + \sin^2(\theta_1 - \theta_2))}\right) - mgl\left(2\cos\theta_1 - \cos\theta_2\right). \tag{7.15}$$

The Hamiltonian is expressed entirely in constants, coordinates, and momenta, so we can use Hamilton's canonical equations.

$$\dot{\theta}_1 = \frac{\partial H}{\partial p_1} = \frac{p_1 - p_2\cos(\theta_1 - \theta_2)}{2ml^2\left(1 + \sin^2(\theta_1 - \theta_2)\right)} \tag{7.16}$$

$$\dot{\theta}_2 = \frac{\partial H}{\partial p_2} = \frac{2p_2 - p_1\cos(\theta_1 - \theta_2)}{ml^2\left(1 + \sin^2(\theta_1 - \theta_2)\right)} \tag{7.17}$$

$$\dot{p}_1 = -\frac{\partial H}{\partial \theta_1} = -2mgl\sin\theta_1 - A + B \tag{7.18}$$

$$\dot{p}_2 = -\frac{\partial H}{\partial \theta_2} = -mgl\sin\theta_2 + A - B \tag{7.19}$$

where

$$A = \frac{p_1 p_2 \sin(\theta_1 - \theta_2)}{ml^2(1+\sin^2(\theta_1-\theta_2))} \tag{7.20}$$

$$B = \frac{p_1^2 + 2p_2^2 - p_1 p_2 \cos\theta_1 - \theta_2}{2ml^2(1+\sin^2(\theta_1-\theta_2))^2}\sin[2(\theta_1-\theta_2)]. \tag{7.21}$$

These differential equations form the components of the Hamiltonian field. The initial conditions $(\theta_1(0), p_1(0), \theta_2(0), p_2(0))$ serve as a point in phase space; starting from here, we can integrate the differential equations numerically to find the phase flow, which fully describes the motion of the double pendulum. The complexity of all four equations gives rise to the double pendulum's notorious *chaotic motion*; a small change in the initial conditions will generate very large changes in the phase flow.

Acknowledgments. I owe a great debt of gratitude to my mentor, Aaron Marcus, for the countless hours he spent with me offering much-needed guidance, introducing and reviewing material, and revising my work. Without his constant help and advice, this paper would certainly not have been possible.

References

[1] V. I. Arnold. *Mathematical Methods of Classical Mechanics.* Trans. K. Vogtmann and A. Weinstein. Springer-Verlag. 1978.

[2] R. B. Lindsay. *Physical Mechanics.* D. Van Nostrand Company. 1933.