Mechanics of Nonholonomic Hybrid Lagrangian Systems

Evan Ortiz
Professor Anthony Bloch

July 2022

1 Introduction

The goal of any mechanics problem is to determine how things move. This text will consider the Lagrangian and Lagrange d'Alembert approach using variations and constraints rather than the traditional Newtonian approach of forces. This allows for insight into more complicated systems including those with nonholonomic constraints such as the rolling without slipping condition. This approach will be applied to 2 scenarios, that of the collision a planar ellipse makes with a wall, and the frictional collision of a circular disk against a wall, describing the effect of spin on the rebound trajectory.

2 Preliminaries

2.1 Lagrangian Mechanics

Lagrangian Mechanics uses variational techniques on the path of a moving body to determine its equation of motion. This is done by optimizing the path integral of the Lagrangian function; typically this is given by L = T - U, where T is the kinetic energy of the body, and U its potential energy, although a more rigorous definition will be provided later.

Definition 2.1 The *Configuration Space* of a given mechanical system is the topological manifold, denoted here as Q, describing the position of the system.

For example, a sphere moving in space might have the configuration space

$$Q = \mathbb{R}^3 \times SO(3)$$

where $\mathbb{R}^3$ describes its position in space, and $SO(3)$ describes its rotational orientation.

Definition 2.2 *Hamilton's Principle* states that a body will move such that

$$\delta \int_a^b L dt = 0$$

where the variations are taken over smooth curves in Q.

Hence we see here the optimization of the Lagrangian.

From Hamilton's Principle, using principles from the calculus of variations, one can derive a solution to this optimization called an *Euler-Lagrange equation* which becomes the equation of motion of the system. Given a generalized coordinate $q = (q_1, q_2, \ldots, q_n)$ this gives n equations of the form

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

In fact, this formulation is equivalent to Newton's second law $\overrightarrow{F} = m\overrightarrow{a}$, an alternative method for determining the equation of motion, which we will show with an example.

Consider a point mass moving in Earth's atmosphere, neglecting air resistance. The configuration space $Q = \mathbb{R}^3$ with a generalized coordinate $q = (x, y, z)$, where (x, y) represents the planar position along the ground, and z represents the height of the ball. Thus the Lagrangian formalism becomes

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz$$

And the Euler-Lagrangian equations for each *x,y,z*

$$\ddot{x} = 0,\ \ddot{y} = 0,\ \ddot{z} + g = 0$$

Compare this to the Newtonian Formulation

$$\overrightarrow{F}_{net} = m \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix}$$

Both formulations thus result in an equivalent equation of motion.

2.2 Relevant Topological Concepts

Definition 2.3 Given a differentiable manifold M, the *tangent space* at a point $q \in M$, denoted T_qM is a vector space consisting of all tangent vectors to q.

For example, if $M = \{(x, y) \in \mathbb{R}^2 : y = x^2\}$, then $T_{(0,0)}M$ is the tangent line at (0,0), a one dimensional vector space. This happens to be the x-axis.

Definition 2.4 The *Tangent Bundle* of a Manifold *M*, denoted *TM* is the set of ordered pairs $(q, \dot{q})$ such that $q \in M$ and $\dot{q} \in T_q M$.

The notion of a tangent bundle allows for a definition of a generalized coordinate for a system as well as a generalized velocity. Given a system with a configuration space Q, a generalized coordinate describes the instantaneous state of the system. Hence for a moment in time the generalized coordinate describing that state is some $q \in Q$. The generalized velocity of the system, defined as $\dot{q} = \frac{d}{dt}q$, is some vector in $T_q Q$.

For example, a flat disk in the plane might have a configuration space as $\mathbb{R}^2 \times S^1$, and at a given moment in time might have a generalized coordinate $q = (x, y, \theta)$. The generalized velocity would then be some $\dot{q} = (\dot{x}, \dot{y}, \omega)$ where $(\dot{x}, \dot{y})$ describes the linear velocity and ω the angular velocity.

Definition 2.5 A *Riemann Manifold* is a differentiable manifold, M, such that the tangent space at each point $q \in M$, is also endowed with an inner product g_q, where

$$g_q : T_q M \times T_q M \to \mathbb{R}$$

The significance of this in mechanics is that it allows for a definition of v^2, as in $T = \frac{1}{2}mv^2$.

Earlier we mentioned that frequently the Lagrangian is take to be $L = T - U$. T is often a function of generalized velocity and U a function of the generalized coordinate. Hence, L is actually a function of the tangent bundle of the configuration space Q.

Definition 2.6 The *Lagrangian,* L is a smooth function $L : TQ \to \mathbb{R}$, having units of energy. In particular, the Lagrangian is called *natural* if $L = \frac{1}{2}g_q(\dot{q}, \dot{q}) - U(q)$ and (Q, g_q) forms a Riemann Manifold.

For example, a flat disk in a plane has a lagrangian of

$$L = \tfrac{1}{2}(m\dot{x}^2 + m\dot{y}^2 + I\dot{\theta}^2) - U(q)$$

where m is the mass of the disk, I is the moment of inertia through its canonical axis, and U is some potential function. This can be written as

$$L = \tfrac{1}{2}\dot{q}^T \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & I \end{bmatrix} \dot{q} - U(q)$$

Thus this lagrangian is natural.

For the remainder of this text we will only consider natural lagrangians.

2.3 Constraint Distributions

Definition 2.7 A constraint distribution, Δ, in a mechanical system is a subset of TQ where Q is the configuation space of the system, such that $\Delta_q \subseteq T_qQ$ is a subspace.

Essentially a constraint distribution restricts either the generalized position or velocity of the system. Many constraints are linear in velocity. Thus they are of the form

$$\sum_{k=1}^{n} a_k^j(q)\dot{q}_k$$

Where $j = 1, \ldots, m$, m is the number of constraints, and n is the number of generalized coordinates.

For example, a sphere rolling without slipping over a plane is a rigid body, meaning the distances of each point on the sphere are fixed relative to every other point on the sphere. I.e. the sphere maintains its shape regardless of where it is located and how it is moving. This is an example of a *holonomic* constraint because it is a constraint solely on the positions of each point on the sphere. As the sphere moves it also exhibits the *nonholonomic* constraint of rolling without slipping. This is given by the equation below, which is linear in velocity.

$$\dot{x} + R\omega_y = 0$$
$$\dot{y} - R\omega_x = 0$$

Where ω_x and ω_y represent the angular velocity of the ball around the x and y axes respectively.

This constraint is nonholonomic because it cannot be integrated to a constraint on position. Intuitively, the ball is known to be rolling but that provides no indication of where the ball is located in the plane.

To account for constraints, modifications to the Lagrange-Euler equations are necessary. This requires a generalization of Hamilton's Principle.

Definition 2.8 The *Lagrange-d'Alembert Principle* states that a body will move such that

$$\delta \int_a^b L dt = 0$$

and that the virtual displacements δq will satisfy the constraints.

Without constraints the principle reduces to Hamilton's case. Like Hamilton's Principle the Lagrange-d'Alembert Principle still takes variations across the *full range of curves* within the configuration space. The resulting curve is then projected onto the constraint distribution, ensuring any virtual displacement, δq, satisfies them. An important note is that these two operations are distinct and noncommuting. It would be incorrect to take variations only across the constraint subspace. Thus the original Lagrangian is preserved and the Euler-Lagrange equations of motion are generalized to the *Lagrange-d'Alembert equations of motion* or alternatively named *Dynamical Nonholonomic equations of motion*. An important assumption in the derivation of this is that the constraint force do no work on the system. That is they satisfy

$$F_1 \delta q_i + \cdots + F_n \delta q_n = 0$$

Using linear algebra analogous to the Lagrange multiplier theorem, it can be shown that each constraint force thus is of the form

$$F_i = \lambda_1 a_i^1 + \cdots + \lambda_m a_i^m$$

where $i = 1, \ldots, n$, and a_i^j is from the constraint equation, given that the vectors $(a_1^1, \ldots, a_n^1), \ldots, (a_1^m, \ldots, a_n^m)$ are linearly independent.

Thus the *Lagrange-d'Alembert Equations* become

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial q_i} = \sum_{j=1}^{m} \lambda_j a_i^j$$

To clarify this, i is the index of the generalized coordinate, j is the index of the constraint. And there are n generalized coordinates and m constraint equations (i.e. Lagrange Multipliers).

2.4 Hybrid Lagrangian Systems

The previous sections have taken variations over smooth curves within the configuration space Q, and therefore assumes only continuous motion of the system. *Hybrid* Lagrangian systems involve discontinuous motion such as a collision.

Definition 2.9 The impact surface, $S \subset Q$, is a smooth embedded submanifold of codimension 1.

This can be thought of as a boundary at which a body may collide against. The key to describing a discontinuous motion is to determine an impact map $(q, \dot{q}^-) \in TQ \mapsto (q, P(q, \dot{q}^-)) \in TQ$.

Determining an explicit formula for this map will require further understanding of the topology of Q.

Definition 2.10 Given a differentiable manifold, M, and a point $q \in M$ the *Cotangent Space* at q, $T_q^* M$, is the dual space of $T_q M$. Further, the *Cotangent Bundle* of M, T^*M, is the set of all ordered pairs (q, f) s.t. $q \in M$ and $f \in T_q^* M$.

If the system in question involves a natural Lagrangian, then there is a natural isomorphism between the tangent and cotangent bundles of Q, called the musical isomporphisms, $\flat$ and $\sharp$.

$$\flat : TQ \to T^*Q,\ (q,v) \mapsto (q, g(v,\cdot))$$
$$\sharp : T^*Q \to TQ,\ \sharp = \flat^-$$

Definition 2.11 Let $h : Q \to \mathbb{R}$ be smooth. The *gradient* of h is defined as $\nabla h = dh^\sharp$.

The *Weierstrass-Erdman* corner conditions allow for a variational approach to determining an impact map. The impact surface, S, can be locally described as the level set of some function $h : Q \to \mathbb{R}$, i.e. locally S can be described as $q \in Q : h(q) = 0$. Defining $p_i = \frac{\partial L}{\partial \dot{q}}$ as the generalized momentum, and $H = T + U$, the *Hamiltonian* (total energy), these conditions can be formulated as such

$$p^+ = p^- + \alpha \cdot dh$$
$$H^+ = H^-$$

Note this is not a rigorous definition of the Hamiltonian, but it will suffice for the purposes of this text. Intuitively the corner conditions given here can be interpreted as the change in momentum of the collision is normal to the impact surface at the point of contact, per the multiplier α, and the energy of the system is conserved. Solving the system given by the corner conditions and the assumption that the constraint forces do no work thus gives the impact map.

Theorem 2.12 Given a natural nonholonomic hybrid Lagrangian system, the fully elastic impact map, P, is given by

$$P(q,\dot{q}) = \dot{q} - 2\frac{dh(\dot{q})}{g(\nabla h, \nabla h)}\nabla h$$

Using the notation $dh(\dot{q}) = \frac{\partial h}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial h}{\partial q_n}\dot{q}_n$

3 Worked Examples

3.1 Bouncing Ellipse

The situation being modelled here is a uniform planar elliptic disk colliding with a one dimensional wall, which will be set as the y-axis. Before the collision the motion is easily described, assuming no potential forces and a frictionless plane. Let m be the mass of the ellipse, I be the moment of inertia with respect to the z-axis, (x, y) be the position of the center of mass of the ellipse, and θ be the rotation of the ellipse assumed to be homogeneous.

$$Q = \mathbb{R}^2 \times S^1$$
$$L = \tfrac{1}{2}(m\dot{x}^2 + m\dot{y}^2 + I\dot{\theta}^2)$$
$$= \tfrac{1}{2}\dot{q}^T M \dot{q}$$

with the matrix of the inner product, M, given as

$$M = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & I \end{bmatrix}$$

Thus the Euler-Lagrange equations are given as

$$\ddot{q} = 0$$
$$\implies q = \dot{q}_0 t + q_0$$

For each coordinate of $q = (x, y, \theta)$.

To determine an impact map, first we need an impact surface, S, which we will use as $S = \{(x, y) \in \mathbb{R}^2 : h(x, y) = 0\}$, with $h(x, y) = x$.

Next we must find the point along the ellipse boundary that will make contact with the wall. Thus we introduce a new variable, ϕ, to parameterize the boundary of the ellipse. For a given θ, setting $\theta = 0$ as the upright ellipse

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a\cos\phi \\ b\sin\phi \end{bmatrix} + \begin{bmatrix} x \\ y \end{bmatrix}$$

The point of the ellipse making contact will occur when $\frac{\partial X}{\partial \phi} = 0$. Denote this critical angle $\phi_c(\theta)$.

$$\phi_c(\theta) = \arctan\left(-\tfrac{b}{a}\tan\theta\right)$$

Thus the impact will occur when

$$h(X, Y) = x + a\cos(\phi_c(\theta))\cos\theta - b\sin(\phi_c(\theta))\sin\theta = 0$$

This is the equation that is used to determine the impact map.

$$\frac{\partial h}{\partial x} = 1$$
$$\frac{\partial h}{\partial y} = 0$$
$$\frac{\partial h}{\partial \theta} = \Phi(\theta)$$

with $\Phi(\theta)$ given by

$$\Phi(\theta) =$$
$$\frac{b^2\sin\theta}{a(\frac{b^2\tan^2\theta}{a^2}+1)^{\frac{1}{2}}} - \frac{a\sin\theta}{(\frac{b^2\tan\theta^2}{a^2}+1)^{\frac{1}{2}}} + \frac{b^2\tan\theta\sec\theta}{a(\frac{b^2\tan\theta^2}{a^2}+1)^{\frac{1}{2}}} - \frac{b^2\tan\theta\sec\theta}{a(\frac{b^2\tan\theta^2}{a^2}+1)^{\frac{3}{2}}} - \frac{b^4\tan^3\theta\sec\theta}{a^3(\frac{b^2\tan\theta^2}{a^2}+1)^{\frac{3}{2}}}$$

Thus

$$dh(\dot{q}) = \dot{x} + \Phi(\theta)\dot{\theta}$$
$$\nabla h = \begin{bmatrix} \frac{1}{m} & 0 & \frac{1}{I}\Phi(\theta) \end{bmatrix}$$
$$\implies g(\nabla h, \nabla h) = \frac{1}{m} + \frac{1}{I}\Phi(\theta)^2$$

Which yields an impact map of

$$\dot{x}^{+} = \dot{x}^{-} - 2\tfrac{C}{m}$$
$$\dot{y}^{+} = \dot{y}^{-}$$
$$\dot{\theta}^{+} = \dot{\theta}^{-} - 2\tfrac{C}{I}\Phi(\theta)$$

Where C is given by

$$C = \frac{\dot{x}+\Phi(\theta)\dot{\theta}}{\frac{1}{m}+\frac{1}{I}\Phi(\theta)^2}$$

3.2 Planar Circular Disk with Frictional Collision

The previous section relied on the assumption that the wall the ellipse collided with was frictionless, and thus there was no reactionary force tangential to the ellipse at the moment of contact. For this section we will focus on a frictional contact, and therefore will restrict it to the case of a circular spinning disk.

Prior to impact, the spinning disk has the same lagrangian and equations of motion as above.

$$L = \tfrac{1}{2}(m\dot{x}^2 + m\dot{y}^2 + I\dot{\theta}^2)$$
$$q = \dot{q}_0 t + q_0$$

for each coordinate in $q = (x, y, \theta)$

And the impact surface is given by the level curve of the same function $h(x, y) = x$.

In a perfectly inellastic situation, the impact will absorb the entirety of the $\dot{x}$ coordinate of the generalized velocity vector, and the $\dot{y}$ and $\dot{\theta}$ coordinates will be orthogonally projected onto the rolling without slipping condition, along the y-axis, given by

$$\dot{y} + R\dot{\theta} = 0$$

The subspace induced by this constraint is equivalent to the kernel of the 1-form

$$\omega = dy + Rd\theta$$

or in operator form

$$\begin{bmatrix} 0 & 1 & R \end{bmatrix}$$
$$ker(\omega) = span\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -R \\ 1 \end{bmatrix}\right)$$

These basis vectors can then be converted to an orthonormal basis under the inner product in the lagrangian. This yields

$$\hat{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{m}} \\ 0 \\ 0 \end{bmatrix}, \hat{e}_2 = \frac{1}{\sqrt{I+mR^2}} \begin{bmatrix} 0 \\ -R \\ 1 \end{bmatrix}$$

Under the Parrallel Axis Theorem, the quantity $I + mR^2$ is the moment of inertia around the point of contact with the wall. Denote this as I_1.

Thus we can orthogonally project the generalized velocity vector onto this subspace.

$$\begin{bmatrix} 0 \\ \dot{y}_0 \\ \dot{\theta}_0 \end{bmatrix} \mapsto \frac{I\dot{\theta} - m\dot{y}R}{I_1} \begin{bmatrix} 0 \\ -R \\ 1 \end{bmatrix} := \alpha \begin{bmatrix} 0 \\ -R \\ 1 \end{bmatrix}$$

In the perfectly inelastic case, the disk will continue to roll along the wall in this manner.

In the perfectly elastic situation, we can use conservation of energy to determine the resulting rebound of the disk. At exactly the moment of impact, the $\dot{x}$ coordinate is still 0, in the elastic case. Thus the same projection onto the rolling without slipping constraint distribution will occur identically to the inelastic case, seen here again.

$$\begin{bmatrix} 0 \\ \dot{y}_0 \\ \dot{\theta}_0 \end{bmatrix} \mapsto \frac{I\dot{\theta} - m\dot{y}R}{I_1} \begin{bmatrix} 0 \\ -R \\ 1 \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ -R \\ 1 \end{bmatrix}$$

The conservation of energy is now applied in order to obtain the resulting $\dot{x}_1$ coordinate, denoting the post impact velocity coordinate.

$$E_0 = const.$$
$$E_0 = \tfrac{1}{2}(m\dot{x}_0^2 + m\dot{y}_0^2 + I\dot{\theta}_0^2) = \tfrac{1}{2}(m\dot{x}_1^2 + m\dot{y}_1^2 + I\dot{\theta}_1^2)$$
$$\dot{x}_1^2 = \dot{x}_0^2 + \dot{y}_0^2 + \tfrac{I}{m}\dot{\theta}_0^2 - \alpha^2 R^2 - \tfrac{I}{m}\alpha^2$$

Where α is the multiplier from above, given by

$$\alpha = \frac{I\dot{\theta} - mR\dot{y}}{I_1}$$

We can use a couple of explicit examples to demonstrate this finding. Firstly, the most basic case, is when $\dot{q}_0 = (\dot{x}_0, 0, 0)$. The formula returns

$$\dot{x}_1^2 = \dot{x}_0^2$$

We can then use intuition from the problem to show $\dot{x}_1 = -\dot{x}_0$ as would be expected. This serves as a good sanity check.

A more interesting example would be if the ball was moving normal to the wall still (i.e. $\dot{y}_0 = 0$) but had nonzero initial rotation. So now let $\dot{q}_0 = (\dot{x}_0, 0, \dot{\theta}_0)$. The impact map becomes

$$\dot{x}_1 = -[\dot{x}_0^2 + (\frac{I}{m} - \frac{I^2R^2}{I_1^2} - \frac{I^3}{mI_1^2})\dot{\theta}_0^2]^{\frac{1}{2}}$$
$$\dot{y}_1 = -R\alpha = -\frac{RI\dot{\theta}_0}{I_1}$$
$$\dot{\theta}_1 = \frac{I}{I_1}\dot{\theta}_0$$

A few important notes from this result: For the x velocity coordinate, the negative root is taken because intuitively the disk will reflect off the wall, and hence change direction. The y velocity coordinate initially is 0, but becomes nonzero afterwards. This is expected in a collision where the rotation of the disk interacts with the boundary wall. The θ velocity coordinate is reduced in magnitude as this is what propels the disk in the y direction. ($I < I_1$) Thus some rotational kinetic energy is transferred to translational, which is again what is expected.

4 Further Study

A natural generalization of these kinds of problems is to change the shape of either the impact surface or of the moving body. In particular, the problem considering a frictional impact may have extensive applications is sports modeling such as modelling a tennis ball's path, basketball rebounds, or other incidents with a colliding ball. Additionally further insight may also be gained from changing the properties of the configuration space to something entirely non-euclidean, such as a situation that my be found in strong gravitational fields.

5 References

[1] A. Bloch, J. Baillieul, P. Crouch, and J. E. Marsden, *Nonholonomic mechanics and control. Interdisciplinary Applied Mathematics.* Springer New York, 2008.

[2] W. Clark and A. Bloch, *The Bouncing Penny and Nonholonomic Impacts.* arXiv, 2019