Noether's First Theorem Part One

Larry Susanka

November 29, 2019

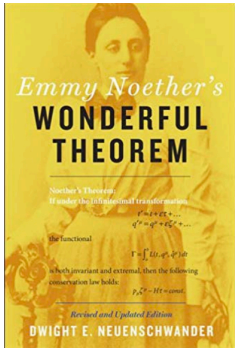
Introduction

Emmy Noether was recruited by Einstein and Hilbert to help them understand how conservation of energy worked in the new theory of General Relativity. She succeeded with this problem in 1918, proving theorems of broad generality that could be applied to many physical theories, including theories such as the modern Standard Model of particle physics.

The aim of this talk is to describe the first of these results and set the stage for later applications. We will concentrate on Noether's Theorem One with one independent variable.

The primary source for this particular talk will be Dwight Neuenschwander's *Emmy Noether's Wonderful Theorem 2nd Edition.*[1]

This is a great book (with well-organized exercises) that not only presents proofs of the theorems at a *reasonably* rigorous mathematical level but deals too with historical context, the physical meaning of these laws, examples and Emmy Noether herself.



[1] Note Second Edition!

When as an undergraduate I first met the elegant connections between symmetries and conservation laws through Lagrangian mechanics, I changed my major to physics and never looked back. In due time I learned how those connections in mechanics are special cases of a deeply profound theorem published by Emmy Noether in 1918. Noether's theorem as a work of mathematical physics is also a work of poetic beauty.

Dwight Neuenschwander
Southern Nazarene University

Noether's Theorems are mathematical theorems which can have physical context and physical consequences.

We will not shy away from physical examples, but will be most concerned with mathematical questions and the statements and contents of the theorems themselves.

For instance, we might discuss a damped oscillator and an application of Noether's First Theorem to that situation, but we will not develop or justify in any way the physical principles that lead to the model.

VARIABLES

There are three types of quantities we will use today to define the state of a physical system:

- time, a real number denoted by t
- geometrical variables generally denoted by individual coordinates $q^1, \ldots, q^N$, or simply the configuration vector $q \in \mathbb{R}^N$.

  (In more general situations than those considered today q could correspond to a point in a coordinate patch of an N-dimensional manifold, the configuration manifold.)
- dynamical variables generally denoted by $\dot{q}^1, \ldots, \dot{q}^N$, or simply $\dot{q}$, also in $\mathbb{R}^N$.

  (Such a $\dot{q}$ could represent a vector in the tangent space at a point q of a configuration manifold.)

If a function has name q the symbols $q(t)$ indicate one of its values, not the whole function. However it is standard to use the locution "the function $q(t)$" to emphasize the symbol we intend to use for the independent variable, and I don't intend to fight this practice. You simply must deduce whether $q(t)$ is to be a point or the whole function from context.

A prediction of the evolution of a physical system will correspond to a time-dependent path $q(t)$ through various "configuration vector" values.

Associated with this path will be a tangent vectors $\frac{dq}{dt}(t)$ which will correspond to a "dynamical vectors" $\dot{q}(t)$.

Using such a path, at each time we can produce a point

$$(t, q(t), \dot{q}(t)) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N.$$

Consistent

Specific coordinate systems are a convenience, not part of nature. Many coordinate systems can be used and an "answer," a number or vector such as momentum produced to describe something real using these coordinates, should be consistent.

That is to say: physical law is to be associated with a collection of preferred coordinate systems, and must be organized in such a way that when predictions are made using the law in two of the preferred coordinate systems and *after interpretation to account for different coordinates*, both calculations produce an equivalent "answer."

Maxwell's Theory of Electromagnetism with Galilean coordinate changes—Galilean Relativity—notoriously fails this test.

Invariant

The phrase *invariant under an "admissible" coordinate change* here means that the number or vector, though calculated in coordinates, need not be "interpreted." It is actually left unaltered when you change to this new "admissible" coordinate system.

The "admissible" coordinate changes we have in mind will be translations, or rotations, or a change in the origin of time, though our results will apply in other cases too. These collections of coordinate changes will, each, form groups that can be described by continuous parameters and represent, geometrically, symmetries in the space on which they act.

(Actually, the full Lie group structure is not used in our work. Only coordinate changes corresponding to very small or "infinitesimal" values of the parameter were used by Noether and needed by us.)

Conserved

The word *conserved* on the other hand refers to equal numerical measurements that are obtained within a single reference frame over a time interval, such as the total momentum of particles before and after some interaction.

They don't imply that this number will be the same in different coordinates.

Noether's Theorems tell us when and how invariance of a functional called the action under a continuous symmetry group of coordinate changes is related to conservation laws on physically realized paths in the configuration space.

A Functional

All functions will be presumed to have whatever degree of continuity/differentiability required for a given calculation.

Consider the function

$$L\colon \mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^N\to\mathbb{R}.$$

In the simplest applications L might just be $T-V$, the difference between kinetic and potential energy at a certain instant in the evolution of a system of particles.[2]

In that example L would have units of energy, $[\mathrm{kg\,m^2\,s^{-2}}]$.

[2] Our Lagrangians will depend on time, position and velocity of a path in configuration space. Examples do exist—though they are much less common—requiring higher derivatives. Techniques similar to the ones we will develop make progress there too.

We will let $\mathcal{S}$ be a specified set of functions of the form

$$q\colon (a,b) \to \mathbb{R}^N$$

where (a,b) is a generic (i.e. not fixed for all members of $\mathcal{S}$) interval in $\mathbb{R}$.

Members of $\mathcal{S}$ are potential configuration paths[3] and for the simplest examples the coordinates of q might have units of $[\mathrm{m}]$.

The derivatives of the functions in $\mathcal{S}$, of the form

$$\dot{q}\colon (a,b) \to \mathbb{R}^N,$$

would then have units $[\mathrm{m\,s^{-1}}]$.

$\mathcal{S}$ must be a pretty rich class of functions, containing many vector spaces of functions.[4]

[3] These configuration paths are not restricted to interpretation as rectangular coordinate paths. Kinetic energy therefore *cannot be presumed* to be given in terms of $\dot{q}\cdot\dot{q}$!

[4] For instance, the twice continuously differentiable functions on various domain time intervals.

We define $\Gamma\colon \mathcal{S} \to \mathbb{R}$ to be

$$\Gamma(q) = \int_a^b L(t, q(t), \dot{q}(t))\, dt$$

where $\dot{q}(t) = \frac{dq}{dt}(t)$.

The time dependence of q and $\dot{q}$, and even a reminder that there is an explicit path from $q(a)$ to $q(b)$, may be suppressed as in

$$\Gamma = \int_a^b L(t, q, \dot{q})\, dt = \int_a^b L\, dt.$$

Γ is called a *functional* (never linear!) and $\Gamma(q)$ is sometimes called the *action* of q on the interval $[a,b]$.

In our simplest example, action can have units of $[\mathrm{kg\,m^2\,s^{-1}}]$.

One very common type of physical theory, of concern to us here, provides a differential equation that must be satisfied by a physical system as it moves from one geometrical-variable combination to another.

Alternatively, and equivalently, we can produce a specific function L defined on solution paths at each time called the *Lagrangian* together with the assertion that the *actual* path taken must have stationary (usually minimum) action.

Hamilton's Principle tells us that in many situations in mechanics this Lagrangian should be the difference between total kinetic and total potential energies in a physical scenario, leaving it up to us to find representations for these.

Differential equations are purely local and build a solution by specifying incremental changes in configuration.

Specifying the solution as the path with least action is a global condition on the solution path and neighboring paths. We will not attempt to justify this principle here, but there are various situations in which this global condition can be deduced.

For instance, a classical description of particle motion can be seen as a limiting case (as Planck's constant $h \to 0$) of the quantum formalism of path integration. Stationary paths are selected in this limit as a result of interference of amplitudes along all possible solutions. Stationary paths are the only paths that remain (in the limit) after interference wipes out the rest.

There have, historically, been many versions of this *Principle of Least Action* in one context or another.

The earliest is *Fermat's principle* (1662) and then *Maupertuis' principle* (1744) and *Euler's principle* (also 1744) and finally *Hamilton's principle* (1834).

What we are doing now does not produce the Physical Law and Lagrangian. Those are the product of the inspiration and genius of Physicists.

But given a Physical Law and Lagrangian this principle has lots of consequences.

First we must determine what we mean by "stationary action."

Notation Clarification

We have used variables t, q and $\dot{q}$ and a function $L(t, q, \dot{q})$ and functions $q(t)$ and $\frac{dq}{dt}(t) = \dot{q}(t)$.

If you see in this or similar situations a notation such as $\frac{\partial}{\partial t}$ or $\frac{\partial}{\partial \dot{q}^i}$ it means the function to which it applies is assumed *for the purposes of the initial calculation* to have t or $\dot{q}^i$ as one of its variables and other variables are constant. After that partial derivative is found, a path $q(t)$ and its derivative $\dot{q}(t)$ might be substituted to create a function of t.

$$\frac{\partial L}{\partial t} = \frac{\partial}{\partial t} L(t, q, \dot{q}) \quad \text{and} \quad \frac{\partial}{\partial \dot{q}^i} L(t, q, \dot{q})$$

are presumed for purposes of the calculation to have t dependency *only* in its first coordinate and $\dot{q}^i$ dependency in the $i + N + 1$ coordinate only, even if you have a path $q(t)$ in mind!

Notation Clarification

However a total derivative such as $\frac{dL}{dt} = \dot{L}$ will indicate that a path $q(t)$ has been selected (even if it is not shown here!) and fed to L before differentiating that composite function:

$$\dot{L}(t) = \frac{d}{dt} L(t, q(t), \dot{q}(t)).$$

This function of t must be calculated by the chain rule:

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{\partial L}{\partial q^i}\dot{q}^i + \frac{\partial L}{\partial \dot{q}^i}\ddot{q}^i$$

where we use the Einstein summation convention to indicate summation over all values of an index variable in terms containing a product of exactly two terms with the same index, one "high" and the other "low."

The Euler-Lagrange Equations

Suppose q is a path, a potential description of a physical evolution from time a to time b , and ζ is a path with $\zeta(a) = 0 = \zeta(b)$.

Define, for small real ϵ, the neighboring path Q_ϵ to be $q + \epsilon\,\zeta$.

We can define

$$\Gamma_\epsilon = \int_a^b L(t, Q_\epsilon, \dot{Q}_\epsilon)\,dt = \int_a^b L(t, q + \epsilon\,\zeta, \dot{q} + \epsilon\,\dot{\zeta})\,dt.$$

So $q = Q_0$, and of course then $\Gamma = \Gamma_0$.

We say q is a stationary path with respect to ζ provided $\frac{d}{d\epsilon}\big|_{\epsilon=0}\Gamma_\epsilon = 0$, and say

q is a stationary path if it is stationary with respect to every ζ.

Assuming differentiability everywhere, we have by Leibniz's Integral Rule

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=\varepsilon_0} \Gamma_\epsilon = \left.\frac{d}{d\epsilon}\right|_{\epsilon=\varepsilon_0} \int_a^b L(t, Q_\epsilon, \dot{Q}_\epsilon)\, dt$$
$$= \int_a^b \left.\frac{d}{d\epsilon}\right|_{\epsilon=\varepsilon_0} L(t, Q_\epsilon, \dot{Q}_\epsilon)\, dt$$

and this equality will hold whether the derivative is 0 or not, and whether ε_0 is 0 or not, though the case $\varepsilon_0 = 0$ is actually the one that concerns us.

We need to calculate the interior derivative by the chain rule from multi-variable calculus and, eventually, find out what it implies if that derivative is 0 for this (and every) ζ.

Let's change notation briefly. (It can't hurt to be double-clear about what this notation means.) Define

$$x = (x^0, x^1, \ldots, x^{2N}) = (t, Q_\epsilon, \dot{Q}_\epsilon)$$
$$= (t, q^1 + \epsilon\zeta^1, \ldots, q^N + \epsilon\zeta^N, \dot{q}^1 + \epsilon\dot{\zeta}^1, \ldots, \dot{q}^N + \epsilon\dot{\zeta}^N).$$

t (and so x^0) is independent of ϵ, but the other coordinates *do* depend on ϵ. Note that when $\epsilon = 0$ we have $x = (t, q, \dot{q})$.

Let's fix t for the next few lines. Avoiding the Einsten summation convention for the next few slides we have

$$\frac{dL \circ x}{d\epsilon}(\varepsilon) = (D_0 L)(x(\varepsilon))\frac{dx^0}{d\epsilon}(\varepsilon) + \sum_{i=1}^N (D_i L)(x(\varepsilon))\frac{dx^i}{d\epsilon}(\varepsilon) + \sum_{i=1}^N (D_{N+i} L)(x(\varepsilon))\frac{dx^{N+i}}{d\epsilon}(\varepsilon)$$

First of all, $\frac{dx^0}{d\epsilon}(\varepsilon) = 0$ for every ε so the first term in the expansion can be ignored. Also, for $i = 1, \ldots, N$

$$\frac{dx^i}{d\epsilon}(\varepsilon) = \frac{d}{d\epsilon}(q^i + \epsilon\zeta^i)(\varepsilon) = \zeta^i$$

and

$$\frac{dx^{N+i}}{d\epsilon}(\varepsilon) = \frac{d}{d\epsilon}(\dot{q}^i + \epsilon\dot{\zeta}^i)(\varepsilon) = \dot{\zeta}^i.$$

So these derivatives don't actually depend on ε.

Substituting those into the derivative expression gives ……

$$\frac{dL \circ x}{d\epsilon}(\varepsilon) = \sum_{i=1}^N (D_i L)(x(\varepsilon))\zeta^i + \sum_{i=1}^N (D_{N+i} L)(x(\varepsilon))\dot{\zeta}^i.$$

And then evaluating at $\varepsilon = 0$ we have

$$\left.\frac{dL \circ x}{d\epsilon}\right|_{\varepsilon=0} = \sum_{i=1}^N (D_i L)(t, q, \dot{q})\zeta^i + \sum_{i=1}^N (D_{N+i} L)(t, q, \dot{q})\dot{\zeta}^i$$

and in typical shorthand notation we have, for each t,

$$\left.\frac{dL \circ x}{d\epsilon}\right|_{\varepsilon=0} = \sum_{i=1}^N \left(\frac{\partial L}{\partial q^i}\zeta^i + \frac{\partial L}{\partial \dot{q}^i}\dot{\zeta}^i\right).$$

This is the integrand of the derivative of the action on functions that vary from q by small multiples of ζ.

Going back to the compact Einstein summation convention we have

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \Gamma_\epsilon = \int_a^b \left(\frac{\partial L}{\partial q^i}\zeta^i + \frac{\partial L}{\partial \dot{q}^i}\dot{\zeta}^i \right) dt.$$

The partial derivatives indicated here simply refer to the partial with respect to the ith (or $(i+N)$th) "slot" of the Lagrangian function L. This is then *evaluated* at $(t, q(t), \dot{q}(t))$ and *only then* does each become a function of t.

Note: *These partial derivative expressions do not depend on ζ!*

We want to rewrite some of the terms using integration by parts.

$$\int_a^b \frac{\partial L}{\partial \dot{q}^i}\dot{\zeta}^i\, dt = \left.\frac{\partial L}{\partial \dot{q}^i}\zeta^i\right|_a^b - \int_a^b \zeta^i \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\, dt.$$

Since $\zeta(a) = \zeta(b) = 0$ the first term on the right is discarded and we have

$$\int_a^b \frac{\partial L}{\partial \dot{q}^i}\dot{\zeta}^i\, dt = -\int_a^b \zeta^i \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\, dt.$$

We then substitute this into the formula for $\left.\frac{d}{d\epsilon}\right|_{\epsilon=0}\Gamma_\epsilon$.

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \Gamma_\epsilon = \int_a^b \left(\frac{\partial L}{\partial q^i}\zeta^i + \frac{\partial L}{\partial \dot{q}^i}\dot{\zeta}^i \right) dt = \int_a^b \left(\frac{\partial L}{\partial q^i}\zeta^i - \zeta^i \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \right) dt$$
$$= \int_a^b \zeta^i \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \right) dt.$$

If q is a stationary path then this derivative is 0 for every ζ. Since each component of ζ is taken, independently of the other components, from a dense subset of the integrable functions, a standard result from integration theory tells us that the factors

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \qquad i = 1, \ldots, N$$

must be the zero function for each i.

With the stationarity condition we have the Euler-Lagrange System of Differential Equations, or "ELE" for short.

The Euler-Lagrange System of Equations in classical notation

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0 \qquad i = 1, \ldots, N$$

contains partial derivatives and a total time derivative.

A specific function $q(t)$ must be in mind for these equations to make sense.

The partial derivatives are evaluated at $(t, q(t), \dot{q}(t))$ to produce functions of time and this is a system of N second order ordinary differential equations in the N functions $q_1, \ldots, q_N$ of variable t.

The ELE system of *ordinary* differential equations:

$$\frac{\partial L}{\partial q^i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \qquad i = 1, \ldots, N$$

can, with appropriate boundary conditions, be solved for q numerically or, in a few very-special but often-presented cases, exactly.

Expanding $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}$ we have explicit second order system

$$\frac{\partial L}{\partial q^i} = \frac{\partial^2 L}{\partial t \partial \dot{q}^i} + \frac{\partial^2 L}{\partial q^\mu \partial \dot{q}^i}\dot{q}^\mu + \frac{\partial^2 L}{\partial \dot{q}^\mu \partial \dot{q}^i}\ddot{q}^\mu \qquad i = 1, \ldots, N.$$

All these steps are reversible, so if the ELE system holds q will be stationary with respect to Lagrangian L, establishing the equivalence of this particular system of second order *ordinary* differential equations with the stationarity condition on the integral of the Lagrangian.

The equations

$$\frac{\partial L}{\partial q^i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \qquad i = 1, \ldots, N$$

suggest that the rate of change of $p_i = \frac{\partial L}{\partial \dot{q}^i}$ is important and we call this quantity, which in our simplest example case has units $[\mathrm{kg\,m\,s^{-1}}]$, a generalized momentum.

p_i is also said to be the canonical momentum *conjugate* to the generalized coordinate q^i.

With this definition, the ELE System in "p dot" form is

$$\frac{\partial L}{\partial q^i} = \dot{p}_i \qquad i = 1, \ldots, N.$$

The total time derivative of the composition $L(t, q, \dot{q})$ on a solution path q is given by

$$\begin{aligned}\frac{dL}{dt} &= \frac{\partial L}{\partial t} + \frac{\partial L}{\partial q^i}\dot{q}^i + \frac{\partial L}{\partial \dot{q}^i}\ddot{q}^i \\ &= \frac{\partial L}{\partial t} + \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\right)\dot{q}^i + p_i\ddot{q}^i \\ &= \frac{\partial L}{\partial t} + \dot{p}_i\dot{q}^i + p_i\ddot{q}^i\end{aligned}$$

Isolating $\frac{\partial L}{\partial t}$ we have

$$\frac{\partial L}{\partial t} = \frac{dL}{dt} - \dot{p}_i\dot{q}^i - p_i\ddot{q}^i = -\frac{d}{dt}\left(p_i\dot{q}^i - L\right).$$

$$\frac{\partial L}{\partial t} = -\frac{d}{dt}\left(p_i\dot{q}^i - L\right).$$

So the time derivative of

$$H = p_i\dot{q}^i - L$$

seems to be important.

We can see that this number is preserved on a solution path precisely when L is independent of t on that path.

H is called the Hamiltonian of the path in the physical scenario described by L.

We have the following "H dot" implication of the ELE system, applied to the Hamiltonian and evaluated along a solution path:

$$\frac{\partial L}{\partial t} = -\dot{H}.$$

The Hamiltonian *function*

$$H = p_i \dot{q}^i - L = \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i - L$$

can be defined on all of $\mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N$ without regard to any path.

The definition of the Hamiltonian function *is therefore not* the sum of potential and kinetic energies, though you might have heard it described this way. It is this particular function, created from the Lagrangian alone.

When composed with a path $q(t)$, as we will often do, it is a function of time which *may* be an energy of *something* along that path, depending on the specific context.

The two functions L and

$$H = p_i \dot{q}^i - L = \left(\frac{\partial L}{\partial \dot{q}^i} \right) \dot{q}^i - L$$

are related to each other by what is called a Legendre transformation.

Lagrangian→Hamiltonian is just one example of a pair related this way in applications.

For instance, in thermodynamics the Helmholtz free energy, A, and Gibbs energy, G, are obtained by performing Legendre transforms on the internal energy and enthalpy, respectively,

Other examples are found in microeconomics, probability and the solution of differential equations in several variables, according to Wikipedia.

What Noether's Theorems Say

The ELE "H dot" and "p dot" equations

$$\frac{\partial L}{\partial t} = -\dot{H} \quad \text{and} \quad \frac{\partial L}{\partial q^i} = \dot{p}_i \quad i = 1, \ldots, N.$$

tell us when the Hamiltonian and canonical momenta will be constants along solution paths. This will happen when time is not explicitly present in the Lagrangian, or when L is independent of one of the configuration coordinates. This independence corresponds to continuous translational symmetries in time or the configuration space under which L is invariant and we have found here special cases of Noether's first theorem.

Noether's Theorem tells us when and how conserved quantities corresponding to continuous symmetries such as these can be realized as combinations of H and conjugate momenta.

Simple Harmonic Oscillator

We will explore several examples from Part One of these talks. Our purpose here is to illustrate the notation in use, not delve deeply into the physics of the particular case.

Here we examine a simple harmonic oscillator.

$$L(x, \dot{x}) = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} k x^2.$$

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \Rightarrow -kx = \frac{d}{dt} (m\dot{x}) \Rightarrow -kx = m\ddot{x} \Rightarrow \quad \ddot{x} = -\frac{k}{m} x.$$

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}. \qquad (\textit{Not Conserved because } \frac{\partial L}{\partial x} \neq 0.)$$

$$H = p\dot{x} - L = m\dot{x}^2 - \frac{1}{2} m \dot{x}^2 + \frac{1}{2} k x^2$$

$$= \frac{1}{2} k x^2 + \frac{1}{2} m \dot{x}^2. \qquad (\textit{Conserved because } \frac{\partial L}{\partial t} = 0.)$$

Pendulum, Fixed Length String

$$L(\theta,\dot\theta) = \frac{1}{2}mR^2\dot\theta^2 - mgR(1-\cos(\theta)).$$

$$\begin{aligned}\frac{\partial L}{\partial\theta} = \frac{d}{dt}\frac{\partial L}{\partial\dot\theta} &\Rightarrow -mgR\sin(\theta) = \frac{d}{dt}\left(mR^2\dot\theta\right)\\ &\Rightarrow -mgR\sin(\theta) = mR^2\ddot\theta\\ &\Rightarrow \quad \ddot\theta = -\frac{g}{R}\sin(\theta).\end{aligned}$$

Pendulum, Fixed Length String

$$p = \frac{\partial L}{\partial\dot\theta} = mR^2\dot\theta.$$

$$\begin{aligned}H = p\dot\theta - L &= mR^2\dot\theta^2 - \frac{1}{2}mR^2\dot\theta^2 + mgR(1-\cos(\theta))\\ &= \frac{1}{2}mR^2\dot\theta^2 + mgR(1-\cos(\theta)).\end{aligned}$$

$$\begin{aligned}0 = \frac{\partial L}{\partial t} = -\frac{dH}{dt} &= -mR^2\dot\theta\ddot\theta - mgR\sin(\theta)\dot\theta\\ \Longrightarrow \quad \ddot\theta &= -\frac{g}{R}\sin(\theta).\end{aligned}$$

Pendulum on a Spring with Relaxed Length R

$$L(r,\dot r,\theta,\dot\theta) = \frac{1}{2}m\dot r^2 + \frac{1}{2}m(R+r)^2\dot\theta^2 - mg(R+r)(1-\cos(\theta)) - \frac{1}{2}kr^2.$$

$$\begin{aligned}\frac{\partial L}{\partial\theta} = \frac{d}{dt}\frac{\partial L}{\partial\dot\theta} &\Rightarrow -mg(R+r)\sin(\theta) = \frac{d}{dt}\left(m(R+r)^2\dot\theta\right)\\ &\Rightarrow -mg(R+r)\sin(\theta) = m(R+r)^2\ddot\theta + 2m(R+r)\dot\theta\dot r.\end{aligned}$$

$$\begin{aligned}\frac{\partial L}{\partial r} = \frac{d}{dt}\frac{\partial L}{\partial\dot r} &\Rightarrow 2m(R+r)\dot\theta^2\dot r - mg\dot r(1-\cos(\theta)) - kr\dot r = \frac{d}{dt}(m\dot r)\\ &\Rightarrow 2m(R+r)\dot\theta^2\dot r - mg\dot r(1-\cos(\theta)) - kr\dot r = m\ddot r.\end{aligned}$$

Pendulum on a Spring with Relaxed Length R

$$L(r,\dot r,\theta,\dot\theta) = \frac{1}{2}m\dot r^2 + \frac{1}{2}m(R+r)^2\dot\theta^2 - mg(R+r)(1-\cos(\theta)) - \frac{1}{2}kr^2.$$

$$\begin{aligned}p_\theta &= \frac{\partial L}{\partial\dot\theta} = -mg(R+r)\sin(\theta)\\ p_r &= \frac{\partial L}{\partial\dot r} = mg\dot r\end{aligned}$$

$$\begin{aligned}H = p_\theta\dot\theta + p_r\dot r - L &= -mg(R+r)\sin(\theta)\dot\theta + mg\dot r^2 - L\\ &= -mg(R+r)\sin(\theta)\dot\theta + \frac{1}{2}m\dot r^2\\ &\quad - \frac{1}{2}m(R+r)^2\dot\theta^2 + mg(R+r)(1-\cos(\theta)) + \frac{1}{2}kr^2\end{aligned}$$

Since $0 = \frac{\partial L}{\partial t} = -\frac{dH}{dt}$ the Hamiltonian is constant on solutions.

Central Force Motion

Spherical coordinates identify points in space by triples (r, θ, ϕ) where r is the distance from an origin, θ is an angle from a z axis to the position vector, and ϕ is the angle from an x axis to the projection of the position vector on the xy plane.

So the rectangular coordinates of this point are

$$(r\sin(\theta)\cos(\phi), r\sin(\theta)\sin(\phi), r\cos(\theta)).$$

Squared velocity of a point moving in these coordinates (combine and simplify 22 terms!) is

$$\vec{v}\cdot\vec{v} = \dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2(\theta)$$

The Lagrangian here will be of the form

$$L(r,\theta,\dot{r},\dot{\theta},\dot{\phi}) = \frac{1}{2}m\vec{v}\cdot\vec{v} - U(r) = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2(\theta)\right) - U(r)$$

for some potential function U.

Central Force Motion

$$L(r,\theta,\dot{r},\dot{\theta},\dot{\phi}) = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2(\theta)\right) - U(r)$$

Since L has no time dependence the Hamiltonian will be constant. And since it has no dependence on ϕ the canonical momentum p_ϕ will also be a constant of the motion.

$$p_\phi = \frac{\partial L}{\partial\dot{\phi}} = mr^2\dot{\phi}\sin^2(\theta).$$

This constant is the z component of the angular momentum, and $mr^2\sin^2(\theta)$ is the moment of inertia about the z axis.

Constraints on the Configuration Space

Many physical systems feature constraints: the motion is confined to a subset of the configuration space and is not all of $\mathbb{R}^N$. One way of thinking of this is to imagine we are confined to a configuration manifold, $M \subset \mathbb{R}^N$.

These constraints are often (usually) specified by one or more equations of the form $h(t,q) = 0$ for smooth h.

These may be handled, without dealing with the configuration manifold directly, by the introduction of extra terms to produce a new Lagrangian. We "add multiples of zero" to the old Lagrangian:

$$L_c = L + \lambda_k h^k.$$

These extra terms involve Lagrange multipliers, one for each constraint equation. These multipliers can be solved for to find additional equations linking the geometrical variables.

A Bead on a Wire

For instance consider a bead constrained to slide on a frictionless wire in the shape of a parabola, $y = x^2$.

The unconstrained Lagrangian for a bead moving under the influence of gravity is

$$L(x,y,\dot{x},\dot{y}) = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) - mg.$$

Adding the constraint $x^2 - y = 0$ produces constrained Lagrangian

$$L_c(x,y,\dot{x},\dot{y},\lambda) = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) - mgy + \lambda(x^2 - y).$$

A Bead on a Wire

$$L_c(x, y, \dot{x}, \dot{y}, \lambda) = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) - mgy + \lambda(x^2 - y).$$

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} \Rightarrow 2\lambda x = m\ddot{x} \Rightarrow \frac{\lambda}{m} = \frac{\ddot{x}}{2x}.$$

$$\frac{\partial L}{\partial y} = \frac{d}{dt}\frac{\partial L}{\partial \dot{y}} \Rightarrow -\lambda - mg = m\ddot{y} \Rightarrow \frac{\lambda}{m} = -g - \ddot{y}.$$

$$\ddot{x} = 2x\left(-g - \ddot{y}\right).$$

A Bead on a Wire

$$\frac{\partial L}{\partial \lambda} = \frac{d}{dt}\frac{\partial L}{\partial \dot{\lambda}} = 0 \Rightarrow x^2 - y = 0 \Rightarrow \ddot{y} = 2\dot{x}^2 + 2x\ddot{x}.$$

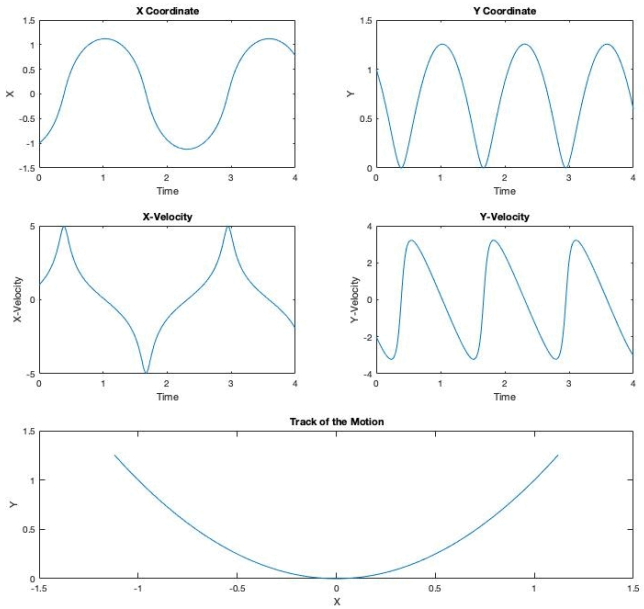
$$\ddot{x} = 2x\left(-g - \ddot{y}\right) \Rightarrow \ddot{x} = \frac{-4x\dot{x}^2 - 2xg}{1 + 4x^2}.$$

This produces the system

$$\dot{x} = w$$
$$\dot{w} = \frac{-4xw^2 - 2xg}{1 + 4x^2}$$

which can be solved, with initial x and $\dot{x}$ values, to whatever level of accuracy is required by numerical (Runge-Kutta, or even Euler) methods.

Bead on a Frictionless Parabolic Wire



To Be Continued …

In the central force motion example, spherical coordinates do not easily reveal other possible constants of the motion.

And how *do* you change coordinates so that L is constant along one axis, and is that necessary to find constants of the motion?

To Be Continued…

References

Dwight Neuenschwander, *Emmy Noether's Wonderful Theorem 2nd Edition*

Yvette Kosmann-Schwarzbach, *The Noether Theorems*

Gennadi Sardanashvily, *Noether's Theorems*

André Mercier, *Variational Principles of Physics*

Jacob Schwichtenberg, *Physics From Symmetry 2nd Edition*

Wolfgang Yourgrau and Stanley Mandelstam, *Variational Principles in Dynamics and Quantum Theory*