APPLICATIONS OF VARIATIONAL PRINCIPLES TO DYNAMICS AND CONSERVATION LAWS IN PHYSICS

DANIEL J OLDER

Abstract. Much of physics can be condensed and simplified using the principle of least action from the calculus of variations. After introducing some basic concepts such as a functional, the variation of a functional and the condition required for a differentiable functional to have an extremum, the action, Lagrangian and Hamiltonian of a physical system will be introduced. Applying the principle of least action to a physical system, not only do the Euler-Lagrange equations (the dynamic equations for particles or field equations for fields) come directly from the differential of the action but so do all of the conservation laws associated with the physical system by Noether's theorem. Applying the main results to particles, we will derive Newton's second law of motion and the conservation of energy and momentum in classical mechanics. Then, applying the results to fields, Maxwell's equations for electromagnetism will be derived as well as the energy-momentum tensor for a classical field.

Contents

Date: August 12, 2012.

1. Introduction

Almost all physical systems, whether they be single particles or systems of particles, scalar fields or vector fields, can be characterized by a functional called the action. The action is a function of all the possible states that the physical system could be in, and it maps each state to a real number. Any physical system with an action can have its physical laws, ordinary or partial differential equations governing the evolution of the system defined by the principle of least action. That is, one can derive the laws of physics by finding the states of the system that minimize the action. When sufficient initial conditions are imposed on the system, along with the principle of least action, the system is defined entirely as far as classical physics is concerned. An interesting thing happens, however, when we make a transformation to the underlying spacetime. Any system must be defined on some set of coordinates. When a change is made to those coordinates like when the space is translated or rotated and the physical system remains unchanged, this transformation defines a symmetry of the physical system. It turns out that each symmetry of a physical system determines a unique conservation law, and this is known as Noether's theorem, a mathematical theorem of fundamental importance to all of modern physics.

2. Calculus of Variations Basics

In order to understand and work with this concept of action, which takes any possible state of a physical system and maps it to a real number, we first have to understand it in a more rigorous mathematical way. The state of a physical system between time t_0 and t_1 is defined by a set of functions. Looking at a system of n particles, those functions would be maps from time to position in space and thus, any plausible state of the system in a given time interval would be defined by n functions or n paths through space. An action is a type of function known as a functional. Its domain is a set of functions, and its codomain is a real number. To begin with, let us define this domain.

Definition 2.1. *A function space is a normed vector space in which the elements are functions from $\mathbf{R}^n$ to $\mathbf{R}^m$.*

For our purposes, the function spaces to be dealt with will be $C^\infty(\mathbf{R};\mathbf{R})$ for our systems of particles, which we will denote by $\mathbf{P}$ and $C^\infty(\mathbf{R}^4;\mathbf{R})$ for fields, which we will denote by $\mathbf{F}$. The derivative of a function $f \in \mathbf{P}$ will be denoted by a dot over it $\dot{f}$ when convenient to simplify notation. An element in $\mathbf{R}^4$ will have four components x_μ where the index μ runs from 0 to 3. Partial differentiation of a function $f \in \mathbf{F}$ with respect to x_μ will be denoted $\partial_\mu f$. For n particles moving in 3 dimensions, our function space will be $\mathbf{P}^{3n}$. For n scalar fields, our function space will be $\mathbf{F}^n$.

Example 2.2. *$L^1[a, b]$ and $C^1[a, b]$ denote integrable functions and differentiable functions on $[a, b]$, respectively.*

Definition 2.3. *A functional is a map $f: X \to \mathbf{R}$ where X is a function space.*

Example 2.4. *Suppose $a < b$. The map from f to its integral on $[a, b]$ defines a functional.*

$$J : L^1[a, b] \to \mathbf{R}$$
$$J(f) = \int_a^b f(x)\, dx.$$

Example 2.5. *Suppose $a < b$. The map taking $f \in C^1[a, b]$ to the length of it's graph between a and b is a functional that sends*

$$J : C^1[a, b] \to \mathbf{R}.$$
$$J(f) = \int_a^b \sqrt{1 + f'(x)^2}\, dx.$$

Definition 2.6. *Let y be a fixed function in a function space X. Let h be another arbitrary function in X and let ϵ be a real number. A functional J defined on X is differentiable at y if:*

$$\Delta J = J(y + \epsilon h) - J(y) = \delta J(\epsilon h) + o(\epsilon) \quad \forall h \in X, \forall \epsilon \in \mathbf{R}.$$

Here, δJ denotes a linear functional where linear is defined in the usual way and $o(\epsilon)$ denotes a function with all terms of order greater than 1 with respect to ϵ.

Remark 2.7. *At any fixed y, if J is differentiable, then it is not difficult to show that the linear functional δJ is uniquely defined. We call this linear functional δJ the differential of J at y.*

Definition 2.8. *A functional J defined on function space X has a relative minimum at $y = y_0$ if there exists a $\delta > 0$ such that*

$$0 < |\epsilon| < \delta \Rightarrow J(y_0) \leq J(y_0 + \epsilon h) \quad \forall h \in X.$$

Theorem 2.9. *If a functional J defined on X is differentiable and has a minimum at y_0, then $\delta J(\epsilon h) = 0$ for all $h \in X$ and for all ϵ at the point y_0,.*

Proof. By definition of J having a minimum at y_0, we have

$$J(y_0 + \epsilon h) - J(y_0) = \delta J(\epsilon h) + o(\epsilon) \geq 0$$

for all ϵ satisfying $0 < |\epsilon| < \delta$ for some $\delta > 0$. This implies that $\delta J(\epsilon h) \geq 0$ when ϵ is small enough since $\Delta J(\epsilon h) \sim \delta J(\epsilon h)$. The $\sim$ denotes equality of principle linear parts of two functions or functionals. We will be using this notation a lot in the next section. Since epsilon can be positive or negative we can also obtain $\delta J(-\epsilon h) = -\delta J(\epsilon h) \geq 0$. Therefore, it must be that $\delta J(\epsilon h) = 0$ at the point y_0. □

From now on, whenever a functional J has a minimum at some point, we will simply denote its differential as δJ since it will be identically zero and not depend on the variation from that point.

Theorem 2.10. *Let α be a continuous function in* $\mathbf{P}$ *defined on the interval* $[t_0, t_1]$. *If*

$$\int_{t_0}^{t_1} \alpha(t)h(t)\, dt = 0$$

for any continuous function h on $[t_0, t_1]$ satisfying the conditions that $h(t_0) = h(t_1) = 0$, then $\alpha(t) = 0$ for all $t \in [t_0, t_1]$.

Proof. Suppose α is nonzero at some point in the interval $[t_0, t_1]$. Then this α must be nonzero in some interval (a, b) contained in $[t_0, t_1]$. Let us define a continuous function $h(t) = \left|\left|t - \left(a + \frac{b-a}{2}\right)\right| - \frac{b-a}{2}\right|$ if $t \in (a, b)$ and $h(t) = 0$ otherwise. Thus αh is completely positive or negative on (a, b) so therefore

$$\int_{t_0}^{t_1} \alpha(t)h(t)\, dt \neq 0.$$

which is a contradiction. Consequently, $\alpha(t) = 0$ for all $t \in [t_0, t_1]$ □

Theorem 2.11. *Let ψ be continuous function in* $\mathbf{F}$ *defined on a compact region* $R \subset \mathbf{R}^4$. *If*

$$\int_R \psi(x_\mu)\Omega(x_\mu)\, dx_\mu = 0$$

for any arbitrary continuous function Ω in R satisfying the conditions $\Omega(x_\mu) = 0$ for all $x_\mu \in \Gamma$ where Γ is the boundary of R, then $\psi(x_\mu) = 0$ for all $x_\mu \in R$.

Proof. Suppose ψ is nonzero at some point $\hat{x}_\mu \in R$. Then this ψ must be nonzero in some ball $B_\epsilon(\hat{x}_\mu) \subset R$. Let us define $\Omega(x_\mu) = ||x_\mu - \hat{x}_\mu| - \epsilon|$ for all $x_\mu \in B_\epsilon(\hat{x}_\mu)$ and $\Omega(x_\mu) = 0$ elsewhere. $\psi\Omega$ is completely positive or negative on $B_\epsilon(\hat{x}_\mu)$ so therefore

$$\int_R \psi(x_\mu)\Omega(x_\mu)\, dx_\mu \neq 0.$$

which is a contradiction. Consequently, $\psi(x_\mu) = 0$ for all $x_\mu \in R$ □

Definition 2.12. *Suppose we have a function space X and functional J such that*

$$X = \{f \colon f \colon \mathbf{R}^m \to \mathbf{R}\}.$$
$$J : X^n \to \mathbf{R}$$
$$J(f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m)) = J(f_i(x_j)) = \int_R F\Big(x_j, f_i(x_j), \frac{\partial f_i}{\partial x_j}\Big)\, dx.$$

Here, R is a compact region in $\mathbf{R}^m$. Let us introduce a transformation of x_j and f_i as such

$$x_j^* = x_j^*(x_j, f_i(x_j), \frac{\partial f_i}{\partial x_j}, \epsilon). \qquad f_i^*(x_j^*) = f_i^*(x_j, f_i(x_j), \frac{\partial f_i}{\partial x_j}, \epsilon).$$

$$\textit{If} \quad J(f_i^*(x_j^*)) = \int_{R^*} F(x_j^*, f_i^*(x_j^*), \frac{\partial f_i^*(x_j^*)}{\partial x_j^*})\, dx^* = J(f_i(x_j))$$

Then we say that J is invariant under the given transformation.

3. The Action, Lagrangian and Lagrangian Density

3.1. **Particles.** In this section, the action of a general system of particles and its differential will be considered. The Lagrangian L of a physical system is a function of time t, each of the particles' positions q_i and each of the particles' velocities $\dot{q}_i$. The action I of a physical system is the integral of the Lagrangian over a time interval $[t_0, t_1]$. It is a functional mapping the particles' paths in the time interval to a real number. The importance of action as a way of formulating physical laws is that the dynamics of the particles as well as the conservation laws come directly from the principle of least action. This principal involves simply finding the minimum of the action functional. Suppose we have a physical system with a Lagrangian L and we want to find the minimum of the action functional. The action for a system of particles with n degrees of freedom is denoted as

$$I(q_1, \ldots, q_n) = I(q_i) = \int_{t_0}^{t_1} L(q_i(t), \dot{q}_i(t))\, dt = \int_{t_0}^{t_1} L\, dt. \tag{1}$$

Let us introduce the transformation

$$t^* = t^*(t, q_i, \dot{q}_i, \epsilon). \qquad q_i^*(t^*) = q_i^*(t, q_i, \dot{q}_i, \epsilon). \tag{2}$$

$$t = t^*(t, q_i, \dot{q}_i, 0). \qquad q_i(t) = q_i^*(t, q_i, \dot{q}_i, 0). \tag{3}$$

The action after this transformation is then,

$$I(q_i^*(t^*)) = \int_{t_0^*}^{t_1{}^*} L(q_i^*(t^*), \frac{d(q_i{}^*(t^*))}{dt^*}\, dt^*.$$

By Taylor's theorem, we know

$$t^* = t + \frac{\partial t^*}{\partial \epsilon}\Big|_{\epsilon=0} \epsilon + o(\epsilon) \qquad q_i^*(t^*) = q_i(t) + \frac{\partial q_i^*(t^*)}{\partial \epsilon}\Big|_{\epsilon=0} \epsilon + o(\epsilon). \tag{4}$$

The second terms in both of these equations will be denoted δt and δq_i respectively. They represent the principal linear parts of the differences $\Delta t = t^* - t$ and $\Delta q_i =$

$q^*(t^*) - q_i(t)$. We then have that $\Delta t \sim \delta t$ and $\Delta q_i \sim \delta q_i$ Also, we can write

$$\Delta \dot{q}_i = \frac{d(q_i^*(t^*))}{dt^*} - \frac{d(q_i(t))}{dt} \sim \frac{d(q_i^*(t^*))}{dt} - \frac{d(q_i(t))}{dt^*} = \frac{d(\Delta q_i)}{dt} \sim \frac{d(\delta q_i)}{dt} = \delta \dot{q}_i \tag{5}$$

where $\delta \dot{q}_i$ is the principal linear part of the difference $\Delta \dot{q}_i$. This follows from the fact that

$$\frac{d}{dt} = \frac{d}{dt^*}\frac{dt^*}{dt} \sim \frac{d}{dt^*}\Big[\frac{d}{dt}(t + \delta t)\Big] = \frac{d}{dt^*}\Big[1 + \frac{d(\delta t)}{dt}\Big] \sim \frac{d}{dt^*}. \tag{6}$$

Our goal is to find the action differential δI. First, we must write out what the difference ΔI is explicitly.

$$\begin{aligned}\Delta I &= I(q_i^*(t^*)) - I(q_i(t) \\ &= \int_{t_0^*}^{t_1^*} L(q_i(t) + \Delta q_i, \dot{q}_i(t) + \Delta \dot{q}_i)\, dt^* - \int_{t_0}^{t_1} L(q_i(t), \dot{q}_i(t))\, dt. \end{aligned} \tag{7}$$

In order to get from this to the differential δI, we must eliminate all terms of order higher than 1 with respect to ϵ. First, expanding the transformed Lagrangian using Taylor's theorem and eliminating higher order terms, we get the expression

$$\Delta I \sim \int_{t_0^*}^{t_1^*} L + \sum_i^N \frac{\partial L}{\partial q_i}\delta q_i + \sum_i^N \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\, dt^* - \int_{t_0}^{t_1} L\, dt \tag{8}$$

Now we will put everything under the same integral and eliminate more higher order terms. What we arrive at is the actual action differential

$$\begin{aligned}\Delta I &\sim \int_{t_0}^{t_1} \Big[L + \sum_i^N \frac{\partial L}{\partial q_i}\delta q_i + \sum_i^N \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\Big]\frac{dt^*}{dt} - L\, dt \\ &\sim \int_{t_0}^{t_1} \Big[L + \sum_i^N \frac{\partial L}{\partial q_i}\delta q_i + \sum_i^N \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\Big]\Big[1 + \frac{d(\delta t)}{dt}\Big] - L\, dt \\ &\sim \int_{t_0}^{t_1} \sum_i^N \frac{\partial L}{\partial q_i}\delta q_i + \sum_i^N \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i + L\frac{d(\delta t)}{dt}\, dt \\ &= \delta I. \end{aligned} \tag{9}$$

Now we have an expression for the action differential but we can write it in a way that reveals much more about the dynamics of the particles and the conservation laws for the system. Our first manipulation to the expression above is to make a substitution for the terms δq_i and $\delta \bar{q}_i$. Let us denote $\Delta \bar{q}_i = q_i^*(t) - q_i(t)$ and let the principal linear difference be $\delta \bar{q}_i$. Then we can see that

$$\begin{aligned}\Delta q_i &= q_i^*(t^*) - q_i^*(t) + q_i^*(t) - q_i(t) \\ &\sim \frac{dq^*(t^*)}{dt}\Delta t + \Delta \bar{q}_i \\ &\sim \frac{dq^*(t^*)}{dt}\delta t + \delta \bar{q}_i \\ &\sim \frac{d(q(t) + \delta q_i)}{dt}\delta t + \delta \bar{q}_i \\ &\sim \dot{q}_i \delta t + \delta \bar{q}_i = \delta q_i \end{aligned} \tag{10}$$

Also, we can see that

$$\delta \dot{q}_i \sim \frac{d}{dt}(\dot{q}_i \delta t + \delta \bar{q}_i) = \frac{d\dot{q}_i}{dt}\delta t + \frac{d(\delta \bar{q}_i)}{dt}. \tag{11}$$

Plugging these new expressions for δq_i and $\delta \dot{q}_i$ into differential of the action and distributing out, we get

$$\delta I = \int_{t_0}^{t_1} \sum_i^n \frac{\partial L}{\partial q_i}\frac{dq_i(t)}{dt}\delta t + \sum_i^n \frac{\partial L}{\partial q_i}\delta \bar{q}_i + \sum_i^n \frac{\partial L}{\partial \dot{q}_i}\frac{d\dot{q}_i}{dt}\delta t + \underbrace{\sum_i^n \frac{\partial L}{\partial \dot{q}_i}\frac{d(\delta \bar{q}_i)}{dt}}_{1} + \underbrace{L\frac{d(\delta t)}{dt}}_{2}\, dt. \tag{12}$$

Then we can also notice that by the product rule, the bracketed terms 1 and 2 can be rewritten as

$$\underbrace{\sum_i^n \frac{\partial L}{\partial \dot{q}_i}\frac{d(\delta \bar{q}_i)}{dt}}_{1} = \frac{d}{dt}\Big(\sum_i^n \frac{\partial L}{\partial \dot{q}_i}\delta \bar{q}_i\Big) - \sum_i^n \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big). \tag{13}$$

$$\underbrace{L\frac{d(\delta t)}{dt}}_{2} = \frac{d(L\delta t)}{dt} - \frac{dL}{dt}. \tag{14}$$

Plugging these into the expression for the action differential above, moving around terms and separating the integral into three integrals, we obtain the expression

$$
\begin{aligned}
\delta I = &\int_{t_0}^{t_1} \sum_i^n \frac{\partial L}{\partial q_i}\delta \bar{q}_i - \sum_i^n \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big)\delta \bar{q}_i \, dt \\
&+ \underbrace{\int_{t_0}^{t_1} \Big[\overbrace{\sum_i^n \frac{\partial L}{\partial q_i}\frac{dq}{dt}\delta t + \sum_i^n \frac{\partial L}{\partial \dot{q}_i}\frac{d\dot{q}}{dt}}^{1} - \overbrace{\frac{dL}{dt}}^{2}\Big]\delta t \, dt}_{*} \\
&+ \int_{t_0}^{t_1} \frac{d}{dt}\Big(\sum_i^n \frac{\partial L}{\partial \dot{q}_i}\delta \bar{q}_i + \frac{d(L\delta t)}{dt}\Big)\, dt.
\end{aligned}
\tag{15}
$$

Within the starred bracket, the first and second terms are in fact equal so we arrive at the simpler expression

$$
\delta I = \int_{t_0}^{t_1} \sum_i^n \Big[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big)\Big]\delta \bar{q}_i \, dt + \int_{t_0}^{t_1} \frac{d}{dt}\Big(\sum_i^n \frac{\partial L}{\partial \dot{q}_i}\delta \bar{q}_i + L\delta t\Big)\, dt. \tag{16}
$$

Finally, substituting back in for $\delta \bar{q}_i$ using formula (10) and applying the fundamental theorem of calculus to the second integral, we get our final version of the action differential

$$
\delta I = \underbrace{\int_{t_0}^{t_1} \sum_i^n \Big[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big)\Big]\delta \bar{q}_i \, dt}_{\text{dynamics}} + \underbrace{\Big(\sum_i^n \frac{\partial L}{\partial \dot{q}_i}\delta q_i - \Big[-L + \sum_i^n \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i\Big]\delta t\Big)\Big|_{t_0}^{t_1}}_{\text{conservation laws}}. \tag{17}
$$

When the action differential is written in this form and set equal to zero as dictated by the principle of least action, it becomes of fundamental importance to physics. Equation (17) gives us key insights into the dynamics and conservation laws of physical systems. The first term, labeled "dynamics", gives rise to the dynamical equations governing physical systems. We shall discover how that is in Section 4. The second part, labeled "conservation laws", dictates how each symmetry in a physical system gives rise to a conservation law. This is due to Noether's theorem. We shall prove this and discuss its implications in Section 5. Presently, we will discover an analogous expression for the differential of the action of a field, and the derivation will follow in a series of analogous steps.

3.2. **Fields.** A physical field is a function of space and time. Fields can be scalars, vectors or tensors. Any field can be represented as a number of scalar fields which are its independent components. A few good examples of physical fields include the mass density of a fluid, the velocity vector field of a fluid, and the electric and magnetic fields. With a field we can associate a Lagrangian density $\mathfrak{L}$ which is a function of the 4 spacetime coordinates x_μ, the n scalar field components ϕ_i, and the $4n$ partial derivatives $\partial_\mu \phi_i$. Here, $x_0 = c\,t$ where t represents time and c is the speed of light. x_1, x_2, x_3 are the three spacial coordinates. The action of a field I is the integral over a region of spacetime of the Lagrangian density $\mathfrak{L}$. Our region of spacetime R is always of the form $[t_0, t_1] \times \mathfrak{R}$ where $[t_0, t_1]$ is our time interval and $\mathfrak{R}$ is our region of space. The integral of $\mathfrak{L}$ over $\mathfrak{R}$ is simply the Lagrangian L so likewise for fields as for particles, the integral of the Lagrangian over a time interval gives the action. Applying the principle of least action to the action I of a field, the field equations of the particular field which govern the field's dynamics as well as conservation laws associated with the field come right out of the derivation of the action differential. To proceed with the derivation of the differential of the action of a field, the action itself is

$$I(\phi_1(x_\mu), \ldots, \phi_n(x_\mu)) = I(\phi_i(x_\mu)) = \int_R \mathfrak{L}\Big(\phi_i(x_\mu), \partial_\mu \phi_i(x_\mu)\Big)\, dx_\mu = \int_R \mathfrak{L}\, dx_\mu. \quad (18)$$

Introducing a transformation to the spacetime and the field components as such

$$x^*_\mu = x^*_\mu(x_\mu, \phi_i, \partial_\mu \phi_i, \epsilon) \qquad \phi^*_i(x^*_\mu) = \phi^*_i(x_\mu, \phi_i, \partial_\mu \phi_i, \epsilon) \quad (19)$$

$$x_\mu = x^*_\mu(x_\mu, \phi_i, \partial_\mu \phi_i, 0) \qquad \phi_i(x_\mu) = \phi^*_i(x_\mu, \phi_i, \partial_\mu \phi_i, 0) \quad (20)$$

the transformed action functional then looks like

$$I(\phi^*_i(x^*_\mu)) = \int_{R^*} \mathfrak{L}\Big(\phi^*_i(x^*_\mu), \partial^*_\mu \phi^*_i(x^*_\mu)\Big)\, dx^*_\mu. \quad (21)$$

By Taylor's theorem, the transformed variables are simply

$$x^*_\mu = x_\mu + \frac{\partial x^*_\mu}{\partial \epsilon}\Big|_{\epsilon=0} \epsilon + o(\epsilon) \qquad \phi^*_i(x^*_\mu) = \phi_i(x_\mu) + \frac{\partial \phi^*_i(x^*_\mu)}{\partial \epsilon}\Big|_{\epsilon=0} \epsilon + o(\epsilon). \quad (22)$$

The second terms in each equation will be called δx_μ and $\delta \phi_i$ respectively. Then $\Delta x_\mu = x^*_\mu - x_\mu \sim \delta x_\mu$ and $\Delta \phi_i = \phi^*_i(x^*_\mu) - \phi_i(x_\mu) \sim \delta \phi_i$. Similarly,

$$\Delta \partial_\mu \phi_i = \partial^*_\mu \phi^*_i(x^*_\mu) - \partial_\mu \phi_i(x_\mu) \sim \delta \partial_\mu \phi_i.$$

We shall now move on to discover the differential of the action for a field as was done previously for particles. First of all, the action difference is

$$\begin{aligned}\Delta I \\ &= I(\phi^*_i(x^*_\mu)) - I(\phi_i(x_\mu)) \\ &= \int_{R^*} \mathfrak{L}\Big(\phi_i(x_\mu) + \Delta\phi_i(x_\mu), \partial_\mu \phi_i + \Delta \partial_\mu \phi_i)\, dx^*_\mu - \int_R \mathfrak{L}\Big(\phi_i(x_\mu), \partial_\mu \phi_i\Big)\, dx_\mu \quad (23)\end{aligned}$$

To get the action differential from this expression, all terms of order higher than 1 with respect to ϵ can be eliminated from the expression. First, we use Taylor's theorem to expand out the transformed Lagrangian and eliminate higher order terms. We end up with

$$\Delta I \sim \int_{R^*} \mathfrak{L} + \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} \delta\phi_i + \sum_{\mu=0}^{3} \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \delta\partial_\mu\phi_i \, dx^*_\mu - \int_R \mathfrak{L} \, dx_\mu \tag{24}$$

Then, to get everything under one integral, we transform the variables in the first integral to our original coordinates and change the region of integration back to R. We must multiply the integrand by the determinant of the Jacobian matrix in order to do this. The Jacobian matrix, denoted $\left[\frac{\partial(x^*_0,\ldots,x^*_3)}{\partial(x_0,\ldots,x_3)}\right]$, has entries of the form $\partial_\mu x^*_\gamma$. Looking at these terms more closely, we see that

$$\partial_\mu x^*_\gamma \sim \partial_\mu(x_\gamma + \delta x_\gamma) = \delta_{\mu\gamma} + \partial_\mu \delta x_\gamma. \tag{25}$$

From this, it easily follows

$$\det\left[\frac{\partial(x^*_0,\ldots,x^*_3)}{\partial(x_0,\ldots,x_3)}\right] \sim 1 + \sum_{\mu=0}^{3} \partial_\mu \delta x_\mu \tag{26}$$

Therefore, our expression now becomes

$$\Delta I \sim \int_R \left[\mathfrak{L} + \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} \delta\phi_i + \sum_{\mu=0}^{3} \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \delta\partial_\mu\phi_i\right] \det\left[\frac{\partial(x^*_0,\ldots,x^*_3)}{\partial(x_0,\ldots,x_3)}\right] - \mathfrak{L} \, dx_\mu \tag{27}$$

$$\sim \int_R \left[\mathfrak{L} + \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} \delta\phi_i + \sum_{\mu=0}^{3} \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \delta\partial_\mu\phi_i\right] \left[1 + \sum_{\mu=0}^{3} \partial_\mu \delta x_\mu\right] - \mathfrak{L} \, dx_\mu \tag{28}$$

Distributing everything out finally and eliminating more higher order terms, we arrive at an expression for the action differential itself.

$$\delta I = \int_R \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} \delta\phi_i + \sum_{\mu=0}^{3} \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \delta\partial_\mu\phi_i\Big) + \mathfrak{L} \sum_{\mu=0}^{3} \partial_\mu \delta x_\mu \, dx_\mu. \tag{29}$$

Now that we have an action differential, we can manipulate it into a form that reveals what fields will actually minimize the action and what quantities will be conserved in time for these fields. First we must make substitutions for $\delta\phi_i$ and $\delta\partial_\mu\phi_i$. We will use the notation $\Delta\bar{\phi}_i = \phi^*_i(x_\mu) - \phi_i(x_\mu)$ and its differential shall be denoted $\delta\bar{\phi}_i$. For

$\delta\phi_i$, we find that

$$\begin{aligned}
\Delta\phi_i &= \phi_i^*(x_\mu^*) - \phi_i^*(x_\mu) + \phi_i^*(x_\mu) - \phi_i(x_\mu) \qquad (30)\\
&\sim \sum_{\mu=0}^{3} \partial_\mu \phi_i^* \Delta x_\mu + \Delta\bar{\phi}_i \\
&\sim \sum_{\mu=0}^{3} \partial_\mu (\phi_i + \delta\phi_i)\delta x_\mu + \delta\bar{\phi}_i \\
&\sim \sum_{\mu=0}^{3} \partial_\mu \phi_i \delta x_\mu + \delta\bar{\phi}_i = \delta\phi_i. \qquad (31)
\end{aligned}$$

For $\delta\partial_\mu\phi_i$, we find that

$$\begin{aligned}
\Delta\partial_\mu\phi_i &= \partial_\mu^*(\phi_i^*(x_\mu^*) - \phi_i(x_\mu^*)) + \partial_\mu(\phi_i(x_\mu^*) - \phi_i(x_\mu)) + (\partial_\mu^* - \partial_\mu)\phi_i(x_\mu^*) \qquad (32)\\
&\sim \partial_\mu^*(\Delta\bar{\phi}_i(x_\mu^*)) + \partial_\mu \sum_{\gamma=0}^{3} (\partial_\gamma\phi_i)\Delta x_\gamma + \Big(\partial_\mu^* - \sum_{\gamma=0}^{3} \partial_\mu x_\gamma^* \partial_\gamma^*\Big)\phi_i(x_\mu^*) \\
&\sim \partial_\mu(\delta\bar{\phi}_i(x_\mu)) + \partial_\mu \sum_{\gamma=0}^{3} (\partial_\gamma\phi_i)\delta x_\gamma + \Big(\partial_\mu^* - \sum_{\gamma=0}^{3} (\delta_{\gamma\mu} + \partial_\mu \delta x_\gamma)\partial_\gamma^*\Big)\phi_i(x_\mu^*). \\
&\sim \partial_\mu(\delta\bar{\phi}_i(x_\mu)) + \sum_{\gamma=0}^{3} (\partial_\mu\partial_\gamma\phi_i)\delta x_\gamma - \sum_{\gamma=0}^{3} (\partial_\mu(\delta x_\gamma))\partial_\gamma^*\phi_i(x_\mu). \\
&\sim \partial_\mu(\delta\bar{\phi}_i) + + \sum_{\gamma=0}^{3} (\partial_\mu\partial_\gamma\phi_i)\delta x_\gamma \\
&= \delta\partial_\mu\phi_i.
\end{aligned}$$

Therefore, plugging our new expressions into the differential and distributing, we get

$$\begin{aligned}\delta I = \int_R \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} \sum_{\mu=0}^{3} \partial_\mu \phi_i \delta x_\mu + \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} \delta \bar{\phi}_i & \qquad (33)\\ + \underbrace{\sum_{\mu=0}^{3} \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \partial_\mu (\delta \bar{\phi}_i)}_{1} + \sum_{i=1}^{n} \sum_{\mu,\gamma=0}^{3} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} (\partial_\mu \partial_\gamma \phi_i) \delta x_\gamma & \\ + \underbrace{\mathfrak{L} \sum_{\mu=0}^{3} \partial_\mu (\delta x_\mu)}_{2} \; dx_\mu & \qquad (34)\end{aligned}$$

Now we can also make substitutions for terms 1 and 2 in this new expression when we notice that by the product rule

$$\underbrace{\sum_{\mu=0}^{3} \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \partial_\mu (\delta \bar{\phi}_i)}_{1} = \sum_{\mu=0}^{3} \sum_{i=1}^{n} \partial_\mu \Big(\frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \delta \bar{\phi}_i \Big) - \sum_{\mu=0}^{3} \sum_{i=1}^{n} \partial_\mu \Big(\frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \Big) \delta \bar{\phi}_i \tag{35}$$

$$\underbrace{\mathfrak{L} \sum_{\mu=0}^{3} \partial_\mu (\delta x_\mu)}_{2} = \sum_{\mu=0}^{3} \partial_\mu (\mathfrak{L}\, \delta x_\mu) - \sum_{\mu=0}^{3} \partial_\mu \mathfrak{L}\, \delta x_\mu \tag{36}$$

Substituting in these two expressions, redistributing the terms and breaking up the integral, we get

$$\begin{aligned}\delta I = \int_R \Big[\sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} - \sum_{\mu=0}^{3} \sum_{i=1}^{n} \partial_\mu \Big(\frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \Big) \Big] \delta \bar{\phi}_i \, dx_\mu & \qquad (37)\\ + \int_R \sum_{\mu=0}^{3} \underbrace{\Big[\overbrace{\sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} \partial_\mu \phi_i \delta x_\mu + \sum_{\gamma=0}^{3} \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\gamma \phi_i)} (\partial_\mu \partial_\gamma \phi_i) \delta x_\mu}^{1} - \overbrace{\partial_\mu \mathfrak{L}}^{2} \Big]}_{*} \delta x_\mu \, dx_\mu & \\ + \int_R \sum_{\mu=0}^{3} \partial_\mu \Big[\sum_{i=1}^{n} \Big(\frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \delta \bar{\phi}_i \Big) + \Big(\mathfrak{L}\, \delta x_\mu \Big) \Big] \, dx_\mu &\end{aligned}$$

Noticing that within the starred bracket, the first term and second term are actually equal, we find that the action differential becomes simply

$$\delta I = \int_R \Big[\sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} - \sum_{\mu=0}^{3} \sum_{i=1}^{n} \partial_\mu \Big(\frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \Big) \Big] \delta \bar{\phi}_i \, dx_\mu$$
$$+ \int_R \sum_{\mu=0}^{3} \partial_\mu \Big(\sum_{i=1}^{n} \Big(\frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \delta \bar{\phi}_i \Big) + \Big(\mathfrak{L} \, \delta x_\mu \Big) \Big) \, dx_\mu. \tag{38}$$

Finally, substituting back in for $\delta \bar{\phi}_i$ using equation (31), we arrive at our final expression for the action differential

$$\delta I = \int_R \underbrace{\Big[\sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial \phi_i} - \sum_{\mu=0}^{3} \sum_{i=1}^{n} \partial_\mu \Big(\frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \Big) \Big]}_{\text{dynamics}} \delta \bar{\phi}_i \, dx_\mu$$
$$\underbrace{+ \int_R \sum_{\mu=0}^{3} \partial_\mu \Big(\sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \delta \phi_i - \Big[-\mathfrak{L} + \sum_{i=1}^{n} \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \sum_{\mu=0}^{3} \partial_\mu \phi_i \delta x_\mu \Big] \delta x_\mu \Big) \, dx_\mu}_{\text{conservation laws}} \tag{39}$$

4. Canonical Variables, the Euler-Lagrange Equations and Dynamics in Physics

Now that we have the differential of an action for both particles and fields, the first thing we want to do is set it equal to zero and find what particle paths or fields satisfy the principle of least action. First of all, we will introduce a new set of variables, called canonical variables, which will make notation and presentation of the mathematics much clearer and give insight into the physics as well. Then, we will derive the Euler-Lagrange equations of the action for both particles and fields, which follows almost immediately from the first integral of the action differential. The Euler-Lagrange equations are the ordinary differential equations governing dynamics in the case of particles and the partial differential equations known as the field equations or wave equations in the case of fields. Finally, we will give two important examples of Euler-Lagrange equations from physics: Newton's second law of motion for classical particles and the field equations for electromagnetism.

4.1. **Particles.** At the outset, let us introduce the canonical variables associated with the Lagrangian for particles. Our current variables are $t, q_1, \ldots, q_n, \dot{q_1}, \ldots, \dot{q_n},$ and L.

Let us change to the following variables $t, q_1, \ldots, q_n, p_1, \ldots, p_n,$ and H, where we define p_i, the momentum, and H, the Hamiltonian, as follows

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{40}$$

$$H = -L + \sum_i^n p_i \dot{q}_i \tag{41}$$

Some important relations with the Hamiltonian to note are that

$$\frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} \qquad \text{and} \qquad \frac{\partial H}{\partial p_i} = \frac{dq_i}{dt} \tag{42}$$

Rewriting the action in terms of these new canonical variables and applying the principle of least action, we get

$$\delta I = \int_{t_0}^{t_1} \sum_i^n \Big[-\frac{\partial H}{\partial q_i} - \frac{dp_i}{dt} \Big] \delta \bar{q}_i \, dt + \Big(\sum_i^n p_i \delta q_i - H \delta t \Big) \Big|_{t_0}^{t_1} = 0. \tag{43}$$

Having introduced the canonical variables, let us derive the Euler-Lagrange equations. In our initial derivation of the action differential, we considered a general transformation in the paths of the particles, q_i, as well as a transformation of time itself. Let us consider the special case where we keep time fixed, $t^* = t$, and we fix the end points of the particles at time t_0 and t_1. That is to say, we consider all particles in some fixed time frame $[t_0, t_1]$ and we require that the boundary conditions of the particles, their positions at the beginning and end of the time interval be the same no matter what paths we consider them to take. This corresponds to saying that for all q_i, $q_i^*(t^*) = q_i^*(t) = q_i(t)$ for $t = t_0, t_1$. From this, it is clear that

$$\delta t = 0, \qquad \delta q_i = \delta \bar{q}_i \qquad \text{and} \qquad \delta q_i(t) \Big|_{t_0}^{t_1} = 0. \tag{44}$$

This means, that for this special case, the second integral vanishes from equation (43) and we get

$$\delta I = \sum_i^n \int_{t_0}^{t_1} \Big[-\frac{\partial H}{\partial q_i} - \frac{dp_i}{dt} \Big] \delta \bar{q}_i \, dt = 0. \tag{45}$$

which we shall call the restricted action differential. Now the term in the brackets is actually just a function of t and in order for the q_i's to be extremal values, the integral must equal zero for all admissible $\delta \bar{q}_i$. Therefore, applying Theorem 2.10, it follows that the expression in brackets must be zero for all extremal values of q_i.

Thus, we get the Euler-Lagrange equations

$$-\frac{\partial H}{\partial q_i} - \frac{dp_i}{dt} = 0 \qquad \text{or} \qquad -\frac{\partial H}{\partial q_i} = \frac{dp_i}{dt} \tag{46}$$

These are differential equations for each path $\{q_i, p_i\}$ that determine, along with initial conditions, the complete dynamics of the system. Together with the right side of equation (42), we have

$$-\frac{\partial H}{\partial q_i} = \frac{dp_i}{dt} \qquad \text{and} \qquad \frac{\partial H}{\partial p_i} = \frac{dq_i}{dt}. \tag{47}$$

These together are known as Hamilton's equations. These Euler-Lagrange equations were derived from considering which functions are extremums of the action when the end-points are fixed and time is unchanged. However, any extremum of the action must be an extremal relative to any other point in the function space. Therefore, any extremum must be an extremal relative to all functions with fixed endpoints and so any extremum of the action must in turn satisfy the condition of the restricted action differential. Equivalently, any extremum must be a solution to the Euler-Lagrange equations.

As an example of where this equation arises in physics, consider a system of n classical particles. Our Lagrangian in classical mechanics is the total potential energy U subtracted from the total kinetic energy T. If we have n particles each with mass m_i and the path of the ith particle is $(x_i(t), y_i(t), z_i(t))$, then we have

$$T = \sum_i^n \frac{1}{2} m_i ({\dot{x}_i}^2 + {\dot{y}_i}^2 + {\dot{z}_i}^2). \tag{48}$$

$$U = U(x_1, \ldots, x_n, y_1, \ldots, y_n, z_1, \ldots, z_n) = U(x_i, y_i, z_i). \tag{49}$$

$$L(x_i, y_i, z_i, \dot{x}_i, \dot{y}_i, \dot{z}_i) = T - U. \tag{50}$$

$$I(x_i, y_i, z_i) = \int_{t_0}^{t_1} L\, dt. \tag{51}$$

We shall presently convert our variables $t, x_i, y_i, z_i, \dot{x}_i, \dot{y}_i, \dot{z}_i,$ and L to canonical variables $t, x_i, y_i, z_i, p_{x_i}, p_{y_i}, p_{z_i},$ and H. The momentum variables are not to be confused with partial derivative notation. They are defined as follows

$$\begin{aligned} p_{x_i} &= m_i \dot{x}_i. \\ p_{y_i} &= m_i \dot{y}_i. \qquad \text{and} \qquad H = T + U. \\ p_{z_i} &= m_i \dot{z}_i. \end{aligned} \tag{52}$$

The new variables are nothing but the x, y and z components of the momentum of each particle and the Hamiltonian is simply the total energy of the system. By the principle of least action, the particles' paths satisfy the Euler-Lagrange equations. Before writing them out though, let us find out what the two terms in each Euler-Lagrange equation really mean.

$$-\frac{\partial H}{\partial x_i} = -\frac{\partial U}{\partial x_i} = F_{x_i}, \qquad -\frac{\partial H}{\partial y_i} = -\frac{\partial U}{\partial y_i} = F_{y_i}, \qquad -\frac{\partial H}{\partial z_i} = -\frac{\partial U}{\partial z_i} = F_{z_i} \tag{53}$$

$F_{x_i}, F_{y_i},$ and F_{z_i} are none other than the components of the force on the ith particle. The Euler-Lagrange equations are

$$-\frac{\partial H}{\partial x_i} = \frac{dp_{x_i}}{dt}, \qquad -\frac{\partial H}{\partial y_i} = \frac{dp_{y_i}}{dt}, \qquad -\frac{\partial H}{\partial z_i} = \frac{dp_{z_i}}{dt} \tag{54}$$

Equivalently they are

$$F_{x_i} = \frac{dp_{x_i}}{dt}, \qquad F_{y_i} = \frac{dp_{y_i}}{dt}, \qquad F_{z_i} = \frac{dp_{z_i}}{dt} \tag{55}$$

or simply just

$$\mathbf{F_i} = \frac{d\mathbf{p_i}}{dt}. \tag{56}$$

where $\mathbf{F_i}$ is the force vector on the ith particle and $\mathbf{p_i}$ is the ith particle's momentum. The Euler-Lagrange equations are thus equivalent to Newton's laws of motion in classical mechanics.

4.2. **Fields.** The variables for fields are $x_\mu, \phi_1(x_\mu), \ldots, \phi_n(x_\mu), \partial_0\phi_1, \partial_0\phi_2, \ldots, \partial_0\phi_n, \ldots,$ $\partial_3\phi_1, \partial_3\phi_2, \ldots, \partial_3\phi_n,$ and $\mathfrak{L}$. We must first transform these to the canonical variables for fields, which are are $x_\mu, \phi_1(x_\mu), \ldots, \phi_n(x_\mu), \Pi_0^1, \Pi_0^2, \ldots, \Pi_0^n, \ldots, \Pi_3^1, \Pi_3^2, \ldots, \Pi_3^n,$ and, H_μ^γ. We define Π_μ^i and H_μ^γ as follows

$$\Pi_\mu^i = \frac{\partial \mathfrak{L}}{\partial(\partial_\mu \phi_i)} \qquad \text{and} \qquad H_\mu^\gamma = -\mathfrak{L}\,\delta_\mu^\gamma + \sum_i^n \Pi_\gamma^i \partial_\mu \phi_i. \tag{57}$$

Now rewriting the action differential using canonical variables and applying the principle of least action, we get

$$\begin{aligned} \delta I = &\int_R \sum_{i=1}^n \Big[\frac{\partial \mathfrak{L}}{\partial \phi_i} - \sum_{\mu=0}^3 \frac{\partial}{\partial x_\mu}\Pi_\mu^i\Big]\delta\bar{\phi}_i\, dx_\mu \\ &+ \int_R \sum_{\mu=0}^3 \partial_\mu\Big(\sum_{i=1}^n \Big(\Pi_\mu^i \delta\phi_i\Big) - \sum_{\gamma=0}^3 H_\mu^\gamma \delta x_\mu\Big)\, dx_\mu = 0 \end{aligned} \tag{58}$$

To make notation simpler, let us say that

$$\epsilon\, G_\mu = \sum_{i=1}^{n} \Big(\Pi^i_\mu \delta\phi_i\Big) - \sum_{\gamma=0}^{3} H^\gamma_\mu \delta x_\mu \tag{59}$$

Then applying Green's theorem to the second integral, we get

$$\delta I = \int_R \sum_{i=1}^{n} \Big[\frac{\partial \mathfrak{L}}{\partial \phi_i} - \sum_{\mu=0}^{3} \frac{\partial}{\partial x_\mu}\Pi^i_\mu\Big] \delta\bar{\phi}_i \, dx_\mu \tag{60}$$
$$+ \int_\Gamma \sum_{\mu=0}^{3} \epsilon\, G_\mu v_\mu \, dx_\mu = 0$$

Here Γ denotes the boundary of R and v_μ denotes the outward unit normal vector to Γ. Consider the special case of a transformation such that the spacetime remains unchanged and all the scalar fields must take on the same values on the boundary Γ. These conditions are equivalent to the following:

$$\delta x_\mu = 0, \qquad \delta\phi_i = \delta\bar{\phi}_i, \qquad \text{and} \qquad \delta\phi_i(x_\mu) = 0 \quad \text{for all } x_\mu \in \Gamma. \tag{61}$$

That means $\epsilon G_\mu = 0$ and the action differential becomes just the restricted action differential which looks like this:

$$\delta I = \int_R \sum_{i=1}^{n} \Big[\frac{\partial \mathfrak{L}}{\partial \phi_i} - \sum_{\mu=0}^{3} \frac{\partial}{\partial x_\mu}\Pi^i_\mu\Big] \delta\bar{\phi}_i \, dx_\mu = 0.$$

The quantity in brackets is just a function of x_μ, and since the integral is zero for all admissible $\delta\bar{\phi}_i$, it follows from Theorem 2.11 that

$$\frac{\partial \mathfrak{L}}{\partial \phi_i} - \sum_{\mu=0}^{3} \frac{\partial}{\partial x_\mu}\Pi^i_\mu = 0. \tag{62}$$

These equations are the Euler-Lagrange equations for fields, otherwise known as the field equations for the field $(\phi_1, ..., \phi_n)$. Along with initial conditions, these equations govern the complete dynamics of the field.

4.2.1. *Example.* The most familiar field in physics is the electromagnetic field so we shall derive the field equations for electromagnetism from the principle of least action. The electromagnetic field is completely determined by the 4-potential A_μ which has 4 scalar components $A_0, A_1, A_2,$ and A_3. These scalar components of the 4-potential will be in place of our ϕ_i's of the general case. $A_0 = -\frac{V}{c}$, where V is the electric potential and $\mathbf{A} = (A_1, A_2, A_3)$ is the magnetic vector potential. The electric and magnetic fields relate to the 4-potential by the following formulas

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}. \tag{63}$$

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{64}$$

The field equations we will be deriving will be Maxwell's equations which are

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}. \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}. \tag{65}$$

$$\nabla \cdot \mathbf{B} = 0. \qquad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{66}$$

The top two equations (65) are known as Maxwell's inhomogeneous equations, and the bottom two equations (66) are Maxwell's homogeneous equations. The homogeneous equations follow directly from our definitions of the electric and magnetic fields. To see this, first we take the divergence of both sides of equation (64). Since the divergence of a curl is always zero we get the left homogeneous equation, Gauss's law for magnetism. Then, to get the second homogeneous equation we take the curl of both sides of equation (63). The curl of a gradient is always zero and so the first term on the right side of equation (63) goes away. The second term is evidently the negative time derivative of the magnetic field. The inhomogeneous equations we will derive from a the Lagrangian density but we first must introduce the electromagnetic field tensor before defining the Lagrangian density. The electromagnetic field tensor is a useful notation for consolidating information about the electromagnetic field. It is defined as follows

$$F_{\mu\gamma} = \partial_\mu A_\gamma - \partial_\gamma A_\mu. \tag{67}$$

With all of these quantities now defined, we will define the lagrangian density of the electromagnetic field as

$$\mathfrak{L} = -\frac{1}{4}\sum_{\gamma,\mu=0}^{3} F_{\mu\gamma}F_{\mu\gamma} - \mu_0 \sum_{\gamma=0}^{3} A_\gamma J_\gamma. \tag{68}$$

Our field equations will look like this:

$$\Pi_\mu^\gamma = \frac{\partial \mathfrak{L}}{\partial(\partial_\mu A_\gamma)} \qquad \frac{\partial \mathfrak{L}}{\partial A_\alpha} - \sum_{\mu=0}^{3} \partial_\beta \Pi_\beta^\alpha = 0. \tag{69}$$

Evaluating the terms in the Euler-Lagrange equation further, we realize that

$$\frac{\partial \mathfrak{L}}{\partial A_\alpha} = -\mu_0 J_\alpha \tag{70}$$

and also that

$$\Pi^\alpha_\beta = \frac{\partial}{\partial(\partial_\beta A_\alpha)}\Big(-\frac{1}{4}\sum_{\gamma,\mu=0}^{3} F_{\mu\gamma}F_{\mu\gamma}\Big) \tag{71}$$

$$= -\frac{1}{2}\sum_{\gamma,\mu=0}^{3} F_{\mu\gamma}\frac{\partial F_{\mu\gamma}}{\partial(\partial_\beta A_\alpha)} \tag{72}$$

$$= -\frac{1}{2}\Big(F_{\alpha\beta}\frac{\partial F_{\alpha\beta}}{\partial(\partial_\beta A_\alpha)} + F_{\beta\alpha}\frac{\partial F_{\beta\alpha}}{\partial(\partial_\beta A_\alpha)}\Big) \tag{73}$$

$$= -\frac{1}{2}\Big(F_{\alpha\beta}\frac{\partial F_{\alpha\beta}}{\partial(\partial_\beta A_\alpha)} + -F_{\alpha\beta}\frac{\partial(-F_{\alpha\beta}}{\partial(\partial_\beta A_\alpha)}\Big) \tag{74}$$

$$= -\Big(F_{\alpha\beta}\frac{\partial F_{\alpha\beta}}{\partial(\partial_\beta A_\alpha)}\Big) \tag{75}$$

$$= -F_{\alpha\beta}. \tag{76}$$

Thus our field equations look like this:

$$\sum_{\beta=0}^{3} \partial_\beta F_{\alpha\beta} = \mu_0 J_\alpha. \tag{77}$$

These are four scalar equations ($\beta = 0, 1, 2, 3$). When $\beta = 0$, equation (77) correspond's to Coulomb's Law, the left inhomogeneous equation. Equation (77) corresponds to the 3 scalar components of the right inhomogeneous equation, Ampere's Law, when $\beta = 1, 2, 3$. Thus, through the 4-potential definitions of the electric and magnetic fields and the Euler-Lagrange equations for the electromagnetic Lagrangian density, we arrive at Maxwell's four field equations for electromagnetism.

5. Noether's Theorem and Conservation Laws

We have seen that from the principle of least action, the first term in the action differential gives rise to the Euler-Lagrange equations which govern dynamics in physical systems. We shall see presently how the second term in the action differential gives rise to conservation laws via Noether's Theorem.

5.1. **Particles.**

Theorem 5.1.1. *Suppose we have the action functional*

$$I(q_i) = \int_{t_0}^{t_1} L(q_i(t), \dot{q}_i(t))\, dt.$$

and suppose it is invariant under the transformation in equation (2)*. If the action functional I is invariant under such a transformation, then*

$$\sum_i^n p_i \frac{\partial q_i^*}{\partial \epsilon}\Big|_{\epsilon=0} - H\frac{\partial t^*}{\partial \epsilon}\Big|_{\epsilon=0} = \textit{constant} \qquad \textit{for all } t \in [t_0, t_1]. \tag{78}$$

Proof. First of all, if I is invariant under such a transformation, then

$$I(q_i^*(t^*)) - I(q_i(t)) = 0 = \Delta I = \delta I. \tag{79}$$

In other words, if I is invariant under the transformation, then $(q_1, \ldots, q_n)$ is an extremum. This implies that

$$\delta I = \Big(\sum_i^n p_i \delta q_i - H\delta t\Big)\Big|_{t_0}^{t_1} = 0. \tag{80}$$

The first integral of the action differential from equation (17) vanishes since the q_i's satisfy the Euler-Lagrange equations. Recalling equation (4) and the definitions of the differentials δt and δq_i, we get

$$\delta I = \epsilon\Big(\sum_i^n p_i \frac{\partial q_i^*}{\partial \epsilon}\Big|_{\epsilon=0} - H\frac{\partial t^*}{\partial \epsilon}\Big|_{\epsilon=0}\Big)\Big|_{t_0}^{t_1} = 0. \tag{81}$$

Since this is true for arbitrary ϵ, it must be that

$$\Big(\sum_i^n p_i \frac{\partial q_i^*}{\partial \epsilon}\Big|_{\epsilon=0} - H\frac{\partial t^*}{\partial \epsilon}\Big|_{\epsilon=0}\Big)\Big|_{t_0}^{t_1} = 0. \tag{82}$$

It follows from this that what is in the parenthesis remains constant over time since the end points t_0 and t_1 were arbitrarily chosen. □

Any transformation that leaves the action functional I invariant is called a symmetry because if the action remains the same then the laws of physics (the Euler-Lagrange equations) remain the same in the transformed coordinates. Therefore, Noether's theorem gives an exact formula for a constant of motion, a conservation law, for every symmetry transformation. To see how truly remarkable this is, we will derive the laws of conservation of energy and momentum for our old example of classical mechanics. Our conservation laws will have the form

$$\sum_i^n \Big[p_{x_i}\frac{\partial x_i^*}{\partial \epsilon} + p_{y_i}\frac{\partial y_i^*}{\partial \epsilon} + p_{z_i}\frac{\partial z_i^*}{\partial \epsilon}\Big]\Big|_{\epsilon=0} - H\frac{\partial t^*}{\partial \epsilon}\Big|_{\epsilon=0} = \text{constant}. \tag{83}$$

Suppose we have the following two transformations (a time translation and 3 space translations)

$$t^* = t + \epsilon, \qquad x_i^*(t^*) = x_i(t), \qquad y_i^*(t^*) = y_i(t), \qquad z_i^*(t^*) = z_i(t). \tag{84}$$

$$t^* = t, \qquad x_i^*(t^*) = x_i(t) + \epsilon_x, \qquad y_i^*(t^*) = y_i(t) + \epsilon_y, \qquad z_i^*(t^*) = z_i(t) + \epsilon_z. \tag{85}$$

It is easy to see the action is invariant under both of these transformations. From the first one, we see that

$$\frac{\partial t^*}{\partial \epsilon}\bigg|_{\epsilon=0} = 1 \qquad \frac{\partial x_i^*}{\partial \epsilon}\bigg|_{\epsilon=0} = 0 \qquad \frac{\partial y_i^*}{\partial \epsilon}\bigg|_{\epsilon=0} = 0 \qquad \frac{\partial z_i^*}{\partial \epsilon}\bigg|_{\epsilon=0} = 0. \tag{86}$$

From this, we get that $H =$ constant. We have already shown that H is simply the total energy so we see that time translation symmetry gives rise to energy conservation by Noether's theorem.

Looking at the second transformation, if we set $\epsilon_y = \epsilon_z = 0$ then we get a transformation of one parameter. We find that

$$\frac{\partial t^*}{\partial \epsilon}\bigg|_{\epsilon=0} = 0 \qquad \frac{\partial x_i^*}{\partial \epsilon}\bigg|_{\epsilon=0} = 1 \qquad \frac{\partial y_i^*}{\partial \epsilon}\bigg|_{\epsilon=0} = 0 \qquad \frac{\partial z_i^*}{\partial \epsilon}\bigg|_{\epsilon=0} = 0. \tag{87}$$

From this, we get

$$\sum_i^n p_{x_i} = P_x = \text{constant}. \tag{88}$$

P_x is the total momentum in x-direction. Likewise, we discover that

$$\sum_i^n p_{y_i} = P_y = \text{constant} \qquad \sum_i^n p_{z_i} = P_z = \text{constant}. \tag{89}$$

if we let the only nonzero parameters be ϵ_y and ϵ_z respectively. Hence the vector $\mathbf{P} = (P_x, P_y, P_z)$ is a constant of motion and we see that space translational symmetry gives rise to the conservation of linear momentum in classical mechanics by Noether's theorem.

5.2. **Fields.**

Theorem 5.2.1. *Suppose we have the action functional*

$$I(\phi_i) = \int_R \mathfrak{L}(\phi_i(x_\mu), \partial_\mu \phi_i(x_\mu))\, dx_\mu.$$

and suppose it is invariant under the transformation given in equation (19)*. If the action functional I is invariant under such a transformation, then*

$$\sum_{\mu=0}^{3} \partial_\mu G_\mu = 0. \tag{90}$$

Proof. First of all, if I is invariant under such a transformation, then

$$I(\phi_i^*(x_\mu^*)) - I(\phi_i(x_\mu)) = 0 = \Delta I = \delta I. \tag{91}$$

If I is invariant under the transformation, then $(\phi_1, \ldots, \phi_n)$ is an extremum. This implies that

$$\begin{aligned} \delta I &= \int_R \sum_{\mu=0}^{3} \partial_\mu \Big[\sum_{i=1}^{n} \Pi_\mu^i \delta\phi_i - \sum_{\gamma=0}^{3} H_\mu^\gamma \delta x_\mu \Big] \, dx_\mu \\ &= \int_R \sum_{\mu=0}^{3} \partial_\mu \Big[\sum_{i=1}^{n} \Pi_\mu^i \delta\phi_i - \sum_{\gamma=0}^{3} H_\mu^\gamma \delta x_\mu \Big] \, dx_\mu = 0. \end{aligned} \tag{92}$$

since the ϕ_i's satisfy the field equations. Recalling equation (22) and the definitions of the differentials δx_μ and $\delta\phi_i$, we get

$$\begin{aligned} \delta I &= \epsilon \Big[\int_R \sum_{\mu=0}^{3} \partial_\mu \Big[\sum_{i=1}^{n} \Big(\Pi_\mu^i \frac{\partial \phi_i^*}{\partial \epsilon}\Big|_{\epsilon=0} \Big) - \sum_{\gamma=0}^{3} H_\mu^\gamma \frac{\partial x_\mu^*}{\partial \epsilon}\Big|_{\epsilon=0} \Big] \, dx_\mu \Big] \, dx_\mu \\ &= \epsilon \Big[\int_R \sum_{\mu=0}^{3} \partial_\mu G_\mu \, dx_\mu \Big] \, dx_\mu \\ &= 0. \end{aligned} \tag{93}$$

Since this is true for arbitrary ϵ, it must be the case that

$$\int_R \sum_{\mu=0}^{3} \partial_\mu G_\mu \, dx_\mu = 0 \tag{94}$$

and so it follows that

$$\sum_{\mu=0}^{3} \partial_\mu G_\mu = 0. \tag{95}$$

□

Equation (95) above, the main result of Noether's theorem, states that the divergence of G_μ is zero. From this, we will derive the conservation laws. First of all, by Green's theorem we can get that

$$\int_R \sum_{\mu=0}^{3} \partial_\mu G_\mu \, dx_\mu = \int_\Gamma \sum_{\mu=0}^{3} G_\mu v_\mu \, dx_\mu = 0. \tag{96}$$

Recalling that $R = [t_0, t_1] \times \mathfrak{R}$ where $\mathfrak{R}$ is a region of space (R^3), we know that the boundary of our region $\Gamma = \{t_0\} \times \mathfrak{R} \cup \{t_1\} \times \mathfrak{R} \cup (t_0, t_1) \times \partial\mathfrak{R}$. Here $\partial\mathfrak{R}$ is the surface of the region $\mathfrak{R}$. Now we have that

$$\int_\Gamma \sum_{\mu=0}^{3} G_\mu v_\mu \, dx_\mu = \int_{\mathfrak{R}} G_0(-1)\, dx_k \Big|_{t_0} + \underbrace{\int_{\partial\mathfrak{R}} G_\mu v_\mu \, dx_\mu}_{*} + \int_{\mathfrak{R}} G_0(1)\, dx_\mu = 0 \Big|_{t_1}. \tag{97}$$

Suppose the region $\mathfrak{R} = B(0, r)$. If we let $r \to \infty$ and require that $G_\mu = 0$ at infinity as is required for physical fields, we get that the starred term above goes to zero. Thus we arrive at the formula

$$\int_{R^3} G_0 \, dx_\mu \Big|_{t_0}^{t_1} = 0. \tag{98}$$

The same logic holds if $\mathfrak{R}$ is an arbitrary compact region of space and we expand it indefinitely. We have here an integral over all of space and since t_0 and t_1 were arbitrarily chosen, we discover that

$$\int_{R^3} G_0 \, dx_\mu = \text{constant}. \tag{99}$$

This is our equation for the conservation laws of a field. Expanded out, conservation laws are of the form

$$\int_{R^3} \Big[\sum_{i=1}^{n} \Big(\Pi_0^i \frac{\partial \phi_i^*}{\partial \epsilon}\Big|_{\epsilon=0} \Big) - \sum_{\gamma=0}^{3} H_0^\gamma \frac{\partial x_0^*}{\partial \epsilon}\Big|_{\epsilon=0} \Big] \, dx_\mu = \text{constant}. \tag{100}$$

5.2.1. *Example.* Let us consider an arbitrary field and the following sets of transformations

$$x_\mu^* = x_\mu + \epsilon_\mu \qquad \phi_i^*(x_\mu^*) = \phi_i(x_\mu). \tag{101}$$

We are concerned with the four transformations in which only one spacetime coordinate is translated and the other three remain unchanged. The first one, corresponding to ϵ_0, is a time translation. The other three, corresponding to $\epsilon_1, \epsilon_2,$ and ϵ_3, are space translations in the x_1, x_2 and x_3 directions, respectfully. Any physical field is going to have its action be invariant under such transformations. Let us see what conservation law to associate with each of these four symmetries. We can readily see that

$$\frac{\partial \phi_i^*}{\partial \epsilon_\mu}\Big|_{\epsilon_\mu=0} = 0 \qquad \frac{\partial x_\mu^*}{\partial \epsilon_\gamma}\Big|_{\epsilon_\gamma=0} = \delta_\gamma^\mu. \tag{102}$$

The constants corresponding to these transformations are

$$\int_{\mathfrak{R}} -\sum_{\gamma=0}^{3} H_0^\gamma \delta_\gamma^\mu \, dx_k = \int_{\mathfrak{R}} H_0^\mu \, dx_k = \text{constant}. \tag{103}$$

The term H_μ^γ is the Hamiltonian tensor but it is also known as the energy-momentum tensor for a field. H_0^0 corresponds to the energy density of the field and (H_0^1, H_0^2, H_0^3) corresponds to the momentum density of the field. Therefore the time translational symmetry corresponds to conservation of energy in fields and space translational symmetry corresponds to conservation of momentum just like it did in the case of particles. The 4-vector $(H_0^0, H_0^1, H_0^2, H_0^3)$ corresponds to the 4-momentum density and thus, a general translation of spacetime gives rise to the conservation of 4-momentum for fields via Noether's theorem.

5.3. **Acknowledgements.** I would like to thank my two mentors, Andrew Lawrie and Casey Rodriguez, for picking a great book for reference in writing this paper, *Calculus of Variations* by Gelfand and Fomin. They were there every week to answer any questions I had, give me pointers on presenting the material clearly and edit my drafts. I would also like to thank my girlfriend Paula Aguti for meticulously going through my paper with me to make sure the writing was clear and grammatically sound.

References

[1] Gelfand, I.M. and Fomin, S.V. Calculus of Variations. Mineola, New York, Dover Publications, Inc., 1963.

[2] Rund, Hanno The Hamilton-Jacobi theory in the calculus of variations; its role in mathematics and physics. New York, New York, Van Nostrand, 1966.