LAGRANGIAN FORMULATION OF THE ELECTROMAGNETIC FIELD

THOMAS YU

Abstract. This paper will, given some physical assumptions and experimentally verified facts, derive the equations of motion of a charged particle in an electromagnetic field and Maxwell's equations for the electromagnetic field through the use of the calculus of variations.

Contents

1. Introduction

In introductory physics classes students obtain the equations of motion of free particles through the judicious application of Newton's Laws, which agree with empirical evidence; that is, the derivation of such equations relies upon trusting that Newton's Laws hold. Similarly, one obtains Maxwell's equations from the application of Coulomb's Law, special relativity, and other ancillary laws that agree with empirical evidence. Arriving at such equations through an exploration of various laws and relationships is usually the main goal of introductory electromagnetism classes. However, with the calculus of variations, one can derive all of these equations neatly with a few physical assumptions and a single variational principle: the principle of stationary action.

With regard to modeling physical phenomena, the functionals used to describe systems in nature usually consist of an integral over time of a function called the Lagrangian. In the words of John Baez, a noted mathematical physicist, "The Lagrangian measures something we could vaguely refer to as the 'activity' or 'liveliness' of the system."[4] The arguments of the Lagrangian are those functions we are interested in for use in modeling the behavior of the system; for instance, in the modeling of the behavior of a free particle, the Lagrangian's arguments consist of the particle's position and velocity functions. When we attempt to find extrema of these functionals with respect to the arguments, in the case of the free particle, we are locating from the set of all possible position and velocity functions those particular position and velocity functions which lead to zero variation in the

Date: July 16, 2012.

functional(this will be explained later; for now, think of these particular functions as akin to critical points in differential calculus). The principle of stationary action expresses the experimental fact that systems in nature tend to favor behavior, expressed in terms of the argument functions, that are critical points of the functional describing them. In terms of John Baez's words, we can intuitively express this fact with the phrase"Nature is lazy"; in essence, nature acts to minimize its 'activity' or liveliness' as expressed in terms of the functional and the Lagrangian. Through the calculus of variations we will define these notions rigorously and apply them to electromagnetism with the aim of deriving the equations of motion in an electromagnetic field as well as Maxwell's equations.

2. Preliminaries

Definition 2.1. A functional is a function which maps functions to $\mathbb{R}$. Let $\mathcal{C}$ be a vector space of functions with norm $\|\|$. We say $\varphi[h]$, where $h \in \mathcal{C}$, is a linear functional if $\forall h \in \mathcal{C}$

$$\alpha\varphi[h] = \varphi[\alpha h] \tag{2.2}$$

$$\varphi[h_1 + h_2] = \varphi[h_1] + \varphi[h_2] \tag{2.3}$$

$$\varphi[h] \text{ is continuous with respect to the norm on } \mathcal{C} \tag{2.4}$$

Now, let $J[y]$ be a functional defined on $\mathcal{C}$. We call

$$\Delta J = J[y+h] - J[y] \tag{2.5}$$

the increment of J, where h is an increment function $\in \mathcal{C}$.

Example 2.6. Let y $\in C^1[a, b]$. The arc length of the graph of y between the points a and b is

$$\int_a^b \sqrt{1 + (y')^2}\mathrm{d}x \tag{2.7}$$

This is a functional on $C^1[a, b]$ since it takes functions as arguments and returns an associated real number, in this case the arc length.

Notation 2.8. For the sake of brevity, we will denote partial derivatives of the form $\frac{\partial f}{\partial x_i}$ as $\partial_{x_i} f$.

Remark 2.9. While there are many types of functionals, this paper will be solely concerned with functionals of the type

$$J = \int_X L(x_0, \ldots, x_n, u_1, \ldots, u_m, \partial_{x_0} u_1, \ldots, \partial_{x_n} u_1, \ldots, \partial_{x_0} u_m, \ldots, \partial_{x_n} u_m)\mathrm{d}x_0 \ldots \mathrm{d}x_n \tag{2.10}$$

Where $u_1, \ldots, u_m$ are functions of the variables $x_0, \ldots, x_n$ and X is an $n+1$ dimensional region.

The crux of the calculus of variations lies in analyzing functionals and those functions which minimize or maximize the value of the functionals through the variation of the functional. For example, equipped with the concept of variation, it is possible to find the function which gives the shortest distance in the functional above(a line as one might imagine), or finding the function for the tunnel through two given points on the Earth which minimizes transit time.

To that end, we shall now explore the concept of the variation of a functional.

Remark 2.11. We can think of the increment function h as an analogue to the h in the definition of the derivative, $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$.

If the increment of J can be expressed as

$$\Delta J = \varphi[h] + \epsilon(\|h\|) \tag{2.12}$$

where $\varphi[h]$ is a linear functional and as $\|h\| \to 0$, $\frac{\epsilon(\|h\|)}{\|h\|} \to 0$, then $\varphi[h]$ is defined to be the variation of J at y, δJ.

Remark 2.13. Analogously with functions of real variables, a functional $J[y]$ has an extremum at y if $J[y] \geq J[y*]$ or $J[y] \leq J[y*]$ for all $y*$ in some open ball(with respect to the norm of the function space) centered at y.

It is easy to show that a necessary condition for y to be an extremal of $J[y]$ is that

$$\delta J = 0 \tag{2.14}$$

for all admissible increments h about y. This is because for small h, the variation determines the sign of the increment. Therefore, the variation must be identically zero since if there were an h such that the variation was not zero, the linear properties of the variation would allow for the sign of the increment to be arbitrary.

Thus, intuitively, we are simply perturbing the functional about a certain point in its domain(a function), obtaining an expression for the change in the value of the functional due to the perturbation, then throwing away negligible terms to retain only the most significant change, the variation of the functional. Negligible means the term is of order 2 or greater with respect to $\|h\|$ or ϵ.

Construction 2.15. We will now derive the variation of functionals of the form

$$J[u] = \int_X L(x_0, \ldots, x_n, u, \partial_{x_0} u, \ldots, \partial_{x_n} u) \mathrm{d}x_0 \ldots \mathrm{d}x_n \tag{2.16}$$

The case of $u_1 \ldots u_n$ is clear from the case of just one u.

Notations 2.17. For the sake of brevity, we will denote

$$\nabla u \equiv (\partial_{x_0} u, \ldots, \partial_{x_n} u) \tag{2.18}$$

$$\overrightarrow{x} = (x_0, \ldots, x_n) \tag{2.19}$$

So that we can write (2.16) as

$$J[u] = \int\limits_R L(\overrightarrow{x}, u, \nabla u)\mathrm{d}\overrightarrow{x} \tag{2.20}$$

where R is the variable region over which the functional is defined.

We now consider a set of transformations

$$x_i^* = \Phi_i(\overrightarrow{x}, u, \nabla u; \epsilon) \tag{2.21}$$

$$u^* = \Psi(\overrightarrow{x}, u, \nabla u; \epsilon) \tag{2.22}$$

which depend on the parameter ϵ, are differentiable with respect to ϵ infinitely many times, and are equal to the identity transformation when $\epsilon = 0$. Let $\sim$ signify equality except for terms of order higher than 1 relative to ϵ; the reason for this will become clear momentarily.

We want to calculate the variation of the functional by finding the principal linear part, relative to ϵ of

$$\Delta J = J[u^*(\overrightarrow{x^*})] - J[u(\overrightarrow{x})] \tag{2.23}$$

In essence, we are perturbing the functional with the transformations of u and adding the condition of the variability of the region through tranformations of $\overrightarrow{x}$.

Calculation of δJ 2.24. By definition,

$$\begin{aligned}\Delta J &= \int_{R^*} L(\overrightarrow{x^*}, u^*, \nabla^* u^*)\mathrm{d}\overrightarrow{x^*} - \int\limits_R L(\overrightarrow{x}, u, \nabla u)\mathrm{d}\overrightarrow{x} \\ &= \int_R [L(\overrightarrow{x^*}, u^*, \nabla^* u^*)\frac{\partial(x_0^*, \dots, x_n^*)}{\partial(x_0, \dots, x_n)} - L(\overrightarrow{x}, u, \nabla u)]\mathrm{d}\overrightarrow{x} \end{aligned} \tag{2.25}$$

where

$$\frac{\partial(x_0^*, \dots, x_n^*)}{\partial(x_0, \dots, x_n)}$$

is the determinant of the Jacobian of the set of transformations $\Phi_i(\overrightarrow{x}, u, \nabla u; \epsilon)$. Note that we are using change of variables for multiple variables.

Let us now examine the transformations of x_i and u. Assuming ϵ is small, we can use Taylor's Theorem to expand each transformation about $\epsilon = 0$ so that

$$\begin{aligned} x_i^* &= \Phi_i(\overrightarrow{x}, u, \nabla u; 0) + \partial_\epsilon \Phi_i(\overrightarrow{x}, u, \nabla u; \epsilon)|_{\epsilon=0}\epsilon + O(\epsilon^2) \\ &= x_i + \epsilon\phi_i + O(\epsilon) \end{aligned} \tag{2.26}$$

$$\begin{aligned} u^* &= \Psi(\overrightarrow{x}, u, \nabla u; 0) + \partial_\epsilon \Psi(\overrightarrow{x}, u, \nabla u; \epsilon)|_{\epsilon=0}\epsilon + O(\epsilon^2) \\ &= u + \epsilon\psi + O(\epsilon) \end{aligned} \tag{2.27}$$

where $O(\epsilon^2)$ represents higher order terms and

$$\phi_i = \partial_\epsilon \Phi_i(\overrightarrow{x}, u, \nabla u; \epsilon)|_{\epsilon=0}$$
$$\psi = \partial_\epsilon \Psi_i(\overrightarrow{x}, u, \nabla u; \epsilon)|_{\epsilon=0}$$

Thus, if we eliminate the higher order terms with respect to ϵ of the transformations of x_i, then we can write the Jacobian of that set of transformations as

$$B = \begin{pmatrix} 1+\epsilon\frac{\partial\phi_0}{\partial x_0} & \epsilon\frac{\partial\phi_1}{\partial x_0} & \cdots & \epsilon\frac{\partial\phi_n}{\partial x_0} \\ \epsilon\frac{\partial\phi_0}{\partial x_1} & 1+\epsilon\frac{\partial\phi_1}{\partial x_1} & \cdots & \epsilon\frac{\partial\phi_n}{\partial x_1} \\ \vdots & \vdots & \cdots & \vdots \\ \epsilon\frac{\partial\phi_0}{\partial x_n} & \epsilon\frac{\partial\phi_1}{\partial x_n} & \cdots & 1+\epsilon\frac{\partial\phi_n}{\partial x_n} \end{pmatrix} \tag{2.28}$$

$$\begin{aligned} det(B) &\sim (1+\epsilon\frac{\partial\phi_0}{\partial x_0})\dots(1+\epsilon\frac{\partial\phi_n}{\partial x_n}) \\ &\sim 1+\epsilon\sum_{i=0}^{n}\frac{\partial\phi_i}{\partial x_i} \end{aligned} \tag{2.29}$$

The above follows from the definition of the determinant as a sum of products and discarding the higher order terms.

Thus, we can write

$$\Delta J \sim \int_R [(L(\overrightarrow{x^*}, u^*, \nabla^* u^*)(1+\epsilon\sum_{i=0}^{n}\frac{\partial\phi_i}{\partial x_i}) - L(\overrightarrow{x}, u, \nabla u)]\mathrm{d}\overrightarrow{x} \tag{2.30}$$

We now use Taylor's Theorem to expand $L(\overrightarrow{x^*}, u^*, \nabla^* u^*)$ about $(\overrightarrow{x}, u, \nabla u)$ and write

$$L^* - L \sim \sum_{i=0}^{n}(\partial_{x_i}L)\delta x_i + (\partial_u L)\delta u + \sum_{i=0}^{n}(\frac{\partial L}{\partial\frac{\partial u}{\partial x_i}})\delta(\partial_{x_i}u) \tag{2.31}$$

where $\delta x_i, \delta u, \delta(\partial_{x_i}u)$ represent the significant terms, with respect to ϵ, of $x_i^* - x, u^* - u, (\partial_{x_i^*}u^*) - (\partial_{x_i}u)$. From (2.26) and (2.27), it is clear that $\delta x_i = \epsilon\phi_i$ and $\delta u = \epsilon\psi$.

Now consider the increment $\overline{\Delta u} = u^*(x) - u(x)$ where we see the change in the two functions at the same coordinates. It can be shown from the definition of δu that

$$\delta u = \overline{\delta u} + \sum_{i=0}^{n}(\partial_{x_i}u)\delta x_i \tag{2.32}$$

$$\overline{\delta u} = \epsilon\overline{\psi} \tag{2.33}$$

where $\overline{\psi} = \psi - \sum_{i=0}^{n}(\partial_{x_i}u)\phi_i$.

Analogously, it can be shown that

$$\delta(\partial_{x_i}\overline{u}) = \delta(\partial_{x_i}u) - \sum_{k=0}^{n}(\partial_{x_i}\partial_{x_k}u)\delta x_k \tag{2.34}$$

The equations above follow intuitively due to similarities with the chain rule, but can be proved rigorously through some manipulation of the terms; for example, $\overline{\Delta u} = u^*(x) - u(x) = (u^*(x) - u^*(x^*)) + (u^*(x^*) - u(x))$. Expanding the first term around x, using (2.27) for the second term, and getting rid of negligible resulting terms, we arrive at (2.32).

Using our expressions in (2.26), (2.27), (2.33), and (2.34), we can now write (2.31) as

(2.35)
$$\delta J = \int_R \sum_{i=0}^{n} (\partial_{x_i} L)\delta x_i + (\partial_u L)\overline{\delta u} + (\partial_u L)\sum_{i=0}^{n} (\partial_{x_i} u)\delta x_i + \sum_{i=0}^{n} (\frac{\partial L}{\partial \frac{\partial u}{\partial x_i}})(\partial_{x_i}\overline{\delta u})$$
$$+ \sum_{i,k=0}^{n} (\frac{\partial L}{\partial \frac{\partial u}{\partial x_i}})(\partial_{x_k}\partial_{x_i} u)\delta x_k + L\sum_{i=0}^{n} \partial_{x_i}(\delta x_i)\mathrm{d}\overrightarrow{x}$$

We can simplify (2.35) by observing that terms within it can be replaced by $\sum_{i=0}^{n} \partial_{x_i}(L\delta x_i)$ and $\sum_{i=0}^{n}(\frac{\partial L}{\partial \frac{\partial u}{\partial x_i}})(\partial_{x_i}\overline{\delta u})$.

Thus, replacing the δ terms with their function counterparts, we finally arrive at

$$\delta J = \epsilon \int_R (\partial_u L - \sum_{i=0}^{n} \partial_{x_i}(\frac{\partial L}{\partial \frac{\partial u}{\partial x_i}}))\overline{\psi}\mathrm{d}\overrightarrow{x}$$
$$+ \epsilon \int_R \sum_{i=0}^{n} \partial_{x_i}(\frac{\partial L}{\partial \frac{\partial u}{\partial x_i}})\overline{\psi} + L\phi_i)\mathrm{d}\overrightarrow{x} \tag{2.36}$$

Remark 2.37. $\partial_u L - \sum_{i=0}^{n} \partial_{x_i}(\frac{\partial L}{\partial \frac{\partial u}{\partial x_i}}) = 0$ are called the Euler-Lagrange Equations; in functional problems with fixed endpoints, boundaries, etc. the variation of the functional reduces to the first term of (2.38). We know that in order for a function to be an extremal of a functional, the variation must be identically zero for all admissible increment functions. In this case, that would be $\overline{\psi}$. Since $\overline{\psi}$ is arbitrary, it follows that $\partial_u L - \sum_{i=0}^{n} \partial_{x_i}(\frac{\partial L}{\partial \frac{\partial u}{\partial x_i}}) = 0$ for all $\overrightarrow{x}$ if a function is to be an extremal.

The above reasoning is formalized for functions of one variable in the following lemma; note the boundary conditions on the increment functions:

Lemma 2.38. *Let $f : [a,b] \subset \mathbb{R} \to \mathbb{R}$ be a k-times continuously differentiable function. Suppose that for all $h \in C_0^k[a,b]$*

$$\int_a^b f(x)h(x)\mathrm{d}x = 0 \tag{2.39}$$

Then f(x) =0 for all $x \in$ [a,b].

Proof. Suppose there is a $c \in [a,b]$ such that $f(c) \neq 0$. Since f is continuous, there is an $\epsilon > 0$ such that f has the same sign as $f(c)$ on the interval $[c-\epsilon, c+\epsilon] \subset [a,b]$ Now let $h(x) = [x-(c-\epsilon)][(c+\epsilon)-x)$ on the interval $[c-\epsilon, c+\epsilon]$ and 0 everywhere else. Then $h(x)$ satisfies our conditions.

Then clearly

$$\int_{c-\epsilon}^{c+\epsilon} f(x)[x-(c-\epsilon)][(c+\epsilon)-x]\mathrm{d}x \neq 0 \tag{2.40}$$

This is a contradiction to our hypothesis that

$$\int_a^b f(x)h(x)\mathrm{d}x = 0 \tag{2.41}$$

for all $h \in C_0^k[a,b]$ □

Due to the ubiquitous nature of the Euler-Lagrange equations in variational problems, this lemma is called the Fundamental Lemma of the Calculus of Variations.

Remark 2.42. In order to deal with the Electromagnetic Field, we will need to work with functionals depending on $u_1, \ldots, u_m$ as in (2.10). To that end, it is easy to see that (2.36) generalizes to

(2.43)
$$\delta J = \epsilon \int_R \sum_{j=1}^m [(\partial_{u_j} L - \sum_{i=0}^n \partial_{x_i}(\frac{\partial L}{\partial \frac{\partial u_j}{\partial x_i}}))\overline{\psi_j}]\mathrm{d}\overrightarrow{x} + \epsilon \int_R \sum_{i=0}^n \partial_{x_i}(\sum_{j=1}^m \frac{\partial L}{\partial \frac{\partial u_j}{\partial x_i}}\overline{\psi_j} + L\phi_i)\mathrm{d}\overrightarrow{x}$$

where

$$\overline{\psi_j} = \psi_j - \sum_{i=0}^n (\partial_{x_i} u_j)\phi_i \tag{2.44}$$

We will end the preliminaries with a discussion of the principle of stationary action and its meaning in physical and variational terms.

Definition 2.45. The action of a physical system is a functional whose arguments are all the possible paths of the physical system.

The principle of stationary action states that a physically accurate trajectory of the physical system will be an stationary function of the functional; that is, the variation of the action at the trajectory will be 0. Note that the trajectory need not induce a relative minimum or maximum; hence stationary instead of extremal.

Example 2.46. Let us examine a system of one free particle with mass m in three dimensional space. Through some physical assumptions about space and time, we can arrive at an action for this system. See [1]:

$$J = \int_{t_0}^{t_1} T(x'(t), y'(t), z'(t)) - U(x,y,z)\mathrm{d}t \tag{2.47}$$

T corresponds to the kinetic energy of the system, and U corresponds to the potential energy of the system. An extremal of this action must satisfy the Euler-Lagrange Equations.

We can write the kinetic energy explicitly as

$$T = \frac{1}{2}m((x'(t)^2 + y'(t)^2 + z'(t)^2)$$

Our space of paths is the set of $(x(t), y(t), z(t)$ that have the same values at t_0 as well as at t_1. Note that we have three independent functions of one variable.

From (2.37) we can generate three equations that correspond to the Euler-Lagrange equations for this functional.

$$\partial_x L = \partial_t \frac{\partial L}{\partial \frac{\partial x}{\partial t}}$$
$$\partial_y L = \partial_t \frac{\partial L}{\partial \frac{\partial y}{\partial t}}$$
$$\partial_z L = \partial_t \frac{\partial L}{\partial \frac{\partial z}{\partial t}}$$

These equations correspond to

$$mx''(t) = -\partial_x U$$
$$my''(t) = -\partial_y U$$
$$mz''(t) = -\partial_z U$$

Since free particles are only under the influence of conservative forces, and conservative forces can be written as the negative of the spatial derivative of a scalar potential energy function, the above equations reduce to

$$m\mathbf{a} = \mathbf{F}$$

Which is just Newton's second law in vector form.

In the next section, we will derive the equations of motion for a particle in an electromagnetic field in more detail as a stepping stone to deriving the field equations.

The integrand of the action is called the Lagrangian of the system, L if the integration is with respect to t. If the action is an integral over all of space time(t, x, y, z), then the integrand is called the Lagrangian density, $\mathfrak{L}$. In the above example, J would be the action of the system and $T - U$ would be the Lagrangian. We note that finding Lagrangians that accurately model physical situations is not a simple process but rather one of guessing and checking using some reasonable physical assumptions.

3. Derivation of the Electromagnetic Field Equations

Remark 3.1. Due to the principle in physics that physical laws should have the same form in any inertial reference frame, physicists developed the covariant formulation of the laws of electromagnetism with tensors and 4-vectors in order to easily show how the form of the laws remained intact under Lorentz transformations(the frames of special relativity). From here on, we shall be working in $\mathbb{R}^{1+3}$, dealing with three dimensional space with the addition of time.

Notations 3.2. In the parlance of special relativity, spacetime coordinates are written in terms of (ct, x, y, z), where c is the speed of light; we will refer to these variables as $(x_0, x_1, x_2, x_3) = (x_0, \overrightarrow{x}) = \mathbf{x}$.

In addition note that both the electric Field and the magnetic Field are functions $f, g : \mathbb{R}^{1+3} \to \mathbb{R}^3$

The Electric field $\mathbf{E} = (E_1, E_2, E_3)$ and the Magnetic field $\mathbf{B} = (B_1, B_2, B_3)$ are expressed in terms of a 4-vector called the 4-Potential $\mathbf{A} = (\frac{\phi}{c}, -\overrightarrow{A}) = (A_0, A_1, A_2, A_3)$

by the following relations:

$$\mathbf{E} = -\nabla\phi - c(\partial_{x_0}\overrightarrow{A}) \tag{3.3}$$

$$\mathbf{B} = \nabla \times \overrightarrow{A} \tag{3.4}$$

ϕ is called the scalar potential, and $\overrightarrow{A}$ is called the vector potential. This formulation for some 4-potential $\mathbf{A}$ compacts several empirically derived observations.

Derivation of the Electromagnetic Field Tensor 3.5. Let us now examine the Lagrangian for a single particle with mass m and charge e in an electromagnetic field. See [3] to see how this is equivalent to T-U as in the case of a free particle.:

$$L = [\frac{1}{2}m(\partial_{x_0}\mathbf{x}) \cdot (\partial_{x_0}\mathbf{x}) + e\sum_{i=0}^{3} A_i(\partial_{x_0}x_i)] \tag{3.6}$$

In this case, the trajectory of the system consists of $\mathbf{x}$ since we are deriving the equations of motion for a charged particle in a given electromagentic field. Thus, we know that x_0, x_1, x_2, x_3 are functions of time, or $\frac{x_0}{c}$;

Let us now fit our general expression for the variation of a functional in (2.43) to our current situation:

In this case, we only have one independent variable, x_0, and 4 functions of one variable, which correspond to the u_j. Thus (2.43) reduces to

$$\begin{aligned} \delta J = \epsilon \int_{t_0}^{t_1} \sum_{i=0}^{3}((\partial_{x_i}L - \partial_{x_0}(\frac{\partial L}{\partial\frac{\partial x_i}{\partial x_0}}))\overline{\psi_i})\mathrm{d}x_0 \\ + \epsilon\int_{t_0}^{t_1} \partial_{x_0}(\sum_{i=0}^{3}\frac{\partial L}{\partial\frac{\partial x_i}{\partial x_0}}\overline{\psi_i} + L\phi)\mathrm{d}x_0 \end{aligned} \tag{3.7}$$

where we are integrating with respect to time as in (2.47). By the principle of stationary action, we know that the physically accurate trajectories of the system/equations of motion of the particle must satisfy $\delta J = 0$ for all admissible $\overline{\psi_i}$ and ϕ which are the increment functions. It follows that if a trajectory satisfies $\delta J = 0$ for all admissible increment functions, it must satisfy $\delta J = 0$ for a subset of the admissible increment functions, namely those $\overline{\psi_i}$ and ϕ such that

$$\overline{\psi_i}(t_1) = \overline{\psi_i}(t_0) = \phi(t_1) = \phi(t_0) = 0 \tag{3.8}$$

The second term of (3.7) then simplifies to 0 since

$$\epsilon\int_{t_0}^{t_1} \partial_{x_0}(\sum_{i=0}^{3}\frac{\partial L}{\partial\frac{\partial u_j}{\partial x_0}})\overline{\psi_i} + L\phi)\mathrm{d}x_0 = \epsilon(\sum_{i=0}^{3}\frac{\partial L}{\partial\frac{\partial u_i}{\partial x_0}})\overline{\psi_i} + L\phi)\Big|_{t_0}^{t_1} = 0 \tag{3.9}$$

by (3.8).

Therefore, we know that all physically accurate trajectories must satisfy

$$\epsilon\int_{t_0}^{t_1} \sum_{i=0}^{3}((\partial_{x_i}L - \partial_{x_0}(\frac{\partial L}{\partial\frac{\partial x_i}{\partial x_0}}))\overline{\psi_i})\mathrm{d}x_0 = 0 \tag{3.10}$$

for all admissible $\overline{\psi_i}$. By the Fundamental Lemma of the Calculus of Variations, it follows that our desired equations satisfy the Euler-Lagrange Equations:

$$\sum_{i=0}^{3}((\partial_{x_i} L - \partial_{x_0}(\frac{\partial L}{\partial \frac{\partial x_i}{\partial x_0}})) = 0 \tag{3.11}$$

We now substitute our Lagrangian

$$L = [\frac{1}{2}m(\partial_{x_0}\mathbf{x}) \cdot (\partial_{x_0}\mathbf{x}) + e\sum_{i=0}^{3} A_i(\partial_{x_0} x_i)]$$

for a particle in an electromagnetic field into (3.11):

Since the first term of (3.6) only contains the terms $\frac{1}{2}m(\partial_{x_0} x_i)^2$ and we know $A_i = A_i(\mathbf{x})$ we can see that

$$\partial_{x_i} L = e\sum_{j=0}^{3}(\partial_{x_i} A_j)(\partial_{x_0} x_j) \tag{3.12}$$

$$\partial_{x_0}(\frac{\partial L}{\partial \frac{\partial x_i}{\partial x_0}}) = [m(\partial_{x_0}^2 x_i) + e(\partial_{x_0} A_i)] \tag{3.13}$$

Thus from the Euler-Lagrange Equations, we can conclude that

$$\begin{aligned} m(\partial_{x_0}^2 x_i) &= e(\sum_{j=0}^{3}[(\partial_{x_i} A_j)(\partial_{x_0} x_j)] - \partial_{x_0} A_i) \\ &= e\sum_{j=0}^{3}(\partial_{x_i} A_j - \partial_{x_j} A_i)(\partial_{x_0} x_j) \end{aligned} \tag{3.14}$$

where we used that $\partial_{x_0} A_i = \sum_{j=0}^{3}(\partial_{x_j} A_i)(\partial_{x_0} x_j)$, which follows from the chain rule.

Note that the left hand side of (3.14) is a rough analogue of the net Newtonian force ($F = ma$), where $a = \partial_{x_0}^2 x_i$. We say roughly since we are in the realm of special relativity, and thus must deal with 4 vectors for force and momentum. In any case, (3.14) represents the equations of motion for a particle in an electromagnetic field in the i^{th} coordinate; for all four equations, we merely need to sum from$i = 0$ to$i = 3$ as expressed in (3.11). These equations express the Lorentz Force Law, which consists of a contribution from the electric Field and the magnetic Field, encoded through $\mathbf{A}$. Thus, we have demonstrated how variational principles can be used to derive fundamental equations of motions.

However, the particularly important aspect of this derivation was to derive

$$F_{ij} = (\partial_{x_i} A_j - \partial_{x_j} A_i) \tag{3.15}$$

Since we sum over $i, j = 0$ to$i, j = 3$, it is clear that F_{ij} represents a 4x4 matrix. More importantly, F_{ij} is called the covariant electromagnetic field tensor; its importance lies in the fact that it allows for a compact, covariant forumlation of Maxwell's equations. We will use F_{ij} in order to calculate a Lagrangian density that will allow us to obtain Maxwell's equations.

We can represent F_{ij} in matrix form using the relations between $\mathbf{A}, \mathbf{E}$, and $\mathbf{B}$ expressed in (3.3) and (3.4) as follows:

$$F_{ij} = \begin{pmatrix} 0 & \frac{E_1}{c} & \frac{E_2}{c} & \frac{E_3}{c} \\ -\frac{E_1}{c} & 0 & -B_3 & B_2 \\ -\frac{E_2}{c} & B_3 & 0 & -B_1 \\ -\frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} \tag{3.16}$$

Derivation of Maxwell's Equations 3.17. Maxwell's equations given in differential form with respect to our variables in $\mathbf{x}$ are as follows:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{3.18a}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.18b}$$

$$\nabla \times \mathbf{E} = c(\partial_{x_0}\mathbf{B}) \tag{3.18c}$$

$$\nabla \times \mathbf{B} = \mu_0 \overrightarrow{J} + \frac{\partial_{x_0}\mathbf{E}}{c} \tag{3.18d}$$

where μ_0 and ϵ_0 are physical constants, $\overrightarrow{J} = (J_1, J_2, J_3)$ is called the current density, and ρ is the volume charge density.

It is easy to see that (3.18b) and (3.18c) are direct consequences of our formulation of $\mathbf{E}$ and $\mathbf{B}$ in terms of $\mathbf{A}$; indeed, it is precisely these laws that motivated such a construction.

Using the vector calculus identities that the divergence of a curl and the curl of a gradient are both identically zero, it is easy to see that (3.3) and (3.4) imply (3.18b) and (3.18c).

Thus, we need only derive (3.18a) and (3.18d) using variational principles in order to complete a comphrehensive formulation of electromagnetism through the calculus of variations.

Definition 3.19. We define the 4 current $\mathbf{J}$

$$\mathbf{J} = (J_0, J_1, J_2, J_3) = (c\rho, \overrightarrow{J}) \tag{3.20}$$

Now, we need a suitable Lagrangian density. Since we have our covariant electromagnetic tensor F_{ij}, let us raise indices to create the contravariant tensor F^{ij}. Readers not familiar with tensors must accept that this new tensor is given by

$$F^{ij} = \begin{pmatrix} 0 & -\frac{E_1}{c} & -\frac{E_2}{c} & -\frac{E_3}{c} \\ \frac{E_1}{c} & 0 & -B_3 & B_2 \\ \frac{E_2}{c} & B_3 & 0 & -B_1 \\ \frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} \tag{3.21}$$

We now examine $F^{ij}F_{ij}$.

$$F^{ij}F_{ij} = 2(\frac{\mathbf{E}\cdot\mathbf{E}}{c^2} - \mathbf{B}\cdot\mathbf{B}) \tag{3.22}$$

This equality, easily verified from our constructions of the tensors as well as the definition of inner products for tensors, is a Lorenz invariant; that is, it is a quantity that is invariant under the Lorenz transformations of special relativity. This quality is one of the reasons this inner product will be used in the Lagrangian density for the electromagnetic field.

We now write the Lagrangian density for the electromagnetic field given a certain charge density and current density.

$$\mathfrak{L} = -\frac{1}{4\mu_0}F^{ij}F_{ij} - \sum_{i=0}^{3} A_i J_i \tag{3.23}$$

The second term of the Lagrangian density is analogous to the potential energy term of the Lagrangian of the charged particle.

In this case, our trajectory for the system are the possible values or configurations of **A** which in turn correspond to the possible configurations of **E** and **B** through their relations. Since charge and current densities are given, we are examining the four functions A_0, A_1, A_2, A_3, which are functions of $\mathbf{x}$. In addition, we are integrating over all of spacetime.

Thus, we need to generalize (2.43) to our system of four functions each relying on 4 variables, where the A_i correspond to the u_i. Changing the j index of (2.43) to match our current relativistic notation, we have that the variation of our functional is

$$\delta J = \epsilon \int_R \sum_{j=0}^{3}((\partial_{A_j}\mathfrak{L} - \sum_{i=0}^{3}\partial_{x_i}(\frac{\partial \mathfrak{L}}{\partial \frac{\partial A_j}{\partial x_i}}))\overline{\psi_j})\mathrm{d}\mathbf{x} + \epsilon \int_R \sum_{i=0}^{3}\partial_{x_i}(\sum_{j=0}^{3}\frac{\partial \mathfrak{L}}{\partial \frac{\partial A_j}{\partial x_i}}\overline{\psi_j} + \mathfrak{L}L\phi_i)\mathrm{d}\mathbf{x} \tag{3.24}$$

where R is all of spacetime.

Notice that the second term of (3.24) is an integral of a gradient; by the generalized Stoke's theorem,

$$\epsilon \int_R \sum_{i=0}^{3}\partial_{x_i}(\sum_{j=0}^{3}\frac{\partial \mathfrak{L}}{\partial \frac{\partial A_j}{\partial x_i}}\overline{\psi_j} + L\phi_i)\mathrm{d}\mathbf{x} = \epsilon \int_{\partial R}\sum_{j=0}^{3}\frac{\partial \mathfrak{L}}{\partial \frac{\partial A_j}{\partial x_i}}\overline{\psi_j} + \mathfrak{L}\phi_i \mathrm{d}x^* \tag{3.25}$$

The second term is now an integral over ∂R, which is the boundary of all of spacetime. We now invoke a physical principle that the magnitude of the electromagnetic field goes to zero sufficiently "fast" as one moves towards the boundary of spacetime(infinity) so that the second term of (3.24) is essentially zero

With respect to the first term, we invoke the same arguments as in the derivation of the electromagnetic tensor so that we may conclude that any physically meaningful configuration of the electromagnetic field must be such that the Euler-Lagrange equations for this system hold.

Thus, we conclude that

$$\sum_{j=0}^{3}(\partial_{A_j}\mathfrak{L} - \sum_{i=0}^{3}\partial_{x_i}(\frac{\partial \mathfrak{L}}{\partial \frac{\partial A_j}{\partial x_i}})) = 0 \tag{3.26}$$

Calculating $F^{ij}F_{ij}$ explicity in terms of **A** using our relations for **E**, **B**, and **A**, we get that:

$$\begin{aligned} F^{ij}F_{ij} = 2[&(\partial_{x_1}A_0 - \partial_{x_0}A_1)^2 + (\partial_{x_2}A_0 - \partial_{x_0}A_2)^2 + (\partial_{x_3}A_0 - \partial_{x_0}A_3)^2 \\ &- (\partial_{x_2}A_3 - \partial_{x_3}A_2)^2 - (\partial_{x_3}A_1 - \partial_{x_1}A_3)^2 - (\partial_{x_1}A_2 - \partial_{x_2}A_1)^2] \end{aligned} \tag{3.27}$$

Calculating the $j = 0$ term yields

$$\partial_{A_0}\mathfrak{L} = -J_0 \tag{3.28}$$

and

$$\begin{aligned}
\sum_{i=0}^{3}\partial_{x_i}\Big(\frac{\partial\mathfrak{L}}{\partial\frac{\partial A_0}{\partial x_i}}\Big) &= -\frac{1}{\mu_0}[0+\partial_{x_1}(\partial_{x_1}A_0-\partial_{x_0}A_1)+\partial_{x_2}(\partial_{x_2}A_0-\partial_{x_0}A_2)+\partial_{x_3}(\partial_{x_3}A_0-\partial_{x_0}A_3)]\\
&= -\frac{1}{\mu_0}(\partial_{x_0}(0)+\partial_{x_1}(\frac{E_{x_1}}{c})+\partial_{x_2}(\frac{E_{x_2}}{c})+\partial_{x_3}(\frac{E_{x_3}}{c})\\
&= -\frac{1}{\mu_0}(\partial_{x_0}(F^{00})+\partial_{x_1}(F^{10})+\partial_{x_2}(F^{20})+\partial_{x_3}(F^{30})\\
&= -\frac{1}{\mu_0}\sum_{k=0}^{3}\partial_{x_k}F^{k0}
\end{aligned} \tag{3.29}$$

which follows from our constructions of F^{ij} and F_{ij}.

Thus, substituting (3.28) and (3.29) into (3.26), we can see that the first of the four Euler Lagrange Equations of the system corresponds to

$$\sum_{k=0}^{3}\partial_{x_k}F^{k0} = \mu_0 J_0 \tag{3.30}$$

For the $j=1$ term of (2.26) we have

$$\partial_{A_1}\mathfrak{L} = -J_1 \tag{3.31}$$

$$\begin{aligned}
\sum_{i=0}^{3}\partial_{x_i}\Big(\frac{\partial\mathfrak{L}}{\partial\frac{\partial A_1}{\partial x_i}}\Big) &= -\frac{1}{\mu_0}[-\partial_{x_0}(\partial_{x_1}A_0-\partial_{x_0}A_1)+\partial_{x_1}0+\partial_{x_2}(\partial_{x_1}A_2-\partial_{x_2}A_1)-\partial_{x_3}(\partial_{x_3}A_1-\partial_{x_1}A_3)]\\
&= -\frac{1}{\mu_0}[\partial_{x_0}(\partial_{x_0}A_1-\partial_{x_1}A_0)+\partial_{x_1}0+\partial_{x_2}(\partial_{x_1}A_2-\partial_{x_2}A_1)+\partial_{x_3}(\partial_{x_1}A_3-\partial_{x_3}A_1)]\\
&= -\frac{1}{\mu_0}[\partial_{x_0}(\frac{-E_{x_1}}{c})+\partial_{x_1}0+\partial_{x_2}(B_{x_3})+\partial_{x_3}(-B_{x_2})]\\
&= -\frac{1}{\mu_0}[\partial_{x_0}(F^{01})+\partial_{x_1}(F^{11}+\partial_{x_2}(F^{21})+\partial_{x_3}(F^{31})]\\
&= -\frac{1}{\mu_0}\sum_{k=0}^{3}\partial_{x_k}F^{k1}
\end{aligned} \tag{3.32}$$

Thus, our second Euler-Lagrange Equation reduces to

$$\sum_{k=0}^{3}\partial_{x_k}F^{k1} = \mu_0 J_1 \tag{3.33}$$

It is now easy to see that we can generalize the entire set of Euler-Lagrange equations to

$$\sum_{k=0}^{3} \partial_{x_k} F^{ki} = \mu_0 J_i \tag{3.34}$$

for $i = 0, 1, 2, 3$.

We only need one more physical fact before we derive (3.18a) and (3.18d), completing our derivation of Maxwell's Equations from variational principles. The physical constants ϵ_0, μ_0, and c are related in the following way:

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \tag{3.35}$$

Let us now derive (3.18a). From the definition of the 4 current $\mathbf{J}$, $J_0 = c\rho$.

$$\begin{aligned}\sum_{k=0}^{3} \partial_{x_k} F^{k0} &= \partial_{x_0}(0) + \partial_{x_1}(\frac{E_{x_1}}{c})) + \partial_{x_2}(\frac{E_{x_2}}{c}) + \partial_{x_3}(\frac{E_{x_1}}{c}) \\ &= \frac{\nabla \cdot \mathbf{E}}{c} \end{aligned}\tag{3.36}$$

Thus, it follows from (3.35) and (3.36) that

$$\begin{aligned}\nabla \cdot \mathbf{E} &= c^2 \mu_0 \rho \\ &= \frac{\mu_0}{\epsilon_0 \mu_0} \rho \\ &= \frac{\rho}{\epsilon_0}\end{aligned}\tag{3.37}$$

just as in (3.18a).

(3.18d) comes from examining the $i = 1, 2, 3$ terms of (3.34):

$$\begin{aligned}\partial_{x_0}(-\frac{E_{x_1}}{c})) + \partial_{x_1}(0) + \partial_{x_2}(B_{x_3}) + \partial_{x_3}(-B_{x_2}) &= \mu_0 J_1 \\ \partial_{x_0}(-\frac{E_{x_2}}{c})) + \partial_{x_1}(-B_{x_3}) + \partial_{x_2}(0) + \partial_{x_3}(B_{x_1}) &= \mu_0 J_2 \\ \partial_{x_0}(-\frac{E_{x_3}}{c})) + \partial_{x_1}(B_{x_2}) + \partial_{x_2}(-B_{x_1}) + \partial_{x_3}(B_{x_1}) &= \mu_0 J_3\end{aligned}$$

Shifting the $\mathbf{E}$ terms to the right side, we obtain

$$\begin{aligned}[\partial_{x_2}(B_{x_3}) - \partial_{x_3}(B_{x_2})] &= \partial_{x_0}(\frac{E_{x_1}}{c})) + \mu_0 J_1 \\ [\partial_{x_3}(B_{x_1}) - \partial_{x_1}(B_{x_3})] &= \partial_{x_0}(\frac{E_{x_2}}{c})) + \mu_0 J_2 \\ [\partial_{x_1}(B_{x_2}) - \partial_{x_2}(B_{x_1})] &= \partial_{x_0}(\frac{E_{x_3}}{c})) + \mu_0 J_3\end{aligned}$$

this set of three equations is equivalent to the vector equation

$$\nabla \times \mathbf{B} = \mu_0 \vec{J} + \frac{\partial_{x_0} \mathbf{E}}{c} \tag{3.38}$$

which is exactly (3.18d).

4. Concluding Remarks

Using the calculus of variations, we developed the necessary mathematical tools such as the general variation of a relevant functional as well as the Euler Lagrange equations for a variational construction of modeling physical phenomena. We then applied these tools specifically to electromagnetism in the context of special relativity, deriving the equations of motion in an electromagnetic field as well as Maxwell's Equations.

Acknowledgments. I would like to thank my mentors Casey Rodriguez and Andrew Lawrie for reviewing the drafts of this paper, recommending several helpful and interesting texts, and, most importantly, for their invaluable assistance and patience while I struggled to grasp the nuances of the mathematical material of this paper.

References

[1] L.D. Landau, E.M. Lifschitz, *Mechanics.* Butterworth-Heinemann, Oxford, 3rd Edition, 1993.
[2] L.D. Landau, E.M. Lifschitz, *The Classical Theory of Fields.* Butterworth-Heinemann, Oxford, 3rd Edition, 1993.
[3] Herbert Goldstein, *Classical Mechanics.* Addison-Wesley Press, MA, 1st Edition, 1950.
[4] John C. Baez, *Lectures on Classical Mechanics.* 2005.
[5] I.M. Gel'fand, S.V. Fomin, *Calculus of Variations.* Prentice-Hall, NJ, 1963