HAMILTON FORMALISM AND VARIATIONAL PRINCIPLE CONSTRUCTION

P. VÁN AND B. NYÍRI

Abstract. It is widely accepted that a variational principle cannot be constructed for an arbitrary differential equation; a rigorous mathematical condition shows which equations can have a variational formulation. On the other hand, the importance of variational principles in various fields of physics resulted in several methods to circumvent this condition and to construct another type of variational principles for any differential equation. In this paper the common origin of the considered methods is investigated, and a generalized Hamiltonian formalism is formulated. Additionally, constructive algorithms are given by different methods to construct variational principles. Simple examples are presented to make construction methods more transparent: several Lagrangians are constructed for the different forms of the Maxwell equations and for the extended heat conduction equation.

1. Introduction

The existence of variational principles for different physical theories is a great mystery in physics. A variational formulation of a physical problem can be useful because we can use special solution methods, we can have a machinery to handle symmetries and a deeper insight into the structure of a theory. Several books and papers suggest the view that a variational formulation of a problem is an important achievement in a physical theory, because the variational principles present a simple and straightforward method to summarize a whole theory in a single formula.

On the other hand it is well known that the possibility of formulating a physical theory by a variational principle is a restriction; we cannot give a variational form for an arbitrary differential equation. A strict mathematical theorem tells us the condition of the existence of a variational principle for a given differential (or almost any kind of) equation (see for example in [**?**, **?**]). The application of that theorem shows clearly that a lot of important equations in physics do not have a variational formulation.

For example if a differential equation contains a first order time derivative then we cannot construct a variational principle without any further ado. The parabolic transport equations of nonequilibrium thermodynamics or the Maxwell equations are good examples here. Fortunately we know well that for the Maxwell equations the introduction of new variables (the scalar and the vector potential) can help in some cases. Several other ideas and tricks exist to circumvent the mathematical condition. Especially the field of nonequilibrium thermodynamics is a parade ground of the different methods. A collection of them is given for example by Ichiyanagi [**?**] or by Ván and Muschik [**?**]. The second paper concentrates on the

Date: August 31, 2001.

mathematical structure of these variational principles and methods and proposes the following classification (Fig. 1):

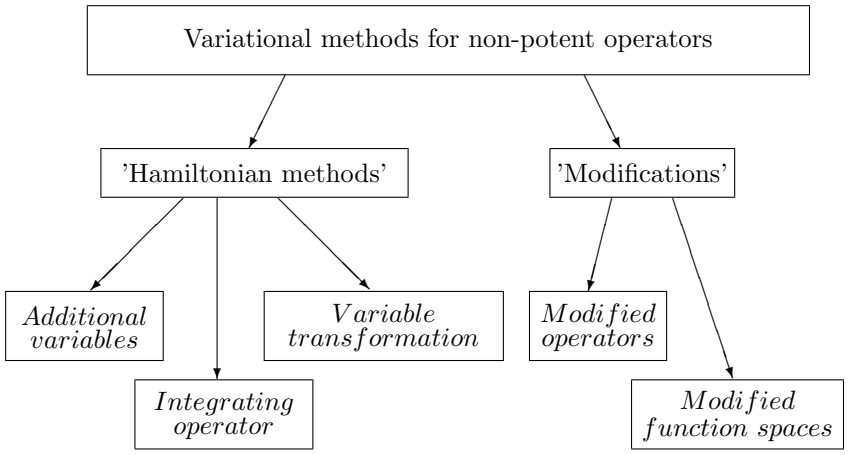

Figure 1

There are two main construction schemes:

- The *Hamiltonian methods* result in 'hamiltonian' or 'true' variational principles for an equation that is (almost) equivalent with the original one. They are called 'true' because they result in a Lagrangian from where we can get Euler-Lagrange equations applying the standard methods without any further mathematical or physical trick.
- The *Modifications* construct variational principles for a modified equation which, instead of being equivalent, is closely related to the original one. Therefore the principles coming from these methods are sometimes called 'quasi-variational' ones.

Methods from both groups can be useful tools in diverse physical theories.

In this paper we will show that all of the 'Hamiltonian methods': the method of *additional variable*, *integrating operator* and *variable transformations* have a common origin, they can be considered as particular cases of a generalized Hamiltonian formalism.

In the second section we summarize the essence of 'true' variational procedures and for the sake of better understanding we use a special form of the starting operator (and equation). In the third section we will show the common origin of these methods. The fourth and fifth section emphasizes some other aspects of the common origin. The sixth section gives a construction algorithm. The seventh contains an example from a field of physics where some variational principles are well know but the original equations contain first order time derivatives; we present different variational principles for the Maxwell equations. The eights section contains different variational formulations for the extended heat conduction equation, because this equation means always a challenge for any construction method. In the last section there is a discussion.

2. Construction methods

A convenient mathematical frame of our investigations can be if we suppose that the basic variables of our starting equation and of the following variational principles are elements of a Banach space. This is a restriction, but enables a concise mathematical formulation and we can concentrate on the physical outcome.

Here and in the following $\mathcal{B}$ denotes a real Banach space, $\mathcal{B}^*$ is its dual (the space of continuous linear functionals on $\mathcal{B}$); the bilinear map of duality will be denoted by $\mathcal{B}^* \times \mathcal{B} \to \mathbb{R}, (p, x) \mapsto \langle p, x \rangle$.

Let $\Theta : \mathcal{B} \rightarrowtail \mathcal{B}^*$ be a continuous mapping. A variational principle shall be formulated for the equation

$$(\varphi \in \mathcal{B})? \qquad \Theta(\varphi) = 0.$$

Let us recall some necessary mathematical tools:

Definition 2.1. Let U be an open subset of $\mathcal{B}$. A mapping $\Theta : U \to \mathcal{B}^*$ is called **potent** on U, if there exists a differentiable function $S : U \to \mathbb{R}$ such, that $DS(\varphi) = \Theta(\varphi)$ for every $\varphi \in U$. Then S is called the **potential** of Θ.

Here DS denotes the strong or Fréchet derivative of the functional S. If $\mathcal{B}$ is a function space than we call a potential **variational**.

Theorem 2.1. *(Helmholtz-Volterra-Vainberg) If $U \subset \mathcal{B}$ is starlike, the function $\Theta : U \to \mathcal{B}$ is continuously differentiable and $D\Theta(\varphi) \in Lin(\mathcal{B}, \mathcal{B}^*)$ is symmetric for all $\varphi \in U$, then Θ is potent on U.*

This theorem can be considered as a generalization of the well know fact, that a 'force field' is 'conservative' if and only if its curl vanishes, that is a differentiable $\mathbb{R}^3 \rightarrowtail \mathbb{R}^3$ function has a potential if and only if its derivative is symmetric on an appropriate set.

The theorem (2.1) give us a method to decide whether we can formulate a variational principle for a given equation or not. We should take the operator in the equation and if its derivative is symmetric then there is a hope to find a variational principle for the equation.

This is only a hope, because a variational principle in physics is an intuitive notion from several different points of view:

- The equation itself is a source of uncertainty. We consider two equations equivalent if they have the same solutions. From that point of view two operators with the same kernel are equivalent, too.
- One usually does not specify exactly the corresponding function spaces, the domain and the range of the operator in the equation. If the corresponding equation is a partial differential equation then the particular problem fixes the boundary or/and initial conditions, therefore the proper domain and range of the operator.
- The topology on the function space, the exact meaning of the differentiation is undefined, too. If we have a Hilbert space instead of a Banach space, the situation becomes simpler. Some considerations of the mathematical meaning and the correspondence between the different derivatives and variations, the role of the Hilbert space, etc. is given in [**?**]. A lot of other versions of the theorem (2.1) are possible.

The construction methods do not care about the lack of a rigorous mathematical formulation. If we are able to accomplish the derivations, and construct the corresponding variational potential, then according to the known and well elaborated examples we can strongly hope that this potential will be the core of any future stronger mathematical treatment. In the following we avoid the usage of very much mathematics, and restrict ourselves to the level of exactness accepted in the treatment of variational principles in physics.

In this section we suppose that our starting equation for which we want to give a variational formulation is given in the next simple form:

$$\Theta(\varphi) = L_i f_i(\varphi) = 0 \tag{2.1}$$

where $\varphi \in \mathcal{B}$, $n \in \mathbb{N}$ and f_i: $\mathcal{B} \rightarrowtail \mathcal{B}$ are functions, $L_i : \mathcal{B} \rightarrowtail \mathcal{B}^*$ are linear operators. Here and in the following we use the convention of Einstein for summation with double indexes. Let us remark that we do not distinguish in the notation the functions defined on function spaces, and the functions on real numbers. In this paper the functions deserve more attention therefore we usually give their domain.

We will see in the next sections, that this form of the equation is general enough to construct variational principles for a lot of important differential operators, and its special form gives some serious advantage in the practical applications.

Now let us suppose, that Θ does not have a symmetric derivative, therefore we need some tricks to get a variational principle. Three basic tricks are known:

2.1. **Method of variable transformation.** [**?**, **?**]

In this method we exploit that from a physical point of view two equations are equivalent if they have the same solutions.

We introduce a new variable $\xi \in \mathcal{B}$ into our equation and after that we try to eliminate our original variable φ considering some special conditions: we require that our original variable be a function of the new independent variable ξ, $\varphi = \varphi(\xi)$, and we want $\Theta \circ \varphi$ to be a potent operator. The following considerations give us ideas to find an appropriate φ. We are looking for a real valued function S of the two variables φ and ξ such that the variational potential have the form $\tilde{S}(\xi) := S(\xi, \varphi(\xi))$.

Moreover, we require that

$$\partial_1 S(\xi, \varphi) = L_i f_i(\varphi) \tag{2.2}$$

$$\partial_2 S(\xi, \varphi(\xi)) = 0, \tag{2.3}$$

for all possible φ and ξ. Here ∂_i means a partial (strong) differentiation according to the ith variable of the function. The general solution of the first simple "partial differential equation" gives that:

$$S(\xi, \varphi) = \langle L_i^* \xi, f_i(\varphi) \rangle - F(\varphi). \tag{2.4}$$

where F is an arbitrary (differentiable) function, L_i^* denotes the transpose of L_i. In this case equation (2.3) determines $\varphi(\xi)$ by the implicit relation

$$\langle L_i^* \xi, Df_i(\varphi) \rangle - DF(\varphi) = 0. \tag{2.5}$$

If we are able to determine φ from (2.5) (with an appropriate choice of F), then

$$\tilde{S}(\xi) = S(\xi, \varphi(\xi)) = \langle L_i^* \xi, f_i(\varphi(\xi)) \rangle - F(\varphi(\xi))$$

is a variational potential of a "hamiltonian" variational principle for ξ and if we require the stationarity (vanishing derivatives) we get the following equation:

$$L_i f_i(\varphi(\xi)) = 0,$$

which can be equivalent to (2.1) if φ is injective.

The solution of (2.5) can be simplified if the elements of our Banach space $\mathcal{B}$ are functions of n real variables, the functions on $\mathcal{B}$ are given by composition, i.e. (by a little abuse of notations) $f_i(\varphi) = f_i \circ \varphi$ and the variational potential can be given in the following integral form

$$S(\xi, \varphi) = \int (L_i^* \xi(x) \cdot f_i(\varphi(x)) - \mathcal{H}(\varphi(x)))\, dx, \tag{2.6}$$

where $\cdot$ denotes the scalar product in $\mathbb{R}^n$ and $\int \mathcal{H}(\varphi(x))dx = F(\varphi)$. We can introduce the Lagrangian

$$\mathcal{L}(\xi, \varphi) = L_i^* \xi \cdot f_i(\varphi) - \mathcal{H}(\varphi) \tag{2.7}$$

This special case is general enough to treat almost all differential equations of classical physics, where the physical quantities are represented by functions of several real numbers. Usually one introduces a Lagrangian l in a narrower sense than $\mathcal{L}$:

$$\mathcal{L}(\xi, \varphi) = l \circ (L_1^* \xi, ..., L_n^* \xi, \varphi).$$

This composition form is important to preserve the form of the Euler-Lagrange equations, but it would be too cramped for our further purposes.

In the following we omit the symbol $\circ$ of composition and we will call the left hand side of the Euler-Lagrange equations as *'variation'* and *'partial variation'* of the Lagrangians, denoting them by a "δ". The partial variation of the Lagrangian (2.7) by φ results in the special form of the condition (2.5)

$$\delta_\varphi \mathcal{L}(\xi, \varphi) = L_i^* \xi \cdot D f_i(\varphi) - \delta \mathcal{H}(\varphi) = 0, \tag{2.8}$$

for determining φ. A 'variation' is not a derivation because the elementary algebraic properties of a derivation (Leibnitz rule) can be violated.

We emphasize again that here we applied mathematical formulas in a loose sense; their exact meaning would require a bit of more investigations.

A simple example helps us to grasp the meaning behind the formulas:

Example 1. If our equation for which a variational principle to be constructed is:

$$x \in C_2([t_1, t_2], \mathbb{R})? \qquad \ddot{x} + \dot{x} = 0, \tag{2.9}$$

then the appropriate form, how a variational potential is looked for is

$$S_1(\xi, x) = \int_{t_1}^{t_2} \mathcal{L}_1(\xi, x) dt = \int_{t_1}^{t_2} (x(\dddot{\xi} - \dot{\xi}) - h_1(x)) dt. \tag{2.10}$$

Here $h_1 : \mathbb{R} \rightarrowtail \mathbb{R}$ differentiable function. Now the condition (2.8) to determine $x(\xi)$ can be written as

$$\delta_x \mathcal{L}_1(x, \xi) = \dddot{\xi} - \dot{\xi} - h_1'(x) = 0$$

If we choose $h_1(x) := \frac{x^2}{2}$ then a solution follows immediately: $x(\xi) = \dddot{\xi} - \dot{\xi}$.

Substituting into (2.10), we get the following Lagrangian as a result of the construction

$$\tilde{\mathcal{L}}_1(\xi) = \tilde{l}_1(\dot{\xi}, \ddot{\xi}) = \frac{(\dddot{\xi} - \dot{\xi})^2}{2}.$$

Now the corresponding Euler-Lagrange equation is:

$$\xi^{(4)} - \ddot{\xi} = \ddot{x}(\xi) - \dot{x}(\xi) = 0$$

Let us remark that here a derivation and a partial integration results together the 'variation' of the Lagrangian. Here we applied the usual condition that we do not vary the final and the initial time, more properly: $Dom\tilde{\mathcal{L}}_1 \subset C_2([t_1, t_2], \mathbb{R}|\xi(t_1) = \xi_1, \xi(t_2) = \xi_2, \xi_1, \xi_2 \in \mathbb{R})$.

We can observe here a Lagrangian with second time derivatives. Here this does not hurt the causality, because the Euler-Lagrange equation does not contain a third order time derivative. This is because of the special form of the Lagrangian.

2.2. **Method of integrating operator.** [**?**]

In this case we follow a pattern similar to the method of variable transformation. We introduce a new variable again, denoted by u, but now we try to eliminate this new variable, not the original one, in the form $u = u(\varphi)$. The variational potential has the shape $S(u(\varphi), \varphi)$. Then we require that:

$$\partial_1 S(u, \varphi) = L_i f_i(\varphi), \tag{2.11}$$

$$\partial_2 S(u(\varphi), \varphi) = 0. \tag{2.12}$$

The general solution of the first simple "partial differential equation" gives the same form as in the previous method:

$$S(u, \varphi) = \langle L_i^* u, f_i(\varphi)\rangle - F(\varphi). \tag{2.13}$$

Here F is an arbitrary function as before. In this case equation (2.12) is used to determine $u(\varphi)$:

$$\langle u, L_i D f_i(\varphi)\rangle - DF(\varphi) = 0. \tag{2.14}$$

If we are able to determine u from this equation (F should be chosen to make it easy), we can get a 'hamiltonian' variational principle with the variational potential

$$\tilde{S}(\varphi) = S(u(\varphi), \varphi) = \langle L_i^* u(\varphi), f_i(\varphi)\rangle - F(\varphi),$$

and the following condition of the extremum (as Euler-Lagrange equation):

$$\langle Du(\varphi), L_i f_i(\varphi)\rangle = 0.$$

This can be equivalent with (2.1) under some conditions. However, the solution of (2.14) is far from easy in most of the particular cases.

Example 2. An important example of the application of this procedure is the *method of least squares*. There we choose our arbitrary function F in the following manner:

$$F_2(\varphi) = \frac{1}{2}\langle L_i f_i(\varphi), L_i f_i(\varphi)\rangle.$$

In this case the condition (2.14) gives that $u(\varphi) = L_i f_i(\varphi)$ and the corresponding variational potential is:

$$\tilde{S}_2(\varphi) = \frac{1}{2}\langle L_i f_i(\varphi), L_i f_i(\varphi)\rangle,$$

and the Euler-Lagrange equation

$$L_i^* L_j f_j(\varphi) = 0.$$

You can observe that now the integrating operator is L_i^*, and the corresponding operator in the equation above is symmetric, if $L_i = L_i^{**}$. (For example if the domain of the L_i-s is finite dimensional.) Otherwise other tricks are necessary. ■

Another example can be the differential equation (2.9) in the previous method:

Example 3. Now we are looking for a variational potential in the form:

$$S_3(u,x) = \int_{t_1}^{t_2} \mathcal{L}_3(u,x)dt = \int_{t_1}^{t_2} (u(\ddot{x}+\dot{x}) - h_3(x))dt \tag{2.15}$$

The condition to determine $u(x)$ is

$$\delta_x \mathcal{L}_3(u,x) = \ddot{u} - \dot{u} - Dh_3(x) = 0$$

If we take $h_3(x) := \frac{(\ddot{x}+\dot{x})^2}{2}$ then we can get easily $u(x) = \ddot{x}+\dot{x}$. The corresponding Lagrangian

$$\tilde{\mathcal{L}}_3(x) = \frac{(\ddot{x}+\dot{x})^2}{2},$$

and the Euler-Lagrange equation

$$x^{(4)} - \ddot{x} = (\frac{d^2}{dt^2} - \frac{d}{dt})(\ddot{x}+\dot{x}) = 0.$$

You can recognize that here $\frac{d^2}{dt^2} - \frac{d}{dt}$ is the 'integrating operator'. ■

Of course the 'integrating operator' can be a multiplication with a simple elementary function, an 'integrating multiplier', too.

Example 4. Let us consider the previous differential equation (2.9), and the variational potential given in the previous example (2.15). Now let $h_4(x) = \frac{e^t\dot{x}^2}{2}$. In this case the equation to determine u is

$$\delta_x \mathcal{L}_4(u,x) = \ddot{u} - \dot{u} - e^t(\ddot{x}+\dot{x}) = 0.$$

An immediate solution of this equation is $u(x) = e^t x$. Therefore the Lagrangian from (2.15) is

$$\tilde{\mathcal{L}}_4(x) = e^t x(\ddot{x}+\dot{x}) - \frac{e^t\dot{x}^2}{2},$$

and the corresponding Euler-Lagrange equation:

$$e^t(\ddot{x}+\dot{x}) = 0.$$

The integrating multiplier is e^t. ■

As we emphasized before some difficulties can arise when we leave the formal ground. A stronger mathematical treatment of Tonti [**?**] can give us a feeling on the mathematical problems and their importance in practical applications of this method.

2.3. **Method of additional variables.** [**?**]

Here, unlike the previous two cases the introduced additional variables are not eliminated, and we construct a variational potential with the help of them. The additional variables are usually elements of the dual of our original vector space. A possible variational potential for (2.1) is:

$$S(\varphi, \varphi^*) = \langle \varphi^*, L_i f_i(\varphi) \rangle + F(\varphi). \tag{2.16}$$

Another possible form to get similar Euler-Lagrange equations can be:

$$S(\varphi, \varphi^*) = \frac{1}{2} \left(\langle \varphi^*, L_i f_i(\varphi) \rangle + \langle L_i^* \varphi^*, f_i(\varphi) \rangle \right), \tag{2.17}$$

where $\varphi^* \in \mathcal{B}^*$. Differentiating by φ^* we get back the original equation, differentiating by φ we get the "dual" equation that is thrown away in most of the applications [**?**, **?**, **?**]. Because of that both previous forms are equally acceptable, furthermore a simple multiplication of the original equation (the first term in the potentials above) gives the same result. Sometimes other essentially equivalent variational potentials are used putting together the two terms and taking into account the properties of the linear operators L_i and L_i^*.

Example 5. This method is so trivial as it si seem to be. A variational potential for the equation (2.9) is

$$S_5(x, x^*) = \int_{t_1}^{t^2} x^*(\ddot{x} + \dot{x}) dt.$$

The corresponding Euler-Lagrange equations are

$$\ddot{x} + \dot{x} = 0 \qquad \ddot{x^*} - \dot{x^*} = 0,$$

the first one is the original equation and we do not need the second one. ■

In spite of its simplicity this method is extensively used in particle physics. The equations of motion for free quantum mechanical particles are first order differential equations (Schrödinger, Dirac equations, etc..). A Lagrangian of a quantum field theory usually contains these equations multiplied by the complex conjugate of the corresponding quantum field (complex Hilbert spaces are considered). However, there the adjoint equation sometimes has a physical meaning, because there the field itself is not a direct measurable quantity, we need to construct the corresponding probability densities.

3. Operator Hamiltonians

Let us observe the similar patterns: every method introduces new variables, one of the variables is eliminated in the first two ones. We eliminate the newly introduced variable in the method of *integrating multipliers* and the original variable in the method of *variable transformations*. Moreover, the assumptions on the variational potential had similar structure in the two cases. We will see that the background of these similarities is a kind of Hamiltonian formalism. But before a more systematic comparison let us repeat here the basic structure of the Hamiltonian formalism of classical mechanics of masspoints. We do not give a systematic treatment but try to emphasize some important issues from our point of view.

3.0.1. *Mechanical Hamiltonians.* The Lagrangian of the Hamilton equations of a mechanical system with the coordinate $x \in \mathbb{R}^k$ and momentum $p \in \mathbb{R}^k$ ($k \in \mathbb{N}$) is

$$L(x,p) = l(x,\dot{x},p) = \dot{x}\cdot p - H(x,p), \tag{3.1}$$

where $H : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ is the Hamiltonian. The Hamilton equations follow after partial variations of the Lagrangian as Euler-Lagrange equations:

$$\delta_p L(x,p) = \dot{x} - \frac{\partial H}{\partial p}(x,p) = 0, \tag{3.2}$$

$$\delta_x L(x,p) = -\dot{p} - \frac{\partial H}{\partial x}(x,p) = 0. \tag{3.3}$$

If we can eliminate p from the equation (3.2) and can substitute the resulted $p(x,\dot{x})$ function into (3.3), we can get the 'original' Newtonian equation of motion.

$$\dot{p}(x,\dot{x}) - \partial_1 H(x,p(x,\dot{x})) = 0. \tag{3.4}$$

Putting the previous expression of momentum into the function (3.1) we get the following Lagrangian:

$$\tilde{L}(x) = \tilde{l}(x,\dot{x}) = \dot{x}p(x,\dot{x}) - H(x,p(x,\dot{x})), \tag{3.5}$$

This Lagrangian differs from (3.1) in its variables, but the form of the function is the same $\tilde{L}(x) = L(x,p(x,\dot{x}))$. The corresponding Euler-Lagrange equation of $\tilde{L}$ is (3.4). Remember all the four different forms of the Lagrangian $\tilde{L}, L, \tilde{l}, l$. The functional form is similar, but the domain, therefore the function itself, is different.

Based on the classical formulation above we modify a little bit the previous scheme to fit our further purposes.

3.0.2. *Modified mechanical Hamiltonians.* Let us modify (3.1) in the following manner

$$L_m(x,\pi) = l(x,\dot{x},\pi) = \dot{x}\cdot f(\pi) - H_m(\pi), \tag{3.6}$$

where $H_m : \mathbb{R}^k \to \mathbb{R}$ and $f : \mathbb{R}^k \to \mathbb{R}^k$. Now the Euler-Lagrange equations give the Hamilton equations

$$\delta_\pi L_m(x,\pi) = \dot{x}\cdot Df(\pi) - DH_m(\pi) = 0, \tag{3.7}$$

$$\delta_x L_m(x,\pi) = -Df(\pi)\cdot\dot{\pi} = 0, \tag{3.8}$$

and the Lagrangian for the Newton equation is

$$\tilde{L}_m(x) = \tilde{l}_m(x,\dot{x}) = \dot{x}\cdot f(\pi(x,\dot{x})) - H_m(\pi(x,\dot{x})). \tag{3.9}$$

Let us remark, that if we use the traditional definition of the momentum, we get

$$p = \frac{\partial \tilde{l}_m(x,\dot{x},p)}{\partial \dot{x}} = f(\pi),$$

therefore the whole modification can be transformed into a 'normal', 'mechanical' formalism, where the Hamiltonian does not depend on x.

$$L(x,p) = \dot{x}\cdot p - H(p).$$

Here $H(f(\pi)) = H_m(\pi)$.

■

After these simple examples let us turn back our attention to the construction methods.

3.1. **Variable transformation.** Considering the equation (2.1) with $n = 1$ we can observe the similarity of the Lagrangians (2.7) and (3.6).

$$\mathcal{L}(\xi, \varphi) = L^*\xi \cdot f(\varphi) - \mathcal{H}(\varphi) \quad \leftrightarrow \quad L_m(x, \pi) = \left(-\frac{d}{dt}x\right)(-f(\pi)) - H_m(\pi),$$

the linear operator L corresponds to the time derivative, and the operator $\mathcal{H}$ to H_m the Hamiltonian in (3.6). Moreover we can see that the similarity is more than formal, (3.6) is a special case of (2.7) Furthermore we can get the following 'operator momentum'

$$p := \frac{\partial l(\varphi, \xi, L^*\xi)}{\partial L^*\xi} = f(\varphi).$$

Now it is easy to recognize in the equations of the construction scheme a Hamiltonian like formalism and transform them into the same form. Let us use the correspondence with the second, modified mechanical system. Here ξ corresponds to x, φ to π, and our arbitrary function $\langle(\varphi)$ is an 'operator Hamiltonian'. The Hamilton equations for the Lagrange function (2.7) are

$$\begin{aligned} \delta_\xi \mathcal{L}(\xi, \varphi) &= L_i f_i(\varphi) = 0 \\ \delta_\varphi \mathcal{L}(\xi, \varphi) &= L_i^*\xi \cdot Df_i(\varphi) - \delta\mathcal{H}(\varphi) = 0 \end{aligned}$$

This simple observation shows us that in the variational principle construction method *variable transformations* our starting equation (2.1) for which a variational principle needs to be constructed can be taken as the first Hamilton equation (arising from the variation of the appropriate Lagrangian by the 'generalized coordinate' ξ) in a Hamiltonian formalism. Therefore in this method we try to construct the Lagrangian formalism from a 'half' Hamiltonian one. From the second equation, arising from the variation by $\pi \equiv \varphi$ we determine how the 'momentum' depends on the 'coordinate' ξ (if it is possible), and we can get back (if we are clever) the Lagrangian of a 'true' variational problem, with the newly introduced variable ξ.

There is one important difference between this generalized Hamiltonian formalism and the one in the classical mechanics: there is no formal way from the Lagrangian formalism to the Hamiltonian one: we cannot determine our momentum(s) $(f_i(\varphi))$ from the Lagrangian without the knowledge of the Hamiltonian. The reason is that the role of the 'time derivation' is not fixed and eventually it is based on the special properties of a generalized Legendre transformation. But before sinking in the technical details let us turn our attention to the other two construction schemes.

3.2. **Integrating operator.** In this case we can recognize similar analogies than in the previous method. The Lagrangians in the method of variable transformation and in the method of integrating multipliers are very similar. Therefore we can see that φ corresponds to π, the operator Hamiltonian is $\mathcal{H}$ and the Lagrangian of the Hamilton equations is

$$\mathcal{L}(u, \varphi) = \sum_{i=1}^{n} L_i^* u \cdot f_i(\varphi) - \mathcal{H}(\varphi). \tag{3.10}$$

The Hamilton equations themselves are

$$\begin{aligned} \delta_u \mathcal{L}(u, \varphi) &= L_i f_i(\varphi) = 0, \\ \delta_\varphi \mathcal{L}(u, \varphi) &= L_i^* u \cdot Df_i(\varphi) - \delta\mathcal{H}(\varphi) = 0. \end{aligned}$$

Here u plays the role of a generalized coordinate. This method does not have a mechanical counterpart, here we should solve the second Hamilton equation for the 'coordinate' u to get the function $u(\varphi)$ and put it into the Lagrangian (not into the second Hamilton equation). In mechanics it would mean that we eliminate the coordinate from the two Hamilton equations to get one unified equation for the momentum.

For the modified mechanical Hamilton equations (3.7)-(3.8) with the application of this method we should get a function $\dot{x}(\pi)$ from (3.7) and put it into the Lagrangian (3.6). The resulted function is the 'real' Lagrangian of the method

$$\tilde{L}_m(\pi) = \dot{x}(\pi)f(\pi) - H_m(\pi).$$

The corresponding Euler-Lagrange equation is:

$$\delta\tilde{L}_m(\pi) = -Dx(\pi)Df(\pi)\dot{\pi} = 0.$$

The equation is different from (3.8) as it is expected in this method; an integrating operator, $Dx(\pi)$ appeared.

3.3. **Additional variables.** Here the analogy is not too strict, we can call any of the variables as a momentum, and our original equation can be any of the Euler-Lagrange equations, only the even number of the variables gives the feeling of a Hamiltonian structure. Let us remember, here none of the variables are eliminated, but we use only one of the final (operator) Hamilton equations.

4. Legendre transformation

We can get a better understanding of the procedure if we investigate the transformation properties of the variables. The root of the method of variable transformation and integrating operator is a generalized Legendre transformation. Generalized in the sense that we are not in a finite dimensional vector space as usual but in an infinite dimensional one. In finite dimension the Legendre transformation is defined with a restriction in the space of variables $(x_1, x_2) \in \mathbb{R}^k \times \mathbb{R}^k$ $(k \in \mathbb{N})$ by the next system of formulas:

$$\begin{aligned} f_1(x_1) + f_2(x_2) &= x_1 \cdot x_2, \\ Df_1(x_1) = x_2, &\qquad Df_2(x_2) = x_1 \end{aligned}$$

where $f_1, f_2 : \mathbb{R}^k \rightarrowtail \mathbb{R}$ are two times differentiable functions and possibly $D^2 f_1 \neq 0$ and $D^2 f_2 \neq 0$. These relation results that the x_1, x_2 variables and the f_1, f_2 functions cannot be independent. x_1 and x_2 are called the Legendre transformed forms of each other, so do the functions f_1 and f_2. If f_1 and/or f_2 are not two times differentiable or $D^2 f_1 = 0$ or/and $D^2 f_2 = 0$ somewhere in their domain then the usual properties of the Legendre transformation can be violated, for example cannot be involutive anymore (e.g. in this case we cannot apply the implicit function theorem). This formulation shows well the symmetric character of the transformation.

In classical mechanics the Legendre transformation is partial only, the Hamiltonian is the Legendre transformation of the Lagrangian in the time derivative of

the coordinates:

$$\tilde{l}(x,\dot{x}) + h(x,p) = p\cdot\dot{x}, \tag{4.1}$$

$$\frac{\partial\tilde{l}}{\partial\dot{x}}(x,\dot{x}) = p \tag{4.2}$$

$$\frac{\partial h}{\partial p}(x,p) = \dot{x}, \tag{4.3}$$

where $\tilde{l}, h : \mathbb{R}^k \times \mathbb{R}^k \rightarrowtail \mathbb{R}$, and according to our formulas the variables cannot be independent we should take either $\dot{x}(x,p)$ or $p(x,\dot{x})$. The transformation of variables results the Legendre transformed form of the corresponding functions. For example $l(x,p) = p\dot{x}(x,p) - h(x,p)$ is the Legendre transformed form of $\tilde{l}(x,\dot{x})$.

We should not mix the different forms of the Lagrangian:

$$\tilde{\mathcal{L}}(x) = \mathcal{L}(x,p(x,\dot{x})) = \tilde{l}(x,\dot{x}) = l(x,\dot{x},p(x,\dot{x})) = p(x,\dot{x})\dot{x} - h(x,p(x,\dot{x})).$$

The functions given on function spaces are denoted by capitals, and the ones on the space of coordinates by tilde (they are independent of the momentums). Hamilton equations are the consequence of the variational properties of the Lagrangian $l(x,\dot{x},p) = p\dot{x} - h(x,p)$. The variation of the Lagrangian $\tilde{l}(x,\dot{x})$ gives the Newton equation.

In the construction methods we can recognize a generalized Legendre transformation. Let us start from Lagrangian (2.7). In this case we can give a Legendre transformation by the next formulas

$$\tilde{\mathcal{L}}(\xi) + \mathcal{H}(\varphi) = L_i^*\xi\cdot f_i(\varphi), \tag{4.4}$$

$$\delta_\xi\tilde{\mathcal{L}}(\xi) = L_i f_i(\varphi), \tag{4.5}$$

$$\delta_\varphi\mathcal{H}(\varphi) = L_i^*\xi\cdot Df_i(\varphi), \tag{4.6}$$

You can observe that here the derivations are replaced by variations (see 2.1). We get an easy and more 'mathematical' application if we apply the Legendre transformation to the variational potential instead of the Lagrangian. In that case the derivations are normal or 'strong' derivations. However, this definition is more transparent from a 'physical' point of view, but in this case the change of the variables requires a more profound investigation, the two times 'variability' does not give any condition.

Now it is easy to see that the Lagrangian $\tilde{\mathcal{L}}$ alone does not determine a momentum: we should know the linear operators L_i (and the functions f_i), too. Moreover, here, contrary to the mechanics, the parallel usage of the derivation-variation does not appear: only variation is used for the change of variables and also to get the equations. Therefore there are only two kinds of Lagrangians: $\tilde{\mathcal{L}}$ and $\mathcal{L}$.

Here the notation of the method of variable transformations has been used, if we replace the variable ξ by u then the same can be said on the structures in the integrating multiplier method.

5. A scheme of construction

The previous observations on the Legendre transformations enables us to give a simple way of the potential construction with variable transformation.

Our starting equation is again (2.1). Let us search the Lagrangian of the variational potential as a function of the new variable ξ in the next form:

$$\tilde{\mathcal{L}}(\xi) = \tilde{l}(L_i^*\xi).$$

If we suppose that this Lagrange function was given by (traditional) Legendre transformation, we can get the following relations

$$D\tilde{l}(L_i^*\xi) = \partial_i\tilde{l}(L_i^*\xi) = f_i(\varphi). \tag{5.1}$$

Here $\tilde{l} : \mathbb{R}^k \rightarrowtail \mathbb{R}$ and $\tilde{\mathcal{L}} : \mathcal{B}_1 \to \mathcal{B}_2$. These relations are necessary to give back the original (2.1), as you can check by the variation of the Lagrange function

$$\delta_\xi\tilde{l}(L_i^*\xi) = L_i\partial_i\tilde{l}(L_i^*\xi)$$

Now we do not need to solve the equation (2.8), the condition of contact, we can construct the solution immediately. (That is actually a way of the solution.) The variables of the Lagrangian l are not independent any more, the relations (5.1) results $n-1$ constraints. If the functions f_i are invertible these constraints are given by $\varphi = f_1^{-1}(\partial_1\tilde{l}(L_1^*\xi)) = ... = f_n^{-1}(\partial_n\tilde{l}(L_n^*\xi))$.

Example 6. Let us see again our simple example equation (2.9). In this case we are looking for a Lagrangian in the form $\tilde{l}(-\dot{\xi}, \ddot{\xi})$. The variation by ξ gives that

$$\frac{d\partial_1\tilde{l}}{dt} + \frac{d^2\partial_2\tilde{l}}{dt^2} = 0.$$

This can be the original example equation (2.9) if $\partial_1\tilde{l} = \partial_2\tilde{l} = x$. Applying this constraint to $\tilde{l}$ we get

$$\tilde{l}(-\dot{\xi}, \ddot{\xi}) = \tilde{l}(-\dot{\xi} + \ddot{\xi}).$$

The form of the function $x(\xi)$ depends on the choice of the otherwise arbitrary $\tilde{l}$:

$$x(\xi) = (D\tilde{l})(\ddot{\xi} - \dot{\xi}),$$

where the dash denotes the differentiation. If $\tilde{l} = id^2_{\mathbb{R}}/2$ we can get back the Lagrangian in the example 1.

■

Remark 5.1. *Nonlinear equations* The practical applicability of the variable transformation method is not so easy as you can think after the previous considerations. The special form of the starting equation means a serious advantage in the construction.

Remark 5.2. *Method of integrating multipliers* The similar structure of the equations give us a feeling that an appropriate Legendre transformation can result a direct procedure to give the Lagrangian as in the method of variable transformation. However, in this case the final equation is unknown, therefore the interpretation of the different terms in the Lagrangian (3.6) does not yield a similar simplification. But we usually do not need it at all, because the solution of (2.5) is usually more easy for u (or ξ in the variable transformation method) than for φ.

6. General considerations

Let us see shortly what can be said in the general case when our starting equation is $\Theta(\varphi) = 0$.

Now a general form of a variational potential is [**?**]:

$$S(\xi, \varphi) = \langle \xi, \Theta(\varphi) \rangle - F(\varphi), \tag{6.1}$$

for both methods. If we want the derivatives of the function S to be zero according to its different variables we get

$$\partial_\xi S(\xi, \varphi) = 0 \quad \rightarrow \quad \Theta(\varphi) = 0, \tag{6.2}$$

$$\partial_\varphi S(\xi, \varphi) = 0 \quad \rightarrow \quad \langle \xi, D\Theta(\varphi) \rangle - DF(\varphi) = 0. \tag{6.3}$$

In the method of variable transformations we want to get a function $\varphi(\xi)$ from (6.3). If we have it then our variational potential will be:

$$\tilde{S}_{vt}(\xi) = S(\xi, \varphi(\xi)) = \langle \xi, \Theta(\varphi(\xi)) \rangle - F(\varphi(\xi)).$$

The Euler-Lagrange equation for this potential is (6.2).

In the method of integrating multipliers we want to get a function $\xi(\varphi)$ from (6.3). If we have it then our variational potential will be:

$$\tilde{S}_{im}(\varphi) = S(\xi(\varphi), \varphi) = \langle \xi(\varphi), \Theta(\varphi) \rangle - F(\varphi).$$

In this case we can get the Euler-Lagrange equation from

$$\langle D\xi(\varphi), \Theta(\varphi) \rangle = 0,$$

according to the derivation of S_{im}.

In both cases it is easy to recognize a generalized Legendre transformation in the background, where one of the variables is ξ and the other variable, the corresponding 'generalized momentum' is $\Theta(\varphi)$. In the special 'almost linear' case treated in the previous sections we can find a geometric meaning of the procedures: the envelope of hyperplanes is connected to the Legendre transformation.

7. Variational principles for the Maxwell equations

In this section we give example applications of the variable transformation method. We will show how the choice of the vector and scalar potentials and the choice of a gauge in the Maxwell equations can help to transform them into a system of equations that have a variational formulation.

First of all we remark, that there are a lot of different forms and interpretations of the Maxwell equations in the different fields of physics. For example we know well that the selection of the independent variables depends on the problem considered, we can use the equations to determine the sources (charge density and the electric current) from the fields, or the fields from the sources. Moreover we know well how the problem of the interaction of the matter and the electromagnetic field in classical electromagnetism is confused and poorly understood [**?**, **?**, **?**]. Furthermore, even if we are dealing with a problem where the sources are given and the fields are the independent variables, we can find different forms of the equations, depending on the considered polarization phenomena. In this case we know that the Maxwell equations are first order non-potent equations in their original form. Therefore traditional variational principles introduces the scalar and vector potentials and usually do not consider any polarization. The variational principle that appears in this field is the best example for the (unintentional) use of the method of variable

transformations. However, the method can be extended for more general equations to consider different polarization effects [**?**].

In the following we will construct variational principles for three different forms of the Maxwell equations with the help of variable transformation method. First it is done for the equations with sources with linear polarization, to see the relation between the 'potentials' introduced for variational purposes and the traditional scalar and vector potentials. The second example gives variational principles for the whole system of Maxwell equations with linear polarization. In the last example we construct variational principles for the Maxwell equations without any polarization to show that the construction method does not require a closed or soluble system of equations.

7.1. **Traditional potentials and variational principles.** In this case we consider the next system of equations for which a variational principle is to be constructed:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon}, \tag{7.1}$$

$$\nabla \times \mathbf{B} = \mu\mathbf{j} + \epsilon\mu\frac{\partial \mathbf{E}}{\partial t}. \tag{7.2}$$

Here the electric field strength $\mathbf{E}$ and magnetic induction $\mathbf{B}$ are considered as function $\mathbb{R}\times\mathbb{R}^3 \rightarrowtail \mathbb{R}^3$, where spacetime is represented by $\mathbb{R}\times\mathbb{R}^3$. t denotes the time, ϵ and μ are the caonstant scalar electric permittivity and the magnetic permeability respectively. The charge density ρ and the current density $\mathbf{j}$ are spacetime functions. We remark that here and in the following all of the equations are given by an inertial observer [**?**, **?**].

A little rearrangement of the terms is necessary to get the traditional potentials directly:

$$-\nabla \cdot (\epsilon\mathbf{E}) + \rho = 0, \tag{7.3}$$

$$\nabla \times \left(\frac{\mathbf{B}}{\mu}\right) - \mathbf{j} - \frac{\partial \epsilon\mathbf{E}}{\partial t} = 0. \tag{7.4}$$

Let us recall that only the solutions of the Maxwell equations are interesting from a physical point of view. This solution does not change, with the previous transformation but the operator of the equation will be different if we change the sign or multiply the equation with a number!

Now we apply the method of variable transformations according to the procedure described in section 5 . We introduce a scalar valued function ϕ for the equation (7.3) and a vector valued function $\mathbf{A}$ for the equation (7.4). After the recognition of the different linear operators we are looking for a Lagrangian in the form

$$\tilde{F}(\phi, \mathbf{A}) = \tilde{f}(\nabla\phi, \phi, \nabla \times \mathbf{A}, \mathbf{A}, \frac{\partial \mathbf{A}}{\partial t}), \tag{7.5}$$

where $\tilde{F}$ is a multivariable function on the appropriate domain. Let us remember that here the different variables were introduced for the different terms in the equations (7.3)-(7.4) according to the basic rules given below. The following adjoint differential operators are used: $(-\nabla\cdot)^* = \nabla$, $(-\frac{\partial}{\partial t})^* = \frac{\partial}{\partial t}$ and $(\nabla\times)^* = \nabla\times$. Now