Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

On Noether's theorems and gauge theories in hamiltonian formulation

Relatore:
Prof. Fiorenzo Bastianelli

Presentata da:
Guglielmo Coloretti

Anno Accademico 2018/2019

Abstract

Nella tesi presente si propone una trattazione esaustiva sui teoremi di Noether, cardine delle più moderne ed avanzate teorie di gauge. In particolare si tenta di fornirne una misura matematica rigorosa senza allontanarsi dalla cruciale intuizione fisica che celano: la ricerca di simmetrie nella natura e la volontà di descrivere le interazioni conosciute con un singolo modello. Più avanti, trovando i caratteri dominanti e l'ispirazione nelle pubblicazioni di Noether, si affrontano i tratti generali della formulazione hamiltoniana delle teorie di gauge, presentando la struttura dell'azione per una particella relativistica, la teoria elettromagnetica e la teoria della relatività generale; si pongono infine alcuni interrogativi sui valori di contorno che emergono dal formalismo adottato. Inoltre, per ottenere un'esposizione più efficace e meno oscura, si accompagna ogni risultato con esempi opportuni.

Contents

Chapter 1

Preamble

There must in fact be some substance, one or more, from which the rest is generated while remaining unchanged.

Aristotle

1.1 Overview

In classical mechanics, field theory and in a certain way also in quantum mechanics we appeal to an object which encodes all the information about the system we are facing: the action integral. For the simple case of a particle, this is a functional depending on the path of the particle, defined as [1]

$$I[q(t)] = \int L(q(t), \dot{q}(t), t)\, dt. \tag{1.1}$$

We will deal with infinitesimal transformations that leave the action unchanged, which means we will treat such transformations which make the variation of the action vanish (at least up to a boundary term [2]). We also recall the method from which we obtain the equations of motion, namely the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0. \tag{1.2}$$

[1] Since the very beginning, let us be clear with notation. Whenever a function is indicated as $q(t)$ we refer to all the set of its components $q^i(t)$, while when we write $q^i(t)$ we obviously indicate its single component in tensorial formulation. Moreover we also point out that we use from now on the Einstein convention for the summation on repeated index. Finally let us remind that, since $q(t)$ represents the set of lagrangian generalized coordinates, each component of it stands for a different generalized coordinates of a particle.

[2] In this case of a single particle, the integral of a total time derivative.

It consists in performing a variation on the action with fixed end points at which the variation does not occur, and requiring the action to be an extreme value (usually a minimum, the "least action principle").

Combining these particular infinitesimal transformations that does not change the action and the conditions which lead to the equations of motion we deduce Noether's first theorem, which can be read quite naively as follow: *to every such transformations is associated a certain quantity that stands still on the path of motion.*

Although we will later discuss and furnish a far better explanation of this result it might be interesting wondering on the physical meaning of it. If in the description of nature we find a conserved quantity, we might say that we have found a constant, something that we can follow as a lighthouse during our searching. This reminds to the very first studying of the realm surrounding us, which took place in the ancient Greek world when the first query was towards uncovering the common root of our manifold nature, a feature that holds its own identity whether everything was swirling or not. Leaving these fascinating and perhaps less physical interpretations aside, we cannot avoid to recognize the powerful approach of lagrangian formalism and Noether's theorems, which provide us with an elegant weapon we can tackle our studying with. The deepest conquer of her theorems lays in establishing a connection between symmetries and conservation laws that elevates the latter beyond useful empirical regularities. The influence of this insight is pervasive in physics, also because it contains the seeds of gauge theories. Indeed there is another theorem established by the genius of Emmy Noether. This is the so called second theorem and it has been abandoned for a long time or at least no one has been given credits to Noehter for it or understood the great importance of her works. This theorem allows a more complete vision of interactions in nature: it finds a connection between equations that describe the same phenomena but looks mathematically completely different and, in a certain way, arbitrary. Indeed, under the correct conditions, the Lagrangian produces equations that connect the dynamical fields and thus reduce the degree of freedom of the system considered. Further this equations depend from arbitrary functions, hence the different manner to obtain a correct interpretations of the same phenomenon. This, as mentioned before, gives birth to gauge theories which may be the most powerful instrument in physics.

The first chapter is dedicated to Noether's theorems. We present the concept of symmetries and on-shell variations, providing some examples to better understand them. We then report Noether's first theorem with an intuitive approach and then extend its treatment to field theory. Once we have introduced all the necessary mathematical tools, we present again Noether's first theorem but with a more rigorous formalism. To extend rigid symmetries to arbitrary ones, we focus for a while on Lie derivatives in order to introduce the topic of the next section: Noether's second theorem, reported and demonstrated in the very original version. We then give some application of the two theorems inspecting the energy momentum tensor and some action integrals. In the second chapter we deal with gauge symmetries, cast them in hamiltonian form and then use this formalism to analyze three lagrangians: the one for special relativity, the one for electromagnetism and the ADM method for general relativity. Finally in the last chapter we give a brief overview on all the boundary terms we neglected in our computations. We also add three appendix to give reason and better explain the mathematical models adopted.

Before the very beginning of the topic, let us say a word on Emmy Noether.

1.2 Emmy Noether: a brief biography

Since it is passed little more than a century from the publication of Noether's theorem, we dedicate few lines to her biography. Amalie Emmy Noether was born on the 23rd of March, 1882 in Erlangen, a University town a bit north of Nürnberg. She was called by her middle name because both her mother and grandmother were also named Amalie. Emmy's father, Max Noether, was a professor of Mathematics of some distinction at the university of Enlargen: he was a member of the Academies of Berlin, Göttingen, Munich, Budapest, Copenaghen, Turin, Accademia dei Lincei, Institut de France and the London Mathematical Society.

Emmy Noether attended a school which should provide a preparation for the life of a lady in which, if you had a profession at all, it would be teaching English and French to other young ladies. After completing her first cicle of studies, she could not be accepted in the University of Erlangen, since ladies were not allowed. It was possible, however, to apply for special permission to listen to lectures. She also took private lessons in Mathematics, preparing for universities studies. In 1903 she passed the university qualification but she could not be admitted to the University of Erlangen. The University of Gottingen was a little bit more open-minded: she went there for a semester, during which she heard lectures from Minkowski, Klein, Hilbert and Schwarzschild. After one semester, Erlangen saw the error of its ways and began admitting women, precisely two out of a class of about a thousand, and so she was able to enroll at the University of Erlangen as a student of mathematics. She finally managed to obtain her Ph.D and it appears that Noether was the second woman Ph.D in Europe, following Sofia Kovalevskaya who received her Ph.D in Göttingen. During all her life she taught mathematics, especially algebra, group theory and differential equation, in Göttingen helped by Klein and Hilbert who tried to find her a position in the University, although could not make it to be regularly paid. She was surrounded by a group of brilliant students who found exciting inspirations in her lectures.

When the Second World War was about to break out, she moved to Pennsylvania at Bryn Mawr College where she was occasionally allowed to give lessons. During spring semester, she had abdominal surgery which was expected to be routine; she seemed to be recuperating well, but suffered complications and died within a few days.

Although we cannot report all the interesting discoveries in her life, we only point

out some quotations from the major scientists published in her honor at her death.
" *She was the most significant creative mathematical genius thus far produced since the higher education of women was begun. In the realm of algebra, in which the most gifted mathematicians have been busy for centuries, she discovered methods which have proved of enormous importance in the development of the present-day younger generation of mathematicians.*"

Albert Einstein, letter of eulogy to the Times. New York Times, page 12, 4 May 1935. https://nyti.ms/2GJc4o1.

" *This entirely non-visual and noncalculative mind of hers was probably one of the main reasons why her lectures were difficult to follow. She was without didactic talent, and the touching efforts she made to clarify her statements, even before she had finished pronouncing them, by rapidly adding explanations, tended to produce the opposite effect. And yet, how profound the impact of her lecturing was. Her small, loyal audience, usually consisting of a few advanced students and often of an equal number of professors and guests, had to strain enormously in order to follow her. Yet those who succeeded gained far more than they would have from the most polished lecture. She almost never presented completed theories; usually they were in the process of being developed. Each of her lectures was a program. And no one was happier than she herself when this program was carried out by her students. Entirely free of egotism and vanity she never asked anything for herself but first of all fostered the work of her students.* [3]"

Bartel van der Waerden. [4]

To honor the almost anniversary of the publication of her famous theorems, we reported the first page of her outstanding masterpiece in Fig.1.1.

[3]Van der Waerden has written elsewhere that when they went walking in Göttingen, as she did with her students at Bryn Mawr, Emmy Noether would talk so rapidly and with such excitement as to be utterly incomprehensible. It came to him that if he led her on several laps around the city, she became, by the third lap, slightly short of breath and spoke slowly enough that he could understand her.

[4]Reprinted in Auguste Dick's Emmy Noether, 1882-1935, pp. 100–111.

Invariante Variationsprobleme.

(F. Klein zum fünfzigjährigen Doktorjubiläum.)

Von

Emmy Noether in Göttingen.

Vorgelegt von F. Klein in der Sitzung vom 26. Juli 1918[1]).

Es handelt sich um Variationsprobleme, die eine kontinuierliche Gruppe (im Lieschen Sinne) gestatten; die daraus sich ergebenden Folgerungen für die zugehörigen Differentialgleichungen finden ihren allgemeinsten Ausdruck in den in § 1 formulierten, in den folgenden Paragraphen bewiesenen Sätzen. Über diese aus Variationsproblemen entspringenden Differentialgleichungen lassen sich viel präzisere Aussagen machen als über beliebige, eine Gruppe gestattende Differentialgleichungen, die den Gegenstand der Lieschen Untersuchungen bilden. Das folgende beruht also auf einer Verbindung der Methoden der formalen Variationsrechnung mit denen der Lieschen Gruppentheorie. Für spezielle Gruppen und Variationsprobleme ist diese Verbindung der Methoden nicht neu; ich erwähne Hamel und Herglotz für spezielle endliche, Lorentz und seine Schüler (z. B. Fokker), Weyl und Klein für spezielle unendliche Gruppen[2]). Insbesondere sind die zweite Kleinsche Note und die vorliegenden Ausführungen gegenseitig durch einander beein-

---

1) Die endgiltige Fassung des Manuskriptes wurde erst Ende September eingereicht.

2) Hamel: Math. Ann. Bd. 59 und Zeitschrift f. Math. u. Phys. Bd. 50. Herglotz: Ann. d. Phys. (4) Bd. 36, bes. § 9, S. 511. Fokker, Verslag d. Amsterdamer Akad., 27./1. 1917. Für die weitere Litteratur vergl. die zweite Note von Klein: Göttinger Nachrichten 19. Juli 1918.

In einer eben erschienenen Arbeit von Kneser (Math. Zeitschrift Bd. 2) handelt es sich um Aufstellung von Invarianten nach ähnlicher Methode.

Kgl. Ges. d. Wiss. Nachrichten. Math.-phys. Klasse. 1918. Heft 2. 17

Figure 1.1: First page of Emmy Noether's masterpiece on symmetries and groups.

Chapter 2

Noether's theorems

We will have to abandon the philosophy of Democritus and the concept of elementary particles. We should accept instead the concept of elementary symmetries.

Werner Heisenberg

2.1 Noether's first theorem

2.1.1 Noether's symmetries

We now focus on the theorems mentioned earlier. We remind that one of their aims is to associate a conserved charge to every given symmetry.

Since we are about to deal with them, let us introduce the concept of symmetry for the action integral. Let q be a set of generalized coordinates; then a transformation of the coordinates $q(t)$ to new coordinates $q'(t)$ is a symmetry if the action functional does not change when we evaluate it in these two different sets of coordinates, i.e. $I[q'(t)] = I[q(t)]$. We will denote this change of coordinates as $q(t) \to q'(t)$. Above all, we will be interested in a Lie group of symmetries, i.e. symmetries that depend on continuous parameters, so that one can consider infinitesimal symmetries which are sufficiently close to the identity $q'(t) = q(t) + \delta_s q(t)$, and we can assume that the action integral behaves in the same way as to when we deal with symmetries, namely we have that $I[q(t) + \delta_s q(t)] = I[q(t)]$. Thus, symmetries are directions spanned by the q's on which the action does not change. However, the function $\delta_s q(t)$ could have a very complicated expression, depending on $q(t)$ and its derivatives as well. Noether's theorems will characterize Lie groups of symmetries.

We actually defined only a strong version of a symmetry, requiring that the action is strictly invariant. Noether's theorems extend to weaker variations, namely those ones that leave the action invariant up to a boundary term. Therefore we define a (infinitesimal) symmetry as a (infinitesimal) function $\delta_s q(t)$ such that, *for any* $q(t)$, the action is invariant up to boundary terms:

$$\delta_s I[q(t)] \equiv I[q(t) + \delta_s q(t)] - I[q(t)] = \int dt \frac{dK}{dt}; \tag{2.1}$$

where we denote the boundary term with K. We remark that the variation $\delta_s I$ is the variation of the action under the symmetry and so a function of both the configuration $q(t)$ and the symmetry $\delta_s q(t)$.

Any function $\delta_s q(t)$ that satisfies (2.1) belongs to a symmetry; hence a symmetry is defined as a variation that changes the action integral by the integral of a total time derivative: this is of great importance, since the action obtained taking account of this type of variations furnishes the same structure for the equations of motion. Therefore symmetries identify coordinate systems in which the equation of motion look in the same way and thus frames of reference that are equivalent in this view. Eq. (2.1) can be understood as an equation for $\delta_s q(t)$. If, for a given action $I[q(t)]$, we find all the variations $\delta_s q(t)$ satisfying (2.1), then we have found all possible symmetries of the problem, which help in solving the equations of motion (we will provide two examples below). We underline, once again, that the notion of symmetry is defined through variations linked to arbitrary $q(t)$.

2.1.2 Examples

We now provide a couple of examples in order to acquire a better understand of this theorem.

The first one is the invariance under rotations of a system involving a central potential, namely

$$I[r(\vec{t})] = \int dt \left(\frac{m}{2} \dot{\vec{r}}^{\,2} - V(r) \right), \tag{2.2}$$

where the symmetry is

$$\vec{r} \to \vec{r}\,'(t) = R\,\vec{r}(t) \tag{2.3}$$

and R is a constant orthogonal matrix, i.e. $R^T = R^{-1}$. To present this transformation as an infinitesimal Noether's symmetry we consider small angles α, for which one has $R\,\vec{r} = \vec{r} + \vec{\alpha} \times \vec{r}$ and therefore

$$\vec{r}\,'(t) = \vec{r}(t) + \vec{\alpha} \times \vec{r}(t) \Rightarrow \delta_s \vec{r}(t) = \vec{\alpha} \times \vec{r}(t). \tag{2.4}$$

Due to the orthogonality between $\vec{\alpha} \times \vec{r}(t)$ and $\vec{r}(t)$ and the smallness of the angle α this transformation turns out to be a symmetry, that is $I[r(\vec{t}) + \vec{\alpha} \times \vec{r}(t)] = I[\vec{r}(t)]$. We note that for this symmetry the boundary value $K = 0$.

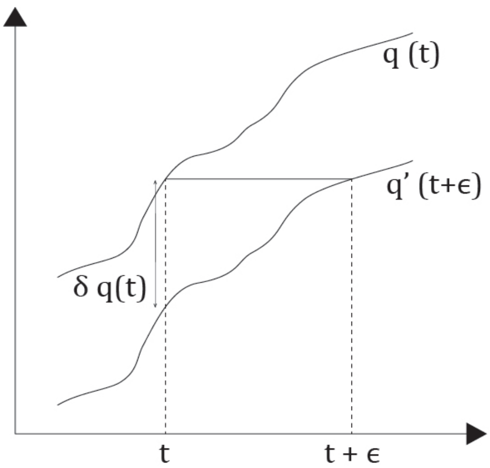

Figure 2.1: The paths $q(t)$, $q'(t)$ related by a time translation $t+\epsilon$. At any given time, their difference yields the variation $\delta q(t)$.

Another example we provide is the invariance of the same action (2.2) under a different symmetry:

$$\delta_s \vec{r}(t) = -\epsilon \dot{\vec{r}}(t), \tag{2.5}$$

where ϵ is constant. By direct calculation,

$$\begin{aligned} \delta_s I[\vec{r}, \delta_s \vec{r}] &= \int dt \bigg(m\dot{\vec{r}} \cdot \delta_s \dot{\vec{r}} - \nabla V \delta_s \vec{r} \bigg) \\ &= \int dt \epsilon (-m\dot{\vec{r}} \cdot \ddot{\vec{r}} + \nabla V \cdot \dot{\vec{r}}) \\ &= \int dt \frac{d}{dt} \bigg(-\epsilon \frac{m}{2} \dot{\vec{r}}^{\,2} + \epsilon V(\vec{r}) \bigg). \end{aligned} \tag{2.6}$$

We see that the action integral is therefore invariant up to the boundary term $K = -\epsilon\bigg(\frac{m}{2}\dot{\vec{r}}^{\,2} - V(\vec{r}) \bigg)$. Physically, we can relate this symmetry to time translations. Indeed consider two function $q(t)$, $q'(t)$ representing the coordinate which are linked by a time translation of magnitude ϵ.

As shown in Fig.2.1, we find that the values of $q'(t)$ are related to those of $q(t)$ via:

$$q'(t+\epsilon) = q(t). \tag{2.7}$$

If ϵ is seen as an infinitesimal quantity, this equation can be cast as

$$q'(t) + \epsilon \dot{q}(t) = q(t) \Rightarrow \delta_s q(t) = -\epsilon \dot{q}(t) \tag{2.8}$$

where we obviously expand in Taylor's series $q'(t+\epsilon)$ around t. Here $\delta_s q(t) = q'(t) - q(t)$ is the difference of the two functions evaluated at the *same* argument t. Before proceeding, we should say something about the symmetry $\delta_s q(t)$:

- The symmetry is represented by the function $\delta_s q(t) = -\epsilon \dot{q}(t)$ which involves only one time. We have transmuted the time translation into a deformation of the function $q(t)$.
- Since the symmetry is simply the difference of two functions evaluated at the *same* time t, it follows directly that $\delta_s \left(\frac{d}{dt} q(t) \right) = \frac{d}{dt}(\delta_s q(t))$.
- We stress a fundamental interpretation of the transformations adopted: *symmetries will always be deformations of the fields, not the coordinates.*

2.1.3 On shell variations

We now focus on another type of variation of the action: the on-shell variation. This kind of variations is in a certain way the opposite to a symmetry. Indeed for symmetries, the variations $\delta_s q(t)$ are forced to satisfy an equation, i.e. (2.1), while the fields $q(t)$ are totally arbitrary; on the other hand for on shell variations, the fields $q(t)$ are forced to satisfy their Euler-Lagrange equations (as it is well known from Analytical Mechanics) while the variations $\delta q(t)$ are arbitrary and not obliged to be symmetries (i.e. to satisfy (2.1)).

Let $\delta q(t)$ be an arbitrary infinitesimal deformation of the variable $q(t)$ [1]. Then, for an action of the form $I[q(t)] = \int dt L(q(t), \dot{q}(t))$ the variation $\delta I[q(t)] = I[q(t) + \delta q(t)] - I[q(t)]$ can be written as

$$\begin{aligned} \delta I[q(t), \delta q(t)] &= \int dt \left(\frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \delta \dot{q}^i \right) \\ &= \int dt \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) \right) \delta q^i + \int dt \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \delta q^i \right) \end{aligned} \tag{2.9}$$

[1] We aim to indicate precisely what we mean with a deformation, i.e. we specify what we indicate as a variation. Given a path $q(x)$ depending on the coordinates $x = x^\mu$ in use (where μ runs over the range desired, i.e. the space on which our functions are defined can be any space), we consider the family of curves $q(x, \alpha)$ labeled by the Greek later α. The initial path is then $q(x, 0)$ obviously. If we select any function $\eta(x)$, a possible varied path could be:

$$q(x, \alpha) = q(x, 0) + \alpha \eta(x),$$

where we define the arbitrary function $\eta(x)$ as the variation of our initial path. The functions $\eta(x)$, as well as the family of curves $q(x, \alpha)$, are assumed to be as regular as required. The variations is thus defined as $\delta q(x) = \frac{d}{d\alpha} q(x, \alpha) \Big|_{\alpha=0} = \eta(x)$.

where we have suppressed the dependence from t of the fields $q(t)$ in order to lighten the notation and we have performed an integration by parts in the second line. We realize that if the fields satisfy the Euler-Lagrange equations of motion associated to it (namely, (1.2) for each index i), the bulk contribution vanishes and the variation reduces to a total time derivative:

$$\delta_o I[q(t)] \equiv I[\underline{q}(t) + \delta q(t)] - I[\underline{q}(t)] = \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\delta q^i\right). \tag{2.10}$$

The bar under $\underline{q}$ indicates that this is the variation evaluated on a solution of the Euler-Lagrange equations of motion and the variation computed in this way are denoted by the symbol δ_o (where the subscript o stands for on-shell). We enlighten once again that (2.10) is valid for *any* fields δq. This variation is taken for a particular chosen solution $\underline{q}$ and for *arbitrary* variations δq.

2.1.4 First theorem

Combining on-shell variation with symmetries we are ready to develop Noether's first theorem. We have now build two different equations that contain the variation of the action:

symmetries δ_s **:** $\qquad \delta_s I[q(t)] = I[q(t) + \delta_s q(t)] - I[q(t)] = \int dt \frac{dK}{dt}$

on shell variations δ_o **:** $\qquad \delta_o I[q(t)] = I[\underline{q}(t) + \delta q(t)] - I[\underline{q}(t)] = \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\delta q^i\right)$

Both terms are boundary terms but each one for very different reason. In the first term $\delta_s q(t)$ satisfies a particular equation, while $q(t)$ is completely arbitrary; on the other hand, in the second term $\underline{q}(t)$ satisfies a particular equation, while $\delta q(t)$ is completely arbitrary.

Putting $q(t) = \underline{q}(t)$ into the symmetry equation and $\delta q(t) = \delta_s q(t)$ into the on shell equation we note that the left sides of these equations is exactly the same. Subtracting the second equation from the first one, the left hand sides cancel out and from the right hand sides we obtain a conservation law:

$$\boxed{\frac{d}{dt}Q = 0 \qquad with \qquad Q = K - \frac{\partial L}{\partial \dot{q}^i}\delta_s q^i} \tag{2.11}$$

Here we are: this is Noether's first theorem: *given a symmetry of the action $\delta_s q(t)$, the combination Q showed in (2.11) is conserved on the paths of the equations of motion.*

It is customary to refer to the conserved quantity Q as *charge*. Due to the elegance of this equation, someone might suppose to fully understand its meaning, captured by the sharp intuition of it. Anyway it is necessary to see some examples to definitely master the complete power of Noether's first theorem.

In the beginning, we could refer to the first two examples of actions reported. The action (2.2) for the central force potential is invariant under rotation with $K = 0$. Our conserved quantity is therefore, by Noether's theorem:

$$\begin{aligned} Q_\alpha &= -m\dot{\vec{r}} \cdot \delta_s \vec{r} \\ &= -m\dot{\vec{r}} \cdot (\vec{\alpha} \times \vec{r}) \\ &= -\alpha \cdot (m\vec{r} \times \dot{\vec{r}}), \end{aligned} \tag{2.12}$$

where we recall that α is a constant angle. Hence we can conclude that the angular momentum defined as the vector $\vec{L} = m\vec{r} \times \dot{\vec{r}}$ is conserved along the trajectories of motion determined by the action (2.2). In this way we have put the conservation of this quantity in a more theoretical setting: it is not a matter of coincidence, but it is due to the symmetry under rotations that the angular momentum happens to be conserved. We found a constant through which we can probe nature.

The same action is invariant again under a different symmetry, namely time translations (2.5); we found that the boundary is $K = -\epsilon\left(\frac{m}{2}\dot{\vec{r}}^{\,2} - V(\vec{r})\right) = -\epsilon L$, where L s the Lagrangian of the action considered. Using Noether's theorem we immediately find that the charge is:

$$Q_\epsilon = -\epsilon\left(\frac{m}{2}\dot{\vec{r}} - V\right) + \epsilon m\dot{\vec{r}}^{\,2} = \epsilon E, \tag{2.13}$$

where $E = \frac{m}{2}\dot{\vec{r}}^{\,2} + V(\vec{r})$ is the total energy. Once again we have related the empirical conservation of a quantity to a theoretical explanation hiding under our vision of nature: energy is conserved due the symmetry of the action under time translations.

The four conserved charges E and $\vec{L}$ for the central field action allow for simplification in the search of solutions to the equation of motion $m\ddot{\vec{r}} = \nabla V$. But we can go a little further, with the so called example of the conformal particle, in order to see how we can completely solve the equation of motion using only conserved charges, if we have enough of them. We solve the problem in one dimension, since it is straightforward to extend the treatment in more dimensions.

We focus our attention on the action

$$I[x] = \int dt \left(\frac{1}{2}m\dot{x}^2 - \frac{\alpha}{x^2}\right) \tag{2.14}$$

and show how we can completely integrate the dynamics only looking at symmetries.

The Euler-Lagrange equation of motion is

$$m\ddot{x} = \frac{2\alpha}{x^3}. \tag{2.15}$$

We know from the previous example that this action is invariant under time translations and that this implies the conservation of energy, namely

$$E = \frac{1}{2}m\dot{x}^2 + \frac{\alpha}{x^2}. \tag{2.16}$$

This equation states an algebraic relation between $x(t)$ and $\dot{x(t)}$.

Further we notice that the potential $V(x) = \frac{\alpha}{x^2}$ is a homogeneous function of the coordinates (i.e. it satisfies $V(\gamma x) = \gamma^k V(x)$). So we have that through the transformation

$$x \to x'(t') = \gamma x(t) \qquad t \to t' = \beta t \tag{2.17}$$

where $\beta = \gamma^{1-\frac{k}{2}}$, the action is changed only by an overall multiplying factor α^k. Since the degree of homogeneity of the potential is -2 we find the Weyl symmetry

$$x \to x'(t') = \sqrt{\lambda}x(t) \qquad t \to t' = \lambda t, \tag{2.18}$$

where, actually, we have consider an overall factor $\gamma = \sqrt{\mu}$. Indeed, under this transformation, $dx/dt \to d(\sqrt{\lambda}x)/d(\lambda x) = \frac{1}{\sqrt{\lambda}}dx/dt$, and the action remains unchanged

$$I \to \int \lambda dt \left(\frac{1}{2}m\frac{\dot{x}^{\,2}}{\lambda} - \frac{\alpha}{\lambda x^2}\right) = I. \tag{2.19}$$

To make use of Noether's theorem we should put the transformation (2.18) in a suitable form, an infinitesimal variation acting on $x(t)$ at some time t. Let $\lambda = 1 + \epsilon$ with $\epsilon \ll 1$ and expand the transformation for $x(t)$ in (2.18) to first order in ϵ :

$$x'((1+\epsilon)t) \approx \left(1 + \frac{\epsilon}{2}\right)x(t) \Rightarrow x'(t) + \dot{x}(t)\epsilon t \approx x(t) + \frac{\epsilon}{2}x(t) \tag{2.20}$$

by which we extract

$$\delta_s x(t) = x'(t) - x(t) = -\epsilon t\dot{x}(t) + \frac{\epsilon}{2}x(t). \tag{2.21}$$

We notice that this transformation act only on $x(t)$ and is a symmetry for the action:

$$\begin{aligned}
\delta_s I[x] &= \int dt \left(\frac{1}{2}m\delta(\dot{x}^{\,2}) - \alpha\delta\left(\frac{1}{x^{\,2}}\right)\right) \\
&= \epsilon \int dt \left[-m\left(\frac{1}{2}\dot{x}^{\,2} + t\dot{x}\ddot{x}\right) + \alpha\frac{x - 2t\dot{x}}{x^3}\right] \\
&= \epsilon \int dt \frac{d}{dt}\left[-m\left(\frac{t\dot{x}^{\,2}}{2} + \frac{\alpha t}{x^{\,2}}\right)\right] \\
&= \epsilon \int dt \frac{d}{dt}\left[-tL\right],
\end{aligned} \tag{2.22}$$

where we see that the boundary term is $K = -tL$. Hence the Weyl symmetry imposes the conservation of the quantity

$$Q = \frac{1}{2}mx\dot{x} - \left(\frac{1}{2}mt\dot{x}^2 + \frac{\alpha t}{x^2}\right) \tag{2.23}$$

up to a sign and an ϵ factor which are both constants and so can be considered irrelevant. Equations (2.16) and (2.23) are two *algebraic* equations [2] for $x(t)$ and $\dot{x}(t)$. Solving them we obtain $x(t)$ as a function of time and two integration constants Q and E, as it should be for a second order differential equation; indeed the solution obtained in this way solves completely the equation of motion (2.15).

We can explicitly check the conservation of the charge:

$$\begin{aligned} \frac{dQ}{dt} &= \frac{1}{2}m(\dot{x}^{\,2} + x\dot{x}) - \frac{1}{2}m(\dot{x}^{\,2} + 2t\dot{x}\ddot{x}) + \alpha\left(\frac{x^{\,2} - 2tx\dot{x}}{x^4}\right) \\ &= (x - 2t\dot{x})\left[\frac{1}{2}m\ddot{x} - \frac{\alpha}{x^3}\right] = 0 \end{aligned} \tag{2.24}$$

due to the equation of motion (2.15).

2.1.5 Hamiltonian mechanics and Lie algebras

Noether's theorem can be used on any functional, not only the action integral in lagrangian form. Indeed, when we use a general action in the hamiltonian formalism (i.e. in phase-space)

$$I[p(t), q(t)] = \int dt(p_i\dot{q}^i - H(p,q)) \tag{2.25}$$

we find an extra structure in it.

Consider the Poisson bracket defined in phase space

$$\{F, G\} = \frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i} \tag{2.26}$$

where $F(p,q,t)$, $G(p,q,t)$ are functions of the whole set of coordinates and their conjugate momenta and could have an explicit time dependence; its associated equations of motion are

$$\begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i} = \{q^i, H\} \\ \dot{p}_i &= -\frac{\partial H}{\partial q^i} = \{p_i, H\}. \end{aligned} \tag{2.27}$$

[2] Note that, in general, at least one of the conserved charges must be an explicit function of time, otherwise there would be no dynamics.

Recall that the total time derivative of a function of the same variables of F or G can be expressed as

$$\frac{dF(p,q,t)}{dt} = \{F, H\} + \frac{\partial F}{\partial t}, \tag{2.28}$$

Whenever the function F considered has vanishing total time derivative, it turns into a conservation quantity along the paths of the equations of motion. In many situations the function F has no explicit time dependence, so that, in order to be a conserved charge, (2.28) requires it has zero Poisson bracket with the Hamiltonian.

We now provide two important results:

1. *Noether's inverse theorem*: If Q is a conserved charge, then the following transformations
$$\delta_s q^i = \{q^i, \epsilon Q\} = \epsilon \frac{\partial Q}{\partial p_i} \qquad , \qquad \delta_s p_i = \{p_i, \epsilon Q\} = -\epsilon \frac{\partial Q}{\partial q^i}, \tag{2.29}$$
where ϵ is a constant infinitesimal parameter, are symmetries of the action. This is called inverse theorem since we first assume the existence of the charge and then verify it give rise to a symmetry.

2. *The Lie algebra of symmetries*: The set of all conserved charges $Q_a (a = 1, 2, \cdots N)$ satisfies a closed Lie algebra:
$$\{Q_a, Q_b\} = f_{ab}{}^c Q_c. \tag{2.30}$$

The proof of these theorems is as follow.

Taking a conserved charge for which (2.28) vanishes and varying the action (2.25) we find:

$$\begin{aligned}
\delta_s I[q(t), p(t)] &= \int dt \Big(\delta_s p_i \dot{q}^i + p_i \frac{d}{dt} \delta_s q^i - \frac{\partial H}{\partial p_i} \delta_s p_i - \frac{\partial H}{\partial q^i} \delta_s q^i \Big) \\
&= \int dt \Big(-\epsilon \frac{\partial Q}{\partial q^i} \dot{q}^i + \frac{d}{dt}(p_i \delta_s q^i) - \epsilon \dot{p}_i \frac{\partial Q}{\partial p_i} + \epsilon \frac{\partial H}{\partial p_i}\frac{\partial Q}{\partial q^i} - \epsilon \frac{\partial H}{\partial q^i}\frac{\partial Q}{\partial p_i} \Big) \\
&= \int dt \Big(\epsilon \Big(-\frac{dQ}{dt} + \frac{\partial Q}{\partial t} + [Q, H] \Big) + \frac{d}{dt}(p_i \delta_s q^i) \Big) \\
&= \int dt \frac{d}{dt} \Big(-\epsilon Q + p_i \delta_s q^i \Big),
\end{aligned} \tag{2.31}$$

which indeed is a total time derivative as required for a symmetry. We notice that we took (2.28) to vanish in the last line. We also stress that we never made use of the equations of motion. In order to obtain a charge, we have to compute the boundary term for the on-shell variation:

$$\begin{aligned}\delta_o I[q(t),p(t)] &= \int dt\bigg(\delta p_i \dot{q}^i + p_i\delta\dot{q}^i - \frac{\partial H}{\partial q^i}\delta q^i - \frac{\partial H}{\partial p_i}\delta p_i\bigg)\\ &= \int dt\bigg(\delta p_i\Big(\dot{q}^i - \frac{\partial H}{\partial p_i}\Big) - \delta q^i\Big(\dot{p}_i + \frac{\partial H}{\partial q^i}\Big)\bigg) + \int dt\frac{d}{dt}(p_i\delta q^i),\end{aligned} \tag{2.32}$$

where we performed an integration by parts in the last line. Due to the fact that we are varying the action on the solutions of the equations of motion (2.27), the first integral in the last line vanishes and we are left with a total time derivative i.e. a boundary term.

Noether's theorem applied to (2.25) with (2.29) as symmetry yields:

$$\tilde{Q} = K - p_i\delta_s q^i = \epsilon Q + p_i\delta_s q^i - p_i\delta_s q^i = \epsilon Q, \tag{2.33}$$

a conserved charge up to an irrelevant infinitesimal constant ϵ.

To prove the second statement (2.30) suppose we have two conserved charges Q_1 and Q_2 (i.e. (2.28) vanishes for both of them). Then we can show that also the Poisson brackets of them is a conserved charge:

$$\frac{d}{dt}\{Q_1,Q_2\} = \{\frac{\partial Q_1}{\partial t} + \{H,Q_1\},Q_2\} + \{Q_1,\frac{\partial Q_2}{\partial t} + \{H,Q_2\}\} = 0, \tag{2.34}$$

where we used the Jacobi's identity[3] for the Poisson brackets. Being a conserved charge, the commutator $\{Q_1,Q_2\}$ generates another symmetry. For instance it can be zero, it may be a new charge, it may be an old one or at least a combination of the charges involved. In any case, the conclusion is that a complete set of conserved charge $Q_a = Q_1, Q_2, \cdots$ must satisfies a Lie algebra (2.30). For further inspections on Lie algebras and Poisson brackets, see Appendix B.

We should now give a brief overview of the conformal particle, studied in the above examples, in hamiltonian formulation. The action integral (2.14) has two charges that completely solves the equations of motion; in hamiltonian form it is embedded with an extra structure. Performing a Legendre transformation [4] of the Lagrangian in (2.14) we obtain the action integral in hamiltonian formalism, namely

$$I[q(t),p(t)] = \int dt\bigg(p\dot{q} - \Big(\frac{p^{\,2}}{2m} + \frac{\alpha}{q^{\,2}}\Big)\bigg). \tag{2.35}$$

We obtain, applying Noether's theorem as explained before in (2.33), three conserved

[3] Namely: $\{h,\{f,g\}\} + \{f,\{g,h\}\} + \{g,\{h,f\}\} = 0$, where f,g,h are arbitrary functions of the phase-space.

[4] We mean the Legendre transformation of $L(q,\dot{q},t)$ with respect to $\dot{q}$.

charges:

$$
\begin{aligned}
H &= \frac{p^{\,2}}{2m} + \frac{\alpha}{q^{\,2}}, \\
Q &= -tH + \frac{1}{2}pq, \\
K &= t^{\,2}H + 2tQ - \frac{m}{2}q^{\,2}.
\end{aligned} \tag{2.36}
$$

Anyway, this system cannot have more than two integration constants. Indeed the following relation can be found among the three charges above

$$
2KH + 2Q^{\,2} + m\alpha = 0. \tag{2.37}
$$

[5]

2.1.6 Noether's first theorem in field theory

Noehter's theorem is perhaps more useful when applied to field theory, which is the formalism used in describing the fundamental forces of Nature. So we aim to briefly introduce the concepts that allow to extend lagrangian and hamiltonian formalism to continuous systems and fields.

We focus on the most common fields, that is fields depending on the three *independent* position coordinates (x, y, z) and time t. Our goal is to translate from the notation ϕ_i to the different one $\phi(x, y, z)$, where just as every subscript i stands to indicate a different generalized coordinate of the system, the same happens to every value of the position coordinates (x, y, z) that stands for a different *generalized coordinate.* The difference lays in the continuous nature of the index labeled with the position coordinate and the discrete feature of the subscripts. We stress for the last time that the position coordinates (x, y, z) are not lagrangian generalized coordinate: as we indicate a generalized discrete coordinate with $\phi_i(t) = q_i(t)$, we indicate a continuous generalized coordinate with $\phi(x, y, z, t)$. The continuous nature arise when we consider a system made of a large number of particles [6] that closely fills our space, so that we deal no more with a discrete set but with a continuous one. The following classical example will clarify all our words.

[5] Another interesting aspect, although it may seem not to be very appropriate for our topic of interest, is that H, Q, K satisfy the $SL(2, \mathbb{R})$ Lie algebra:

$$
\begin{aligned}
\{Q, H\} &= H, \\
\{Q, K\} &= -K, \\
\{H, K\} &= 2Q.
\end{aligned} \tag{2.38}
$$

[6] In field theory we let better tend this number to infinite.

Consider an infinitely long elastic rod that can undergo small oscillatory displacements along the x-axis of the particles it is made of. A system composed of discrete particles that can approximate the continuous rod is an infinite chain of equal mass points spaced a distance a apart and connected by uniform massless springs having force constant k. Assume further that the mass points can move only along the length of the chain. Its Lagrangian is

$$L = T - V = \frac{1}{2}\sum_{i=1}^{\infty}[m\dot{\phi}_i^{\,2} - k(\phi_{i+1} - \phi_i)^{\,2}], \tag{2.39}$$

where ϕ denotes the displacement of the ith particle from its equilibrium position. Relation (2.39) shows clearly that, when we take a number of particles that grows to infinite, the integer index i identifying the particular mass point becomes the continuous position coordinate x: instead of the variable ϕ_i we have $\phi(x)$ and we should better perform an integration over our continuous generalized coordinates instead of a summation.

We aim to obtain a suitable lagrangian formulation for continuous systems. We have just seen that in dealing with non discrete set of point masses, our Lagrangian looks more like an integral over an object to which we could give the name of lagrangian density since it gives the Lagrangian when we perform its integration over the position coordinates. In order to extend our treatment to a three dimensional space we make the following assumptions: a Latin letter index refers only to the three coordinates of the physical space, a Greek letter index refers to all four coordinates, where the first three are the space ones and the fourth is the time variable; a derivative of the field quantities with respect to any one of the four coordinates x^μ will be denoted by the relative index of the variable we are taking the derivative with respect to, separated from the index that stands from the component of the field by a comma, or with the symbol ∂_ν; namely

$$\phi_{\rho,\nu} \equiv \partial_\nu \phi_\rho \equiv \frac{d\phi_\rho}{dx^\nu}. \tag{2.40}$$

We also denote the field component with the subscript ρ, assuming it is a tensor of any order and that the Greek letter could then stands for a multiple index. We further notice that in the hamiltonian principle of least action the time variable plays the same role of the three position coordinates and that all these four variables are to be treated as completely independent. So we are tempted to write the action as the integral in 4-dimensions of the lagrangian density:

$$\begin{aligned} I[\phi(x)] &= \int L dt \\ &= \int \left(\int \mathcal{L}(\phi,\, \phi_{,\nu},\, x) d^3x \right) dt \equiv \int \mathcal{L}(\phi,\, \phi_{,\nu},\, x) d^4x. \end{aligned} \tag{2.41}$$

We notice that in general the lagrangian density depends also from the derivative of the field ϕ with respect to the position coordinates and not only to the time coordinate;

that is we allow also this dependence of the lagrangian density in order to treat all four coordinates equally. To derive the equations of motion in field theory from Hamilton principle we take the variation of the action and make it vanish. Recall that the variation of the field vanishes at the bounding surface $\mathcal{S}$ of the region of integration, similar to the case of integration along time variable only previously treated. We also point out that the functions that should be varied are the ones from which the lagrangian density depends on, namely ϕ_ρ and $\phi_{\rho,\nu}$; that is we have to take the variation with respect to the field and its partial derivatives. This could be written as

$$\delta I[\phi(x)] = \int \left(\frac{\partial \mathcal{L}}{\partial \phi_\rho} \delta \phi_\rho + \frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \delta \phi_{\rho,\nu} \right) d^4x, \tag{2.42}$$

integration by parts yields

$$\delta I[\phi(x)] = \int \left[\frac{\partial \mathcal{L}}{\partial \phi_\rho} - \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \right) \right] \delta \phi_\rho (d^4x) + \int \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \delta \phi_\rho \right) d^4x. \tag{2.43}$$

When the variation of the Lagrangian vanishes on the surface $\mathcal{S}$ the second term makes no contribution. Since the variation of the fields are totally independent and arbitrary we obtain the equations of motion in field theory:

$$\frac{\partial \mathcal{L}}{\partial \phi_\rho} - \partial_\nu \left(\frac{\mathcal{L}}{\partial \phi_{\rho,\nu}} \right) = 0, \tag{2.44}$$

which holds for each component of the field labeled with ρ as explained above. Comparing it with the discrete case, since we deal with continuous spatial index x^i, each of (2.44) corresponds to an entire set of discrete equations of motion. In discrete systems the Lagrangian is uncertain up to a total time derivative of an arbitrary function of the generalized coordinates and time; with continuous systems the lagrangian density $\mathcal{L}$ is uncertain up to a term of the form

$$\partial_\nu F_\nu(\phi, x) \tag{2.45}$$

where F, as indicated, is any four differentiable functions of the field quantities and the coordinates. Indeed its variation is zero, so that it does not change the equations of motion:

$$\delta \int \partial_\nu F_\nu(\phi, x)(d^4x) = \delta \int F_\nu(\phi, x) d\sigma^\nu = 0 \tag{2.46}$$

where $d\sigma^\nu$ represents the components of an element of surface oriented along the direction of the outward normal; we specify that this variation vanishes since we assumed that on the surface $\mathcal{S}$ the variations of the fields are zero.

The transition from discrete to field theory notation can be summarize as follow:

$$\begin{aligned} L &\to \mathcal{L} \\ t &\to x^\mu \\ q_i &\to \phi_\rho \\ \dot{q}_i &\to \phi_{\rho,\nu} \end{aligned} \tag{2.47}$$

We are now ready to derive Noether's theorem in field theory. Clearly the on-shell variation, i.e. the variation of the action integral along the path described by the equations of motion for arbitrary variation of the field and its partial derivative, are to be computed as

$$\delta_o I[\phi(x)] \equiv I[\underline{\phi}(x) + \delta\phi(x)] - I[\underline{\phi}(x)] = \int \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \delta\phi_\rho \right) d^4x \tag{2.48}$$

where, as before, with the underlying bar $\underline{\phi}$ we indicate that we are taking the variations on the trajectories of motion.

The set of symmetries for the field theory action is defined as the set of all infinitesimal functions $\delta_s\phi(x)$ such that, for *arbitrary* $\phi(x)$:

$$\delta_s I[\phi(x)] \equiv I[\phi(x) + \delta_s\phi(x)] - I[\phi(x)] = \int \partial_\nu K^\nu d^4x. \tag{2.49}$$

That is: symmetries are variations of the fields such that the variation of the action integral vanishes up to a 4-divergence of a boundary term K^ν. As in the discrete case, symmetries identify sets of coordinates in which the structure of the equation of motions remain unchanged. Let us stress out the following details:

- (2.49) is an equation for the function $\delta_s\phi$, not for ϕ: $\delta_s\phi$ is a symmetry provided that (2.49) holds for *all* ϕ.
- The variations act directly on the fields: coordinates play no role, following the tensorial formalism adopted.

Since (2.44) is valid for all variations $\delta\phi$ (and in particular for $\delta_s\phi$) and (2.49) is valid for all fields ϕ (and for $\underline{\phi}$, too), thus, inserting $\delta_s\phi$ into (2.44) and $\underline{\phi}$ into (2.49), we notice that the left hands sides are equal; subtracting both equations we obtain the conserved current equation

$$\boxed{\partial_\mu J^\mu \qquad with \qquad J^\mu \equiv \frac{\partial \mathcal{L}}{\partial \phi_{\rho,\mu}} \delta_s\phi_\rho - K^\mu.} \tag{2.50}$$

Thus, infinitesimal symmetries enjoyed by a lagrangian density $\mathcal{L}$ give rise to conserved charges, which can be used to drive our understanding of the model described by the lagrangian density $\mathcal{L}$.

2.1.7 Noether's first theorem: a more precise formalism

Although we have already introduced Noether's first theorem, we would like to put it in a more precise form in order to shape it in a better formalism which suit a comparison with Noether's second theorem, which we will report in the next chapter once we will have provided some necessary mathematical tools. We also leave the proof of this theorem to the section dedicated to Noether's second theorem. We will deal with the original versions of these theorems, namely the ones introduced by Emmy Noether in her article, even though they detach a little from the ones we can find in today literature as well as the one we introduced earlier for the first theorem.

We will apply Noether's theorems to lagrangian densities depending on an arbitrary number of fields ϕ_ρ, where $\rho = 1, ..., N$. We will assume our space to depend on four coordinates, three space-like x^i and one time-like t, and we indicate them with the compressed notation x^μ (recall that the Greek index run from 0 to 3). Further we let the lagrangian density depends from the fields and their first derivatives and also the coordinates, that is $\mathcal{L} = \mathcal{L}(\phi, \phi_{,\mu}, x^\mu)$. Both theorems are derived from the variational problem of the action integral $\int d^4x\mathcal{L}$. We finally remind the reader to the definition of symmetry: it is a transformation rule on the fields that makes the action invariant up to boundary terms.

Theorem: Noether's first theorem. *Let G_η be a finite continuous group of transformations $\delta_s\phi_\rho$ depending in a differentiable manner on η constant parameters ω_α ($\alpha = 1, ..., \eta$). If the action integral is invariant under G_η, then the following relations hold:*

$$\sum_{\rho=1}^{N} [\Phi]_\rho \frac{\partial(\delta_s\phi_\rho)}{\partial(\Delta\omega_\alpha)} = \partial_\mu j^\mu_\alpha \tag{2.51}$$

where

$$j^\mu_\alpha = -\left[\mathcal{L}\frac{\partial(\delta x^\mu)}{\partial(\Delta\omega_\alpha)} + \sum_{\rho=1}^{N} \frac{\partial\mathcal{L}}{\partial\phi_{\rho,\mu}} \frac{\partial(\delta_s\phi_\rho)}{\partial(\Delta\omega_\alpha)}\right] \tag{2.52}$$

Let us clarify our notation:

- $[\Phi]_\rho$ is an expression associated to the field ϕ_ρ and defined as:

$$[\Phi]_\rho = \frac{\partial\mathcal{L}}{\partial\phi_\rho} - \partial_\mu \frac{\partial\mathcal{L}}{\partial\phi_{\rho,\mu}}, \tag{2.53}$$

  so that in this compact notation the Euler-Lagrange equations for the ρ-th field are written as $[\Phi]_\rho = 0$.

- With $\Delta\omega_\alpha$ we indicate an infinitesimal displacement from ω_α; we are allowed to do so since the group G_η considered is continuous. Further, since the group of transformations depends in a differentiable manner from the constant parameters ω_α, the infinitesimal variations δx^μ, $\delta_s\psi_i$ depend from the infinitesimal parameters $\Delta\omega_\alpha$ in the same way and we can write them as:

$$\delta x^\mu = \sum_{\alpha=1}^{\rho} \frac{\partial(\delta x^\mu)}{\partial(\Delta\omega_\alpha)}\Delta\omega_\alpha \qquad \delta_s\phi_i = \sum_{\alpha=1}^{\rho} \frac{\partial(\delta_s\psi_i)}{\partial(\Delta\omega_\alpha)}\Delta\omega_\alpha, \tag{2.54}$$

as we have highlighted in (2.52).

- The minus sign in the expression for the current j^μ_α is somehow strange; however this fact will be clearer when we will deal with Noether's second theorem. Further, when we assume that we are on the paths of motion, i.e. when the Euler-Lagrange equations of motion are satisfied ($[\Phi]_\rho = 0$), this current turns out to vanish and we recover the important result of having η conserved quantities. These indeed are:

$$Q_\alpha = \int d^3x j^0_\alpha(x^i, t) \tag{2.55}$$

2.1.8 Hamiltonian formulation

We should dig further in Noehter's first theorem in field theory formalism, but first we would like to introduce the hamiltonian formulation of it, since it will be of great usefulness later. We define the canonical momentum densities as

$$\pi^\rho(x) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}_\rho}, \tag{2.56}$$

where as usual the superposed dot stands for partial derivation with respect to time variable. The quantities $\phi_\rho(x^i, t), \pi^\rho(x^i, t)$ together define the infinite-dimensional *phase space* describing the field and its development. As in the discrete case, it is straightforward to see that when the Lagrangian has some cyclic generalized coordinate ϕ^ρ (i.e. it does not contain ϕ^ρ explicitly) we find a conserved quantity, namely

$$\partial_\mu(\frac{\partial\mathcal{L}}{\phi_{\rho,\mu}}) = \frac{d\pi^\rho}{dt} + \partial_i(\frac{\partial\mathcal{L}}{\phi_{\rho,i}}) = 0. \tag{2.57}$$

Thus it follows that if ϕ^ρ is cyclic, there is an integral conserved quantity

$$\Pi^\rho = \int \pi^\rho(x^i, t)d^3x. \tag{2.58}$$

We can perform a Legendre transformation on the lagrangian density $\mathcal{L}(\phi_\rho,\, \phi_{\rho,\nu},\, x^\nu)$ with respect to $\dot{\phi}_\rho$, obtaining the hamiltonian density

$$\mathcal{H}(\phi,\, \phi_{,i},\, \pi,\, x) = \pi^\rho\dot{\phi}_\rho - \mathcal{L}. \tag{2.59}$$

From this definition we can easily obtain the first equation of motion by

$$\frac{\partial \mathcal{H}}{\partial \pi^\rho} = \dot{\phi}_\rho + \pi^\mu \frac{\partial \dot{\phi}_\mu}{\partial \pi^\rho} - \frac{\partial \mathcal{L}}{\partial \dot{\phi}_\mu} \frac{\partial \dot{\phi}_\mu}{\partial \pi^\rho} = \dot{\phi}_\rho. \tag{2.60}$$

Unfortunately, the equation for π^ρ requires a bit more calculation. To this purpose we notice that $\mathcal{H}$ in terms of the canonical variables is a function of ϕ_ρ through the explicit dependence of $\mathcal{L}$, and through $\dot{\phi}_\rho$. Therefore we have:

$$\frac{\partial \mathcal{H}}{\partial \phi_\rho} = \pi^\mu \frac{\partial \dot{\phi}_\mu}{\partial \phi_\rho} - \frac{\partial \mathcal{L}}{\partial \dot{\phi}_\mu} \frac{\partial \dot{\phi}_\mu}{\partial \phi_\rho} - \frac{\partial \mathcal{L}}{\partial \phi_\rho} = -\frac{\partial \mathcal{L}}{\partial \phi_\rho}. \tag{2.61}$$

With the help of Euler-Lagrange equations (2.44), this can be cast as

$$\frac{\partial \mathcal{H}}{\partial \phi_\rho} = -\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \phi_{\rho,\mu}} \right) = -\dot{\pi}^\rho - \partial_i \left(\frac{\partial \mathcal{L}}{\partial \phi_{\rho,i}} \right). \tag{2.62}$$

Since we still have the presence of $\mathcal{L}$, this is not a useful form. However, by an exactly parallel derivation, we find

$$\frac{\partial \mathcal{H}}{\partial \phi_{\rho,i}} = \pi^\mu \frac{\partial \dot{\phi}_\mu}{\partial \phi_{\rho,i}} - \frac{\partial \mathcal{L}}{\partial \dot{\phi}_\mu} \frac{\partial \dot{\phi}_\mu}{\partial \phi_{\rho,i}} - \frac{\partial \mathcal{L}}{\partial \phi_{\rho,i}} = \frac{\partial \mathcal{L}}{\partial \phi_{\rho,i}}. \tag{2.63}$$

Inserting the above equation into (2.62) we obtain the second canonical equation:

$$\frac{\partial \mathcal{H}}{\partial \phi_\rho} - \partial_i \left(\frac{\partial \mathcal{H}}{\partial \phi_{\rho,i}} \right) = -\dot{\pi}^\rho. \tag{2.64}$$

Canonical equations (2.60) and (2.64) can be written in a notation more closely approaching Hamilton's equations for a discrete system by introducing the notion of a *functional derivative*, defining it as:

$$\frac{\delta}{\delta \psi_\rho} = \frac{\partial}{\partial \psi_\rho} - \partial_i \left(\frac{\partial}{\partial \psi_{\rho,i}} \right). \tag{2.65}$$

The functional derivative is clearly linear, obeys Leibniz rule and chain rule.

Due to the independence of $\mathcal{H}$ from ${\pi^\rho}_{,i}$, the equations of motion can be written as

$$\dot{\phi}_\rho = \frac{\delta \mathcal{H}}{\delta \pi^\rho}, \qquad \dot{\pi}^\rho = -\frac{\delta \mathcal{H}}{\delta \phi_\rho}. \tag{2.66}$$

With this symbolism the equations of motion (2.44) take the form

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}_\rho} \right) - \frac{\delta \mathcal{L}}{\delta \phi_\rho} = 0. \tag{2.67}$$

Anyway, although it sketches an elegant parallelism with discrete system, the functional derivative suppressed the parallel treatment of time and space variables and here is the reason why it is used carefully. We notice that the variational problem can be put in a more precise and formal structure in the following way. The problem of finding an extreme value for the action integral can be computed via the functional derivative as:

$$\delta I = \int d^4x^{\frac{\delta \mathcal{L}}{\delta \phi_\rho}} \delta\phi_\rho = 0. \tag{2.68}$$

Assuming the variations being arbitrary and vanishing at the borders, we find the extreme value is given by the equations of motion above, namely Euler-Lagrange equations for field theory (2.67).

Finally, we briefly introduce the notion of Poisson brackets in field theory, as we will make use of them later. We aim to derive them as a natural generalization from the discrete case we are at ease with. The subtle difference to grasp is, once again, the continuous nature of generalized coordinates in field theory. Indeed we can associate to every point of coordinates x^μ a value of the generalized coordinates $\phi_\rho(x^\mu)$, in analogy to the example of the infinite elastic rod tackled at the beginning. Hence we should have a relation of the form (at fixed time):

$$\begin{gathered} \{\phi_\rho(x), \pi_\rho(y)\} = \begin{cases} 1 & if \quad x = y \\ 0 & elsewhere \end{cases} \\ \{\phi_\rho, \phi_\nu\} = 0, \end{gathered} \tag{2.69}$$

where in the last equation we suppressed the dependence of the field from the position coordinates since the result is unaffected by them. We are temped to express this relation in distributional sense, since it fits its formalism quite neatly; that is

$$\{\phi_\rho(x), \pi_\nu(y)\} = \delta(x - y)\delta_{\rho,\nu}, \tag{2.70}$$

which corresponds to the well known equation $\{q_i(t), p_j(t)\} = \delta_{ij}$ [7]. Regarding the similar equation for the associated momenta and the relation that holds between them and the generalized coordinates, we have:

$$\begin{gathered} \{\phi_\rho(x), \pi_\nu(y)\} = \delta(x - y)\delta_{\rho\nu} \\ \{\phi_\rho, \phi_\nu\} = \{\pi_\rho, \pi_\nu\} = 0. \end{gathered} \tag{2.71}$$

We can now furnish a definition of the Poisson brackets for continuous systems, based on the previous relations.

[7] $\delta(x, y)$ is the delta Dirac distribution, δ_{ij} is the Kronecker delta.

Poisson brackets for fields. *Given two quantities obtained by their respective densities*

$$\begin{aligned} A &= \int \mathcal{A}(\phi,\, \phi_{,\nu},\, \pi,\, \pi_{,\nu},\, x^i, t) d^3x \\ B &= \int \mathcal{B}(\phi,\, \phi_{,\nu},\, \pi,\, \pi_{,\nu},\, x^i, t) d^3x, \end{aligned} \tag{2.72}$$

where ϕ stands for the set of generalized coordinates adopted and π for their associated momenta, their Poisson brackets are defined as:

$$\{A, B\} = \int \sum_\rho \left(\frac{\delta \mathcal{A}}{\delta \phi_\rho} \frac{\delta \mathcal{B}}{\delta \pi_\rho} - \frac{\delta \mathcal{B}}{\delta \phi_\rho} \frac{\delta \mathcal{A}}{\delta \pi_\rho} \right) d^3x \tag{2.73}$$

where $\frac{\delta}{\delta\psi}$ is the functional derivative (2.65). We notice that this definition states that the Poisson brackets for the fields considered has to be taken at a given time, at which each field is evaluated.

We notice that this definition reduces to (2.71), as expected. Indeed, we immediately realize that the functional derivative depends on the point indicated with position coordinates x^μ; that is the functional derivative

$$\frac{\delta F(x)}{\delta \psi_\rho(y)} = \left[\frac{\partial F(x)}{\partial \psi_\rho(y)} - \partial_i \left(\frac{\partial F(x)}{\partial \psi_{\rho,i}(y)} \right) \right] \delta(x-y) \tag{2.74}$$

is actually a distribution due to its behavior under integration through space coordinates. We recall the definition of a distribution: given a space of functions Ω a distribution is a linear functional on this space which is also continuous [8].

Once we have developed this formalism we can easily show that

$$\frac{dF(x)}{dt} = \{F, H\} + \frac{\partial F(x)}{\partial t}. \tag{2.75}$$

Therefore we have recovered the fundamental results obtained for the discrete case, namely the two theorems tackled in (2.29) and (2.30), with the relative transitions in notation for field theory.

[8] Actually we should be more precise than this. Indeed, as it is well known, we should give a topology on the space Ω in order to define continuity. Further, the space in use identifies the type of distributions. We will deal with spaces of functions which have compact support or with the so called distribution with compact support. We could just only ask the space Ω to be the Schwartz's one and deal with temperate distributions (as the Fourier's transform).

2.2 Noether's second theorem

2.2.1 Lie derivatives and background fields

In Noether's theorem it is of high importance to see variations as transformations acting *on* fields instead of a change of coordinates: therefore we see transformations as active operations instead of passive ones. As we have emphasized, this is the natural interpretation of a symmetry. Indeed the form of the action may change if the coordinates are changed, but not its value. Moreover we would like to tackle every kind of transformations, not only those ones performed through constant parameters. These are the leading ideas that drive the notion of Lie derivative, that is a derivation along congruences on a manifold. We try to extend the variation used in Noether's first theorem, namely transformations that depends on a *constant* parameter, to transformations that depends on a variable which can varies instead. Indeed, Noether's second theorem copes with this transformations and being able to deal with them is fundamental in order to understand the meaning of the theorem. We stress for the last time that for Noether's theorems one never need to change the coordinates; only the fields transform.

We should be more precise about this topic; however we assume as well known the notion of manifold and coordinates (namely, charts) on it. The idea is to find a set of curves that behave as what we intuitively individuate as coordinates: when we follow one of this line all the parameters that build the other coordinate lines do not change. A congruence is, roughly speaking, a foliation of the manifold we deal with through parametric lines. For instance, let us consider a two-dimensional manifold $\mathcal{M}$, a curve Σ_0 and a vector field $\vec{U} = \frac{d}{d\mu}$ [9]. We recall that a vector field in an open set $A \subseteq \mathcal{M}$ is an application which maps each point $P \in A$ into a vector $\vec{u}$ belonging to the tangent space T_A of A. Indeed a vector $\vec{u}$ is a linear operator defined at a point on the manifold $\mathcal{M}$ acting on scalar functions [10] and which gives the derivative of the function evaluated at the particular point considered.

We call a congruence of the vector field $\vec{U}$ the family of integral curves of $\vec{U}$ that start from the curve Σ_0 (along which $\lambda = \lambda_0$), covering the manifold $\mathcal{M}$ and which have the vectors defined by the vector fields as derivatives at each point. By covering, we here mean that there is one (and only one) integral curve of $\vec{U}$ on the manifold $\mathcal{M}$. Moving a point $P(\lambda_0)$ from Σ_0 along the corresponding congruence to the point $P(\lambda_0 + \Delta\lambda)$ is called *push forward* or *Lie dragging*. This is the active interpretation of a change of coordinates.

Spacetime translations $x^\mu \to x'^\mu = x^\mu + \epsilon^\mu$ are seen as transformations of the field as follow: given $\phi(x)$ one builds a new field, namely the Lie dragged field $\phi'(x)$, whose

[9] It turns out that, in order to perform the Lie derivative along this curves, the vector field $\vec{U}$ need to be at least of class $\mathcal{C}^1$.

[10] A scalar function f is a function defined from the manifold $\mathcal{M}$ to $\mathbb{R}$.

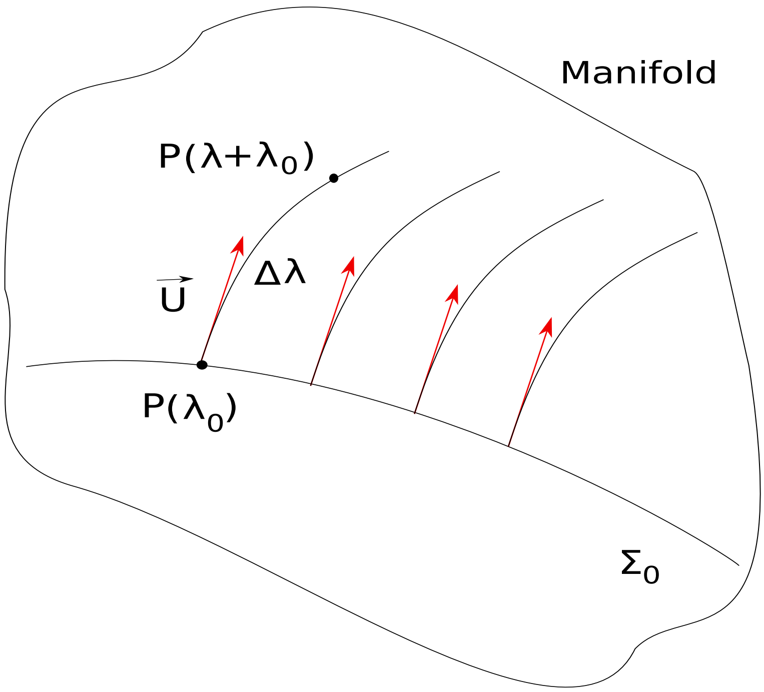

Figure 2.2: Lie dragging on a congruence. The vector field is denoted by $\vec{U}$ and generates the integral curves depending on the parameter λ. The curve from which all of the parametric lines depart is Σ_0, along which the parameter λ takes the constant value λ_0. We also try to depict the Lie dragging of the point P from Σ_0 by a displacement $\Delta\lambda$ along the congruence.

values are

$$\begin{aligned}\phi'(x) &= \phi(x-\epsilon) \\ &\simeq \phi(x) - \epsilon^\mu \partial_\mu \phi(x),\end{aligned} \tag{2.76}$$

where we retained first order in ϵ^μ. The variation of the field which is associated to a translation of coordinates is then

$$\delta\phi(x) = \phi'(x) - \phi(x) = -\epsilon^\mu \partial_\mu \phi(x). \tag{2.77}$$

There is a formal way to obtain the Lie derivative of a vector field $\vec{U}$. This is based on the concept of Lie dragging and on the existence of coordinate lines. However, these topics better fit the field of differential geometry; therefore we aim to introduce the notion of Lie derivative in a more intuitive way, namely through infinitesimal displacements.

First of all we must recall that Lie derivative is defined on tensorial quantities and that tensors transform according to precise law under a change of coordinates, which is understood to be seen as an active operation. Moreover we stress that Lie derivative obeys the Leibniz rule:

$$\pounds_{\vec{V}}(\tilde{w}(\vec{W})) = (\pounds_{\vec{V}}\tilde{w})(\vec{W}) + \tilde{w}(\pounds_{\vec{V}}\vec{W}), \tag{2.78}$$

where with $\pounds_{\vec{V}}$ we indicate the Lie derivative along the congruence build on the vector field $\vec{V}$ and with $\tilde{w}$ a one-form (i.e. a covariant tensor with components w_ρ). To compute the correct formulae for tensors of any order we only miss the correct active transformation of the jacobian that individuates the change of coordinates of tensors. We stress once again that transformations are to be considered as active operations. Hence we have:

$$x'^\mu = x^\mu + \zeta^\mu(x) \tag{2.79}$$

$$\Rightarrow \frac{\partial x'^\mu}{\partial x^\nu} = \delta^\mu_\nu + \partial_\nu \zeta^\mu(x) \quad , \quad \frac{\partial x^\mu}{\partial x'^\nu} = \delta^\mu_\nu - \partial_\nu \zeta^\mu(x'). \tag{2.80}$$

We have provided all the tools to furnish the Lie derivative of tensors of any order. We have always considered only transformations encoded in constant parameters; we aim to extend our possibilities to any kind of traformations: Lie derivative is the operator that allows us to do so. We recall that we are interested in infinitesimal transformations.

- *Scalar fields* $\rightarrow$
  These are tensors of order zero, for which the transformation law for a change of coordinates yields

  $$\phi'(x') = \phi(x). \tag{2.81}$$

  Expanding $\phi'(x+\zeta) = \phi(x)$ to linear order in ζ we have

  $$\delta\phi(x) = \phi'(x) - \phi(x) = -\zeta^\nu(x)\partial_\nu \phi(x) = \pounds_{\vec{\zeta}}[\phi]. \tag{2.82}$$

  This looks like the one we derive earlier but now the displacement $\zeta^\mu(x)$ is arbitrary.

- *Vector fields* $\rightarrow$
  These are contravariant tensors of order (1,0), for which the transformation law for a change of coordinates implies

$$V'^{\mu}(x') = \frac{\partial x'^{\mu}}{\partial x^{\nu}} V^{\nu}(x). \tag{2.83}$$

  Expanding right and left side to first order we find:

$$V'^{\mu}(x) + \zeta^{\nu}(x)\partial_{\nu}V^{\mu}(x) = V^{\mu}(x) + (\partial_{\nu}\zeta^{\mu}(x))V^{\nu}(x), \tag{2.84}$$

  from which we obtain

$$\delta V^{\mu}(x) = V'^{\mu}(x) - V^{\mu}(x) = (\delta_{\nu}\zeta^{\mu}(x))V^{\nu}(x) - \zeta^{\nu}(x)\partial_{\nu}V^{\mu}(x) = \pounds_{\vec{\zeta}}[\vec{V}] \tag{2.85}$$

  In the usual formalism this is written:

$$\pounds_{\vec{\zeta}}[\vec{V}] = [\vec{\zeta}, \vec{V}] = \left(\zeta^{i}\frac{\partial v^{j}}{\partial x^{i}} - v^{i}\frac{\partial v^{j}}{\partial x^{i}}\right)\mathbf{e}_{j}, \tag{2.86}$$

  where the square brackets do not stand for the Poisson ones but for the commutator of the two vectors [11].

- *One-forms* $\rightarrow$ These are covariant tensors of order (0,1), for which the transformation law for a change of coordinates implies

$$A'_{\mu}(x') = \frac{\partial x^{\nu}}{\partial x'^{\mu}} A_{\nu}(x). \tag{2.87}$$

  Expanding both sides we obtain:

$$A'_{\mu}(x) + \zeta^{\nu}(x)\partial_{\nu}A_{\mu}(x) = A_{\mu}(x) - A_{\nu}(x)\partial_{\mu}\zeta^{\nu}(x), \tag{2.88}$$

  hence it follows

$$\delta A_{\mu}(x) = A'_{\mu}(x) - A_{\mu}(x) = -\zeta^{\nu}(x)\partial_{\nu}A_{\mu}(x) - A_{\nu}(x)\partial_{\mu}\zeta^{\nu}(x) = \pounds_{\vec{\zeta}}[\tilde{A}]. \tag{2.89}$$

  This can be expressed in a more traditional notation by means of (2.78), namely [12]

$$\pounds_{\vec{\zeta}}[\tilde{A}] = \left(\zeta^{j}\frac{\partial A_{i}}{\partial x^{j}} - A_{j}\frac{\partial \zeta^{j}}{\partial x^{i}}\right)\mathbf{e}_{i} \tag{2.90}$$

[11] Lie brackets (i.e. the commutator) are defined on two operators and the space the operators belong to is closed under their action. Indeed, as known from differential geometry, vectors are to be considered as (linear) operators acting on fields (scalar functions defined from the manifold to $\mathbb{R}$) of the manifold, which give the derivative of them along the integral curve they define at the point considered. This is the reason why we expanded them on the vectors $\mathbf{e}_j$, since they form a vectorial space and we picked the canonical basis. In fact, as Lie brackets "remain" in the space they act on, the commutator of two vectors is itself a vector.

[12] Again, the same concepts are endorsed. In particular the Lie derivative of a one-form is itself a one-form and the set of all one-forms (linear functional acting on vector spaces) is a linear space.

Inspecting the transformation laws for tensors of higher order or exploiting the Leibniz rule for Lie derivative, it is possible to obtain the variation of any tensors with the same technique used above.

2.2.2 Second theorem

We can now deal with variations depending not only from constant parameters [13], which was the case for Noether's first theorem, but also from arbitrary functions. This is the case treated in Noether's second theorem, which endorses a deeper insight in symmetries in physics. Noether's second theorem did not receive enough attention, although it is of great importance. Gauge theories find their roots in it, but literature does not give much credit to this theorem. As for the first one, we report a formulation which is as close as possible to the original version and we also provide a proof of the results obtained. We still deal with the variation of integral action and we keep the same notation as for Noether's first theorem (see section (2.1.7)).

Theorem: Noether's second theorem. *Let $G_{\infty\eta}$ be an* <u>*infinite*</u> *continuous group of transformations $\delta_s\phi_\rho$ depending in a differentiable manner from η arbitrary functions $p_\alpha(x^\mu)$, $(\alpha = 1, ..., \eta)$ and their first derivatives. If the action integral is invariant under $G_{\infty\eta}$, then the following relations hold:*

$$\sum_{\rho=1}^{N}[\Phi]_\rho a_{\alpha-\rho} = \sum_{\rho=1}^{N}\partial_\mu\{[\Phi]_\rho b^\mu_{\alpha-\rho}\} \tag{2.91}$$

where $a_{\alpha-\rho}(x^\mu, \phi, \partial_\mu\phi)$ e $b^\mu_{\alpha-\rho} = b^\mu_{\alpha-\rho}(x^\mu, \phi, \partial_\mu\phi)$ are arbitrary functions whose explicit form depends from the particular transformations used.

We now give the proof of these results. We split it in two parts: in the first one, which provides a proof also for Noether's first theorem, we furnish the solution for the variational problem of the action integral; in the second one, assuming that we are dealing with a group $G_{\infty\eta}$, we demonstrate the results of Noether's second theorem. To make it simpler, we focus only with a single field ϕ, since it is easier to extend the treatment to more fields.

Proof. Step 1. As we did in the above sections, we consider an infinitesimal transformation of the coordinates and we write it as a transformation of the fields, namely:

$$\begin{aligned} x^\mu &\to x'^\mu = x^\mu + \delta x^\mu \\ \phi &\to \phi'(x'^\mu) = \phi(x^\mu) + \delta\phi \end{aligned} \tag{2.92}$$

[13] Literature usually refers to them as rigid transformations.

Recalling that we have:

$$\begin{aligned} \delta\phi &= \phi'(x') - \phi(x) = \phi'(x+\delta x) - \phi(x) \\ &= \phi'(x) - \phi(x) + \delta x^\mu \partial_\mu \phi' = \delta_\sim \phi + \delta x^\mu \partial_\mu \phi, \end{aligned} \tag{2.93}$$

where we put $\delta_\sim \phi = \phi'(x) - \phi(x)$ and we exploit the relation $\partial_\mu \phi'(x) = \partial_\mu \phi(x)$ at first order. Thus we can formally write $\delta = \delta_\sim + \delta x^\mu \partial_\mu$ and for symmetries $\delta = \delta_s + \delta x^\mu \partial_\mu$

Let now I be the action integral in field theory formalism associated to the lagrangian density $\mathcal{L}$. Varying it and considering variations that are symmetries, we obtain:

$$\delta_s I = \int d^4x \delta\mathcal{L} + \int \delta(d^4x)\mathcal{L} = 0 \tag{2.94}$$

Let us focus on the first term. We compute the variation of the lagrangian density.

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial x^\mu}\delta x^\mu + \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\delta(\partial_\mu\phi) \\ &= \frac{\partial\mathcal{L}}{\partial x^\mu}\delta x^\mu + \frac{\partial\mathcal{L}}{\partial\phi}(\delta_s\phi + \delta x^\mu\partial_\mu\phi) + \frac{\partial\mathcal{L}}{\partial\phi_\mu}(\delta_s + \delta x^\nu\partial_\nu)(\partial_\mu\phi) \\ &= \delta x^\nu\left[\frac{\partial\mathcal{L}}{\partial x^\nu} + \frac{\partial\mathcal{L}}{\partial\phi}\partial_\nu\phi + \frac{\partial\mathcal{L}}{\partial\phi_\mu}\partial_\mu(\partial_\nu\phi)\right] + \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta_s\phi + \frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\delta_s(\partial_\mu\phi)\right] \\ &= \delta x^\mu(\partial_\mu\mathcal{L}) + \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\right]\delta_s\phi + \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial\phi_\mu}\delta_s\phi\right] \end{aligned} \tag{2.95}$$

We then focus on the second term of (2.94). We have to compute the variation of the infinitesimal volume element of the four coordinates d^4x. For an arbitrary transformation we have $d^4x' = |J(x^\mu)|d^4x$, where with the symbol $|J(x^\mu)|$ we mean the absolute value of the determinant of the jacobian of the transformation, i.e. $J(x^\mu) \equiv det(\frac{\partial x'^\mu}{\partial x^\nu})$. The transformation we are using is $x'^\mu = x^\mu + \delta x^\mu$; therefore:

$$J(x^\mu) = \delta^\mu{}_\nu + \partial_\nu \delta x^\mu. \tag{2.96}$$

We then define the matrix $M^\mu{}_\nu = \delta^\mu{}_\nu + \partial_\nu\delta x^\mu$. Exploiting the Levi-Civita symbol $\epsilon^{\mu\nu\rho\sigma}$, we can write

$$detM = \epsilon^{\mu\nu\rho\sigma} M^0{}_\mu M^1{}_\nu M^2{}_\rho M^3{}_\sigma. \tag{2.97}$$

Computing this determinant we find:

$$\begin{aligned}
detM &= \epsilon^{\mu\nu\rho\sigma} M^0{}_\mu M^1{}_\nu M^2{}_\rho M^3{}_\sigma \\
&= \epsilon^{\mu\nu\rho\sigma}(\delta^0{}_\mu + \partial_\mu \delta x^0)(\delta^1{}_\nu + \partial_\nu \delta x^1)(\delta^2{}_\rho + \partial_\rho \delta x^2)(\delta^3{}_\sigma + \partial_\sigma \delta x^3) \\
&= \epsilon^{0123}(\delta^0{}_0 + \partial_0 \delta x^0)(\delta^1{}_1 + \partial_1 \delta x^1)(\delta^2{}_2 + \partial_2 \delta x^2)(\delta^3{}_3 + \partial_3 \delta x^3) + o((\delta x)^2) \\
&\approx (1 + \partial_0 \delta x^0)(1 + \partial_1 \delta x^1)(1 + \partial_2 \delta x^2)(1 + \partial_3 \delta x^3) \\
&= (1 + \partial_0 \delta x^0 + \partial_1 \delta x^1 + \partial_0 \delta x^0 \partial_1 \delta x^1)(1 + \partial_2 \delta x^2 + \partial_3 \delta x^3 + \partial_2 \delta x^2 \partial_3 \delta x^3) \\
&= [1 + \partial_0 \delta x^0 + \partial_1 \delta x^1 + o((\delta x)^2)][1 + \partial_2 \delta x^2 + \partial_3 \delta x^3 + o((\delta x)^2)] \\
&= 1 + \partial_0 \delta x^0 + \partial_1 \delta x^1 + \partial_2 \delta x^2 + \partial_3 \delta x^3 + o((\delta x)^2) \\
&\approx 1 + \partial_\mu \delta x^\mu.
\end{aligned} \tag{2.98}$$

Therefore $\delta(d^4x) = (\partial_\mu \delta x^\mu) d^4x$. We see that in the third equality, putting μ, ν, ρ, σ different from $0, 1, 2, 3$ respectively, we obtain a product of infinitesimal quantities we assume we can neglect, since it is higher than first order. We can finally write the variation of the action integral in the following form:

$$\begin{aligned}
\delta_s I &= \int d^4x \left\{ [\Phi]\delta_s \phi + \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta_s \phi \right] + \delta x^\mu (\partial_\mu \mathcal{L}) + \mathcal{L} \partial_\mu \delta x^\mu \right\} \\
&= \int d^4x \left\{ [\Phi]\delta_s \phi + \partial_\mu \left[\mathcal{L} \delta x^\mu + \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta_s \phi \right] \right\} = 0.
\end{aligned} \tag{2.99}$$

Since this equality must be true for every arbitrary function p_α and all quantities used depends in a differentiable manner from these arbitrary functions, we obtain [14]

$$[\Phi]\delta_s \phi = \partial_\mu B^\mu, \tag{2.101}$$

where we put $-B^\mu = \mathcal{L}\delta x^\mu + \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta_s \phi$. This is the solution to the variational problem for the action integral. We can see that when we can recover the results of Noether's first theorem. Indeed in Noether's first theorem instead of arbitrary functions $p_\alpha(x)$ we have constant parameters ω_α that we can extract from every derivations in x^μ. Therefore if we insert the equations (2.54) in (2.101) we obtain the result (2.52) of Noether's first theorem. We also find the answer to the minus sign that shook our curiosity before.

Step 2. We now prove the second result of the theorem. We assume that the transformations of the group $G_{\infty\eta}$ depend in a differentiable manner not only from the functions

[14] Recall that the dependence of the quantities of the theorem can be put as follow:

$$\delta\psi = \sum_{\alpha=1}^{\eta} \left\{ \frac{\partial \psi}{\partial(\Delta p_\alpha)} \Delta p_\alpha + \frac{\partial \psi}{\partial(\Delta p_{\alpha,\mu})} \partial_\mu (\Delta p_\alpha) \right\}, \tag{2.100}$$

where ψ is either the coordinate x or the function ϕ. Further we explicitly dropped dependence of any quantity from the second derivative of the functions $p_\alpha(x)$ since they are second order terms.

$p_\alpha(x)$ but also from their first derivatives. Therefore the variation $\delta_s\phi$ and the term B^μ are linear in the infinitesimal variations $\Delta p_\alpha(x)$. Thus we can write:

$$\delta_s\phi = \sum_{\alpha=1}^{\eta}\{a_\alpha \Delta p_\alpha + b^\mu_\alpha \partial_\mu[\Delta p_\alpha]\}, \tag{2.102}$$

where

$$\begin{aligned} a_\alpha &= \frac{\partial \phi}{\partial(\Delta p_\alpha)} \\ b^\mu_\alpha &= \frac{\partial \phi}{\partial(\Delta p_{\alpha,\mu})} \end{aligned} \tag{2.103}$$

We then exploit the following identity:

$$[\Phi]b^\mu_\alpha \partial_\mu(\Delta p) = \partial_\mu([\Phi]b^\mu_\alpha \Delta p_\alpha) - \partial_\mu([\Phi]b^\mu_\alpha)\Delta p_\alpha = \partial_\mu A^\mu - \partial_\mu([\Phi]b^\mu_\alpha)\Delta p_\alpha, \tag{2.104}$$

where we put $A^\mu = ([\Phi]b^\mu_\alpha \Delta p_\alpha)$ and we use the Einstein convention for summation on repeated index. Therefore we have:

$$\sum_\alpha \{a_\alpha[\Phi] - \partial_\mu([\Phi]b^\mu_\alpha)\}\Delta p_\alpha = \partial_\mu(B^\mu - A^\mu). \tag{2.105}$$

Now we aim to exploit the arbitrariness of the infinitesimal variations of the variables:

- We choose δx^μ so that they vanish at the bound of the domain.
- We choose Δp_α so that they vanish at the bound of the domain as well as their derivatives.

We then perform an integral over the whole space, obtaining:

$$\sum_\alpha \int_{\mathcal{D}} \{a_\alpha[\Phi] - \partial_\mu([\Phi]b^\mu_\alpha)\}\Delta p_\alpha d^4x = \int_{\mathcal{D}} \partial_\mu(B^\mu - A^\mu)d^4x. \tag{2.106}$$

We also write A^μ taking account of(2.102):

$$A^\mu = \mathcal{L}\delta x^\mu + \frac{\mathcal{L}}{\partial\phi_{,\mu}}\sum_\alpha \{a_\alpha \Delta p_\alpha + b^\mu_\alpha(\partial^\mu \Delta p_\alpha)\} \tag{2.107}$$

in order to understand the role the hypothesis assumed play. Indeed:

1. Since we assume the variations vanish at the bound of the domain, we can exploit Green's theorem in (2.106) and evaluate the integrand at the borders, where it vanishes.

2. We are thus left with

$$\sum_\alpha \int_{\mathcal{D}} \{a_\alpha[\Phi] - \partial_\mu([\Phi]b^\mu_\alpha)\}\Delta p_\alpha d^4x = 0 \tag{2.108}$$

We recall that the quantities Δp_α are arbitrary functions. Therefore [15] the term in braces identically vanishes. In other words, we found the η relations we were looking for:

$$[\Phi]a_\alpha = \partial_\mu\{[\Phi]b^\mu_\alpha\} \Rightarrow \sum_{\rho=1}^{N}[\Phi]_\rho a_{\alpha-\rho} = \sum_{\rho=1}^{N}\partial_\mu\{[\Phi]_\rho b^\mu_{\alpha-\rho}\} \tag{2.109}$$

where in the last equation we wrote the result for an arbitrary number N of fields. Therefore to every parameter α of the group $G_{\infty\eta}$ we can extract a conserved quantity without requiring equations of motion to be satisfied. □

2.2.3 Energy-momentum tensor

Noether's theorems provide also an analogous equation for field theory to the Jacobi's integral for discrete systems. Inspecting this conservation equation we can extract a Noether's conserved current.

To begin with, let us take the total derivative of $\mathcal{L}$ with respect to x^μ instead of the total time derivative of L with respect to t; we have [16]

$$\frac{d\mathcal{L}}{dx^\mu} = \frac{\partial\mathcal{L}}{\partial\phi_\rho}\phi_{\rho,\mu} + \frac{\partial\mathcal{L}}{\phi_{\rho,\nu}}\phi_{\rho,\nu\mu} + \frac{\partial\mathcal{L}}{\partial x^\mu}. \tag{2.110}$$

Using the equations of motion (2.44), this can be expressed as

$$\begin{aligned}\frac{d\mathcal{L}}{dx^\mu} &= \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial\phi_{\rho,\nu}}\right)\phi_{\rho,\mu} + \frac{\partial\mathcal{L}}{\partial\phi_{\rho,\nu}}\phi_{\rho,\mu\nu} + \frac{\partial\mathcal{L}}{\partial x^\mu}\\ &= \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial\phi_{\rho,\nu}}\right) + \frac{\partial\mathcal{L}}{\partial x^\mu},\end{aligned} \tag{2.111}$$

where we assumed that we can change the order of derivation in the second derivative of the field. This can be written as:

$$\partial_\nu\left[\frac{\partial\mathcal{L}}{\partial\phi_{\rho,\nu}}\phi_{\rho,\mu} - \mathcal{L}\delta_{\mu\nu}\right] = -\frac{\partial\mathcal{L}}{\partial x^\mu}. \tag{2.112}$$

[15] This is the so called Du Bois-Reymond Lemma.

[16] As it could be guessed, $\phi_{\rho,\nu\mu}$ stands for $\frac{\partial^2\phi_\rho}{\partial x^\nu\partial x^\mu}$.

Whenever the lagrangian density does not depend explicitly upon x^μ, the quantity in square brackets is conserved. Indeed, as we are about to show, it is a Noether's current. First we should notice that this quantity is a tensor, called energy-momentum tensor:

$$T^\nu{}_\mu = \frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \phi_{\rho,\mu} - \mathcal{L} \delta^\nu{}_\mu, \tag{2.113}$$

In particular, since we deal with conservation in time when we refer to Noether's currents, we should write the conservation equation (2.112) as a continuity equation, which, after splitting its space and time components and if we assume that the lagrangian density does not depend explicitly upon x^μ, can be written as:

$$T^\nu{}_{\mu,\nu} = \frac{\partial T^0{}_\mu}{\partial t} + \frac{\partial T^i{}_\mu}{\partial x^i} = \frac{\partial T^0{}_\mu}{\partial t} + \nabla \cdot T^i{}_\mu = 0. \tag{2.114}$$

So that in this way we can find charges P_μ (the well known energy and momentum four vector) conserved in time [17]

$$P_\mu = \int T^0{}_\mu d^3x. \tag{2.115}$$

Since it represents the conservation of four independent currents, the energy-momentum tensor must be associated to some symmetry in Noether's theorem. Indeed, it can be defined as the conserved current under constant spacetime translations [18], which actually forces the Lagrangian to be invariant under this transformations. In other words, constant space-time translations must be a symmetry for the action considered. Since we are dealing with constant translations, the current derived must be linear with respect to them and thus can be written as

$$J^\mu = \epsilon^\nu T^\mu{}_\nu \tag{2.116}$$

[17] We stress that what is actually yielded by the continuity equation (2.114) is the conservation of the density of a physical property.

[18] Energy-momentum tensor is better defined as $T_{\mu\nu} = \frac{1}{\sqrt{g}} \frac{\delta S}{\delta g^{\mu\nu}}$, where $g_{\mu\nu}$ is the metric of the manifold considered. However this is not our topic and for our purpose the definition given above produces the same results; in fact the different definition is only used when the Noether's current associated with the space-time translation symmetry is not symmetric, and this happens only in the presence of a spin source. This issue is recovered by introducing the Belinfante's tensor (we will used this result in the next chapter inspecting Einstein-Hilbert action). Just to be as clear as possible, we recall that the physical meaning of the energy-momentum tensor is the following: $T^{\mu\nu}$ represents the flux of the μ component of the four-momentum in special relativity frame across a surface of constant x^ν. Therefore:

$$\begin{array}{ll} T^{00} & energy\ \ density \\ T^{0i} & energy\ \ flux\ \ across\ \ x^i\ \ surface \\ T^{i0} & i\ \ momentum\ \ density \\ T^{ij} & flux\ \ of\ \ i\ \ momentum\ \ across\ \ x^j\ \ surface \end{array}$$

where the coefficient $T^\mu{}_\nu$ of the current are the components of the energy momentum tensor, conserved since the current J^μ is. We now aim to examine two cases in particular and to derive the relative energy-momentum tensor.

Scalar fields

If the Lagrangian does not depend on x^μ explicitly, then it is invariant under spacetime constant translations, namely

$$\delta_s \phi(x) = -\epsilon^\mu \partial_\mu \phi(x), \tag{2.117}$$

which, as shown earlier, is its Lie derivative under constant translations. Computing the variations, we obtain:

$$\begin{aligned} \delta_s I[\phi(x)] &= \frac{\partial \mathcal{L}}{\partial \phi_\mu} \delta_s \phi_\mu + \frac{\partial \mathcal{L}}{\partial \phi_{\mu,\rho}} \delta_s \phi_{\mu,\rho} \\ &= -\epsilon^\nu \left[\frac{\partial \mathcal{L}}{\partial \phi_\mu} \phi_{\mu,\nu} + \frac{\partial \mathcal{L}}{\partial \phi_{\mu,\rho}} \phi_{\mu,\rho\nu} \right] = -\partial_\nu (\epsilon^\nu \mathcal{L}), \end{aligned} \tag{2.118}$$

which is actually a boundary term. We note that we exploit the independence of the lagrangian density from x^μ explicitly. Here we are: the conserved Noether's current is

$$J^\mu = \epsilon^\rho \left[\frac{\partial \mathcal{L}}{\partial \phi_{\mu,\nu}} \phi_{\mu,\nu\rho} - \delta^\mu{}_\rho \mathcal{L} \right] \equiv \epsilon^\rho T^\mu{}_\rho, \tag{2.119}$$

from which we derive the same energy-momentum tensor (2.113) computed above.

Maxwell's electromagnetism

The action in electromagnetism is:

$$I[A_\mu(x)] = -\frac{1}{4} \int F_{\mu\nu} F^{\mu\nu} d^4x, \tag{2.120}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This theory is invariant under spacetime translations $x^\mu \to x^\mu + \epsilon^\mu$ with constant ϵ^μ from which we can compute its associated energy-momentum tensor. Furthermore it is also invariant under gauge transformations $\delta A_\mu = \partial_\mu \lambda(x)$; hence we can compute an improved energy-momentum tensor, called Belinfante tensor, which looks pretty nice: it is symmetric and gauge invariant. This property suggests that this action is invariant under a larger group.

The variation of a one-form under constant translation is computed through its Lie derivative, namely $\delta_{in} A_\mu = -\epsilon^\nu \partial_\nu A_\mu$, where we use the subscript in δ_{in} because we are about to improve it with the additional invariance of Maxwell's theory (the subscript should mean "initial"). This variation is a symmetry, since

$$\begin{aligned}\delta_{in} I[A_\mu] &= \int \delta_{in} F^2 d^4x \\ &= \int 2F^{\mu\nu}\delta_{\mu\nu} d^4x \\ &= \int 2F^{\mu\nu}\epsilon^\alpha(\partial_\mu\partial_\nu A_\alpha - \partial_\nu\partial_\mu A_\alpha) d^4x\end{aligned} \tag{2.121}$$

is a boundary term due to the sufficient continuity requirements on the second derivatives of A_μ we can impose without any difficulty. However we immediately notice that this variation is not gauge invariant under Lorentz transformations; therefore the current we could obtain does not behave nicely under Lorentz transformations as it is not gauge invariant. Hence we exploit the extended symmetry of the action to write a better variation which is a combination of a constant spacetime translation together with a particular gauge transformation:

$$\delta_s A_\mu = -\epsilon^\alpha\partial_\alpha A_\mu + \partial_\mu(\epsilon^\alpha A_\alpha) = F_{\mu\alpha}\epsilon^\alpha, \tag{2.122}$$

where ϵ^α is constant. This variation is invariant under Lorentz transformations since the derivatives of A_μ are involved only through the tensor $F_{\mu\nu}$.

In order to find the energy-momentum tensor we compute the variation of the action under this variation [19] :

$$\begin{aligned}\delta_s F^2 &= 2F^{\mu\nu}\delta_s F_{\mu\nu} = 2F^{\mu\nu}\epsilon^\alpha(\partial_\mu F_{\nu\alpha} - \partial_\nu F_{\mu\alpha}) = 2F^{\mu\nu}\epsilon^\alpha(\partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu}) \\ &= -2F^{\mu\nu}\epsilon^\alpha\partial_\alpha F_{\mu\nu} = \partial_\alpha(-\epsilon^\alpha F^2) = \partial_\alpha(-\epsilon^\alpha\mathcal{L}),\end{aligned} \tag{2.123}$$

which is a boundary term. Hence the variation (2.122) is a symmetry for the action. We stress that we used Bianchi's identity for the third equivalence, namely $\partial_\mu F_{\nu\lambda} + \partial_\lambda F_{\mu\nu} + \partial_\nu F_{\lambda\mu} = 0$.

Therefore the conserved Noether's current is

$$\begin{aligned}J^\mu &= \frac{\partial\mathcal{L}}{\partial A_{\rho\mu}}\delta A_\rho - K^\mu = -F^{\mu\rho}\epsilon^\sigma F_{\rho\sigma} + \epsilon^\mu\mathcal{L} \\ &= \epsilon^\sigma[-F^{\mu\rho}F_{\rho\sigma} + \delta^\mu{}_\sigma\mathcal{L}] = -\epsilon^\sigma\left[F^{\mu\rho}F_{\rho\sigma} + \frac{1}{4}\delta^\mu{}_\sigma F^{\alpha\beta}F_{\alpha\beta}\right]\end{aligned} \tag{2.124}$$

from which we derive the electromagnetic energy-momentum tensor

$$T^\mu{}_\sigma = -F^{\mu\rho}F_{\sigma\rho} + \frac{1}{4}\delta^\mu{}_\sigma F^{\alpha\beta}F_{\alpha\beta}. \tag{2.125}$$

Since $F_{\mu\nu}$ is Lorentz covariant, this tensor also is. Moreover it has zero trace, as it can be showed by direct computation or by the argument we are about to provide below.

[19] Recall that $\delta I[A_\mu] = \delta\int F^2 d^4x = \int \delta F^2 d^4x$.

2.2.4 Further inspections

We aim to inspect two particular actions in order to underline some interesting applications provided by Noether's theorems.

Maxwell's theory, once again

We now focus once again on the action of Maxwell's electrodynamics. Exploiting its wider group of invariance we would like to try to extend the variation (2.122) from constant ϵ^μ to arbitrary $\zeta^\mu(x)$, i.e. we combine the transformation $x'^\mu \to x^\mu + \zeta^\mu(x)$ with a particular gauge transformation. Thus we compute the Lie derivative $\delta A_\mu(x)$ and then add to it a gauge parameter to cast it in a similar form to (2.122), that is

$$\begin{aligned}\delta_s A_\mu = \delta A_\mu &= -\zeta^\nu \partial_\nu A_\mu - \delta_\mu \zeta^\nu A_\nu + \partial_\mu(\zeta^\nu A_\nu) \\ &= F_{\mu\nu}\zeta^\nu.\end{aligned} \tag{2.126}$$

This choice could be argued since adding an extra term may vary the value of the current which should be conserved. However, as we will see in the next chapter, gauge transformations are better related to constraints and the charges associated to them are always zero. Therefore in adding a gauge transformation we do not alter the conserved current.

However, Maxwell's theory is not invariant under an arbitrary $\zeta^\mu(x)$ [20], but only to vector fields satisfying the equation

$$\zeta_{\mu,\nu} + \zeta_{\nu,\mu} = \frac{1}{2}\eta_{\mu\nu}\zeta^\rho_{,\rho}, \tag{2.127}$$

where $\eta_{\mu\nu}$ is the $(3+1)$ Minkowski's metric. These infinitesimal vector fields represent a particular group of transformations called the *conformal group*. Any vector fields $\zeta^\mu(x)$ that belong to it is a symmetry of the action. We derive this result and its associated currents computing them directly.

Our aim is to write the variation of the Lagrangian as $\delta\mathcal{L} = \partial_\mu K^\mu + f(\zeta)$, for some function f which depends on ζ but *does not depend on the field* A_μ, since we do not want to change the dynamical fields but only to impose some restriction over the kind of transformations of coordinates involved (namely, that they satisfy (2.127)). Hence we

[20] We add that invariance under a general $\zeta^\mu(x)$ is the symmetry of general relativity.

have:

$$
\begin{aligned}
\delta_s\mathcal{L} &= \delta_s(\frac{1}{4}F^2) \\
&= F^{\mu\nu}\partial_\mu(\zeta^\rho F_{\nu\rho}) \\
&= F^{\mu\nu}F_{\nu\rho}\partial_\mu\zeta^\rho + F^{\mu\nu}\zeta^\rho\partial_\mu F_{\nu\rho} \\
&= F^{\mu\nu}F_\nu{}^\rho\partial_\mu\zeta_\rho + \frac{1}{2}F^{\mu\nu}\zeta^\rho(\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu}) \\
&= -F^{\mu\nu}F^\rho{}_\nu\partial_\mu\zeta_\rho - \frac{1}{2}F^{\mu\nu}\zeta^\rho\partial_\rho F_{\mu\nu} \\
&= -F^{\mu\nu}F^\rho{}_\nu\partial_\mu\zeta_\rho - \frac{1}{4}\partial_\rho(F^{\mu\nu}F_{\mu\nu})\zeta^\rho \\
&= -\frac{1}{2}F^{\mu\nu}F^\rho{}_\nu(\partial_\mu\zeta_\rho + \partial_\rho\zeta_\mu) + \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\partial_\alpha\zeta^\alpha - \frac{1}{4}\partial_\rho(\zeta^\rho F^2) \\
&= -\frac{1}{2}F^{\mu\nu}F^\rho{}_\nu\bigg(\partial_\mu\zeta_\rho + \partial_\rho\zeta_\mu - \frac{1}{2}\eta_{\mu\rho}\partial_\alpha\zeta^\alpha\bigg) - \partial_\rho(\zeta^\rho\mathcal{L}),
\end{aligned}
\tag{2.128}
$$

where we exploited Bianchi's identity in manipuLating the tensor $\partial_\lambda F_{\mu\nu}$ and the anti-symmetry of the tensor $F_{\mu\nu}$. Therefore, when the vector field ζ^μ solves (2.127), (2.126) is a symmetry of Maxwell's action, since $-\partial_\rho(\zeta^\rho\mathcal{L})$ is a boundary term. We then find the Noether's current associated to it, by only applying Noether's procedure:

$$
\begin{aligned}
J^\rho &= \bigg(\frac{\partial\mathcal{L}}{\partial A_{\mu,\rho}}\bigg)\delta_s A_\mu - K^\mu \\
&= F^{\rho\alpha}\zeta^\beta F_{\alpha\beta} + \frac{1}{4}\zeta^\rho F^2 \\
&= \zeta^\beta\bigg(F^{\rho\alpha}F_{\alpha\beta} + \frac{1}{4}\delta^\rho{}_\beta F^2\bigg),
\end{aligned}
\tag{2.129}
$$

with $\partial_\rho J^\rho = 0$ and ζ^μ satisfying (2.127).

Finally, to provide a brief overview, we list the possible vector fields that solves (2.127):

- *Constant translations (4 generators).* This is the simplest solution, namely

$$\zeta^\mu(x) = \zeta_0^\mu. \tag{2.130}$$

- *Lorentz transformations (6 generators).*

$$\zeta^\mu(x) = \epsilon^\mu{}_\nu x^\nu, \tag{2.131}$$

  where $\epsilon^\mu{}_\nu$ are constant parameters such that $\epsilon^{\mu\nu} = -\epsilon^{\mu\nu}$.

- *Dilatations (1 generators).* A constant rescaling by λ, namely

$$\zeta^\mu(x) = \lambda x^\mu. \tag{2.132}$$

- *Special conformal transformations (4 generators).*

$$\zeta^\mu(x) = 2x^\nu b_\nu x^\mu - b^\mu x^\nu x_\nu, \tag{2.133}$$

  where b^μ are constant parameters.

(2.127) is an equation whose solutions belong to the four group listed above. It is called conformal Killing's equation and its solutions conformal Killing's vector fields. Physically, they represent curves along which the Lie derivative of the quantities involved in the action combine as to keep the variation of the action equal to zero.

Schrödinger's equation
We now aim to focus on Schrödinger's equation inspecting the action it is derived from, and see what physical meaning we could attribute to the conserved charge Noether's theorem provides.

The action we are looking for can be found in field theory and written as

$$I[\psi(x), \psi^*(x)] = \int dt \int d^3x \Big(i\hbar\psi^*\dot{\psi} - \frac{\hbar^2}{2m}\nabla\psi^* \cdot \nabla\psi - V(x^i)\psi^*\psi \Big). \tag{2.134}$$

Taking the extreme value of the action we find the equations of motion. Before going through direct computations we notice that since ψ is complex, ψ and ψ^* can be thought as independent and are to be considered so. Varying with respect to ψ^* we obtain:

$$\begin{aligned} \delta I[\psi(x), \psi^*(x)] &= \int dt \int d^3x \Big[i\hbar\delta\psi^*\dot{\psi} - \frac{\hbar^2}{2m}\nabla\delta\psi^* \cdot \nabla\psi - V\delta\psi^*\psi \Big] \\ &= \int dt \int d^3x \delta\psi^* \Big[i\hbar\dot{\psi} + \frac{\hbar^2}{2m}\nabla^2\psi - V\psi \Big] \\ &\quad - \frac{\hbar^2}{2m} \int dt \int_{\partial\Omega} d^3x (\delta\psi^*(\nabla\psi \cdot \vec{n})), \end{aligned} \tag{2.135}$$

where we exploit the first Green's identity on the function $\nabla\delta\psi^* \cdot \nabla\psi$, and $\partial\Omega$ is the surface that bounds the volume of integration with normal outward direction $\vec{n}$. If we assume that the function ψ evaluated at spacial infinite on the bound of the whole 4-space vanishes, for arbitrary variations of ψ^* we have

$$i\hbar\dot{\psi} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = H\psi, \tag{2.136}$$

that is: the Schrödinger's equation. Proceeding in the same way for the variations of $\psi(x)$ we obtain the conjugate of the Schrödinger's equation (2.136). (2.134) is clearly invariant under phase transformation for constant α:

$$\psi \to \psi e^{i\alpha}. \tag{2.137}$$

Let us find the associated Noether's charge. First we need to write it as a suitable infinitesimal transformation up to leading order in α:

$$\begin{aligned} \delta_s\psi &= i\alpha\psi \\ \delta_s\psi^* &= -i\alpha\psi^*. \end{aligned} \tag{2.138}$$

It is straightforward to see that Schrödinger's action is strictly invariant under these transformations (i.e. its boundary term vanishes). Hence Noether's charge follows only from the on-shell variation, by replacing the variation with the symmetry, namely [21]

$$\begin{aligned} \delta_{s,o}I[\psi(x),\psi^*(x)] &= \int dt \int d^3x \frac{\partial\mathcal{L}}{\partial\psi_{,\mu}}\delta_s\psi_{,\mu} \\ &= \int dt \int d^3x i\hbar\psi^*\delta_s\dot{\psi} \\ &= \int dt \int d^3x \frac{i\hbar}{2}\frac{d}{dt}(\psi^*\delta_s\psi) \\ &= -\frac{\alpha\hbar}{2}\int dt\frac{d}{dt}\int d^3x(\psi^*\psi). \end{aligned} \tag{2.139}$$

where with $\delta_{s,o}$ we mean that we are taking a symmetry variation for the action on the equations of motion. The associated Noether's charge is thus

$$Q = -\int d^3x\,\frac{\hbar}{2}\psi^*\psi = -\int d^3x\,\frac{\hbar}{2}\rho\,, \tag{2.140}$$

or

$$Q = \int d^3x\,\psi^*\psi = \int d^3x\,|\psi|^2 = \int d^3x\,\rho\,, \tag{2.141}$$

that is the total probability of finding the particle in space. This is of great importance: since this quantity is conserved, we can fix its value to be equal to 1 as we normally does in quantum mechanics and thus refer to it as a probability. We stress that here we considered only those functions ψ which vanish sufficiently rapidly at infinity; the same argument can be carried out with non vanishing functions on a finite border obtaining an analogous equation of conservation of a certain charge, turning out in defining a

[21] Recall that $\frac{d}{dt}(\psi^*\psi) = \dot{\psi}^*\psi + \psi^*\dot{\psi} = 2Re[\psi^*\dot{\psi}]$.

probability and a current of probability. Indeed in computing the on shell variation we are left with a boundary term that we do not neglect this time:

$$\int dt \int_{\partial\Omega} d^3x[-\frac{\hbar^2}{2m}\delta\psi^*(\nabla\psi\cdot\vec{n})]\,; \tag{2.142}$$

using the symmetry provided above and via Gauss theorem, it can be cast in the form

$$\frac{i\hbar^2}{2m}\int dt \int d^3x \nabla\cdot[\psi^*(\nabla\psi)]. \tag{2.143}$$

We can thus interpret the quantity $\vec{j}=\frac{i\hbar}{m}\psi^*(\nabla\psi)$ as a probability current; the coefficient adopted will find an explanation in the following operation. Indeed Noether's theorem thus implies, using (2.139) and (2.140):

$$\frac{\partial\rho}{\partial t}+\nabla\cdot\vec{j}=0, \tag{2.144}$$

that is: an equation for the conservation of probability density in a finite volume along the time evolution of the system. We finally notice that, since ρ is a real quantity and Noether's theorem is proved to be true, the probability current is better defined as $\vec{j}=\frac{\hbar}{m}Im[\psi^*(\nabla\psi)]$.

Chapter 3

Gauge theories in hamiltonian formulation

Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty and perfection.

Hermann Weyl

3.1 Introduction

All three fundamental interactions (electroweak theory, strong interaction and the theory of gravity) are described by Lagrangian possessing a gauge invariance. Therefore it is clear how the idea of gauge invariance is fundamental in developing physical models. Moreover gauge theories provide a fulfilling theoretical interpretations of nature: all is ruled by symmetries; that is: symmetries furnish a way to unveil the underground correspondences and find the fundamental laws of interactions. Further all gauge theories can be understood within the general hamiltonian structure we provide in this chapter. This stands as a leading flag in the developing of a theory of great conceptual clarity that is able to trace common features in nature.

Gauge theories share common aspects. First, they involve a gauge symmetry, namely a transformation containing *arbitrary* functions of spacetime that leaves the variation of the action integral equal to zero, at least up to a boundary term. Therefore it follows readily that the equations of motion must be related among each other in a certain way; thus the equations of motion do not empty all degrees of freedom of the action integral.

Hence (at least in a Hamiltonian description) a gauge theory has *constraints* which act as the generators of the corresponding gauge symmetries. We are about to be more precise and meticulous in the following sections. The key is that we are dealing with theories that seem to miss something, since their Lagrangians furnish less equations that unkowns but that connect their dynamical fields someway. This strange behavior suggests that we should exploit this particular attitude and find an elegant way to tune the physical reality in a suitable mathematical model.

At first sight, these theories may look a bit presumptuous and not so useful, since theories whose equations of motion do not fully determine the evolution of their variables do not sound very physical and are intuitively perceived as pathological. However, digging deeper, the great discovery of gauge symmetries uncovered the concept of equivalent class of configurations: fields that may differ in their mathematical presentation but represent the *same physical reality*. For instance Maxwell's equations, which fully described the behavior of electromagnetic field, are invariant by adding the derivative of an *arbitrary* scalar function. Therefore, although mathematically they are two distinct and different functions, physically they furnish the same fields and model the same physical reality. Historically, Maxwell's theory has been the first gauge theory discovered.

3.1.1 A quick insight

We should first provide the rigorous definition and description of gauge theories. Sometimes this turns out to not to be the more intuitive approach. Indeed, gauge theories tend to be too much linked to a mathematical trick than to physical considerations. They are built first by looking at the mathematics beyond Lagrangians and then through interpreting the results in a physical manner. Therefore we aim to underline gauge features with a simple standard action integral, waiting for the next section to provide their complete description.

We then begin with the following action:

$$I[A_0(t), \psi(t)] = \frac{1}{2} \int dt (\dot{\psi} - A_0)^2. \tag{3.1}$$

Looking at this simple mechanical model, we find all the main characteristics of a gauge theory.

1. *Gauge symmetry:* the action is invariant under the transformations

$$\psi \to \psi + \epsilon(t) \qquad and \qquad A_0 \to A_0 + \dot{\epsilon}(t), \tag{3.2}$$

   where $\epsilon(t)$ is an *arbitrary* function of time. We should call it gauge symmetry, in order to distinguish it from global or Noether's symmetries. This gauge symmetry imposes the following properties.

2. *Non independence of the equations of motion:* when computing the variation with respect to these transformations we obtain the equations of motion, namely:

$$\begin{aligned} \frac{d}{dt}\frac{\partial L}{\partial \dot{A}_0} - \frac{\partial L}{\partial A_0} &= (\dot{\psi} - A_0) = 0 \\ \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{\psi}} - \frac{\partial L}{\partial \psi} &= \frac{d}{dt}(\dot{\psi} - A_0) = 0, \end{aligned} \tag{3.3}$$

hence we can see that the equation for the field A_0 already contains the equation for ψ. Therefore the equations are not all independent: there is only one equation, not two. Hence the solution contains arbitrary functions. We stress once again that we have less equations than unknowns.

3. *The general solution contains arbitrary functions:* the solution of the equations of motion is

$$\psi(t) = f(t), \qquad A_0(t) = \dot{f(t)}, \tag{3.4}$$

and contains an arbitrary function of time. Given any initial condition, one can always modify the evolution of the system at later times.

4. *The Hamiltonian obeys constraints:* we shall first write the Hamiltonian. Once computed the momenta $p_{A_0} = \frac{\partial L}{\partial \dot{A}_0} = 0$ and $p_\psi = \frac{\partial L}{\partial \dot{\psi}} = \dot{\psi} - A_0$, we can write it as

$$\begin{aligned} H(p_\psi(t), \psi(t), A_0(t)) &= p_\psi \dot{\psi} - L \\ &= \frac{1}{2}p_\psi^2 + A_0 p_\psi. \end{aligned} \tag{3.5}$$

The corresponding action in hamiltonian form is

$$I[p_\psi(t), \psi(t), A_0(t)] = \int \left(p_\psi \dot{\psi} - \frac{1}{2}p_\psi^2 - A_0 p_\psi \right) dt. \tag{3.6}$$

The equations of motion (Hamilton's equations in this formalism) for the ψ, p_ψ variables are:

$$\begin{aligned} \dot{\psi} &= \frac{\partial H}{\partial p_\psi} = p_\psi + A_0 \\ \dot{p_\psi} &= -\frac{\partial H}{\partial \psi} = 0, \end{aligned} \tag{3.7}$$

while for the A_0 variable looks a bit different. Indeed this variable appears as a *Lagrange's multiplier* in the Hamiltonian and its associated equation is an example of *constraint*

$$\frac{\partial H}{\partial A_0} = p_\psi = 0, \tag{3.8}$$

that is an equation involving no time derivatives.

There is one last thing to underline, linked to gauge fixing in the arbitrariness of the function of time contained in the solution of the equations of motion. We said that the equation of motion of ψ is already contained in the equation for A_0, but not the other way round. This means that one can dispose of $\psi(t)$ and no information will be lost because its equation is already there. Conversely, one cannot dispose of $A_0(t)$. To understand the reason why a parallel treatment is not possible let us consider the following argument.

If we fix the gauge $\psi(t) = 0$ in the action we have:

$$I_{\psi=0}[A(t)] = \int dt\, A_0^2 \to A_0(t) = 0, \tag{3.9}$$

where on the right hand side of the arrow we reported the equation of motion which the action considered lead to. We uniquely fixed $A_0(t)$, and showed that the pure-gauge action (3.1) in fact has no degrees of freedom at all. On the other hand, let us now fix $A_0(t) = 0$ instead; we find an action of the form

$$I_{A_0=0}[\psi(t)] = \int dt\dot{\psi}^2 \to \ddot{\psi} = 0, \tag{3.10}$$

which represent a free field $\psi(t)$ carrying one degree of freedom. This problem is quite subtle and the answer could be find in whether a variable is involved in the hamiltonian's action as a true variable or as a lagrangian's multiplier. Let us be precise in next sections.

3.1.2 Gauge theories and Noether's theorems

Although we only aim to deal with a particular formalism of gauge theories (the hamiltonian form), we would like to underline the relations between gauge theories and Noether's theorems. We remind to sections (2.1.7) and (2.2.2) for the notation we will use.

Let us assume that the *local* group $G_{\infty\eta}$ has a non trivial *global* subgroup G_η. Further let us assume that there exists a certain subgroup of the local group for which we have $\Delta p_\alpha(x) = \Delta\omega_\alpha$ [1]; therefore we can write equation (2.102) for $\delta_s\phi$ as:

$$\delta_s\phi = \sum_{\alpha=1}^{\eta}\{a_\alpha\Delta\omega_\alpha + b_\alpha^\mu\partial_\mu(\Delta\omega_\alpha)\} = \sum_{\alpha=1}^{\eta} a_\alpha\Delta\omega_\alpha \tag{3.11}$$

If we put this expression for $\delta_s\phi$ in the equation for Noether's first theorem (2.51), we obtain:

$$\sum_\rho [\Phi]_\rho a_{\alpha-\rho} = \partial_\mu j_\alpha^\mu \tag{3.12}$$

[1] That is: a subgroup of infinitesimal arbitrary functions p_α that equals the infinitesimal constant parameters ω_α of the group.

Here we are: if Euler-Lagrange equations are satisfied for all fields, (namely if $[\Phi]_\rho = 0,\ \forall\rho$) we immediately obtain an equation for the conservation of the quantities j^μ_α. However, if instead we assume that for some ρ or all of them we have $[\Phi]_\rho \neq 0$, it is possible to find a conserved quantity as well. Indeed for Noether's second theorem we have

$$\sum_\rho [\Phi]_\rho a_{\alpha-\rho} = \sum_\rho \partial_\mu \{[\Phi]_\rho b^\mu_{\alpha-\rho}\} \tag{3.13}$$

from which it follows that:

$$\partial_\mu j^\mu_\alpha = \sum_\rho \partial_\mu \{[\Phi]_\rho b^\mu_{\alpha-\rho}\} \Longrightarrow \partial_\mu J^\mu_\alpha = 0, \tag{3.14}$$

where

$$J^\mu_\alpha \equiv \left\{ j^\mu_\alpha - \sum_\rho ([\Phi]_\rho b^\mu_{\alpha-\rho}) \right\}. \tag{3.15}$$

This brief treatment shows that, when a local group of transformations has a non trivial global subgroup, it is possible to find some conserved quantities without requiring Euler-Lagrange equations of motion to be satisfied and where arbitrary functions are involved. This conservation laws are called *improper laws of conservation* and we can see the seeds of gauge theories hiding in them. On the other hand, when Euler-Lagrange equations of motions are satisfied we immediately find some conserved quantities which are called *proper conservation laws.*

Since we are about to focus on gauge theories and improper transformations, we would like first to introduce a brief example of a proper transformation. We provide the easiest one, namely we focus on the so called Galileo's group. This is a Lie group of ten parameters (which means that the arbitrary functions of Noether's second theorem are ten constants) to which we expect we can extract ten conserved quantities. The action integral from which we start is the one for a free particle $\int dt\ m\frac{\dot{x}^i\dot{x}^i}{2}$.

Galileo's proper group

<u>*Space translations (3 parameters)*</u>: we choose the transformation of the group to be

$$\delta x^i(t) = a^i(t) \tag{3.16}$$

where $a^i(t)$ is an arbitrary function. An explicit calculation gives

$$\delta S[x] = \int dt\, m\dot{x}^i\dot{a}^i \tag{3.17}$$

The term that multiplies the arbitrary function is the conserved quantity, namely the i-th component of the momentum $p^i = m\dot{x}^i$. To prove it we only have to consider the variation of the action on the paths of motion (i.e. we assume that the Euler-Lagrange equations of motion are satisfied). Indeed the variation now vanishes and after an integration by parts we obtain:

$$0 = \delta_o S[x^i(t)] = \int dt\, p^i(t)\dot{a}^i(t) = -\int dt\, \dot{p}^i(t)\dot{a}^i(t) \Rightarrow \dot{p}^i(t) = 0. \tag{3.18}$$

That is: due to invariance of the action under space translations, momentum is conserved. Further we notice that if we consider $a^i(t)$ to be constant we obtain a symmetry of the action. *Time translation (1 parameter)*: we now consider an arbitrary translation in the time coordinate

$$t \to t' = t - \epsilon(t) \tag{3.19}$$

Computing the variation of the functions $x^i(t)$ we find a result obtained earlier, namely:

$$\delta x^i(t) = \epsilon(t)\dot{x}^i(t) \tag{3.20}$$

The variation of the action integral assumes the following form:

$$\delta S[x] = \int dt\, m\dot{x}^i \partial_t(\epsilon \dot{x}^i) = \int dt\left[\partial_t\left(\frac{\epsilon m}{2}\dot{x}^i\dot{x}^i\right) + \dot{\epsilon}\frac{m}{2}\dot{x}^i\dot{x}^i\right] = \int dt\, \dot{\epsilon}\left(\frac{m}{2}\dot{x}^i\dot{x}^i\right) \tag{3.21}$$

where we neglect total time derivatives, assuming the boundary terms involved vanish. The term $E = \left(\frac{m}{2}\dot{x}^i\dot{x}^i\right)$ is the energy of the free particle. Here we are: if we evaluate the variation of the action integral on the equations of motion $\delta_o S = 0$, we obtain a conserved quantity, that is the energy. Further we notice that if we consider $\epsilon(t)$ to be constant we obtain a symmetry of the action. *Space rotations (3 parameters)*: the variations of the functions $x^i(t)$ under arbitrary space rotations $w^i(t)$ are:

$$\delta x^i(t) = \epsilon^{ijk}\omega^j(t)x^k(t), \tag{3.22}$$

where we used the Levi-Civita symbol ϵ^{ijk}. The variation of the action integral reads

$$\delta S[x] = \int dt\, \dot{\omega}^i \epsilon^{ijk} x^j \dot{x}^k \tag{3.23}$$

Again, if we consider the functions $w^i(t)$ to be constant we obtain a symmetry, while if we evaluate the variation of the action integral on the equations of motion we obtain three conserved quantities, namely the component of the angular momentum $L^i = \epsilon^{ijk}x^j p^k$. *Proper Galilean transformations (3 parameters)*: with proper Galilean transformations we mean transformations between two inertial frames of reference in relative constant motion with constant relative velocity v^i. The transformations of the dynamical variables for arbitrary time dependent velocities $v^i(t)$ are:

$$\delta x^i(t) = v^i(t)t \tag{3.24}$$

Computing the variation of the action integral we find:

$$\delta S[x] = \int dt\, \dot{v}^i (m\dot{x}^i t - m x^i) \tag{3.25}$$

from which we can see that when the functions v^i are constant we obtain a symmetry; on the other hand when we evaluate the variation on the equations of motion ($\delta_o S = 0$) we find three conserved quantities, i.e. $G^i = m\dot{x}^i t - m\dot{x}^i$.

3.2 Gauge theories in hamiltonian form: general structure

Here we should present some of the most important gauge theories. We stress that we deal only with the hamiltonian form and unfortunately we do not provide a straightforward method of finding them. The general method is called the Dirac's procedure and its hard treatment is left to better manuals of a wide literature. However we include these treatments in hamiltonian form in these pages since they arise from Noether's second theorem.

The most important gauge theory actions (i.e. the ones which belong to the fundamental interactions) in hamiltonian formalism assume the following form:

$$\boxed{I[p_i(t), q^i(t), \lambda^a(t)] = \int dt [p_i \dot{q}^i - H_0(p_i, q^i) + \lambda^a \theta_a(p_i, q^i)]}. \tag{3.26}$$

Let us analyze it in detail. The functions $p_i(t), q^i(t), \lambda^a(t)$ are independent fields varied in the action. The total Hamiltonian is composed of two terms: H_0, which denotes the part of the Hamiltonian that is not a constraint; $-\lambda^a\theta_a$, which instead includes the contributions from the constraints θ_a. Further it is important to notice that this action contains no derivatives of the λ ' s. To compute the equations of motion we have to take arbitrary variations with respect to all three fields the action depends from, finding:

$$\begin{aligned} \dot{q}^i &= \frac{\partial H_0}{\partial p_i} - \lambda^a \frac{\partial \theta_a}{\partial p_i}, \\ \dot{p}_i &= -\frac{\partial H_0}{\partial q^i} + \lambda^a \frac{\partial \theta_a}{\partial q^i}, \\ \theta_a(p_i, q^i) &= 0. \end{aligned} \tag{3.27}$$

This system of equations reveals its gauge nature in the presence of the Lagrange's multipliers $\lambda^a(t)$ and the constraints $\theta_a(p_i, q^i)$.

The first two equations determine the evolution of p, q given initial conditions $p(0)$, $q(0)$. However, these initial conditions cannot be totally arbitrary since they must satisfy the equation for θ_a which can be clearly interpreted as constraints; they define an hypersurface in the phase space on which the variables of motion p, q must lie along the evolution of the system. Further, to actually integrate the equations of motion, the expressions of the Lagrange's multipliers $\lambda^a(t)$ have to be found.

Therefore there must be some indication to follow in order to find out a way to chose the functions $\lambda^a(t)$; moreover we have to check that, given initial conditions, the solutions p, q keep satisfying the equation for the constraints θ_a.

These issues can be handled looking at the constraints, which exactly contain the meaning of gauge theories. Using Poisson brackets to compute time derivatives, we obtain:

$$\frac{d}{dt}\theta_a(p_i, q^i) = \{\theta_a, H_0\} - \{\theta_a, \theta_b\}\lambda^b. \tag{3.28}$$

The equation for the constraints θ_a must always hold, which means that the evolution of the system has to follow a path that keep its time derivative equal to zero, i.e.

$$\{\theta_a, H_0\} - \{\theta_a, \theta_b\}\lambda^b \approx 0. \tag{3.29}$$

The symbol $\approx$ (*weakly zero*) implies that we do not require $\dot{\theta}_a$ to be strictly zero everywhere, but it is enough if it vanishes on the surface defined by the constraints, namely when $\theta_a = 0$. Equation (3.29) is a consistency condition for the time evolution of the system. Calling $G_{ab} = \{\theta_a, \theta_b\}$, the following situations may occur:

1. **Non-gauge theories.** When the matrix G_{ab} is invertible, (3.29) fixes completely the functions $\lambda^a(t)$ to the following value

$$\lambda^a(t) = G^{ab}\{\theta_a, H_0\}. \tag{3.30}$$

   Therefore the functions $\lambda^a(t)$ plays the role of preserving the third equation of (3.27), that is it defines the dynamic variables $p(t)$, $q(t)$ in order to keep them on the surface $\theta_a = 0$ along the time evolution of the system.

2. **Gauge theories.** If the matrix G_{ab} is zero, or at least if it is weakly zero (i.e. if it vanishes on the surface $\theta_a = 0 : G_{ab} \approx 0$), then (3.29) imposes no restrictions on the functions $\lambda^a(t)$ (as before, on the surface $\theta_a = 0$) which therefore remained undetermined by the equations of motion. Here lies the first sign of a gauge theory: the equations of motion contain arbitrary functions of time. We are interested in the so called *first class* constraints, namely a set of constraints such that

$$\{\theta_a, H_0\} = G_a{}^b\theta_b \quad , \quad \{\theta_a, \theta_b\} = G_{ab}{}^c\theta_c, \tag{3.31}$$

   which are both weakly zero since the conditions required on G_{ab}. Indeed, if $\theta_a \approx 0$, also $\{\theta_a, \theta_b\} \approx 0$ and $\{\theta_a, H_0\} \approx 0$.

The reason why we pick such conditions is remarkably: if the Hamiltonian and the constraints satisfy (3.31), then the action (3.26) is invariant under the following variations:

$$\begin{aligned} \delta_G q^i &= \{q^i, \theta_a\}\epsilon^a(t) \\ \delta_G p_i &= \{p_i, \theta_a\}\epsilon^a(t) \\ \delta_G \lambda^c &= \dot{\epsilon}^c(t) + \epsilon^a(t) G_a{}^c - \lambda^a \epsilon^b(t) G_{ab}{}^c, \end{aligned} \tag{3.32}$$

where $\epsilon^a(t)$ are arbitrary functions of time and with the symbol δ_G we highlight the particular type of variations considered. Here we are: these transformations are the gauge symmetries of the action. Therefore all actions of the form (3.26), where the Hamiltonian and the constraints satisfy (3.31) have a gauge symmetry (namely (3.32)). This is the reason why we write this variation as δ_G: they are symmetries but of a particular type.

3. **Mixed case.** Although we completely avoid its treatment, we mention it for the sake of completeness. This case endorses sets of constraints such that some of them satisfy (3.31), thus related to gauge symmetries, and some other different conditions related to the non-gauge symmetries as in the first case showed above. [2].

We should now demonstrate the invariance of the action (3.26) under gauge symmetries (3.32). We begin with the variations of the canonical variables:

$$\delta_G q_i = \epsilon^a \frac{\partial \theta_a}{\partial pi} \quad , \quad \delta_G p_i = -\epsilon^a \frac{\partial \theta_a}{\partial q^i}. \tag{3.33}$$

Hence the variation of the action is:

$$\begin{aligned} \delta_G I[p_i(t), q^i(t), \lambda^a(t)] &= \delta_G \int dt (p_i \dot{q}^i - H_0 - \lambda^a \theta_a) \\ &= \int dt \Bigg[-\epsilon^a \frac{\partial \theta_a}{\partial q^i} \dot{q}^i - \dot{p}^i \epsilon^a \frac{\partial \theta_a}{\partial p_i} - \frac{\partial H_0}{\partial q^i} \epsilon^a \frac{\partial \theta_a}{\partial p_i} + \frac{\partial H_0}{\partial p_i} \epsilon^a \frac{\partial \theta_a}{\partial q_i} + \\ &\quad - \delta_G \lambda^a \theta_a - \lambda^a \left(\frac{\partial \theta_a}{\partial q^j} \epsilon^b \frac{\partial \theta_b}{\partial p_j} - \frac{\partial \theta_a}{\partial p_j} \epsilon^b \frac{\partial \theta - b}{\partial q^j} \right) \Bigg] + \int dt \frac{d}{dt}(p_i \delta_G q^i) \\ &= \int dt \left[-\epsilon^a \frac{d}{dt}\theta_a - \epsilon^a [H_0, \theta_a] - \delta_G \lambda^a \theta_a - \lambda^a \epsilon^b [\theta_a, \theta_b] \right] + \int dt \frac{d}{dt}(p_i \delta_G q^i). \end{aligned} \tag{3.34}$$

[2] It is worth to underline that there is a much more mathematical and precise manner to determine gauge theories, that is the one highlighted by Dirac's procedure. However this procedure sometimes turns out to be difficult and tricky and we preferred to focus only on the type of gauge theories mentioned, since they cover almost all fundamental examples. Further, starting with the Hamiltonian we avoided all the tedious computations of the standard method in order to derive them.

If θ_a are all first class constraints (i.e. they satisfy (3.31)), the variation above reduces to:

$$\begin{aligned}
\delta_G I &= \int dt[\dot{\epsilon}^a\theta_a + \epsilon^a G_a{}^b\theta_b - \delta_G\lambda^a\epsilon^b G_{ab}{}^c\theta_c] + \int dt\frac{d}{dt}(p_i\delta_G q^i - \epsilon^a\theta_a) \\
&= \int dt(\dot{\epsilon}^c + \epsilon^a G_a{}^c - \delta_G\lambda^c - \lambda^a\epsilon^b G_{ab}{}^c)\theta_c + \int dt\frac{d}{dt}(p_i\delta_G q^i - \epsilon^a\theta_a) \\
&= \int dt\frac{d}{dt}(p_i\delta_G q^i - \epsilon^a\theta_a).
\end{aligned} \tag{3.35}$$

where we collect in the second integral all the boundary values rising from integration by parts. Some comments regarding this kind of variation are worth to be addressed:

- *No charge:* since the gauge symmetry is a symmetry of the action we can compute its associated Noether's charge, namely:

$$Q = K - p_i\delta_G q^i = p_i\delta_G q^i - \epsilon^a\theta_a - p_i\delta_G q^i = -\epsilon^a\theta_a = 0, \tag{3.36}$$

  since the constraints obey $\theta_a = 0$ at any time [3].

- *Degrees of freedom:* a gauge theory contains $2N$ canonical variables corresponding to the phase space for N degrees of freedom (namely p, q corresponding to $\frac{1}{2}\times 2$ degrees of freedom) plus the additional constraints θ_a. Suppose that subscripts of θ_a runs in the range $a =, \ldots, g$. The equations that the dynamical variables must satisfy $\theta_a(p_i, q^i) = 0$ subtract g of them. Therefore there are g gauge symmetries which implies that not all of the canonical variables p, q are physically meaningful by themselves, since there are some combinations of them that produce some others. Gauge symmetries subtract g physically meaningful variables and the total number of degrees of freedom is $\frac{1}{2}(2N - 2g)$.

- *Time evolution:* in gauge theories the time evolution of the canonical variables is the usual one $\dot{q} = \partial_p H_0$, $\dot{p} = -\partial_q H_0$ plus an additional term $\lambda^a\partial_p\theta_a$, $\lambda^a\partial_q\theta_a$, respectively, which is a gauge transformation with the gauge parameters being the Lagrange's multipliers.

In the next sections we provide some fundamental examples and applications of gauge theories, trying to underline their common features. Our procedure will always be the same: introduce lagrangian form, derive its hamiltonian action, and finally show how the constraints generate the gauge symmetry.

[3] This topic is actually subtler than this, since it turns out that whenever the field considered is a complex function, a gauge charge does exist and is different from zero. Anyway we will deal only with real fields.

3.2.1 Special relativity: point particle

The action for a spinless relativistic point particle parameterized as $x^\mu(\tau)$ is

$$I[x^\mu(\tau)] = -m \int ds = -m \int d\tau \sqrt{-\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\eta_{\mu\nu}}, \tag{3.37}$$

where τ indicates any parameter describing the curve in a Minkowski's space with its well known related metric $\eta_{\mu\nu}$. We immediately find a transformation that leaves the action unchanged, that is: a symmetry. This transformation is a bit nonphysical and actually is a reparameterization:

$$\tau' = \tau'(\tau) \quad , \quad x'^\mu(\tau') = x^\mu(\tau). \tag{3.38}$$

In order to apply Noether's theorem, we aim to express it in term of an infinitesimal transformation of the field x^μ rather than of the parameter (coordinate) τ. Taking $\tau' = \tau + \epsilon(\tau)$, we can write this transformation as $\delta x^\mu(\tau) = -\epsilon(\tau)\dot{x}^\mu$. It is a gauge symmetry as the infinitesimal parameter $\epsilon(\tau)$ is an arbitrary function of τ. Computing explicitly its variation we obtain:

$$\begin{aligned} \delta_s I[x^\mu(\tau)] &= -m \int d\tau \frac{-\dot{x}^\mu \delta\dot{x}_\mu}{(-\dot{x}^2)^{\frac{1}{2}}} \\ &= -m \int d\tau \frac{d}{d\tau}\Big(\epsilon(\tau)(-\dot{x}^2)^{\frac{1}{2}}\Big), \end{aligned} \tag{3.39}$$

where with the superposed dot we indicate a derivative with respect to the parameter τ and we use $\delta\dot{x^\mu} = \frac{d}{dt}(\delta x^\mu) = -\dot{\epsilon}(\tau)\dot{x}^\mu - \epsilon(\tau)\ddot{x}^\mu$. We also stress in the notation that the function $\epsilon(\tau)$ is not constant. The transformation considered is a symmetry, since the variation of the action is the boundary term $K = \epsilon(\tau)\Big(-(\dot{x}^2)^{\frac{1}{2}}\Big) = \epsilon(\tau)L$. Then Noether's conserved charge is:

$$Q = K - \frac{\partial L}{\partial \dot{x}^\mu}\delta_s x^\mu = K - m\frac{\dot{x}_\mu}{(-\dot{x}^2)^{\frac{1}{2}}}\epsilon(\tau)\dot{x}^\mu = K - \epsilon(\tau)L = 0. \tag{3.40}$$

As expected for a gauge theory, the Noether's charge associated to this transformation vanishes. We recognize again that we are dealing with a gauge symmetry. To relate to the hamiltonian formulation of our treatment we must find the hamiltonian action. Further we aim to find it in the form (3.26). However we obtain a momentum

$$p_\nu = \frac{m\dot{x}_\nu}{\sqrt{-\dot{x}^\mu\dot{x}_\mu}}, \tag{3.41}$$

and we are incapable of solving $\dot{x}^\mu$ in terms of p_μ. This issue could be imputed to a random occurrence, but instead is another sign of a gauge theory: (3.41) seem to represent four independent equations although there are only three. Indeed there is a constraint, rising when contracting (3.41):

$$\theta \equiv p_\nu p^\nu + m^2 = 0, \tag{3.42}$$

satisfied without the equations of motion. To find a formulation which suits our general structure we exploit a particular action called the Polyakov action. This action is derived introducing an auxiliary variable, the *einbein* $e(\tau)$ which is treated as a dynamical field. The Polyakov action is defined as

$$I_P[x^\mu(\tau), e(\tau)] = \frac{1}{2}\int \left(\frac{1}{e}\dot{x}^\mu \dot{x}_\mu - em^2\right) d\tau. \tag{3.43}$$

This action is precisely of the form (3.1), taken as an introductory example. Notice that its fundamental property is that it does not contain any derivatives of the einbein, so that we can use its equation of motion as a constant:

$$\frac{\partial L}{\partial e} = 0 \quad \Rightarrow \quad e(x^\mu) = \frac{1}{m}\sqrt{-\dot{x}^\mu}, \tag{3.44}$$

and replace it in the Polyakov action, which gives back the relativistic action (3.37). Therefore these two actions are completely equivalent and generate the same equations of motion but the Polyakov's one possesses a useful advantage: it is quadratic in the field x^μ. Since they are equivalent, the Polyakov action must be invariant under the same reparameterization used above under which the einbein becomes $e'(\tau') = e(\tau)\frac{d\tau}{d\tau'}$. The infinitesimal version of this transformation is:

$$e'(\tau + \epsilon) = e(\tau)(1 - \dot{\epsilon}(\tau)) \quad \Rightarrow \quad \delta e(\tau) = -\frac{d}{d\tau}(\epsilon(\tau)e(\tau)). \tag{3.45}$$

The variation of the Polyakov action is thus:

$$\begin{aligned} I_P[x^\mu(\tau), e(\tau)] &= \frac{1}{2}\int d\tau \left(\frac{2e\dot{x}^\mu \delta\dot{x}_\mu - \dot{x}^2\delta e}{e^2} - m^2\delta e\right) \\ &= -\frac{1}{2}\int d\tau \frac{d}{d\tau}\left(\epsilon\left(\frac{\dot{x}^2}{e} - m^2 e\right)\right), \end{aligned} \tag{3.46}$$

that is, a boundary term $K = -\frac{1}{2}\epsilon(\tau)\left(\frac{\dot{x}^2}{e} - m^2 e\right) = -\epsilon(\tau)L$.

The Hamiltonian for the Polyakov action is easily found to be

$$H(p_\mu, x^\mu, e) = \frac{1}{2}e(p_\mu p^\mu + m^2), \tag{3.47}$$

therefore the corresponding action in hamiltonian formalism reads as:

$$\boxed{I_P[p_\mu(\tau), x^\mu(\tau), e(\tau)] = \int d\tau \left[p_\mu \dot{x}^\mu - \frac{1}{2}e(p_\mu p^\mu + m^2)\right]} \tag{3.48}$$

This action has precisely the form underlined in the general structure:

- Free Hamiltonian: $H_0 = 0$; this fact is due to the gauge invariance: the Hamiltonian is not conjugate to any physical time, because we are using the arbitrary parameter τ to describe the evolution of the system.
- Lagrangian multiplier: e, the einbein.
- Constraint: $\theta = \frac{1}{2}(p^2 + m^2)$, which represents the Einstein's equation for conservation of energy-momentum in natural units. This set of constraints, since it is composed by only one equation, is trivially first class, namely $[\theta, \theta] = 0$.

We can demonstrate that the constraint θ defines the gauge transformation $\delta_G x^\mu$, $\delta_G p_\mu$ through Poisson brackets by direct calculations, and $\delta_G e$ by the formula given in the last line of (3.32). We obtain:

$$\begin{aligned}\delta_G x^\mu &= \{x^\mu, \epsilon(\tau)\frac{1}{2}(p^2 + m^2)\} = \epsilon(\tau)p^\mu \\ \delta_G p_\mu &= \{p_\mu, \epsilon(\tau)\frac{1}{2}(p^2 + m^2)\} = 0 \\ \delta_G e &= \dot{\epsilon}(\tau).\end{aligned} \tag{3.49}$$

The variation of the Polyakov action under this symmetries is therefore:

$$\begin{aligned}\delta_G I_P[p_\mu, x^\mu, e] &= \int d\tau \left[p_\mu \delta_G \dot{x}^\mu - \frac{1}{2}\delta_G e(p^2 + m^2)\right] \\ &= \int d\tau \frac{d}{d\tau}\left[\frac{1}{2}\epsilon(\tau)(p^2 - m^2)\right],\end{aligned} \tag{3.50}$$

that is: the variation of the action under this transformation vanishes up to the boundary term $K = \frac{1}{2}\epsilon(\tau)(p^2 - m^2)$. Its associated Noether's charge, as expected, vanishes once again on the surface described by the constraint $\theta = 0$:

$$Q = K - p_\mu \delta_G x^\mu = \epsilon(\tau)\left(\frac{p^2}{2} - \frac{m^2}{2} + m^2\right) = \epsilon(\tau)\theta. \tag{3.51}$$

3.2.2 Electromagnetism

Let us focus on the Lagrangian of electromagnetism, namely (2.120). Since we aim to define an action of the form (3.26) and put it in a canonical structure, we first need to split it into its space and time components:

$$\begin{aligned}
I_{EM}[A_\mu(x)] &= -\frac{1}{4}\int d^4x F^{\mu\nu}F_{\mu\nu} \\
&= \int d^4x\left[-\frac{1}{2}F^{0i}F_{0i} - \frac{1}{4}F^{ij}F_{ij}\right] \\
&= \int d^4x\left[\frac{1}{2}(\dot{A}_i - \partial_i A_0)(\dot{A}^i - \partial^i A_0) - \frac{1}{4}F_{ij}F^{ij}\right] \\
&= \int d^4x\left(\frac{1}{2}\dot{A}_i\dot{A}^i - \dot{A}_i\partial^i A_0 + \frac{1}{2}\partial_i A_0\partial^i A_0 - \frac{1}{4}F_{ij}F^{ij}\right),
\end{aligned} \tag{3.52}$$

where we exploited the antisymmetry of the tensor [4] $F^{\mu\nu}$. Further we recall that the Latin index run over the three numbers $1, 2, 3$ and stand for the space variables, while the Greek ones run over all the numbers that represent the dimension of the spacetime variety, including the first one for the time coordinate, namely $0, 1, 2, 3$. We then define the conjugate momenta. We stress that in hamiltonian formalism they are defined only for those variables that appear with time derivatives. Therefore we have:

$$\pi_i = \frac{\partial L}{\partial \dot{A}_i} = \dot{A}_j - \partial_i A_0 \Rightarrow \dot{A}_i = \pi_i + \partial_i A_0. \tag{3.53}$$

We see that the canonical momenta π_i are the components of the electric field. The Hamiltonian is thus, using (3.53):

$$\begin{aligned}
H(p_i, A^i) &= \pi_i\dot{A}^i - L \\
&= \pi_i(\pi^i + \partial^i A_0) - \left[\frac{1}{2}(\pi_i + \partial_i A_0)(\pi^i + \partial^i A_0) - (\pi_i + \partial_i A_0)\partial^i A_0 + \frac{1}{2}\partial_i A_0\partial^i A_0 - \frac{1}{4}F_{ij}F^{ij}\right] \\
&= \frac{1}{2}\pi_i\pi^i + \frac{1}{4}F_{ij}F^{ij} + \pi_i\partial^i A_0.
\end{aligned} \tag{3.54}$$

When we integrate the Hamiltonian in order to find the expression of the action we can cast the last term in a different form exploiting integration by parts, namely:

$$\int d^4x(\pi_i\partial^i A_0) = \int d^4x(-A_0\partial_i\pi^i) + (A_0\pi^i)\Big|_{\partial\Omega} \tag{3.55}$$

[4] Just to be clear, we recall that this tensor reads, as it is well known:

$$\begin{bmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{bmatrix}$$

and we assume that the term evaluated at the border of the variety vanishes.

Hence the action in hamiltonian formalism for electromagnetism is:

$$\boxed{I_{EM}[A_i(x),\pi_i(x),A_0(x)] = \int d^4x \left[\pi_i \dot{A}^i - \left(\frac{1}{2}\pi_i\pi^i + \frac{1}{4}F_{ij}F^{ij}\right) + A_0\partial_i\pi^i\right].} \tag{3.56}$$

Here we are: this action has already the form (3.26). Indeed:

- Free Hamiltonian: $H_0 = \frac{1}{2}\pi_i\pi^i + \frac{1}{4}F_{ij}F^{ij} = \frac{1}{2}(\vec{E}^2 + \vec{B}^2)$, which represents the dynamical contribution to the Hamiltonian.
- Lagrange's multiplier: A_0.
- Constraint: $\theta = \partial_i\pi^i = \nabla \cdot \vec{E} = 0$, that is: Gauss' law, simply.

Here we must answer the same question: given initial condition for π_i and A_i that satisfy $\theta = 0$, do their time evolution keep on respecting this condition? We know that this is so, since Gauss' law is never violated. However we aim to compute it directly in gauge theories frame, using the formalism we provided in the sections above. We notice that now we are dealing with what we called continuous generalized coordinates and momenta, i.e. we are dealing with the field theory formalism. Hence we recall all the mathematical notation introduced in section (2.1.8), and in particular the definition of Poisson brackets. Let us impose the conservation of the constraint $\theta = \partial_i\pi^i$. Since we have only one constraint, the Poisson brackets $\{\theta,\theta\} = 0$ and we are left with:

$$\begin{aligned}
\frac{d\theta}{dt} &= \{\theta(x^\mu), H(x'^\mu)\} \\
&= \{\partial_i\pi^i(x^\mu), \frac{1}{2}\pi_i(x'^\mu)\pi^i(x'^\mu) + \frac{1}{4}F_{ij}(x'^\mu)F^{ij}(x'^\mu) - A_0(x'^\mu)\partial_i\pi^i(x'^\mu)\} \\
&= 2\partial_k\partial_i'\{\pi_k(x^\mu), A_j(x'^\mu)\}F^{ij}(x'^\mu) \\
&= -2(\partial_k\partial_i'\delta^3(x^\mu - x'^\mu))F^{ik}(x') = 0,
\end{aligned} \tag{3.57}$$

where in the last equivalence we exploited the symmetry of the first tensor (due to the definition of a distribution and its derivatives [5]) and the antisymmetry of the latter. Therefore the constraint θ is conserved along the equations of motion and must generate the gauge symmetry through Poisson brackets. We note that in field theory, as defined

[5] Indeed, providing the correct space of functions for the function ϕ, the derivative of a distribution T is defined as

$$\langle T', \phi(x)\rangle = -\langle T, \phi'\rangle. \tag{3.58}$$

in previous chapters, the Poisson brackets are a functional. Hence they can act on an arbitrary function $\Lambda(x)$; we obtain:

$$\begin{aligned}
\delta_G A_j(x)\Lambda(x,t) &= \{A_j, \partial_i \pi^i\}\Lambda(x,t) \\
&= \int dx' \sum_i \left(\frac{\delta A_j}{\delta A_i} \frac{\delta(\partial_k \pi^k)}{\delta \pi_i} - \frac{\delta(\partial_k \pi^k)}{\delta A_i} \frac{\delta A_j}{\delta \pi_i} \right) \delta(x-x')\Lambda(x,t) \\
&= \int dx' \partial_k \sum_i \left(\frac{\delta A_j}{\delta A_i} \frac{\delta \pi^k}{\delta \pi_i} - \frac{\delta \pi^k}{\delta A_i} \frac{\delta A_j}{\delta \pi_i} \right) \delta(x-x')\Lambda(x,t) \\
&= \int dx' \delta_{ik} \partial_j \delta(x-x')\Lambda(x,t) \\
&= -\int dx' \partial_j \Lambda(x',t)\delta(x-x') \\
&= -\partial_j \Lambda(x,t),
\end{aligned} \tag{3.59}$$

where we used the definition of derivative of a distribution and dropped the subscripts in x^i due to practical needs. Here the function Λ has any degree of regularity desired. Similarly we have:

$$\begin{aligned}
\delta_G \pi_j(x,t)\Lambda(x,t) &= \{\pi_j, \partial_k \pi^k\}\Lambda(x,t) \\
&= -\int dx' \sum_i \left(\frac{\delta \pi_j}{\delta \pi_i} \frac{\delta \pi_k}{\delta A_i} - \frac{\delta \pi_k}{\delta A_i} \frac{\delta \pi_j}{\delta \pi_i} \right) \delta(x-x')\Lambda(x',t) \\
&= 0,
\end{aligned} \tag{3.60}$$

which shows how the momenta π_i are gauge invariant, as they should be since they represent the electric field. To compute the variation for the Lagrange multiplier A_0 we see from the third equation of (3.32) that it consists only in the time derivative of the arbitrary function ϵ^a, since ${G_{ab}}^c = 0$ due the unique constraint θ. Therefore we determine the variation of this quantity by requiring that it provides a gauge symmetry for the action. We obtain:

$$\begin{aligned}
\delta_G I_{EM}[A_i, \pi_i, A_0] &= \int d^4x [\pi_i \delta_G \dot{A}^i + \delta_G A_0 \theta] \\
&= \int d^4x [-\pi_i \partial^i \partial_0 \Lambda + \delta_G A_0 \partial_i \pi^i] \\
&= \int d^4x [\partial_i(-\pi^i \partial_0 \Lambda) + \partial_i \pi^i \partial_0 \Lambda + \delta_G A_0 \partial_i \pi^i] \\
&= \int d^4x [\partial_\mu K^\mu + \partial_i \pi^i (\delta_G A_0 + \partial_0 \Lambda)],
\end{aligned} \tag{3.61}$$

where $K^\mu = (0, -\pi^i \partial_0 \Lambda)$ is a four vector, according to our notation. It is worth to notice that we used the fact that π_i, θ, H_0 are all gauge invariant i.e. their variations vanish.

Further let us stress that the prefactor of $\delta_G A_0 + \partial_0 \Lambda$ does not vanish since we varied the action for certain particular variations of the variables but for completely arbitrary functions, namely we are not on the equations of motion (therefore neither on the surface of constraints), as we explained in Thatcher (2). Hence we need $\delta_G A_0 = -\partial_0 \Lambda$ in order for $\delta_G I_{EM}$ to be a boundary term (namely, K^μ).

We then conclude that the gauge transformation of the field A_μ generated by the constraint $\partial_i \pi^i = 0$ is indeed the expected one, i.e. $\delta_G A_\mu = -\partial_\mu \Lambda$. Moreover we could also check that the conserved charge implied by Noether's theorem vanishes. Indeed applying Noether's theorem, we find:

$$\begin{aligned} J^\mu &= K^\mu - \pi_i \delta_G A^0 \\ &= \pi^i[-\partial_0 \Lambda + \partial_0 \Lambda] = 0 \end{aligned} \tag{3.62}$$

3.2.3 Generalisation in curved spaces: background fields and general relativity

To see how the theory of general relativity can be read as a gauge theory we need to take our treatment one step further. Indeed we realized that we dealt only with flat spaces, that is with spaces that has a constant metric. The theory of general relativity finds its roots just in the possibility to imagine a curved space-time with a metric as a function of the coordinates of the manifold considered. As we could guess, the only difference is that we have to take account the variation of the metric too, since it is is not constant anymore. The subtle point is that we should define a sort of derivative in spaces that are no more flat. To this aim we need to introduce the so called covariant derivative and develop a formalism which is not straightforward at all. We remind the reader to the Appendix C, where we briefly give an overview of the most important results, especially in writing the Lagrange and Hamilton equations in covariant form.

We provide the derivation of the gauge features of general relativity exploiting the ADM method. This method starts from the Lagrangian that provides Einstein's equations, neglecting its term related to matter contributions. To be as clear as possible we briefly introduce this Lagrangian. It enters the so-called Einstein-Hilbert action, which reads as

$$I[g_{\mu\nu}(x), \Gamma(x), \mathcal{L}_M(x)] = \int \left[\frac{1}{2}R + \mathcal{L}_M\right]\sqrt{-g}d^4x \tag{3.63}$$

where the range of the index is $\{0, 1, 2, 3\}$, $g_{\mu\nu}$ is the metric tensor, g its determinant [6] R the Ricci scalar, Γ the connection (say, the Christoffel's symbols of the second kind since we are considering a metric) and $\mathcal{L}_M$ stands for the Lagrangian of matter fields appearing in the theory. Further, we notice we used the natural units system. We want

[6] We exploited the Minkowski's metric with signature (1, -1, -1, -1) so that we must put a minus sign in order to compute the square root.

to prove that the choice made for the action integral is appropriate by showing that it leads to the correct equations of motion for the gravitational field, namely the Einstein's one. Finding the paths of motion means finding the paths that make the variation of the action integral vanish, so that we are computing an on-shell variation and we thus use the symbol $\delta_o I$ to point the variation, in order to emphasize that we are computing a variation on paths that solve the Euler-Lagrange equations of motion. So that performing the variation of the action we find:

$$\begin{aligned}\delta_o I[g^{\mu\nu}(x), R(x), \mathcal{L}_M(x)] &= \int \left[\frac{1}{2}\frac{\delta(\sqrt{-g}R)}{\delta g^{\mu\nu}} + \frac{\delta(\sqrt{-g}\mathcal{L}_M)}{\delta g^{\mu\nu}}\right]\delta g^{\mu\nu} dx \\ &= \int \left[\frac{1}{2}\left(\frac{\delta R}{\delta g^{\mu\nu}} + \frac{R}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}}\right) + \frac{1}{\sqrt{-g}}\frac{\delta(\sqrt{-g}\mathcal{L}_M)}{\delta g^{\mu\nu}}\right]\delta g^{\mu\nu}\sqrt{-g}d^4x \\ &= 0\end{aligned} \tag{3.64}$$

Since this equation holds for arbitrary variations of the metric $\delta g_{\mu\nu}$, the integrand vanishes:

$$\frac{\delta R}{\delta g^{\mu\nu}} + \frac{R}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}} = \frac{1}{\sqrt{-g}}\frac{\delta(\sqrt{-g}\mathcal{L}_M)}{\delta g^{\mu\nu}}. \tag{3.65}$$

These are the equations of motion. To show that they coincide with the Einstein one we only have to compute explicitly the variations of each terms. First we notice that the right hand side is just the energy-momentum tensor (since the usual Lagrangian is $\mathcal{L} = \sqrt{-g}\mathcal{L}_M$ and the energy-momentum tensor is given by $\frac{1}{\sqrt{g}}\frac{\delta L}{\delta g^{\mu\nu}}$). The computation of the other two variations is much more laborious, but only a matter of calculus.

In order to find the expression of the variation of the Ricci's scalar variation we initially seek to obtain the variation of the Riemann's tensor and then of the Ricci's tensor. Recall that these quantities are defined as

$$\begin{aligned} R^\rho_{\sigma\mu\nu} &= \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} \\ R &= g^{\sigma\nu}R_{\sigma\nu}\ , \end{aligned} \tag{3.66}$$

respectively, where $R_{\sigma\nu}$ comes from the contraction $R_{\sigma\nu} = R^\mu_{\sigma\mu\nu}$.

The variation of the Riemann's tensor is computed by noticing that although the Christoffel's symbol does not transform as a tensor, the difference $\Gamma^\rho_{\nu\sigma} - \Gamma^\rho_{\mu\sigma}$ does instead [7]. Thus we are allowed to compute its covariant derivative:

$$\nabla_\mu(\delta\Gamma^\rho_{\nu\sigma}) = \partial_\mu(\delta\Gamma^\rho_{\nu\sigma}) + \Gamma^\rho_{\mu\lambda}\delta\Gamma^\lambda_{\nu\sigma} - \Gamma^\lambda_{\mu\nu}\delta\Gamma^\rho_{\lambda\sigma} - \Gamma^\lambda_{\mu\sigma}\delta\Gamma^\rho_{\nu\lambda}. \tag{3.68}$$

[7] Recall that the law of transformation for the connections are:

$$\Gamma^\rho_{\nu\sigma} = \Gamma^r_{mn}X^\rho_r X^m_\nu X^n_\sigma + X^r_{\nu\sigma}X^\rho_r \tag{3.67}$$

where with Latin index we refer to a certain system of coordinates and with Greek ones a different system of coordinates; further $X^m_{\nu\cdots\sigma} = \frac{\partial x^m}{\partial x^\nu\cdots\partial x^\sigma}$. Thus the quantity $C^\rho_{\nu\sigma}(x) = \Gamma^\rho_{\nu\sigma}(x) - \Gamma'^\rho_{\nu\sigma}(x) = \delta\Gamma^\rho_{\nu\sigma}(x)$ transforms as a tensor.

We clearly see that the variation of the Riemann's tensor is exactly the difference of covariant derivatives of connections, that is:

$$\delta R^{\rho}_{\sigma\mu\nu} = \nabla_{\mu}(\delta\Gamma^{\rho}_{\nu\sigma}) - \nabla_{\nu}(\delta\Gamma^{\rho}_{\mu\sigma}). \tag{3.69}$$

To obtain the variation of the Ricci's tensor, we only have to contract two index of (3.69):

$$\delta R_{\sigma\nu} \equiv \delta R^{\rho}_{\sigma\rho\nu} = \nabla_{\rho}(\delta\Gamma^{\rho}_{\nu\sigma}) - \nabla_{\nu}(\delta\Gamma^{\rho}_{\rho\sigma}). \tag{3.70}$$

This is called PaLatini's identity.

The Ricci's scalar is defined as $R = g^{\sigma\nu}R_{\sigma\nu}$; hence its variation yields:

$$\begin{aligned} \delta R &= R_{\sigma\nu}\delta g^{\sigma\nu} + g^{\sigma\nu}\delta R_{\sigma\nu} \\ &= R_{\sigma\nu}\delta g^{\sigma\nu} + \nabla_{\rho}(g^{\sigma\nu}\delta\Gamma^{\rho}_{\nu\sigma} - g^{\sigma\rho}\delta\Gamma^{\mu}_{\mu\sigma}), \end{aligned} \tag{3.71}$$

where in the second line we exploited the so called compatibility of the metric with the connection, that is the requirement $\nabla_{\rho}g^{\sigma\nu} = 0$, which allows us to transport the metric inside the covariant derivative [8]. Since the quantity $(g^{\sigma\nu}\delta\Gamma^{\rho}_{\nu\sigma} - g^{\sigma\rho}\delta\Gamma^{\mu}_{\mu\sigma}) = B^{\rho}$ is a vector field, we can cast the covariant derivative in a different form. Indeed $\sqrt{-g}B^{\rho}$ is a contravariant vector density [9], and we obtain:

$$\nabla_{\rho}(\sqrt{-g}B^{\rho}) = \frac{\partial(\sqrt{-g}B^{\rho})}{\partial x^{\rho}}. \tag{3.73}$$

Finally, when we integrate this quantity in the variation of the action, applying Stokes' theorem [10] we can write it as a boundary term and thus neglect it, at least if we push the

[8] We recall that the requirement of compatibility of the metric forces the manifold to acquire a common and intuitive feature: taken two vector transported in parallel, it makes the angle between them remain constant during the parallel transportation. For further details see Appendix C.

[9] Recall that if T^n is a relative vector of weight W, then

$$\nabla_n T^n = \sqrt{g}^{(W-1)}\frac{\partial}{\partial x^n}[\sqrt{g}^{(1-W)}T^n]. \tag{3.72}$$

See [10] chapter VII, sections 7.1,7.2.

[10] Recall that Stokes's theorem yields:

$$\int_{R_M} T_{k_1\cdots k_{M-1},k_M} d\tau^{k_1\cdots k_m}_{(M)} = \int_{R_{M-1}} T_{k_1\cdots k_{M-1}} d\tau^{k_1\cdots k_{M-1}}_{(M-1)}, \tag{3.74}$$

where R_M is a space of dimension M with the space R_{M-1} as boundary, with the comma, as usual, we indicate the partial derivative with respect to the index that stands on its right, and with $d\tau^{k_1\cdots k_M}_M, \tau^{k_1\cdots k_{M-1}}_{M-1}$ we denote the M-cell, $M-1$-cell of the space R_M, R_{M-1} respectively (the concept of the M-cell is the one due to Grassmann, 1842). We also stress that this theorem holds since, although the left hand side is not a tensor quantity, when contracted with the skew symmetry cell yields a tensor. Further, this theorem is valid in non metrical spaces too, and, upon introducing a metric a_{mn}, boundary to infinite. We should pay attention when asking whether this term vanishes at the boundary or not, since the integrand depends also on the derivative of the metric; we will furnish a quick inspection on this topic in the next chapter.

However, when there is no boundary or the variation of the metric vanishes in a neighborhood of the boundary, this term does not contribute to the variation of the action. We therefore obtain

$$\frac{\delta R}{\delta g^{\mu\nu}} = R_{\mu\nu}. \tag{3.76}$$

We are only left with the variation of the determinant. This is quite easy if we exploit Jacobi's formula for differentiating a determinant, namely $\delta g = \delta(det(g_{\mu\nu})) = g g^{\mu\nu}\delta g_{\mu\nu}$. We thus have:

$$\begin{aligned} \delta\sqrt{-g} &= -\frac{1}{2\sqrt{-g}}\delta g = \frac{1}{2}\sqrt{-g}(g^{\mu\nu}\delta g_{\mu\nu}) \\ &= -\frac{1}{2}\sqrt{-g}(g_{\mu\nu}\delta g^{\mu\nu}) \end{aligned} \tag{3.77}$$

where we use the rule for varying the inverse of a matrix $\delta g^{\mu\nu} = -g^{\mu\sigma}(\delta g_{\sigma\lambda})g^{\lambda\nu}$. Hence, we conclude that:

$$\frac{1}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}} = -\frac{1}{2}g_{\mu\nu}. \tag{3.78}$$

Combing all variations computed above, i.e. putting (3.71), (3.78) in (3.65) we obtain Einstein's equation:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = T_{\mu\nu}. \tag{3.79}$$

This proves the correctness of the Einstein-Hilbert action.

We now focus on the ADM method, which begins with the Einstein-Hilbert action just introduced above and cast its terms in a different way, in order to obtain an action in hamiltonian form as the one of gauge theories. Just as in electrodynamics we separated the field into A_0, A_i, here we shall work with g_{00}, g_{i0}, g_{ij}. Everything else is completely equivalent. This is called the ADM method. The purpose is the following: as we obtained directly the correct formulation of the action in electromagnetism using this procedure, we expect to reach the same conclusions also in general relativity.

In general relativity we work with a smooth but non trivial manifold $\mathcal{M}$, which we think as being composed by the set of 3-dimensional surfaces Σ^3_t one for each constant

---

it becomes a generalization of Green's theorem, namely:

$$\int_{R_M} \frac{\partial}{\partial x^r}[\sqrt{a}T^r]\sqrt{a}dv_N = \int_{R_{N-1}} T^r n_r dv_{M-1}, \tag{3.75}$$

where a is the determinant of the metric, n_r is the unit normal to V_{M-1}, and with $dv_N = \sqrt{a}|dx_s x^k|$ we denote the volume element of the cell; the quantity $|dx_s x^k|$ is the determinant of the matrix that defines the cell. See [10], chapter VII, sections 7.3-7.6.

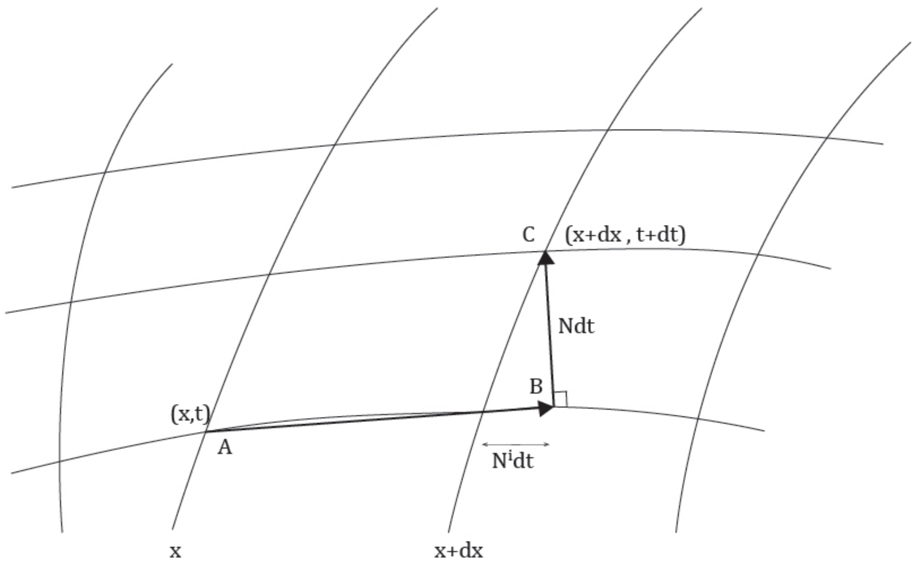

Figure 3.1: Elements used in the ADM method, defined on a manifold.

time t, with a metric that we denote by $g_{ij}(\vec{x}, t)$. In order to relate foliations at infinitesimally close times, we define $N(\vec{x}, t)$, the "lapse" function, such that starting from $\vec{x}$ at t, if we advance a distance $N(\vec{x}, t)dt$ in the (hyper)direction normal to Σ_t^3 at $\vec{x}$, we would reach exactly the surface Σ_{t+dt}^3. We must also define $N^i(\vec{x}, t)$ such that $N^i(\vec{x}, t)dt$ measures the "shift" produced, at constant time, between $\vec{x} + d\vec{x}$ and the point that will eventually hit $(\vec{x} + d\vec{x}, t + dt)$ by projecting with Ndt (see Fig.3.1). The relation between the metric components and the lapse and shift function N and N^i is obtained by simply writing the space-time interval between the point A and C in Fig.3.1 in both forms, and a short calculation shows [11]

$$g_{00} = -N^2 + N_i N^i \quad , \quad g_{0i} = N_i, \tag{3.80}$$

where spatial index are raised/lowered using the spatial metric g_{ij} and its inverse g^{ij}.

Exchanging g_{00}, g_{0i} in favor of N, N^i, which is after all only a matter of computation, the Einstein-Hilbert action takes the following form:

$$I_{EH}[N, N^i, g_{ij}] = \int d^4x N \sqrt{-^{(3)}g} (^{(3)}R - K^2 + K^{ij} K_{ij}) + B, \tag{3.81}$$

where B is a boundary term we assume we are allowed to neglect, and

$$\begin{aligned} K_{ij} &= \frac{1}{2N}[-\dot{g}_{ij} + N_{i|j} + N_{j|i}] \\ K &= g^{ij} K_{ij} \end{aligned} \tag{3.82}$$

[11] See [16] section 3.2.

where with $N_{i|j}$ we denote the covariant derivative of N_i with respect to x^j along the surface $\Sigma_t^{(3)}$, and with ${}^{(3)}g$, ${}^{(3)}R$ indicates the determinant and the Ricci's scalar with respect to the three dimensional metric g_{ij}, respectively. Thus, and this is important to keep in mind since here lies all the effort of ADM method, ${}^{(3)}R$ contains no time derivative; the only term with time derivatives is $\dot{g}_{ij}$ contained in K_{ij}. Therefore, as it will be clear in the hamiltonian form of the action, the true dynamical field is g_{ij}, while N, N^i are lagrangian multipliers. Indeed g_{00}, g_{0i} are not fixed and their combinations N, N^i play the role of lagrangian multipliers.

In order to put the action in hamiltonian form we need to explicitly obtain the Hamiltonian from the Einstein-Hilbert Lagrangian; that is we need to find the conjugate momenta. Actually we also seek a sort of definition of them. We define them with respect to the only field appearing with time derivatives in the action (2.27):

$$\Pi^{ij} = \frac{\partial \mathcal{L}}{\partial \dot{g}_{ij}}. \tag{3.83}$$

We can compute them from the variation of the action with respect to $\dot{g}_{ij}$; thus, recalling the definition (3.82), we find:

$$\begin{aligned} \delta I &= \int d^3x \sqrt{|g|}[-2K\delta K + 2K^{ij}\delta K_{ij}] \\ &= \int d^3x \sqrt{|g|}[Kg^{ij} - K^{ij}]\delta \dot{g}_{ij}, \end{aligned} \tag{3.84}$$

so that:

$$\Pi_{ij} = \sqrt{|g|}[Kg^{ij} - K^{ij}]. \tag{3.85}$$

Since we want the Hamiltonian only in terms of generalized coordinates and conjugate momenta, we must invert this relation and solve for the terms containing $\dot{g}_{ij}$ (namely K_{ij}) and write them as functions of Π^{ij}, g_{ij}. Taking the trace of last equation we find:

$$K = \frac{\Pi}{2\sqrt{|g|}}, \tag{3.86}$$

where $\Pi \equiv \Pi^{ij} g_{ij}$. Hence we can cast K_{ij} as

$$K^{ij} = \frac{1}{\sqrt{|g|}}\left(-\Pi^{ij} + \frac{\Pi}{2}g^{ij}\right). \tag{3.87}$$

Then, from (3.82), we solve for $\dot{g}_{ij}$ and write the Hamiltonian:

$$\begin{aligned}
H &= \int d^3x(\Pi^{ij}\dot{g}_{ij} - \mathcal{L}) \\
&= \int d^3x\Big[\Pi^{ij}\Big(-\frac{2N}{\sqrt{|g|}}\Big(\frac{\Pi}{2}g_{ij} - \Pi_{ij}\Big) + N_{i|j} + N_{j|i}\Big) - N\sqrt{|g|}(R - K^2 + K^{ij}K_{ij})\Big] \\
&= \int d^3x\Big[-\frac{2N}{\sqrt{|g|}}\Big(\frac{\Pi^2}{2} - \Pi^{ij}\Pi_{ij}\Big) + 2\Pi^{ij}N_{i|j} - N\sqrt{|g|}R + N\frac{\Pi^2\sqrt{|g|}}{4|g|} + \\
&- N\frac{\sqrt{|g|}}{|g|}\Big(\frac{3\Pi^2}{4} - 2\frac{\Pi}{2}\Pi + \Pi^{ij}\Pi_{ij}\Big)\Big] \\
&= \int d^3x\Big[N\Big[\frac{\Pi^{ij}\Pi_{ij}}{\sqrt{|g|}} - \frac{1}{2}\frac{\Pi^2}{\sqrt{|g|}} - \sqrt{|g|}R\Big] - 2\Pi^{ij}_{|j}N_i\Big] \; + \; B
\end{aligned} \tag{3.88}$$

where with $\Pi^{ij}_{|j}$ we always denote the covariant derivative of Π^{ij} along x^j on Σ^3_t and with B a boundary term. Therefore we obtained the Hamiltonian in terms of hamiltonian densities; indeed we can write it as:

$$H = \int d^3x(N\mathcal{H} + N_i\mathcal{H}^i), \tag{3.89}$$

where

$$\begin{aligned}
\mathcal{H} &= \frac{1}{\sqrt{|g|}}\Big(\Pi^{ij}\Pi_{ij} - \frac{\Pi^2}{2}\Big) - \sqrt{|g|}R \\
\mathcal{H}^i &= -2\Pi^{ij}_{|j}.
\end{aligned} \tag{3.90}$$

Here we are: the hamiltonian formulation of the Einstein-Hilbert action $I = \int p\dot{q} - H_0 + \lambda^a\phi_a$ reads as follow

$$\boxed{I_{EH-ADM}[g_{ij}(x), \Pi^{ij}(x), N(x), N_i(x)] = \int d^4x[\Pi^{ij}\dot{g}_{ij} - N\mathcal{H} - N_i\mathcal{H}^i]} \tag{3.91}$$

We readily recognize the elements of a gauge theory:

- Free Hamiltonian: $H_0 = 0$; this is typical of theories which are invariant under generalized coordinate transformations (similarly to the Polyakov action).
- Lagrange's multipliers: there are four of them, namely N and N_i.
- Constraints: also, we have four constraints, namely $\mathcal{H} = 0$, $\mathcal{H}^i = 0$.

We should stress some comments. The degrees of freedom in principle would have $6+6$ integration constants coming from Π^{ij} and g_{ij} (these are in fact symmetric matrices); however we also have 4 constraints and 4 gauge symmetries that reduce the number of degrees of freedom. Indeed we are left with $12-4-4=4$ degrees of freedom, the same as in the Einstein-Hilbert action formulation.

We aim to determine the gauge symmetries directly computing the Poisson brackets of the field and the constraints. Recall that the Poisson brackets in field theory is a distribution. Taking an arbitrary function $\zeta(x)$, we thus find:

$$\begin{aligned}\delta_G g_{ij}(x)\zeta(x) &= \{g_{ij}(x), \mathcal{H}^i(x)\}\zeta_i(x)\\ &= -2\{g_{ij}, \Pi^{ij}_{|j}\}\zeta_i(x)\\ &= -2(\{g_{ij}, \Pi^{ij}_{|j}\})_{|j}\zeta_i(x) + 2\{g_{ij|j}, \Pi^{ij}\}\\ &= 2\{g_{ij}, \Pi^{ij}\}\zeta_{i|j}(x)\\ &= 2\int dx'\frac{1}{2}(\delta^k_i\delta^l_j + \delta^l_i\delta^k_j)\delta(x-x')\zeta_{i|j}(x')\\ &= \zeta_{i|j}(x) + \zeta_{j|i}(x) = \pounds_\zeta[g_{ij}],\end{aligned} \tag{3.92}$$

where we used the compatibility of the connection with the metric and Green's theorem as explained in the above footnotes (see Appendix C for further details). We notice that (3.92) corresponds exactly to the Lie derivative of the field g_{ij} along the three-dimensional surface Σ^3_t. In order to recover this result we proceed as follow. Since $g_{ij}v^iw^j = \tilde{g}(\vec{V}, \vec{W})$ is a scalar, its Lie derivative and its covariant derivative coincide and are a simple derivative with respect to the parameter that indicates the curve along which we are moving. Therefore we can write:

$$\pounds_{\vec{V}}[\tilde{g}(\vec{A}, \vec{B})] = \nabla_{\vec{V}}[\tilde{g}(\vec{A}, \vec{B})]. \tag{3.93}$$

where we refer to the metric as a 2-form acting on two vectors $\vec{A}$, $\vec{B}$ of the manifold and denote it with a superposed tilde (with a slight abuse of notation, we still denote it as a covariant tensor of the second rank when we use its components). Applying Leibniz rule on both sides of this equation and recalling the compatibility of the metric with the connection, we find

$$\pounds_{\vec{V}}[\tilde{g}](\vec{A}, \vec{B}) + \tilde{g}(\pounds_{\vec{V}}[\vec{A}], \vec{B}) + \tilde{g}(\vec{A}, \pounds_{\vec{V}}[\vec{B}]) = \tilde{g}(\nabla_{\vec{V}}\vec{A}, \vec{B}) + \tilde{g}(\vec{A}, \nabla_{\vec{V}}\vec{B}). \tag{3.94}$$

Using this relation onto the basis vectors $\vec{e}_i$, $\vec{e}_j$ we obtain [12]:

$$
\begin{aligned}
(\pounds_{\vec{V}}[\tilde{g}])_{ij} &= \pounds_{\vec{V}}[\tilde{g}](\vec{e}_i, \vec{e}_j) \\
&= (\partial_i V^k g_{kj} + V^k \Gamma^l_{ik} g_{lj}) + (\partial_j V^k g_{ki} + V^k \Gamma^l_{jk} g_{li}) \\
&= V_{j|i} + V_{i|j},
\end{aligned}
\tag{3.95}
$$

as we wanted to show. When we ask this derivative to be zero we obtain the so called Killing's equation, that is an isometry of the metric g_{ij}; in other words we have an equation that gives the vectors of the integral curves along which the Lie derivative of the metric vanishes.

Similarly, as in (3.92), we can obtain the gauge transformations of $\delta\Pi^{ij}$, δN, δN^i, exploiting the method introduced for a general gauge theory.

[12] Recall that

$$
\begin{aligned}
\pounds_{\vec{V}}[\vec{e}_i] &= -[\vec{e}_i, \vec{V}] = -\partial_i V^k \vec{e}_k \\
\nabla_{\vec{V}} \vec{e}_i &= V^k \nabla_{\vec{e}_k} \vec{e}_i = V^k \Gamma^l_{ik} \vec{e}_l
\end{aligned}
$$

Chapter 4

Comments on boundary terms

The principle of least action is able to describe nature in an elegant formalism that provides the keys to find the equations of motion for a system and inspect its deep symmetries. A huge amount of phenomena is mastered by just one statement: the action must be stationary under arbitrary variations of the dynamical variables. Since physical interactions are described by fields, this powerful formulation is the core foundation of classical and quantum mechanics. However, this principle requires that the initial and final condition must be held fixed; this issue needs to be inspected with a careful treatment of boundary condition at infinity. In this brief chapter, we do not aim to find a solution to this problem, but only to show the reasonable questions from which it arises.

To appreciate this difficulty, let us first focus on non-gauge theory. Take the simplest case of a single real scalar field on some manifold $\mathcal{M}$ (which, after all, we consider to be non-compact) and that possesses a set of four coordinates divided in a time coordinate and the usual three space coordinates:

$$I[\phi(x)] = \int_{\mathcal{M}} \mathcal{L}(\phi,\, \phi_{,\nu},\, x) d^4x. \tag{4.1}$$

Its variation is:

$$\begin{aligned}\delta I[\phi(x)] &= \int_{\mathcal{M}} \left(\frac{\partial \mathcal{L}}{\partial \phi_\rho} \delta\phi_\rho + \frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \delta\phi_{\rho,\nu} \right) d^4x \\ &= \int_{\mathcal{M}} \left[\frac{\partial \mathcal{L}}{\partial \phi_\rho} - \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \right) \right] \delta\phi_\rho d^4x + \int_{\mathcal{M}} \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \delta\phi_\rho \right) d^4x.\end{aligned} \tag{4.2}$$

Via Green's theorem we can cast the boundary term as

$$B = \int_{\mathcal{M}} \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial \phi_{\rho,\nu}} \delta\phi_\rho \right) d^4x = \int_{\mathcal{M}} \frac{\partial \mathcal{L}}{\partial \phi_{\rho,\mu}} \delta\phi_\rho d\Sigma_\mu. \tag{4.3}$$

The principle of least action requires that this term vanishes. However, consider a

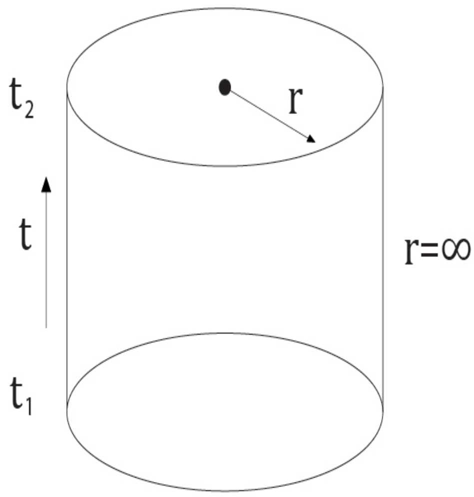

Figure 4.1: In order to depict the manifold, we reduced it to three dimensions, although it actually possesses four coordinates.

manifold as in Fig.4.1. The boundary of $\mathcal{M}$ has three pieces: the two covers at constant times t_1 and t_2, where $d\Sigma_\mu = d^3x$ pointing upwards and downwards in time respectively; and the cylinder at $r \to \infty$ where $d\Sigma_\mu = r^2 d\Omega dt \hat{r}$ ($d\Omega$ stands for the solid angle). The boundary term is thus:

$$B = \int d^3x \frac{\partial \mathcal{L}}{\partial \phi_{\rho,0}} \cancel{\delta\phi_\rho \Big|_{t_1}^{t_2}} + \int \frac{\partial \mathcal{L}}{\partial \phi_{\rho,r}} \delta\phi_\rho r^2 d\Sigma dt \Big|_{r\to\infty}. \tag{4.4}$$

The first term, evaluated at t_1 and t_2, vanishes because $\delta\phi(t_1) = \delta\phi(t_2) = 0$, i.e. the initial and final states are fixed. This is in full consistency with the equations of motion that require initial and final conditions for a unique problem; this corresponds to fix the two constants of integration in the equations of motion. The last term is evaluated for large r and one cannot assume that ϕ is also fixed there. If one fixes the field for large r, the equations of motion may have no solution at all.

In non-gauge field theories one normally deals with fields with compact support where $\phi(r) \to 0$ fast enough for large r; indeed a massive field typically exhibits an exponential decay. This means that $\frac{\partial \mathcal{L}}{\partial \phi_{\rho,r}}$ falls fast enough at infinity (it suppresses the growth of r^2) making the boundary term vanishes, so this does not become an issue and it is safe to omit the discussion.

For a gauge theory the situation is quite different. First we could deal with non massive fields, which means long range interactions and that we cannot assume a compact support neither a fast decay. Further, and this turns out to be subtler, even if fields do vanish fast enough asymptotically, the presence of Lagrange's multipliers makes this analysis delicate, since there are no dynamical equations restricting them.

For instance, let us consider first Maxwell's electrodynamics on a similar manifold with time and three space coordinates. To derive Maxwell's equations, we compute the variation of $I[A_\mu(x)] = -\frac{1}{4}\int F^2$:

$$\begin{aligned} \delta I[A_i(x)] &= -\int_{\mathcal{M}} d^4x F^{\mu\nu}\partial_\mu \delta A_\nu \\ &= \int_{\mathcal{M}} d^4x(\partial_\mu F^{\mu\nu})\delta A_\nu - \int_{\mathcal{M}} d^4x \partial_\mu(F^{\mu\nu}\delta A_\nu). \end{aligned} \tag{4.5}$$

The second term is the boundary term we are interested in. Via Green's theorem we can cast is as:

$$\begin{aligned} B &= \int_{\partial\mathcal{M}} F^{\mu\nu}\delta A_\mu d\Sigma_\nu \\ &= \int_{\partial\mathcal{M}} d^3x F^{i0}\cancel{\delta A_i\Big|_{t_1}^{t_2}} + \int_{\partial\mathcal{M}} dt d\Sigma_i F^{0i}\delta A_0 + \int_{\partial\mathcal{M}} dt d\Sigma_j F^{ij}\delta A_i. \end{aligned} \tag{4.6}$$

As before, the first term vanishes due to fixed initial and final time condition, $\delta A_i(t_1, t_2) = 0$. However the second and third terms need to be considered in a different way. The second one further involves A_0, a Lagrange's multiplier as we saw in the previous chapter. Since this field does not satisfy any equation of motion it can in principle take any value. That is, it could be zero, finite or even infinity. One may be tempted to declare simply that A_0 must be such that this term vanish. Actually, a large literature on this topic shows that this is not a wise choice.

Anyway, boundary terms problem can be solved in order to recover all the results and formalism adopted in previous chapters. Our aim is not to show all the procedure, but to underline how this issue is not straightforward and need to be studied carefully.

Appendix A

Calculus of variations

We aim to specify what we mean with the variation of a functional, namely an operator that acts on functions and gives back a real number. Formally a functional is defined as follow.

Functional. *A functional is a map* $I : \mathcal{F} \to \mathbb{R}$ *that associates a real number to every element of a set of functions.*

We tackle the issue in a way that is more physical and useful for our purposes. The problem of calculus of variations arises from the study of optimizing paths, that is finding the function that take the extreme value of a particular integral. To put it in a mathematical form, the integral we refer to is

$$I[f(x,y,y')] = \int dx f(x,y,y'), \tag{A.1}$$

where f is the functions that stands for the quantity we need to vary, $y \equiv y(x)$ is a function of the x variable and $y' \equiv y'(x)$ its derivative. Our treatment will deal with function of a single variable, since extending it to vectorial functions of many variables is straightforward (if we pay attention to perform independent variations for each variables, as it will be clear in the following lines). The unknown element to be obtained for the purpose, is the form of y as a function of the variable x. We try to compute the value of the integral on functions that differ a little from y, which we denote by $y + \delta y$. This naif measure of the smallness of the difference could be put in a more definite form:

$$\delta y = \epsilon v \tag{A.2}$$

where ϵ is an arbitrary constant so small in magnitude that any positive integral power of I is unimportant relative to every lower positive integral power. Also, v is any regular uniform function of x within the range of the integral and all its derivatives also are regular uniform functions of x within that range. Further, v is an arbitrary function and we moreover assume that v *is independent from* ϵ. We will denote the quantity (A.2) with δy and call it variation of the function y. We stress that is nothing but a function. Writing the value of the functional on y and on $y + \delta y$ with I and J, respectively, we have

$$J - I = \int dx \Big[f(x, y + \epsilon v, y' + \epsilon v') - f(x, y, y') \Big] \tag{A.3}$$

the integration extending between the fixed constant limits. We further assume that the function f is of such a form that $f(x, y + \epsilon v, y' + \epsilon v')$ can be expanded in a uniformly converging series, which proceeds in powers of ϵ. Thus we have, after this expansion,

$$J - I = \epsilon I_1 + \frac{1}{2}\epsilon^2 I_2 + R_3; \tag{A.4}$$

where

$$\begin{aligned} I_1 &= \int \left(v\frac{\partial f}{\partial y} + v'\frac{\partial f}{\partial y'} \right) dx, \\ I_2 &= \int \left(v^2 \frac{\partial^2 f}{\partial y^2} + 2vv'\frac{\partial^2 f}{\partial y \partial y'} + v'^2 \frac{\partial^2 f}{\partial y'^2} \right) dx, \end{aligned} \tag{A.5}$$

with the same range of integration as for I and where R_3 denotes the aggregate of terms that involve third and higher powers of ϵ. It is clear that the quantity ϵI_1 (called the first variation) when non vanishing, dominates the value of the right-hand side and that R_3 is the term of lower interest, since it gives the smallest contribute with respect to the others.

We know provide a way to determine the function that takes the extreme value of the functional considered under some important conditions, namely that the variation of the function f vanishes at the end points of the range of integration. This is quite natural to assume, since we recall that we deal with a problem of optimizing paths and we must fix a common feature of these paths, that is they have to begin and end at the same points. Therefore we can cast the first variation in a different form. Indeed we can write:

$$\int dx \;\; v'\frac{\partial f}{\partial y'} = [v\frac{\partial f}{\partial y'}] - \int dx \;\; v\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right), \tag{A.6}$$

where the quantity in square brackets is to be taken at the end points of the range of integration. Since we assumed that no variation occurs there, the function v vanishes when evaluated at the end points as well as the term in square brackets. Hence we can write the first variation in the following form:

$$I_1 = \int dx \;\; v\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]. \tag{A.7}$$

Since the quantity ϵI_1 (when non vanishing) dominates the value of $J - I$, a change of sign for ϵ changes the sing of ϵI_1, that is, changes the sign of the value of $J - I$: in other words, one variation lead to an increase, and another variation to a decrease. As a maximum needs to be characterized by a decrease for all variations, and a minimum by an increase for all variations, the preceding possibility must be excluded: consequently the quantity ϵI_1, and therefore the integral I_1 must vanish. Further, since the function v is completely arbitrary, the only way to make I_1 vanish is that the term $\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)$ vanishes itself. Here we are: we have obtained an equation for the function f that makes the value of the functional a maximum or a minimum. We note that this treatment is equivalent as dealing with a Lagrangian f and an action I. We thus provided a method for computing the equations of motions that happens to represent the paths that take the extreme value (namely, a maximum or a minimum) of the action integral.

Modern literature provides a standard formal way to treat variations. Indeed the variation of a functional can be written in a formal way as

$$\delta I \equiv \int dx \frac{\delta I}{\delta f(x)} \delta f(x) \tag{A.8}$$

where $f(x) = f(x, y(x), y'(x))$, and $\frac{\delta}{\delta f(x)}$ is the functional derivative. Therefore the problem of the variation of the action integral reads:

$$\int dx \frac{\delta I}{\delta f(x)} \delta f(x) = 0 \tag{A.9}$$

Note that the functional derivative introduced before when dealing with field theory reduces exactly to this one.

Appendix B

Symplectic group

We now provide an overview on the symplectic manifolds in order to give a formal meaning to the Poisson brackets. We begin with some definitions and results.

Let $\mathcal{M}^{2n}$ be an even-dimensional manifold. A *symplectic structure* is a closed nondegenerate differential 2-form w^2 on $\mathcal{M}^{2n}$:

$$dw^2 = 0 \qquad and \qquad \forall \zeta \neq 0,\ \exists \eta :\ w^2(\zeta, \eta) \neq 0 \tag{B.1}$$

where $\zeta, \eta \in TM_x$, and TM_x is the tangent vector space of the manifold in the point x and dw^2 is the exterior derivative of the 2-form. The pair $(\mathcal{M}^{2n}, w^2)$ is called a *symplectic manifold*.

Symplectic structures arise naturally from the Lagrangian function of a system. Consider then a Lagrangian with configuration space V. The set of generalized velocities $\dot{q}$ is easily considered as a tangent vector space to the configuration manifold V; the generalized momentum $p = \frac{\partial L}{\partial \dot{q}}$, thus, could be considered as cotangent vectors, that is one-forms that act on the tangent vector space to the manifold. We define the one-form $w^1 = pdq$ and the 2-form taking its exterior derivative $dw^1 = w^2 = dp \wedge dq = \sum_i dp_i \wedge dq_i$, which is therefore non degenerate and closed. That is: we have defined a symplectic structure [1].

Another definition is to be mentioned. Indeed a symplectic structure established an isomorphism between the space of tangent vectors and its associated one-forms. In order to define such isomorphism, to each vector ζ tangent to a symplectic manifold $(\mathcal{M}^2, w^2)$ at the point x, we associate a 1-form w^1_ζ on TM_x by the formula [2]:

$$w^1_\zeta(\eta) = w^2(\eta, \zeta) \qquad \forall \eta \in TM_x \tag{B.2}$$

[1] This property is somewhat more general, since it is a result carried from a general theorem involving tangent bundle and cotangent bundle.

[2] This is the so called Fischer-Riesz theorem, we are allowed to consider if we assume that the one-form ω^1 is continuous as a functional.

The vectorial space of the 1-forms defined on a tangent space TM_x to a manifold $\mathcal{M}$ represent the dual space of that tangent space and is denoted by TM_x^*. We denote by J the isomorphism $J : T^*M_x \to TM_x$ constructed above. Recall that given a function H on the manifold, its differential dH is a 1-form on $\mathcal{M}$ so that the isomorphism considered could produce a vector field JdH on $\mathcal{M}$.

We only miss the last concept, namely the phase flow $g_H^t(x)$ of a function $H : \mathcal{M}^{2n} \to \mathbb{R}$. It is defined as the one parameter group of diffeomorphism $g^t : \mathcal{M}^{2n} \to \mathcal{M}^{2n}$ provided by the vector field JdH:

$$\left.\frac{d}{d\tau}\right|_{\tau=0} g^\tau(x) = JdH(x), \tag{B.3}$$

Let us be more explicit. Consider the symplectic manifold $\mathcal{M}^{2n} = \mathbb{R}^{2n} = \{(p,q)\}$ where q and p are the generalized coordinates and associated momenta of a certain hamiltonian function H. Let q_t, p_t be the canonical coordinates evaluated at the instant t and $q_{t+\tau}$, $p_{t+\tau}$ be their values evaluated at the instant $t+\tau$. These latter are functions of the first ones and the time interval τ taken as a parameter:

$$q_{t+\tau} = q(q_t, p_t, \tau), \qquad p_{t+\tau} = q(q_t, p_t, \tau). \tag{B.4}$$

If we consider these formulae as transformations from the variables q_t, p_t to the variables $q_{t+\tau}$, $p_{t+\tau}$, this transformation is canonical. This is a map of the trajectories of the solutions of the system depending continuously to the parameter τ. The generating function [3]that encodes this transformation is the action integral (upon a minus sign) viewed as depending only on $q_{t+\tau}$ and q_t, namely $S(q_t, q_{t+\tau})$; indeed the differential of the action integral in hamiltonian formalism $S = \int_{t_0}^{t_1} pdq - Hdt$ is exact [4] and when evaluated at the extreme values reads (along the equations of motion):

$$\begin{aligned} dS &= \sum_i (p_i(t+\tau)q_i(t+\tau) - H(t+\tau)d(t+\tau) - p_i(t)q_i(t) + H(t)dt) \\ &= \sum_i (p_i(t+\tau)q_i(t+\tau) - p_i(t)q_i(t)) \end{aligned} \tag{B.6}$$

Therefore the generating function $F = -S(q_t, q_{t+\tau})$ produces the one parameter group diffeomorphism in (B.3) and maps the points q_t, p_t in the points $q_{t+\tau}$, $p_{t+\tau}$ along the

[3] We are dealing with a transformation of the first kind: given some canonical coordinates q, p, aiming to find different coordinates Q, P that verify $\dot{Q} = \frac{\partial H'}{\partial P}$, $\dot{P} = -\frac{\partial H'}{\partial Q}$ where H' is the Hamiltonian in these different coordinates (that is: aiming to define some other canonical coordinates), a generating function of the first kind is defined as a function $F(q, Q, t)$ that gives:

$$p_i = \frac{\partial F}{\partial q_i}, \qquad P_i = -\frac{\partial F}{\partial Q_i}, \qquad H' = H + \frac{\partial F}{\partial t}. \tag{B.5}$$

[4] Indeed, from the definition of the action integral we have $S = \int dtL$.

equations of motion on the manifold:

$$S^\tau(p(t), q(t)) = (p(t+\tau), q(t+\tau)), \tag{B.7}$$

Indeed, using the notation of (B.3), we obtain:

$$\frac{d}{d\tau}\bigg|_{\tau=0} S^\tau(p(t), q(t)) = (\dot{p}, \dot{q}) = JdH, \tag{B.8}$$

clearly in canonical basis, since in this case $J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$ where I is the identity matrix and $dH = (\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p})$.

We now focus on Poisson brackets. We have just showed that to any function H on the symplectic manifold we can define a one parameter group $g_H^\tau : \mathcal{M}^{2n} \to \mathcal{M}^{2n}$ of canonical transformations of $\mathcal{M}^{2n}$. We called it the phase flow of the hamiltonian function equal to H. Let F be another given function on the symplectic manifold. We define their Poisson brackets as

$$\{F, H\}(x) = \frac{d}{d\tau}\bigg|_{\tau=0} F(g_H^\tau(x)), \tag{B.9}$$

where obviously $x = (p, q)$ [5]. Therefore the Poisson brackets of two functions is again a function. In other words, we see that the Poisson brackets of two functions F, H is the derivative of F along the direction of the phase flow with hamiltonian function H. Hence, we can give the Poisson brackets of two fields F, H on a manifold $\mathcal{M}$ a deeper meaning: they are the *Lie derivative of the function F along the phase flow with hamiltonian function H* [6].

In order to the deduce the standard form of the Poisson brackets we use the isomorphism J between 1-forms and vector fields on a symplectic manifold $(\mathcal{M}^{2n}, w^2)$. This isomorphism is defined, exploiting (B.2), by the relation

$$w^2(\eta, Jw^1) = w^1_{J\omega^1}(\eta) \tag{B.10}$$

where $J\omega^1$ is a vector of the manifold. Exploiting the definition (B.3), we can cast the Poisson brackets as:

$$\{F, H\} = dF(JdH) \tag{B.11}$$

that is, the Poisson brackets of the functions F and H is equal to the value of the 1-form dF on the vector fields IdH of the phase flow with hamiltonian function H. Further, exploiting (B.10), we obtain:

$$\{F, H\} = w^2(JdH, JdF), \tag{B.12}$$

[5] As it should be clear, in the whole treatment of this section when we write (p, q) we mean $(p_1, \cdots, p_n, q_1, \cdots, q_n)$, as well as when we write $(\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p})$ we mean $(\frac{\partial H}{\partial q_1}, \cdots, \frac{\partial H}{\partial q_n}, \frac{\partial H}{\partial p_1}, \cdots, \frac{\partial H}{\partial p_n})$.

[6] We see that $g_H^\tau(x)$ is the push-forward of the point x on the manifold along the congruence spanned by the integral curves of the equations of motion furnished by the Hamiltonian H.

namely, the Poisson brackets of the functions F and H is equal to the skew scalar product of the vector field of the phase flows with hamiltonian functions H and F.

We have now reached the results we were looking for: *the Poisson brackets of the functions F and H is a skew-symmetric bilinear function of F and H, that is*

$$\begin{aligned} \{F, H\} &= -\{H, F\} \\ \{H, \lambda_1 F_1 + \lambda_2 F_2\} &= \lambda_1\{H, F_1\} + \lambda_2\{H, F_2\} \quad (\lambda_1 \in \mathbb{R}). \end{aligned} \tag{B.13}$$

For instance, let us compute the Poisson brackets of two functions F and H in the canonical coordinate space $\mathbb{R}^{2n} = (p, q)$

$$w^2(\zeta, \eta) = < J\zeta, \eta >= \zeta^T J^T \eta. \tag{B.14}$$

By (B.12) we have:

$$\{F, H\} = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\frac{\partial F}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial F}{\partial p_i}, \tag{B.15}$$

since the matrix J is the canonical one used above. We thus recovered the usual standard formula of Poisson brackets.

Moreover it is easily checked that the Poisson brackets verifies the following properties:

Skew-symmetry: $\{F, H\} = -\{H, F\}$;

Jacobi identity: $\{F, \{G, H\}\} + \{H, \{F, G\}\} + \{G, \{H, F\}\} = 0$;

Bilinearity: $\{H, \lambda_1 F_1 + \lambda_2 F_2\} = \lambda_1\{H, F_1\} + \lambda_2\{H, F_2\}$;

Closure: $\{F, H\}$ is another function defined on the manifold.

Hence they define a Lie algebra. When dealing with first integral, namely with conserved charges as explained in above sections, we notice that the corresponding functions constitute a subalgebra, that is the subalgebra of the conserved charges (see 2.29).

Appendix C

Tensors

We present here a subtle problem, that is the problem of independence of the equations from the set of coordinates used. Actually this issue have been studied deeply and the answer lies in using tensors. Indeed, tensorial equations, due to their transformations laws among different sets of coordinates, keep their form in all coordinate systems: if an equation involves some physical quantities in one coordinate system, when evaluated in other coordinate systems it still involves the same quantities in the same manner, obviously evaluated in the different coordinate systems. However, before facing this problem, we should have defined what a coordinate system is. This is of great importance and not so easy; the whole tools of differential geometry have been built to cope with this problem. Although we cannot provide all the results, we only would like to bring up some of them in order to introduce the concepts of background fields and show their features.

The hardest part in dealing with tensors is defining a derivative, since the usual one breaks their transformation laws. One type of derivative that could be provided is the Lie ones, but we shall focus on another one, that is the so called covariant or absolute derivative. This derivative turns out to be useful since it gives a law for the parallel transport also in spaces (manifolds) in which it is puzzling to do so. For instance, it is clear when two lines are parallel in a flat euclidean space, but it is not along the surface of a sphere; further, if we do not have the possibilities to immerse the sphere in a higher dimensional space, it is much more difficult to define what is parallel or not at first sight. All this treatment is tied tightly to Lagrangian systems since the principle of least action can be put in a geometrical form; indeed the equations of motion turn out to be the paths that minimize the distance between two points in the configuration space, namely the geodesics (actually, when energy is conserved).

The covariant derivative is defined as follow. Let $\mathcal{M}$ be a manifold and $q^m(u)$ the coordinates of a curve parameterized by u. Given a (contravariant) tensor T^r, where r is an index running in a specific range $r_1, \cdots r_n$, consider the tensor evaluated on the curve

$T^r(q(u))$, then the covariant derivative $\frac{\delta T_r}{\delta u}$ of this tensor is defined as:

$$\begin{aligned}\frac{\delta T_r}{\delta u} &\equiv \frac{dT^r}{du} + \left\{ {r \atop mn} \right\} T^m \frac{dq^n}{du} \\ &= \left(\frac{\partial T^r}{\partial x^n} + \left\{ {r \atop mn} \right\} T^m \right) \frac{dq^n}{du}\end{aligned} \tag{C.1}$$

where the symbol $\left\{ {r \atop mn} \right\}$ is the Christoffel's symbol (of the second kind), also indicated as Γ^r_{mn}, and defined as:

$$\left\{ {r \atop mn} \right\} = a^{rs}[mn, s] = a^{rs}\frac{1}{2}\left(\frac{\partial a_{sm}}{\partial q^n} + \frac{\partial a_{sn}}{\partial q^m} - \frac{\partial a_{mn}}{\partial q^s} \right), \tag{C.2}$$

where a_{mn} is the metric tensor of the manifold [1]. We note that we decomposed the expression of the covariant derivative in two parts: the one in round brackets are the components of the tensor obtained, the other is the basis namely the vector field with parameter u. When the covariant derivative of a tensor vanishes, we say that it is propagated in parallel along the curve $q^m(u)$. The covariant derivative of a covariant tensor S_r can be obtained exploiting its contraction with the tensor above. In fact the covariant derivative of a scalar is the usual one and vanishes since it is a constant. When we contract the tensor S_r with a tensor T^r propagated in parallel we obtain an equation that yields the covariant derivative of the tensor S_r. Iterating this process we can obtain the covariant derivative of a tensor of any rank [2].

We can now provide the Euler-Lagrange equations of motion in the so called covariant form, namely in a form that it is invariant under change of coordinate system. When the Lagrangian contains only the kinetic element, namely $L = \frac{1}{2}a_{mn}(q)\dot{q}^m\dot{q}^n$, we can vary the action $\int dtL(q, \dot{q})$ with fixed limit of integration obtaining the equations of motion:

$$\frac{d^2q^r}{du^2} + \left\{ {r \atop mn} \right\} \frac{dq^m}{du}\frac{dq^n}{du} = 0. \tag{C.3}$$

Since the tensor $\frac{dq^r}{du}$ represents the tangent vector field of the curve q^r, we have obtained an equation that yields that the tangent vector to the curve is propagated in parallel

[1] The covariant derivative is also denoted by $\nabla_{\vec{V}}T^r = \frac{\delta T_r}{\delta u}$ where $\vec{V} = \frac{d}{du}$ is the vector fields with parameter u; its components are also denoted by $\nabla_n T^r = \frac{\partial T^r}{\partial x^n} + \left\{ {r \atop mn} \right\} T^m$.

[2] There is a more general way to introduce this topic. Indeed covariant derivative is defined also in spaces that do not carry a metric tensor. It is defined methodically, requiring that it satisfies certain properties. Upon introducing a metric tensor, we require that the covariant derivative of the tensor vanishes; this is assumed in order to let the scalar product (induced by the metric) of two tensors propagated in parallel remain unchanged under parallel propagation. This requirement leads to have the affine connection coefficients equal to the Christoffel's symbols computed above. Further the expression of the covariant derivative also reduces to the one given above. See [10], chapter VIII.

along the curve itself: we have thus obtained a geodesic, namely a "straight" line in our curved space. We can also derive the hamiltonian form of them, by inspecting the Hamiltonian deduce from the Lagrange of a free particle

$$L = \frac{1}{2} m a_{ij}(q) \dot{q}^i \dot{q}^j. \tag{C.4}$$

The conjugate momenta are

$$p_i \equiv \frac{\partial L}{\partial \dot{q}^i} = m a_{ij} \dot{q}^j, \tag{C.5}$$

hence the Hamiltonian reads

$$H(p, q) = p_i \dot{q}^i - L(q, \dot{q}) = \frac{1}{2m} a^{ij}(q) p_i p_j. \tag{C.6}$$

Therefore the equations of motion are:

$$\begin{aligned} \dot{q}^i &= \{q^i, H\} = \frac{1}{m} a^{ij}(q) p_j \\ \dot{p}_i &= \{p_i, H\} = -\frac{1}{2m} (\partial_i a^{kl}) p_k p_l. \end{aligned} \tag{C.7}$$

We aim now to show directly that the free Lagrangian $L = \frac{1}{2} a_{mn}(q) \dot{q}^m \dot{q}^n$ is invariant under a change of coordinates in a curved space, that is that the motion of the free particle is independent from the coordinate system. Recall that we are dealing with a transformation of the form:

$$\begin{aligned} q^i &\rightarrow \quad {q'}^i(q) = q^k \frac{\partial q^i}{\partial {q'}^k} \\ a_{ij} &\rightarrow \quad a'_{ij}(q') = a_{kl}(q) \frac{\partial q^k}{\partial {q'}^i} \frac{\partial q^l}{\partial {q'}^j} \end{aligned} \tag{C.8}$$

where we used the transformation laws for tensors. In order to find the variation of this terms, we need to compute their Lie derivatives. For a dragging along the congruence defined by the integral curve $\zeta^i(q)$ on the manifold, exploiting the results in section (2.2.1), we find:

$$\begin{aligned} \delta q^i &= {q'}^i - q^i = -\zeta^i(q) \partial_i q^i = -\zeta^i(q) \delta_i^j = -\zeta^i(q) = \\ &= \pounds_{\vec{\zeta}}[(q^i)] \\ \delta a_{ij} &= a'_{ij}(q) - a_{ij} - \zeta k \partial_k a_{ij} - a_{ik}(q) \delta_j \zeta^k - a_{kj}(q) \partial_i \zeta^k \\ &= \pounds_{\vec{\zeta}}[(a_{ij})] \end{aligned} \tag{C.9}$$

Performing the variation of the action $\delta I = \int dt \delta L$, and substituting the variations using (C.9), we obtain:

$$\delta L(q, \dot{q}) = m a_{ij} \delta \dot{q}^i \dot{q}^j + \frac{m}{2} (\delta q^k \partial_k a_{ij}) \dot{q}^i \dot{q}^j + \frac{m}{2} \delta a_{ij} \dot{q}^i \dot{q}^j = 0 \tag{C.10}$$

This symmetry, technically speaking, is different from the others, since we transformed not only the dynamical variables but also the metric which describes the action of external forces. To see how this is possible we have to adjust the principle of least action including the potential in the metric. Let us show this procedure.

The Lagrangian of a mechanical system composed by a single particle of unit mass is defined in general as the sum of two terms, the kinetic energy $T = \frac{1}{2}a_{mn}(q)\dot{q}^m\dot{q}^n$ and the potential $U(q)$. If the total energy $E = T + U$ is conserved (a case that in theoretical physics occurs most of the time, at least under suitable approximations) we can find the so called abbreviate action. Indeed the principle of least action reads

$$\delta I(p,q) = \int \delta(pdq) - \delta H dt = \delta \int pdq = \delta \int p_i q^i dt, \tag{C.11}$$

since the Hamiltonian H is a constant function when the total energy is conserved. We can write the generalized momenta as

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = a_{ij}\dot{q}^j, \tag{C.12}$$

and the constant energy as

$$E = \frac{1}{2}a_{ij}\dot{q}^i\dot{q}^k + U(q). \tag{C.13}$$

From last equation we find:

$$dt = \sqrt{\frac{a_{ij}dq_i dq_j}{2[E - U(q)]}}, \tag{C.14}$$

When we insert this result in (C.12), (C.11) yields:

$$\delta I(p,q) = \int \sqrt{2[E - U(q)]a_{ij}(q)dq_i dq_j}. \tag{C.15}$$

If we use the arc-length $ds^2 = a_{ij}(q)dq_i dq_j$ and redefine the $d\sigma^2 = 2[E - U(q)]ds^2$, we can cast the principle of least action as

$$\delta I(p,q) = \delta \int d\sigma = 0, \tag{C.16}$$

that is: to find the equations of motion we need to find the path that minimize the action; this path is the shortest one, namely the geodesic of the metric $d\sigma^2 = 2[E - U(q)]ds^2$ which is deformed along the manifold by the potential $U(q)$.

To recover the equations of motion for a free particle found in (C.3) we readily see that

$$T = \frac{1}{2}a_{mn}(q)\dot{q}^m\dot{q}^n = \frac{1}{2}\left(\frac{ds}{dt}\right)^2$$

$$\begin{aligned} I(q,\dot{q}) &= \int T dt = \int \frac{T ds}{\sqrt{2T}} = \frac{1}{\sqrt{2}}\int \sqrt{T} ds \\ &= \frac{1}{\sqrt{2}}\int \sqrt{E-U} ds = \int d\sigma, \end{aligned} \tag{C.17}$$

therefore this two ways of finding the correct paths of motion are completely equivalent.

Bibliography

[1] E. Noether. ≪*Invariante Variationsprobleme*≫. In: Nachrichten von der Kőnigl. Gesellschaft der Wissenschaften zu Gőttingen (1918), pp. 235–257.

[2] Max Banados, *A short review on Noether's theorems, gauge symmetries and boundary terms* url: https://arxiv.org/pdf/1601.03616v3.pdf.

[3] A. K. Hadler, Andronikos Paliathanasis, P.G.L. Leach, *Noether's Theorem and Symmetry*, https://www.mdpi.com/2073-8994/10/12/744.

[4] K. Brading e H. R. Brown, *Noether's Theorems and Gauge Symmetries. Ago. 2000.* url: https://arxiv.org/pdf/hep-th/0009058.pdf.

[5] H. R. Brown e K. Brading. *General covariance from the perspective of Noether's theorems. 2002.* url: http://philsci-archive.pitt.edu/821/1/TorrettiB&B.pdf.

[6] A. A. Deriglazov, *On singular lagrangian underlying the schrödinger equation*, Physics Letters A, vol. 373, no. 43, pp. 3920–3923, 2009.

[7] Fiorenzo Bastianelli, *Appunti per il corso di Fisica Teorica 1 - a.a. 2017/18.*

[8] N. Byers. *The Life and Times of Emmy Noether. Nov. 1994.* url: https://arxiv.org/pdf/hep-th/9411110.pdf.

[9] Chris Quigg, *Colloquium: A Century of Noether's Theorem.*

[10] J. L Synge, A. Schild, *Tensor Calculus* (Dover Publications, inc. New York).

[11] S. Lang, *Differential and Riemannian Manifolds* (Springer Verlag).

[12] H. Goldstein, C. P. Poole, J. L. Safko, *Classical Mechanics, third edition* (Addison Wesley).

[13] V. I. Arnol'd, *Mathemathical Methods of Classical Mechanics, second edition*, (Springer Verlag).

[14] L. D. Landau, E. M. Lifŝits, *Fisica Teorica 1, Meccanica* (Editori Riuniti, Roma).

[15] L. D. Landau, E. M. Lifŝits, *Fisica Teorica 2, Teoria dei Campi* (Editori Riuniti, Roma).

[16] R. L. Arnowitt, S. Deser, and C. W. Misner, *The Dynamics of general relativity*, url: https://arxiv.org/pdf/gr-qc/0405109.pdf.

[17] Bernard F. Schutz, *A first course in general relativity*, (Cambridge University Press. 1980).

[18] Bernard F. Schutz, *Geometrical methods of mathematical physics*, (Cambridge University Press. 1980).