UNIVERSITÉ D'ORLÉANS

ÉCOLE DOCTORALE
Énergie - Matériaux - Sciences de la Terre et de l'Univers (EMSTU)
Laboratoire de Physique et Chimie de L'Environnement et de l'Espace

THÈSE présentée par :

Abedennour DIB

soutenue le : **19 Juin 2026**

pour obtenir le grade de : **Docteur de l'Université d'Orléans**

Discipline/ Spécialité : Physique

Théories Maxwelliennes Généralisées :
Structure théorique

THÈSE dirigée par :
M. SPALLICCI Alessandro D.A.M. PR émérite, Université d'Orléans - CNRS - LPC2E

RAPPORTEURS :
M. MAVROMATOS Nikolaos E. PR, Ethnikó Metsóvio Polytechneío, Athína - CERN Genève - King's College, London
M. POTTING Robertus PR, Universidade do Algarve, Faro

JURY :
M. CELESTIN Sébastien PR, Université d'Orléans, CNRS – LPC2E, Président du jury
M. VAN HOLTEN Jan-Willem PR émérite, Universiteit Leiden – Nationaal Instituut voor Kernfysica en Hoge-Energiefysica, Amsterdam
M. CAPOZZIELLO Salvatore PR, Università degli Studi di Napoli Federico II – Scuola Superiore Meridionale, Napoli
M. HELAYËL-NETO José A. DR, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro
M. LÄMMERZAHL Claus PR émérite, Universität Bremen

INVITÉS :
M. PROKOPEC Tomislav Dr, Universiteit Utrecht

Contents

1 Theoretical structure of Electrodynamics **6**

1.1 Classical electrodynamics 11

1.1.1 Historical overview 11

1.1.2 Maxwell's equations 14

1.1.3 Potential formulation of Maxwell's equations 17

1.1.4 Covariant formulation of Maxwell's theory 19

1.1.5 Hamiltonian picture 28

1.2 Quantum Electrodynamics (QED) 31

1.2.1 Quantisation of the free electromagnetic field 31

1.2.2 QED and Feynman rules for electrodynamics 36

1.2.3 The recipe 42

2 Extended Theories of Electrodynamics **44**

2.1 Massive Extensions 48

2.1.1 de Broglie-Proca Action 48

2.1.2 Stueckelberg Trick 53

2.1.3 Bopp-Landé-Thomas-Podolsky Model 55

2.1.4 Maxwell-Chern-Simons Theory and topological mass generation . 58

2.1.5 Bounds on the photon mass 63

2.2 Non-Linear Theories 64

2.2.1 Born-Infeld model 64

2.2.2 Heisenberg-Euler theory 66

2.3 Lorentz Symmetry Violations 68

2.3.1 Carroll-Field Jackiw Lagrangian 70

2.3.2 Photon sector of the Standard Model extension 73

3 Generalised Non Linear Electrodynamics **76**

3.1 Background Field Formalism 80

3.2 Construction of the action, equations of motion and energy-momentum-tensor . . . 85
3.3 Dispersion relations and group velocities . . . 93
3.3.1 $\mathcal{L}(\mathcal{F})$ with constant $\vec{E}$ and $\vec{B}$ . . . 93
3.3.2 $\mathcal{L}(\mathcal{F},\mathcal{G})$ with a purely magnetic field . . . 96
3.3.3 $\mathcal{L}(\mathcal{F},\mathcal{G})$ with constant $\{\vec{E},\vec{B}\}$ . . . 99

4 Mass generation in Generalised Non-Linear Electrodynamics 103
4.1 Field redefinition . . . 107
4.2 Equations of motion and dispersion relations . . . 111
4.3 Hamiltonian picture . . . 115
4.4 A general approach . . . 119

5 Perspectives and Conclusions 132
5.1 Other research work . . . 133
5.2 Perspectives . . . 135
5.3 Conclusion . . . 136

Appendices 138

A Field Theory in a nutshell 139

B Constrained Hamiltonian Systems 142
B.0.1 Classification of constraints . . . 145
B.0.2 Dirac Bracket . . . 147

C Green's functions and Propagators 150

Acknowledgments

This thesis has been the culmination of over three years of hard work, sleepless nights, and unending discussions. While the author slot bears my name, I am eternally indebted to many people for its success.

I will start by thanking Prof. Nikolaos Mavromatos, Prof. Robertus Potting, Prof. Salvatore Capozziello, Prof. Claus Lämmerzahl, Prof. Sébastien Célestin, Prof. Jan Willem Van Holten and Prof. Tomislav Prokopec for accepting to be a part of my jury and granting me the title of doctor. Their advices, remarks and feedbacks have been instrumental in devising this version of the manuscript. I will insist on thanking Prof. Van Holten who also acted as my scientific advisor for the past three years, from whom I learnt a lot, and Prof. Prokopec for the countless hours of lively (and lovely) discussions we exchanged over the past few months.

I will then continue by thanking my two academic fathers, Prof. Alessandro Spallicci and Prof. José Helayel-Neto. I cannot count the hours we spent talking, laughing and learning all together, in what we called in the beginning the connection Oran-Paris-Rio, but all that I can say for certainty is that I am the scientist I am today thanks to the two of you. Your advice, patience and kindness allowed me to grow in one of the best possible environments, and I will be eternally indebted to you. When I stated that you are my academic fathers, it was an understatement, and I should've said that you were my spiritual fathers instead.

I want to also thank all of the friends and colleagues I met along the way. Starting with the LPC2E, in which I cannot cite everyone, but will nonetheless explicitly name Yanis (also known as our supreme commander), Pietro, Nicolas (Jr), Nicolas (Sr), Panisara, Mehul, Rima, Nina, Chirelle, Hannah, Sébastien, Clara and all of those I am forgetting. Thank you all for being a part of this adventure, and for having to play babyfoot with me. You made the lab a great place to be at, and I shared with all of you amazing memories. I will also thank the many people I met in conferences,

visits, and other tribulations: Altin, Volker, Claus, Salvatore, Emmanuele, Francesco, Filippo, Pietro, Maria, Francesca, Damos, Ryan, Ilia and all of those I forgot. Our exchanges (scientific, or not so scientific) helped me grow a lot, and get a better, fuller picture of the landscape in which I existed. I would also like to thank the original Rusty boys: Carlos, Sebastian, Ross, Luke and Gabriel for being there for me when things got tough, and taking me out when I wanted to curl up at home. I will miss all of you dearly.

I will conclude by thanking those who stayed behind and my loved ones. I will start with those who contributed to my academic journey, Prof. Iddir, Prof. Lagraa, Dr. Rahmouni and first and foremost Dr. Nazim Djeghloul. Thank you for believing in me when no one did, and when I was a rather mediocre student. You've given me a chance and I will not let it go to waste. Then of course, I will thank my mother Nadia, my father Allal, and my sister Doria. You have been always there for me, you pushed me to be where I am and who I am, thank you for the life you gave me. And of course, thank you Sabrina. I don't have the words to summarise all of what you did, so I will say the following: you have been with me since the beginning of my academic journey, and have seen all of my struggles and successes. You were always there, never faulting, and I will never, ever be able to thank you enough for this. When people were laughing when I stated that I'd be a theoretical physicist, you'd shout at them, and today, we've proved them wrong.

Luck was a spark, and hard work kept the fire going.

Chapter 1

Theoretical structure of Electrodynamics

Résumé

La théorie de l'électromagnétisme de Maxwell est un des piliers de la physique moderne. Au delà des phénomènes électriques et magnétiques, elle permet surtout de décrire la lumière, le messager favori de la nature. Dans ce chapitre, nous explorons la structure théorique de l'électrodynamique classique ainsi que quantique, afin d'établir la liste des ingrédients fondamentaux de cette théorie. Le point de départ est constitué par les équations de Maxwell, qui relient les champs électrique $\vec{E}$ et magnétique $\vec{B}$ aux densités de charge ρ et de courant $\vec{J}$

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \ , \tag{1.1}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}, \tag{1.2}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \ , \tag{1.3}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \ . \tag{1.4}$$

Toutefois, ces équations ne doivent pas être envisagées comme un système homogène: elles contiennent en réalité deux types de relations distinctes. D'une part, les équations de divergence jouent le rôle de contraintes agissant sur le système et d'autre part, les équations faisant intervenir les rotationnels gouvernent l'évolution temporelle des champs. Ceci indique que toutes les composantes des champs $\vec{E}$ et $\vec{B}$ ne sont pas indépendantes, et semble suggérer une structure sous-jacente plus riche.

Qui plus est, en prenant la divergence de la loi d'Ampère Maxwell (1.20) and en utilisant la loi de Gauss(1.19), nous obtenons une équation de continuité qui décrit la conservation des charges dans un volume fermé:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \ . \tag{1.5}$$

Autrement dit, la forme même des équations impose des restrictions sur les sources admissibles. La théorie ne décrit donc pas uniquement l'évolution des champs: elle encadre également la manière dont les charges et les courants peuvent exister et évoluer.

En l'absence de sources, les équations se simplifient et révèlent la dynamique propre du champ électromagnétique. On obtient alors des équations d'onde, dont les solutions décrivent la propagation de perturbations à une vitesse finie (celle de la lumière)

$$\omega^2 = c^2 k^2 \ . \tag{1.6}$$

Nous savons néanmoins que les ondes lumineuses possèdent deux degrés de liberté,

chose qui n'est pas trivial à démontrer dans le cas de la formulation champs de l'électromagnétisme. C'est dans ce contexte que nous introduisons la formulation potentielle de la théorie. Les champs électrique et magnétique sont exprimés à partir d'un potentiel scalaire φ et d'un potentiel vecteur $\vec{A}$

$$\vec{E} = -\vec{\nabla}\varphi - \frac{\partial \vec{A}}{\partial t} \quad , \tag{1.7}$$

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad , \tag{1.8}$$

ce qui permet de satisfaire automatiquement une partie des équations de Maxwell et met en évidence que la description en termes de champs comporte des redondances. Les potentiels apparaissent alors non pas comme des outils auxiliaires, mais comme les véritables variables fondamentales de la théorie.

Cependant, cette description introduit une ambiguïté essentielle: les potentiels ne sont pas définis de manière unique. En effet, une transformation appropriée laisse les champs physiques invariants

$$\varphi' = \varphi - \frac{\partial \lambda}{\partial t} \ , \tag{1.9}$$

$$A' = A + \vec{\nabla}\lambda \ . \tag{1.10}$$

Cette propriété, appelée invariance de jauge, révèle que la théorie possède une redondance locale. L'invariance de jauge joue un rôle central de la structure théorique, et est l'un des ingrédients les plus importants de la théorie. Elle implique que toutes les composantes des potentiels ne correspondent pas à des degrés de liberté physiques indépendants et qu'il est nécessaire d'imposer des conditions supplémentaires pour éliminer les variables superflues.

La reformulation relativiste de la théorie permet de clarifier cette structure. Les potentiels sont rassemblés en un unique quadrivecteur $A^\mu(\varphi, \vec{A})$, et les champs électromagnétiques sont décrits par un tenseur antisymétrique $F^{\mu\nu}$ construit à partir de ceux-ci . Cette écriture unifie les composantes électrique et magnétique et rend explicite l'invariance de Lorentz. Elle montre que l'électromagnétisme est naturellement une théorie relativiste, dont la forme ne dépend pas du choix du référentiel.

La dynamique du champ est ensuite dérivée à partir d'un lagrangien unique, invariant à la fois sous transformations de Lorentz et de jauge

$$\mathcal{L} = -\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu} - A_\mu J^\mu \ . \tag{1.11}$$

L'ensemble des équations de Maxwell en découle par application du principe de moindre

action. Cette formulation présente l'avantage de concentrer toute la structure théorique dans un seul objet scalaire, dont les propriétés de symétrie déterminent entièrement la dynamique, ce qui facilite l'extension future de la théorie.

Le passage à la formulation hamiltonienne met en évidence un aspect fondamental jusque-là implicite: la présence de contraintes. En calculant les variables conjuguées, on constate que la composante π^0 ne possède pas de dynamique, indiquant la nature auxiliaire du champ A^0, ce que nous appelerons contrainte primaire. L'évolution dans le temps de la contrainte primaire mène à une contrainte secondaire, la loi de Gauss. En utilisant la classification de Dirac, nous montrons que ces deux contraintes sont dites de première classe, ce qui signifie qu'elle génèrent les transformations de jauge. Cela a pour mérite d'immédiatement réduire les nombres de degrés de liberté physique à deux, sans avoir à imposer de jauge. Nous insistons à ce moment sur l'importance du formalisme Hamiltonien contraint pour discuter des théories de jauge.

La quantification du champ met en lumière des difficultés supplémentaires. Une procédure directe conduit à l'apparition d'états de norme négative, ce qui est incompatible avec une interprétation probabiliste. Ce problème reflète le fait que la redondance de jauge n'a pas été complètement éliminée au niveau quantique. Pour y remédier, une condition supplémentaire est imposée sur les états physiques, permettant de sélectionner un sous-espace cohérent tout en conservant l'invariance relativiste, ce que l'on appelle la quantification à la Gupta-Bleuler.

L'extension de la théorie à l'électrodynamique quantique repose sur l'introduction d'un couplage minimal entre le champ électromagnétique et les champs de matière. Ce couplage n'est pas arbitraire: il est imposé par l'invariance de jauge. Ainsi, l'interaction électromagnétique apparaît comme une conséquence directe de la symétrie fondamentale de la théorie.

Enfin, l'analyse des corrections quantiques révèle la présence de divergences dans les calculs perturbatifs. La cohérence de la théorie est assurée par des identités reliant les différentes amplitudes (tel l'identité de Ward), qui garantissent la conservation du courant et la préservation des propriétés fondamentales du photon. Ces résultats montrent notamment que, pour que le photon puisse posséder une masse dite physique, il faudrait que le pôle du propagateur soit décalé à tous les ordres de la série perturbative lors du calcul de la polarisation du vide. Nous concluons donc ce chapitre en établissant la liste d'ingrédients clés de l'électromagnétisme, en commençant par les ingrédients émergeants de la théorie classique

- La linéarité.
- L'invariance de Lorentz .
- L'invariance de jauge sous les transformations du groupe $U(1)$.

- L'absence de masse pour le photon.
- Ainsi que l'absence de monopoles magnétiques.

Nous notons que la quantification de la théorie, et l'introduction d'intéractions avec les fermions, tue directement le premier ingrédient! La linéarité de l'électromagnétisme n'est donc qu'un artéfact dû à une approximation à basse énergie. Quant aux ingrédients issus de la théorie quantique

- Les états physiques ne sont représentés que par les deux polarisations transverses.
- L'unitarité (Ou absence de ghosts).
- La renormalisabilité de la théorie.
- La présence des identités de Ward.
- La présence de divergences infrarouges.
- Absence de décalage du pôle après renormalisation (La théorie doit rester sans masse).

Cette liste nous permet notamment d'établir les trois grandes familles d'extensions que nous allons aborder dans le chapitre suivant: les théories massives, non-linéaires classiquement, ainsi que les théories violants l'invariance de Lorentz.

1.1 Classical electrodynamics

1.1.1 Historical overview

Lightening and electricity have played a central role in mankind's history. Every pantheon possessed a deity responsible for the generation of lightening, and most of them were considered of high rank. Whether it was Zeus, Odin or Raijin, they all represented the fascination and awe that electric phenomena inspired to their respective populations. Other, less impressive, electric phenomena such as static electricity have also been known since Thales of Miletus. Indeed ancient greeks knew that by rubbing an amber stone (pronounced as ēlektron in ancient Greek), light objects would get attracted. They naively assumed this property to be similar to the Magnesian stone, also known as the Lodestone.

The behaviour of Lodestone was not due to electricity, but to its less deified twin, magnetism. Naturally occurring magnetic phenomena were less impressive than a lightening bolt hitting the ground and as such did not extend every major pantheon with a magnetic deity. Nonetheless, it played a key role in the development of compasses which allowed humanity to explore more freely the seas and earth.

The fusion of the two however, did inspire poets, mythologies and writers alike, as light came to be a central theme in arts and notably painting. First associated with the sun, it quickly became a synonym for all that is good and pure, a symbol of the divine. The study of light (or optics) is almost as ancient as humanity itself, as the first traces of lenses we have date back to 2000 BC. The field evolved very quickly through the Greeks, Arabs and culminated with Fermat's and Descartes work birthing geometrical optics.

While all three are ancient, and despite sharing the same speed of propagation in a vacuum, the scientific fields did not evolve at the same pace. The first one to bloom being optics, as its applications in lenses, glasses, and such were critical to the advancement of civilisation. In the context of optics, one essential contribution came from Sir Isaac Newton, when he decomposed white light through a prism and proved that it is nothing but the sum of colors. With this observation, he expanded Descartes's work and created the corpuscular picture of light.

To Newton, light was composed of particles moving at infinite speed from point A to B. This view clashed with Hooke's and Huygens representation of light waves, and the feud between Newton and Hooke last until the passing of the latter. Newton being Newton, his detractors were quickly dismissed and corpuscular light was the leading model until the 19^{th} century when Thomas and Fresnel conducted their famous interference and diffraction experiments proving the wave-like nature of light.

In the meantime, the twin brothers caught up, and scientists began to discuss more

electricity and magnetism. Citing an exhaustive summary of all the different contributions to the field of electromagnetism is beyond the scope of this thesis[1]; we shall keep this brief. Probably the most important early contribution was due to Charles-Auguste Coulomb with his torsion balance experiment. Using the setup illustrated below

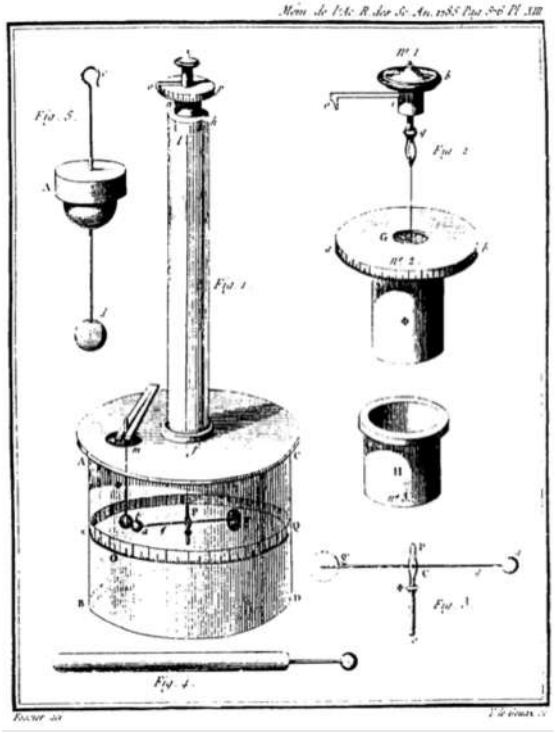


Figure 1.1: Coulomb's original setup

he discovered that the force exerted between two charges is proportional to the intensity of said charges and inversely proportional to the squared distance separating them:

$$\vec{F}_{Coulomb} = k\frac{q_1 q_2}{r^2}\vec{n} \tag{1.12}$$

where q_1, q_2 are charges, r is the distance between them, $\vec{n}$ is the unit vector indicating the direction between the charges, and k is a constant known today as Coulomb's constant. Credit where credit is due; Coulomb was not the first scientist to measure such a proportionality. Indeed, Robinson and Cavendish devised different experiments that obtained similar (and better) results before Coulomb, but did not publish their results, as their setups was more complicated than Coulomb's and thus gave results that were harder to interpret .

[1]For a more extensive read on the history of electromagnetism, I recommend The Lightning Tamers by Kathy Joseph

After Coulomb's empirical work, scientists became intrigued with electric and magnetic phenomena, and multiplied the experiments to understand better this new field. Moving away from static charges, Kirchhoff devised his laws for electric circuits laying the foundations of electrokinetics, Kelvin studied the discharges in condensers and more importantly Ørsted discovered that the flow of electric currents caused the needle of compasses to move, which was the first quantitative evidence of an interaction between electricity and magnetism. This culminated with Ampère's theory, which was the corner stone of electrodynamics.

However, the true revolution of experimental electromagnetism happened a bit later in time, thanks to the experiments of Michael Faraday, which ruled conclusively that electric and magnetic phenomena were not independent of each other, and could be induced by one another. Unfortunately, Faraday was not well versed in mathematics, and he could not go beyond the empirical results he obtained.

On the mathematical side, the contributions of Gauss, Ostrogradski, Stokes and later on Green, laid the foundations on which Maxwell built upon to construct his unified theory of electrodynamics. Gauss's investigations into electrostatics led him to formulate what is now known as Gauss's law, which relates the total electric flux through a closed surface to the amount of electric charge enclosed within it. Although originally expressed in integral form, this law reflects a deeper geometric idea: that the sources of a field can be detected by examining how the field "flows" through surrounding boundaries. Continuing, Green provided a more general analytical framework in his essay on potential theory. Green introduced what is now known as Green's theorem, which connects a line integral around a closed curve to a double integral over the region it encloses. He extended Gauss's pioneering work by showing how boundary conditions determines behaviour inside an enclosed region, a central idea for electrostatics.

The next major step came from Ostrogradski, who generalised these boundary–volume relationships into three dimensions. His divergence theorem (often called the Gauss–Ostrogradski theorem) rigorously established that the flux of a vector field through a closed surface equals the integral of its divergence over the volume inside. This result lead to the differential form of Gauss's law, transforming it from a global statement about flux into a local statement about charge density. This shift—from integral to differential formulations—was crucial for the later development of field theory. Finally, Stokes extended these ideas further with his theorem, which relates the circulation of a vector field around a closed loop to the curl of the field across the surface bounded by that loop. This theorem unified the earlier results (including Green's theorem) and became essential in expressing the laws of electromagnetism—such as Faraday's law of induction and Ampère's law—in both integral and differential forms.

Maxwell's contributions went far beyond formalising Faraday's observations, as he

showed that not only electricity and magnetism were one and the same, but also that light is a result of the electromagnetic field.

1.1.2 Maxwell's equations

Maxwell's theory of electromagnetism was in its original formulation composed of a set of 20 equations [1,2] and were brought in their more familiar form by Oliver Heaviside [3], which contributed greatly to the emerging field of electrodynamics.

The set of equations written down by Heaviside was akin to the following more modern formulation [4]:

$$\vec{\nabla} \cdot \vec{D} = \rho_f \ , \tag{1.13}$$

$$\vec{\nabla} \times \vec{H} = \vec{J_f} + \frac{\partial \vec{D}}{\partial t} \ , \tag{1.14}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \ , \tag{1.15}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \ , \tag{1.16}$$

where $\vec{D}$ is the displacement vector, $\vec{H}$ is the magnetic field strength (also known as electromagnetic excitation), $\vec{E}$ is the electric field, $\vec{B}$ the magnetic field, ρ the free electric charge and $\vec{J}$ the free electric current density. This is known as the macroscopic formulation of the Maxwell equations, where the influence of matter is incorporated within the $\vec{D}$ and $\vec{H}$ fields, leaving only the free charges and currents.

Naively, it is possible to interpret these equations as describing a set of four independent vector (and pseudo-vector) fields. Thankfully, the vectors $\vec{D}$ and $\vec{H}$ are auxiliary fields and can be linked to the vectors $\vec{E}$ and $\vec{B}$ through the following constitutive relations [4]:

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P} \ , \tag{1.17}$$

$$\vec{H} = \frac{1}{\mu_0} \vec{B} - \vec{M} \ , \tag{1.18}$$

where $\vec{P}$ is the polarisation vector, $\vec{M}$ is the magnetisation vector, ε_0 is the vacuum permittivity, and μ_0 the vacuum permeability.
The constitutive relations of electrodynamics play a key role in describing the matter propagation of electromagnetic waves, and are open to change depending on the type of media (inhomogeneous, non-linear, etc.). These different constitutive relations, their origin and interpretation, although not discussed in this thesis (as our work will focus

on the microscopic (also known as vacuum) formulation of Maxwell and its extensions), possess an extensive literature [5–8].
In the absence of a medium, the constitutive relations become a one-to-one equivalence between the excitations and the fields, and we can thus rewrite the four equations in their most familiar form

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \ , \tag{1.19}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \ , \tag{1.20}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \ , \tag{1.21}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \ , \tag{1.22}$$

these are then referred to as the microscopic formulation of the Maxwell equations. They can be further subdivided into two groups: the source equations and the circulation equations.

The source (or divergence) equations (1.19) and (1.21), are static equations tell us about the origin of the fields. They are referred to as the electric and magnetic Gauss laws, and give us two key informations:

- The source of the electric field are the electric charges that exist in nature. This expresses itself more formally in the sense of electric monopoles, of positive or negative sign.
- There is no magnetic monopoles. That is, we can never split a magnet into two independent north and south poles. The source of the magnetic fields that we observe in nature is always the reciprocal electric fields. This particular feature will be discussed in more details later in this chapter.

Meanwhile, the circulation (or curl) equations (1.20) and (1.22), are all about the dynamics of the theory. They show the intertwined nature of electric and magnetic fields. The time variation of one, induces the second and vice versa. These equations were the true revolution that stemmed from Maxwell: we should not be looking at the phenomena as due to either electric and magnetic fields, but as a singular electromagnetic field. By taking the divergence of equation (1.20) and using (1.19) we obtain a continuity equation that describes the conservation of charges in a closed volume

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \ . \tag{1.23}$$

Following the historical route, we can search for wave solutions to these equations. We shall assume the simpler case where the wave is propagating in a totally empty

space, take ρ and $\vec{J}$ to be null, and write down our equations.

$$\vec{\nabla} \cdot \vec{E} = 0 \ , \tag{1.24}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \ , \tag{1.25}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \ , \tag{1.26}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \ , \tag{1.27}$$

we then continue by taking the curl of the curl equations and using the identity

$$\vec{\nabla} \times \left(\vec{\nabla} \times \vec{V} \right) = \vec{\nabla} \left(\vec{\nabla} \cdot \vec{V} \right) - \nabla^2 \vec{V} \ , \tag{1.28}$$

we show that

$$\mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2 \vec{E} = 0 \ , \tag{1.29}$$

$$\mu_0 \varepsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} - \nabla^2 \vec{B} = 0, \tag{1.30}$$

which are standard wave equations. Choosing a standing wave solution for each field

$$\vec{E}(t, \vec{r}) = \vec{E_0} \sin(\omega t - \vec{k} \cdot \vec{r}) \ , \tag{1.31}$$

$$\vec{B}(t, \vec{r}) = \vec{B_0} \sin(\omega t - \vec{k} \cdot \vec{r}) \ , \tag{1.32}$$

we obtain twice the same dispersion relation

$$\omega^2 = \frac{1}{\mu_0^2 \varepsilon_0^2} k^2 \ , \tag{1.33}$$

where ω is the wave frequency and $k = |\vec{k}|$ is the amplitude of the wave vector. From the dispersion relation, it was then a straightforward step to the group velocity

$$v_g = \frac{\partial \omega}{\partial k} \quad \Rightarrow v_g = \frac{1}{\mu_0 \varepsilon_0} \qquad . \tag{1.34}$$

We can draw two conclusions from the above results:

- First, the electric and magnetic fields propagate with the same velocity and share the same dispersion relation, implying that they are indeed two facets of the same entity.
- Second, we can estimate their velocity by measuring $\mu_0 \epsilon_0$, which is exactly what Maxwell did. And he very quickly realized that the value he obtained for the

velocity of the waves, was extremely close to that of what was known at the time as the speed of light. From now on, we shall refer to $\mu_0\varepsilon_0$ as $\frac{1}{c}$ where c is the speed of light in a vacuum.

In hindsight, it might seem simple to obtain these results, but they were revolutionary for their era. Maxwell achieved what was at that time the greatest unification of natural sciences, he merged in one single theory electricity, magnetism, and optics.

1.1.3 Potential formulation of Maxwell's equations

When it was understood that light waves were nothing more than electromagnetic waves, a problem emerged with the degrees of freedom of the theory. Strictly speaking, both the electric and magnetic field propagate three degrees of freedom, giving a total of 6 for the theory. Taking into account the divergence equations, which act as constraints, we can reduce the total to 4. However,in optics, it was known due to Malus's law, that light propagates only two polarisations, leaving us with two additional degrees of freedom on our hands.

There are of course ways around this issue with the fields formulation, but we shall stray from the historical approach to introduce a simpler solution to the issue at hand. We start by introducing the electric potential φ also known as scalar potential, and the magnetic potential $\vec{A}$ otherwise called vector potential [4]. We can then find the associated fields to the potentials and get

$$\vec{E} = -\vec{\nabla}\varphi - \frac{\partial \vec{A}}{\partial t} \quad , \tag{1.35}$$

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad , \tag{1.36}$$

which automatically satisfy both equation (1.21) and (1.22) due to the two following identities

$$\vec{\nabla} \cdot \vec{\nabla} \times \vec{V} = 0 \ , \tag{1.37}$$

$$\vec{\nabla} \times \vec{\nabla} f = 0 \ . \tag{1.38}$$

This reduces the dynamics of the model to the only inhomogeneous equations giving

$$\nabla^2\varphi + \frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{A}) = -\frac{\rho}{\varepsilon_0} \quad , \tag{1.39}$$

$$\left(\nabla^2 \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2}\right) - \vec{\nabla}\left(\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2}\frac{\partial \varphi}{\partial t}\right) = -\mu_0 \vec{J} \quad , \tag{1.40}$$

where the first equation is a scalar, and the second a vector, totalling 4 degrees of freedom a priori.
These two equations possess a very intriguing property. Indeed, if we redefine the potentials by adding an arbitrary function λ in the following manner

$$\varphi' = \varphi - \frac{\partial \lambda}{\partial t} \ , \tag{1.41}$$

$$A' = A + \vec{\nabla}\lambda \ , \tag{1.42}$$

we see that the equations are unchanged. This feature is referred to as gauge invariance, and highlights a fundamental property of electrodynamics: electric and magnetic fields are physical observables, potentials are not as they can always be rescaled by an arbitrary fashion. Gauge invariance as a concept will be expand further later in this thesis, but for the moment we shall describe it only as a redundancy in our tools. There are more degrees of freedom in our equations than what is physically relevant. By using this property, we can *fix* a gauge condition that will greatly simplify our equations of motion.
Two popular choices of gauge fixings are the Lorenz gauge and the Coulomb gauge: The Lorenz gauge (and not **Lorentz** gauge) is defined as follows

$$\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2}\frac{\partial \varphi}{\partial t} = 0 \ , \tag{1.43}$$

The advantage of this gauge fixing, is that aside from preserving covariance (an important feature we will discuss later) it gives to the equations for the potentials a proper wave equation form

$$\frac{1}{c^2}\frac{\partial \varphi}{\partial t} - \nabla^2 \varphi = \Box \varphi = \frac{\rho}{\varepsilon_0} \ , \tag{1.44}$$

$$\frac{1}{c^2}\frac{\partial \vec{A}}{\partial t} - \nabla^2 \vec{A} = \Box \vec{A} = \mu_0 \vec{J} \ , \tag{1.45}$$

where $\Box$ is the d'Alembert operator. This gauge fixing reduces the total degrees of freedom by one, leaving us with three. We can however see that the Lorenz gauge does not totally fix the potential, as functions that verify

$$\Box \lambda = 0 \ , \tag{1.46}$$

non trivially (such as harmonic function) leave (1.43) invariant, also called residual gauge. We must then fix this residual gauge by a condition that is compatible with the initial one (in our case (1.43). In this situation a common choice of gauges is the

Coulomb gauge which simply consists on killing the electric potential

$$\varphi = 0 \ , \tag{1.47}$$

leaving only the magnetic potential. When implemented in the Lorenz gauge, it will automatically lead to

$$\vec{\nabla} \cdot \vec{A} = 0 \ , \tag{1.48}$$

fixing an additional degree of freedom, leading to a total of two. Additionally, (1.46) directly leads to the transverse property of electromagnetic waves in a vacuum. Starting from the Lorenz fixed equations of motion in a vacuum

$$\Box\varphi = 0 \ , \tag{1.49}$$

$$\Box\vec{A} = 0 \ . \tag{1.50}$$

By then fixing the Coulomb gauge, we reduce the dynamics to only

$$\Box\vec{A} = 0 \ , \tag{1.51}$$

which admits standing waves solutions

$$\vec{A}(\vec{x},t) = \vec{\epsilon}\Big(a(k)e^{i(\omega t - \vec{k}\cdot\vec{x})} + a^{+}(k)e^{-i(\omega t - \vec{k}\cdot\vec{x})}\Big) \ , \tag{1.52}$$

where $\vec{\epsilon}$ is the polarisation vector, and $a^{+}(k)$ is the complex conjugate of the wave amplitude $a(k)$ to ensure the reality of $\vec{A}(\vec{x},t)$. By replacing the solution in the equation of motion we obtain the dispersion relation

$$\omega^2 = c^2 \ \vec{k}\cdot\vec{k} \ , \tag{1.53}$$

and by injecting (1.52) in (1.46) we show that

$$\vec{\epsilon}\cdot\vec{k} = 0 \ \Rightarrow \vec{\epsilon} \perp \vec{k} \ , \tag{1.54}$$

meaning that the polarisation is perpendicular to the direction in which the wave is propagating, confirming that we are dealing with transverse waves.

1.1.4 Covariant formulation of Maxwell's theory

So far we have only discussed the structural features of Maxwell's theory, but have not considered its dynamical features, notably when changing frames. Going down this road

will lead us to the covariant formulation of electromagnetism, but beforehand we shall briefly introduce special relativity. In classical mechanics, reference changes obey the Galilean transformations [9] which read

$$t' = t \tag{1.55}$$
$$x' = x - vt \tag{1.56}$$
$$y' = y \tag{1.57}$$
$$z' = z \ , \tag{1.58}$$

which assume that space and time are two distinct, absolute objects. We can immediately see some problems with the Maxwell equations when considering these transformations. Indeed, it is easy to show that the conjectural equations (*i.e.* Ampère-Maxwell and Gauss law) are not invariant under the Galilean transformations. Worse! We can also show rather easily that the wave operator $\Box = \frac{\partial^2}{\partial r^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$ is not either an invariant. This would mean that under a coordinate change, the dispersion relation of the wave would not be the same, making it by essence unphysical. However, by relaxing the absolute nature of space and time, and accepting the fact that the speed of light is an invariant, we can achieve what is known as the Lorentz-Poincaré transformations (Often abridged as Lorentz transformations)

$$\begin{aligned} t' &= \gamma\left(t - \frac{vx}{c^2}\right) \\ x' &= \gamma\left(x - vt\right) \\ y' &= y \\ z' &= z \end{aligned} \quad , \tag{1.60}$$

which preserve the nature of the Maxwell equations in every reference frame, while treating the wave operator as a true relativistic invariant. But we may then ask, what is truly a relativistic invariant? Let us shed some light on the idea!

Suppose that our displacement is represented by the vector $d\vec{x} = (dx^1, dx^2, dx^3) = (dx, dy, dz)$. In Newtonian mechanics, the distance travelled can be trivially obtained through the Pythagorean theorem, such as

$$dS^2 = dx^2 + dy^2 + dz^2 \ , \tag{1.59}$$

this shows that even though two observers can describe the event in two independent

coordinate systems, the amplitude of the latter is untouched[2]. This result does not hold in special relativity, as we have to take into account the absolute speed of light for every observer; we must then recast it at an instant t in the following form:

$$dS^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2) = c^2dt^2 - d\vec{x}^2 = c^2d\tau^2 \ , \tag{1.60}$$

where τ is the proper time. It is very simple to notice that the above expression is invariant under Lorentz transformations. Through the space-time interval, we can define two objects: the 4-vector x^μ where $\mu = 0, 1, 2, 3$ and the Minkowski metric $\eta^{\mu\nu}$.

The metric tensor is a geometrical object that defines the distances and angles between two points in a vector space. An arbitrary metric has the following properties

$$\begin{aligned} &g(v_1 \cdot v_2) = g(v_2 \cdot v_1) && \text{Symmetry} \ , \\ &g(\alpha a + \beta b, c) = \alpha g(a \cdot c) + \beta g(b \cdot c) && \text{Linearity} \ , \end{aligned}$$

where α, β are arbitrary scalars and a, b, c elements of the vector space E. If $\forall a \in E$

$$g(a, b) = 0 \Rightarrow b = 0 \ ,$$

the metric is said to be non-degenerate and positive definite. In an orthonormal reference frame, the $g_{ij} = \delta_{ij}$.

Using the Minkowski metric, we can rewrite our space-time interval as

$$dS^2 = \eta_{\mu\nu}dx^\mu dx^\nu \ , \tag{1.61}$$

where we introduced the Einstein notation[3] on μ et ν; $\eta^{\mu\nu}$ can be expressed as a diagonal matrix of elements $(1, -1, -1, -1)$ which is also known as its signature.

A four-dimensional space possessing a coordinate system x^μ and a Minkowski metric $\eta^{\mu\nu}$ is called a Minkowski space-time. The Minkowski metric will allow us to move trivially from covariant to contravariant vector, and as such we will not explicit the operation every time we use it. The metric also defines the trace and dot product operation on the vector space.

Before moving to the definition of scalars, vectors and tensor in the Lorentz sense, we will introduce one additional key ingredient: the Levi-Civita tensor. The Levi-Civita

[2]This is also known as the Euclidean metric.

[3]The Einstein notation simply states that if two indices are repeating there is an implicit sum $V_iA^i = \sum_{i=0}^{i=n} V_iA_i$

tensor (symbol) is a totally antisymmetric object (which we will define below) that allows us to move to the dual basis of a theory. In an index free notation, it is equivalent to a Hodge product. It will notably allow us to transform electric **vector** fields to magnetic **pseudo-vector** fields. In 4 dimensions the Levi-Civita tensor[4]is defined as follows

$$\epsilon_{\mu\nu\rho\sigma} = \begin{cases} 1 & \text{if } (\mu,\nu,\rho,\sigma) \text{ is an even permutation of } (0,1,2,3) \\ -1 & \text{if } (\mu,\nu,\rho,\sigma) \text{ is an odd permutation of } (0,1,2,3) \\ 0 & \text{for any other case} \end{cases} . \tag{1.62}$$

We will now close this section on special relativity by introducing the definitions of scalars, vectors, and tensors. We define the following transformations:

$$\frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu \ ,$$

also known as coordinate transformations, they are strictly equivalent to the Lorentz transforms and are considered as invertible. A contravariant vector will then be defined as

$$V'^\mu(x') = \frac{\partial x'^\mu}{\partial x^\alpha} V^\alpha(x) \ . \tag{1.63}$$

A covariant vector on the other hand will be defined as

$$V'_\mu(x') = \frac{\partial x^\alpha}{\partial x'^\mu} V_\alpha(x) \ , \tag{1.64}$$

and one can move from a covariant to contravariant basis using a metric tensor

$$V^\mu = \eta^{\mu\nu} V_\nu \ . \tag{1.65}$$

A scalar is defined as an invariant in the Lorentz sense, as such

$$\phi'(x) = V'^\mu(x') W'_\mu(x') = \frac{\partial x'^\mu}{\partial x^\alpha} V^\alpha(x) \frac{\partial x^\beta}{\partial x'^\mu} W_\beta(x) = V^\alpha(x) W_\alpha(x) = \phi(x) \ . \tag{1.66}$$

An arbitrary rank 2 contravariant tensor will be defined as

$$T'^{\mu\nu}(x') = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} T^{\alpha\beta}(x) \ , \tag{1.67}$$

so on and so forth for higher-order objects. For all intents and purposes we shall treat scalars and vectors as rank0 and rank 1 tensors, respectively.

[4]This is an abuse, as the Levi-Civita is not only pseudo-tensorial (as it projects on the dual basis) but more strictly speaking not a tensor at all since it does not obey to the Lorentz transformations.

A rank-2 tensor can be symmetric if $T^{\mu\nu} = T^{\nu\mu}$, skew or antisymmetric if $T^{\mu\nu} = -T^{\nu\mu}$, or asymmetric if $T^{\mu\nu} \neq T^{\nu\mu}$. We can, however, always separate the symmetric and antisymmetric part of any arbitrary tensor $T^{\mu\nu}$, which gives for the symmetric part

$$S^{\mu\nu} = \frac{1}{2}\left(T^{\mu\nu} + T^{\nu\mu}\right) \ , \tag{1.68}$$

and for its anti symmetric part

$$A^{\mu\nu} = \frac{1}{2}\left(T^{\mu\nu} - T^{\nu\mu}\right) \ . \tag{1.69}$$

From the $\vec{E}$ and $\vec{B}$ fields we can construct a tensorial object that will represent the electromagnetic field. Also known as electromagnetic field strength tensor it is defined as

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} \tag{1.70}$$

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix} \ , \tag{1.71}$$

this tensor is totally antisymmetric (and thus traceless), and geometrically equivalent to a 2-form. We may then express the fields $\vec{E}$ and $\vec{B}$ as components of $F^{\mu\nu}$

$$E_i = cF_{0i} \ , \tag{1.72}$$

$$B_i = -\frac{1}{2}\epsilon_{ijk}F^{jk} \ , \tag{1.73}$$

with $i = 1, 2, 3$.

In the potential formulation of electrodynamics, we can reduce the potential scalar φ and potential vector $\vec{A}$ into a 4-vector such as $A^\mu(\phi, \vec{A})$. Since both formulations are equivalent, it is possible to express the electromagnetic field strength as a function of the 4-potential

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ . \tag{1.74}$$

From this tensor, we can construct the first Maxwell invariant in SI units

$$F^{\mu\nu}F_{\mu\nu} = 2\left(\frac{\vec{E}^2}{c^2} - \vec{B}^2\right) \ . \tag{1.75}$$

By contracting $F^{\mu\nu}$ with the Levi-Civita symbol, we can obtain the dual tensor

$$G^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0 \end{pmatrix} \ , \tag{1.76}$$

from which we can obtain the second Maxwell invariant by contracting it with the EM field strength

$$F^{\mu\nu}G_{\mu\nu} = -\frac{4}{c}\vec{E}\cdot\vec{B} \ , \tag{1.77}$$

which is related with the determinant of $F^{\mu\nu}$

$$\det(F) = (F^{\mu\nu}G_{\mu\nu})^2 = \frac{1}{c^2}(\vec{E}\cdot\vec{B})^2 \ . \tag{1.78}$$

With the electromagnetic tensor, we can now move to the Lagrangian formulation of electrodynamics. Proving that the action is unique is beyond the topic of this thesis, and can be shown in most advanced electrodynamics books [10]. In classical field theory, actions are defined as follows

$$S = \int d^4x \ \mathcal{L} \ , \tag{1.79}$$

where $\int d^4x$ represents an integration over the entire space-time and $\mathcal{L}$ is the Lagrangian density, which is related to the usual Lagrangian in the following fashion $L = \int d^3\mathcal{L}$. We then define the Maxwell Lagrangian density as follows[5]

$$\mathcal{L} = -\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu} - A_\mu J^\mu \ , \tag{1.80}$$

where J^μ is a 4-current. A legitimate question that arises is whether it is possible to also construct an action from the dual invariant. And while it is absolutely possible to do so, and that would retrieve the proper equations of motions, that term is topological in nature, and would lead to the propagation of no degrees of freedom. We shall expand on this point on the section about Chern-Simons and topological electrodynamics later on.

Using the principle of stationary action

$$\delta S = 0 \Rightarrow \int d^4x \ \delta\mathcal{L} = 0 \ , \tag{1.81}$$

we will be able to obtain our field equations. The field strength not being fundamental, but rather a result of the 4-potentials we will thus apply the variations with respect to the latter.

$$\frac{\partial \mathcal{L}}{\partial A_\mu} = 0 \ , \tag{1.82}$$

$$\frac{1}{\mu_0} \partial_\mu F^{\mu\nu} - J^\nu = 0 \ , \tag{1.83}$$

which gives the two inhomogeneous Maxwell equations. As for the homogeneous Maxwell equations, they are obtained by contracting the first set of equations with the Levi-Civita symbol

$$\partial_\alpha F_{\mu\nu} + \partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu} = 0 \ , \tag{1.84}$$

which is nothing but the expression of the Ricci Curbastro-Bianchi identity. It is also important to notice that the density (1.80) does display gauge invariance. By transforming the potentials as follows $A'^\mu(x) = A^\mu(x) + \partial^\mu \Lambda(x)$ we can easily show that the action and equations of motions are left unchanged. The invariance by $U(1)$ gauge transformations is a key ingredient of electrodynamics, which will lead to complications when dealing with the quantisation of the system. Because of this, if we wished to obtain wave equations, we have to fix a gauge. Choosing the Lorenz gauge $\partial^\mu A_\mu = 0$

[5]It can also be proven that this action uniquely reproduces the Maxwell equations as shown in [10].

we show that

$$\Box A^\nu = \mu_0 J^\nu \ , \tag{1.85}$$

taking the sources to be absent

$$\Box A^\nu = 0 \ , \tag{1.86}$$

and taking a phase-exponential ansatz $A^\nu(x) = \tilde{A}^\nu(k) e^{i\omega t - kx}$ we obtain the dispersion relations

$$k^\mu k_\mu = 0 \Rightarrow \omega = ck \ . \tag{1.87}$$

From this point on, we shall use the natural unit system in which $\mu_0 = \epsilon_0 = \hbar = c = 1$ unless stated otherwise. Consider the equations of motion with sources before gauge fixings

$$\partial_\mu F^{\mu\nu} = J^\nu \quad \Rightarrow \left(\eta^{\mu\nu}\Box - \partial^\mu\partial^\nu\right) A_\mu = J^\nu \ , \tag{1.88}$$

which in momentum space reads

$$\left(-\eta^{\mu\nu}k^2 + k^\mu k^\nu\right) \tilde{A}_\mu = \tilde{J}^\nu \ . \tag{1.89}$$

Solutions to this equation requires the use of Green functions[6], which will reduce to the following

$$\left(-\eta^{\mu\nu}k^2 + k^\mu k^\nu\right) D_{\nu\alpha}(k) = \eta^\mu{}_\alpha \ \Rightarrow \Pi^{\mu\nu} D_{\nu\alpha}(k) = \eta^\mu{}_\alpha, \tag{1.90}$$

which immediately shows a problem. When taking the fully gauge invariant formulation, the matrix Π is not invertible as its determinant is identically zero, stopping us from finding solutions to the equation. Gauge fixings are thus a necessary ingredient to describe correctly our system, as they should lift the singularity. If one considered instead Eq. (1.85) the solution would read

$$(\Pi^{-1})^{\mu\nu} = -\frac{\eta^{\mu\nu}}{k^2} \ , \tag{1.91}$$

which we will now refer to as the propagator. There is however a way to derive a much more general expression, consisting on adding a general gauge fixing term to the action

[6]For a brief introduction, refer to Appendix C.

such as

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2\xi}\left(\partial^{\mu}A_{\mu}\right)^{2} - A^{\mu}J_{\mu} \ . \tag{1.92}$$

By then deriving the equations of motion, and solving in the presence of sources, we may derive the following propagator

$$(\Pi^{-1})^{\mu\nu} = \frac{-\eta^{\mu\nu} + (1-\xi)\frac{k^{\mu}k^{\nu}}{k^{2}}}{k^{2}} \ , \tag{1.93}$$

which gives us the liberty to pick the parameter ξ. Notably, when choosing what is referred to (rather ambiguously) as the Feynman gauge, we retrieve the same results as Eq.(1.91), meaning that they are classically equivalent.

An integral component of electrodynamics and our work is the stress-energy-momentum tensor (abbreviated as EMT for energy-momentum tensor). The EMT can be obtained through the Noether procedure, but will lead to a non-gauge-invariant quantity. It can also be obtained through different means (such as the Hilbert method) which will yield gauge-invariant objects, and we will consider these results instead. It will be defined in a vacuum as

$$T^{\mu\nu} = \frac{1}{\mu_{0}}\left[F^{\mu\alpha}F^{\nu}{}_{\alpha} - \frac{1}{4}\eta^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta}\right] \ , \tag{1.94}$$

and is both symmetric and traceless. The traceless aspect of the EMT indicates the presence of scale invariance in the model, which highlights another key ingredient of classical electrodynamics: **masslessness**. It is important to insist on the following: if the EMT had a trace, it would indeed imply that the theory holds at a certain scale. However, the presence of a scale does not *necessarily* imply the massive nature of a model, it could be for example due to a screening[7] mechanism in the model. Additionally it is very easy to show that the EMT is conserved by estimating its four divergence

$$\partial_{\mu}T^{\mu\nu} = 0 \ . \tag{1.95}$$

To conclude the classical theory of electrodynamics, we shall derive the Hamiltonian picture of the theory, which will allow us to quantize the model later on.

[7]Screening is the damping of electric fields caused by the presence of mobile charge carriers.

1.1.5 Hamiltonian picture

The Hamiltonian picture of electromagnetism is key to understanding the gauge structure of the theory. While the definition of the Hamiltonian is unchanged $H = \int d^3x \ \pi^\mu \dot{A}_\mu - \mathcal{L}$, the Legendre transform will be singular. Let us start by breaking the manifest covariance of the action (1.80)

$$\mathcal{L} = -\frac{1}{2} F^{0i} F_{0i} - \frac{1}{4} Fij F_{ij} \ . \tag{1.96}$$

We then continue by defining our conjugate momenta through the Legendre transformation:

$$\pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\mu)} = F^{\mu 0} \ , \tag{1.97}$$

which shows explicitly that the Legendre transform is singular since $\pi^0 = 0$. Electromagnetism is in this sense a constrained Hamiltonian system[8] [11] doted of the primary constraint $\pi^0 = 0$. We also explicitly see that the conjugate momentum of the magnetic potential $A^i(x)$ is nothing but the electric field $E^i(x)$. Before writing down the Hamiltonian, we recall the definition of the Poisson bracket between two arbitrary functions of the phase space in terms of A^μ fields and their conjugate momenta π^μ

$$\{F, G\} = \frac{\delta F}{\delta A_\mu} \frac{\delta G}{\delta \pi^\mu} - \frac{\delta F}{\delta \pi^\mu} \frac{\delta G}{\delta A_\mu} \ , \tag{1.98}$$

to then introduce the fundamental Poisson brackets (PB) of the phase space, which will read:

$$\{A_\mu(t, \vec{x}), \pi^\nu(t, \vec{y})\} = \delta_\mu{}^\nu \delta^3(\vec{x} - \vec{y}) \ , \tag{1.99}$$

$$\{\pi_\mu(t, \vec{x}), \pi_\nu(t, \vec{y})\} = 0 \ , \tag{1.100}$$

$$\{A_\mu(t, \vec{x}), A_\nu(t, \vec{y})\} = 0 \ . \tag{1.101}$$

From the above definition of conjugate momenta, the Hamiltonian is written as

$$\begin{aligned} H &= \int d^3\vec{x} \ \left(\pi^i(x)\partial_0 A_i(x) - \mathcal{L}(x)\right) \\ &= \int d^3\vec{x} \ \left(\pi^i(x)\partial_0 A_i(x) - \pi^i(x)\partial_i A_0(x) + \pi^i(x)\partial_i A_0(x) + \frac{1}{2}\pi^i(x)\pi_i(x) + \frac{1}{4}F^{ij}(x)F_{ij}(x)\right) \\ &= \int d^3\vec{x} \ \left(-A_0(x)\partial_i \pi^i(x) - \frac{1}{2}\pi^i(x)\pi^i(x) + \frac{1}{4}F^{ij}(x)F_{ij}(x)\right) \ , \end{aligned} \tag{1.102}$$

[8] Refer to Appendix B for a brief introduction.

where the first term of the last line is obtained through an integration by parts. We then get the total Hamiltonian

$$H_T = H + \int d^3\vec{x}\ \lambda_0(x)\pi^0(x)\ , \tag{1.103}$$

where $\lambda_0(x)$ is a Lagrange multiplier that enforces π^0 as a constraint. Since constraints have to act as constants of movement, we will now apply the Dirac-Bergmann algorithm [12, 13] which can be briefly summarised as follows:

- 1) Evaluate the Poisson Bracket (also known as the consistency of constraints) between the primary constraint and the total Hamiltonian.
- 2) If the result is another constraint, repeat step one with this additional constraint.
- 3) Stop once there no new additional constraints.

The consistency of the primary constraint

$$\begin{aligned}\{\pi^0(t,\vec{x}), H_T(t)\} &= \int d^3\vec{y}\ \ \{\pi^0(t,\vec{x}), -A_0(t,\vec{y})\ \partial_i\pi^i(t,\vec{y})\} \\ &= \int d^3\vec{y}\ \ \{A_0(t,\vec{y}), \pi^0(t,\vec{x})\}\ \partial_i\pi^i(t,\vec{y}) \\ &= \int d^3\vec{y}\ \ \delta^3(\vec{x}-\vec{y})\ \partial_i\pi^i(t,\vec{y}) = \partial_i\pi^i(t,\vec{x}) \approx 0\ ,\end{aligned} \tag{1.104}$$

leads to a secondary constraint; which is nothing but Gauss's law. and the consistency of the secondary constraint

$$\begin{aligned}\{\partial_i\pi^i(t,\vec{x}), H_T(t)\} &= \int d^3\vec{y}\ \ \{\partial_i\pi^i(t,\vec{x}), \frac{1}{4}F^{jk}F_{jk}\} \\ &= \int d^3\vec{y}\ \ \{A_k(t,\vec{y}), \pi^i(t,\vec{x})\}\partial_i\partial_j F^{jk}(t,\vec{y}) \\ &= \int d^3\vec{y}\ \ \partial_i\partial_j F^{jk}\delta_k{}^i\delta(\vec{x}-\vec{y}) \\ &= \partial_i\partial_j F^{ij}(t,\vec{x}) = 0\ ,\end{aligned} \tag{1.105}$$

there is thus no additional constraints that appear. We may now classify these constraints by computing their PB

$$\{\pi^0(x), \partial_i\pi^i(y)\} = 0\ , \tag{1.106}$$

which shows that the model is composed of two first class constraints. Applying the master formula for counting physical degrees of freedom we get

$$\begin{aligned} 2\times DOF = &\text{ Phase Space variables} - 2\times\text{First Class Constraints} \\ &- \text{ Second Class Constraints} \\ 2\times DOF = 8 - 2\times 2 - 0 \quad &\Rightarrow DOF = 2 \ , \end{aligned} \tag{1.107}$$

showing that indeed electromagnetic waves are endowed of two polarisations. However, our phase space is still too large, and to reduce it we must study the dynamics of the scalar potential A^0. Since the Hamiltonian does not possess the conjugate momentum π^0, the PB between A^0 and H is trivially zero, and its time evolution is given by

$$\{A^0(t,\vec{x}), H_T(t)\} = \int d^3\vec{x}\ \{A^0(x), \pi^0(y)\}\lambda_0(y) = \lambda_0(x) \ , \tag{1.108}$$

showing that the evolution of $A^0(x)$ is arbitrary. We are thus free to set $A^0(x) = 0$, $\pi^0(x) = 0$ and consider the phase space reduced to the fields $A^i(x)$ and their conjugate momenta $\pi^i(x)$. However, we are left with an arbitrary parameter $\Lambda(x)$ in the total Hamiltonian. Taking the variation of the field A^i

$$\delta A^i(x) = \{A^i(x), \int d^3\vec{y}\ \ \Lambda(y)\partial_i\pi^i(y)\} = -\partial_i\Lambda(x) \ , \tag{1.109}$$

we notice that this is equivalent to a local gauge transformation. We shall thus consider that first-class constraints generate gauge transformations[9] [13].
From the classical theory of electrodynamics, we can conclude that the fundamental variable of electrodynamics is the 4-vector $A^\mu(x)$ that possesses 2 degrees of freedom and therefore is massless. Said 4-vector has to be included in a *linear* theory that exhibits both *Lorentz invariance* and *gauge invariance* under $U(1)$ transformations. Gauge invariance and masslessness are usually assumed to be linked, since adding an explicit mass term breaks gauge invariance and changes the constraint structure. We will, however, show that this is not necessarily the case in the next chapters. Additionally, gauge invariance leads to a lot of peculiarities when it comes to the quantum field of electrodynamics. In the following section, we will pursue our endeavour by quantising canonically the electromagnetic field, before deriving the Feynman rules and drawing a final list of the ingredients of electrodynamics.

[9]Whether all first class constraints (also known as the Dirac conjecture) or only primary first class constraints generate gauge transformations is up to debate, and we refer to [11] for more information.

1.2 Quantum Electrodynamics (QED)

1.2.1 Quantisation of the free electromagnetic field

To consider photons it is imperative to quantise the electromagnetic field. However, when taking into account the fully covariant theory, some concerns arise. Indeed, since the electromagnetic potential is a 4-vector, and the theory exhibits both Lorentz invariance and gauge invariance we end up dealing with more states than physically relevant at the quantum level. The main issue stems from the time-like polarisation due to A^0, which will cause the metric of our Hilbert space to be indefinite. We thus lose probabilistic interpretation of quantum mechanics since negative probabilities would be possible within our model [14]. We could of course set $A^0 = 0$ since we showed in the Hamiltonian formalism that the conjugate $\pi^0 = 0$ but this would be at the cost of the manifest Lorentz invariance. In the interest of dealing with our ghost states in a manifestly covariant way, we will introduce the Gupta-Bleuler method. The reason we pick this quantisation scheme is that it is not only historically relevant, often forgotten about, but also that it is well defined. It can notably be very well generalised to non-abelian theories through the BRST cohomology [11, 15] (which we will not address in this thesis).

We start by considering the gauge-fixed action (1.92) and perform its Legendre transform

$$\pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\mu)} = F^{\mu 0} - \frac{1}{\xi}\eta^{\mu 0}\partial_\nu A^\nu \ , \tag{1.110}$$

which shows that the system is not singular. Since the observables should not depend on ξ, we will set it to 1 also known as the Feynman gauge. Our fundamental Poisson Brackets at equal time will then read

$$\{A_\mu(x), \pi^\nu(y)\}_{x^0=y^0} = \delta_\mu{}^\nu \delta(\vec{x} - \vec{y}) \tag{1.111}$$

$$\{\pi_\mu(x), \pi_\nu(y)\}_{x^0=y^0} = 0 \quad ; \{A_\mu(x), A_\nu(y)\}_{x^0=y^0} = 0 \ . \tag{1.112}$$

In terms of derivatives of the 4-potential, we get for spatial components

$$\{A_i(x), \dot{A}_j(y)\}_{x^0=y^0} = \delta_{ij}\ \delta^3(\vec{x} - \vec{y}) \ , \tag{1.113}$$

which seems equivalent to three independent scalar fields. For time components we obtain

$$\{A_0(x), \dot{A}_0(y)\}_{x^0=y^0} = -\delta^3(\vec{x} - \vec{y}) \ , \tag{1.114}$$

which is similar to a single scalar field. Together, they can be summed as

$$\{A_\mu(x), \dot{A}_\nu(y)\}_{x^0=y^0} = -\eta_{\mu\nu}\ \delta^3(\vec{x}-\vec{y})\ , \tag{1.115}$$

which explicitly shows what we stated above, the time-like component has an opposite sign to the space-like component, meaning that our system has a ghost. The same procedure between the conjugate momenta leads to

$$\{\dot{A}_\mu(x), \dot{A}_\nu(x)\}_{x^0=y^0} = 0\ . \tag{1.116}$$

The Hamiltonian is written as

$$\begin{aligned}\mathcal{H} &= \int d^3\vec{x}\,(\pi^\mu \partial_0 A_\mu - \mathcal{L}) \\ &= \int d^3\vec{x}\left(\pi^0\dot{A}_0 + \pi^i\dot{A}_i + \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}(\partial\cdot A)^2\right)\ ,\end{aligned} \tag{1.117}$$

and after some manipulations reads

$$\mathcal{H} = \int d^3\vec{x}\left(-\frac{1}{2}\left(\dot{A}_0^2 + \vec{\nabla}A_0\cdot\vec{\nabla}A_0\right) + \frac{1}{2}\left(\dot{A}^i\cdot\dot{A}^i + \vec{\nabla}A^i\cdot\vec{\nabla}A^i\right)\right)\ , \tag{1.118}$$

which as stated above in the brackets is equivalent to four massless scalar fields [10] with a negative sign on the one associated with A^0.

Since we are pursuing a covariant description of photons, we have to consider more polarisations than physically relevant. Usually, light's degrees of freedom are described by the two polarisation ϵ^λ where $\lambda = \{1,2\}$, which are transverse to the wave vector $\vec{k}$ such as $\vec{\epsilon}^\lambda\cdot\vec{k} = 0$. In our case, we shall extend ϵ^λ to include both time-like and longitudinal polarisations so that we can write the solution to the wave equation as

$$A_\mu(x) = \int d\tilde{k}_B\ \sum_{\lambda=0}^{3}\left(a^{(\lambda)}(k)\epsilon_\mu^{(\lambda)}(k)e^{-ikx} + a^{\dagger(\lambda)}(k)\epsilon_\mu^{*(\lambda)}(k)e^{ikx}\right) \tag{1.119}$$

where ϵ^λ with $\lambda = 0,1,2,3$ are the polarisation 4-vectors that we consider real and orthonormal

$$\eta^{\mu\nu}\epsilon_\mu^{(\lambda)}\epsilon_\nu^{*(\lambda')} = \epsilon^{(\lambda)}\cdot\epsilon^{*(\lambda')} = g^{\lambda\lambda'}\ , \tag{1.120}$$

where $g^{00} = 1\ , g^{i0} = 0\ , g^{ij} = -\delta^{ij}$. In a referential where $k^\mu = (k^0, 0, 0, k^3)$, the four

polarisation vectors would be

$$\epsilon_\mu^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \ , \ \epsilon_\mu^{(1)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \ , \ \epsilon_\mu^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } \epsilon_\mu^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \ ,$$

where ϵ^1 and ϵ^2 are the transverse polarisations, ϵ^0 is the time-like polarisation (or ghost) and ϵ^3 is the longitudinal polarisation. From these polarisations we can show that following relation always holds

$$\frac{\epsilon_\mu^{(\lambda)} \epsilon_\nu^{*(\lambda)}}{g^{\lambda\lambda}} = \epsilon_\mu^{(\lambda)} \epsilon^{*(\lambda)} g_{\lambda\lambda} = \eta^{\mu\nu} \ . \tag{1.121}$$

By then using the correspondence principle, we can turn the Poisson brackets to the equal time commutation relations (ECTR) [10]

$$\Big[A_\mu(x) \ , \ \dot{A}_\nu(x) \Big]_{x^0=y^0} = -i\eta^{\mu\nu} \delta^3(\vec{x} - \vec{y}) \tag{1.122}$$

$$[A_\mu(x) \ , \ A_\nu(x)]_{x^0=y^0} = 0 \ \ , \Big[A_\mu(x) \ , \ \dot{A}_\nu(x) \Big]_{x^0=y^0} = 0 \qquad . \tag{1.123}$$

This results in $a^{(\lambda)}(k)$ and $a^{\dagger(\lambda)}(k)$ becoming operators which obey

$$\Big[a^{(\lambda)}(k), a^{\dagger(\lambda')}(k') \Big] = -2k^0 g^{\lambda\lambda'} (2\pi)^3 \delta^3(\vec{k} - \vec{k}') \tag{1.124}$$

$$\Big[a^{\dagger(\lambda)}(k), a^{\dagger(\lambda')}(k') \Big] = 0 \ \ , \ \ \Big[a^{(\lambda)}(k), a^{(\lambda')}(k') \Big] = 0 \qquad . \tag{1.125}$$

Through these operators we derive the quantised form of the Hamiltonian

$$\mathcal{H} = \int d\tilde{k}_B \ k^0 \left(-a^{\dagger(0)}(\vec{k}) a^{(0)}(\vec{k}) + \sum_{\lambda=1}^{3} a^{\dagger(\lambda)}(\vec{k}) a^{(\lambda)}(\vec{k}) \right) \ . \tag{1.126}$$

The Fock space of photons results from the creation operators $a^{\dagger(\lambda)}$ acting on the vacuum state $|0\rangle$. Said vacuum state must be anihilated by the ladder operators $a^{(\lambda)}$ such as $a^{(\lambda)}(k)\,|0\rangle = 0$.

From this we deduce that the state $a^{\dagger(\lambda)}(k)\,|0\rangle$ characterises a photon of momentum

k and polarisation λ. The norm of a polarisation state

$$|\lambda\rangle = \int d\tilde{k}_B \; f(k) a^{\dagger(\lambda)}(k) \, |0\rangle \;\; , \tag{1.127}$$

gives

$$\begin{aligned} \langle\lambda|\lambda\rangle &= \int d\tilde{k}'_B \int d\tilde{k}_B \; f^*(k') f(k) \, \langle 0| \, a^{(\lambda')}(k') a^{\dagger(\lambda)}(k) \, |0\rangle \\ &= - \, g^{\lambda\lambda'} \int d\tilde{k}_B \; |f(k)|^2 \end{aligned} \tag{1.128}$$

which shows that even at the quantum level, the indefinite nature of the metric persists. In the free case, this is not necessarily an issue, but when dealing with interactions, we may excite unwanted states and end up with negative energies [14].
To deal with this awkward situation, let us first introduce what an arbitrary state for photons would look like. Consider $|\psi\rangle$ such as

$$|\psi\rangle = |\psi\rangle_T \, |\phi\rangle \;\; , \tag{1.129}$$

where $|\psi\rangle_T$ is the physical state due to transverse polarisations

$$|\psi\rangle_T = \prod_{i=1}^{N} a^{\dagger(1)}(k_i) \prod_{j=1}^{M} a^{\dagger(2)}(q_j) \, |0\rangle \tag{1.130}$$

and $|\phi\rangle$ represents the unphysical scalar and longitudinal modes

$$|\phi\rangle = \prod_{i=1}^{M} a^{\dagger(0)}(k_i) \prod_{j=1}^{K} a^{\dagger(3)}(q_j) \, |0\rangle \;\; . \tag{1.131}$$

However, it is important to remind ourselves that the action we started with is not the Maxwell action! It is missing a key ingredient that arises from the dynamics, in the Lorenz gauge. As such we wish that physical states are those that obey the constraint

$$\partial_\mu A^\mu = 0 \; . \tag{1.132}$$

In this form, the condition cannot be enforced as an operator identity, and to this end we will consider that physical states are those that satisfy

$$\langle\psi| \, \partial_\mu \hat{A}^\mu(x) \, |\psi\rangle \;\; . \tag{1.133}$$

We decompose the solution (1.119) into positive and negative frequency components

$$\hat{A}_\mu(x) = \hat{A}^{(+)}_\mu(x) + \hat{A}^{(-)}_\mu(x) \tag{1.134}$$

where

$$\hat{A}^{(+)}_\mu(x) = \int d\tilde{k}_B \sum_{\lambda=0}^{3} a^{(\lambda)}(k)\epsilon^{(\lambda)}_\mu(k)e^{-ikx} \tag{1.135}$$

$$\hat{A}^{(-)}_\mu(x) = \int d\tilde{k}_B \sum_{\lambda=0}^{3} a^{\dagger(\lambda)}(k)\epsilon^{*(\lambda)}_\mu(k)e^{ikx} \ , \tag{1.136}$$

enforcing

$$\langle\psi|\, \partial^\mu \hat{A}^{(-)}_\mu(x) = 0 \tag{1.137}$$

$$\partial^\mu \hat{A}^{(+)}_\mu(x)\, |\psi\rangle = 0 \ . \tag{1.138}$$

Using only the positive frequency part, we can obtain

$$\partial^\mu \hat{A}^{(+)}_\mu(x)\, |\psi\rangle = -i \int d\tilde{k}_B \sum_{\lambda=0}^{3} a^{(\lambda)}(k)\ k^\mu \epsilon^{(\lambda)}_\mu(k)e^{-ikx}\, |\psi\rangle = 0 \ . \tag{1.139}$$

Considering that transverse polarisation states naturally obey to the condition $k^\mu\epsilon^{(1)}_\mu = k^\mu\epsilon^{(2)}_\mu = 0$, the expression above reduces to

$$-i \int d\tilde{k}_B \Big(\ a^{(0)}(k)\ k^\mu\epsilon^{(0)}_\mu(k) + a^{(3)}(k)\ k^\mu\epsilon^{(3)}_\mu(k)\Big)\, e^{-ikx}\, |\psi\rangle = 0 \ , \tag{1.140}$$

which has to be valid $\forall\ x$, showing that the Lorenz condition acts on unphysical states as

$$\Big(\ a^{(0)}(k)\ k^\mu\epsilon^{(0)}_\mu(k) + a^{(3)}(k)\ k^\mu\epsilon^{(3)}_\mu(k)\Big)\, |\phi\rangle = 0 \ . \tag{1.141}$$

Since both $k^\mu\epsilon^{(0)}_\mu$ and $k^\mu\epsilon^{(3)}_\mu$ are scalars, we can compute them in a referential where $k^\mu = (k^0, 0, 0, k^3)$ leading to $k^\mu\epsilon^{(0)}_\mu = k^0$ and $k^\mu\epsilon^{(3)}_\mu = -k^3$. Imposing that $k^0 = k^3$, we can summarise the Lorenz condition on unphysical states as the following expression between ladder operators

$$\Big(a^{(0)}(k) - a^{(3)}(k)\Big)\, |\phi\rangle = 0 \ . \tag{1.142}$$

It is then straight forward to show that the expectation value of the Hamiltonian in

an arbitrary state $|\psi\rangle$ will be

$$\begin{aligned}\langle\psi|\,\mathcal{H}\,|\psi\rangle &= \int d\tilde{k}_B\; k^0 \left({}_T\langle\psi| \sum_{\lambda=1}^{2} a^{\dagger(\lambda)}(k)a^{(\lambda)}(k)\,|\psi\rangle_T \,\langle\phi|\phi\rangle + \right. \\ &\qquad \left. {}_T\langle\psi|\psi\rangle_T \,\langle\phi|\, a^{\dagger(3)}(k)a^{(3)}(k) - a^{\dagger(0)}(k)a^{(0)}(k)\,|\phi\rangle \right) \\ &= \int d\tilde{k}_B\; k^0 \left({}_T\langle\psi| \sum_{\lambda=1}^{2} a^{\dagger(\lambda)}(k)a^{(\lambda)}(k)\,|\psi\rangle_T \,\langle\phi|\phi\rangle \right) \,, \end{aligned} \tag{1.143}$$

where we used (1.142) to get rid of the second term. The total energy will then be

$$E = \frac{\langle\psi|\,\mathcal{H}\,|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\int d\tilde{k}_B\; k^0 \left({}_T\langle\psi| \sum_{\lambda=1}^{2} a^{\dagger(\lambda)}(k)a^{(\lambda)}(k)\,|\psi\rangle_T \right)}{{}_T\langle\psi|\psi\rangle_T} \,, \tag{1.144}$$

which explicitly shows that the total energy is carried entirely by the transverse modes.

1.2.2 QED and Feynman rules for electrodynamics

So far we have described only the dynamics of the free electromagnetic field, ignoring the nature of currents and charges. This however is unrealistic, as in nature electromagnetic field constantly interact with charged particles, and as such they must be added. Describing the dynamics of the free Dirac field is not relevant for this thesis, we will thus introduce directly the full action of quantum electrodynamics [16], including the interaction between fermions and photons

$$\mathcal{L}_{QED} = \bar{\psi}\left(\frac{i}{2}\overleftrightarrow{\partial\!\!\!/} - m\right)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2 - e\bar{\psi}\gamma^\mu\psi A_\mu \tag{1.145}$$

where ψ and $\bar{\psi}$ are Dirac spinors that represent respectively fermions and anti-fermions, γ^μ is a Dirac matrix that obeys the usual Clifford algebra [14], $\partial\!\!\!/ = \gamma_\mu\partial^\mu$ according to the Feynman convention and e is the electric charge. The interaction term is obtained through the minimal coupling procedure, by replacing the partial derivatives in the Dirac action by the covariant derivative $D^\mu = \partial^\mu + ieA^\mu$.

Interactions in particle physics are of extremely short durations (order of 10^{-28} seconds) and makes measuring what is *really* happening impossible. Mathematically, the interaction picture of field theories is at best ill defined, and at worst should not exist (according to Haag's theorem [17], but this did not stop it from being extremely successful in its prediction, yield in the case of QED the most precise measurements in physics in the electron $g-2$ moment. In quantum field theory, we consider that initial states (or ingoing states) labelled $|in\rangle$ and final states (or outgoing states) labelled $|out\rangle$ are both representing free field configurations. They are related by what is known as

the scattering matrix (or S-matrix) which rules the unitary evolution from one state to another. As such, the probability amplitude from *in* states to *out* states is defined as $S_{io} = \langle out| \mathcal{S} |in\rangle$ where $\mathcal{S}$ is the S-matrix.

Introducing and demonstrating the exact development leading to the derivation of cross sections is far beyond the reach of this thesis, and are extensively discussed in most graduate textbooks [14, 16, 18]. The process by which we compute interactions relies fundamentally on Wick's theorem [19], which will allow us to pertubartively expand the S matrix into a sum of terms that we will translate into Feynman diagrams. In that sense, Feynman diagrams are nothing but a visual technique to count properly the terms added by each order of the Wick expansion.

An additional key ingredient to the cross section computations is the LSZ reduction formula, which allows us to obtain amplitudes from correlation functions [18]. In extremely vulgarised terms, we could say that $\langle in|$ and $|out\rangle$ states are connected by the Green functions (propagators) of the fields.

The Feynman rules can then be obtained straightforwardly from the action, by distinguishing between 1-point functions (or external lines) and 2-point functions (internal lines) and respecting the shape of the interaction term (vertices). In QED, fermions and anti-fermions will be represented by straight lines with an arrow pointing in the chosen time direction for Fermions, and the opposite for anti-fermions (As such positrons can be interpreted as electrons going back in time in Feynman diagrams). Photons will be drawn as arrowless squiggly lines (since the photon is its own anti-particle), and vertices will always be composed of an electron, a positron and a photon. Regrouped in a table they read

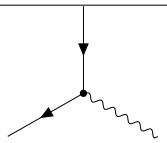
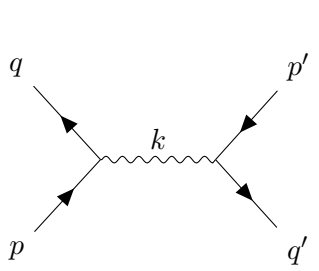
Element	Symbol	Mathematical expression
Fermion internal line (propagator)		$\int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p}+m)}{p^2-m^2+i\epsilon}$
Photon propagator		$\int \frac{d^4k}{(2\pi)^4} \frac{-i\eta^{\mu\nu}}{k^2+i\epsilon}$
Vertex		$-ie\gamma^\mu(2\pi)^4\delta(\sum incoming\ momenta)$
Ingoing fermion		$u(p)$
Ingoing antifermion		$\bar{u}(p)$
Outgoing fermion		$v(p)$
Outgoing antifermion		$\bar{v}(p)$
Ingoing and outgoing photons		$\epsilon^\mu(k)$ (ingoing) or $\epsilon^{\mu*}(k)$ (outgoing)

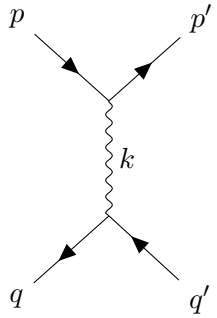
Table 1.1: Feynman's rules for Quantum Electrodynamics in momentum space and in the Feynman gauge $\xi = 1$

Where $u(p)$ and $v(p)$ are Dirac spinors, $\epsilon^\mu(k)$ is the photon polarisation vector, m is the fermion's mass and $\not{p} = \gamma^\mu p_\mu$.

We can see that internal lines carry an integral over the momentum of the virtual particle, and vertices enforce the conservation of momentum. From this we can get the tree levels diagrams in a very straight forward way and derive the amplitudes. Taking the Bhabha scattering as an example, the two leading order diagrams will be:



(a) Annihilation process



(b) Scattering process

Figure 1.2: Leading order contributions to the Bhabha scattering.

Which when translated into mathematical form gives the following amplitude for the process

$$\mathcal{M} = e^2 \left(\bar{v}_q \gamma^\nu u_p\right) \frac{1}{(p+q)^2} \left(\bar{u}_{p'} \gamma_\nu v_{q'}\right) - e^2 \left(\bar{v}_q \gamma^\mu v_{q'}\right) \frac{1}{(q-q')^2} \left(\bar{u}'_p \gamma_\mu u_p\right) \ . \tag{1.146}$$

Tree level diagrams (also known as order 0) are usually well behaved, and do not lead to any type of difficulties. However, This is not the case as soon as we go to the next order in perturbations, and start encountering loops. Let us illustrate our point with the following simple process known as the fermion self-energy

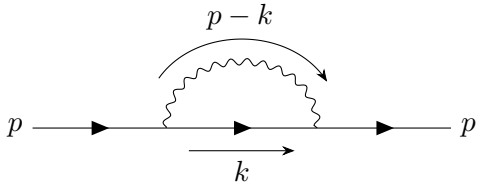


Figure 1.3: Fermion's self-energy

This diagram can be read as a fermion emitting a virtual photon before reabsorbing it. This process involves two vertices like those described above, but it will however lead to serious issues. Indeed, unlike the simpler Bhabha processes that we discussed, the self-energy diagram puts in play two internal lines, one fermionic and one photonic. When computing the amplitude, we quickly notice that we are left with an integral over the momenta which will lead to divergences. Without going into details, the electron self-energy function will read [14, 18]

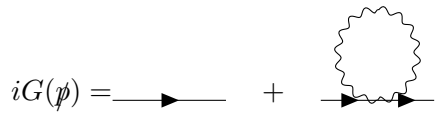
$$-i\Sigma(p) = (-ie)^2 \int \frac{d^4k}{(2\pi)^4} \gamma^\mu \frac{i\left(\not{k} + m\right)}{k^2 - m^2 + i\varepsilon} \gamma^\nu \frac{-i\eta_{\mu\nu}}{(p-k)^2 + i\varepsilon} \ , \tag{1.147}$$

which is clearly divergent due to the lack of upper bounds on the momentum k. It could

then be tempting to discard said divergence due to its obvious unphysical nature, but we must remember that our approach is perturbative, and as such it has to be accounted for in our calculations. To this end we will proceed with regularisation, which consists on reducing the degree of divergence (from power-law divergence to logarithmic) and isolating the source of infinities. This process is extremely long and tedious, even for one loop, and there are various schemes that could be used (Pauli-Villars, Dimensional Regularisation, Renormalisation group methods etc), and will not be discussed in this thesis, but is detailed in every advanced QFT book [14, 16, 18].

Once the regularisation is achieved, the next step is to renormalise the theory, by "hiding" the infinities in a measurable parameter. In reality, this process is better known as **dressing** the parameter, and will lead us to understand the order 0 of theories as represented by bare grandeurs, that are not physical. The example of the fermion self-energy notably leads to hiding the infinities in the fermion's mass, yielding the physical, measurable mass.

Once the renormalisation process finished, we must add the proper counter terms to the action to ensure that the action is well behaved at the order in question. By doing so, we effectively modify the propagator of the particles through their interactions, and for the electron this would lead at one loop to something of the form

$iG(\not{p}) =$  $+ \cdots$

$$= \frac{i}{\not{p} - m} + \frac{i}{\not{p} - m} i\Sigma(\not{p}) \frac{i}{\not{p} - m} + \dots \,, \tag{1.148}$$

which can be shown to be equivalent to a geometric series [20], leading to the pole of the propagator to be shifted by the self-energy function. In this sense, we will define the difference between *effective* and *physical* mass as the correction by loops. A field could be massless at order 0, but acquire a physical mass once corrected by self-energy diagrams. In the same vein, a model could be massive at order 0, but lose its mass once renormalised, showing that the mass was only effective, *i.e.* valid at lower energy scales.

Following a similar logic, we could study what a photon's "self-energy" would yield as results. The diagram in question known as the vacuum polarisation

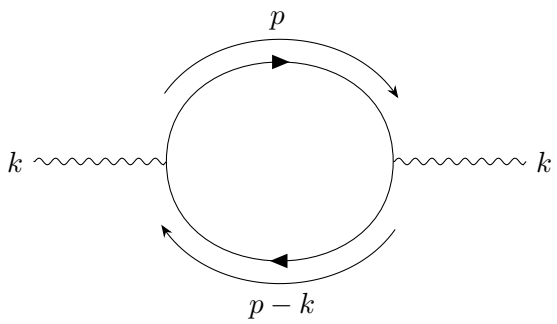


Figure 1.4: Vacuum polarisation diagram

Which in mathematical form reads

$$i\Pi^{\mu\nu}(k) = -(-ie)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{(k-p)^2 - m^2 + i\varepsilon} \frac{i}{p^2 - m^2 + i\varepsilon} \cdot Tr\left(\gamma^\mu \left(\not{p} - \not{k} + m\right) \gamma^\nu (\not{p} + m)\right) \ . \tag{1.149}$$

This has to undergo the same regularisation process as the self-energy diagram, with one notable difference: the gauge invariant nature of photons. Gauge invariance at the quantum level is expressed in terms of the Ward identity [21]

$$k_\mu \mathcal{M}^\mu = 0 \ , \tag{1.150}$$

which will be upheld on physical states. Naturally, any regularisation scheme (or renormalisation) of photons must conserve their gauge invariant nature, and due to this fact we can show that there is no mass renormalisation for photons in QED [14], as the pole remains at $k^2 = 0$. In the case of the vacuum polarisation diagram, the renormalised parameter ends up being the electric charge. The fact that gauge invariance and masslessness appear to be deeply connected is only true in flat spaces, as discussed in [22], which could lead to the pole for photons propagators to be shifted in constant curvatures.

The symmetries of QED do not end there, as they give us another important result in Furry's theorem [23]. According to it, any loop with an odd number of fermionic lines such as
will automatically possess a null contribution to the amplitude. A corollary of this theorem, is that a single photon cannot be absorbed or emitted from the vacuum.

So far, we have only discussed UV divergences for electrodynamics. But the massless nature of photons in QED leads to divergence of another nature, infrared (IR) divergences. Since photons do not possess a minimal energy, there can always be an infinite number of photons of lower energy contributing to any process. This will lead

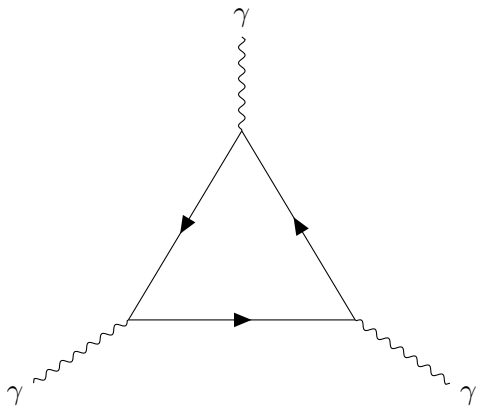


Figure 1.5: Triangle diagram

to a divergence in the sense of the number of particles contributing, rather than the energy scale of said particles. Dealing with infrared divergences, is a necessity of QED, and while Weinberg showed how to deal formally through his soft photon (graviton) theorem [24], a very simple way to circumvent it, is to endow the photon with a small μ such as $\mu << m_e$ [14].

Quantum electrodynamics goes of course, much beyond what we discussed here, but from this small foray into the (quantum) fields, we can complete the list of ingredients for electrodynamics.

1.2.3 The recipe

To conclude this chapter, we will list the ingredients of electrodynamics, and for the rest of this work, we will check on this list regularly, to see if our extended models are healthy, and if not, if they could be treated. As stated in the previous section, classically the list is composed of

- Linearity.
- Lorentz invariance.
- Gauge invariance under (1) transformations.
- Masslessness.
- And something we did not discuss, the absence of magnetic monopoles.

Interestingly, the quantum theory already kills the first ingredient! We will thus consider the linearity of electrodynamics to be only effective. Regardless, the ingredients from the quantum side are

- Contributions to the physical states from only two polarisations (No time-like or longitudinal contribution).
- Unitarity.
- Renormalisasbility.
- Ward identities.
- Infrared divergences (soft photons).
- No pole shifting after renormalisation (still no mass).

In this thesis, we will only discuss the classical aspects of extended theories of electrodynamics. We will however discuss certain features that relate to the quantisation or renormalisation process when it is relevant. We will split the extended theories into three large categories: **massive**, **Lorentz violating** and **classically non-linear** theories of electrodynamics.

Chapter 2

Extended Theories of Electrodynamics

Résumé

Dans ce chapitre nous introduirons différentes extensions à la théorie de Maxwell, en commençant par les modèles massifs, puis en explorant les théories non linéaires et enfin les violations de la symétrie de Lorentz. L'objectif est de comprendre les implications physiques, mathématiques et structurelles de ces modifications, ainsi que de dresser un portrait unifié de chacune de ces trois familles de modèles.

La première extension étudiée repose sur l'ajout d'un terme de masse pour le photon. Bien que cela semble en contradiction avec les principes fondamentaux établis précédemment, notamment l'invariance de jauge, cette modification permet de résoudre certains problèmes comme les divergences infrarouges et reste la modification la plus simple possible de la théorie. Le modèle le plus simple est celui de de Broglie-Proca, qui consiste à ajouter un terme de masse directement dans l'action, en le couplant au potentiel vecteur A^μ. Cette modification entraîne immédiatement la perte de l'invariance de jauge, ce qui signifie que les potentiels électromagnétiques acquièrent un caractère physique mesurable, contrairement au cas standard où ils sont redondants et nécessitent une Fixation de jauge.

Les équations du mouvement issues de ce modèle montrent que le champ obéit à une équation de type Klein-Gordon, impliquant l'existence d'un référentiel de repos pour le photon. Les relations de dispersion sont modifiées, ce qui entraîne des vitesses de phase et de groupe dépendantes de la fréquence. Cela implique notamment des délais entre différentes composantes spectrales d'un signal. Une autre conséquence majeure est la modification du potentiel électrostatique, qui devient de type Yukawa et décroît exponentiellement avec la distance. Cette décroissance impose que la masse du photon soit extrêmement faible afin de rester compatible avec les observations expérimentales, mais tue fondamentalement la portée supposée infinie de l'électromagnétisme.

L'analyse hamiltonienne du modèle révèle que la théorie possède trois degrés de liberté, ce qui correspond à un champ vectoriel massif. La présence de contraintes de seconde classe confirme l'absence d'invariance de jauge. Le formalisme de Dirac permet de construire un Hamiltonien réduit défini positif et d'éviter les instabilités associées aux polarisations temporelles.

Cependant, la perte d'invariance de jauge pose un problème conceptuel important, notamment de par l'importance que cette dernière joue au niveau quantique. Pour y remédier, on introduit le mécanisme de Stueckelberg, qui consiste à ajouter un champ scalaire supplémentaire. Ce champ restaure l'invariance de jauge en compensant la transformation du potentiel. Le modèle obtenu est équivalent au modèle de Proca dans une jauge particulière, ce qui montre que ce dernier peut être interprété comme une fixation de jauge du modèle de Stueckelberg. Cette approche préfigure le mécanisme de

Higgs, dans lequel un champ scalaire est absorbé par un champ vectoriel pour lui conférer une masse. L'analyse hamiltonienne confirme que le système possède alors quatre degrés de liberté, correspondant à un champ vectoriel massif et un champ scalaire.

Une autre approche pour introduire une masse consiste à utiliser des dérivées d'ordre supérieur, comme dans le modèle de Bopp-Podolsky. Ce modèle modifie le potentiel électrostatique de manière à éviter les divergences au voisinage de l'origine. Il conduit à une superposition d'un potentiel de Coulomb et d'un potentiel de Yukawa. L'équation du mouvement révèle la présence de deux modes de propagation, l'un massif et l'autre sans masse, ce qui correspond à deux types de photons. Cependant, cette théorie souffre d'instabilités graves, connues sous le nom d'instabilités d'Ostrogradski, liées aux dérivées d'ordre supérieur introduites précédemment. Ces instabilités rendent la quantification problématique en introduisant des énergies non bornées inférieurement, ou pire, des probabilités négatives (perte d'unitarité).

Nous continuons ensuite avec la théorie de Maxwell-Chern-Simons, qui introduit une masse d'origine topologique dans un espace de dimension réduite. Dans ce cadre, la masse n'est pas associée à une énergie au sens habituel, mais à une propriété globale de l'espace. Le terme de Chern-Simons est invariant de jauge et conduit à une équation du mouvement de premier ordre. L'analyse hamiltonienne montre que ce terme pur ne possède aucun degré de liberté, confirmant sa nature topologique. En revanche, lorsqu'il est combiné avec le terme de Maxwell, il donne naissance à une théorie décrivant une excitation massive unique. Cette masse apparaît dans les relations de dispersion, bien qu'elle ne figure pas explicitement dans le tenseur énergie-impulsion. Nous soulignerons le fait que les masses d'origines topologiques doivent être accommodées sur les polarisations transverses du photon.

Nous concluons en abordant brièvement les contraintes expérimentales sur la masse du photon. Différentes méthodes, allant des observations astrophysiques aux expériences de laboratoire, permettent d'établir des limites extrêmement faibles, confirmant que si le photon possède une masse, celle-ci est presque nulle avant de résumer dans un tableau toutes les masses en dessous de $10^{-50}kg$.

La deuxième partie du chapitre s'intéresse aux théories non linéaires de l'électrodynamique.Plus particulièrement, nous introduisons les modèles appartenant à la classe de Plebański, caractérisée par des actions dépendant uniquement des invariants de Maxwell et conservant l'invariance de Lorentz.

Le modèle de Born-Infeld constitue un exemple majeur de théorie non linéaire. Il a été introduit pour résoudre le problème des divergences du champ électrique au voisinage des charges ponctuelles. L'action est construite de manière analogue à la relativité restreinte, en introduisant une limite maximale pour le champ électrique. Cette modification empêche les divergences et conduit à une dynamique dans laquelle le vide se

comporte comme un milieu non linéaire. Les équations du mouvement montrent que les champs électriques et magnétiques sont couplés de manière non triviale, modifiant les propriétés optiques du vide. Malgré ces non-linéarités, le modèle conserve deux degrés de liberté et est un des rares modèles non-linéaires ne présentant pas de biréfringence du vide.

La théorie de Heisenberg-Euler adopte une approche différente, issue de la théorie quantique des champs. Il s'agit d'une action effective obtenue en intégrant les degrés de liberté fermioniques dans un champ électromagnétique constant. Cette théorie capture des effets quantiques comme la polarisation du vide et la création de paires. Elle est valide à basse énergie et prédit notamment des phénomènes comme la diffusion photon-photon et la biréfringence du vide. Dans la limite des champs faibles, l'action se simplifie et met en évidence des corrections non linéaires aux équations de Maxwell.

Nous closons le chapitre en explorant les violations de la symétrie de Lorentz, également connues comme l'extension du modèle standard. Cette symétrie bien que fondamentale en théorie des champs, pourrait être brisée à haute énergie. Le cadre du Standard Model Extension (SME) permet donc d'étudier ces violations à basse énergie.

Le modèle de Carroll-Field-Jackiw introduit un terme qui brise explicitement l'invariance de Lorentz en couplant le champ électromagnétique à un vecteur constant. Ce terme est invariant de jauge mais viole la symétrie CPT. Les équations du mouvement modifiées conduisent à des relations de dispersion anisotropes et à des effets comme la biréfringence du vide. L'analyse hamiltonienne montre que la théorie conserve deux degrés de liberté, malgré l'apparition d'un comportement semblable à celui d'un champ massif.

Le chapitre se conclut par la présentation du secteur photonique complet de l'extension du modèle standard, incluant à la fois les termes CPT pairs et impairs. Ces termes introduisent des tenseurs de violation de Lorentz qui modifient la dynamique du champ électromagnétique. Une classification finale résume les différentes extensions selon leurs propriétés fondamentales, telles que l'invariance de jauge, la masse, la linéarité et l'invariance de Lorentz.

2.1 Massive Extensions

The concept of photon mass may seem a bit ill advised with what was laid out in the previous chapter. Indeed, we showed that gauge invariance at the quantum level leads to strong constraints on our structures and by extension, observations. We, nonetheless, also discussed the fact that masslessness leads to infrared divergences in the theory, and a simple way to circumvent this issue is by endowing the photon with a very small mass. In this section, we will discuss what this would mean for the theories. More specifically, we will track the presence and absence of gauge invariance across these models and briefly discuss the phenomenological implications of said theories.

2.1.1 de Broglie-Proca Action

The first and simplest massive extension of electrodynamics is the de Broglie-Proca (dBP) action [25–28]. It is simply obtained by coupling a mass parameter to a quadratic term in the action, to reproduce a Klein-Gordon equation for the propagation $(\Box + m^2)\phi(x) = 0$. Since the fundamental variables of electrodynamics are the potentials and not the field-strengths, then, in a vacuum, we obtain an action of the following form [10]

$$\mathcal{L}_{dBP} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2 A^\mu A_\mu \ , \tag{2.1}$$

we could imagine what would happen if we tried to couple the mass to the field strength tensor, but we can very quickly see that a quadratic term of the form m^2F^2 would be incorrect dimensionally, leading to a dimensionful action in the natural units system, unless we made use of higher derivatives on the field strengths like in the Bopp-Landé-Thomas-Podolsky model which is discussed later on. The immediate consequence of adding a massive term to the action is the loss of gauge invariance.

Indeed, considering the transformation $\delta A^\mu = \partial^\mu \Lambda$, we can show that

$$\delta_\Lambda \mathcal{L}_{dBP} = m^2 A^\mu \delta A_\mu = m^2 A^\mu \partial_\mu \Lambda = \left(-m^2 \partial^\mu A_\mu\right) \Lambda \ , \tag{2.2}$$

meaning that the action is not gauge invariant anymore. Additionally, by taking the variation of the action with respect to the 4-potential $\dfrac{\delta \mathcal{L}_{dBP}}{\delta A_\mu} = 0$, we can derive the equations of motion

$$\partial_\mu F^{\mu\nu} + m^2 A^\nu = 0 \ , \tag{2.3}$$

while the Ricci Curbastro-Bianchi [29] identity remains untouched $\partial_\mu \tilde{F}^{\mu\nu} = 0$. By

taking the divergence of (2.3)

$$\partial_\mu \partial_\nu F^{\mu\nu} + m^2 \partial_\nu A^\nu = 0 \Rightarrow \ \partial_\nu A^\nu = 0 \ , \tag{2.4}$$

we can obtain a subsidiary condition that is nothing more than the Lorenz gauge. Said subsidiary condition also shows that our loss of gauge-invariance is simply due to the fact that the system is already fully determined by its dynamics, there is no redundancy in the description. This fact will be better highlighted in the Hamiltonian treatment of the theory later on. The intuition we can naively derive from this, is that the presence of a mass shifts the nature of potentials from mathematical artifacts, to physically measurable quantities.

For now, we can proceed with the dispersion relations for the theory. From (2.3) and using the subsidiary condition (2.4) we get

$$\Box A^\nu + m^2 A^\nu = 0 \tag{2.5}$$

which is nothing more but a Klein-Gordon equation for a vector field. By then using the phase exponential ansatz $\partial^\mu = -ik^\mu$ we obtain

$$k^\mu k_\mu = m^2 \quad \Rightarrow \quad \omega^2 = k^2 + m^2 \tag{2.6}$$

which explicitly shows that there exists now a rest frame for photons. More interestingly, we can show that the phase and group velocities at first order, will be

$$v_p = \frac{\omega}{k} \approx \left(1 - \frac{m^2}{\omega^2}\right)^{-\frac{1}{2}} \ , \tag{2.7}$$

$$v_g = \frac{\partial \omega}{\partial k} \approx \left(1 - \frac{m^2}{\omega^2}\right)^{\frac{1}{2}} \ , \tag{2.8}$$

which explicitly shows that at the first order, the velocities would be $(\frac{1}{\omega})^2$ functions, implying that there would be a delay between high and low frequency emissions.

A very important prediction from dBP theory that holds for every massive model is the rapid decay of the range for electromagnetic interactions. To show this let us obtain the Gauss law from (2.3)

$$\partial_i F^{i0} + m^2 A^0 = J^0 \ , \tag{2.9}$$

$$\vec{\nabla} \cdot \vec{E} + m^2 \varphi = \rho \ , \tag{2.10}$$

by taking the static case, the equations reduces to

$$-\Delta\varphi + m^2\varphi = \rho \ , \tag{2.11}$$

expressing the Laplacian in spherical coordinates, we can show that the equation admits a solution of the form

$$\varphi(r) = q\frac{e^{-mr}}{r} \tag{2.12}$$

where q is the point charge. From the above, we draw the conclusion that the potential decreases exponentially with the distance, with a decrease proportional to the mass. The larger the mass, the quicker the decrease, which immediately puts into perspective that for our massive models to work, we need to have an incredibly small mass to be consistent with all of the laboratory, astrophysical and cosmological observations. A discussion about the various upper limits will be provided later in the manuscript. This modified Coulomb potential is also known in the literature as the Yukawa [18] potential, and is a key feature of massive particles.

To complete the discussion we will now turn to the Hamiltonian picture to count the physical degrees of freedom, and get more insight on the gauge structure of the theory. Starting from (2.1), we break the manifest covariance of the action

$$\mathcal{L}_{dBP} = -\frac{1}{2}F^{0i}F_{0i} - \frac{1}{4}FijF_{ij} + m^2\left(A^iA_i + A^0A_0\right) \ , \tag{2.13}$$

and then perform the Legendre transform

$$\pi^\mu = \frac{\partial\mathcal{L}_{dBP}}{\partial(\partial_0 A_\mu)} = F^{\mu 0} \ , \tag{2.14}$$

which yields exactly the same result as the Maxwellian case. We thus get a primary constraint $\phi_1 = \pi^0 = 0$ that will have to be added to the total Hamiltonian. Said total Hamiltonian will be defined as

$$\mathcal{H}_T = \mathcal{H} + \int d^3\vec{x}\ \lambda_0(x)\pi^0(x) \ , \tag{2.15}$$

where λ_0 is a Lagrange multiplier enforcing the primary constraint and $\mathcal{H}$ is

$$\mathcal{H} = \int d^3\vec{x}\ \left(-A_0(x)\partial_i\pi^i(x) + \frac{1}{2}\pi^i(x)\pi_i(x) + \frac{1}{4}F^{ij}(x)F_{ij}(x) - \frac{1}{2}m^2\left(A^0A_0 + A^iA_i\right)\right) \ . \tag{2.16}$$

Proceeding with the Dirac-Bergmann algorithm[1] [12, 13, 30], the consistency of constraints leads to

$$\{\pi^0, \mathcal{H}_T\} = \partial_i \pi^i + m^2 A^0 \approx 0 \ , \tag{2.17}$$

which is a secondary constraint $\phi_2 = \partial_i \pi^i + m^2 A^0 \approx 0$ representing the modified Gauss law for this model.The consistency of the constraint will give

$$\{\phi_2, \mathcal{H}_T\} = m^2 \left(\partial_i A^i + \lambda_0\right) \approx 0 \ , \tag{2.18}$$

this results in the multiplier λ_0 getting fixed such as

$$\lambda_0 = -\partial_i A^i \ , \tag{2.19}$$

there are thus no additional constraints. Computing the Poisson bracket between ϕ_1 and ϕ_2

$$\{\phi_1(t, \vec{x}) \ , \phi_2(t, \vec{y})\} = -m^2 \delta(\vec{x} - \vec{y}) \ , \tag{2.20}$$

we observe that it is not vanishing. More importantly, it forms a skew-symmetric matrix

$$\phi_{ij}(\vec{x}, \vec{y}) = \begin{pmatrix} 0 & -m^2 \\ m^2 & 0 \end{pmatrix} \delta^3(\vec{x} - \vec{y}) \ , \tag{2.21}$$

that is invertible since $m^2 \neq 0$. We will use this example to ensure that our models are properly described, as if the constraint matrix turns out to be non-invertible, it is indicating the presence of first class constraints, and thus some remnants of gauge invariance [11]. The matrix $\phi_{ij}(\vec{x}, \vec{y})$ also verifies [11]

$$\int d^3\vec{z} \sum_{k=1}^{2} C_{ik}(\vec{x}, \vec{z}) \ C^{-1}_{kj}(\vec{z}, \vec{y}) = \delta_{ij} \delta^3(\vec{x} - \vec{y}) \ . \tag{2.22}$$

Using the master formula

$$2 \times DOF = \text{ Phase Space variables} - 2 \times \text{First Class Constraints} - \text{ Second Class Constraints} \ , \tag{2.23}$$

[1] Introduced in the first chapter, refer to Appendix B for more details.

we show that the model possesses three degrees of freedom

$$2 \times DOF = 8 - 0 - 2 = 6 \quad \Rightarrow DOF = 3 \ . \tag{2.24}$$

Since the model is endowed of second-class constraints, we must now construct the Dirac bracket [13] of theory to reduce the phase space. For any function $F(t)$ and $G(t)$ of the phase space, the Dirac bracket will read

$$\{F(t),\ G(t)\}_D = \{F(t),\ G(t)\} - \sum_{i,j=1}^{2} \int d^3\vec{z}\ d^3\vec{x}\ \{F(t),\ \phi_i(t,\vec{z})\}\ \phi^{-1}_{ij}(\vec{z},\vec{u})\{\phi_j(t,\vec{u}),\ G(t)\}\ . \tag{2.25}$$

By definition, the Dirac bracket of constraints with any phase space functions cancels out [11]

$$\begin{aligned}\{F(t),\ \phi_k(t,\vec{x})\}_D &= \{F(t),\ \phi_k(t,\vec{x})\} - \sum_{i,j=1}^{2} \int d^3\vec{z}\ d^3\vec{x}\ \{F(t),\ \phi_i(t,\vec{z})\}\ \phi^{-1}_{ij}(\vec{z},\vec{u})\{\phi_j(t,\vec{u}),\ \phi_k(t,\vec{x})\} \\ &= \{F(t),\ \phi_k(t,\vec{x})\} - \sum_{i,j=1}^{2} \int d^3\vec{z}\ d^3\vec{x}\ \{F(t),\ \phi_i(t,\vec{z})\}\delta_{ik}\delta^3(\vec{z}-\vec{x}) = 0\ ,\end{aligned} \tag{2.26}$$

which means that now we can take our constraints as strong equalities and solve them. By replacing $\pi^0 = 0$ and $\partial_i\pi^i + m^2 A_0 = 0$ in the Total Hamiltonian, we obtained the so-called Reduced Hamiltonian

$$\mathcal{H}_{Red} = \int d^3\vec{x}\ \frac{1}{2}\left(\vec{\pi}\cdot\vec{\pi} + \vec{B}\cdot\vec{B} + + m^2\vec{A}\cdot\vec{A} + \frac{1}{m^2}(\vec{\nabla}\cdot\vec{\pi})^2\right)\ , \tag{2.27}$$

which is explicitly a function of only π^i and A^i, implying that the configuration space carries only three degrees of freedom, corresponding to the three components of spin $s = 1$ as expected from the vector representation of the Lorentz group and in agreement with the prediction made by (2.23). We can also see that through the use of the Dirac bracket, we circumvent the presence of ghosts induced by time-like polarisations, simplifying greatly the quantisation of the theory.

An important remark is that if the model is endowed with second class constraints, we must ensure that their number is pair. While in the literature [31], the possibility of half degrees of freedom is discussed, in the general structure of massless and massive Maxwellian models [32], the presence of odd numbers of second class constraints would lead to instabilities. As such, we shall use the canonical picture systematically in our model building to ensure the following conditions

- *The Hamiltonian picture of electrodynamical models possess at least one primary constraint, in the form of* $\pi^0 = 0$.
- *Said primary constraints lead to at least one secondary constraint, of which one acts as a modified Gauss law.*
- *If the number of constraints is odd , they must either be first class constraints, or a mix of an odd number of first class and even number of second class.*
- *If the model possesses an odd number of second class constraints, it will be considered as pathological.*
- *If after the Dirac bracket, the Reduced Hamiltonian of the theory is not definite positive, the model will be considered as unstable and require additional care when dealt with.*

We may however raise the following concern: in the previous section, we stated that gauge invariance is a very strong condition that plays an important role at the quantum level, notably this is what allows us to cancel the triangle diagrams in QED. Additionally, gauge invariance imposes the unphysical aspect of potentials, which is tightly confirmed experimentally (only gauge invariant quantities have ever been measured). We must then be very careful in how we introduce and deal with mass in electrodynamical models. An easy way to escape the loss of gauge invariance is through the Stueckelberg action.

2.1.2 Stueckelberg Trick

In 1957 [33] developed a method that would restore gauge invariance for vector fields. Now, while this method is usually referred to as the Stueckelberg trick, it is important to underline that is far from a simple trick. The Stueckelberg action (or trick or field) [34] is very well-defined. To highlight this fact let us make the following consideration: suppose that the action (2.1) is a function of a deformed 4-potential $\tilde{A}^\mu(x)$.

$$\mathcal{L}_{Stueck} = -\frac{1}{4}\tilde{F}^{\mu\nu}\tilde{F}_{\mu\nu} + \frac{1}{2}m^2\tilde{A}^\mu\tilde{A}_\mu \ . \tag{2.28}$$

Consider now that said 4-potential is nothing but a shift of the original one such as

$$\tilde{A}^\mu = A^\mu + \frac{1}{m}\partial^\mu\phi \ , \tag{2.29}$$

where $\phi(x)$ is a scalar field. By then considering the following gauge transformations

$$A'^\mu = A^\mu + \partial^\mu\Lambda \tag{2.30}$$

$$\phi' = \phi - \Lambda \ , \tag{2.31}$$

we can show that the action

$$\mathcal{L}_{Stueck} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2\left(A^\mu + \frac{1}{m}\partial^\mu\phi\right)^2 , \tag{2.32}$$

is gauge invariant. Roughly speaking, this is saying that the field $\tilde{A}^\mu$ is no longer the same one as described by Maxwell, but that the Maxwell potential A^μ is a part of $\tilde{A}^\mu$ and that it acquired a partner ϕ that enforces gauge invariance. When fully expanded the action reads

$$\mathcal{L}_{Stueck} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}\partial^\mu\phi\partial_\mu\phi + \frac{1}{2}m^2A^\mu A_\mu + mA^\mu\partial_\mu\phi , \tag{2.33}$$

showing that the mass mixes the longitudinal and transverse modes respectively represented by ϕ and A^μ. We can then immediately notice that by fixing $\phi = 0$, we retrieve the action (2.1), showing that the de Broglie-Proca action is nothing but a gauge fixing of the Stueckelberg action. This is a precursor of the Higgs mechanism, since we could say that the scalar field ϕ was a Goldstone Boson that was swallowed by the photon A^μ. This explains why the Proca theory, while not gauge invariant is still perfectly renormalisable and unitary at the quantum level, which is not true of non-abelian gauge theories [14, 18].

Let us turn our attention to the Hamiltonian picture of the theory. Following the same procedure as before, we write the action in a manifestly non covariant form

$$\mathcal{L}_{Stueck} = -\frac{1}{2}F^{0i}F_{0i} + \frac{1}{2}(\partial^0\phi\partial_0\phi + \partial^i\phi\partial_i\phi) - \frac{1}{4}FijF_{ij} + m^2\left(A^iA_i + A^0A_0\right) + m(A^0\partial_0\phi + A^i\partial_i\phi) . \tag{2.34}$$

Performing the Legendre transform on our two fields A^μ and ϕ we respectively get

$$\pi^\mu = \frac{\partial\mathcal{L}_{Stueck}}{\partial\dot{A}_\mu} = \begin{cases} \pi^0 = 0 \\ \\ \pi^i = F^{i0} \end{cases} , \tag{2.35}$$

and

$$P = \frac{\partial\mathcal{L}_{Stueck}}{\partial\partial_0\phi} = \partial^0\phi + mA^0 , \tag{2.36}$$

where we have one primary constraint $C_1 = \pi^0 \approx 0$. We then construct the total

Hamiltonian

$$\mathcal{H}_T = \int d^3\vec{x}\ \left[\frac{1}{2}P^2 - A_0(x)\partial_i\pi^i(x) + \frac{1}{2}\pi^i(x)\pi_i(x) + \frac{1}{4}F^{ij}(x)F_{ij}(x) - \frac{1}{2}m^2\left(\frac{1}{m^2}\partial^i\phi\partial_i\phi + A^iA_i\right)\right.$$
$$\left. -m\left(PA^0 + A^i\partial_i\phi\right)\right] + \int d^3\vec{x}\ C_1\lambda_1, \tag{2.37}$$

where λ_1 is a Lagrange multiplier enforcing the constraint. Evaluating the Poisson bracket of the primary constraint C_1 with the Hamiltonian we get the modified Gauss Law

$$\{C_1, \mathcal{H}_T\} = \{\pi^0, -A_0\partial_i\pi^i - mPA^0\} = \partial_i\pi^i + mP \approx 0\ , \tag{2.38}$$

which will give the modified Gauss law for the theory, which we will relabel as $C_2 = \partial_i\pi^i + mP \approx 0$. The consistency of the constraint C_2

$$\{C_2, \mathcal{H}_T\} = 0, \tag{2.39}$$

leads to no additional constraints. The Poisson bracket between the primary and secondary constraint

$$\{C_1, C_2\} = 0, \tag{2.40}$$

trivially vanishes, showing that they are first class.

A quick computation of the degrees of freedom for the system

$$2 \times DOF = (8+2) - 2 \times 2 - 0 = 6 \quad \Rightarrow DOF = 3\ . \tag{2.41}$$

confirming that the model is indeed equivalent to a massive scalar field, while preserving the gauge invariance.

2.1.3 Bopp-Landé-Thomas-Podolsky Model

As we saw in the Stueckelberg section, one way to add mass and to conserve gauge invariance is to add external fields (notably scalar). Another approach, which is more problematic consists on adding higher derivatives. The earliest example (and probably simplest) would be the Bopp-Landé-Thomas-Podolsky (BLTP) model [35–37] often only referred to as Bopp-Podolsky.

Before ruining the party by enumerating the problems that stem from higher derivatives, we shall offer a partial outlook on the theory, focusing on what makes it so appealing (at least at the classical level). The core idea behind BLTP is to regulate the

electron self-energy classically, by modifying the electrostatic potential in a way that does not diverge anymore at the origin in contrast to the Coulomb potential. This has the clear advantage of dealing with the radiation reaction problem in electromagnetism without any of the divergences inherent to the Maxwellian theory. The modification to the electrostatic potential consists on a superposition of the Coulombic potential and a Yukawa potential such as

$$\phi(r) = \frac{e^2}{r}\left(1 - e^{-\frac{r}{l}}\right) \ , \tag{2.42}$$

also known as a Hellmann potential [38], where l is a length parameter taken to be very small implying that the reciprocal parameter $m = \frac{1}{l}$ should be very large, which we will show is not a problem in itself. For $l = 0$ or $r >> 0$ we retrieve the usual Maxwellian theory.

Taking all of these considerations into account, we will be using the following form of the action

$$\mathcal{L}_{BLTP} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{l^2}{2}\partial_\mu F^{\mu\nu}\partial^\sigma F_{\sigma\nu} \ , \tag{2.43}$$

which is manifestly gauge invariant under U(1) transformations $\delta A^\mu = \partial^\mu \Lambda$ since it composed only of field strength tensors. The presence of higher derivatives has the added benefit of improving the renormalisation process [39], even though this benefit is in appearance only, as we will discuss later down the line. Before ruining our fun, let us take the variation of the action (2.43) with respect to A^μ

$$\delta\mathcal{L}_{BLTP} = -F^{\mu\nu}\partial_\mu \delta A_\nu + l^2 \partial_\mu F^{\mu\nu}(\Box \delta A_\nu - \partial^\nu \partial^\sigma \delta A_\sigma) \ , \tag{2.44}$$

where the third term will vanish after one integration by parts due to the contraction of $F^{\mu\nu}$ with $\partial_\mu\partial_\nu$. This leads to

$$\partial^\mu F_{\mu\nu} + l^2 \Box \partial^\mu F_{\mu\nu} = 0 \ . \tag{2.45}$$

By setting the Lorenz gauge condition $\partial^\mu A_\mu = 0$ we can simplify further the equations

$$(1 + l^2\Box)\Box A_\mu = 0 \ , \tag{2.46}$$

indicating the propagation of two independent modes

$$\left(\frac{1}{l^2}+\Box\right)A^\mu=0\ , \tag{2.47}$$

$$\Box A^\mu=0\ . \tag{2.48}$$

Using the phase exponential ansatz $\partial^\mu=-ik^\mu$, we get

$$k^\mu k_\mu=\frac{1}{l^2}\ , \tag{2.49}$$

$$k^\mu k_\mu=0\ , \tag{2.50}$$

which confirms the presence of two independent modes, one massive and one massless. In this sense, BLTP is a theory that describes two photons. In the presence of a current, equation (2.46) will read

$$(1+l^2\Box)\Box A_\mu=J_\mu\ , \tag{2.51}$$

and will be no longer be separable. We can however introduce two auxiliary potentials (and their respective Lorenz gauge) [7]

$$\tilde{A}_\mu=l^2\Box A_\mu\ ,\qquad \partial^\mu\tilde{A}_\mu=0 \tag{2.52}$$

$$\hat{A}_\mu=A_\mu+l^2\Box A_\mu\ ,\qquad \partial^\mu\hat{A}_\mu=0\ \ , \tag{2.53}$$

which then leads to two independent equations of motion for each potential

$$\Box\hat{A}_\mu=J_\mu \tag{2.54}$$

$$\Box\tilde{A}_\mu+\frac{1}{l^2}\tilde{A}_\mu=J_\mu\ , \tag{2.55}$$

which confirms the presence of two photons, one massive and the other massless.

This fact can also be made explicit in the action: Let us consider the following

$$\mathcal{L}=-\frac{a}{4}F^{\mu\nu}F_{\mu\nu}-\frac{b}{2}\partial_\mu F^{\mu\nu}\partial^\alpha F_{\alpha\nu}+\frac{b}{2c^2}\left(B_\nu-c\partial^\mu F_{\mu\nu}\right)\left(B^\nu-c\partial^\alpha F_{\alpha\nu}\right)\ , \tag{2.56}$$

where a,b,c are arbitrary coefficients and B^ν is an auxiliary field for which the equation of motion reads

$$B^\nu=c\partial_\mu F^{\mu\nu}\ , \tag{2.57}$$

that once plugged into the action kills the third term and retrieves the BLTP theory.

Hence, action (2.56) and (2.43) are equivalent (up to coefficients). We may then proceed with (2.56) and expand the third term

$$\mathcal{L} = -\frac{a}{4}F^{\mu\nu}F_{\mu\nu} + \frac{b}{2c^2}B_\nu B^\nu - \frac{b}{c}B_\nu \partial_\mu F^{\mu\nu} \ , \tag{2.58}$$

by then integrating by parts the third term, and completing the square we get

$$\mathcal{L} = -\frac{a}{4}\left(F^{\mu\nu} - \frac{b}{ac}B^{\mu\nu}\right)\left(F_{\mu\nu} - \frac{b}{ac}B_{\mu\nu}\right) + \frac{b^2}{4ac^2}B^{\mu\nu}B_{\mu\nu} + \frac{b}{2c^2}B_\nu B^\nu \ , \tag{2.59}$$

where $B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu$. From the first term we can deduce the presence of a new gauge field C^μ defined as

$$C^\mu = A^\mu - \frac{b}{ac}B^\mu \ , \tag{2.60}$$

and its associated field strength $C^{\mu\nu} = \partial^\mu C^\nu - \partial^\nu C^\mu$. Replacing them in (2.59), we get

$$\mathcal{L} = -\frac{a}{4}C^{\mu\nu}C_{\mu\nu} + \frac{b^2}{4ac^2}B^{\mu\nu}B_{\mu\nu} + \frac{b}{2c^2}B_\nu B^\nu \ , \tag{2.61}$$

explicitly showing that the model is composed of two photons, one massless, and one massive with a mass squared equal to $\frac{b}{c^2}$ with opposite kinetic signs, indicating the presence of ghosts.

By studying the Hamiltonian picture, we could confirm that the model indeed propagates five degrees of freedom but, something else is confirmed: The Hamiltonian is not bound from below. This is known as an Ostrogradski type instability, an artifact of higher derivatives, and one of the worst problems one could encounter for quantisation. Indeed, Ostrogradski instabilities force us to pick between unitarity and the positivity of energy. And particles with negative energies would be a disaster, as their lifetime would be zero since their vacuum state would be unstable (we could hypothetically create photons from nothing for example). Despite this apparent issue, BLTP is a remarkably well-behaved theory classically, and is still able to make predictions in the semi-classical limit [38, 40].

2.1.4 Maxwell-Chern-Simons Theory and topological mass generation

The case of Maxwell-Chern-Simons [41] theory requires a bit more attention, as we are dealing with what is known as a topological field theory [42]. Topological field theories are a family of gauge theories, offering an alternative construction to Yang-Mills theories [14]. A common misconception is that topological spaces are not endowed with metrics. This is incorrect, and a correct statement would be that topological spaces

are not necessarily endowed with a metric. However, topological properties of said spaces should not depend on a metric, leading to the notion of topological invariants. A purely topological theory should carry zero physical degrees of freedom, as its variation with respect to the metric tensor $g^{\mu\nu}$ is identically null[2], leading to a vanishing energy-momentum tensor.

We may then wonder what mass as a topological invariant would look like, and immediately draw a conclusion: it can only be inertial, not gravitational. One class of topological invariants are known as the Chern-Simons invariants, and can be constructed in the case of electrodynamics only in 2+1 dimensions. Indeed, the Chern-Simons term exists on the boundary of the space, and in the case of Minkowski, the space is not compact, making it impossible to have such boundary terms exist.

To begin, let us recall some features of electrodynamics in 2+1 D. In 4D, the magnetic field is a pseudo(axial)-vector defined by its relation with the electromagnetic field strength as $B^i = -\frac{1}{2}\varepsilon^{ijk}F_{jk}$. The dimensional reduction transforms it into a pseudo-scalar defined as $B = -\frac{1}{2}\varepsilon^{ij}F_{ij}$, while the electric field remains a vector albeit a two dimensional one. This means that the 2-form F possess only three components[3].

As stated above, in three dimensions we can construct the Chern-Simons term, which also happens to be a Lorentz invariant and local[4]. Said term is defined as

$$\mathcal{L}_{CS} = \frac{k}{2}\varepsilon^{\mu\nu\sigma}A_\mu\partial_\nu A_\sigma \ , \tag{2.62}$$

where k is a constant of positive mass dimension. The explicit dependence on the potential A^μ may raise some concerns around the gauge invariance, and we shall ensure its presence. Consider the transformation $\delta A^\mu = \partial^\mu\Lambda$ acting on $\mathcal{L}_{CS}$

$$\begin{aligned}\delta\mathcal{L}_{CS} =& \frac{k}{2}\varepsilon^{\mu\nu\sigma}(\delta A_\mu\partial_\nu A_\sigma + A_\mu\partial_\nu\delta A\sigma) \\ =& \frac{k}{2}\epsilon^{\mu\nu\sigma}(\partial_\mu\Lambda\partial_\nu A_\sigma + A_\mu\partial_\nu\partial_\sigma\Lambda) \ ,\end{aligned} \tag{2.63}$$

the second term vanishes identically, while the first one requires an integration by part. We are only left with a divergence term $\partial_\mu(\Lambda\varepsilon^{\mu\nu\sigma}\partial_\nu A_\sigma)$ which is taken to be zero at the boundary. The Chern-Simons term is thus gauge invariant.

[2]This fact holds only for classical topological field theories, as topological quantum field theories can propagate degrees of freedom post-quantisation, as shown for example by Witten in the case of 2+1 D gravity [43]

[3]This result is trivial, as any anti symmetric tensor possess $\frac{D(D-1)}{2}$ components where D is the dimension of the space it exists in.

[4]We will see later that, constructing a Chern-Simons term in four dimensions lead to Lorentz Symmetry Violations.

Taking the variation of (2.62) with respect to the potential A^μ we get

$$\varepsilon^{\mu\nu\sigma}\partial_\nu A_\sigma = 0 \ , \tag{2.64}$$

which is a first order equation, making it fundamentally different to the Maxwell theory.

Performing the Legendre transform

$$\pi^\mu = \frac{\partial \mathcal{L}_{CS}}{\partial \dot{A}_\mu} = \begin{cases} \pi^0 = 0 \\ \\ \pi^i = \frac{k}{2}\varepsilon^{0ij}A_j \end{cases} \ , \tag{2.65}$$

shows the presence of 4 primary constraints which we will name $\phi_0 = \pi^0 \approx 0$ and $\phi^i = \pi^i - \frac{k}{2}\varepsilon^{0ij}A_j \approx 0$ and then leads us to the following Hamiltonian

$$\begin{aligned} \mathcal{H} &= \int d^2\vec{x} \ \ \pi^i \partial_0 A_i - \mathcal{L}_{CS} \\ &= -\int d^2\vec{x} \ \ \frac{k}{2} A_0 \varepsilon^{0ij}\partial_i A_j \ , \end{aligned} \tag{2.66}$$

which does not contain any momenta (and thus no kinetic energy). Adding the primary constraints to the Hamiltonian

$$\mathcal{H}_T = \mathcal{H} + \int d^2\vec{x} \ \lambda_0\phi^0 + \lambda_i\phi^i \ , \tag{2.67}$$

and proceeding with the Dirac Bergmann algorithm

$$\{\pi^0, \mathcal{H}_T\} = \frac{k}{2}\varepsilon^{0ij}\partial_i A_j \approx 0 \ \ , \tag{2.68}$$

Continuing with the algorithm,

$$\{\phi_2, \mathcal{H}_T\} = \{\frac{k}{2}\varepsilon^{0ij}\partial_i A_j, \lambda_i\pi^i\} \Rightarrow \lambda_i = 0 \ , \tag{2.69}$$

yielding no additional constraint. The Poisson bracket between ϕ_0 and all the other constraints vanishes, making it first class, while the Poisson bracket between ϕ_i and ϕ_2 does not, making them second class. This leaves us with a total of one first class constraint and four second class, reducing the number of degrees of freedom of the phase space to zero, confirming the topological nature of the term.

However, a model that does not propagate any degrees of freedom, and that fails to reproduce any known results of electrodynamics in any limit is uninteresting. We shall turn our attention to a slightly different model, known as Maxwell-Chern-Simons

theory. The action reads

$$\mathcal{L}_{MCS} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{k}{2}\varepsilon^{\mu\nu\sigma}A_\mu\partial_\nu A_\sigma \ , \tag{2.70}$$

and its variation with respect to the potential reads

$$\partial_\mu F^{\mu\nu} + \frac{k}{2}\varepsilon^{\nu\alpha\beta}F_{\alpha\beta} = 0 \ , \tag{2.71}$$

defining the Maxwell dual as $\tilde{F}^\nu = \frac{1}{2}\varepsilon^{\nu\alpha\beta}F_{\alpha\beta}$ and contracting the above equation with a Levi-Civita tensor, we get

$$\partial^\nu \tilde{F}^\mu - \partial^\mu \tilde{F}^\nu + kF^{\mu\nu} = 0 \ . \tag{2.72}$$

By then taking the divergence of the above equation and then using (2.71) we finally get

$$(\partial^\mu\partial_\mu + k^2)\tilde{F}^\nu = 0 \ , \tag{2.73}$$

which describes a single massive degree of freedom with mass k. From this equation, we can also get the dispersion which will be of the form $E^2 = P^2 + k^2$.

Topological mass is a rather peculiar one. Even though we stated above that it does not appear in the energy-momentum tensor, we can clearly see that it does indeed show up in the dispersion relations. We can however, from the EMT define the momentum and energy functions using $E = \int d^3x \ T^{00}$ and $P^i = \int d^3x \ T^{0i}$, and since the mass is not present in $T^{\mu\nu}$ we may naively assume that we are running in a contradiction. However, the integration over the whole space plays a crucial role here, and allows us to deduce that the mass lays on the boundary of the space.

So far, we've only discussed topological mass generation in three dimensions. Constructing a Chern-Simons invariant in even dimensional spaces being impossible, we might naively believe that this type of mechanism is bound to be for lower (or higher) dimensional models. It is nevertheless important to remind the reader that Chern-Simons invariants are not the only type of topological invariants, and we can in fact generate topological masses in even dimensions through the introduction of additional fields, and topological interactions [44, 45]. It can be shown that in these cases, the mass term has the added property of being carried by the transverse part of the propagators [46].
The attentive reader may then wonder the relevance of this section, where we discuss a $(2+1)$ theory, while the rest of the thesis only discusses four dimensional models. The

point we wished to get across is that, certain mechanisms for mass generation can shift poles, and generate rest frames without altering the number of degrees of freedom of the theory, as well as their transversality. This will be of extreme relevance in the last chapter of this thesis, when we will introduce a mechanism that seems very similar to the one discussed above, without however, being topological.

2.1.5 Bounds on the photon mass

While in the above sections we discussed the structural differences between models, and potential mechanisms for mass generation in electrodynamics, we were very careful to not mention anything about the bounds that exist on photon masses. By essence, massive electrodynamics being a deformation of Maxwell, we expect the mass parameter to be extremely small[5]. A rather high number of methods and techniques have been used to set these bounds, and instead of carrying a detailed analysis and explanation of each of these (which is wildly beyond the scope of this thesis)[6], we shall summarize some of them in the following non-exhaustive table:

Author and Reference	**Value (kg)**	**Method**
de Broglie [27]	10^{-53}	Dispersion relations (first historical estimate)
Barnes and Scargle [53]	3×10^{-57} 3×10^{-56}	Observations of Crab Nebula magnetohydrodynamic waves
Ryutov [54]	1.4×10^{-54}	Solar wind magnetohydrodynamic model (40 AU), official PDG limit
de Bernadis et al. [55]	2.9×10^{-54}	Cosmic microwave background dipole anisotropy
Yang and Zhang [56]	6.3×10^{-53} 9.6×10^{-53}	Pulsar spin-down observations
Ryutov [57]	10^{-52}	Solar wind magnetohydrodynamic model (1 AU)
Adelberger et al. [58], Yamaguchi [59], Byrne and Burman [60, 61], Chibisov [62], Byrne [63]	3×10^{-63} 3.6×10^{-52}	Galactic potential models
Lakes [64], Luo et al. [65], Tu et al. [66]	9×10^{-56} 3.3×10^{-52}	Torsion pendulum experiments
Füllekrug [67]	3.6×10^{-52}	Lightning discharge propagation speed in the troposphere
Spallicci et al. [68]	1.1×10^{-51} 2.8×10^{-51}	Angular momentum law using MMS satellite data
Bonetti et al. [69, 70], Xing et al. [71], Wang et al. [72]	3.1×10^{-51} 5.2×10^{-50}	Fast Radio Burst observations

Table 2.1: Photon mass upper limits below 10^{-50}**kg**.

[5] With the added exception of Bopp-Podolsky Electrodynamics where it is extremely large, so large in fact that it cannot propagate.

[6] For full reviews see [47–52]

2.2 Non-Linear Theories

As discussed in chapter one, quantum electrodynamics is inherently a non-linear theory in opposition to electromagnetism. In the same vein of extensions that General relativity was to Newtonian gravitation, a full non-linear realisation of electromagnetism at the classical level may yield some new powerful predictions. In this context, we can underline two types of non-linearities, those that are classical, *i.e.* the model is fundamentally non-linear, and those stemming from quantum effects, implying that we are dealing with effective actions.

Regardless of the origin of the non-linearities, those considered in this section share a common point: they belong to the Plebański class of electrodynamics [73]. This family of models includes all the actions that are strictly functions of the Maxwell invariants $\mathcal{L}(\mathcal{F},\mathcal{G})$ as well as Lorentz invariant, and are proven to be systematically massless. In what follows, we shall introduce the two best established representative of said class: the Born-Infeld and Heisenberg-Euler theories. Before going further let us recall the definition of the two Maxwell invariants

$$\mathcal{F} = -\frac{1}{4}F^2 = -\frac{1}{4}F_{\sigma\tau}F^{\sigma\tau} = \frac{1}{2}\left(\vec{E}^2 - \vec{B}^2\right) \ , \tag{2.74}$$

$$\mathcal{G} = -\frac{1}{4}F_{\sigma\tau}\tilde{F}^{\sigma\tau} = -\frac{1}{4}F_{\sigma\tau}G^{\sigma\tau} = \vec{E}\cdot\vec{B} \ , \tag{2.75}$$

2.2.1 Born-Infeld model

The earliest formulation of a non-linear theory of electromagnetism dates back to 1912, when Gustave Mie attempted to handle in a consistent fashion point charges [74, 75]. However, the theory breaking gauge invariance was disregarded, and non-linear electromagnetism had to wait until 1933 before Born [76] offered a second wind to the field, which was then improved upon by Infeld and himself [77] birthing the Born-Infeld model.

To motivate the modification, let us go back to the Maxwell action:

$$\mathcal{L}_{Max} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \ , \tag{2.76}$$

from which we can derive Coulomb's law, as discussed previously. Something we can immediately notice is that by taking the limit $r \to 0$, the electric field blows up to infinity, implying that the theory is ill-defined at the position of the particle. To circumvent this issue, Born and Infeld offered a procedure similar to special relativity, where the kinetic energy $\frac{1}{2}mv^2$ is replaced by $mc^2\left(1-\sqrt{1-\frac{v^2}{c^2}}\right)$, which when applied to the

Maxwell action gives

$$\mathcal{L}_{BI} = b^2 \left(1 - \sqrt{1 - \frac{F^{\mu\nu} F_{\mu\nu}}{2b^2}} \right) , \tag{2.77}$$

where b is a large constant of dimension $b = [M]^2$. Maxwell's theory can be straightforwardly recovered by Taylor expanding around $b^2 >> \mathcal{F}$, but now the full action has an upper bound on the electric field $\vec{E}$. In the case where $\vec{B} = 0$ and $|\vec{E}| > b^2$, the action becomes imaginary, and therefore abhorrent.

Equation (2.77) is nothing but a naïve ansatz, and the full expression for the Born-Infeld action in terms of Maxwell invariant is given as

$$\mathcal{L}_{BI} = b^2 \left(1 - \sqrt{1 - \frac{\mathcal{F}}{b^2} - \frac{\mathcal{G}^2}{b^4}} \right) , \tag{2.78}$$

and its equivalent form in Minkowski spacetime

$$\mathcal{L}_{BI} = b^2 \left(1 - \sqrt{- \det \left(\eta^{\mu\nu} + \frac{1}{b} F^{\mu\nu} \right)} \right) , \tag{2.79}$$

which is going to be very useful to derive the equations of motion. The full demonstration of the least action principle can be found in [78], and leads to the following equations of motion

$$\partial_\mu \left[\frac{1}{\sqrt{\kappa}} \left(F^{\mu\nu} - \frac{\mathcal{G}}{b^2} \tilde{F}^{\mu\nu} \right) \right] = 0 , \tag{2.80}$$

where κ is defined as

$$\kappa = 1 + \frac{1}{b^2} \left(\frac{1}{2} \mathcal{F} \right) - \frac{1}{b^4} \det(F^{\mu\nu}) . \tag{2.81}$$

While the Ricci Curbastro-Bianchi identity remains untouched[7]. If we decide to rewrite the equations in a not manifestly covariant form, using the macroscopic formu-

[7] As it should in any theory of electromagnetism that does not include magnetic monopoles.

lation of electromagnetism we can get the following set of formula

$$\vec{\nabla} \cdot \vec{D} = 0 \tag{2.82}$$

$$\vec{\nabla} \times \vec{H} = -\frac{\partial \vec{D}}{\partial t} \tag{2.83}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{2.84}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \ , \tag{2.85}$$

where the displacement vector and induction vector are redefined as

$$\vec{D} = \frac{1}{\kappa}\left(\vec{E} - \frac{\mathcal{G}\vec{B}}{b^2}\right) \tag{2.86}$$

$$\vec{H} = \frac{1}{\kappa}\left(\vec{B} - \frac{\mathcal{G}\vec{E}}{b^2}\right) \ . \tag{2.87}$$

This shows that even in a vacuum, waves in Born-infeld theory behave as in a medium. The non-linearities inherent to the model lead to non-trivial permeability and permittivity, as well as modified refractive indexes for the vacuum. This feature is a keystone of non-linear theories of electromagnetism, which usually leads to vacuum birefringence. Born-Infeld however is one of the very few models that does not exhibit such a feature [79].

Working out the Hamiltonian picture and constraints for non-linear theories is a rather difficult process. Usually carried out in the first order (also known as Plebanski [73]) formalism, we will restrain ourselves from diving into it in this part of the thesis. We will simply refer to Dirac's work that studied the fully constrained Hamiltonian of the theory [80] and obtained two propagating degrees of freedom, which is to be expected from a Plebanski-type model.

2.2.2 Heisenberg-Euler theory

While Born and Infeld's theory offered a fundamentally non-linear theory, in hope to regularise point charges at the classical level, Heisenberg and Euler's attempt was radically different in essence. Instead of constructing a classical action, Heisenberg [81], Euler [82], and often too forgotten Weisskopf [83], laid the ground work for what would become effective field theories.

Heisenberg and Euler derived their theories by searching exact solutions to the Dirac equation in a constant background field $F^{\mu\nu}$[8], but a more modern approach is by

[8]This is not equivalent as a constant potential A^{μ}.

studying QED in a background. Indeed, we have discussed in the first chapter the non-linear nature of QED, as well as certain properties of it such as vacuum polarisation. To deal with such things without too much hassle, let us turn our attentions towards effective actions.

By definition, an effective action is non-renormalizable (possessing thus a limited range of validity), and contains fewer degrees of freedom, in essence it is a simplification of the full theory. However, this simplification includes all of the effects that the fully theory should contain within the validity range of the effective theory. Historically, the best representation of effective theories, is Fermi's theory of beta decay. It worked remarkably well to describe the weak interaction, but its non renormalizability led to divergences in higher energy regimes, which were understood much later-on with the advent of electroweak theory.

In this sense, the Heisenberg-Euler theory is described by the action $\mathcal{L}_{eff}(A)$ when A^μ is associated with a constant electromagnetic field. The reason for its popularity is the vast array of things we can recover from it: Light-by light scattering in the low energy limit, the Schwinger pair creation, the QED β function and much more! In what follows, we will introduce the full action as well as its weak field limit, but we shan't go through the whole derivation, as it can be found in [20], and further discussions around its various extensions can be found in [84].

We shall simply state that the Heisenberg-Euler action is the renormalized effective action arising from integrating out a massive fermion for constant $F^{\mu\nu}$ [20], and its expression is as follows

$$\mathcal{L}_{EH} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{e^2}{32\pi^2}\int_0^\infty \frac{ds}{s}e^{is\varepsilon}e^{-sm^2}\left[\frac{\mathrm{Re}\cosh(es\sqrt{2(\mathcal{F}+i\mathcal{G})})}{\mathrm{Im}\cosh(es\sqrt{2(\mathcal{F}+i\mathcal{G})})}\mathcal{G} - \frac{4}{e^2s^2} - \frac{2}{3}F^{\mu\nu}F_{\mu\nu}\right] , \tag{2.88}$$

where s is Schwinger's proper time. There exists several equivalent forms to this action, but all of them lead to the same feature, a critical field, which unlike the one of Born-Infeld does not set an upper bound on the validity of the theory, but rather a bound on where the Heisenberg-Euler model takes precedence over standard electromagnetism. The critical field $\mathcal{E}$ is defined as

$$\mathcal{E} = \frac{m^2c^3}{e\hbar} \approx 4.4\ 10^{13}G\ . \tag{2.89}$$

Taking the weak field limit of the theory, we can get the greatly simplified expression

$$\mathcal{L}_{EH} = \mathcal{F} + \frac{2\alpha^2}{45m^4}\left(4\mathcal{F}^2 + 7\mathcal{G}^2\right)\ , \tag{2.90}$$

where we will from now on set $\frac{2\alpha^2}{45m^4} = \beta$. Since the theory derives directly from QED, a study of the degrees of freedom through the Hamiltonian will not be relevant, as we can immediately deduce that they are a total of two. Heisenberg-Euler theory does possess vacuum birefringence, and as such makes a prediction that has not yet been verified by experiments. Searches for vacuum birefringence induced by non-linearities are still extremely active, whether in laboratory settings, or in astrophysical and cosmological cases [85, 86].

These two non-linear models will be mentioned again in the next chapter, where we will introduce a generalised framework to deal with non-linear electromagnetism in a background.

2.3 Lorentz Symmetry Violations

The standard model of particle physics is an extremely well tested theory, that tends to shine by its precision, and as some may say, is the crown jewel of modern physics. Structurally speaking, the SM includes the three fundamental interactions (*i.e*, abelian (commutative) and non-abelian gauge theories) as well as fermions. As such, its composed of particles of spin 1 and $\frac{1}{2}$ with the sole exception of the Higgs field, which is spin 0.

Fermions (Spin $\frac{1}{2}$)	
Quarks	Up (u), Down (d), Charm (c), Strange (s), Top (t), Bottom (b)
Leptons	Electron (e), Muon (μ), Tau (τ), Neutrinos (ν_e, ν_μ, ν_τ)
Bosons ($SU(3) \times SU(2) \times U(1)$)	
Gauge Bosons (Spin 1)	Photon (γ), Gluon (g), $W^{\pm}$, Z^0
Scalar Boson (Spin 0)	Higgs (H)

Table 2.2: Standard Model of particle physics

The standard model does not include any particles of higher spin, despite their possible existence. Indeed, fields being representations of the Lorenz group, allow the existence of particles up to spin 2. However, the inclusions of particles of higher spin (notably the graviton) requires the addition of supersymmetry to express them in any non-trivial way due to the Coleman-Mandula theorem[9]. Despite the absence of gravitons, the standard model is still incredibly successful, and I would argue that most of

[9]To be more specific, representations of spin 3/2 are absolutely allowed within the framework of standard QFT (c.f the Rarita-Schwinger equation) but having been never observed before, there is no point in adding them by hand to the standard model.

its success is due to the remarkable (or unreasonable) power of relativistic field theory. Whether classical or quantum, field theory relies on some key ingredients, the most important of which is Lorentz Symmetry invariance.

For a relativistic theory to hold, we expect to obey exactly to the symmetries generated by the Lorentz (and by extension the Poincaré) group, under passive and active transformations. Invariance under said transformations imply some key symmetries: invariance by rotations generate the conservation of angular momentum, and spin. Invariance by translation generate the conservation of energy-momentum. And additionally, we can show that there exist an invariance under the discrete CPT (Charge, Parity, Time) reversal symmetry, known as the CPT theorem. Now, there is very little doubt around the fact that the standard model of particle physics is Lorentz invariant, the larger (and much more exciting) question at hand is whether the Lorentz invariance is a fundamental symmetry of nature, or simply an effective one. As such, when considering Beyond the Standard Model physics (BSM), we can look for theories that may exhibit such a property. In 1989 Kostelecky and Samuel [87–89] showed that in the context of bosonic string theory, tensor fields could condensate spontaneously breaking the Lorentz symmetry at high energies.

However, discussing of Lorentz violations in scenarii of extreme energy regimes is a tricky exercise, as our current experimental setups are in the impossibility to reach the required energies.

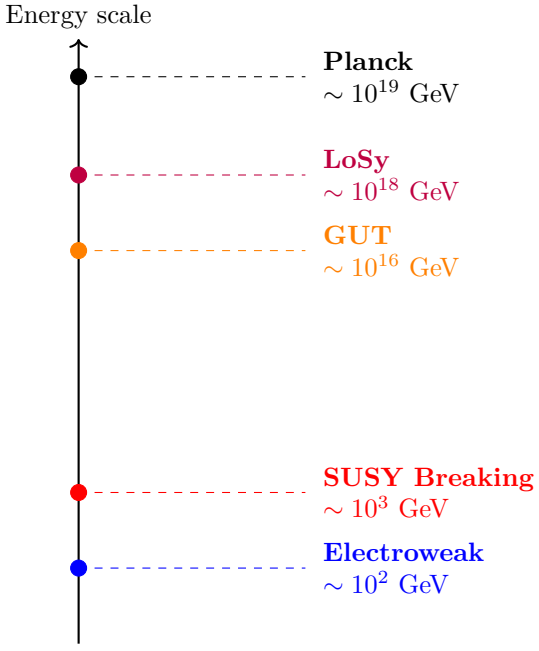


Figure 2.1: Energy scales at which symmetries are expected to break.

This leads us to the introduction of the Standard Model Extension (SME), which is the framework that deals with low energy manifestations of Lorentz Symmetry breaking. The SME is a vast and complex topic to discuss, and in what follows we will focus only on the photon sector of said framework. We will start by introducing the simplest CPT-odd (Or CPT breaking) action, in the form of the Carroll Field Jackiw (CFJ) model, before discussing the full photon sector (CPT-odd and CPT-even). Additionally, we will not consider any supersymmetric effects acting on the SME, nor glance at the fermionic sector.

2.3.1 Carroll-Field Jackiw Lagrangian

In their original paper, Sean Carroll, George Field and Roman Jackiw [90] offer an extension of Maxwell's theory by adding a Lorentz violation parameter in a similar fashion to the mass in the dBP theory. Indeed, naïve mass terms in electromagnetism will act as a gauge invariance violation parameters, and as discussed above, is well constrained. To consider Lorentz violations, they modified the Maxwell action with a Chern-Simons term. However, as showed previously, Chern-Simons terms can only exist in odd dimension with a scalar parameter. In even dimensions, said parameter must be of tensorial nature. Otherwise stated

$$\mathcal{L}_{CFJ} = \mathcal{L}_{Maxwell} + \mathcal{L}_{CS} \ , \tag{2.91}$$

with

$$\mathcal{L}_{CS} = -\frac{1}{2}\kappa^{\mu} A^{\nu} \tilde{F}_{\mu\nu} \ , \tag{2.92}$$

where $\tilde{F}^{\mu\nu}$ is the electromagnetic dual tensor. Coupling the theory to an undefined 4-vector raises the questions of gauge invariance, and by taking the transformation of the action under the usual $U(1)$ with respect to the gauge parameter Λ such as $\delta A^{\mu} = \partial^{\mu}\Lambda$, we get

$$\delta_{\Lambda}\mathcal{L}_{CS} = \frac{1}{4}\Lambda \tilde{F}^{\mu\nu} \left(\partial_{\mu}\kappa_{\nu} - \partial_{\nu}\kappa_{\mu}\right) \ , \tag{2.93}$$

up to a four-divergence. For gauge invariance to be upheld,(2.93) must vanish, requiring that κ^{μ} be a constant, automatically violating the Lorentz invariance. Indeed, any constant 4-vector implies a preferred direction, breaking the invariance under boosts and rotations of the theory. Additionally, the presence of the dual electromagnetic tensor explicitly breaks the CPT symmetry of the theory (thus CPT odd).

We may then write down the full action as

$$\mathcal{L}_{CFJ} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\kappa_\mu A_\nu F_{\rho\sigma} \ , \tag{2.94}$$

and its variation with respect to A^μ yielding

$$\partial_\mu F^{\mu\nu} + \frac{1}{2}\varepsilon^{\nu\mu\rho\sigma}\kappa_\mu F_{\rho\sigma} = 0 \ , \tag{2.95}$$

which are the modified equations of motion, while the Ricci Curbastro-Bianchi identity remained (of course) unaltered.

Expressing the equations in terms of potentials, and using the gauge invariance to our advantage to impose the Lorenz gauge condition $\partial^\mu A_\mu = 0$ we get

$$\Box A^\nu + \varepsilon^{\nu\mu\rho\sigma}\kappa_\mu\partial_\rho A_\sigma = 0 \tag{2.96}$$

$$\Rightarrow (\eta^{\nu\sigma}\Box + \varepsilon^{\nu\mu\rho\sigma}\kappa_\mu)A_\sigma = 0 \ , \tag{2.97}$$

and by then using the phase exponential ansatz $\partial_\mu = ik_\mu$ we express the latter equation in its momentum space form

$$(-\eta^{\nu\sigma}k^\mu k_\mu + \varepsilon^{\nu\mu\rho\sigma}\kappa_\mu k_\rho)\,\tilde{A}_\sigma(k) = 0 \tag{2.98}$$

$$\Rightarrow D^{\nu\sigma}\tilde{A}_\sigma(k) = 0 \ , \tag{2.99}$$

where $D^{\nu\sigma} = -\eta^{\nu\sigma}k^\mu k_\mu + \varepsilon^{\nu\mu\rho\sigma}\kappa_\mu k_\rho$. To obtain the dispersion relations (or equivalently the propagator) in this case, we must solve the matrix $D^{\nu\sigma}$ by finding its determinant and ensuring that it is not trivially zero. After some computations we find for the determinant

$$\det(D) = (k^\mu k_\mu)^2\left[(k^\mu k_\mu)^2 + \kappa^2 k^2 - (\kappa^\mu k_\mu)^2\right] \ , \tag{2.100}$$

and by extension for the dispersion relation

$$(k^\mu k_\mu)^2 + \kappa^2 k^2 - (\kappa^\mu k_\mu)^2 = 0 \ . \tag{2.101}$$

Assuming nothing around κ^μ we could set our dispersion relation to the rest frame (*i.e* $k^i = 0$) and see whether our minimal energy has shifted. Doing so gives

$$\omega^2(\omega^2 - \kappa^i\kappa_i) = 0 \ , \tag{2.102}$$

implying the existence of two modes, one trivial and one with non-zero vacuum energy (*i.e* massive like behaviour). However, non-zero rest energy does not mean that the

theory is necessarily massive, and we need to count the degrees of freedom. To do so we shall now move to the Hamiltonian representation of the theory.

As always we start by breaking the manifest covariance of the action

$$\mathcal{L}_{CFJ} = -\frac{1}{2}F^{0i}F_{0i} - \frac{1}{4}F^{ij}F_{ij} + \frac{1}{2}\varepsilon^{0ijk}\kappa_0 A_i F_{jk} - \frac{1}{2}\varepsilon^{0ijk}\kappa_i A_0 F_{jk} + \varepsilon^{0ijk}\kappa_i A_j F_{0k} \ , \tag{2.103}$$

we then continue with the Legendre transform

$$\pi^\mu = \frac{\partial \mathcal{L}_{CFJ}}{\partial \dot{A}_\mu} = \begin{cases} \pi^0 = 0 \\ \\ \pi^i = F^{i0} + \varepsilon^{0ijk}\kappa_j A_k \end{cases} \ , \tag{2.104}$$

where $\pi^0 = 0$ is the usual primary constraint, labelled as $\phi_1 = \pi^0 \approx 0$. Then, for the canonical Hamiltonian we get

$$\begin{aligned} \mathcal{H} &= \int d^3\vec{x} \ \ \pi^i \partial_0 A_i - \mathcal{L} \\ &= \int d^3\vec{x} \ \ - A_0\left(\partial_i \pi^i + \frac{1}{2}\varepsilon^{0ijk}F_{ij}\kappa_k\right) - \frac{1}{2}\pi^i\pi_i + \varepsilon^{0ijk}\pi_i\kappa_j A_k \\ &\qquad + \frac{1}{2}\left(\kappa_i\kappa^i A_j A^j - \kappa_i A^i \kappa_j A^j\right) + \frac{1}{4}F^{ij}F_{ij} + \frac{1}{2}\varepsilon^{0ijk}\kappa_0 A_i F_{jk} \ . \end{aligned} \tag{2.105}$$

which when augmented by the primary constraints gives the total Hamiltonian

$$\mathcal{H}_T = \mathcal{H} + \int d^3\vec{x}\lambda_a\phi^a(x) \tag{2.106}$$

By then proceeding with the Dirac Bergmann algorithm, we can show the presence of a secondary constraint

$$\{\phi_1, \mathcal{H}_T\} = \{\pi^0, \mathcal{H}_T\} = \partial_i\pi^i + \frac{1}{2}\varepsilon^{0ijk}F_{ij}\kappa_k \approx 0 \ , \tag{2.107}$$

namely the modified Gauss law for the theory now labeled as

$$\phi_2 = \partial_i\pi^i + \frac{1}{2}\varepsilon^{0ijk}F_{ij}\kappa_k \approx 0 \tag{2.108}$$

which shows that the violation vector couples to the magnetic field. Continuing with the algorithm, we can see that no additional constraints arise, nor Lagrange multiplier gets fixed, we may then proceed with the classification.

The Poisson bracket between the constraints identically vanishes

$$\{\phi_1, \phi_2\} = 0 \ , \tag{2.109}$$

endowing the theory with two first class constraints, and in the absence of any additional fields, reducing the number of physical degrees of freedom to two. This means that despite the presence of a rest frame (*i.e* minimal energy), the theory is not naïvely massive. There is however in the literature extensive discussions around this fact, and it can be proven that the CFJ action, can be indeed described as a massive theory of electrodynamics [91–93].

From the dispersion relations, we can show that the model possess vacuum birefringence [94], similarly to non-linear models. On a cosmological level, this should influence the anisotropies of the CMB, offering a new source of measurements for low energy remnants of LSV [95, 96].

2.3.2 Photon sector of the Standard Model extension

So far we have considered only a CPT-odd version of the Lorentz violation, and we shall now turn our attention to a CPT-even case. To preserve the CPT invariance, the violation cannot be induced by a vector field, and we must instead add a rank four tensor that will couple to the field strength tensors. The cpt-even term will thus read

$$\mathcal{L}_{CPT-even} = \kappa_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma} \tag{2.110}$$

where the tensor $\kappa_{\mu\nu\rho\sigma}$ exhibits the same symmetries as the tensor $P^{\mu\nu\rho\sigma} = F^{\mu\nu}F^{\rho\sigma}$ *i.e*, it is double traceless and $\kappa_{\{\mu,\nu\}\{\rho,\sigma\}}$ where $\{,\}$ indicates an antisymmetry between the indices and a symmetry between the pair of indices. This would lead to a total of 19 independent Lorentz violating parameter, describing the most general case possible for a CPT-even sector.

However, one particular case is if the LSV to arise from the condensation of chiral fermions, which would allow us to rewrite the violation tensor in terms of a single bosonic field ξ^μ, which will have the added benefit of avoiding super-fields of higher spin post supersymmetrisation. Skipping the details of the demonstration, we can show that

$$\kappa_{\mu\nu\rho\sigma} = \frac{1}{2} \left(\eta_{\mu\alpha}\kappa_{\nu\beta} - \eta_{\mu\beta}\kappa_{\nu\alpha} + \eta_{\nu\beta}\kappa_{\mu\alpha} - \eta_{\nu\alpha}\kappa_{\mu\beta} \right) \ , \tag{2.111}$$

where $\kappa_{\mu\nu} = \xi^\mu\xi^\nu - \frac{1}{4}\eta^{\mu\nu}\xi^\alpha\xi_\alpha$, which in turns re-expresses the cpt-even term as

$$\mathcal{L}_{CPT-even} = \frac{1}{4}\left(\frac{1}{2}\xi^\mu\xi_\nu F_{\mu\alpha}F^{\alpha\nu} + \frac{1}{8}\xi^\rho\xi_\rho F^{\mu\nu}F_{\mu\nu}\right) , \tag{2.112}$$

confirming the possibility of linking the CPT-even sector to the same source as the CPT-odd one. The disperion relations of the CPT-even sector and their derivations can be found in [94]. As such the photon sector of the standard model extension must include all of the terms leading to LSV, CPT-even or not, giving for the total action the form

$$\mathcal{L}_{LSV} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\kappa_\mu A_\nu F_{\rho\sigma} - \frac{1}{4}\kappa^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} . \tag{2.113}$$

To conclude this chapter, we shall summarise all of the possible extensions in a table with respect to each of the fundamental ingredients of electrodynamics.

$U(1)$ **invariance**	**Masslessness** $(m = 0)$	**Lorentz invariance** $(SO(1,3))$	**Linearity**	**Models**
y	y	y	y	Maxwell
n	n	y	y	de Broglie-Proca
y	n	y	y	Stueckelberg, Chern-Simons, Bopp-Podolsky, Kalb-Ramond etc.
y	y	y	n	Non-Linear Electrodynamics
n	n	y	n	Generalised Proca, Proca Nuevo
y	y	n	y	CPT-even sector before supersymmetrisation
y	n	n	y	Carroll-Field-Jackiw action, CPT-odd sector, CPT-even sector after photino integration
y	n	n	n	Generalised NLED in a background
y	y	n	n	Non Linear LSV

Table 2.3: All known extensions of electrodynamics and their corresponding features/violations

A small aside consists on model that exhibits non of these ingredients. Such a theory is unknown so far, and would, in my opinion be hard to qualify as an electromagnetic theory. If said model manages to properly reduce to Maxwell within a suitable limit, there should be, in principle, no opposition to its existence, but the maximal extension of electromagnetism seems to be the most unlikely candidate for experimental validation.

Chapter 3

Generalised Non Linear Electrodynamics

Résumé

Lors de ce chapitre, nous proposons une approche générale et générique des théories non linéaires de l'électromagnétisme en s'appuyant sur le formalisme du champ de fond. L'objectif est de construire un cadre unifié permettant de décrire les interactions entre un champ électromagnétique intense et des perturbations faibles, interprétées comme des ondes électromagnétiques.

Dans un premier lieu, le formalisme du champ de fond est présenté comme une méthode permettant de séparer un champ en deux contributions: une composante de fond, supposée forte et variant lentement, et une perturbation dynamique. Dans le contexte étudié, cette séparation est appliquée directement aux potentiels ou aux tenseurs de champ électromagnétique. Le champ total est ainsi décomposé en une partie de fond et une fluctuation, ce qui permet d'analyser la propagation des perturbations dans un milieu effectif défini par le champ de fond.

Lorsqu'on applique cette méthode à l'électromagnétisme de Maxwell, nous obtenons un Lagrangien contenant un terme mélangeant le champ de fond et la perturbation. Ce terme peut être réinterprété comme une interaction avec un courant effectif généré par le champ de fond. Ce dernier est trivialement conservé dû aux propriétés d'antisymétrie du champ de fond. Toutefois, même dans ce cadre classique, la présence d'un champ de fond empêche l'existence d'un véritable vide, introduisant ainsi des effets non-linéaires.

Le tenseur énergie-impulsion associé à la perturbation est altéré et nous montrons que dans le cas le plus général sa divergence n'est plus nulle en présence du champ de fond, lorsque ce dernier n'est pas constant. Cette non-conservation traduit un échange d'énergie entre la perturbation et le fond, ce qui peut être interprété comme un décalage ou une dérive de fréquence. L'argument repose sur l'identification de la densité d'énergie avec la composante temporelle du tenseur, et sur son lien avec la fréquence via la relation énergie-fréquence.

En appliquant le formalisme du champ de fond à des théories étendues telle que celle de Broglie-Proca ou dans certaines extensions du modèle standard, nous remarquons que ces variations de fréquences peuvent toujours émerger. Dans ces cas, le tenseur énergie-impulsion contient des contributions supplémentaires dépendant explicitement du champ de fond, et sa non-conservation peut apparaître même pour des champs constants. Les effets de décalage de fréquence deviennent alors plus généraux et peuvent dépendre des dérivées des potentiels de fond ou de paramètres de violation de symétrie.

La seconde partie du chapitre vise à construire une action générale pour les théories non linéaires de l'électromagnétisme. Cette construction repose sur plusieurs hypothèses: la possibilité de décomposer le champ en fond et perturbation (formalisme du champ de fond), la dominance du champ de fond, et l'existence d'une expansion en série des in-

variants électromagnétiques. Les invariants considérés sont les deux quantités scalaires construites à partir du tenseur électromagnétique et de son dual.

En développant l'action autour d'un champ de fond, on obtient une série en puissances de la perturbation. Les coefficients de cette expansion dépendent du fond et caractérisent le degré de non-linéarité de la théorie. À l'ordre le plus bas, le champ de fond agit comme un courant. Aux ordres supérieurs apparaissent des interactions plus complexes: interaction photon-fond, puis processus de création et annihilation de photons, et enfin diffusion photon-photon.

Les coefficients associés aux termes d'ordre élevé possèdent des dimensions d'énergie négatives, ce qui indique que la théorie n'est pas renormalisable. Cependant, au niveau classique, cela n'est pas problématique, car la théorie est interprétée comme une théorie effective valide jusqu'à une certaine échelle d'énergie.

Les équations du mouvement sont obtenues en variant l'action par rapport au potentiel de la perturbation. Elles prennent une forme très générale, contenant des termes linéaires et non linéaires en $f^{\mu\nu}$, ainsi qu'un courant dépendant du champ de fond. Ces équations peuvent être simplifiées dans le cas de champs de fond constants.

À partir de ces équations, on construit le tenseur énergie-impulsion invariant de jauge de la théorie. Celui-ci inclut des contributions d'ordre supérieur en la perturbation, reflétant les interactions non linéaires. Sa divergence contient plusieurs termes dépendant des dérivées du fond, ce qui confirme que la conservation de l'énergie-impulsion est en général violée.

Pour simplifier l'analyse, on se limite ensuite à une approximation quadratique en la perturbation. Cette version linéarisée conserve les effets essentiels tout en rendant les calculs plus tractables. Elle montre que le degré de non-linéarité dépend fortement du champ de fond: un fond purement maxwellien ne permet pas de générer des effets non linéaires pour les perturbations.

Nous mentionnons brièvement la violation des symétries de Poincaré. La non-conservation du tenseur énergie-impulsion implique une perte de l'invariance par translation, et l'analyse du moment angulaire montre également une violation de la symétrie de rotation. Cela remet en question les notions fondamentales de masse et de spin, associées aux opérateurs de Casimir du groupe de Poincaré. Dans ce cadre, il devient possible que ces quantités ne soient plus invariantes, ce qui ouvre la porte à des effets comme l'apparition de masses effectives pour les photons, qui seront explorées au chapitre suivant.

Nous poursuivons ensuite avec l'étude des relations de dispersion et des vitesses de groupe. En partant des équations du mouvement linéarisées, on exprime les équations de Maxwell modifiées en présence d'un champ de fond constant. Ces équations sont ensuite analysées pour différents cas particuliers.

Dans un premier cas, on considère une dépendance uniquement en un invariant $\mathcal{F}$. Les équations se simplifient et permettent d'obtenir une relation de dispersion explicite. Celle-ci révèle l'existence de plusieurs modes de propagation, dont certains correspondent au cas standard, tandis que d'autres dépendent des propriétés du fond. La présence de plusieurs solutions traduit un phénomène de biréfringence du vide.

Dans un second cas, on considère une théorie dépendant des deux invariants $\mathcal{F}$ et $\mathcal{G}$ mais avec une contribution uniquement magnétique. La structure des équations devient plus complexe, et la relation de dispersion se décompose en plusieurs branches. Les coefficients de la relation dépendent du champ magnétique et des constantes de la théorie, et les vitesses de groupe obtenues présentent une dépendance non triviale en la direction de propagation.

Enfin, le cas le plus général avec champs électrique et magnétique est abordé. Les équations sont réécrites en introduisant des vecteurs auxiliaires permettant de regrouper les contributions du fond. Malgré ces simplifications, l'expression finale de la relation de dispersion reste extrêmement complexe, illustrant la difficulté intrinsèque de l'analyse des théories non linéaires complètes.

La dernière partie du chapitre examine comment des théories spécifiques s'inscrivent dans ce cadre général. En développant l'action en série, on retrouve facilement la limite de faible champ de la théorie d'Heisenberg-Euler en choisissant des coefficients appropriés. En revanche, la reconstruction de la théorie de Born-Infeld est plus subtile et nécessite des manipulations supplémentaires ainsi qu'une redéfinition des champs.

3.1 Background Field Formalism

In this section, we aim to discuss in more details[1] the work described in [97]. To do so, we will have to familiarise the reader with the background field formalism, and some of the results previously obtained for Extended Theories of Electromagnetism (ETEMs) through its use.

The background field formalism (or method) [16, 98] is a powerful technique that was initially introduced by DeWitt [99] which allows its user to obtain effective actions around fixed backgrounds (such as the case of the Heisenberg-Euler action) for gauge invariant theories. In the full quantum representation, let $\phi(x)$ be a quantum field. By splitting said field into a classical background Φ and quantum fluctuations φ, we rewrite it as

$$\phi(x) = \Phi(x) + \varphi(x) \ , \tag{3.1}$$

which would then allow the expansion of the action around the aforementioned background. Now, $\Phi(x)$ can be taken to be constant, or as a solution of the classical equations of motion, allowing us to study better the low energy phenomenology of the theory[2].

Within the context of our work, we consider only the classical aspects of the discussed theories, and as such treating the fields as classical backgrounds and quantum fluctuations makes no sense. We thus adapt the method to our systems: we will be expanding our electromagnetic fields into strong external (or background fields) that change very slowly in time, and a perturbation that propagates (*i.e.* an electromagnetic wave). In a gauge invariant theory, we would thus split the field strength into

$$F_T{}^{\mu\nu} = F^{\mu\nu} + f^{\mu\nu} \ , \tag{3.2}$$

where $F^{\mu\nu}$ is the background field and $f^{\mu\nu}$ will represent our perturbation[3]. In what follows we shall always consider the background field $F^{\mu\nu}$ to be non-dynamical, and as such will neglect its equations of motion. While the split appears to be arbitrary, we will assume that the background field is some infrared entity while the electromagnetic wave encompasses the UV behaviour of the theory.

Performing the split at the level of the field strengths is however ill-suited to discuss certain models (as de Broglie-Proca) which explicitly break gauge invariance, and to generalise as much as possible the procedure, we shall consider the split at the level of

[1] This is a euphemism.

[2] By the very idea of effective actions, the theory resulting from such methods will be non-renormalisable and only hold up to specific energy scales.

[3] We will be using the word photon interchangeably with electromagnetic wave when referencing to the perturbation represented by the 4-potentials a^μ.

the potentials:

$$A_T{}^\mu = A^\mu + a^\mu \ , \tag{3.3}$$

where A^μ and a^μ are respectively the background and perturbation 4-potentials. Equation (3.3) has the benefit of reproducing (3.2), while leaving the additional liberty of keeping variable potentials even in the case of a constant field strength. Let us apply the procedure to the Maxwell action and see what we can get

$$\mathcal{L}_{Maxwell} = -\frac{1}{4}F_T{}^{\mu\nu}F_{T\mu\nu} = -\frac{1}{4}f^{\mu\nu}f_{\mu\nu} - \frac{1}{2}F^{\mu\nu}f_{\mu\nu} + \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \ . \tag{3.4}$$

Since we wish to neglect the dynamics of the background, we shall drop all the pure background terms in the action and thus get the following result

$$\mathcal{L}_{Maxwell} = -\frac{1}{4}f^{\mu\nu}f_{\mu\nu} - \frac{1}{2}F^{\mu\nu}f_{\mu\nu} \ , \tag{3.5}$$

which shows a mixing term between the background field and the photon field. Using integrations by parts we can however explicitly show that the background field shall act as a current, both at the level of the action and equations of motion. Indeed since

$$\mathcal{L}_{Maxwell} = -\frac{1}{4}f^{\mu\nu}f_{\mu\nu} + a_\nu(\partial_\mu F^{\mu\nu}) \ , \tag{3.6}$$

we can recast $\partial_\mu F^{\mu\nu} = J_B^\nu$, and by then taking the variation of the equation with respect to the photon 4-potential a^μ, we then obtain for the equations of motion

$$\partial_\mu f^{\mu\nu} + J_B^\nu = 0 \ , \tag{3.7}$$

and for their divergence

$$\partial_\nu\partial_\mu f^{\mu\nu} + \partial_\nu J_B^\nu = 0 \tag{3.8}$$

$$\Rightarrow \partial_\nu J_B^\nu = 0 \tag{3.9}$$

which shows that the current is conserved. In reality, the conservation of the current is a trivial equality since J_B^ν is the divergence of a totally anti-symmetric tensor. We thus conclude that when using the background field formalism, even Maxwell's theory leads to a sort of non-linear effect, since true vacua do not exist anymore due to the background fields.

By taking the equations of motions, and manipulating them (this of course can be obtained through the variation of the action with respect to the metric, but we will use a different procedure which we will detail later on), we can get the gauge invariant

energy-momentum tensor for the photon which will read

$$\theta^{\mu}{}_{\alpha} = f^{\mu\nu} f_{\nu\alpha} + \delta^{\mu}_{\alpha} \left(\frac{1}{4} f^{\beta\nu} f_{\beta\nu} \right) \ . \tag{3.10}$$

This is exactly the Maxwell energy-momentum Tensor (EMT), and this is coherent as this is not an extended theory of electromagnetism. However, its divergence is not equal to zero due to the background contribution

$$\partial_{\mu} \theta^{\mu}{}_{\alpha} = - \left(\partial_{\nu} F^{\nu\rho} \right) f_{\rho\alpha} \ , \tag{3.11}$$

this feature will be reappearing in most of the extended models we will discuss.
An important point to clarify is that at no point are we supposing that the total energy momentum tensor of the theory is not conserved. Indeed, if we decided to include the background action in the mix, and took the variation of said action with respect to the metric, we would get in the most general case, two expressions of the form

$$\partial_{\mu} \theta^{\mu}_{\alpha}(a^{\mu}) = \Gamma(A^{\mu}, a^{\mu}) \tag{3.12}$$

$$\partial_{\mu} \theta^{\mu}_{\alpha}(A^{\mu}) = -\Gamma(a^{\mu}, A^{\mu}) \tag{3.13}$$

which would impose that the total energy momentum tensor is conserved $\partial_{\mu} \theta^{\mu}_{\alpha}(A^{\mu}) + \partial_{\mu} \theta^{\mu}_{\alpha}(a^{\mu}) = 0$. But since we are putting ourselves in a scenario where can only observe the photon propagation, and know little about the background, we are indirectly including the background effects as some non-conservations of the energy momentum of the theory, or otherwise put, we consider that we are dealing with open systems.

In the Maxwellian case, the background field must be considered non-constant for this violation of the EMT to appear, but certain models do not require this. From now on, we shall consider that theories exhibiting non-conservations of the energy-momentum tensor possesses frequency shifts (or drifts)[4]. A naïve way to show this is by considering the following: Since the purely time-like component of the energy-momentum tensor represent the energy density $\mathcal{E}$

$$\theta^{00} = \mathcal{E} \ , \tag{3.14}$$

we can obtain the energy by integrating it over the entire volume

$$\int d^3x \ \ \theta^{00} = E \ . \tag{3.15}$$

[4]The nuance between a frequency shift and drift is due to the cumulative effect expected on (large) distances. A shift is a localised effect while a drift is taken to be resulting from a sum of shifts.

The divergence of the EMT will be for the time component

$$\partial_\mu \theta^{\mu 0} = \partial_0 \theta^{00} + \partial_i \theta^{i0} \ , \tag{3.16}$$

which we can integrate as before

$$\int d^3x \ \partial_0 \theta^{00} + \partial_i \theta^{i0} \ , \tag{3.17}$$

and drop the second term since it is a surface term. We are then left with the integral of the time variation of the energy density. Using the identity on sums (and by extension integrals)

$$\frac{\partial}{\partial x} \sum_y f(x,y) = \sum_y \frac{\partial}{\partial x} f(x,y) \ , \tag{3.18}$$

on $\mathcal{E}(t,x)$ we may get

$$\int d^3x \ \partial_t \mathcal{E} = \partial_t E \ . \tag{3.19}$$

The next step is a bit more conceptual, since we make use of the photon energy formula $E = h\nu$ where ν is the frequency, transforming the above expression into the time variation of the frequency, confirming the presence of frequency shifts (or drifts) if it is different than zero.

Let us return from this digression (or attention drift for the sake of the pun) and focus on the matter at hand. To further illustrate uses of the background field method, we shall apply it to the de Broglie-Proca theory and LSV, as discussed in [100]. In the case of the de Broglie-Proca theory, this yields an energy-momentum tensor of the form

$$\theta^\mu{}_\alpha = f^{\mu\nu} f_{\nu\alpha} + m^2 a^\mu a_\alpha + \frac{1}{2} m^2 (A^\mu a_\alpha + A_\alpha a^\mu) + \delta^\mu_\alpha \left(-\frac{1}{4} f^{\nu\rho} f_{\nu\rho} - m^2 a^\nu a_\nu - m^2 A^\nu a_\nu \right) \ , \tag{3.20}$$

which shows that the background field appears through the potentials in the energy-momentum of the theory. The mass parameter will thus act as a mixing constant, similar to the Stueckelberg theory (without preserving gauge invariance however). Its divergence reads

$$\partial_\mu \theta^\mu{}_\alpha = -\left(\partial_\nu F^{\nu\rho}\right) f_{\rho\alpha} + \frac{1}{2} m^2 \left(\partial_\alpha A^\beta + \partial^\beta A_\alpha \right) a_\beta \ , \tag{3.21}$$

meaning that even in the case of a constant background $F^{\mu\nu}$ energy-momentum is not

conserved. More precisely, the frequency shifts of the theory will be proportional to the derivatives of the background potentials in addition to the Maxwellian shift.

In the case of the photon sector of the SME, as derived in [94], the EMT will be

$$\theta^{\mu}{}_{\alpha} = f^{\mu\nu} f_{\nu\alpha} - \frac{1}{2}\kappa_{\alpha}\tilde{f}^{\mu\nu} a_{\nu} + \kappa^{\mu\rho\sigma\tau} f_{\sigma\tau} f_{\rho\alpha} + \delta^{\mu}_{\alpha}\left(\frac{1}{4} f^{\rho\sigma} f_{\rho\sigma} + \frac{1}{4}\kappa^{\rho\sigma\tau\lambda} f_{\rho\sigma} f_{\tau\lambda}\right) \ , \quad (3.22)$$

where there is no mixed term between background and photon. In the non conservation of the EMT however

$$\begin{aligned}\partial_{\mu}\theta^{\mu}{}_{\alpha} = &- \left(\partial_{\nu} F^{\nu\rho}\right) f_{\rho\alpha} - \left[\frac{1}{2}\left(\partial_{\mu}\kappa_{\alpha}\right)\tilde{f}^{\mu\nu} a_{\nu} - \frac{1}{4}\left(\partial_{\alpha}\kappa^{\mu\nu\rho\sigma}\right) f_{\mu\nu} f_{\rho\sigma}\right] \\ &- \left\{\partial_{\mu}\left(\kappa^{\mu\nu\rho\sigma} F_{\rho\sigma}\right) f_{\nu\alpha} + \kappa_{\mu}\tilde{F}^{\mu\nu} f_{\nu\alpha}\right\} \ , \end{aligned} \quad (3.23)$$

where the terms between straight brackets are totally independent of the background field, and only reflects the behaviour of the theory in the case of non-constant violation parameters (and will thus usually vanish identically), while the terms between curly brackets represent the energy exchange between the background field and the photon. In this case too, notably due to the CPT-odd contribution, there will exist a frequency shift even in the case of a constant background field.

Now that the background field formalism has been introduced, and some of its applications briefly discussed, we will attempt to construct a general (and generic) action to describe non-linear theories of electromagnetism.

3.2 Construction of the action, equations of motion and energy-momentum-tensor

Before proceeding, let us state clearly the assumptions that will lead to the action for the generic non-linear theory even if this seems redundant with the previous section:

- Electromagnetic fields $F^{\mu\nu}$ can always be split into a background and a perturbation.
- The background fields have a much larger amplitude than the perturbation, but evolve very slowly (if at all) so that their dynamics can be neglected.
- All non-linear theories of electromagnetism can be described by a power series expansion of the Maxwell invariants $\mathcal{F}$ and $\mathcal{G}$ modulo proper coefficients.
- The power-series expansion is well-behaved and converges, so that each order contributes less than the previous one.

Formally, this means that it is possible to build an action $\mathcal{L}(\mathcal{F},\mathcal{G})$ which will encompass most, if not all possible non-linear models. As a firs step, let us recall the definition of the two Maxwell invariants

$$\mathcal{F} = -\frac{1}{4}F^2 = -\frac{1}{4}F_{\sigma\tau}F^{\sigma\tau} = \frac{1}{2}\left(\vec{E}^2 - \vec{B}^2\right) \ , \tag{3.24}$$

$$\mathcal{G} = -\frac{1}{4}F_{\sigma\tau}\tilde{F}^{\sigma\tau} = -\frac{1}{4}F_{\sigma\tau}G^{\sigma\tau} = \vec{E}\cdot\vec{B} \ , \tag{3.25}$$

then let's proceed by splitting the field strength and its dual into a background and perturbation

$$F^{\mu\nu}_{Total} = F^{\mu\nu} + f^{\mu\nu} \ , \tag{3.26}$$

$$\tilde{F}^{\mu\nu}_{Total} = \tilde{F}^{\mu\nu} + \tilde{f}^{\mu\nu} \ , \tag{3.27}$$

which recasts the invariants as

$$\mathcal{F}_{Total} = \underbrace{-\frac{1}{4}F^2}_{\mathcal{F}} \underbrace{-\frac{1}{2}Ff}_{} \underbrace{-\frac{1}{4}f^2}_{\mathcal{F}_{\rm f}} = \mathcal{F} \underbrace{-\frac{1}{2}Ff + \mathcal{F}_{\rm f}}_{\delta\mathcal{F}} = \mathcal{F} + \delta\mathcal{F}, \tag{3.28}$$

$$\mathcal{G}_{Total} = \underbrace{-\frac{1}{4}F\tilde{F}}_{\mathcal{G}} \underbrace{-\frac{1}{2}\tilde{F}f}_{} \underbrace{-\frac{1}{4}f\tilde{f}}_{\mathcal{G}_{\rm f}} = \mathcal{G} \underbrace{-\frac{1}{2}\tilde{F}f + \mathcal{G}_{\rm f}}_{\delta\mathcal{G}} = \mathcal{G} + \delta\mathcal{G}, \tag{3.29}$$

where we used the following identity $F_{\rm B}\tilde{f} = \tilde{F}_{\rm B}f$. We shall now expand the action

around a fixed background up to the fourth order in the perturbation $f^{\mu\nu}$

$$\mathcal{L}(\mathcal{F}_{Total},\mathcal{G}_{Total}) = \mathcal{L}(\mathcal{F}_{\rm B}+\delta\mathcal{F},\mathcal{G}_{\rm B}+\delta\mathcal{G}) =$$
$$\mathcal{L}(\mathcal{F}_{\rm B},\mathcal{G}_{\rm B}) + \delta\mathcal{F}\frac{\partial\mathcal{L}}{\partial\mathcal{F}}\bigg|_{\rm B} + \delta\mathcal{G}\frac{\partial\mathcal{L}}{\partial\mathcal{G}}\bigg|_{\rm B} + \frac{\delta\mathcal{F}^2}{2}\frac{\partial^2\mathcal{L}}{\partial\mathcal{F}^2}\bigg|_{\rm B} + \delta\mathcal{F}\delta\mathcal{G}\frac{\partial^2\mathcal{L}}{\partial\mathcal{F}\partial\mathcal{G}}\bigg|_{\rm B} + \frac{\delta\mathcal{G}^2}{2}\frac{\partial^2\mathcal{L}}{\partial\mathcal{G}^2}\bigg|_{\rm B} +$$
$$\frac{\delta\mathcal{F}^3}{6}\frac{\partial^3\mathcal{L}}{\partial\mathcal{F}^3}\bigg|_{\rm B} + \frac{\delta\mathcal{F}^2\delta\mathcal{G}}{2}\frac{\partial^3\mathcal{L}}{\partial\mathcal{F}^2\partial\mathcal{G}}\bigg|_{\rm B} + \frac{\delta\mathcal{F}\delta\mathcal{G}^2}{2}\frac{\partial^3\mathcal{L}}{\partial\mathcal{F}\partial\mathcal{G}^2}\bigg|_{\rm B} + \frac{\delta\mathcal{G}^3}{6}\frac{\partial^3\mathcal{L}}{\partial\mathcal{G}^3}\bigg|_{\rm B} + \frac{\delta\mathcal{F}^4}{24}\frac{\partial^4\mathcal{L}}{\partial\mathcal{F}^4}\bigg|_{\rm B} +$$
$$\frac{\delta\mathcal{F}^3\delta\mathcal{G}}{6}\frac{\partial^4\mathcal{L}}{\partial\mathcal{F}^3\partial\mathcal{G}}\bigg|_{\rm B} + \frac{\delta\mathcal{F}^2\delta\mathcal{G}^2}{3}\frac{\partial^4\mathcal{L}}{\partial\mathcal{F}^2\partial\mathcal{G}^2}\bigg|_{\rm B} + \frac{\delta\mathcal{F}\delta\mathcal{G}^3}{6}\frac{\partial^4\mathcal{L}}{\partial\mathcal{F}\partial\mathcal{G}^3}\bigg|_{\rm B} + \frac{\delta\mathcal{G}^4}{24}\frac{\partial^4\mathcal{L}}{\partial\mathcal{G}^4}\bigg|_{\rm B} \ , \qquad (3.30)$$

and define the following coefficients

$$\frac{\partial\mathcal{L}}{\partial\mathcal{F}}\bigg|_{\rm B} = C_1 \ , \ \frac{\partial\mathcal{L}}{\partial\mathcal{G}}\bigg|_{\rm B} = C_2 \ , \ \frac{\partial^2\mathcal{L}}{\partial\mathcal{F}^2}\bigg|_{\rm B} = D_1 \ , \ \frac{\partial^2\mathcal{L}}{\partial\mathcal{F}\partial\mathcal{G}}\bigg|_{\rm B} = D_2 \ , \ \frac{\partial^2\mathcal{L}}{\partial\mathcal{G}^2}\bigg|_{\rm B} = D_3 \ ,$$
$$\frac{\partial^3\mathcal{L}}{\partial\mathcal{F}^3}\bigg|_{\rm B} = M_1 \ , \ \frac{\partial^3\mathcal{L}}{\partial\mathcal{F}^2\partial\mathcal{G}}\bigg|_{\rm B} = M_3 \ , \ \frac{\partial^3\mathcal{L}}{\partial\mathcal{F}\partial\mathcal{G}^2}\bigg|_{\rm B} = M_4 \ , \ \frac{\partial^3\mathcal{L}}{\partial\mathcal{G}^3}\bigg|_{\rm B} = M_2 \ , \ \frac{\partial^4\mathcal{L}}{\partial\mathcal{F}^4}\bigg|_{\rm B} = N_1 \ ,$$
$$\frac{\partial^4\mathcal{L}}{\partial\mathcal{F}^3\partial\mathcal{G}}\bigg|_{\rm B} = N_3 \ , \ \frac{\partial^4\mathcal{L}}{\partial\mathcal{F}^2\partial\mathcal{G}^2}\bigg|_{\rm B} = N_4 \ , \ \frac{\partial^4\mathcal{L}}{\partial\mathcal{F}\partial\mathcal{G}^3}\bigg|_{\rm B} = N_5 \ , \ \frac{\partial^4\mathcal{L}}{\partial\mathcal{G}^4}\bigg|_{\rm B} = N_2 \ .$$

where we left the subscript B to signify background fields, notably since the coefficients should be computed on the considered background action.

Ignoring the purely background contribution of $\mathcal{L}(\mathcal{F}_{\rm B},\mathcal{G}_{\rm B})$, we get at the fourth order the following Lagrangian density

$$\begin{aligned}\mathcal{L}_4 =& - C_1\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right) - C_2\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right) + \frac{D_1}{2}\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)^2 + \\ & D_2\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right) + \frac{D_3}{2}\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right)^2 - \frac{M_1}{6}\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)^3 - \\ & \frac{M_3}{2}\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)^2\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right) - \frac{M_4}{2}\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right)^2 - \\ & \frac{M_2}{6}\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right)^3 + \frac{N_1}{24}\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)^4 + \frac{N_3}{6}\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)^3\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right) + \\ & \frac{N_4}{3}\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)^2\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right)^2 + \frac{N_5}{6}\left(\frac{1}{2}Ff+\frac{1}{4}f^2\right)\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right)^3 + \\ & \frac{N_2}{24}\left(\frac{1}{2}\tilde{F}f+\frac{1}{4}f\tilde{f}\right)^4 \ , \end{aligned} \qquad (3.31)$$

which after some lengthy manipulations leads to the much nicer

$$\mathcal{L}_4 = -\frac{1}{2}\left(C_1 F^{\mu\nu} + C_2 \tilde{F}^{\mu\nu}\right) f_{\mu\nu} - \frac{1}{4} C_1 f^{\mu\nu} f_{\mu\nu} - \frac{1}{4} C_2 f^{\mu\nu} \tilde{f}_{\mu\nu} + \frac{1}{8} Q^{\mu\nu\kappa\lambda} f_{\mu\nu} f_{\kappa\lambda} + \frac{1}{8} R^{\mu\nu\kappa\lambda\rho\sigma} f_{\mu\nu} f_{\kappa\lambda} f_{\rho\sigma} + \frac{1}{16} S^{\mu\nu\kappa\lambda\rho\sigma\omega\tau} f_{\mu\nu} f_{\kappa\lambda} f_{\rho\sigma} f_{\omega\tau} \ , \tag{3.32}$$

having used

$$Q^{\mu\nu\kappa\lambda} = D_1 F^{\mu\nu} F^{\kappa\lambda} + D_3 \tilde{F}^{\mu\nu} \tilde{F}^{\kappa\lambda} + 2 D_2 \tilde{F}^{\mu\nu} F^{\kappa\lambda} \ , \tag{3.33}$$

where $Q^{\mu\nu\kappa\lambda}$ is anti-symmetric in $[\mu\nu]$ and $[\kappa\lambda]$ but symmetric under the exchange of the pairs $[\mu\nu]$ - $[\kappa\lambda]$;

$$R^{\mu\nu\kappa\lambda\rho\sigma} = D_1 F^{\mu\nu} \eta^{\kappa\rho} \eta^{\lambda\sigma} + \frac{1}{2} D_3 \tilde{F}^{\mu\nu} \varepsilon^{\kappa\lambda\rho\sigma} + \frac{1}{2} D_2 F^{\mu\nu} \varepsilon^{\kappa\lambda\rho\sigma} + D_2 \tilde{F}^{\mu\nu} \eta^{\kappa\rho} \eta^{\lambda\sigma} - \frac{1}{6} M_1 F^{\mu\nu} F^{\kappa\lambda} F^{\rho\sigma} - \frac{1}{6} M_2 \tilde{F}^{\mu\nu} \tilde{F}^{\kappa\lambda} \tilde{F}^{\rho\sigma} - \frac{1}{2} M_3 F^{\mu\nu} F^{\kappa\lambda} \tilde{F}^{\rho\sigma} - \frac{1}{2} M_4 F^{\mu\nu} \tilde{F}^{\kappa\lambda} \tilde{F}^{\rho\sigma} \ , \tag{3.34}$$

$$S_B^{\mu\nu\kappa\lambda\rho\sigma\omega\tau} = \frac{1}{2} D_1 \eta^{\mu\kappa} \eta^{\nu\lambda} \eta^{\rho\omega} \eta^{\sigma\tau} + \frac{1}{8} D_3 \varepsilon^{\mu\nu\kappa\lambda} \varepsilon^{\rho\sigma\omega\tau} + \frac{1}{2} D_2 \eta^{\mu\kappa} \eta^{\nu\lambda} \varepsilon^{\rho\sigma\omega\tau} - \frac{1}{2} M_1 F^{\mu\nu} F^{\kappa\lambda} \eta^{\rho\omega} \eta^{\sigma\tau} - \frac{1}{4} M_2 \tilde{F}^{\mu\nu} \tilde{F}^{\kappa\lambda} \varepsilon^{\rho\sigma\omega\tau} - \frac{1}{4} M_3 F^{\mu\nu} F^{\kappa\lambda} \varepsilon^{\rho\sigma\omega\tau} - M_3 F^{\mu\nu} \tilde{F}^{\kappa\lambda} \eta^{\rho\omega} \eta^{\sigma\tau} - \frac{1}{2} M_4 F^{\mu\nu} \tilde{F}^{\kappa\lambda} \varepsilon^{\rho\sigma\omega\tau} - \frac{1}{2} M_4 \tilde{F}^{\rho\sigma} \tilde{F}^{\omega\tau} \eta^{\mu\kappa} \eta^{\nu\lambda} + \frac{1}{24} N_1 F^{\mu\nu} F^{\kappa\lambda} F^{\rho\sigma} F^{\omega\tau} + \frac{1}{24} N_2 \tilde{F}^{\mu\nu} \tilde{F}^{\kappa\lambda} \tilde{F}^{\rho\sigma} \tilde{F}^{\omega\tau} + \frac{1}{6} N_3 F^{\mu\nu} F^{\kappa\lambda} F^{\rho\sigma} \tilde{F}^{\omega\tau} + \frac{1}{4} N_4 F^{\mu\nu} F^{\kappa\lambda} \tilde{F}^{\rho\sigma} \tilde{F}^{\omega\tau} + \frac{1}{6} N_5 F^{\mu\nu} \tilde{F}^{\kappa\lambda} \tilde{F}^{\rho\sigma} \tilde{F}^{\omega\tau} \ . \tag{3.35}$$

The parameters C, D, Q, R, S depend only on the background field. If the latter is constant, so will they.

Considering the terms in C_1, C_2 to be the first order in our infinite series, we can infer that at first order, the background only acts as a current. At second order, there are interactions between photon and background; at third photon splitting and merging becomes possible and finally at fourth, photon-photon scattering becomes possible.

An important thing to highlight is that the background-dependent tensorial coefficients, R and S, present negative mass dimensions: R with dimension $[E]^{-2}$ and S, with dimension $[E]^{-4}$, meaning that the quantum field theoretical version of the theory is non-renormalisable[5]. This would be an enormous problem, if we were attempting to

[5]This is an abuse of language. The theory is in fact not UV complete, and we will sometimes refer to non-UV completeness as non-renormalisability

describe a UV complete theory. But since we intend to study the model only classically, we will consider treat this as an effective classical field theory that holds below a well defined cut-off, fixed by the physics we investigate. In this case, the effective field theory stretches at most up to the GeV scale, since it is a purely photonic model, without available energy to excite the weak boson mediators (We, however, will not go even remotely close to these energies, as we are looking at the free propagation features of the model, and want to avoid quantisation in what follows).

To proceed with the fields equations, we need to take the variation of the action with respect to the photon four-potential a^μ

$$\begin{aligned}\delta\mathcal{L}_4 = \Bigg[&-\left(C_1 F^{\mu\nu} + C_2\tilde{F}^{\mu\nu}\right) - C_1 f^{\mu\nu} - C_2\tilde{f}^{\mu\nu} + \frac{1}{2}Q^{\mu\nu\kappa\lambda}f_{\kappa\lambda} \\ &+ \frac{3}{4}R^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma} + \frac{1}{2}S^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\Bigg]\partial_\mu\delta a_\nu \ , \end{aligned} \tag{3.36}$$

by then integrating by parts, and discarding all of the surface terms we get

$$\begin{aligned}\delta\mathcal{L}_4 = \Bigg[&\partial_\mu\left(C_1 f^{\mu\nu} + C_2\tilde{f}^{\mu\nu}\right) + \partial_\mu\left(C_1 F^{\mu\nu} + C_2\tilde{F}^{\mu\nu}\right) - \frac{1}{2}\partial_\mu\left(Q^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\right) - \\ &\frac{3}{4}\partial_\mu\left(R^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma}\right) - \frac{1}{2}\partial_\mu\left(S^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\right)\Bigg]\delta a_\nu \ , \end{aligned} \tag{3.37}$$

by then completing the procedure using $\dfrac{\delta\mathcal{L}}{\delta a_\nu} = 0$, the field equations will read

$$\begin{aligned}&\partial_\mu\left(C_1 f^{\mu\nu} + C_2\tilde{f}^{\mu\nu}\right) - \frac{1}{2}\partial_\mu\left(Q^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\right) - \frac{3}{4}\partial_\mu\left(R^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma}\right) \\ &- \frac{1}{2}\partial_\mu\left(S^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\right) + \mathcal{J}^\nu = 0 \ , \end{aligned} \tag{3.38}$$

where $\mathcal{J}^\nu = \partial_\mu\left(C_1 F^{\mu\nu} + C_2\tilde{F}^{\mu\nu}\right)$. These represent the most general case for the field equations, and will simplify greatly in the case of constant background fields.

With the equations of motion, we may now construct the gauge invariant energy-momentum tensor[6]. First, let us recall that due to the symmetries acting on the tensors $Q^{\mu\nu\rho\sigma}$, $R^{\mu\nu\rho\sigma\kappa\lambda}$ and $S^{\mu\nu\rho\sigma\kappa\lambda\alpha\beta}$ due to their contractions with the f fields, the only surviving components are those that are anti-symmetric in $\{\mu,\nu\};\{\rho,\sigma\}$ (so on so forth) while being symmetric on pair permutations. We will then index these remaining com-

[6]The resulting energy momentum tensor is equivalent to the Hilbert EMT, or the canonical one after Belinfante-Rosenfeld improvements.

ponents with a hat: $\hat{Q}$, $\hat{R}$ and $\hat{S}$. Starting off the field equations

$$\partial_\mu\left(C_1 f^{\mu\nu}+C_2\tilde{f}^{\mu\nu}\right)-\frac{1}{2}\partial_\mu\left(\hat{Q}_B^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\right)-\frac{3}{4}\partial_\mu\left(\hat{R}^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma}\right)-$$
$$\frac{1}{2}\partial_\mu\left(\hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\right)=-\mathcal{J}^\nu\ , \tag{3.39}$$

by then contracting the field equations with $f_{\nu\alpha}$ we get

$$\partial_\mu\left(C_1 f^{\mu\nu}+C_2\tilde{f}^{\mu\nu}\right)f_{\nu\alpha}-\frac{1}{2}\partial_\mu\left(\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\right)f_{\nu\alpha}-$$
$$\partial_\mu\left(\hat{R}^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma}\right)f_{\nu\alpha}-\partial_\mu\left(\hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\right)f_{\nu\alpha}=-\mathcal{J}^\nu f_{\nu\alpha}\ . \tag{3.40}$$

We then need to:

- Apply the Ricci Curbastro-Bianchi identity $\partial_\mu f_{\nu\alpha}=-\partial_\nu f_{\alpha\mu}-\partial_\alpha f_{\mu\nu}$ whenever a term of the form $\partial_\mu f_{\nu\alpha}$ appears in our manipulations.
- Make use of the identity $\tilde{f}^{\mu\nu}\partial_\alpha f_{\mu\nu}=\frac{1}{2}\partial_\alpha(\tilde{f}^{\mu\nu}f_{\mu\nu})$.

which will bring the equations of motion into the following form

$$2\partial_\mu\left(C_1 f^{\mu\nu}f_{\nu\alpha}+C_2\tilde{f}^{\mu\nu}f_{\nu\alpha}\right)+\underbrace{\partial_\mu\left(\hat{Q}_B^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\right)f_{\nu\alpha}}_{\mathcal{A}}-\underbrace{\frac{3}{2}\partial_\mu\left(\hat{R}^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma}\right)f_{\nu\alpha}}_{\mathcal{B}}-$$
$$\underbrace{\partial_\mu\left(\hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\right)f_{\nu\alpha}}_{\mathcal{C}}+\partial_\alpha(\frac{1}{2}C_1 f^{\mu\nu}f_{\mu\nu})+\partial_\alpha(\frac{1}{2}\tilde{f}^{\mu\nu}f_{\mu\nu})-\frac{1}{2}(\partial_\alpha C_2)\tilde{f}^{\mu\nu}f_{\mu\nu}=-2\mathcal{J}^\nu f_{\nu\alpha}\ . \tag{3.41}$$

The terms $\mathcal{A},\mathcal{B},\mathcal{C}$ must be expressed as a continuity equations, which we will now do.

For $\mathcal{A}$

$$\partial_\mu\left(\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\right)f_{\nu\alpha}=\partial_\mu\left(\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}f_{\nu\alpha}\right)-\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\left(\partial_\mu f_{\nu\alpha}\right)\ , \tag{3.42}$$

the second term on the right hand side will be decomposed as

$$\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\left(\partial_\mu f_{\nu\alpha}\right)=\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}(\underbrace{\partial_\mu\partial_\nu a_\alpha}_{0}-\partial_\mu\partial_\alpha a_\nu)$$
$$\Rightarrow\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\left(\partial_\mu f_{\nu\alpha}\right)=-\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\partial_\mu\partial_\alpha a_\nu=-\frac{1}{2}\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\partial_\alpha f_{\mu\nu}\ , \tag{3.43}$$

which will give for $\mathcal{A}$

$$\partial_\mu\left(\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}f_{\nu\alpha}\right)+\frac{1}{2}\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\partial_\alpha f_{\mu\nu}\ . \tag{3.44}$$

By virtue of the symmetry of $\hat{Q}$ under the pair exchanges, we may recast the second term in $\mathcal{A}$ as

$$\frac{1}{4}\partial_\alpha\left(\hat{Q}^{\mu\nu\kappa\lambda}f_{\mu\nu}f_{\kappa\lambda}\right)-\frac{1}{4}\left(\partial_\alpha\hat{Q}^{\mu\nu\kappa\lambda}\right)f_{\mu\nu}f_{\kappa\lambda}\ . \tag{3.45}$$

Finally, $\mathcal{A}$ will read

$$\partial_\mu\left(\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}\right)f_{\nu\alpha}=\partial_\mu\left(\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}f_{\nu\alpha}\right)+\frac{1}{4}\partial_\alpha\left(\hat{Q}^{\mu\nu\kappa\lambda}f_{\mu\nu}f_{\kappa\lambda}\right)-\frac{1}{4}\left(\partial_\alpha\hat{Q}^{\mu\nu\kappa\lambda}\right)f_{\mu\nu}f_{\kappa\lambda}\ . \tag{3.46}$$

Following similar steps (and keeping in mind the symmetry under pair exchanges), $\mathcal{B}$ can be expressed as

$$\begin{aligned}\partial_\mu\left(\hat{R}^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma}\right)f_{\nu\alpha}=&\partial_\mu\left(\hat{R}^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma}f_{\nu\alpha}\right)+\frac{1}{6}\partial_\alpha\left(\hat{R}^{\mu\nu\kappa\lambda\rho\sigma}f_{\mu\nu}f_{\kappa\lambda}f_{\rho\sigma}\right)-\\&\frac{1}{6}\left(\partial_\alpha\hat{R}^{\mu\nu\kappa\lambda\rho\sigma}\right)f_{\mu\nu}f_{\kappa\lambda}f_{\rho\sigma}\ ,\end{aligned} \tag{3.47}$$

and $\mathcal{C}$ as

$$\begin{aligned}\partial_\mu\left(\hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\right)f_{\nu\alpha}=&\partial_\mu\left(\hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\nu\alpha}f_{\omega\tau}\right)+\frac{1}{8}\partial_\alpha\left(\hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\mu\nu}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\right)\\&-\frac{1}{8}\left(\partial_\alpha\hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}\right)f_{\mu\nu}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\ .\end{aligned} \tag{3.48}$$

Now that all of the terms have been put under the form of a continuity equation thanks to the manipulation in Eqs.(3.42)-(3.48), it is possible to write down ${\theta^\mu}_\alpha$ and its conservation equation

$$\partial_\mu{\theta^\mu}_\alpha=\mathcal{D}_\alpha\ . \tag{3.49}$$

The gauge invariant energy-momentum tensor for the photon will then read at the fourth order, and in the most general case

$$\begin{aligned}{\theta^\mu}_\alpha=&C_1f^{\mu\nu}f_{\nu\alpha}-\frac{1}{2}\hat{Q}^{\mu\nu\kappa\lambda}f_{\kappa\lambda}f_{\nu\alpha}-\frac{3}{4}\hat{R}^{\mu\nu\kappa\lambda\rho\sigma}f_{\kappa\lambda}f_{\rho\sigma}f_{\nu\alpha}-\frac{1}{2}\hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}f_{\nu\alpha}\\&+{\delta^\mu}_\alpha\left(\frac{1}{4}C_1f^{\beta\nu}f_{\beta\nu}-\frac{1}{8}\hat{Q}^{\beta\nu\kappa\lambda}f_{\beta\nu}f_{\kappa\lambda}-\frac{1}{8}\hat{R}^{\beta\nu\kappa\lambda\rho\sigma}f_{\beta\nu}f_{\kappa\lambda}f_{\rho\sigma}-\frac{1}{16}\hat{S}^{\beta\nu\kappa\lambda\rho\sigma\omega\tau}f_{\beta\nu}f_{\kappa\lambda}f_{\rho\sigma}f_{\omega\tau}\right)\ ,\end{aligned} \tag{3.50}$$

where the terms in C_2 cancelled each after other using $\tilde{f}^{\mu\nu}f_{\nu\alpha}=-\frac{1}{4}\delta^\mu_\alpha\tilde{f}\cdot f$. This is absolutely expected, since the terms in C_2 are topological terms and cannot appear in the energy-momentum tensor.

The divergence of the EMT reads

$$\begin{aligned}\mathcal{D}_\alpha =& -\mathcal{J}^\nu f_{\nu\alpha} + \frac{1}{4}(\partial_\alpha C_1) f^2 + \frac{1}{4}(\partial_\alpha C_2)\tilde{f}\cdot f - \frac{1}{8}\left(\partial_\alpha \hat{Q}^{\mu\nu\kappa\lambda}\right) f_{\mu\nu} f_{\kappa\lambda} \\ & - \frac{1}{8}\left(\partial_\alpha \hat{R}^{\mu\nu\kappa\lambda\rho\sigma}\right) f_{\mu\nu} f_{\kappa\lambda} f_{\rho\sigma} - \frac{1}{16}\left(\partial_\alpha \hat{S}^{\mu\nu\kappa\lambda\rho\sigma\omega\tau}\right) f_{\mu\nu} f_{\kappa\lambda} f_{\rho\sigma} f_{\omega\tau} \ , \end{aligned} \tag{3.51}$$

which shows that in order to obtain frequency shifts in this model, one should consider non-constant background fields. Working with a fourth order non-linear theory is however extremely daunting, and since most higher order effects are expected to be suppressed at lower energy scales, we will turn our attention to the linearised (or second order) action[7]. This will read

$$\mathcal{L}_2 = -\frac{1}{4}C_1 f^{\mu\nu} f_{\mu\nu} - \frac{1}{4}C_2 f^{\mu\nu}\tilde{f}_{\mu\nu} - \frac{1}{2}\left(C_1 F^{\mu\nu} + C_2 \tilde{F}^{\mu\nu}\right) f_{\mu\nu} + \frac{1}{8} Q^{\mu\nu\kappa\lambda} f_{\mu\nu} f_{\kappa\lambda} \ , \tag{3.52}$$

where we denoted $\mathcal{L}_2$ to signify that the action contains only terms that are up to the second order in f as stated above. This simplification will in turn lead to the following equations of motion

$$\partial_\mu\left(C_1 f^{\mu\nu} + C_2 \tilde{f}^{\mu\nu}\right) - \frac{1}{2}\partial_\mu\left(Q^{\mu\nu\rho\sigma} f_{\rho\sigma}\right) = -\mathcal{J}^\nu \ . \tag{3.53}$$

Something that deserves to be highlighted is that the degree of non-linearity for photon interactions in this theory depend explicitly on the non-linearity of the background field. Indeed, in the case where $C_1 = 1$ the theory will be reduced exactly to the Maxwellian case which we described above. And since the next higher order coefficient D_1 can be written as

$$D_1 = \left.\frac{\partial^2 \mathcal{L}}{\partial \mathcal{F}^2}\right|_{\mathrm{B}} = \frac{\partial}{\partial \mathcal{F}} C_1 \ , \tag{3.54}$$

any higher order contributions will be trivially zero. Roughly speaking, this means that on a Maxwellian background, we can only expect a Maxwellian photon, and as such, to be at liberty to discuss truly non-linear behaviours for photons, one would have to consider non-linear background fields.

In [97], we briefly discuss the Poincaré violating aspect of this theory. Indeed, the fact that the energy-momentum tensor is not a conserved quantity anymore hints towards a breaking down of translational symmetry. While a quick computation of the canonical angular momentum tensor

$$\mathcal{M}^{\alpha\mu\nu} = i\left(x^\mu T^{\alpha\nu} - x^\nu T^{\alpha\mu}\right) - \mathcal{S}^{\alpha\mu\nu} \ , \tag{3.55}$$

[7] By linearised, we mean to keep in the action all the terms that are at most quadratic in f

where T is the canonical energy-momentum tensor, and

$$\begin{aligned}\mathcal{S}^{\alpha\mu\nu} =& \mathcal{J}^{\alpha\mu}a^{\nu} - \mathcal{J}^{\alpha\nu}a^{\mu} + C_1(f^{\alpha\mu}a^{\nu} - f^{\alpha\nu}a^{\mu}) + C_2(\tilde{f}^{\alpha\mu}a^{\nu} - \tilde{f}^{\alpha\nu}a^{\mu}) \\ &+ \frac{1}{2}\left(Q^{\alpha\nu\rho\sigma}f_{\rho\sigma}a^{\mu} - Q^{\alpha\mu\rho\sigma}f_{\rho\sigma}a^{\nu}\right) \end{aligned} \tag{3.56}$$

is the spin density tensor, with $\mathcal{J}^{\alpha\mu} = \left(C_1 F_{\mathrm{B}}^{\alpha\mu} + C_2\tilde{F}_{\mathrm{B}}^{\alpha\mu}\right)$, shows that it is not conserved either by virtue of the non-constance of the background field. This implies a violation of rotational symmetry, and as such a full breakdown of the Poincaré group.

The fact that we are dealing with a Poincaré symmetry violating model brings into question multiple issues, but one in particular: the Casimir operators[8] of the group. Indeed, the Poincaré group possesses two invariant operators, which are associated to the mass and the helicity/spin of the states that carry its unitary irreducible representations. In the case described above, it is safe to assume that at least one of these quantities is no longer an invariant, thereby putting into question the formal definition of mass and/or spin. Could a Poincaré symmetry violation lead to the emergence of non-constant masses? What would a variable spin imply physically? How would this fundamentally affect the structure of space-time?

And while we have not found the answer to most of these questions, in the next chapter we will show that within this framework it is absolutely possible for photons to obtain effective masses [101], but before adressing this, let us turn our attention to the propagation features of the theory, most notably its dispersion relations.

[8]The two Casimir operators [14] are $P^2 = P_\mu P^\mu$ and $W^2 = W_\mu W^\mu$, where P is momentum, and W is the Pauli-Lubanski pseudo-vector. In the rest frame, we get $P^2 = mc^2$, that is mass for the first Casimir operator. Instead $W^2 = P_\mu P^\mu$. has to do with spin for massive particles and helicity for massless particles.

3.3 Dispersion relations and group velocities

In what follows, we shall derive the dispersion relations for the theory at second order, considering various cases. Instead of only offering to the reader the most general one, we shall progressively generalise our system so that the general expression can be expressed in terms of intermediate quantities. The reason why we chose this approach will be made clear with the final expression, but the starting point will thus be the linearised field equations (3.53) expressed in terms of the photon fields $\{\vec{e}, \vec{b}\}$ fields and the background fields $\{\vec{E}, \vec{B}\}$, which will give for Gauss's law

$$\begin{aligned}&\vec{\nabla}\cdot\vec{e}+\frac{1}{C_1}\left(\frac{D_1}{c^2}\vec{E}+\frac{D_3}{c}\vec{B}\right)\cdot\vec{\nabla}\left(\vec{E}\cdot\vec{e}-c^2\vec{B}\cdot\vec{b}\right)+\\&\frac{1}{C_1}\left(D_2\vec{B}+\frac{D_3}{c}\vec{E}\right)\cdot\vec{\nabla}\left(\vec{B}\cdot\vec{e}+\vec{E}\cdot\vec{b}\right)=-\vec{\nabla}\cdot\vec{E}\end{aligned} \tag{3.57}$$

and for the Ampère-Maxwell law

$$\begin{aligned}&\vec{\nabla}\times\vec{b}+\vec{\nabla}\times\vec{B}+\frac{1}{C_1}\vec{B}\times\vec{\nabla}\left(D_1\vec{B}\cdot\vec{b}-D_3\vec{E}\cdot\vec{b}-\frac{D_1}{c^2}\vec{E}\cdot\vec{e}-\frac{D_3}{c}\vec{B}\cdot\vec{e}\right)\\&+\frac{1}{C_1}\vec{E}\times\vec{\nabla}\left(-\frac{D_3}{c}\vec{B}\cdot\vec{b}+\frac{D_2}{c^2}\vec{E}\cdot\vec{b}+\frac{D_2}{c^2}\vec{B}\cdot\vec{e}+\frac{D_3}{c}\vec{E}\cdot\vec{e}\right)-\frac{1}{c^2}\partial_t\vec{e}\\&-\frac{1}{c^2C_1}\left(\frac{D_1}{c^2}\vec{E}+\frac{D_3}{c}\vec{B}\right)\partial_t\left(\vec{E}\cdot\vec{e}-c^2\vec{B}\cdot\vec{b}\right)-\frac{1}{c^2C_1}\left(D_2\vec{B}+\frac{\vec{D_3}}{c}\vec{E}\right)\partial_t\left(\vec{B}\cdot\vec{e}+\vec{E}\cdot\vec{b}\right)=0\ ,\end{aligned} \tag{3.58}$$

where we have restored c, μ_0, ϵ_0. The equations related to the Ricci Curbastro-Bianchi identity are of course exactly the same. We immediately notice how daunting the expression for the Ampère-Maxwell equation is, and may assume the difficulties related to finding the corresponding dispersion relation and group velocities.

3.3.1 $\mathcal{L}(\mathcal{F})$ with constant $\vec{E}$ and $\vec{B}$

Let us consider first a simpler case, where the action $\mathcal{L}(\mathcal{F})$, *i.e.* no $\mathcal{G}$ dependence, which automatically sets $C_2 = D_2 = D_3 = 0$. Additionally, we will consider a constant background with both $\vec{E}$ and $\vec{B}$ fields present. This simplifies our equations into the

following form

$$\vec{\nabla}\cdot\vec{e}+\frac{D_1}{C1}\vec{E}\cdot\vec{\nabla}\left(\frac{1}{c^2}\vec{E}\cdot\vec{e}-\vec{B}\cdot\vec{b}\right)=0 \tag{3.59}$$

$$\vec{\nabla}\times\vec{b}+\frac{D_1}{C_1}\vec{B}\times\vec{\nabla}\left(\vec{B}\cdot\vec{b}-\frac{1}{c^2}\vec{E}\cdot\vec{e}\right)-\frac{1}{c^2}\partial_t\vec{e}-\frac{1}{c^2}\frac{D_1}{C_1}\vec{E}\partial_t\left(\frac{1}{c^2}\vec{E}\cdot\vec{e}-\vec{B}\cdot\vec{b}\right)=0\ , \tag{3.60}$$

where $\frac{D_1}{C_1}$ has the dimension $[T]^{-2}$. By considering the plane wave decomposition for the photon fields

$$\vec{e}=\vec{e}_0\ e^{i(\vec{k}\cdot\vec{x}-\omega t)}\ , \tag{3.61}$$

$$\vec{b}=\vec{b}_0\ e^{i(\vec{k}\cdot\vec{x}-\omega t)}\ , \tag{3.62}$$

we find the dispersion relation to be given by

$$\det\left[\left(\frac{\omega^2}{c^2}-k^2\right)\delta_{ij}+k_ik_j+\frac{D_1}{C_1}\Omega_i\Omega_j\right]=0\ , \tag{3.63}$$

where we define Ω_i as

$$\vec{\Omega}=\frac{\omega}{c^2}\vec{E}+\vec{k}\times\vec{B}\ . \tag{3.64}$$

Then, using the fact that, for a general 3×3 matrix,

$$\det M=\frac{1}{6}(\mathrm{trM})^3-\frac{1}{2}(\mathrm{trM}^2)\mathrm{trM}+\frac{1}{3}\mathrm{trM}^3\ , \tag{3.65}$$

for an M_{ij} of the form

$$M_{ij}=a\delta_{ij}+bu_iu_j+cv_iv_j\ , \tag{3.66}$$

then the determinant will read

$$\det M=a\left[(a+bu^2)(a+cv^2)-bc(\vec{u}\cdot\vec{v})^2\right]\ . \tag{3.67}$$

In our case,

$$a = \frac{\omega^2}{c^2} - k^2 \ , \tag{3.68}$$

$$b = 1 \ , \tag{3.69}$$

$$c = \frac{D_1}{C_1} \ , \tag{3.70}$$

giving for the dispersion relation

$$\left(\frac{\omega^2}{c^2} - k^2\right)\omega^2 \left[\left(1 + \frac{D_1}{c^2 C_1}E^2\right)\frac{\omega^2}{c^2} + 2\frac{D_1}{c\ C_1}\vec{E}\cdot(\vec{k}\times\vec{B})\frac{\omega}{c} + \left(-k^2 + \frac{D_1}{C_1}(\vec{k}\times\vec{B})^2 - \frac{D_1}{c^2 C_1}(\vec{k}\cdot\vec{E})^2\right)\right] = 0 \ , \tag{3.71}$$

which we recast as

$$\omega^2\left(\frac{\omega^2}{c^2} - k^2\right)\left(\alpha\frac{\omega^2}{c^2} + \beta\frac{\omega}{c} + \gamma\right) = 0 \ , \tag{3.72}$$

with

$$\alpha \equiv 1 + \frac{D_1}{C_1}\frac{E^2}{c^2} \ , \tag{3.73}$$

$$\beta \equiv 2\frac{D_1}{C_1}\frac{\vec{E}}{c}\cdot(\vec{k}\times\vec{B}) \ , \tag{3.74}$$

$$\gamma \equiv -k^2 - \frac{D_1}{c^2 C_1}(\vec{k}\cdot\vec{E})^2 + \frac{D_1}{C_1}(\vec{k}\times\vec{B})^2 \ . \tag{3.75}$$

The solutions to the dispersion relation are as follows:

$$\omega = 0 \ , \tag{3.76}$$

which is absolutely meaningless, but also

$$\omega^2 = c^2 k^2 \ , \tag{3.77}$$

which stems from the gauge invariance of the theory, and the roots to the equation

$$\alpha\left(\frac{\omega}{c}\right)^2 + \beta\frac{\omega}{c} + \gamma = 0 \ , \tag{3.78}$$

which will in turn read

$$\omega = \frac{c}{2\alpha}\left(-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}\right) \ , \tag{3.79}$$

$$\omega = \frac{c}{1+\frac{D_1}{C_1}\frac{E^2}{c^2}}\left[\frac{D_1}{C_1}\frac{\vec{E}}{c}\cdot(\vec{k}\times\vec{B}) \pm \sqrt{\frac{D_1^2}{C_1^2}\frac{\left(\vec{E}\cdot(\vec{k}\times\vec{B})\right)^2}{c^2} - \left(1+\frac{D_1}{C_1}\frac{E^2}{c^2}\right)\left(-k^2 - \frac{D_1}{c^2 C_1}(\vec{k}\cdot\vec{E})^2 + \frac{D_1}{C_1}(\vec{k}\times\vec{B})^2\right)}\right] \ , \tag{3.80}$$

where the non-linearity of the theory is well apparent as the plus or minus sign highlights clearly the presence of vacuum birefringence. However, the presence of four roots, would imply the propagation of four degrees of freedom, which would be in clear contradiction with the structure of the theory. By looking closely, we can see that the usual Maxwell dispersion relation, is not totally independent from the birefringent root we obtained, and that we can reduce the latter to the former by setting the background fields to zero (which will reduce the coefficients to $\alpha = 1, \beta = 0, \gamma = -k^2$). This implies that the $\omega^2 = c^2k^2$, is a gauge artifact, while the true propagation features of the theory lie within the birefringent root.

We now have to differentiate equation (3.78) with respect to k_i to obtain the group velocity, and it will read

$$\frac{\vec{v_g}}{c} = \frac{1}{\alpha\omega + \frac{\beta}{2}c}\left[(1 - \frac{D_1}{C_1}B^2)c\vec{k} + \frac{D_1}{c\,C_1}(\vec{E}\cdot\vec{k})\vec{E} + c\frac{D_1}{C_1}(\vec{B}\cdot\vec{k})\vec{B} + \frac{D_1}{c\,C_1}\omega(\vec{E}\times\vec{B})\right] \ , \tag{3.81}$$

where ω are the roots we derived above, and we retrieve the Maxwellian limit for $\alpha = 1,\ \beta = 0,\ C_1 = 1,\ D_1 = 0$ and $\omega = ck$.

3.3.2 $\mathcal{L}(\mathcal{F}, \mathcal{G})$ with a purely magnetic field

Let us now consider the case where the action is a function of both invariants, but in the absence of background electric fields. This case is rather relevant for galactic systems for example, as we expect the quasi neutrality hypothesis to hold at these scales.

The set of equations we will work with will then become

$$\vec{\nabla}\cdot\vec{e}-c\frac{D_3}{C_1}\vec{B}\cdot\vec{\nabla}(\vec{B}\cdot\vec{b})+\frac{D_2}{C_1}\vec{B}\cdot\vec{\nabla}(\vec{B}\cdot\vec{e})=0, \tag{3.82}$$

$$\vec{\nabla}\times\vec{b}+\frac{1}{C_1}\vec{B}\times\vec{\nabla}\left(D_1\vec{B}\cdot\vec{b}-\frac{D_3}{c}\vec{B}\cdot\vec{e}\right)-\frac{1}{c^2}\partial_t\vec{e}+\frac{D_3}{c\,C_1}\vec{B}\partial_t(\vec{B}\cdot\vec{b})-\frac{D_2}{c^2C_1}\vec{B}\partial_t(\vec{B}\cdot\vec{e})=0\ . \tag{3.83}$$

By following the same procedure as before and using the plane wave decomposition, we show that the dispersion relation follows from

$$M_{ij}e_{0j}=0\ , \tag{3.84}$$

where the M matrix has the form

$$M_{ij}=\alpha\delta_{ij}+u_iv_j+w_it_j\ , \tag{3.85}$$

with the components defined as

$$\alpha\equiv\frac{\omega^2}{c^2}-k^2, \tag{3.86}$$

$$u_i\equiv\frac{D_1}{C_1}(\vec{B}\times\vec{k})_i+\frac{D_3}{c\,C_1}\omega B_i+c\frac{D_3}{C_1}(\vec{B}\cdot\vec{k})\frac{k_i}{c}\ , \tag{3.87}$$

$$v_j\equiv(\vec{B}\times\vec{k})_j, \tag{3.88}$$

$$w_i\equiv\frac{D_2}{c^2C_1}\omega^2B_i-\frac{D_3}{c\,C_1}\omega(\vec{B}\times\vec{k})_i-\frac{D_2}{C_1}(\vec{B}\cdot\vec{k})k_i \tag{3.89}$$

$$tj\equiv B_j\ . \tag{3.90}$$

For a 3×3 matrix of the form

$$M_{ij}=\alpha\delta_{ij}+\beta u_iv_j+\gamma w_it_j\ , \tag{3.91}$$

the determinant is given by

$$\det M=\alpha\left[(\alpha+\beta\ \vec{u}\cdot\vec{v})(\alpha+\gamma\ \vec{w}\cdot\vec{t})-\beta\gamma(\vec{u}\cdot\vec{t})(\vec{v}\cdot\vec{w})\right]\ . \tag{3.92}$$

Let us now evaluate each term for $\beta = \gamma = 1$.

$$\vec{u}\cdot\vec{v} = \frac{D_1}{C_1}\left((\vec{B}\times\vec{k})^2 = \frac{D_1}{C_1}(B^2k^2 - (\vec{B}\cdot\vec{k})^2\right) \tag{3.93}$$

$$\vec{w}\cdot\vec{t} = \frac{D_2}{C_1}\frac{\omega^2}{c^2}B^2 - \frac{D_2}{C_1}(\vec{B}\cdot\vec{k})^2 \tag{3.94}$$

$$\vec{u}\cdot\vec{t} = \frac{D_3}{C_1}\frac{\omega}{c}B^2 + \frac{c}{\omega}\frac{D_3}{C_1}(\vec{B}\cdot\vec{k})^2 \tag{3.95}$$

$$\vec{v}\cdot\vec{w} = -\frac{D_3}{C_1}\frac{\omega}{c}(\vec{B}\times\vec{k})^2 = -\frac{D_3}{C_1}\frac{\omega}{c}\left(B^2k^2 - (\vec{B}\cdot\vec{k})^2\right)\ . \tag{3.96}$$

$\det M = 0$ results in two branches for the dispersion relations:

$$\alpha = 0 \Rightarrow \omega^2 = c^2k^2\ , \tag{3.97}$$

which are respectively, meaningless, and a gauge artifact; and the roots of

$$P\frac{\omega^4}{c^4} + Q\frac{\omega^2}{c^2} + R = 0\ , \tag{3.98}$$

$$\omega = \pm c\sqrt{\frac{-Q \pm \sqrt{Q^2 - 4PR}}{2P}}\ , \tag{3.99}$$

where we defined

$$P \equiv 1 + \frac{D_2}{C_1}B^2, \tag{3.100}$$

$$Q \equiv -2k^2 + \frac{D_1 - D_2}{C_1}B^2k^2 - \frac{D_1 + D_2}{C_1}(\vec{B}\cdot\vec{k})^2 + \frac{D_1D_2 + {D_3}^2}{{C_1}^2}B^2(\vec{B}\times\vec{k})^2\ , \tag{3.101}$$

$$R \equiv k^4 - \frac{D_1}{C_1}k^2(\vec{B}\times\vec{k})^2 + \frac{D_2}{C_1}k^2(\vec{B}\cdot\vec{k})^2 + \frac{{D_3}^2 - D_1D_2}{{C_1}^2}(\vec{B}\cdot\vec{k})^2(\vec{B}\times\vec{k})^2\ . \tag{3.102}$$

We may now proceed with the group velocities by differentiating (3.98) with respect to $\vec{k}$, paying attention to the dependences of Q and R. We will use

$$d(\vec{B}\cdot\vec{k})^2 = 2(\vec{B}\cdot\vec{k})\vec{B}\cdot d\vec{k}\ , \tag{3.103}$$

$$d(\vec{B}\times\vec{k})^2 = 2B^2\vec{k}\cdot d\vec{k} - 2(\vec{B}\cdot\vec{k})\vec{B}\cdot d\vec{k}\ , \tag{3.104}$$

$$d[k^2(\vec{B}\times\vec{k})^2] = 2(\vec{B}\times\vec{k})^2\vec{k}\cdot d\vec{k} + 2B^2k^2\vec{k}\cdot d\vec{k} - 2k^2(\vec{B}\cdot\vec{k})\vec{B}\cdot d\vec{k}\ , \tag{3.105}$$

$$d[(\vec{B}\cdot\vec{k})^2(\vec{B}\times\vec{k})^2] = 2(\vec{B}\cdot\vec{k})(\vec{B}\times\vec{k})^2\vec{B}\cdot d\vec{k} - 2(\vec{B}\cdot\vec{k})^3\vec{B}\cdot d\vec{k} + 2B^2(\vec{B}\cdot\vec{k})^2\vec{k}\cdot d\vec{k}\ , \tag{3.106}$$

which then gives for the group velocities

$$\begin{aligned}\frac{\vec{v}_g}{c} =& \frac{1}{2P\frac{\omega^2}{c^2}+Q}\vec{k}\left[\frac{\omega}{c}\left(2-\frac{D_1-D_2}{C_1}B^2-\frac{D_1D_2+D_3^2}{C_1^2}B^4\right)+\frac{c}{\omega}\left(-2k^2+2\frac{D_1}{C_1}B^2k^2-\frac{D_1+D_2}{C_1}(\vec{B}\cdot\vec{k})^2\right.\right.\\ &\left.\left.-\frac{D_3^2-D_1D_2}{C_1^2}B^2(\vec{B}\cdot\vec{k})^3\right)\right]+(\vec{B}\cdot\vec{k})\vec{B}\left[\frac{\omega}{c}\left(\frac{D_1+D_2}{C_1}+\frac{D_3^2+D_1D_2}{C_1^2}B^2\right)+\frac{c}{\omega}\left(-\frac{D_1+D_2}{C_1}k^2\right.\right.\\ &\left.\left.-\frac{D_3^2-D_1D_2}{C_1^2}B^2k^2+2\frac{D_3^2-D_1D_2}{C_1^2}(\vec{B}\cdot\vec{k})^2\right)\right]\ . \end{aligned} \tag{3.107}$$

The expression is rather barbaric, and still far from the most general case. Before stepping to the latter, we shall offer a dictionary to translate the system from constant magnetic field to constant electric field (and vice versa). Indeed by performing

$$D_1 \leftrightarrow D_2\ ,\quad \vec{B}\leftrightarrow\frac{\vec{E}}{c}\ ,\quad D_3\leftrightarrow\pm D_3\ , \tag{3.108}$$

we should be able to move from one to the other. Additionally, since in the dispersion relations D_3 only appears as squared, we need not worry about a minus sign. We do not need to recompute any of the quantities above, and can simply replace within the expressions using the dictionary.

3.3.3 $\mathcal{L}(\mathcal{F},\mathcal{G})$ with constant $\{\vec{E},\vec{B}\}$

We will now turn our attention to the most general case[9], but before partaking in the painful computations ahead, let us attempt to write down the Ampère-Maxwell equation in a nicer form. To do so, we shall define the following pure-background vectors

$$\begin{aligned}\vec{\chi}\equiv\frac{D_1}{c^2C_1}\vec{E}+\frac{D_3}{c^2C_1}\vec{B}\quad &,\quad \vec{\lambda}\equiv\frac{D_2}{C_1}\vec{B}+\frac{D_3}{c\,C_1}\ ,\\ \vec{\xi}\equiv-\frac{D_1}{c^2C_1}\vec{B}+\frac{D_3}{c^3C_1}\vec{E}\quad &,\quad \vec{\mu}\equiv\frac{D_2}{c^2C_1}\vec{E}-\frac{D_3}{cC_1}\vec{B}\ ,\end{aligned} \tag{3.109}$$

which then brings the Gauss and Ampère-Maxwell laws into the following form

$$\vec{\nabla}\cdot\vec{e}+\vec{\chi}\cdot\vec{\nabla}\left(\vec{E}\cdot\vec{e}-c^2\vec{B}\cdot\vec{b}\right)+\vec{\lambda}\vec{\nabla}\left(\vec{B}\cdot\vec{e}+\vec{E}\cdot\vec{b}\right)=0 \tag{3.110}$$

$$\begin{aligned}&\vec{\nabla}\times\vec{b}+\vec{\xi}\times\vec{\nabla}\left(\vec{E}\cdot\vec{e}-c^2\vec{B}\cdot\vec{b}\right)+\vec{\mu}\times\vec{\nabla}\left(\vec{B}\cdot\vec{e}+\vec{E}\cdot\vec{b}\right)\\ &-\frac{1}{c^2}\partial_t\vec{e}-\vec{\chi}\frac{1}{c^2}\partial_t\left(\vec{E}\cdot\vec{e}-c^2\vec{B}\cdot\vec{b}\right)-\vec{\lambda}\frac{1}{c^2}\partial_t\left(\vec{B}\cdot\vec{e}+\vec{E}\cdot\vec{b}\right)=0\ ,\end{aligned} \tag{3.111}$$

where we additionally notice that $\vec{\chi}\to\vec{\xi}$ and $\vec{\lambda}\to\vec{\mu}$ whenever we perform the replacements $\vec{E}\to-\vec{B}$ and $\vec{B}\to\frac{1}{c^2}\vec{E}$.

[9] Without considering variable backgrounds, we will discuss this later on.

From equations (3.110) and (3.111) by considering the plane wave expansion (3.61) and recalling that

$$\vec{b}_0 = \frac{\vec{k}}{\omega} \times \vec{e}_0 \ , \tag{3.112}$$

we show that the Gauss law becomes

$$\vec{k} \cdot \vec{e} = -\left(\vec{\chi} \cdot \vec{k}\ \vec{E} + \vec{\lambda} \cdot \vec{k}\ \vec{B}\right) \cdot \vec{e}_0 + \left(c^2 \vec{\chi} \cdot \vec{k}\ \vec{B} - \vec{\lambda} \cdot \vec{k}\ \vec{E}\right) \cdot \vec{b}_0 \tag{3.113}$$

and to simplify it further we define

$$\vec{\alpha} \equiv \vec{\chi} \cdot \vec{k}\ \vec{E} + \vec{\lambda} \cdot \vec{k}\ \vec{B} \ , \tag{3.114}$$

$$\vec{\beta} \equiv c^2 \vec{\chi} \cdot \vec{k}\ \vec{B} - \vec{\lambda} \cdot \vec{k}\ \vec{E} \ , \tag{3.115}$$

so that

$$\vec{k} \cdot \vec{e}_0 = -\vec{\alpha} \cdot \vec{e}_0 + \vec{\beta} \cdot \vec{b_0} \ , \tag{3.116}$$

where we apply the Faraday-Lenz law (3.112) on the second term to get

$$\vec{k} \cdot \vec{e}_0 = \left[-\vec{\alpha} + \left(\vec{\beta} \times \frac{\vec{k}}{\omega}\right)\right] \cdot \vec{e}_0 \ . \tag{3.117}$$

We now turn our attention to the Ampère-Maxwell law, which will read

$$\begin{aligned} &\vec{k} \times \vec{b}_0 + \frac{\omega}{c^2}\vec{e}_0 + \left(\vec{\xi} \times \vec{k} + \frac{\omega}{c^2}\vec{\chi}\right) \vec{E} \cdot \vec{e}_0 + \left(\vec{\mu} \times \vec{k} + \frac{\omega}{c}\vec{\lambda}\right) \vec{B} \cdot \vec{e}_0 \\ &- c^2 \left(\vec{\xi} \times \vec{k} + \frac{\omega}{c^2}\vec{\chi}\right) \vec{B} \cdot \vec{b}_0 + \left(\vec{\mu} \times \vec{k} + \frac{\omega}{c}\vec{\lambda}\right) \vec{E} \cdot \vec{e}_0 = 0 \ , \end{aligned} \tag{3.118}$$

where similarly to the previous computation we will define intermediary quantities to simplify its expression. As such, by setting[10]

$$\vec{\gamma} \equiv \vec{\xi} \times \vec{k} + \frac{\omega}{c^2}\vec{\chi} \ , \tag{3.119}$$

$$\vec{\delta} \equiv \vec{\mu} \times \vec{k} + \frac{\omega}{c}\vec{\lambda} \ , \tag{3.120}$$

[10]the delta vector must not be confused with a Kronecker symbol. The Kronecker symbol will always carry two indices.

the equation takes the much nicer form

$$\vec{k} \times \vec{b}_0 + \frac{\omega}{c^2}\vec{e}_0 + \vec{\gamma}\vec{E} \cdot \vec{e}_0 + \vec{\delta}\vec{B} \cdot \vec{e}_0 - c^2\vec{\gamma}\vec{B} \cdot \vec{b}_0 + \vec{\delta}\vec{E} \cdot \vec{b}_0 = 0 \ , \tag{3.121}$$

the next step consists on applying the Lenz Faraday law on every term involving $\vec{b}_0$. For the first one we get

$$\vec{k} \times \vec{b}_0 = (\vec{k} \cdot \vec{e}_0)\frac{\vec{k}}{\omega} - \frac{k^2}{\omega}\vec{e}_0 \ , \tag{3.122}$$

where we then use the Gauss law (3.117) on the first term to get

$$\vec{k} \times \vec{b}_0 = \left(\frac{\vec{k}}{\omega}\left(\vec{\beta} \times \frac{\vec{k}}{\omega}\right) - \frac{k^2}{\omega} - \frac{\vec{k}}{\omega}\vec{\alpha} \cdot \vec{e}\right) \cdot \vec{e}_0 \ , \tag{3.123}$$

as for the remaining two terms in $\vec{b}_0$

$$\vec{B} \cdot \vec{b}_0 = (\vec{B} \times \frac{\vec{k}}{\omega}) \cdot \vec{e}_0 \ , \tag{3.124}$$

$$\vec{E} \cdot \vec{b}_0 = (\vec{E} \times \frac{\vec{k}}{\omega}) \cdot \vec{e}_0 \ . \tag{3.125}$$

Replacing in (3.121) and multiplying by ω

$$\left(\frac{\omega^2}{c} - k^2\right)\vec{e}_0 - \vec{k}\ \vec{\alpha} \cdot \vec{e}_0 + \frac{\vec{k}}{\omega}\left(\vec{\beta} \times \vec{k}\right) \cdot \vec{e}_0 + \omega\vec{\gamma}\ \vec{E} \cdot \vec{e}_0 + \\ \omega\vec{\delta}\ \vec{B} \cdot \vec{e}_0 - c^2\vec{\gamma}(\vec{B} \times \vec{k}) \cdot \vec{e}_0 + \vec{\delta}(\vec{E} \times \vec{k}) \cdot \vec{e}_0 = 0 \ , \tag{3.126}$$

which we will then express in index notation

$$\left(\frac{\omega^2}{c^2} - k^2\right)\delta_{ij}e_{0j} + k_i\left(-\vec{\alpha} + \vec{\beta} \times \frac{\vec{k}}{\omega}\right)_j e_{0j} + \gamma_i\left(\omega\vec{E} - c^2\vec{B} \times \vec{k}\right)_j e_{0j} + \\ \delta_i\left(\omega\vec{B} + \vec{E} \times \vec{k}\right)_j e_{0j} = 0 \ , \tag{3.127}$$

which in turn let's us express the equation as

$$M_{ij}e_{0j} = 0 \ , \tag{3.128}$$

$$M_{ij} = a\delta ij + k_iK_j + \gamma_i\Gamma_j + \delta_i\Delta_j \ , \tag{3.129}$$

where

$$\vec{K} \equiv -\vec{\alpha} + \vec{\beta} \times \frac{\vec{k}}{\omega} \ , \tag{3.130}$$

$$\vec{\Gamma} \equiv \omega \vec{E} - c^2 \vec{B} \times \vec{k} \ , \tag{3.131}$$

$$\vec{\Delta} \equiv \omega \vec{B} + \vec{E} \times \vec{k} \ . \tag{3.132}$$

All that is left is to work out a relation for

$$\det \left[a\delta_{ij} + b \ k_i K_j + c \ \gamma_i \Gamma_j + d \ \delta_i \Delta_j \right] \ , \tag{3.133}$$

which after some lengthy manipulations will be given as

$$\begin{aligned} &\left(\frac{\omega^2}{c^2} - k^2\right) \left[\left(\frac{\omega^2}{c^2} - k^2\right) \left(\left(\frac{\omega^2}{c^2} - k^2\right) + \vec{k} \cdot \vec{K} + \vec{\gamma} \cdot \vec{\Gamma} + \vec{\delta} \cdot \vec{\Delta} \right) + \left(\vec{k} \cdot \vec{K} \vec{\gamma} \cdot \vec{\Gamma} - \vec{k} \cdot \vec{\Gamma} \vec{K} \cdot \vec{\gamma} \right) \right. \\ &+ \left(\vec{k} \cdot \vec{K} \vec{\delta} \cdot \vec{\Delta} - \vec{k} \cdot \vec{\Delta} \vec{K} \cdot \vec{\delta} \right) + \left(\vec{\gamma} \cdot \vec{\Gamma} \vec{\delta} \vec{\Delta} - \vec{\gamma} \cdot \vec{\Delta} \vec{\Gamma} \cdot \vec{\delta} \right) \Big] + \Big[\left(\vec{k} \cdot \vec{K} \right) \left(\vec{\gamma} \cdot \vec{\Gamma} \vec{\delta} \cdot \vec{\Delta} - \vec{\gamma} \cdot \vec{\Delta} \vec{\Gamma} \cdot \vec{\delta} \right) \\ &\left. + \left(\vec{k} \cdot \vec{\Delta} \right) \left(\vec{K} \cdot \vec{\gamma} \vec{\Gamma} \cdot \vec{\delta} - \vec{\gamma} \cdot \vec{\Gamma} \vec{K} \cdot \vec{\delta} \right) + \left(\vec{k} \cdot \vec{\Gamma} \right) \left(\vec{K} \cdot \vec{\delta} \vec{\gamma} \cdot \vec{\Delta} - \vec{K} \cdot \vec{\gamma} \vec{\delta} \cdot \vec{\Delta} \right) \Big] = 0 \right. \end{aligned} \ . \tag{3.134}$$

Finding the roots of this equation is a daunting work, due to the full expression of the dispersion relation. We can however infer from the two previous simpler cases discussed that this general solution will entail three different roots:

- A meaningless root $\omega = 0$.
- A gauge root $\omega = ck$, which has does not represent the physical propagation of the theory.
- A birefringent root, which can be reduced to $\omega = ck$ within a suitable limit.

And while we do not offer the expression of the last case (as the expression does not appear to possess a closed form for the roots), we may however use the results of section 3.3.2, to describe astrophysical/cosmological scenarii (since the electric field on large scales is expected to be null), or even linear accelerator experiments (where we can naively set $\vec{B} = 0$) after using the dictionary we offered.

Chapter 4

Mass generation in Generalised Non-Linear Electrodynamics

Résumé

Le chapitre s'ouvre sur un problème de la théorie introduite au chapitre précédent: l'absence de photon se propageant librement. À partir de l'action au deuxième ordre $\mathcal{L}_2$, nous négligeons les contributions d'ordre supérieur, liées au coefficients D_1, D_2 et D_3 supposées suppressées par construction afin d'arriver à une forme plus adaptée à la discussion qui suivra. Cela conduit à une action effective où le photon est couplé au fond via un coefficient C_1, pris non constant. En supprimant les termes dual pour simplifier encore plus l'analyse, l'action prend une forme où apparaît un courant effectif dépendant du fond.

Les équations du mouvement montrent immédiatement une différence cruciale entre fond constant et non constant. Dans le premier cas, on retrouve les équations de Maxwell usuelles. Dans le second, le photon reste couplé au fond même en absence de sources externes. Cette dépendance structurelle empêche toute propagation libre: le champ ne peut pas être découplé du fond, ce qui compromet la récupération des relations de dispersion standards.

Ce problème motive l'introduction d'une redéfinition de champ. Une première tentative consiste à redimensionner directement le tenseur électromagnétique, mais cette dernière ignorant la dynamique fondamentale de la théorie, est finalement rejetée. Le potentiel a_μ est alors redéfini en absorbant C_1 dans une fonction $\alpha = \sqrt{C_1}$. Cette transformation modifie profondément la structure du tenseur de champ, qui acquiert des termes supplémentaires dépendant des dérivées de α.

La question de l'invariance de jauge de la théorie est ensuite soulevée. Bien que la redéfinition semble la briser, une analyse plus fine montre que l'invariance est en réalité préservée, fait confirmé plus tard dans le chapitre. Les transformations de jauge des anciens et nouveaux champs restent compatibles, et la variation du tenseur de champ redéfini s'annule, confirmant que la théorie reste construite à partir de quantités invariantes.

En réécrivant l'action avec les nouveaux champs, on fait apparaître des termes supplémentaires. Parmi eux, un terme quadratique en potentiel, qui se comporte comme un terme de masse effectif. Ce terme dépend d'un vecteur V_μ, défini comme le gradient logarithmique de α. Ainsi, la redéfinition de champ induit naturellement une masse pour le photon.

Dans la section suivante, les équations du mouvement sont analysées en détail afin d'en tirer les propriétés de propagation. Une simplification importante est effectuée en premier lieu, consistant à éliminer la partie symétrique et sans trace du tenseur $V_\mu V_\nu$, ce qui reviendra à une forme de fixation de jauge. On suppose également que le fond est non dynamique et purement spatial. Dans ces conditions, nous remarquons que l'action

contient explicitement un terme de masse bien défini, avec un signe correct.

Les équations du mouvement sont ensuite transformées en espace de Fourier pour obtenir les relations de dispersions. La résolution complète étant difficile, et souhaitant continuer la démarche présentée dans un des papiers publiés, une réduction mécanique est adoptée en annulant les composantes spatiales du vecteur d'onde. Cela permet d'obtenir une fréquence propre non nulle, proportionnelle à la masse effective. Cette fréquence est positive, ce qui exclut la présence de modes tachyoniques au niveau classique.

Même si les vitesses de propagation ne sont pas explicitement calculées dans ce cadre réduit, la présence d'un référentiel de repos indique une propagation subluminale. Cette interprétation est cohérente avec une vision optique: le fond agit comme un milieu modifiant l'indice de réfraction du vide. La masse effective reflète alors ce ralentissement de la propagation.

Pour confirmer la nature massive du modèle, une analyse hamiltonienne est entreprise. La contrainte primaire $\pi^0 = 0$ reste la même que dans toute théorie d'électrodyanmique représentant l'absence de polarisation temporelle. L'algorithme de Dirac-Bergmann permet ensuite de déterminer la contrainte secondaire, qui correspondent à une loi de Gauss modifiée.

Le système présentant deux contraintes de seconde classe, permet de réduire l'espace des phases via les crochets de Dirac. L'Hamiltonien réduit est alors exprimé uniquement en termes des variables physiques. Il est borné par le bas, ce qui garantit l'absence d'instabilités de type Ostrogradski.

Le comptage des degrés de liberté montre qu'il y en a trois, ce qui correspond à un champ vectoriel massif. Cela renforce l'interprétation précédente: le photon acquiert une polarisation supplémentaire.

Une approche plus générale est ensuite développée sans les simplifications utilisées précédemment. Les équations du mouvement sont obtenues dans leur forme complète, puis transformées en espace de Fourier. Le déterminant de l'opérateur différentiel est calculé en le décomposant en parties symétrique et antisymétrique, ce qui simplifie considérablement l'analyse.

Deux relations de dispersion émergent. La première donne une relation modifiée entre fréquence et vecteur d'onde, incluant des termes dépendant du fond. La seconde impose une orthogonalité entre le vecteur d'onde et le fond, interprétée comme une condition nécessaire à la stabilité et à l'invariance de jauge.

Dans le cas général, les fréquences peuvent devenir complexes, indiquant un amortissement ou des instabilités. Pour éviter cela, il faut imposer la condition de transversalité. Dans ce cas, la relation de dispersion redevient réelle et correspond à celle d'un champ massif. Les vitesses de groupe associées confirment une propagation subluminale

dans ce régime.

L'étude du propagateur montre la présence de deux pôles: un pôle transverse massif et un pôle longitudinal lié à la structure de jauge. Ce dernier est initialement singulier, ce qui impose l'introduction d'une fixation de jauge pour régulariser la théorie. Le pôle transverse confirme l'existence d'une masse effective induite par le fond.

Une analyse hamiltonienne plus générale, sans fixation préalable, montre cette fois la présence de contraintes de première classe. Cela confirme explicitement que la théorie est bien une théorie de jauge. La transformation de jauge obtenue est modifiée par le fond, ce qui conduit à une dérivée covariante décalée, mais n'influe pas sur la structure de groupe de jauge, gardant la nilpotence de l'opérateur BRST et l'unitarité de la théorie.

La redéfinition de champ apparaît alors comme une brisure de symétrie seulement apparente, analogue à un mécanisme de Higgs au niveau effectif. Le fond, en étant non constant, modifie l'état du vide et peut être interprété comme induisant une brisure spontanée de symétrie.

Enfin, les contributions négligées initialement sont discutées. Les termes d'ordre supérieur ne peuvent annuler la masse, car ils introduisent des structures de dimension différente tandis que les termes duaux peuvent au mieux modifier les relations de dispersion.

Le chapitre se conclut sur une interprétation physique du mécanisme de génération de masse. Celle-ci est vue comme un effet émergent dû au fond, analogue à d'autres mécanismes connus en physique des champs. La question de la nature physique de cette masse reste ouverte, notamment au niveau quantique, où des effets d'anomalies pourraient intervenir.

4.1 Field redefinition

In the previous chapter we introduced the foundations of generalised non-linear electrodynamics, in its most general form (*i.e.* fourth order in photon field strength) and discussed briefly its key features. Something that was however not highlighted is the absence of a free propagating photon in the theory. Let us now consider the following expression of the action at second order

$$\mathcal{L}_2 = -\frac{1}{2}\left(C_1 F^{\mu\nu} + C_2 \tilde{F}^{\mu\nu}\right) f_{\mu\nu} - \frac{1}{4} C_1 f^{\mu\nu} f_{\mu\nu} - \frac{1}{4} C_2 f^{\mu\nu} \tilde{f}_{\mu\nu} + \frac{1}{8} Q^{\mu\nu\rho\sigma} f_{\mu\nu} f_{\rho\sigma} \ , \quad (4.1)$$

and express it as

$$\mathcal{L}_2 = \mathcal{L}_1 + \delta\mathcal{L}_2 \ , \quad (4.2)$$

where

$$\left.\frac{\partial \mathcal{L}}{\partial \mathcal{F}}\right|_{\mathrm{B}} = C_1 \ , \left.\frac{\partial \mathcal{L}}{\partial \mathcal{G}}\right|_{\mathrm{B}} = C_2 \ . \quad (4.3)$$

where all the terms that depend on second derivatives of the background fields are encompassed in $\delta\mathcal{L}_2$. In what follows, we will ignore $\delta\mathcal{L}_2$ to consider only leading order effects. Indeed, by nature of the power expansion, higher order corrections must be suppressed, imposing for the D coefficients to be much smaller than the C coefficients. This reduces the system to what we will now call the "first" order action, since it can also be obtained by doing the background field expansion at first order in $\mathcal{F}$

$$\mathcal{L}_1 = -\frac{1}{2}\left(C_1 F^{\mu\nu} + C_2 \tilde{F}^{\mu\nu}\right) f_{\mu\nu} - \frac{1}{4} C_1 f^{\mu\nu} f_{\mu\nu} - \frac{1}{4} C_2 f^{\mu\nu} \tilde{f}_{\mu\nu} \ . \quad (4.4)$$

For the sake of the discussion, we will also drop dual contributions by setting $C_2 = 0$. This may seem motivated by physical reasons since we may assume that the C_2 terms are topological, but in our case C_2 being a variable breaks this assumption. In our situation, it's a simplification choice to make our argument clearer but we will address some concerns that could arise here later on. This leads then to the following form

$$\mathcal{L}_1 = -\frac{1}{4} C_1 f^{\mu\nu} f_{\mu\nu} \ - \frac{1}{2}\left(C_1 F^{\mu\nu}\right) f_{\mu\nu} \ , \quad (4.5)$$

in which we can integrate by parts the second term and get

$$\mathcal{L}_1 = -\frac{1}{4} C_1 f^{\mu\nu} f_{\mu\nu} + a_\nu \mathcal{J}^\nu \ , \quad (4.6)$$

where $\mathcal{J}^\nu = \partial_\mu (C_1 F^{\mu\nu})$ assumes the role of a conserved current. Taking the variation of $\mathcal{L}_1$ with respect to the photon 4-potential a^μ gives us the equations of motion

$$\partial_\mu (C_1 f^{\mu\nu}) + \mathcal{J}^\nu = 0 \ . \tag{4.7}$$

In the case where the background field is taken to be constant, the solution is of course the standard usual Maxwell theory since

$$C_1 \partial_\mu f^{\mu\nu} = 0 \ , \tag{4.8}$$

is totally independent from the background field. However, if the background is taken to be non constant we get

$$\partial_\mu C_1 f^{\mu\nu} + C_1 \partial_\mu f^{\mu\nu} = -\mathcal{J}^\nu \ , \tag{4.9}$$

which shows two things: first, that the photon is never propagating in a vacuum, despite the absence of external sources due to the background contribution. This has already been discussed previously, and we could naïvely neglect this source term. However, we would still be left with

$$\partial_\mu f^{\mu\nu} = -\frac{1}{C_1} \partial_\mu C_1 f^{\mu\nu} \ , \tag{4.10}$$

which infers that the photon is never freely propagating, as it seems fundamentally coupled with respect to the background field. This is a structural problem in the model, as it seems that we may not retrieve the usual dispersion relations in any limit as long as $C_1 \neq constant$.

This motivates us to look for alternative formulations, that would decouple the photon propagation from the background field. To this end, we will search for suitable field redefinitions that achieve our goal. The source of the Maxwellian photon is the following part of the action

$$\mathcal{L}_0 = -\frac{1}{4} C_1 f^{\mu\nu} f_{\mu\nu} \ \ . \tag{4.11}$$

The simplest possible field redefinition would be to resize the field strength as $f'^{\mu\nu} \to \sqrt{C_1} f^{\mu\nu}$, which would then immediately simplify the action to

$$\mathcal{L}_0 = -\frac{1}{4} f'^{\mu\nu} f'_{\mu\nu} \ . \tag{4.12}$$

This has the benefit of preserving gauge invariance, and seems to not alter too much the structure of the theory. It is however, a bad field redefinition, as redefinitions should

act on the fundamental fields of the theory, which is not the case of the field strength. In electrodynamics, this implies that we must redefine the 4-potential, and keeping in mind that the coefficient C_1 is not taken to be constant, it will radically change the structure of the theory.

Since the action is quadratic in a^μ, we will define

$$\alpha = \sqrt{C_1} \ , \tag{4.13}$$

as an intermediate object to reduce our notations, before applying the following field redefinition

$$a'^\mu \to \alpha \ a^\mu = \sqrt{C_1} \ a^\mu \ , \tag{4.14}$$

which will then define the field strength as

$$f'^{\mu\nu} = \partial^\mu a'^\nu - \partial^\nu a'^\nu = \partial^\mu(\alpha a^\nu) - \partial^\nu(\alpha a^\mu) = \alpha f^{\mu\nu} + a^\nu \partial^\mu \alpha - a^\mu \partial^\nu \alpha \ \ . \tag{4.15}$$

We now rewrite the original field strength in terms of the redefined quantities

$$\alpha f^{\mu\nu} = f'^{\mu\nu} - \frac{1}{\alpha}\left(a'^\nu \partial^\mu \alpha - a'^\mu \partial^\nu \alpha\right) \ \ , \tag{4.16}$$

and we then set $V^\mu = \dfrac{\partial^\mu \alpha}{\alpha}$, getting

$$\alpha f^{\mu\nu} = f'^{\mu\nu} - \left(a'^\nu V^\mu - a'^\mu V^\nu\right) \ \ . \tag{4.17}$$

Continuing further, we can derive

$$\alpha^2 f^{\mu\nu} f_{\mu\nu} = f'^{\mu\nu} f'_{\mu\nu} + 4 a'_\mu V_\nu f'^{\mu\nu} + 2\left(a'^\nu a'_\nu V^\mu V_\mu - a'_\mu a'_\nu V^\mu V^\nu\right) \ \ , \tag{4.18}$$

and get the equivalent form of the action

$$\mathcal{L}'_0 = -\frac{1}{4} f'^{\mu\nu} f'_{\mu\nu} - a'_\mu V_\nu f'^{\mu\nu} - \frac{1}{2}\left(a'^\nu a'_\nu V^\mu V_\mu - a'_\mu a'_\nu V^\mu V^\nu\right) \ \ . \tag{4.19}$$

The action seems to exhibit the presence of a mass term, that can be expressed as

$$-\frac{1}{2}\left(a'^\nu a'_\nu V^\mu V_\mu - a'_\mu a'_\nu V^\mu V^\nu\right) = \frac{1}{4}(\varepsilon^{\mu\nu\rho\sigma} a_\rho V_\sigma)^2 \ , \tag{4.20}$$

but we will prefer the first expression for what is to follow. To conclude this part about field redefinitions, we may worry about the current term that is present in the first action that was drop before the field redefinition. We can, however, show that it still

behaves as a current term post-field redefinition

$$a_\nu \mathcal{J}^\nu \to \frac{1}{\alpha} a'_\nu \mathcal{J}^\nu = a'_\nu \mathcal{J}'^\nu \ , \tag{4.21}$$

with $\mathcal{J}'^\nu = \dfrac{\partial_\mu \left(\alpha^2 F^{\mu\nu}\right)}{\alpha}$.

Having a closer look at the gauge invariance of the action and using the usual gauge transformation for $U(1)$ theories $a''^\mu = a'^\mu + \partial^\mu \chi$ we get,

$$\begin{aligned} \delta\mathcal{L} &= \delta \left(-\frac{1}{4} f'^{\mu\nu} f'_{\mu\nu} - a'_\mu V_\nu f'^{\mu\nu} - \frac{1}{2} \left(a'^\nu a'_\nu V^\mu V_\mu - a'_\mu a'_\nu V^\mu V^\nu \right) \right) \\ \delta\mathcal{L} &= \left(V_\nu \partial_\mu f'^{\mu\nu} + V^2 \partial_\mu a'^\mu - V^\mu V^\nu \partial_\mu a'_\nu \right) \chi \ \ , \end{aligned} \tag{4.22}$$

which holds up to 4-divergences. This result confirms that the action is not gauge invariant under the usual $U(1)$ transformations. We will however show in section 4.4 that the theory is indeed gauge invariant, but under a different set of transformations.

4.2 Equations of motion and dispersion relations

In this section, and the upcoming one, we will take the rather reductive approach that was taken in [101], to highlight the published version of our work. We will however show in the last section that said approach was incorrect, due to an inappropriate gauge fixing. Indeed, at the time that the paper was written, we assumed that massive models must necessarily include three degrees of freedom (which sounds absolutely reasonable), and we were not well versed with topological mass generations. We believe that however, despite the improper[1] gauge fixing that occurred (and thus loss of gauge structure of the theory), the propagation features of the theory to be relatively similar, and this endeavour gave us a first hint towards the stability of the theory (despite its background dependence).

To start, let us remind the reader that any totally symmetric tensor can be decomposed as a totally symmetric part plus a trace, under its fundamental representation

$$S^{\mu\nu} = \mathcal{S}^{\mu\nu} + \frac{1}{4}\eta^{\mu\nu}S \ , \tag{4.23}$$

where $\mathcal{S}^{\mu\nu}$ is the totally symmetric and traceless part. Since in our action, the term $V^\mu V^\nu$ forms a totally symmetric tensor, we may then use the above decomposition to get

$$V^\mu V^\nu = \mathcal{V}^{\mu\nu} + \frac{1}{4}\eta^{\mu\nu}V_\alpha V^\alpha \ \ , \tag{4.24}$$

where $\mathcal{V}^{\mu\nu}$ represents the totally symmetric and traceless part. While totally symmetric and traceless tensors play an important role in gravitation[2], it does not in electromagnetism since the field strength is by definition totally anti-symmetric, we will thus choose to kill this term allowing us to rewrite the action

$$\mathcal{L}_0' = -\frac{1}{4}f'^{\mu\nu}f'_{\mu\nu} - a'_\mu V_\nu f'^{\mu\nu} - \frac{1}{2}\left(a'^\nu a'_\nu V^\mu V_\mu - a'_\mu a'_\nu V^\mu V^\nu\right) \ , \tag{4.25}$$

as

$$\mathcal{L}_0 = -\frac{1}{4}f'^{\mu\nu}f'_{\mu\nu} - a'_\mu V_\nu f'^{\mu\nu} - \frac{3}{8}V^\mu V_\mu a'^\nu a'_\nu \ \ . \tag{4.26}$$

As stated in the previous chapter, even though we may consider non-constant backgrounds, we limit our studies to non-dynamical ones. This imposes for V^μ to be a purely space-like vector since $\partial^0\alpha = 0$. Additionally, we will consider that the back-

[1] This is a language abuse, by improper we meant inappropriate.

[2] TT gauge for gravitational waves [102, 103]

ground vector V^μ is constant, otherwise put the momentum of the background field p^μ is not evolving in time nor space. Now, since V^μ is purely space-like, we have $V^\mu V_\mu = V^0 V_0 + V^i V_i = -V_i V_i = -\vec{V}^2$ which in turn allows us to define $M = \frac{3}{4}\vec{V}^2$ and get

$$\mathcal{L}_0 = -\frac{1}{4} f'^{\mu\nu} f'_{\mu\nu} - a'_\mu V_\nu f'^{\mu\nu} + \frac{1}{2} M a'^\nu a'_\nu \quad , \tag{4.27}$$

ensuring the proper sign for the mass term.

By then taking the variation of the action with respect to the 4-potential,

$$\delta\mathcal{L}_0 = -f'^{\mu\nu}\partial_\mu \delta a'\nu - \delta a'_\mu V_\nu f'^{\mu\nu} - a'_\mu V_\nu \partial^\mu \delta a'^\nu + a'_\mu V_\nu \partial^\nu \delta a'^\mu + M a'^\nu \delta a'_\nu \ , \tag{4.28}$$

the fourth term can be re-expressed as $\frac{1}{2} V_\nu \partial^\nu (a'^\mu a'_\mu)$, and act as a surface term that will be dropped before applying the variations on the action. We then get the expression for the equations of motion

$$\partial_\mu f'^{\mu\nu} + M a'^\nu + V^\nu \partial_\mu a'^\mu - V^\mu \partial^\nu a'_\mu = 0 \quad . \tag{4.29}$$

Rewriting the equation in term of potentials

$$\left(\eta^{\mu\nu}\Box - \partial^\mu\partial^\nu + \eta^{\mu\nu} M - V^\mu \partial^\nu + V^\nu \partial^\mu\right) a'_\mu = 0 \ , \tag{4.30}$$

and using the phase exponential ansatz $\partial^\mu = -ik^\mu$ we may express the equations of motion in Fourier space

$$\left(-\eta^{\mu\nu} k^2 + k^\mu k^\nu + \eta^{\mu\nu} M + i V^\mu k^\nu - i V^\nu k^\mu\right) \tilde{a}'_\mu = 0 \ . \tag{4.31}$$

To find the dispersion relations we have to solve

$$D^{\mu\nu} \ \tilde{a}'_\mu = 0 \ , \tag{4.32}$$

where $D^{\mu\nu}$ is

$$D^{\mu\nu} = -\eta^{\mu\nu} k^2 + k^\mu k^\nu + \eta^{\mu\nu} M + i V^\mu k^\nu - i V^\nu k^\mu \ . \tag{4.33}$$

Computing the determinant of $D^{\mu\nu}$ is quite complicated and in [101], we did not yet find its expression. This will however be addressed later on. The simpler path we chose is via the mechanical reduction of the theory, which consists on setting the space variations to zero. Otherwise put, $k^i = 0$. This, in addition to the fact that the non-dynamical nature of the vector V^μ sets its time component to zero $V^0 = 0$, yields for the equations

of motion

$$-\tilde{a}'_0 M - iV^i k^0 \tilde{a}'_i = 0 \ , \tag{4.34}$$

$$-k^2 \tilde{a}'_i + M\tilde{a}'_i - iV_i k^0 \tilde{a}'_0 = 0 \ , \tag{4.35}$$

From (4.35)

$$\tilde{a}'_i = \frac{iV^i k^0}{-k^2 + M} \tilde{a}'_0 \ \ , \tag{4.36}$$

and (4.34)

$$\tilde{a}'_0 = -\frac{iV^i k^0}{M} \tilde{a}'_i \ \ , \tag{4.37}$$

injecting (4.36) into (4.34) and (4.37) into (4.35), it gives, respectively,

$$\left(\frac{4}{3}k_0^2 M - M^2\right) \tilde{a}'_0 = 0 \ \ , \tag{4.38}$$

$$\left(-\frac{4}{3}k_0^2 + M\right) \tilde{a}'_i = 0 \ \ , \tag{4.39}$$

and as such we get for the frequency

$$k_0 = \omega = \sqrt{\frac{3}{4}M} \ \ , \tag{4.40}$$

which appears to be definite positive due to the definition of M ensuring the absence of tachyonic modes at the classical level. Even though we cannot derive the phase and group velocities from (4.40), we can highlight that the presence of a rest frame indicates a propagation at below c. In a media, the interactions that light has along its path, slow it down yielding what is referred to as the refractive index in the geometrical optics limits. Since we are in a non-linear system, the propagation of light in a background should possess a slower than c velocity, as the background should affect the refractive index of the vacuum. This is coherent with the result we obtained, as M is a function that depends entirely on the background field. In that sense, we can argue that the effect on the propagation of the effective mass present in our modified action, is strictly equivalent to the change in refractive index that the background fields would have induced in the full non-linear picture. This discussion will be further clarified down the line with the full dispersion relations, confirming our initial hunch.

The presence of a rest frame while a solid hint of the presence of mass, is still insufficient, we will thus consider the Hamiltonian picture of the model as it will inform us on the gauge theoretical structure of the theory as well as the number of physical degrees of freedom.

4.3 Hamiltonian picture

As shown in chapter 1 and 2, we start by breaking the manifest covariance of the Lagrangian density (4.27)

$$\mathcal{L}_0 = -\frac{1}{2}f'^{0i}f'_{0i} - \frac{1}{4}f'^{ij}f'_{ij} + \frac{1}{2}M\left(a'^0 a'_0 + a'^i a'_i\right) - a'_0 V_i f'^{0i} - a'_i V_j f'^{ij} \ , \tag{4.41}$$

to then proceed with the Legendre transform

$$\pi^\mu = \frac{\partial \mathcal{L}_\prime}{\partial \dot{a}'_\mu} = \begin{cases} \pi^0 = \dfrac{\partial \mathcal{L}_0}{\partial \partial_0 a'_0} = 0 \\ \\ \pi^i = \dfrac{\partial \mathcal{L}_0}{\partial \partial_0 a'_i} = f'^{i0} - a'^0 V^i \end{cases} \ ,$$

where $\pi^0 = 0$ is the primary constraint that indicates the absence of propagating time-like polarisations. We continue by expressing the velocities as functions of the conjugate momenta

$$f'^{i0} = \pi^i + a'^0 V^i \rightarrow \partial^0 a'^i = \partial^i a'^0 - \pi^i - a'^0 V^i \ , \tag{4.42}$$

which will then have to be injected into the action.

The Hamiltonian will then be defined as

$$\begin{aligned} \mathcal{H} &= \int d^3x \ \left[\pi^i \partial_0 a'_i - \mathcal{L}_0\right] \\ &= \int d^3x \ \left[\pi^i \partial_i a'_0 - \pi^i V_i a'_0 - \pi^i \pi_i - \mathcal{L}_0\right] \\ &= \int d^3x \ \left[-a'_0 \partial^i \pi_i - \pi^i V_i a'_0 - \frac{1}{2}\pi^i \pi_i + \frac{1}{4}f'^{ij}f'_{ij} + a'_i V_j f'^{ij} - \frac{1}{2}\left(M' a'^0 a'_0 + M a'^i a'_i\right)\right] \ , \end{aligned} \tag{4.43}$$

where $M' = \dfrac{7}{4}V^i V_i$.

With the Hamiltonian density in hands, we can now study the constraints. We then apply the Dirac-Bergmann algorithm [12,13,30] and construct the augmented (or total) Hamiltonian which will include all the primary constraints

$$\mathcal{H}_{Aug} = \int d^3x \mathcal{H} + \lambda_i(x)\phi^i(x) \ , \tag{4.44}$$

where ϕ^i represents all the primary constraints and λ_i are the Lagrange multipliers that enforces them.

We have $\pi^0 = 0$ acting as a constraint which we will define as

$$\phi_1 = \pi^0 = 0 \ , \tag{4.45}$$

giving for the augmented Hamiltonian

$$\mathcal{H}_{Aug} = \int d^3x \mathcal{H} + \lambda^1(x)\phi_1(x) \ . \tag{4.46}$$

The Poisson bracket of the primary constraint ϕ_1 with the augmented Hamiltonian reads

$$\{\phi_1 \ , \ \mathcal{H}_{Aug}\} = \{\pi^0 \ , \ \mathcal{H}_{Aug}\} = \partial^i \pi_i + \pi^i V_i + a_0' M' \approx 0 \ \ , \tag{4.47}$$

This new constraint, is nothing more than the modified Gauss law for the system. It will be relabelled to as

$$\phi_2 = \partial^i \pi_i + \pi^i V_i + a_0' M' \approx 0 \ \ . \tag{4.48}$$

We now proceed with the consistency of this secondary constraint

$$\mathcal{H}_{Aug} = \mathcal{H} + \int d^3x \ \lambda^1 \phi_1 \ \ . \tag{4.49}$$

Continuing and computing the Poisson bracket between the new constraint and $\mathcal{H}_{Aug}$ gives

$$\begin{aligned} \{\phi_2 \ , \ \mathcal{H}_{Aug}\} =& \frac{3}{2} V_k \partial_j f'^{jk} + V_k V^k \partial_j a'^j - V^i V^k \partial_k a_i' - M \partial_j a'^j + \lambda_1 M' \approx 0 \\ \Rightarrow \lambda_1 =& -\frac{5}{6M} \partial_j f'^{jk} - \frac{1}{27} \partial^j a_j' \ \ , \end{aligned} \tag{4.50}$$

which only fixes a Lagrange multiplier. There are thus no additional constraints.

As we now possess the full set of constraints, we can derive their algebra following Dirac's classification [13] as we have done previously.

A quick computation gives

$$\{\phi_1(t, \vec{x}) \ , \ \phi_2(t, \vec{y})\} = -M' \delta^3(\vec{x} - \vec{y}) \neq 0 \ \ , \tag{4.51}$$

which indicates that the system is composed of two second class constraints that form

a skew-symmetric invertible matrix

$$\phi^{ij}(\vec{x},\vec{y}) = \begin{pmatrix} 0 & -M' \\ M' & 0 \end{pmatrix} \delta^3(\vec{x}-\vec{y}) \quad , \tag{4.52}$$

obeying the identity

$$\int d^3z \sum_k \phi_{ik}(\vec{x},\vec{z})\ \phi^{-1}_{kj}(\vec{z},\vec{y}) = \delta_{ij}\delta^3(\vec{x}-\vec{y}). \tag{4.53}$$

.

Using the constraints matrix, we can construct the Dirac bracket of our model, which in turn will allow us to reduce our phase space. For the phase space functionals $F(t)$ and $G(t)$, we get

$$\{F(t),G(t)\}_D = \{F(t),G(t)\} - \sum_{i,j=1}^{2} \int d^3\vec{z}\, d^3\vec{u}\ \ \{F(t),\ \phi_i(t,\vec{z})\}\ \phi^{-1}_{ij}(\vec{z},\vec{u})\ \{\phi_j(t,\vec{u}),\ G(t)\}\ . \tag{4.54}$$

Applying (4.53) it is easy to show that the Dirac bracket of the constraints with any phase space function is zero

$$\{F(t),\phi_k(t,\vec{x})\}_D = \{F(t),\phi_k(t,\vec{x})\} - \sum_{i,j=1}^{2} \int d^3\vec{z}\, d^3\vec{u}\ \ \{F(t),\phi_i(t,\vec{z})\}\ \phi^{-1}_{ij}(\vec{z},\vec{u})\ \phi_{jk}(\vec{u},\vec{x}) = 0, \tag{4.55}$$

this in turn allows us to take the constraints ϕ_1 and ϕ_2 as strong equalities and replace them in the total Hamiltonian yielding our reduced Hamiltonian

$$\mathcal{H}_{Red} = \int d^3x \left[\frac{1}{2}\pi_i^2 + \frac{1}{4}{f'_{ij}}^2 + \frac{1}{2M'}\left(\partial^i\pi_i + \pi^i V_i\right)^2 + a'_i V_j f'^{ij} + \frac{1}{2}M{a'_i}^2\right], \tag{4.56}$$

which indicates explicitly the presence of only the vector potentials and their conjugate momenta as phase space variables. It is also important to note that the Hamiltonian appears to be bound from below and as such the model is free of Ostrogradski instabilities, meaning that the model is classically well behaved.

The total number of physical degrees of freedom of the system is then given by the

master formula [11]

$$2\times DOF = \text{ Phase Space variables} - 2\times\text{First Class Constraints} - \text{ Second Class Constraints .} \tag{4.57}$$

The phase space is composed of a'^{μ} and π^{μ} and as a result we get

$$2 \times DOF = 8 - 0 - 2 = 6 \quad \Rightarrow DOF = 3 \ , \tag{4.58}$$

which shows that there are three propagating degrees of freedom within the model.

At this stage, one may now wonder whether we can still consider the model to be a gauge theory, as it is usually accepted that only first class constraints generate gauge transformations. There has been however recent results [104] that show that even in the case of de Broglie-Proca electrodynamics, the second class constraints can generate gauge transformations. This reassures us with the notion that we are still dealing fundamentally with a gauge theory of electromagnetism despite the manifest loss of gauge invariance.

4.4 A general approach

The absence of gauge invariance by the standard $U(1)$ transformations, is a problem that has to be tackled. By setting $V^{\mu\nu}$ to zero, we most likely have proceeded with an improper gauge fixing. To close the discussion around the gauge invariance of the theory, and stray away from the published work, we shall look at a more general case of the action. We will not set V^{μ} to be purely space-like, nor fix any parts of the action à priori. The only remaining condition, is that V^{μ} is taken to be constant. In what follows, we will attempt to derive the dispersion relations for the theory, its group velocity as well as its propagator. If the above results confirm our intuition about the gauge invariance of the theory, we shall move on to the Hamiltonian picture, to confirm the number of degrees of freedom, and find the gauge transformation. Let us go back to our initial action (4.19)

$$\mathcal{L}_0' = -\frac{1}{4} f'^{\mu\nu} f'_{\mu\nu} - a'_{\mu} V_{\nu} f'^{\mu\nu} - \frac{1}{2} \left(a'^{\nu} a'_{\nu} V^{\mu} V_{\mu} - a'_{\mu} a'_{\nu} V^{\mu} V^{\nu} \right) \ , \tag{4.59}$$

we may then get the equations of motion as

$$\partial_{\mu} f'^{\mu\nu} + V^{\nu} \partial^{\mu} a'_{\mu} - V^{\mu} \partial^{\nu} a'_{\mu} - a'^{\nu} V^{\mu} V_{\mu} + a'^{\mu} V^{\nu} V_{\mu} = 0 \ \ , \tag{4.60}$$

and by then using the plane wave ansatz $\partial_{\mu} = -ik_{\mu}$, we get the Fourier space representation

$$\left[-\eta^{\mu\nu} \left(k^2 + V^2 \right) + k^{\mu} k^{\nu} + i V^{\mu} k^{\nu} - i V^{\nu} k^{\mu} + V^{\mu} V^{\nu} \right] \tilde{a}(k)_{\mu} = 0 \ . \tag{4.61}$$

We must now try to estimate the determinant of

$$D^{\mu\nu} = \left[-\eta^{\mu\nu} \left(k^2 + V^2 \right) + k^{\mu} k^{\nu} + i V^{\mu} k^{\nu} - i V^{\nu} k^{\mu} + V^{\mu} V^{\nu} \right] \ , \tag{4.62}$$

which we will rewrite as

$$D^{\mu\nu} = -\alpha \eta^{\mu\nu} + \chi^{\mu\nu} \ , \tag{4.63}$$

where

$$\alpha = k^2 + V^2 \ , \tag{4.64}$$

$$\chi^{\mu\nu} = k^{\mu} k^{\nu} + i V^{\mu} k^{\nu} - i V^{\nu} k^{\mu} + V^{\mu} V^{\nu} = \left(k^{\mu} + i V^{\mu} \right) \left(k^{\nu} - i V^{\nu} \right) \ , \tag{4.65}$$

which then gives

$$D^{\mu\nu} = -\alpha \left(\eta^{\mu\nu} - \frac{1}{\alpha} \chi^{\mu\nu} \right) \ , \tag{4.66}$$

where we will refine the last term as $\xi^{\mu\nu} = \frac{1}{\alpha}\chi^{\mu\nu}$ and then express the full determinant as a binomial expansion

$$\det(D) = (-\alpha)^4 \det(\eta^{\mu\nu} - \xi^{\mu\nu}) \ , \tag{4.67}$$

$$\begin{aligned} \det(\eta - \xi) =& 1 - \xi^{\mu}{}_{\mu} + \frac{1}{2}\left((\xi^{\mu}{}_{\mu})^2 - \xi^{\mu\nu}\xi_{\nu\mu} \right) - \frac{1}{6}\left[(\xi^{\mu}{}_{\mu})^3 - 3\xi^{\mu}{}_{\mu}(\xi^{\rho\nu}\xi_{\nu\rho}) + 2\left(\xi^{\mu\nu}\xi_{\nu\rho}\xi^{\rho}{}_{\mu} \right) \right] \\ &+ \det(\xi) \ . \end{aligned} \tag{4.68}$$

We notice that ξ can be expressed as

$$\xi^{\mu\nu} = S^{\mu\nu} + iA^{\mu\nu} \ , \tag{4.69}$$

where S is a totally symmetric part and A a totally anti-symmetric part given by the expressions

$$S^{\mu\nu} = \frac{1}{\alpha}\left(k^{\mu}k^{\nu} + V^{\mu}V^{\nu} \right) , \tag{4.70}$$

$$A^{\mu\nu} = \frac{1}{\alpha}(V^{\mu}k^{\nu} - V^{\nu}k^{\mu}) \ , \tag{4.71}$$

which greatly simplifies the upcoming computations. Term by term, the determinant will be

$$\xi^{\mu}{}_{\mu} = \frac{k^2 + V^2}{\alpha} \ , \tag{4.72}$$

$$\xi^{\mu\nu}\xi_{\nu\mu} = \frac{(k^2 + V^2)^2}{\alpha^2} \ , \tag{4.73}$$

$$\xi^{\mu\nu}\xi_{\nu\rho}\xi^{\rho}{}_{\mu} = \frac{(k^2 + V^2)^3}{\alpha^3} \ , \tag{4.74}$$

$$\det \xi = 0 \ , \tag{4.75}$$

this reduces the expression to

$$\det(\eta - \xi) = 1 - \xi^{\mu}{}_{\mu} = 1 - \frac{k^2 + V^2}{\alpha} \ , \tag{4.76}$$

yielding for $\det D$

$$\det D = (-\alpha)^4 \det(\eta - \xi) = \alpha^4 \left(1 - \frac{k^2 + V^2}{\alpha}\right) \tag{4.77}$$

$$= \alpha^3(\alpha - k^2 + V^2) = 0 \ , \tag{4.78}$$

showing that it is singular. This confirms our previous intuition about gauge invariance, and will require gauge fixings to obtain the propagator. However, we can consider that our dispersion matrix possess the following Eigenvalue structure:

$$D\{0, -\alpha, -\alpha, -\alpha\} \ , \tag{4.79}$$

and thus our dispersion relation is simply characterised by $\alpha = 0$. Doing so we get

$$k^2 + V^2 = 0 \ , \tag{4.80}$$

which will read

$$\omega = \pm\sqrt{\vec{k}^2 - V^\mu V_\mu} \ , \tag{4.81}$$

indicating the presence of a rest frame, depending on the square of the background vector. We can nevertheless see that in the most general case, this is not necessarily defined to be real, notably if V^μ is defined to be purely time-like. As such, the stability of the theory would impose it to be space-like. We can underline the fact that, in our previous assumptions, this was already accepted to be the case, since we consider that V^μ is a logarithmic gradient of the background fields, and said fields are taken to be non-dynamical (i.e. constant in time), ensuring the healthy behaviour of the model. We can further highlight this by looking at the case where V^μ is purely space-like and get

$$\omega = \pm\sqrt{\vec{k}^2 + \vec{V}^2} \ , \tag{4.82}$$

Before moving on to the propagator, we shall derive the group velocities for the theory. Starting from

$$\omega = \pm\sqrt{\vec{k}^2 + \vec{V}^2} \ , \tag{4.83}$$

and using

$$\vec{v}_g = \frac{\partial\omega}{\partial\vec{k}} \ , \tag{4.84}$$

we get

$$\vec{v}_{g1} = \frac{\vec{k}}{\sqrt{\vec{k}^2 + \vec{V}^2}} \ , \tag{4.85}$$

which indicates sub-luminal propagations, a very reassuring feature of any massive model (no tachyons, as the mechanical reduction hinted towards). Going back to the dispersion matrix (4.66)

$$D^{\mu\nu} = -\alpha \left(\eta^{\mu\nu} - \frac{1}{\alpha} \chi^{\mu\nu} \right) \ , \tag{4.86}$$

we can consider that it defines a sort of modified transverse operator, which we will rename $\mathcal{P}^{\mu\nu}$

$$D^{\mu\nu} = -(k^2 + V^2)\mathcal{P}^{\mu\nu} \tag{4.87}$$

with

$$\mathcal{P}^{\mu\nu} = \left[\eta^{\mu\nu} - \frac{(k^\mu + iV^\mu)(k^\nu - iV^\nu)}{k^2 + V^2} \right] \ . \tag{4.88}$$

We can however quickly see that this operator does not admit k^μ as a null vector, further confirming that the wave vector by itself is not transversal to the propagation. We can however find naïvely the null vector of the model, by comparing it with Maxwell's transversal projector

$$P^{\mu\nu} = \eta^{\mu\nu} - \frac{k^\mu k^\nu}{k^2} \ , \tag{4.89}$$

and noticing that the k^μ vectors, are replaced by their shifted counterparts $k^\mu \pm iV^\mu$. The imaginary factor here, is not indicating a complexified gauge direction, but is simply an artifact of the phase exponential ansatz used earlier on. Doing this replacement we can see that

$$\mathcal{P}^{\mu\nu}(k_\nu + iV_\nu) = 0 \ , \tag{4.90}$$

which indicates that the null vector of said theory is $\mathcal{K}^\mu = k^\mu + iV^\mu$. Furthermore, we can see that

$$\mathcal{K}^2 = (k^\mu + iV^\mu)(k_\mu - iV_\mu) = k^2 + V^2 = 0 \ , \tag{4.91}$$

which furthers confirms the genuine modification of the gauge structure of the theory,

as well as the transversal nature of the mass term. We will now continue our general approach by computing the propagator of the theory from the action detailing the steps taken in [101]. We start from (4.19)

$$\mathcal{L}_0 = -\frac{1}{4} f'^{\mu\nu} f'_{\mu\nu} - a'_\mu V_\nu f'^{\mu\nu} - \frac{1}{2}\left(a'^\nu a'_\nu V^\mu V_\mu - a'_\mu a'_\nu V^\mu V^\nu\right) \quad , \tag{4.92}$$

which we rewrite as

$$\mathcal{L}_0 = -\frac{1}{4} f'^2 + V_\mu a'_\nu \partial^\mu a'^\nu - V_\mu a'_\nu \partial^\nu a'^\mu - \frac{1}{2} V^2 a'^2 + \frac{1}{2} V_\mu V_\nu a'^\mu a'^\nu \ . \tag{4.93}$$

As shown above, the second term can be expressed as a surface term, and we may rewrite the third one as

$$\begin{aligned} -V_\mu a'_\nu \partial^\nu a'^\mu =& -\frac{1}{2} V_\mu a'_\nu \partial^\nu a'^\mu - \frac{1}{2} V_\mu a'_\nu \partial^\nu a'^\mu \\ =& \frac{1}{2} a'^\mu \partial_\nu (V_\mu a'^\nu) - \frac{1}{2} V_\nu a'_\mu \partial^\mu a'^\nu \\ =& \frac{1}{2} a'^\mu (V_\mu \partial_\nu - V_\nu \partial_\mu) a'^\mu \ . \end{aligned} \tag{4.94}$$

By then decomposing the metric as

$$\eta^{\mu\nu} = \theta^{\mu\nu} + \omega^{\mu\nu} \ , \tag{4.95}$$

where $\theta^{\mu\nu}$ is the transverse projector and $\omega^{\mu\nu} = \dfrac{\partial^\mu \partial^\nu}{\Box}$ the longitudinal projector, the first term becomes

$$-\frac{1}{4} f'^2 = \frac{1}{2} a^\mu \Box \theta_{\mu\nu} a^\nu \ , \tag{4.96}$$

since the field strength is a transverse function [14, 18]. After some manipulations the action will read

$$\mathcal{L}_0 = \frac{1}{2} a^\mu \left[\left(\Box - V^2\right) \theta_{\mu\nu} - V^2 \omega_{\mu\nu} + V_\mu \partial_\nu - V_\nu \partial_\mu + V_\mu V_\nu\right] a^\nu \ , \tag{4.97}$$

$$\mathcal{L}_0 = \frac{1}{2} a^\mu M_{\mu\nu} a^\nu \ , \tag{4.98}$$

where $M^{\mu\nu}$ is given by

$$M_{\mu\nu} = \alpha \theta_{\mu\nu} + \beta \omega_{\mu\nu} + \gamma V_\mu \partial_\nu + \delta V_\nu \partial_\mu + \epsilon V_\mu V_\nu \ , \tag{4.99}$$

with

$$\alpha = \Box - V^2 \tag{4.100}$$

$$\beta = -V^2 \tag{4.101}$$

$$\gamma = -\delta = 1 \tag{4.102}$$

$$\epsilon = 1 \ . \tag{4.103}$$

Our task is now to find M^{-1} such as

$$M^{\mu\kappa} M^{-1}_{\kappa\nu} = \delta^\mu_\nu = \theta^\mu{}_\nu + \omega^\mu{}_\nu \ . \tag{4.104}$$

We start by developing the above expression

$$\left(\alpha\theta^{\mu\kappa} + \beta\omega^{\mu\kappa} + \gamma V^\mu \partial^\kappa + \delta V^\kappa \partial^\mu + \epsilon V^\mu V^\kappa\right)\left(x\theta_{\kappa\nu} + y\omega_{\kappa\nu} + zV_\kappa\partial_\nu + wV_\nu\partial_\kappa + tV_\kappa V_\nu\right) = \delta^\mu_\nu \ , \tag{4.105}$$

and by expanding every term on the left hand side

$$\begin{aligned}
&\alpha x\ \theta^\mu{}_\nu + \alpha z\left(V^\mu\partial_\nu - V^\kappa\partial_\kappa\omega^\mu{}_\nu\right) + \alpha t\left(V^\mu V_\nu - \frac{\partial^\kappa V_\kappa}{\Box}\partial^\mu V_\nu\right) + \beta y\omega^\mu{}_\nu + \beta z V^\kappa\partial_\kappa\omega^\mu{}_\nu + \beta w\partial^\mu V_\nu + \\
&\beta t\frac{V^k\partial_k}{\Box}\partial^\mu V_\nu \gamma y V^\mu\partial_\nu + \gamma z V^\kappa\partial_\kappa V^\mu\partial_\nu + \gamma w\Box V^\mu V_\nu + \gamma t V^\kappa\partial_\kappa V^\mu V_\nu + \delta x\left(\partial^\mu V_\nu - V^\kappa\partial_\kappa\omega^\mu{}_\nu\right) + \\
&\delta y V^\kappa\partial_\kappa\omega^\mu{}_\nu + \delta z V^2\Box\omega^\mu{}_\nu \delta w V^\kappa\partial_\kappa\partial^\mu V_\nu + \delta t V^2\partial^\mu V_\nu \epsilon x\left(V^\mu V_\nu - \frac{V^\kappa\partial_\kappa}{\Box}V^\mu\partial_\nu\right) + \epsilon y\frac{V^\kappa\partial_\kappa}{\Box}V^\mu\partial_\nu\partial_\nu \\
&+ \epsilon z V^2 V^\mu + \epsilon w V^\kappa\partial_\kappa V^\mu V_\nu + \epsilon t V^2 V^\mu V_\nu = \theta^\mu{}_\nu + \omega^\mu{}_\nu \ ,
\end{aligned} \tag{4.106}$$

we get the following set of equations

$$\theta^\mu{}_\nu \to \alpha x = 1 \to x = \frac{1}{\alpha} \ , \tag{4.107}$$

$$\omega^\mu{}_\nu \to -\alpha z V^\kappa\partial_\kappa + \beta y + \beta z V^\kappa\partial_\kappa - \delta x V^\kappa\partial_\kappa + \delta y V^\kappa\partial_\kappa + \delta z V^2\Box = 1 \ , \tag{4.108}$$

$$V^\mu\partial_\nu \to \alpha z + \gamma y + \gamma z V^\kappa\partial_\kappa - \epsilon x\frac{V^\kappa\partial_\kappa}{\Box} + \epsilon y\frac{V^\kappa\partial_\kappa}{\Box} + \epsilon z V^2 = 0 \ , \tag{4.109}$$

$$V_\nu\partial^\mu \to -\alpha t\frac{V^\kappa\partial_\kappa}{\Box} + \beta w + \beta t\frac{V^\kappa\partial_\kappa}{\Box} + \delta x + \delta w V^\kappa\partial_\kappa + \delta t V^2 = 0 \ , \tag{4.110}$$

$$V^\mu V_\nu \to \alpha t + \gamma w\Box + \gamma t V^\kappa\partial_\kappa + \epsilon x + \epsilon w V^\kappa\partial_\kappa + \epsilon t V^2 = 0 \ . \tag{4.111}$$

We can further decompose them into two independent systems of linear equations

$$\begin{aligned}(\beta+\delta V^\kappa\partial_\kappa)y+(-\alpha V^\kappa\partial_\kappa+\beta V^\kappa\partial_\kappa+\delta V^2\Box)z&=1+\frac{\delta V^\kappa\partial_\kappa}{\alpha}\\(\gamma+\epsilon\frac{V^\kappa\partial_\kappa}{\Box})y+(\alpha+\gamma V^\kappa\partial_\kappa+\epsilon V^2)z&=\frac{\epsilon V^\kappa\partial_\kappa}{\alpha\Box}\ ,\end{aligned}\tag{4.112}$$

and

$$\begin{aligned}(\beta+\delta V^\kappa\partial_\kappa)w+(-\alpha\frac{V^\kappa\partial_\kappa}{\Box}+\beta\frac{V^\kappa\partial_\kappa}{\Box}+\delta V^2)t&=-\frac{\delta}{\alpha}\\(\gamma\Box+\epsilon V^\kappa\partial_\kappa)w+(\alpha+\gamma V^\kappa\partial_\kappa+\epsilon V^2)t&=-\frac{\epsilon}{\alpha}\ .\end{aligned}\tag{4.113}$$

We must then re-express them as matrices and find their determinant to move forward. We will skip the demonstration, and write the resulting determinant as

$$\det=\alpha\left(\beta+(\gamma+\delta)V^\kappa\partial_\kappa+\epsilon\frac{(V^\kappa\partial_\kappa)^2}{\Box}\right)+\left(\frac{\beta\epsilon}{\Box}-\gamma\delta\right)\left(V^2\Box-(V^\kappa\partial_\kappa)^2\right)\equiv\Delta\ ,\tag{4.114}$$

and finally getting

$$\begin{aligned}M^{-1}_{\mu\nu}=\frac{1}{\alpha}\theta_{\mu\nu}&+\frac{1}{\Delta}\left[\left(\alpha+\gamma V\cdot\partial+\epsilon V^2\right)\left(1+\frac{\delta}{\alpha}V\cdot\partial\right)+\left((\alpha-\beta)V\cdot\partial-\delta V^2\Box\right)\frac{\epsilon}{\alpha}\frac{V\cdot\partial}{\Box}\right]\omega_{\mu\nu}\\&+\frac{V_\mu\partial_\nu}{\Delta}\left[\left(-\gamma-\epsilon\frac{V\cdot\partial}{\Box}\right)\left(1+\frac{\delta}{\alpha}V\cdot\partial\right)+(\beta+\delta V\cdot\partial)\frac{\epsilon}{\alpha}\frac{V\cdot\partial}{\Box}\right]\\&+\frac{V_\nu\partial_\mu}{\Delta}\left[-\left(\alpha+\gamma V\cdot\partial+\epsilon V^2\right)\frac{\delta}{\alpha}-\left(\alpha\frac{V\cdot\partial}{\Box}-\beta\frac{V\cdot\partial}{\Box}-\delta V^2\right)\frac{\epsilon}{\alpha}\right]\\&+\frac{V_\mu V_\nu}{\Delta}\left[(\gamma\Box+\epsilon V\cdot\partial)\frac{\delta}{\alpha}-(\beta+\delta V\cdot\partial)\frac{\epsilon}{\alpha}\right].\end{aligned}\tag{4.115}$$

By now recalling the definition of our coefficients, we can see two distinct poles for the propagator: One on the transverse part

$$\frac{1}{\alpha}=\frac{1}{\Box-V^2}=\frac{1}{k^2-V^2}\ ,\tag{4.116}$$

which further confirms the transversality of the mass. We must however be careful with our backgrounds, as we can clearly see that for time-like backgrounds (not dealt here-in), the theory will be unbound from below and ghost ridden. Luckily, in the space-like case this problem is not present. The second pole is obtained from the longitudinal part

$$\frac{1}{\Delta}\ ,\tag{4.117}$$

with Δ

$$\Delta = (\Box - V^2)\left[-V^2 + \frac{(V\cdot\partial)^2}{\Box}\right] + (-\frac{V^2}{\Box} + 1)(V^2\Box - (V\cdot\partial)^2) = 0 \ , \tag{4.118}$$

showing that the longitudinal part is singular. This is a manifestation of the gauge invariance of the theory, and we must introduce a gauge fixing to solve this. In the action this reads

$$\mathcal{L}_{gf} = \frac{1}{2\xi}\left(\partial_\mu a'^\mu\right)^2 = -\frac{1}{\xi}a'^\mu\Box\omega_{\mu\nu}a'^\nu \ , \tag{4.119}$$

this in turn changes the β coefficient such as to have

$$\alpha = \Box - V^2, \quad \beta = -\frac{\Box}{\xi} - V^2, \quad \gamma = -\delta = \epsilon = 1 \ , \tag{4.120}$$

which finally gives us for Δ

$$\Delta = -\frac{1}{\xi}\left(\Box^2 + (V\cdot\partial)^2\right) \ . \tag{4.121}$$

We may now confidently state that there exist two poles

- One on the transverse part $\theta_{\mu\nu}$ by setting $\alpha = 0$ which gives

$$\alpha = 0 \Rightarrow k^2 = -V^2 = \vec{V}^2 \ . \tag{4.122}$$

- And a second on the longitudinal part $\omega_{\mu\nu}$ by solving $\Delta = 0$ which reads

$$k^2 - (V\cdot k) = k^2 + \vec{V}\cdot\vec{k} = 0 \ . \tag{4.123}$$

The presence of the massive pole on the transverse part is reminiscent of topologically induced masses [41, 44, 45], meaning that the mass generated by the background field has to be accommodated by the null helicity component of the spin-1 carried by the photon.

All of these considerations may raise questions as to our original Hamiltonian treatment. If the decomposition we used truly acted as a gauge fixing, it must have altered the constraint structure of the theory. Let us explore this possibility by considering the most general case, where V^μ can be both space-like and time-like to try and confirm the gauge nature of the model. Starting from

$$\mathcal{L}_0 = -\frac{1}{4}f'^2 + V_\mu a'_\nu \partial^\mu a'^\nu - V_\mu a'_\nu \partial^\nu a'^\mu - \frac{1}{2}V^2 a'^2 + \frac{1}{2}V_\mu V_\nu a'^\mu a'^\nu \ , \tag{4.124}$$

and breaking its manifest covariance

$$\mathcal{L}_0 = -\frac{1}{2}f_{0i}^2 - \frac{1}{4}f_{ij}^2 - (a_0 V_i - a_i V_0)f^{0i} - a_i V_j f^{ij} - \frac{1}{2}\left(a_0^2 V_i^2 + a_i^2 V^\mu V_\mu - 2a_i a_0 V^0 V^i - a_i a_j V^i V^j\right) \ , \tag{4.125}$$

before performing the Legendre transform to obtain the conjugate momenta

$$\pi^\mu = \frac{\partial \mathcal{L}_\prime}{\partial \dot{a}'_\mu} = \begin{cases} \pi^0 = \dfrac{\partial \mathcal{L}_0}{\partial \partial_0 a'_0} = 0 \\ \\ \pi^i = \dfrac{\partial \mathcal{L}_0}{\partial \partial_0 a'_i} = f'^{i0} - (a'^0 V^i - a'^i V^0) \end{cases} \ ,$$

where $\pi^0 = 0$ is the primary constraint, and will be relabelled as $\phi_1 = \pi^0 = 0$.

The total Hamiltonian will then read

$$\mathcal{H}_T = \int d^3\vec{x} \left[-a_0 \partial^i \pi_i - \frac{1}{2}\pi^i \pi_i - \pi\left(a_0 V_i - a_i V_0\right) + \frac{1}{4}f_{ij}f^{ij} + a_i V_j f^{ij} - \frac{1}{2}\left(\left(a_i V^i\right)^2 - \left(a_i V_j\right)^2\right) + \lambda_1 \phi_1 \right] \ . \tag{4.126}$$

From the expression of the Hamiltonian, we can already see that the primary and secondary constraint will commute due to the absence of a quadratic term in a_0. Nonetheless, we will proceed with the Dirac Bergmann algorithm [12, 13, 30] and get

$$\{\phi_1, \mathcal{H}_T\} = \partial^i \pi_i + \pi^i V_i \approx 0 \ , \tag{4.127}$$

which will be the modified Gauss law, relabelled as $\phi_2 = \partial_i \pi^i + \pi^i V_i \approx 0$. The Poisson bracket between the Gauss law and the Hamiltonian will read

$$\{\phi_2, \mathcal{H}_T\} = \{\partial_i \pi^i, \pi^i a_i V_0\} + \{\pi^i V_i, \pi^i a_i V_0\} \ , \tag{4.128}$$

where all of the other brackets cancel each other, leaving us with this combination as the only non trivial contribution. This reduces to

$$\begin{aligned} \{\phi_2, \mathcal{H}_T\} =& V_0 \partial_i \pi^i + \pi^i V_i V_0 \approx 0 \\ =& V_0 \phi_2 \ , \end{aligned} \tag{4.129}$$

showing that this does not generate an additional constraint as it can be expressed as

a previously known constraint, the theory is thus composed of two constraints ϕ_1, ϕ_2.

The classification of the constraints is trivial and will immediately give

$$\{\phi_1, \phi_2\} = 0 \ , \tag{4.130}$$

showing the presence of two first class constraints, confirming the gauge theoretical aspect of the theory. Additionally, this also confirms our earlier hunch, that fixing the totally symmetric and traceless part of the $V^\mu V^\nu$ tensor acted effectively as a gauge fixing.

The extended Hamiltonian, which includes all first class constraints, is however still dependent on two undefined Lagrange multipliers

$$\mathcal{H}_{ext} = \mathcal{H} + \int d^3\vec{x} \ \lambda_1\phi_1 + \lambda_2\phi_2 \ . \tag{4.131}$$

Let us then consider the evolution of the time-like polarisation

$$\{a'_0, \mathcal{H}_{ext}\} = \lambda_1 \ , \tag{4.132}$$

which confirms the arbitrariness of the scalar potential, allowing us to set it to zero. The vector potential a'^i is not fully defined through its evolution due to the remaining multiplier

$$\delta a'^i = \{a'^i, \int d^3\vec{x} \ \lambda_2(\partial_j\pi^j + \pi^j V_j)\} = -(\partial_i - V_i)\lambda_2 \ , \tag{4.133}$$

which is the expression of our local gauge transformation. This shows that the background vector field shifts the derivative acting on the fields by a constant value, letting us express the form of our covariant derivative as

$$D_\mu = \partial_\mu + V_\mu \ , \tag{4.134}$$

which can be further confirmed by verifying Noether's second theorem. Indeed, if the covariant divergence of the equation of motion vanishes identically then the corresponding covariant derivative defines a Noether identity and therefore generates a local gauge symmetry. Considering (4.60)

$$\mathcal{E}^\nu = \partial_\mu f'^{\mu\nu} + V^\nu\partial^\mu a'_\mu - V^\mu\partial^\nu a'_\mu - a'^\nu V^\mu V_\mu + a'^\mu V^\nu V_\mu \ \ , \tag{4.135}$$

we can see that

$$D_\nu \mathcal{E}^\nu = (\partial_\nu + V_\nu)\mathcal{E}^\nu = 0 \ . \tag{4.136}$$

The fact that the local gauge transformation is of the form discussed above (*i.e.* the usual U(1) transformation + constant parameter) is reassuring, as it implies that the algebra formed by the first class constraints still closes, ensuring the nilpotency of the BRST operator [11] and thus, the unitarity of the theory at the quantum level (as long of course as we stick with the space-like case for the background vector [105]). In fact, one could go further, and show that the algebra is still Abelian. Indeed, considering

$$\delta_\Lambda a^\mu = D^\mu \Lambda \ , \tag{4.137}$$

we can compute

$$[\delta_{\Lambda_1}, \delta_{\Lambda_2}]\, a^\mu = \delta_{\Lambda_1}\delta_{\Lambda_2} a^\mu - \delta_{\Lambda_2}\delta_{\Lambda_1} a^\mu \ , \tag{4.138}$$

keeping in mind that the gauge parameters are independent $\delta_{\Lambda_2}\Lambda_1 = 0$, it is trivial to show that the gauge algebra is Abelian

$$[\delta_{\Lambda_1}, \delta_{\Lambda_2}] = 0 \ , \tag{4.139}$$

exactly as Maxwell's case, thus not influencing the BRST structure of the model.

From the above, we confirm what we speculated in [101], which is that the field redefinition described in (4.14) is only a formal symmetry breaking, and only the manifest gauge invariance has been lost, in the same manner that a Higgs mechanism would act on the vacuum state.

In the early sections of this chapter, we made some simplifying assumptions that were two folds: first, we decided to ignore higher order contributions to the action, and second we decided to drop dual terms. With the mass generation effect we uncovered, we might be worried that these terms would lead to modulation, or worst, cancel the mass generated. We will try to give some arguments to reassure the reader that even in the worst case scenario, the mass can only be modified, but never erased by these contributions. Let us begin with D terms, and use for illustration D_1. We remind ourselves that it is defined as

$$D_1 = \left.\frac{\partial^2 \mathcal{L}}{\partial \mathcal{F}^2}\right|_B = \left.\frac{\partial}{\partial \mathcal{F}}\right|_B C_1 = \left.\frac{\partial}{\partial \mathcal{F}}\right|_B \alpha^2 \ . \tag{4.140}$$

This term will be effectively incorporated in the action coupled to

$$\begin{aligned}\delta\mathcal{L} &= \frac{1}{8}D_1 F^{\mu\nu}F^{\rho\sigma}f_{\mu\nu}f_{\rho\sigma} \\ &= \partial_{\mathcal{F}}\alpha^2 F^{\mu\nu}F^{\rho\sigma}f_{\mu\nu}f_{\rho\sigma}\ ,\end{aligned} \tag{4.141}$$

meaning that if there is any contribution to the mass, it should be coming from the field redefinition of $\partial_{\mathcal{F}}\ \alpha^2$ into the potentials, which will yield objects of dimensionality different to that of V^μ. This implies that while this contribution might shift or correct the massive pole, it cannot cancel it identically.

As for the C_2 terms, let us suppose that they remained, and that the first one would be written as

$$-\frac{1}{4}C_2 \tilde{f^{\mu\nu}}f_{\mu\nu} = -\frac{1}{4}\beta^2\tilde{f}f = -\frac{1}{4}\beta^2\varepsilon^{\mu\nu\rho\sigma}f_{\mu\nu}f_{\rho\sigma}\ . \tag{4.142}$$

If we proceed with the same field redefinition as a above, we will get

$$\beta^2\varepsilon^{\mu\nu\rho\sigma}f_{\mu\nu}f_{\rho\sigma} = \varepsilon^{\mu\nu\rho\sigma}f'_{\mu\nu}f'_{\rho\sigma} + 4\varepsilon^{\mu\nu\rho\sigma}a'_\mu B_\nu f'_{\rho\sigma} + 2\varepsilon^{\mu\nu\rho\sigma}\left(a'_\mu a'_\rho B_\nu B_\sigma - a'_\mu a'_\nu B_\rho B_\sigma\right)\ , \tag{4.143}$$

where $B^\nu = \dfrac{\partial^\nu\beta}{\beta}$ similarly to V^μ. We immediately see that the last term is identically zero, while the first one will be topological term. The only remaining component is the second, which is a Carroll-Field-Jackiw term, which may modulate the mass and dispersion relations, but cannot destroy the massive behaviour since our mass term is quadratic in the action.

As closing remarks, let us briefly discuss the mechanism behind the effective mass generation in our system. In modern physics, mass is taken to be an emergent phenomenon in the most general case. Masses can emerge through topological interactions as discussed for example in the Chern-Simons [41,44,45] section of this thesis, but much more commonly through which will spontaneously break one of the symmetries (gauge symmetry, chiral symmetry etc.) of a model, which in turn generates Goldstone Bosons which can be then "swallowed" by the particles to add a third polarisation (which can be longitudinal or transverse depending on the model) [14, 16, 18]. As such, it is generally accepted in physics that symmetry breaking is behind mass generation.

In this model, the field redefinition at play appears to be generating a mass, and it could be argued by some readers that such a redefinition, despite our arguments and despite the fact we proved the gauge invariance of the theory, is unjustified.

For the sake of the argument, consider the action (4.1). The coefficients coupling

to the photon fields are not constants, which means that these dielectric coefficients are functions of space-time (in our case, space more specifically). This in turn has some consequences on the vacuum state of the system, which will develop a different refractive index and break its own symmetry. As such, while the original action seemed to obey two $U(1)$ invariances (photon and background), it is reasonable to believe that due to the non-dynamical nature of the background, a spontaneous symmetry breaking hidden at the quantum level would emerge. By choosing to absorb the order zero background contribution into the potential 4-vector, we are indirectly showing this effect by proceeding to a manifest symmetry breaking, which allows us to describe classically the otherwise subtle effects of the background field. These mechanisms seem very close to those discussed in [93] which is coherent as the constant nature of V^μ indicates the presence of a Lorentz symmetry violation in our model.

The unanswered question here is considering the physicality of the mass. Indeed, we showed that classically, the theory is well behaved (from the dispersion relations, poles, and Hamiltonian in the gauge fixed version), but quantisation can introduce anomalies [11, 18] and might cast shadow on the massiveness of the model. The next step is to couple fermions to the theory, which might not be as trivial as it seems. Indeed, finding a fermionic representation might appear simple, but we have to keep in mind that if the effects of the background on photons are not trivial, they should also be non trivial on fermions. Once fully quantised, and with the proper Feynman rules in hand, we could finally state if our mass is physical, or simply an effective artifact that only holds at the classical level.

Chapter 5

Perspectives and Conclusions

5.1 Other research work

Some endeavours of different nature have been followed and led to publications. We refrained from mentioning them in the core text as these projects were rather adjacent to ETEM (Extended Theories of Electromagnetism) and/or represented applications in plasma physics, astrophysics and cosmology.

A good way to test ETEM is through the measurement of the Ampère-Maxwell law. In [68], we considered the deviations from such law generated by non-linearities or massive behaviours or Lorentz symmetry violations. We then analysed six years worth of data stemming from the Magnetospheric Multiscale Mission (MMS), which is a four-satellite constellation, crossing mostly turbulent regions of magnetic reconnection. The data analysis showed, for the highest time resolution burst data and optimal tetrahedron configurations drawn by the four spacecraft, deviations in a large percentage in the solar wind regime. But unfortunately this regime was not much explored by MMS. Despite our thorough analysis supported by the MMS experts, the deviations are likely due to unaccounted experimental errors. We have set upper limits on the photon mass m_γ of the order of 2.1×10^{-51}kg and on the LSV vector parameter $|\vec{k}|$ of $6 \times 10^{-9}\mathbf{m}$. The next step would be to explore these deviations considering calibration errors. We hope that future missions will provide better measurements to reduce to a minimum the probability of experimental errors, and at least put obtain better constraints on ETEM. Ideally, we will be able to confirm the presence of deviations, In any case the synergy between plasma and fundamental physics could be fruitful to both domains.

In chapter 3, we discuss the emergence of frequency shifts and drifts in non-constant backgrounds. One of the implications of such a mechanism would be that on large distances, there would be an additional static contribution to Doppler and expansion effects. This, in principle, should play an important role in astrophysical and cosmological measurements, and we are currently exploring it in an on-going paper. Keeping in mind that on larger scales, electric fields cannot exist (quasi-neutrality hypothesis) and the general form of frequency shifts for arbitrary non-linear models, we manage to propose an *Ansatz* that produces frequency drifts based on photon-background magnetic fields interaction

$$z^{\mathrm{S}}_{\mathrm{ETEM}} = -\alpha \int_{\mathrm{observer}}^{\mathrm{emitter}} \vec{B} \cdot d\vec{l}_{\mathrm{los}} \ , \tag{5.1}$$

where α is a free parameter, and B is the background magnetic field, and $d\vec{l}_{\mathrm{los}}$ is the path run by the photon. We suppose that this interaction is responsible for the excess Doppler recorded in the stellar and gas rotation curves in galaxies, opposed to the

excess Doppler perceived as a real excess rotational velocity due to Dark Matter (DM) or to Extended Theories of Gravity (ETG). In absence of theoretical and experimental proof of the existence of DM and in the presence of an overly successful theory of general relativity, we postulate that the excess Doppler is not determined by an excess rotational velocity. Indeed, by using the magnetic profile of the Milky way, we produce the Doppler shift contribution due to ETEM on the scale of the galaxy, and show that when summed to the other contributions, it reproduces to a very satisfactory degree the rotation curves of the Milky way with five different data sets. We manage to determine the value of the coupling "magnetic field - photon path" α parameter in the order of 7×10^{-15}. The feebleness of this coupling explains that it went undetected in Earth laboratories which size cannot compare with galactic distances, despite the much larger magnetic field, available on Earth.

By making the coupling even smaller, we could perfectly also consider cases where ETEM shifts superpose to DM and/or ETG contributions. In the paper under work, we are not oblivious that DM has become a cornerstone in modern cosmology and address how ETEM could affect gravitational lensing and the Cosmic Microwave Background. This effort pursed a similar approach taken by my thesis director on dark energy.

So far in the text, when we refer to ETEM, we limit ourselves to classical/quantum field theories. However, gravitational effects should not be neglected in a proper theory of electromagnetism, especially if one has to discuss cosmological and astrophysical implications. From a fundamental point of view, gravitation and electromagnetism are not only the sole interactions with measurable macroscopic effects, but also both considered to be of infinite range and described by free, massless particles. Both are gauge theories (in different senses of course), and both can be described through curvatures and connections.

In this view, we attempted to look into the gravitational self-force acting on a photon [106]. The starting point of this endeavour is the Aichelburg-Sexl [107, 108] metric, which describes the gravitational field acting on a massless particle boosted to the speed of light. This is a suitable approximation, if we consider that photons can be described as massless point particles. But, in their original work, Aichelburg and Sexl only consider radial motion, and neglected angular contributions, which is crucial in astrophysical and theoretical contexts, such as particle collisions with non-zero impact parameters and ultra-relativistic scattering in curved spacetime. To deal with the most general situation, we then generalise the theory to add angular velocities, before deriving the black hole perturbations theory and the Regge-Wheeler-Zerilli equations [109, 110]. We then conclude the work by discussing the self-force [111,112], proposing a conceptual effort to define geodesic deviations for massless particles as effective frequency shifts.

Continuing on the gravitational side of things, ETG predict new behaviours on large

scales, but are very often coupled rather naïvely to Maxwell's theory of electromagnetism to make observational predictions on electromagnetic emissions. In [113], we make the first step for the inclusion of ETEM in conjunction with ETG. Within the framework of $f(R)$ gravity [114], we couple minimally the electromagnetic curvature invariant $f(\mathcal{F})$ to ensure a consistent treatment of the ensuing interactions. We show that this reduces properly to known ETEM in flat space-time, notably the Bopp-Podolsky theory. We also argue that these extensions should have phenomenological consequences in the early universe, or near charged compact objects. This paper is the first step into a larger project, which consists on dealing with both ETEM and ETG, a framework we will refer to as Extended Theories of Long Range Interactions (ETLRI).

5.2 Perspectives

Despite our best attempts, this thesis is not fully representative of what ETEM are, and can be. Indeed, the work we presented is only limited to the classical aspects of said models, and in the case of GNLED, only limited to constant/or non-dynamical backgrounds. In the near future, we intend to study the quantum mechanical picture tied to the mass generation mechanism described in chapter 4, as well as the renormalisation features of such a model. The idea is to rule with certainty on the physicality of the mass, and this can only be achieved by seeing if the massive pole does not get erased at any order in the perturbation series.

On a different front, we will look into the possible inclusion of a de Broglie-Proca mass term in the Bopp-Landé-Thomas-Podolsky theories, to see whether both photon modes can be accommodated to be massive, as well as the stability conditions.

The larger project is to develop the framework of ETLRI. Doing it in a systematic way is currently impossible, but the first step we will partake is the inclusion of torsion to Maxwell's theory. The idea would be to consider Einstein-Cartan spaces in the metrically flat limit coupled to Maxwell, and see whether the symmetry breaking induced by the torsion leads to an unstable phase space. The procedure would then be extended to different types of torsion, and possibly beyond Einstein-Cartan with non-metric theory such as metric affine gravity as well as ETEM. A systematic study of the constraints, and their potential reducible symmetries could then allow us to understand better how (non)minimal coupling to electrodynamics could affect ETG and constrain further more their dynamics. The goal is two folds: one, is to see whether torsion coupling can generate an acceptable mass for electromagnetism, and two to derive the phenomenological and observational implications on astrophysical scales linked to such interactions.

An ambitious idea would be to consider a wide class of extended theories, of either gravitation or electrodynamics, and devise a systematic classification process based

solely on their constraint algebra. The idea being that through their degrees of freedom, and exact constrained Hamiltonian structure, we could construct equivalence classes that would reduce the redundancies that are unavoidable in the usual model building approach. In an ideal world, this would lead us to a unified way of dealing with extended gravitation and electromagnetism, simplifying the path to phenomenology and observations.

Finally, the classic interaction of an electromagnetic background fields with a light wave, becomes at quantum level the interaction among photons and possibly photon and charges. This interaction prompted the consideration that an effect of blurred images would place. This longstanding reserve could be nullified when adopting ETEM for reading long distance sources.

5.3 Conclusion

In this thesis we explored the consequences of background fields on the propagation of electromagnetic waves. We notably showed that in the case of a generic class of non-linear electromagnetism, the presence of non-constant background fields induces violations of the energy-momentum tensor and by extension frequency shifts [97]. We also demonstrated that within the same category of models, the aforementioned background-photon interactions will lead to the generation of an effective mass, which appears to be classically healthy [101].

We are now arriving at the end of this adventure and it is time to conclude. In chapter 1, we introduced the fundamental ideas and structures tied to Maxwell's theory, to derive a list of all the ingredients of electromagnetism.

In chapter 2, we introduced the first modifications to Maxwell's theory by choosing which ingredients to keep from the original list. This lead us to introduce the massive, non-linear and Lorentz violating extensions, and their theoretical structure, closing on the table of extended theories of electromagnetism with their composition in terms of the ingredients derived in chapter 1.

In chapter 3, we presented our work on Generalised Non-Linear Electromagnetism, a generic framework to deal with the propagation of photons in backgrounds, showing the presence of frequency shifts as well as the propagating features of such family of models.

In chapter 4, we expanded upon the work presented in chapter 3 to show that within that same generic framework, it is possible to generate a mass for photons. We showed both a gauge fixed, and general approach, where both agreed that the massive polarisation has to be transverse, a key feature of topologically massive models. We derived the Hamiltonian picture and argued on the stability conditions of the theory,

concluding that the model does not seem to exhibit tachyonic behaviour, and that it should be unitary.

In summary, extended theories of electromagnetism offer a rich framework, both conceptually, theoretically and observationally. Conceptually, electromagnetism being one of the jewels of physics makes it taken for granted, and pushes us to ignore many of its structural subtleties. Breaking away from the consensus allows us to imagine different paths that lead to different effects, and out of the box solutions to open problems in modern physics, such as the darkness of the universe. Theoretically, the tools we must use and develop to deal with these frameworks can and will be used in other cases, enriching the field of physics. Observationally, the possibility to not only, measure new effects, new constraints and new physics, but to also re-interpret the enormous amount of astrophysical and cosmological data amassed across the years makes the field ripe with possibilities. While the field of ETEM is still niche, I truly believe that it is bound to grow in popularity, and will soon find its place in the landscape of theoretical physics.

Appendices

Appendix A

Field Theory in a nutshell

In the thesis it is assumed that the reader is somewhat familiar with field theory and by extension analytical mechanics. This appendix aims to briefly introduce the conceptual steps that generalises point systems to fields. In analytical mechanics, the evolution of systems is described by a trajectory derived through the least action principle, that requires the definition of an integral action S that depends on all possible trajectories. Physical trajectories however will be those that leave said action stationary $\delta S = 0$. Let us define the action S as

$$S = \int dt\ L(q_i(t), \dot{q}_i(t)), \tag{A.1}$$

where $L(q_i, \dot{q}_i)$ is the Lagrangian, which is a functional of q_i and $\dot{q}_i$, respectively the generalised coordinate and the velocity deriving from the generalised coordinate. For systems that depend on equations of motion of higher order than second, it will be necessary to add a dependency on $\ddot{q}_i$, $\dddot{q}_i$... so on so forth. This will describe a physical system of N degrees of freedom, evolving in what is known as configuration space [9].

The equations of motion are obtained through the variation of the action

$$\delta S = \int dt\ \delta L(q_i, \dot{q}_i) = 0, \tag{A.2}$$

which leads to the well known Euler-Lagrange equations

$$\frac{\partial L(t)}{\partial q_i(t)} - \frac{d}{dt}\frac{\partial L(t)}{\partial \dot{q}_i(t)} = 0\ , \tag{A.3}$$

for $1 \leq i \leq N$. In physics, it usually accepted that the Lagrangian (which is fundamen-

tally an abstract mathematical function) can be expressed as

$$L(t) = T - V \ , \tag{A.4}$$

where T is the kinetic energy and V the potential energy. There exist also an equivalent picture, that is a function of the total energy of the system, known as Hamiltonian dynamics, which has the advantage of producing first order differential equations. To move from the Lagrangian prescription to the Hamiltonian one, we have to move from a configuration space to a phase space. We start by stating that the Hamiltonian is defined as

$$H(t) = T + V \ , \tag{A.5}$$

and equivalently

$$H(t) = \sum_{i}^{N} p_i \dot{q}_i - L(t) \ . \tag{A.6}$$

where p_i is the conjugate momenta. The passage from the configuration space to the phase space is through the Legendre transform, which will define the conjugate momenta

$$p_i(t) = \frac{\partial L(t)}{\partial \dot{q}_i(t)} \ . \tag{A.7}$$

The variation of the Hamiltonian with respect to $q_i(t)$ and $p_i(t)$ will lead to $2N$ independent evolution equations known as the Hamilton Jacobi equations

$$\dot{q}_i(t) = \frac{\partial H(q_i, p_i)}{\partial p_i} \tag{A.8}$$

$$\dot{p}_i(t) = \frac{\partial H(q_i, p_i)}{\partial q_i}. \tag{A.9}$$

These two pictures however can only describe systems of finite degrees of freedom, and are thus unable to handle the infinite dimensional physical systems that we deal with in fundamental physics. The formalism must then be expanded to take these into account, as well as respecting the symmetries of the Poincaré group. This is what is known as field theory [14]. These fields can be considered as the continuum of the degrees of freedom $q_i(t)$ through the following mapping

$$q_i(t) = q(t, i) \rightarrow \varphi(t, \vec{x}) = \varphi(x) \ . \tag{A.10}$$

where $\varphi(x)$ is a functional of the space-time position x. Additionally, our usual deriva-

tives have to be generalised to functional derivatives such as

$$\frac{\partial q_i(t)}{\partial q_j(t)} = \delta_{ij} \rightarrow \frac{\delta\varphi(t,\vec{x})}{\delta\varphi(t,\vec{y})} = \frac{\delta\varphi(x)}{\delta\varphi(y)} = \delta^4(x-y) \ , \tag{A.11}$$

where $\delta^4(x-y)$ is the Dirac distribution. Since we are now describing the evolution of systems through the evolution of fields that depend on space-time variables, we have to consider instead of the Lagrangian $L(q_i(t), \dot{q}_i(t)$, the Lagrangian density $\mathcal{L}(\varphi(x), \partial_\mu\varphi(x))$ such as

$$S = \int dt \ L = \int d^4x \ \mathcal{L} \ . \tag{A.12}$$

The equations of motion are obtained by applying the least action principle, and taking the functional derivative of the action with respect to the fields [18] that compose the Lagrangian density. In a similar fashion, the Hamiltonian picture will also be generalised by modifying the definition of the Legendre transform and by considering Hamiltonian densities rather than Hamiltonian. The field theoretic Legendre transform will then be

$$p_i(t) = \frac{\partial L(t)}{\partial \dot{q}_i(t)} \rightarrow \pi_n(x) = \frac{\delta L(t)}{\delta \dot{\varphi}_n(x)} = \frac{\partial \mathcal{L}(x)}{\partial \dot{\varphi}_n(t)} \ , \tag{A.13}$$

leading to the Hamiltonian being defined as

$$\mathcal{H} = \int d^3\vec{x} \ \pi\dot{\varphi} - \mathcal{L} \ . \tag{A.14}$$

The equations of motion can also be derived from the least action principle, and the Hamilton-Jacobi equations can be obtained straightforwardly by following the mapping offered here. We also define the time evolution through the Poisson Bracket as

$$\frac{dF(t)}{dt} = \{F(t), \mathcal{H}(t)\} = \int d^3\vec{x} \left(\frac{\delta F(t)}{\delta\varphi_n(x)} \frac{\delta\mathcal{H}(t)}{\delta\pi_n(x)} - \frac{\delta F(t)}{\delta\pi_n(t)} \frac{\delta\mathcal{H}(t)}{\delta\varphi_n(t)} \right) \ , \tag{A.15}$$

which is antisymmetric, linear and obeys both the Leibniz rule and Jacobi identity [11].

Appendix B

Constrained Hamiltonian Systems

In this appendix we will briefly introduce the notion of Constrained Hamiltonian systems and the Dirac classification of constraints. Constraints will always arise when the Legendre transform is singular. Let us consider an N dimensional configuration space of variables $q_i, \dot{q}_i$. The Legendre transform will replace the velocities $\dot{q}$ with the independent conjugate momenta p_i, demanding that the phase space be $2N$ dimensional. A singular Legendre transform means that some positions are not independent from their conjugate momenta, yielding a phase space of dimension $N' < 2N$.

We shall now consider the case where the Legendre transform yields M constraints

$$p_i(t) = \frac{\partial L}{\partial \dot{q}_i} \Rightarrow \phi_m(q_i, p_i) = 0 \ , \quad 0 \leq m \leq M \tag{B.1}$$

also known as primary constraints. To obtain the variation of the Hamiltonian $H(q_i, p_i) = \sum p_i \dot{q}_i - L$ obeying the constraints $\phi_m = 0$ we need to consider the variations of the unconstrained Hamiltonian

$$\left(\frac{\partial H}{\partial q_i} + \dot{p}_i\right) \delta q_i + \left(\frac{\partial H}{\partial p_i} - \dot{q}_i\right) \delta p_i = 0 \ , \tag{B.2}$$

and the variation of the constraints

$$\frac{\partial \phi_m}{\partial q_i} \delta q_i + \frac{\partial \phi_m}{\partial p_i} \delta p_i = 0 \ , \tag{B.3}$$

which shows that the vector of components $\left(\frac{\partial H}{\partial q_i} + \dot{p}_i\right)$ and $\left(\frac{\partial H}{\partial p_i} - \dot{q}_i\right)$ and the M vectors of components $(\frac{\partial \phi_m}{\partial q_i})$ and $(\frac{\partial \phi_m}{\partial p_i})$ are parallel since they are both perpendicular to the

same vector of components $\delta q_i, \delta p_i$, and as such, of proportional components

$$\left(\frac{\partial H}{\partial q_i} + \dot{p}_i\right) = -\sum_{m=1}^{M} \lambda_m \frac{\partial \phi_m}{\partial q_i} \ , \tag{B.4}$$

$$\left(\frac{\partial H}{\partial p_i} - \dot{q}_i\right) = -\sum_{m=1}^{M} \lambda_m \frac{\partial \phi_m}{\partial p_i} \ , \tag{B.5}$$

where the coefficients λ_m are undefined.

From the previous two equations, we can get the Hamilton-Jacobi equations in the presence of constraints

$$\dot{q}_i = \frac{\partial H}{\partial p_i} + \sum_{m=1}^{M} \lambda_m \frac{\partial \phi_m}{\partial p_i} = \{q_i(t), H(t)\} + \sum_{m=1}^{M} \lambda_m \{q_i(t), \phi_m(t)\} \ , \tag{B.6}$$

$$\dot{p}_i = -\left(\frac{\partial H}{\partial q_i} + \sum_{m=1}^{M} \lambda_m \frac{\partial \phi_m}{\partial q_i}\right) = \{p_i(t), H(t)\} + \sum_{m=1}^{M} \lambda_m \{p_i(t), \phi_m(t)\} \ , \tag{B.7}$$

which for an arbitrary function $f(q_i, p_i)$ will give

$$\frac{d}{dt} f(q_i, p_i) = \{f(t), H(t)\} + \sum_{m=1}^{M} \lambda_m \{f(t), \phi_m(t)\} \ . \tag{B.8}$$

To generalise further, we introduce the notion of weak equality $\approx$, which is an equality modulo all the constraints acting on the theory

$$\frac{d}{dt} f(q_i, p_i) = \{f(t), H_T(t)\} = \{f(t), H(t)\} + \sum_{m=1}^{M} \{f(t), \lambda_m\}\phi_m + \sum_{m=1}^{M} \lambda_m \{f(t), \phi_m(t)\} \tag{B.9}$$

$$\approx \{f(t), H(t)\} + \sum_{m=1}^{M} \lambda_m \{f(t), \phi_m(t)\} \ , \tag{B.10}$$

where the term in $\{f(t), \lambda_m\}$ was dropped since it multiplies a constraint. Regardless, it would be meaningless since the parameters λ_m are totally arbitrary. We may notice that we have used a modified Hamiltonian denoted H_T, which we will define as the total Hamiltonian, and written as

$$H_T = H + \sum_{m}^{M} \lambda_m \phi_m \ . \tag{B.11}$$

We will then require that the constraints be constants of movement, and as such

should not evolve in time. However, writing it as a strong equality wouldn't make much sense, as we should consider all the possible constraints that will reduce our dynamics. As such we state it explicitly in the form of a weak equality

$$\dot{\phi}_n \approx 0 \Rightarrow \{\phi_n, H_T\} = \{\phi_n, H\} + \sum_m^M \lambda_m \{\phi_n, \phi_m\} \approx 0 \ , \tag{B.12}$$

If the above equations leads to a situation where there is an explicit contradiction, we may reject the form of the Lagrangian as abhorrent. Excluding all such cases, the equation can lead to three scenarii:

- 0=0, a trivial result that states that the constraint is coherent.
- It fixes λ_m as a function of the phase space variables.
- It yields a secondary constraint[1] $\chi(q_i, p_i) \approx 0$.

By imposing the consistency condition on all the constraints, we end up with M primary constraints, K secondary constraints and a number of equations that determine part, or all the coefficients λ_m as functions of the phase space. This leads to $M + K = L$ constraints φ_l satisfying the L equations

$$\dot{\varphi}_l = \{\varphi_l, H\} + \sum_m^M \lambda_m \{\varphi_l, \phi_m\} \approx 0 \ , \tag{B.13}$$

where m implies a sum on primary constraints. These L equations can be viewed as a system of linear inhomogeneous equation of unknown μ_m to which the solutions are $\mu_m = U_m(q_i, p_i)$.

If in addition to the solution μ_m, there exists A solutions V_{am} with $1 \leq a \leq A$ of the homogeneous equation

$$V_{am}\{\varphi_l, \phi_m\} = 0 \ , \tag{B.14}$$

then we can write μ_m as a superposition of both

$$\mu_m = U_m(q_i, p_i) + \sum_{a=1}^A v_a V_{am}(q_i, p_i) \ , \tag{B.15}$$

where v_a are arbitrary constants. With this solution the total Hamiltonian can be

[1] The distinction between primary and secondary constraints is simply on the origin of said constraint: the Legendre transform or the equations of motion.

rewritten as

$$\begin{aligned} H_T =& H(q_i, p_i) + \sum_{m=1}^{M} U_m(q_i, p_i)\phi_m(q_i, p_i) + \sum_{a=1}^{A} v_a \Phi_a(p_i, q_i) \\ =& H'(q_i, p_i) + \sum_{a=1}^{A} v_a \Phi_a(p_i, q_i) \ , \end{aligned} \tag{B.16}$$

where

$$\Phi_a = \sum_{m=1}^{M} V_{am}\phi_m \ , \tag{B.17}$$

are linear combinations of constraints and

$$H'(q_i, p_i) = H(q_i, p_i) + \sum_{m=1}^{M} U_m(q_i, p_i)\phi_m(q_i, p_i) \ , \tag{B.18}$$

is the part of the Hamiltonian that does not depend on arbitrary functions.

B.0.1 Classification of constraints

The total Hamiltonian is by definition the generator of time evolution. Now, if said Hamiltonian had to depend on arbitrary parameters, the time evolution of the system would also be arbitrary, regardless of initial conditions $q_i(t_0)$ and $p_i(t_0)$ at time t_0

$$\begin{aligned} q_i(t_0 + \delta t) =& q_i(t_0) + \delta \left.\frac{d}{dt} q(t)\right|_{t_0} \\ =& q_i(t_0) + \delta t \{q_i(t), H_T(q_i, p_i)\}_{t=t_0} \\ =& q_i(t_0) + \delta t \{q_i, H'(q_i, p_i)\}_{t=t_0} + \sum_{a=1}^{A} \delta t v_a \{q_i, \Phi_a(q_i, p_i)\} \ , \end{aligned} \tag{B.19}$$

which contained a determined part, and an arbitrary part since it depends on the v_a parameters. By construction, the Poisson brackets of the constraints Φ_a weakly vanishes with all the constraints φ_l

$$\{\varphi_l, \Phi_a\} = \sum_{m=1}^{M} \{\varphi_l, V_{am}\}\phi_m + \sum_{m=1}^{M} V_{am}\{\varphi_l, \phi_m\} = \sum_{m=1}^{M} \{\varphi_l, V_{am}\}\phi_m \ , \tag{B.20}$$

where we have used (B.14). This leads us to the following definition: A function F of the phase space is considered to be first class if

$$\{F, \varphi_l\} \approx 0 \ , \tag{B.21}$$

otherwise it is referred to as second class. Φ_a is a first class function, and by construction so is H'. As long as the total Hamiltonian depends on arbitrary v_a parameters, for other values v'_a of said parameters, the evolution of $q'_i(t_0 + \delta t)$ will be given as

$$q_i(t_0 + \delta t) = q_i(t_0) + \delta t\{q_i, H'(q_i, p_i)\}_{t=t_0} + \sum_{a=1}^{A} \delta t v'_a \{q_i, \Phi_a(q_i, p_i)\} \ , \tag{B.22}$$

yielding for the difference Δq_i

$$\Delta q_i(t) = \sum_{a=1}^{A} \delta t(v_a - v'_a)\{q_i(t), \Phi_a(t)\}_{t=t_0} \ , \tag{B.23}$$

and a similar result for Δp_i by replacing q_i by p_i. This shows that Φ_a are the generators of transformations that contribute arbitrarily to the evolution of a system without influencing its physical state. This is the very definition of a gauge transformation. We shall define $\varepsilon_a = \delta t\,(v_a - v'_a)$ as the parameters of said transformations that are both arbitrary due to v_a , v'_a and infinitesimal due to δt.

Supposing that the system undergoes two transformations in a row: a transformation generated by Φ_a of parameter ε_a followed by a transformation generated by Φ_b of parameter γ_b, we get

$$\begin{aligned} F \to F' &= F + \varepsilon_a\{F, \Phi_a\} \to F^{"} = F' + \gamma_b\{F', \Phi_b\} \\ &= F + \varepsilon_a\{F, \Phi_a\} + \gamma_b\{F, \Phi_b\} + \gamma_b\varepsilon_a\{\{F, \Phi_a\}, \Phi_b\} \ , \end{aligned} \tag{B.24}$$

while a transformation generated by Φ_b followed by a transformation generated by Φ_a will be

$$F + \gamma_b\{F, \Phi_b\} + \varepsilon_a\{F, \Phi_a\} + \varepsilon_a\gamma_b\{\{F, \Phi_b\}, \Phi_a\} \ . \tag{B.25}$$

Their difference (in other words, their commutator) will read

$$\begin{aligned} \Delta F &= \gamma_b\varepsilon_a\{\{F, \Phi_a\}, \Phi_b\} - \varepsilon_a\gamma_b\{\{F, \Phi_b\}, \Phi_a\} \\ &= \gamma_b\varepsilon_a(\{\{F, \Phi_a\}, \Phi_b\} + \{\{\Phi_b, F\}, \Phi_a\}) \ , \end{aligned} \tag{B.26}$$

which after using the Jacobi identity can be rewritten as

$$\Delta F = \varepsilon_a \gamma_b \{F, \{\Phi_a, \Phi_b\}\} \ . \tag{B.27}$$

Additionally, since Φ_a and Φ_b are first class constraints, their Poisson bracket is also a first class constraint

$$\{\Phi_a, \Phi_b\} = \epsilon_{abc} \Phi_c \ , \tag{B.28}$$

allowing to deduce that the infinitesimal transformations generated by the constraints form a closed algebra

$$\Delta F = \epsilon_{abc} \varepsilon_a \gamma_b \Phi_c \ . \tag{B.29}$$

B.0.2 Dirac Bracket

It is possible that a combination of second class constraints reduces to a first class constraint. We then look for all possible combinations to maximise the number of first class constraints. Whatever remains will form the full ensemble of second class constraints

$$\chi_s(q_i, p_i) \approx 0 \ . \tag{B.30}$$

These constraints are problematic for quantisation. Indeed, in quantum mechanical systems, the phase space variables q_i and p_i become operators $\hat{q}_i$ and $\hat{p}_i$, and their Poisson bracket becomes a commutator between said operators. This in turn changes the fundamental Poisson brackets of the phase space into the canonical commutators

$$[\hat{q}_i, \hat{p}_i] = i\delta_{ij} \ , \quad [\hat{q}_i, \hat{q}_j] = 0 \ , \quad [\hat{p}_i, \hat{p}_j] = 0 \ .$$

This is nothing but the correspondence principle in its simplest expression. These operators will then act on the states $|\psi\rangle$ of the Hilbert space $\mathcal{H}$, and as such constraints cannot be considered as relations between $\hat{p}$ and $\hat{q}$ but as operators that cancel the physical states of the Hilbert space. More precisely, in analytical mechanics, we consider that amongst the solutions to Hamilton's equations, physical solutions are those that verify the constraints $\varphi_l(q_i, p_i) \approx 0$. When quantised, these constraints will become operators acting on $\mathcal{H}$ by selecting the physical states as those verifying $\hat{\varphi}_l |\psi\rangle = 0$.

This works perfectly fine for first class constraints since the classical relation $\{\Phi_a, \Phi_b\} =$

$\epsilon_{abc}\Phi_c$ leads to its quantum counterpart

$$\left[\hat{\Phi}_a, \hat{\Phi}_b\right] = i\epsilon_{abc}\hat{\Phi}_c \left|\psi\right\rangle \ , \tag{B.31}$$

which for physical states will give $0 = 0$. However, second class constraints that classically verify

$$\{\chi_s, \chi_p\} = K_{sp} \neq 0 \ , \tag{B.32}$$

lead to a contradiction at the quantum level

$$[\hat{\chi}_s, \hat{\chi}_p] \left|\psi\right\rangle = 0 \rightarrow \hat{K}_{sp} \left|\psi\right\rangle = 0 \ . \tag{B.33}$$

To avoid this contradiction, Dirac proposed to replace the Poisson Bracket by another bracket which will introduce second class constraints in its very definition. The starting point is Dirac's theorem, which we will not demonstrate [11], stating that the determinant of the $S \times S$ matrix composed of the Poisson bracket of second class constraints

$$\{\chi_s(q_i, p_i), \chi_{s'}(q_i, p_i)\} \ , \tag{B.34}$$

is different than zero. This ensures the invertibility of the matrix

$$\det\{\chi_s, \chi_{s'}\} \neq 0 \Rightarrow \ \sum_{s''}^{S} \{\chi_s, \chi_{s''}\}^{-1} \{\chi_{s''}, \chi_{s'}\} = \delta_{s,s''} \ , \tag{B.35}$$

and that it is of even dimension. From the transpose

$$\{\chi_s, \chi_{s'}\}^T = \{\chi_{s'}, \chi_s\} = -\{\chi_s, \chi_{s'}\} \ , \tag{B.36}$$

we get

$$\det\{\chi_s, \chi_{s'}\} = \det\{\chi_s, \chi_{s'}\}^T = (-1)^S \det\det\{\chi_s, \chi_{s'}\} \ , \tag{B.37}$$

which shows that $(-1)^S = 1$, implying that second class constraints must always be an even number. For functions of the phase space f and g, the Dirac bracket is defined through the Poisson bracket and the second class constraints as

$$\{f(t), g(t)\}_D = \{f(t), g(t)\} - \sum_{s,s'}^{S} \{f(t), \chi_s(t)\} \{\chi_s(t), \chi_{s'}(t)\}^{-1} \{\chi_{s'}, g(t)\} \ . \tag{B.38}$$

The Dirac bracket shares all of the properties of the Poisson Bracket, with one notable addition: the Dirac bracket of a second class constraint with any function of the phase space vanishes, turning the second class constraints into strong equalities, allowing us to solve them and reduce the degrees of freedom of the phase space.

Appendix C

Green's functions and Propagators

In this Appendix we will introduce very succinctly the notion of Green's function. Consider the following

$$\left(\Box + m^2\right)\psi(x) = j(x) \ , \tag{C.1}$$

which is nothing more than the Klein-Gordon equation with source term $j(x)$ [16, 18]. Green's functions are the solutions to the above equation for a localised (or point-like) source

$$\left(\Box + m^2\right)G(x-y) = \delta^4(x-y) \ . \tag{C.2}$$

The general solution for (C.1) will be of the form

$$\psi(x) = \psi_0(x) + \int d^4y \ G(x-y)j(y) \ , \tag{C.3}$$

where $\psi_0(x)$ is the solution to the vacuum Klein-Gordon equation

$$(\Box + m^2)\psi_0(x) = 0 \ . \tag{C.4}$$

To solve equation (C.2), which is a differential equation of constant coefficients, we will make use of the Fourier transform $\tilde{G}(p)$ [4] of $G(x-y)$

$$G(x-y) = \int \frac{d^4p}{(2\pi)^4}\tilde{G}(p)e^{-ip\,\cdot\,(x-y)} \ . \tag{C.5}$$

By then using the Fourier transform of the Dirac distribution

$$\delta^4(x-y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \ , \tag{C.6}$$

we can deduce from (C.2) that

$$\begin{aligned}
&\int \frac{d^4p}{(2\pi)^4}(\Box + m^2)\tilde{G}(p)e^{-ip\cdot(x-y)} = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \\
&\int \frac{d^4p}{(2\pi)^4}(-p^2 + m^2)\tilde{G}(p)e^{-ip\cdot(x-y)} = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \\
&\Rightarrow \tilde{G}(p) = \frac{-1}{p^2 - m^2} \ ,
\end{aligned} \tag{C.7}$$

and by extension that

$$G(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{-1}{p^2-m^2} e^{-ip\cdot(x-y)} \ . \tag{C.8}$$

This expression is divergent for

$$p^2 = p_0^2 - \vec{p}\cdot\vec{p} = m^2 \Rightarrow p_0 = \pm\omega_{\vec{p}} \ , \tag{C.9}$$

where

$$\omega_{\vec{p}} = \sqrt{\vec{p}\cdot\vec{p} + m^2} \ , \tag{C.10}$$

are the two poles of the integral

$$-\int \frac{d^3\vec{p}}{(2\pi)^3 2\omega_{\vec{p}}} e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \ \times \int \frac{dp_0}{2\pi}\left(\frac{1}{p_0 - \omega_{\vec{p}}} - \frac{1}{p_0 + \omega_{\vec{p}}}\right) e^{-ip_0\cdot(x_0-y_0)} \ . \tag{C.11}$$

The integral on p_0 is evaluated through the Residue theorem by considering p_0 to be complex valued. Green's functions will thus depend on how we go around the poles [14, 18].

Publications list

- F. Bajardi, M. Benetti, S. Capozziello, A. Dib, "Extended theories of electrodynamics in $f(R)$ gravity." [113]
- A. Dib, J. A. Helayël-Neto and A. D. A. M. Spallicci, "Generalised Non-Linear Electrodynamics: classical picture and effective mass generation." [101]
- A. Dib, A. Garnier, A.D.A.M. Spallicci, "Generalized Aichelburg–Sexl and self-force for photons." [106]
- A.D.A.M. Spallicci, A. Dib and J. A. Helayël-Neto, "Photon frequency variation in non-linear electro-magnetism." [97]
- A.D.A.M. Spallicci, G. Sarracino, O. Randriamboarison, J. A. Helayël-Neto and A. Dib, "Testing the Ampère–Maxwell law on the photon mass and Lorentz symmetry violation with MMS multi-spacecraft data." [68]

List of Figures

1.1 Coulomb's original setup 12
1.2 Leading order contributions to the Bhabha scattering. 39
1.3 Fermion's self-energy 39
1.4 Vacuum polarisation diagram 41
1.5 Triangle diagram 42

2.1 Energy scales at which symmetries are expected to break. 69

List of Tables

1.1 Feynman's rules for Quantum Electrodynamics in momentum space and in the Feynman gauge $\xi = 1$ 38

2.1 Photon mass upper limits below $10^{-50}\mathbf{kg}$. 63
2.2 Standard Model of particle physics 68
2.3 All known extensions of electrodynamics and their corresponding features/violations 75

Bibliography

[1] C. Maxwell. On physical lines of force. *London Edinburgh Dublin Phys. Mag. and J. .Sci.*, 21:161, 1861.

[2] C. Maxwell. A dynamical theory of the electromagnetic field. *Phil. Trans. R. Soc.*, 155459–512, 1865.

[3] O. Heaviside. Electromagnetic induction and its propagation. *The Electrician*, 1887.

[4] J. D. Jackson. *Classical electrodynamics*. John Wiley & Sons, New York, 1999. 3rd edition.

[5] F. W. Hehl and Y. N. Obukhov. *Foundations of classical electrodynamics*. Birkhäuser, Boston, 2001.

[6] G. O. Schellstede, V. Perlick, and C. Lämmerzahl. Testing non-linear vacuum electrodynamics with Michelson interferometry. *Phys. Rev. D*, 92:025039, 2015.

[7] J. Gratus, V. Perlick, and R. W. Tucker. On the self-force in Bopp-Podolsky electrodynamics. *J. Phys. A: Math. Theor.*, 48:435401, 2015.

[8] V. Perlick. On the hyperbolicity of maxwell's equations with a local constitutive law. *J. Math. Phys.*, 52:042903, 2011.

[9] H. Goldstein. *Classical Mechanics*. Pearson, 2011.

[10] W. Greiner and J. Reinhardt. *Field quantization*. Springer, Berlin, 1996.

[11] M. Henneaux and C. Teitelboum. *Quantization of gauge systems*. Princeton University Press, Princeton, 1992.

[12] J. I. Anderson and P.G. Bergmann. Constraints in covariant field theories. *Phys. Rev.*, 83:1018, 1951.

[13] P. A. M. Dirac. Generalized Hamiltonian dynamics. *Can. J. Math.*, 2:129, 1950.

[14] C. Itzykson and J.-B. Zuber. *Quantum Field Theory*. Dover Publications, Mineola, 2006.

[15] J.W. van Holten. *Aspects of BRST Quantization*, pages 99–166. Springer Berlin Heidelberg, 2005.

[16] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Addison-Wesley, 1995.

[17] R. Haag. On quantum field theories. *Kong. Dan. Vid. Sel. Mat. Fys. Med.*, 29N12:1–37, 1955.

[18] S. Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 6 2005.

[19] G. C. Wick. The evaluation of the collision matrix. *Phys. Rev.*, 80:268–272, Oct 1950.

[20] M. D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 3 2014.

[21] J. C. Ward. An identity in quantum electrodynamics. *Phys. Rev.*, 78:182, 1950.

[22] S. Deser and R. I. Nepomechie. Gauge Invariance Versus Masslessness in De Sitter Space. *Annals Phys.*, 154:396, 1984.

[23] W. H. Furry. A Symmetry Theorem in the Positron Theory. *Physical Review*, 51(2):125–129, January 1937.

[24] S. Weinberg. Infrared photons and gravitons. *Phys. Rev.*, 140:B516, 1965.

[25] L. de Broglie. *Nouvelles Recherches sur la Lumière*, volume 411 of *Actualités Scientifiques et Industrielles*. Hermann & C^{ie}, Paris, 1936.

[26] A. Proca. Sur les photons et les particules charge pure. *Comptes Rendus Hebd. Séances Acad. .Sci. Paris*, 203:709, 1936.

[27] L. de Broglie. Ondes et quanta. *Comptes Rendus Hebd. Séances Acad. Sci. Paris*, 177:507, 1923.

[28] A. Proca. Particules libres photons et particules charge pure. *J. Phys. et Radium*, 8:23, 1937.

[29] F. Toscano. Luigi Bianchi, Gregorio Ricci Curbastro e la scoperta delle identità di Bianchi. In *Atti XX Congr. Naz. Storia della Fisica e dell'Astronomia*, page 353, Milano, 2000. Società Italiana degli Storici della Fisica e dell'Astronomia. 1-3 June 2000 Napoli.

[30] P. G. Bergmann and I. Goldberg. Dirac bracket transformations in phase space. *Phys. Rev.*, 98:531, 1955.

[31] C. de Rham, S. Garcia-Saenz, L. Heisenberg, V. Pozsgay, and X. Wang. To half-be or not to be? *J. High En. Phys.*, 2023:88, 2023.

[32] V. Errasti Díez, B. Gording, J. A. Méndez-Zavaleta, and A. Schmidt-May. Maxwell-Proca theory: definition and construction. *Phys. Rev. D*, 101:045009, 2020.

[33] E. C. G. Stueckelberg. Théorie de la radiation de photons de masse arbitrairement petite. *Helv. Phys. Acta*, 30:209, 1957.

[34] H. Ruegg and M. Ruiz-Altaba. The Stueckelberg field. *Int. J. Mod. Phys. A*, 19:3265, 2004.

[35] F. Bopp. Eine lineare Theorie des Elektrons. *Ann. Phys. (Leipzig)*, 430:345, 1940.

[36] B. Podolsky. A generalized electrodynamics. Part I - non-quantum. *Phys. Rev.*, 62:68, 1942.

[37] A. Landé and L. H. Thomas. Finite self-energies in radiation theory. Part II. *Phys. Rev.*, 60:514, 1941.

[38] A. Shala and V. Perlick. Waves in Bopp–Landé–Thomas–Podolsky generalized electrodynamics. *Annals Phys.*, 484:170286, 2026.

[39] W. Thirring. Regularization as a consequence of higher order equations. *Phys. Rev.*, 77:570, 1950.

[40] H. Carley, M. K.-H. Kiessling, and V. Perlick. On the schrödinger spectrum of a hydrogen atom with electrostatic Bopp–Landé–Thomas–Podolsky interaction between electron and proton. *Int. J. Mod. Phys. A*, 34:1950146, 2019.

[41] S.-S. Chern and J. Simons. Characteristic forms and geometric invariants. *Ann. Math.*, 99:48, 1974.

[42] D. Birmingham, M. Blau, M. Rakowski, and G.Thompson. Topological field theory. *Phys. Rept.*, 209:129–340, 1991.

[43] E. Witten. 2 + 1 dimensional gravity as an exactly soluble system. *Nuclear Physics B*, 311(1):46–78, 1988.

[44] S. Deser., R. Jackiw, and S. Templeton. Topologically massive gauge theories. *Ann. Phys. (N.Y.)*, 140:372, 1982. [Erratum: Ann. Phys. (N.Y.) 185, 406 (1988)].

[45] H. Dvali, E. Jackiw, and S. Pi. Topological mass generation in four dimensions. *Phys. Rev. Lett.*, 96:081602, 2006.

[46] T.J. Allen, M.J. Bowick, and A. Lahiri. Topological mass generation in (3+1)-dimensions. *Mod. Phys. Lett. A*, 6:559–572, 1991.

[47] L.-C. Tu, J. Luo, and G. T. Gillies. The mass of the photon. *Rep. Progr. Phys.*, 68:77, 2005.

[48] A. Accioly, J. Helayël-Neto, and E. Scatena. Upper bounds on the photon mass. *Phys. Rev. D*, 82:065026, 2010.

[49] A. Scharff Goldhaber and M. M. Nieto. Photon and graviton mass limits. *Rev. Mod. Phys.*, 82:939, 2010.

[50] G. Spavieri, J. Quintero, G. T. Gillies, and M. Rodriguez. A survey of existing and proposed classical and quantum approaches to the photon mass. *Eur. Phys. J. D*, 61:531, 2011.

[51] J.-J. Wei and X.-F. Wu. Testing fundamental physics with astrophysical transients. *Front. Phys.*, 16:44300, 2021.

[52] R. I. Workman and the Particle Data Group. Review of particle physics. *Progr. Theor. Exp. Phys.*, page 083C01, 2022.

[53] A. Barnes and J. D. Scargle. Improved upper limit on the photon rest mass. *Phys. Rev. Lett.*, 35:1117, 1975.

[54] D. D. Ryutov. Using plasma physics to weigh the photon. *Plasma Phys. Contr. Fus.*, 49:B429, 2007.

[55] P. de Bernadis, S. Masi, F. Melchiorri, and A. Moleti. Photon mass and cosmic microwave background anisotropy. *Astrophys. J.*, 284:L21, 1984.

[56] Y.-P. Yang and B. Zhang. Tight constraint on photon mass from pulsar spindown. *Astrophys. J.*, 842:23, 2017.

[57] D. D. Ryutov. The role of finite photon mass in magnetohydrodynamics of space plasmas. *Plasma Phys. Contr. Fus.*, 39:A73, 1997.

[58] E. Adelberger, G. Dvali, and A. Gruzinov. Photon-mass bound destroyed by vortices. *Phys. Rev. Lett.*, 98:010402, 2007.

[59] Y. Yamaguchi. A composite theory of elementary particles. *Progr. Theor. Phys. Suppl.*, 11:1, 1959.

[60] J. C. Byrne and R. R. Burman. A photon rest mass and electric currents in the galaxy. *J. Phys. A Math. Nucl. Gen.*, 6:L12, 1973.

[61] J. C. Byrne and R. R. Burman. Of fundamental electrodynamics and astrophysics. *Nat.*, 253:27, 1975.

[62] G. V. Chibisov. Astrophysical upper limits on the photon rest mass. *Sov. Phys. Usp.*, 19:624, 1976. [Usp. Fiz. Nauk, 119 (1976) 591].

[63] J. C. Byrne. Cosmic tests of maxwell's equation. *Astrophys. Space Sci.*, 46:115, 1977.

[64] R. Lakes. Experimental limits on the photon mass and cosmic magnetic vector potential. *Phys. Rev. Lett.*, 80:1826, 1998.

[65] J. Luo, L.-C. Tu, Z.-K. Hu, and E.-J Luan. New experimental limit on the photon rest mass with a rotating torsion balance. *Phys. Rev. Lett.*, 90:081801, 2003. Reply, 91, 149102 (2003).

[66] L.-C. Tu and J. Luo. Experimental tests of Coulomb's law and the photon rest mass. *Metrologia*, 41:S136, 2004.

[67] M. Füllekrug. Probing the speed of light with radio waves at extremely low frequencies. *Phys. Rev. Lett.*, 93:043901, 2004.

[68] A. D. A. M. Spallicci, G. Sarracino, O. Randriamboarison, J. A. Helayël-Neto, and A. Dib. Testing the Ampère-Maxwell law on the photon mass and Lorentz symmetry violation with MMS multi-spacecraft data. *Eur. Phys. J. Plus*, 139:551, 2024.

[69] L. Bonetti, J. Ellis, N. E. Mavromatos, A. S. Sakharov, E. K. Sarkisian-Grinbaum, and A. D. A. M. Spallicci. Photon mass limits from fast radio bursts. *Phys. Lett. B*, 757:548, 2016.

[70] L. Bonetti, J. Ellis, N. E. Mavromatos, A. S. Sakharov, E. K. Sarkisian-Grinbaum, and A. D. A. M. Spallicci. FRB 121102 casts new light on the photon mass. *Phys. Lett. B*, 768:326, 2017.

[71] N. Xing, H. Gao, J.-J. Wei, Z. Li, W. Wang, B. Zhang, X.-F. Wu, and P. Mészáros. Limits on the weak equivalence principle and photon mass with FRB 121102 subpulses. *Astrophys. J. Lett.*, 882:L13, 2019.

[72] B. Wang, J.-J. Wei, X.-F. Wu, and M. López-Corredoira. Revisiting constraints on the photon rest mass with cosmological fast radio bursts. *J. Cosm. Astropart. Phys.*, 2023:02, 2023.

[73] J. Plebañski. *Lectures on non-linear electrodynamics*. Nordita, København, 1970.

[74] G. Mie. Foundations of a theory of matter. First communication. *Annalen Phys.*, 37(3):511–534, 1912.

[75] G. Mie. Foundations of a theory of matter. Second communication. *Annalen Phys.*, 39(11):1–40, 1912.

[76] M. Born. Modified field equations with a finite radius of the electron. *Nature*, 132(3329):282.1, 1933.

[77] M. Born and L. Infeld. Foundations of the new field theory. *Proc. R. Soc. London A*, 144:425, 1934.

[78] Y. F. Alam and A. Behne. Review of Born-Infeld electrodynamics. 11 2021.

[79] L. Mezincescu, J. G. Russo, and P. K. Townsend. Hamiltonian birefringence and Born-Infeld limits. *JHEP*, 02:186, 2024.

[80] P. A. M. Dirac. A reformulation of the born-infeld electrodynamics. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 257(1288):32–43, 1960.

[81] W. von Heisenberg and H. Euler. Folgerungen aus der Diracschen Theorie des Positrons. *Zeitschr. Phys.*, 98:714, 1936.

[82] H. Euler. Über die streuung von licht an licht nach der Diracschen theorie. *Ann. Phys. (Leipzig)*, 26:398, 1936.

[83] V. Weisskopf. The electrodynamics of the vacuum based on the quantum theory of the electron. *Kong. Dan. Vid. Sel. Mat. Fys. Med.*, 14N6(6):1–39, 1936.

[84] G. V. Dunne. Heisenberg-Euler effective Lagrangians: basics and extensions. In M. Shifman, A. Vainshtein, and J. Wheater, editors, *From fields to strings: circumnavigating theoretical physics, Ian Kogan memorial collection*, volume 1, page 445. World Scientific, Singapore, 2004.

[85] A. Cadène, P. Berceau, M. Fouché, R. Battesti, and C. Rizzo. Vacuum magnetic linear birefringence using pulsed fields: status of the bmv experiment. *Eur. Phys. D.*, 68:16, 2014.

[86] G. Zavattini, F. Della Valle, U. Gastaldi, E. Milotti, G. Messineo, R. Pengo, L. Piemontese, and G. Ruoso. The PVLAS experiment for measuring the magnetic birefringence of vacuum. *N. Cim. C*, 36:41, 2013.

[87] V. A. Kostelecký and S. Samuel. Phenomenological gravitational constraints on strings and higher-dimensional theories. *Phys. Rev. Lett.*, 63:224, 1989.

[88] V. A. Kostelecký and S. Samuel. Spontaneous breaking of Lorentz symmetry in string theory. *Phys. Rev. D*, 39:683, 1989.

[89] V. A. Kostelecký and S. Samuel. Gravitational phenomenology in higher-dimensional theories and strings. *Phys. Rev. D*, 40:1886, 1989.

[90] S. M. Carroll, G. B. Field, and R. Jackiw. Limits on a Lorentz- and parity-violating modification of electrodynamics. *Phys. Rev. D*, 41:1231, 1990.

[91] L. Bonetti, L. R. dos Santos Filho, J. A. Helayël-Neto, and A. D. A. M. Spallicci. Effective photon mass from Super and Lorentz symmetry breaking. *Phys. Lett. B*, 764:203, 2017.

[92] B. Altschul. Radiatively induced lorentz-violating photon masses. *Phys. Rev. D*, 73:036005, 2006.

[93] A. P. Baêta Scarpelli, H. Belich, L. P. Boldo, L. P. Colatto, J. A. Helayël-Neto, and A. L. M. A. Nogueira. Remarks on the causality, unitarity and supersymmetric extension of the Lorentz and CPT-violating Maxwell-Chern-Simons model. *Nucl. Phys. B*, 127:105, 2004.

[94] L. Bonetti, L. R. dos Santos Filho, J. A. Helayël-Neto, and A. D. A. M. Spallicci. Photon sector analysis of Super and Lorentz symmetry breaking: effective photon mass, bi-refringence and dissipation. *Eur. Phys. J. C*, 78:811, 2018.

[95] I. Motie, B. Lamine, A. Blanchard, R. Battesti, and C. Rizzo. The optical analogy between a Lorentz-violating cosmos and a magneto-electric medium. *Eur. Phys. J. C*, 85(9):1048, 2025.

[96] I. Motie, J. Khodagholizadeh, S. Mahmoudi, B. Lamine, and A. Blanchard. Polarization of the CMB in the Standard Model Extension. 1 2026.

[97] A. D. A. M. Spallicci, A. Dib, and J. A. Helayël-Neto. Photon frequency variation in non-linear electro-magnetism. *Phys. Lett. B*, 855:138773, 2024.

[98] L. F. Abbott. Introduction to the Background Field Method. *Acta Phys. Polon. B*, 13:33, 1982.

[99] B. S. DeWitt. Quantum Theory of Gravity. 3. Applications of the Covariant Theory. *Phys. Rev.*, 162:1239–1256, 1967.

[100] A. D. A. M. Spallicci, J. A. Helayël-Neto, M. López-Corredoira, and S. Capozziello. Cosmology and the massive photon frequency shift induced by the Standard-Model Extension. *Eur. Phys. J. C.*, 81:4, 2021.

[101] A. Dib, J. A. Helayël-Neto, and A. D. A. M. Spallicci. Generalised non-linear electrodynamics: classical picture and effective mass generation. *Eur. Phys. J. C*, 86:83, 2026.

[102] R. M. Wald. *General relativity.* Univ. Chicago Press, Chicago, 1984.

[103] C. Misner, K. Thorne, and J. Wheeler. *Gravitation.* W. H. Freeman, San Francisco, 1973.

[104] J. B. Pitts. Does a second-class primary constraint generate a gauge transformation? electromagnetisms and gravities, massless and massive. *Ann. Phys. (N.Y.)*, 462:169621, 2024.

[105] A. P. Baeta Scarpelli, H. Belich, J. L. Boldo, and J. A. Helayel-Neto. Aspects of causality and unitarity and comments on vortexlike configurations in an abelian model with a Lorentz breaking term. *Phys. Rev. D*, 67:085021, 2003.

[106] A. Dib, A. Garnier, and A.D.A.M. Spallicci. Generalised aichelburg–sexl and self-force for photons. *Int. J. Geom. Methods Mod. Phys.*, 0(0):2550278, 0.

[107] P. C. Aichelburg ans R. U. Sexl. On the gravitational field of a massless particle. *Gen. Rel. Grav.*, 2:303, 1971.

[108] T. Dray and G. 't Hooft. The gravitational shock wave of a massless particle. *Nucl. Phys. B*, 253:173, 1985.

[109] T. Regge and J.A. Wheeler. Stability of a Schwarzschild singularity. *Phys. Rev.*, 108:1063, 1957.

[110] F. J. Zerilli. Tensor harmonics in canonical form for gravitational radiation and other applications. *J. Math. Phys.*, 11:2203, 1970.

[111] S. E. Gralla and R. M. Wald. A rigorous derivation of gravitational self-force. *Class. Q. Grav.*, 25:205009, 2008. Corrigendum, *ibid.* **28**, 159501 (2011).

[112] Y. Mino, M. Sasaki, and T. Tanaka. *Phys. Rev. D*, 55:3457, 1997. T. C. Quinn and R. M. Wald, Phys. Rev. D **56**, 3381 (1997).

[113] F. Bajardi, M. Benetti, S. Capozziello, and A. Dib. Extended theories of electrodynamics in f(R) gravity. *Annals Phys.*, 489:170443, 2026.

[114] S. Capozziello and M. De Laurentis. Extended theories of gravity. *Phys. Rept.*, 509:167, 2011.

Abedennour DIB

Théories Maxwelliennes Généralisées - Structure théorique

Les observations astrophysiques reposent principalement sur des signaux électromagnétiques, qui sont encore interprétés à l'aide de la théorie de Maxwell datant du XIXe siècle, une théorie linéaire et dépourvue de masse photonique, qui pourrait constituer une approximation d'une théorie plus générale, à l'instar de la gravité newtonienne par rapport à la relativité générale. Nous examinons les théories maxwelliennes généralisées (GMT) qui induisent A) des modifications des lois de Maxwell vérifiables à l'aide des données des sondes spatiales sur le vent solaire ; B) la biréfringence, objet d'expériences en France, pour laquelle l'indice de réfraction dépend de la polarisation de la lumière et de la direction de propagation ; C) la dispersion pour laquelle la vitesse de groupe diffère de c d'un facteur proportionnel à l'inverse du carré de la fréquence, ce qui a un impact sur les signaux des pulsars, les sursauts radio rapides (FRB), l'astronomie multi messagère et les estimations de la masse du graviton ; D) un décalage de fréquence dans le vide, dû à l'échange d'énergie entre les photons et les champs de fond, entraînant une réinterprétation de l'expansion de l'univers et de l'énergie noire. Aller au-delà de Maxwell est une nécessité en électrodynamique quantique pour prendre en compte les corrections non linéaires. Les GMT sont soit massifs (de Broglie-Proca, Bopp, Podolski, Stueckelberg et autres), soit non linéaires (Born-Infeld, von Heisenberg-Euler et autres). En physique des particules moderne, le Modèle Standard (SM), qui respecte la symétrie de Lorentz-Poincaré, n'explique pas les oscillations et les masses des neutrinos, le déséquilibre matière-antimatière et l'univers sombre. Pour le SM, le photon est la seule particule libre sans masse. Les théories étendant le SM sur la base d'une rupture de symétrie (théories SME) conduisent à une masse photonique effective, invariante sous l'action d'un champ de jauge, mais anisotrope. Dans cette thèse, nous avons sélectionné certaines de nos publications et discuté des effets non linéaires induits par les champs de fond sur la propagation du photon libre, en mettant en évidence la présence de décalages de fréquence. Au sein de la même classe de modèles, nous présentons les mécanismes qui conduisent à l'émergence d'une masse effective due aux interactions photon-fond, ainsi que les caractéristiques de stabilité et de propagation de ces extensions massives. Nous montrons notamment la présence d'une masse transversale qui préserve l'invariance de jauge ainsi que l'absence de tachyons dans les cas étudiés.

Mots clés : Photon, Théories massives, Théories non-linéaires, Au-delà du Modèle Standard

Generalised Maxwellian Theories – theoretical structure

Astrophysical observations are mainly based on electromagnetic signals, still read with 19th century Maxwell's theory, linear and without photonic mass, possibly approximating a larger theory, such as Newtonian gravity for general relativity. We examine Generalised Maxwellian Theories (GMT) that induce A) modifications of Maxwell's laws testable using solar wind spacecraft data; B) birefringence, object of experiments in France, for which the refractive index depends on light polarisation and propagation direction; C) dispersion for which the group velocity differs from c by a factor proportional to the inverse of the square of the frequency, impacting pulsar signals, Fast Radio Bursts (FRBs), multi-messenger astronomy and graviton mass estimates; D) a frequency shift in the vacuum, due to the energy exchange between the photon and background fields, causing the reinterpretation of the universe expansion and of dark energy. Going beyond Maxwell is a necessity in quantum electrodynamics for taking into account non-linear corrections. GMT are either massive (de Broglie-Proca, Bopp, Podolski, Stueckelberg and others) or non-linear (Born-Infeld, von Heisenberg-Euler and others). In modern particle physics, the Standard Model (SM), which respects Lorentz-Poincaré symmetry, does not explain the oscillations and masses of neutrinos, the matter-antimatter unbalance, and the dark universe. For the SM, the photon is the only free massless particle. Theories extending the SM based on symmetry breaking (SME theories) lead to an effective, gauge invariant, but anisotropic photon mass. In this thesis, we picked some of our publications and discussed the non-linear effects induced by background fields on the free photon propagation, highlighting the presence of frequency shifts. Within the same class of models, we present the mechanisms that lead to the emergence of an effective mass due to the photon-background interactions, as well as the stability and propagation features of such massive extensions. We notably show the presence of a transversal mass that preserves gauge invariance as well as the absence of tachyons in the studied cases.

Keywords: Photons, Massive theories, Non-linear theories, Beyond the Standard Model

Université d'ORLÉANS

Laboratoire de Physique et Chimie de l'Environnement et de l'Espace

3A Avenue de la Recherche Scientifique

45071 ORLÉANS

LPC2E