Dissertation presented to the Instituto Tecnológico de Aeronáutica, in partial fulfillment of the requirements for the degree of Master of Science in the Graduate Program of Physics, Field of Nuclear Physics.

Gabriel Ronconi Prada

SINGULAR SYSTEMS AND FIRST-ORDER ACTIONS: A COMPARATIVE STUDY

Dissertation approved in its final version by signatories below:

Documento assinado digitalmente
PEDRO JOSE POMPEIA
Data: 08/03/2026 16:43:28-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Pedro José Pompeia
Advisor

Campo Montenegro
São José dos Campos, SP - Brazil
2026

Cataloging-in Publication Data
Documentation and Information Division

Prada, Gabriel Ronconi
Singular systems and first-order actions: a comparative study / Gabriel Ronconi Prada.
São José dos Campos, 2026.
87f.

Dissertation of Master of Science – Course of Physics. Area of Nuclear Physics – Instituto Tecnológico de Aeronáutica, 2026. Advisor: Prof. Dr. Pedro José Pompeia.

1. Teoria quântica de campos. 2. Campos eletromagnéticos. 3. Eletrodinâmica. 4. Equações de Faddeew. 5. Física nuclear. 6. Física. I. Instituto Tecnológico de Aeronáutica. II. Title.

BIBLIOGRAPHIC REFERENCE

PRADA, Gabriel Ronconi. **Singular systems and first-order actions: a comparative study**. 2026. 87f. Dissertation of Master of Science – Instituto Tecnológico de Aeronáutica, São José dos Campos.

CESSION OF RIGHTS

Gabriel Ronconi Prada
Rua Tenente-Brigadeiro do Ar Paulo Victor da Silva
12224-300 – São José dos Campos - SP

SINGULAR SYSTEMS AND FIRST-ORDER ACTIONS: A COMPARATIVE STUDY

Gabriel Ronconi Prada

Thesis Committee Composition:

Prof. Dr.	Brett Vern Carlson	President	-	ITA
Prof. Dr.	Pedro José Pompeia	Advisor	-	ITA
Prof. Dr.	Samuel Augusto Wainer		-	ITA
Prof. Dr.	Bruto Max Pimentel Escobar		-	IFT-UNESP

ITA

To my family.

Acknowledgments

To my family, specially my parents, Paulo and Sueli, as well as my sibling Mari and my grandmother Antônia for their support, love, companionship throughout my life. Without them, I undoubtedly would not have come this far. To my advisor, professor Pedro Pompeia, not only for all the patience, the advice, and the teachings, but also for going above and beyond in making sure my student life was going well and healthy.

To professors Brett Vern Carlson, Samuel Augusto Wainer and Bruto Max Pimentel Escobar for accepting being part of this thesis' committee, as well as their feedback and suggestions regarding this final text. I would also like to thank my ITA professors Odilon Lourenço da Silva Filho, Mariana Dutra da Rosa Lourenço, César Henrique Lenzi and Tobias Frederico for the teachings during this course.

To my undergraduate professors at UNIFESP who were and still are also a very important part of my academic journey and who I will never forget, in special professors Pedro Levit Kaufmann, Cláudia Aline Azevedo dos Santos Mesquita, Grasiele Cristiane Jorge, Vanessa Gonçalves Paschoa and Thiago Castilho de Mello, as well as my undergraduate colleagues Leonardo Farias Santos and Matheus Fideles Leite.

To all my incredible friends who have been with me through the good times and bad times, and whose companionship I am very lucky to have! In particular, I would like to acknowledge, in alphabetical order, Edson Gabriel (AKA Haruki), Luiz (AKA Tark), Juliana (AKA Bibi), Martin Carroll, Rita Lima, Robyn Parkinson, Victor Matos, and Ziggy Swan.

I would also like to thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the financial support.

"Pensamentos que me afligem
Sentimentos que me dizem
Dos motivos escondidos
Na razão de estar aqui
As perguntas que me faço
São levadas ao espaço
E de lá eu tenho todas
As respostas que eu pedi"

— PENSAMENTOS (*Roberto Carlos e Erasmo Carlos*)

Resumo

Sistemas singulares formam uma categoria vasta e importante de sistemas físicos, tanto na mecânica em finitos graus de liberdade quanto na teoria de campos. Entre eles, podemos considerar não apenas sistemas com vínculos cinemáticos fisicamente aparentes, mas também sistemas com simetrias de calibre (*gauge*), caso de muitas teorias de campos como o eletromagnetismo. O desenvolvimento de diferentes formalismos na física teórica também proveram diversos métodos de entender e lidar com as particularidades desses tipos de sistemas. Neste trabalho, compararemos algumas dessas estratégias, estudando sistemas singulares no formalismo Lagrangiano, bem como no formalismo Hamiltoniano através de dois métodos diferentes: o método de Dirac-Bergmann e o método de Faddeev-Jackiw. Começaremos expondo cada procedimento no caso mecânico, depois fazendo a transição para teorias de campos, onde algumas modificações são necessárias para acomodar os infinitos graus de liberdade. Para este fim, usaremos o campo eletromagnético livre como exemplo motivador. Por fim, aplicaremos conceitos adquiridos nessas análises à eletrodinâmica não-linear, uma classe de diversas teorias de campos com amplas aplicações, das quais daremos especial atenção à eletrodinâmica de Born-Infeld e uma generalização recentemente proposta por Kruglov. Para isso, usaremos ainda uma formulação da eletrodinâmica não-linear baseada em quadros de Jordan e variações de Palatini, que nos permitem extrair mais relações na forma de vínculos, reforçando a utilidade de examiná-la como sistema singular.

Abstract

Singular systems are a vast and important category of physical systems, both in mechanics with finite degrees of freedom and in field theories. Among them, we can consider not only systems with physically evident kinematical constraints, but also theories with gauge symmetries, which includes many field theories such as electromagnetism. The development of different formalisms in theoretical physics has also provided various ways to understand and deal with the intricacies of this sort of system. In this work, we will compare some of these methods, studying singular systems in the Lagrangian formalism, and in the Hamiltonian formalism using two different procedures: the Dirac-Bergmann method and the Faddeev-Jackiw method. We begin by showing how each method is applied to a mechanical system, before making the transition to a field, which requires some changes to properly handle the infinite degrees of freedom. To this end, we will use the free electromagnetic field as a motivating example. Lastly, we will apply some of the concepts we acquired in these analyses to nonlinear electrodynamics (NLED), a family of various field theories with a broad range of applications, from which we will give special attention to the specific cases of Born-Infeld electrodynamics and a generalization recently proposed by Kruglov. To do this, we will also use a formulation of NLED based on Jordan frames and Palatini variations, which also allow us to extract more mathematical relations as system constraints, further reaffirming the importance of analyzing it as a singular system.

Contents

1 Introduction

In physics, a system, either mechanical or field theoretical, is said to be singular if it cannot be described in terms of standard equations of motion involving accelerations, but also requires further relations between the variables, which are called constraints. These constraints may represent not only kinematic restrictions the system's movement, but also systems with gauge symmetries, which are ubiquitous in modern physics, being commonly found in electromagnetism, quantum field theory and the standard model of particle physics, being among the most prominent examples of singular systems (SUNDERMEYER, 1982). Throughout the development of modern physics, different methods have been proposed to study the dynamics of singular systems. The main purpose of this work is to examine in a comprehensive manner three well-established methods and how they may be applied.

In chapter 2, we introduce each method in the context of mechanics, starting with the constrained Lagrangian formalism, which is the primordial formalism of analytical mechanics and classical field theory, and requires only the knowledge of the action written in terms of a Lagrangian. Then, we transition to the Hamiltonian formalism, which is the basis for the Dirac-Bergmann and Faddeev-Jackiw methods. These require a Legendre transformation of the Lagrangian, adding more degrees of freedom (the canonical momenta) in exchange for turning the equations of motion into first-order differential equations, and are important for the transition to quantum mechanics, which is also typically usually studied under the perspective of the Hamiltonian formalism. The Dirac-Bergmann method was first proposed by Dirac in 1950 as a way to bring singular systems into quantum mechanics by means of canonical quantization (DIRAC, 1950). The Faddeev-Jackiw method, which is our main focus, was first proposed by L. Faddeev and R. Jackiw in 1988 (FADDEEV; JACKIW, 1988) and complemented by J. Barcelos-Neto and C. Wotzasek in 1992 to allow a more natural way to apply it to singular systems (BARCELOS-NETO; WOTZASEK, 1992a), and requires rewriting the Lagrangian to be linear in the velocities, which we then use to treat all of the phase space (coordinates and momenta alike) in the same way and obtaining equations of motion for all variables. This relies on the mathematical theory of differential forms and symplectic geometry. Brief introductions on each of those subjects, with relevant results for this work, are included in appendices

A and B respectively. Both these Hamiltonian methods allow for a fairly straightforward generalization of canonical quantization, which is the traditional way to bring a classical system into the realm of quantum mechanics (DIRAC, 1964; BARCELOS-NETO; WOTZASEK, 1992b).

In chapter 3, we transition to infinite degrees of freedom in field theory. We do this in a practical manner by using the free electromagnetic field as an example throughout. Its gauge symmetry is well known, and thus it is a very endearing and widespread example of a singular system. We go through each of the 3 methods again, examining how to transition them to the field theoretical context.

Finally, in chapter 4, we give a more involved application of the concepts of the Faddeev-Jackiw method established in the previous chapters to a generalized theory of nonlinear electrodynamics. These types of theories have been proposed to explain a variety of physical effects in electrodynamics that seem to violate the linear superposition principle from classical electrodynamics, including the interfaces between different conducting materials, electric field self-energy, birefringecence; and vacuum polarization in quantum field theory (SOROKIN, 2022; BIALYNICKI-BIRULA, 1983). In the theory of gravitation, actions based on nonlinear functions of the Ricci scalar, such as $f(R)$ theories, can often be modified in a way that adds more degrees of freedom in exchange for simplifying the field equations. This is similar to the Legendre transform in Hamiltonian mechanics and is known as a Jordan frame (SOTIRIOU; FARAONI, 2010). We will use a similar construction for nonlinear electrodynamics, where the action is a nonlinear function of the electromagnetic Lorentz invariant scalar and pseudoscalar. Additionally, we also propose a formulation based on the Palatini variation from gravitation, where relations that are usually taken as definitions (the connection in relativity, or the electromagnetic tensor definitions in our electrodynamics case) are analyzed instead as constraints and dynamical equations resulting from the variational principle (TSAMPARLIS, 1978). This will then form a system where our singular formalisms will prove very useful, as this Palatini variation brings many new constraints to the forefront of the theory.

2 Theory of singular systems

2.1 Overview of analytical mechanics

Before we move into the actual theory of singular systems, we begin by giving a brief recapitulation of the formalism of Lagrangian and Hamiltonian mechanics for regular (nonsingular) systems and some of their properties. Let us have a physical system with n degrees of freedom and a Lagrangian function $L(q_n, \dot{q}_n, t)$. If we take the Hessian matrix of L with respect to $\dot{q}$,

$$\mathbf{W} := \begin{bmatrix} \frac{\partial^2 L}{\partial \dot{q}_1^2} & \cdots & \frac{\partial^2 L}{\partial \dot{q}_1 \partial \dot{q}_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial \dot{q}_n \partial \dot{q}_1} & \cdots & \frac{\partial^2 L}{\partial \dot{q}_n^2} \end{bmatrix}. \tag{2.1}$$

The system is said to be *regular* or *nonsingular* if this matrix is nonsingular, that is, if its determinant is not zero. Otherwise, the system is singular, which implies the existence of constraints acting on it. Singular systems will be the focus of this work from the next chapter on. For now let us focus on the simpler regular case to establish the foundations of our theory.

To obtain the equations of motion (EOM) for our system, we look to the action

$$S[q_j] = \int_{t_1}^{t_2} L(q_j, \dot{q}_j, t)\, dt,$$

which is a linear functional of each q_j. We can define a variational (or functional) derivative for such a functional with respect to a function q_j as

$$\frac{d}{d\epsilon} \left(S[q_j + \epsilon\alpha_j]\right)_{\epsilon=0} = \int_{t_1}^{t_2} \left(\sum_{j=1}^{n} \alpha_j \frac{\delta S}{\delta q_j} \right) dt,$$

where ϵ is a constant and α_j is any function such that $\alpha_j(t_1) = \alpha_j(t_2) = 0, \quad j = 1, \ldots, n$. Hamilton's principle says

$$\frac{\delta S}{\delta q_j} = 0, \quad j = 1, \ldots, n$$

which, if we expand the derivative, leads us to the Lagrangian equations of motion

$$\frac{\partial L}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0, \quad j = 1, \ldots, n$$

however, as we will see, for singular system, these equations may not be uniquely solvable and/or may require further analysis of constraints.

The Hamiltonian formalism provides an alternate way to obtain equations of motion. We define the *canonical momenta* as

$$p_j \coloneqq \frac{\partial L}{\partial \dot{q}_j}, \quad j = 1, \ldots, n.$$

We then treat these momenta as independent variables from the q_j coordinates. As we will see, the main difference between the mathematical properties of regular systems versus singular systems in the Hamiltonian formalism is that in regular systems, each p_j as defined above may be written as an independent function of $\dot{q}_j$'s and q_j's. In a singular system, due to the Hessian matrix (2.1) having linearly dependent rows, not all p_j's may be written as functions involving the velocities $\dot{q}_j$ and thus we cannot relate every generalized coordinate velocity to momenta.

The Hamiltonian is a function defined as the Legendre transformation of L

$$H(q, p, t) = \sum_{j=1}^{n} p_j \dot{q}_j - L(q, \dot{q}, t).$$

For nonsingular systems, we can invert the canonical momenta's definition and write the velocities $\dot{q}_j$ as functions of p_j and other coordinates q_j, and substituting this into the above expression will lead any possible dependence on the velocities on the right-hand side to be replaced by dependence on the momenta. Actually, as we will see later, this is the case even for singular systems where the relation between momenta and velocities is not necessarily invertible, though it is less obvious.

To obtain the equations of motion, we once again use Hamilton's principle. Taking

$$S = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} \left(\sum_{j=1}^{k} \dot{q}_j p_j - H\right) dt,$$

Hamilton's principle states that

$$\frac{\delta S}{\delta q_j} = \frac{\delta S}{\delta p_j} = 0, \quad j = 1, \ldots, n \tag{2.2}$$

which, expanding the derivative, gives us the Hamiltonian equations of motion

$$\begin{cases} \frac{\delta S}{\delta q_j} & = \frac{\partial H}{\partial q_j} - \dot{p}_j = 0 \\ \frac{\delta S}{\delta p_j} & = \frac{\partial H}{\partial p_j} - \dot{q}_j = 0 \end{cases}$$

and applying Hamilton's principle (2.2), we get the equations of motion

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \qquad \dot{p}_j = -\frac{\partial H}{\partial q_j} \tag{2.3}$$

or

$$\dot{q}_j = \{q_j, H\} \qquad \dot{p}_j = \{p_j, H\},$$

where

$$\{u, v\} := \sum_{j=1}^{n} \frac{\partial u}{\partial q_j}\frac{\partial v}{\partial p_j} - \frac{\partial v}{\partial q_j}\frac{\partial u}{\partial p_j} \tag{2.4}$$

is the Poisson bracket.

The Poisson bracket plays a central role in Hamiltonian mechanics. This is in great part because of its algebraic properties. The most important of them are that for any 3 functions f, g, h and scalar α, the Poisson bracket satisfies

- Bilinearity: $\{\alpha f, g\} = \{f, \alpha g\} = \alpha \{f, g\}$.
- Skew-symmetry: $\{f, g\} = -\{g, f\}$.
- Leibniz rule: $\{fg, h\} = f\{g, h\} = \{f, h\} g$
- Jacobi identity: $\{\{f, g\}, h\} . + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$.

There are plenty of ways to prove and understand these properties. For our work, we note that it is particularly useful to define a generalized Poisson bracket with a differential 2-form ω

$$\{f, g\} = \sum_{\mu,\nu=1}^{n} \omega^{\mu\nu} \frac{\partial f}{\partial \eta^\mu}\frac{\partial g}{\partial \eta^\nu},$$

where we use η_μ as the set of all phase space variables, where the first n components are the n coordinates q_n and the remaining components are the n momenta p_n. This defition of a generalized Poisson bracket with a 2-form elegantly ties the properties with the theory of differential forms[1]. Both the Dirac-Bergman method and Faddeev-Jackiw method rely on generalized Poisson brackets, so this also proves immediately useful for our study of singular systems.

Using the same idea of applying differential forms and the set of all phase space variables η^{μ}, there is another way to write the Hamiltonian equations of motion. We write the tautological 1-form (which we derive in more formal mathematical fashion in B.1.5)

$$\theta_{\nu}\left(\eta\right) := \sum_{j=1}^{k} p_j dq^j,$$

that is, $\theta_{\nu}\left(q_j\right) = p_j$ and $\theta_{\nu}\left(p_j\right) = 0$, the Lagrangian is

$$L\left(\eta\right) = \theta_{\nu}\left(\eta\right)\dot{\eta}^{\nu} - H \tag{2.5}$$

where we already employ Einstein's summation convention (which, from now on, will be the standard throughout the rest of this work). Then, the equations of motion are derived from Hamilton's principle as

$$\frac{\delta S}{\delta \eta^{\mu}} = \left(\frac{\partial \theta_{\mu}}{\partial \eta^{\nu}} - \frac{\partial \theta_{\nu}}{\partial \eta^{\mu}}\right)\dot{\eta}^{\nu} - \frac{\partial H}{\partial \eta^{\mu}} = 0. \tag{2.6}$$

This concludes our summary of important results from regular Lagrangian and Hamiltonian mechanics. We are now ready to start dealing with singular systems.

2.2 Singular systems

Now, we will start to study the theory of singular systems, firstly from the point of view of Lagrangian field theory, and then with the Hamiltonian formalism, where we have two ways of dealing with singularities, namely the Dirac-Bergmann method and the Faddeev-Jackiw method.

2.2.1 Lagrangian theory

Let us have a physical system with n degrees of freedom. We start with the action S, defined as the integral of the Lagrangian L

[1] A brief introduction to mathematical aspects of differential forms is given in appendix A

$$S = \int L\left(q_j, \dot{q}_j\right) dt,$$

where $q_j, j = 1, \ldots, n$ are the generalized coordinates of our system. Then, from the principle of least action, we obtain a set of Lagrange equations of motion:

$$\frac{\partial L}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0.$$

Ultimately, our goal will be to solve for generalized accelerations $\ddot{q}_j$. If we expand the second term in the equations above, we have

$$\begin{aligned} \frac{\partial L}{\partial q_j} - \frac{\partial^2 L}{\partial \dot{q}_j \partial q_k}\dot{q}_k - \frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_k}\ddot{q}_k &= 0 \\ \implies \frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_k}\ddot{q}_k &= \frac{\partial L}{\partial q_j} - \frac{\partial^2 L}{\partial \dot{q}_j \partial q_k}\dot{q}_k. \end{aligned}$$

This can be read as a linear system of equations on $\ddot{q}_k$:

$$\mathbf{M}\ddot{\mathbf{q}} = \mathbf{w} \tag{2.7}$$

where $\ddot{\mathbf{q}}$ is the vector of all n generalized accelerations; $\mathbf{w}$ is the vector with components

$$w_j = \frac{\partial L}{\partial q_j} - \frac{\partial^2 L}{\partial \dot{q}_j \partial q_k}\dot{q}_k,$$

and $\mathbf{M}$ is the Hessian matrix of generalized velocities

$$\mathbf{M} = \begin{bmatrix} \frac{\partial^2 L}{\partial \dot{q}_1^2} & \cdots & \frac{\partial^2 L}{\partial \dot{q}_1 \partial \dot{q}_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial \dot{q}_n \partial \dot{q}_1} & \cdots & \frac{\partial^2 L}{\partial \dot{q}_n^2} \end{bmatrix}. \tag{2.8}$$

The singularity of the system depends on the structure of this matrix. If it is singular (i.e. $\det \mathbf{M} = 0$), with rank r, then the system does not have a single well defined set of solutions, which means we cannot obtain all expressions for $\ddot{q}_k$ from the Lagrange equations of motion (2.7). In this case, the matrix has a non-trivial kernel, with a basis $\mathbf{v}_1, \ldots, \mathbf{v}_(n-r)$ given by solutions to the null eigenvalue equation

$$\mathbf{M}\mathbf{v}_j = \mathbf{v}_j^T\mathbf{M} = 0.$$

If we take the inner product of both sides of (2.7) by these $\mathbf{v}_j$, we obtain a set of equations

$$\mathbf{v}_j^T\mathbf{w} = 0, \quad j = 1, \ldots, n-r.$$

As we have seen, the elements of the vector $\mathbf{w}$ are functions of q and $\dot{q}$. This way, we obtain $n-r$ relations between the coordinates and velocities from the equation above. If the system and its Lagrangian are such that these equations are automatically satisfied, then the system has no constraints, and we may use the equations of motion to solve for $n-r$ accelerations $\ddot{q}_j$. The coordinates with unsolved accelerations can be taken as any arbitrary function of time.

Now, it is also possible that at least some of the equations are not satisfied. Let there be $m \leq n-r$ such equations. Then, which we will denote by

$$\gamma_j\left(q,\dot{q}\right) \coloneqq \mathbf{v}_j^T\mathbf{w}.$$

These functions represent constraints of the system. By setting $\gamma_j = 0$, they act reducing the degrees of freedom of the original system, or, geometrically, constraining the movement to a lower-dimensional hypersurface. If we define a new matrix $\mathbf{M}'$ as the matrix $\mathbf{M}$ evaluated on the constrained surface, this new matrix may have a lower rank r' than $\mathbf{M}$, which introduces new null eigenvectors, which by the same process as before may or may not introduce more constraints, and thus restrict our system to an even lower-dimensional hypersurface, and so on. We then repeat this process of obtaining eigenvectors and constraints, getting a new matrix, until we finally have a set of m' constraints γ_j when all further equations

$$\mathbf{v}_j^T\mathbf{w} = 0$$

are automatically satisfied on the hypersurface.

With this set of m' constraints, we proceed by dividing then into two categories: *A-type constraints*, which depend only on the coordinates q, and *B-type constraints*, which depend on the coordinates q and velocities $\dot{q}$. Let m_A be the number of A-type constraints and m_B the number of B-type constraints. The constraint equations are

$$\begin{aligned}\gamma_j^A(q) &= 0. \quad j = 1,\ldots,m_A;\\ \gamma_j^B(q,\dot{q}) &= 0. \quad j = 1,\ldots,m_B.\end{aligned}$$

If we derive the A-type constraints with respect to time, we obtain

$$\frac{\partial\gamma_j^A}{\partial q_k}\dot{q}_k = 0, \quad j = 1,\ldots,m_A,$$

which may give us additional independent constraints, so that we now have $m_A' \geq m_A$ A-type and $m_B' \geq m_B$ B-type constraints. On the other hand, the derivatives of B-type constraints and second derivatives of A-type constraints give us relations between the accelerations

$$\begin{cases} \frac{\partial^2 \gamma_j^A}{\partial q_k \partial q_l}\dot{q}_k\dot{q}_l + \frac{\partial \gamma_j^A}{\partial \dot{q}_k}\ddot{q}_k = 0, \quad j = 1, \ldots, m'_A \\ \frac{\partial \gamma_j^B}{\partial q_k}\dot{q}_k + \frac{\partial \gamma_j^B}{\partial \dot{q}_k}\ddot{q}_k = 0 \quad j = 1, \ldots, m'_B \end{cases}.$$

The $n \times m'_A$ matrix

$$\begin{bmatrix} \nabla_q \gamma_1 \\ \vdots \\ \nabla_q \gamma_{m'_A} \end{bmatrix}$$

has maximal rank, m'_A, and we can rewrite m'_A B-type constraints as

$$\dot{q}_j \frac{\partial \gamma_k^A}{\partial q_j} = 0, \quad k = 1, \ldots, m'_A,$$

Let the remaining B-type constraints be

$$\psi_k^B(q, \dot{q}) = 0, \quad k = 1, \ldots, (m'_B - m'_A)$$

and the $2n \times m'_B$ matrix

$$\begin{bmatrix} \nabla_q \gamma_1 \\ \vdots \\ \nabla_q \gamma_{m'_A} \\ \nabla_{\dot{q}} \psi_1^B \\ \vdots \\ \nabla_{\dot{q}} \psi^B_{m'_B - m'_A} \end{bmatrix}$$

also has maximal rank, m'_B, and the first m'_B equations for accelerations can be written as

$$\begin{cases} \frac{\partial \gamma_k^A}{\partial q_l}\ddot{q}_l + \frac{\partial^2 \gamma_k^A}{\partial q_j \partial q_l}\dot{q}_j\dot{q}_l & = 0, \quad k = 1, \ldots, m'_A; \\ \frac{\partial \psi_k^B}{\partial \dot{q}_l}\ddot{q}_l + \frac{\partial \psi_k^B}{\partial q_l}\dot{q}_l & = 0, \quad k = 1, \ldots, (m'_B - m'_A). \end{cases}$$

The remaining accelerations can be written as

$$\mathbf{M}'\ddot{\mathbf{q}} = \mathbf{w}', \quad k = 1, \ldots, (r' - m'_B)$$

from the equations of motion. The A-type constraints which depend only on q can be used to express exactly m'_A of the q coordinates in terms of other coordinates. Let $\sigma = m'_A + 1, \ldots, n$ such that q_σ is the set of remaining coordinates that are not written as functions of the others. Then the A-type constraints are

$$q_1 = f_1(q_\sigma);$$

$$\vdots$$

$$q_{m'_A} = f_{m'_A}(q_\sigma).$$

We can use these equations to eliminate the m'_A coordinates as they can be simply represented by functions of the remaining coordinates (and as we have seen, we can also obtain their velocities and accelerations, also obtainable from q_σ). So our system can be represented by just $n - m'_A$ coordinates and the $m'_B - m'_A$ B-type constraints:

$$\psi_k^B(q, \dot{q}) = 0, \quad k = 1, \ldots, (m'_B - m'_A)$$

and the accelerations are given by

$$\begin{cases} \frac{\partial \psi_k^B}{\partial \dot{q}_\sigma}\ddot{q}_\sigma + \frac{\partial \psi_k^B}{\partial q_\sigma}\dot{q}_\sigma & = 0, \quad k = 1, \ldots, (m'_B - m'_A)\,; \\ M'_{k\sigma}\ddot{q}_\sigma & = w'_k, \quad k = 1, \ldots, (r' - m'_B)\,. \end{cases}$$

The motion takes place in the constrained hypersurface with dimension $2n - m'_A - m'_B$, on the space of $2(n - m'_A)$ coordinates q_σ,$\dot{q}_\sigma$.

The B-type constraints are satisfied for all t as long as they are satisfied for $t = 0$; thus we only consider them to restrict the possible initial values, while using the acceleration equations to solve the equations of motion. The equations above allow us to write $r' - m'_A$ accelerations in terms of the rest:

$$\ddot{q}_\alpha = g_\alpha(q_A, \dot{q}_A, \ddot{q}_A).$$

There are $n - r'$ coordinates q_A and $r' - m'_A$ coordinates q_α. (If $n = r'$ then there are zero q_A). Together they make up the full q_σ set.

If some accelerations $\ddot{q}_\alpha$ can be solved and written in terms of q_A, then no B-type constraint will depend only on q_A and $\dot{q}_A$. This means we can solve the accelerations by choosing q_As as arbitrary functions of time, and giving initial values to q_α, $\dot{q}_\alpha$ at $t = 0$ obeying the B-type constraints, the q_α are uniquely determined and will obey the constraints. The general solution will then have one arbitrary function of time for each acceleration that could not be solved. For the accelerations that could be solved, they will have the same number of restrictions on initial values and velocities as there are B-type constraints.

Since we eliminated all A-type constraints by changing the coordinates, we can think of the constrained Lagrangian system as involving only B-type constraints in general.

2.2.1.0.1 Example A didactic example of a singular Lagrangian due to (SUNDERMEYER, 1982) is

$$L(q_1, q_2, \dot{q}_1, \dot{q}_2) = \frac{1}{2}\dot{q}_1^2 + q_2\dot{q}_1 + (1-\alpha)q_1\dot{q}_2 + \frac{\beta}{2}(q_1 - q_2)^2, \tag{2.9}$$

where α, β are arbitrary nonzero constants.

The Euler-Lagrange equations are

$$\begin{aligned}
\frac{\partial L}{\partial q_1} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_1} = 0 &\implies (1-\alpha)\,\dot{q}_2 + \beta\,(q_1 - q_2) - \ddot{q}_1 - \dot{q}_2 = 0 \\
&\implies \ddot{q}_1 = \alpha\dot{q}_2 - \beta\,(q_1 - q_2) \\
\frac{\partial L}{\partial q_2} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_2} = 0 &\implies \dot{q}_1 - \beta\,(q_1 - q_2) - (1-\alpha)\,\dot{q}_1 = 0 \\
&\implies \alpha\dot{q}_1 - \beta\,(q_1 - q_2) = 0.
\end{aligned}$$

Notice how the acceleration $\ddot{q}_2$ does not appear. In fact, the Hessian matrix

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

is clearly singular; its kernel is 1-dimensional with basis vector $\begin{bmatrix} 0 & 1 \end{bmatrix}$. This gives us a constraint

$$w_2 = 0 = -\beta\,(q_1 - q_2) + \alpha\dot{q}_1,$$

confirming that the equation of motion for q_2 gives rise to a B-type constraint. The conservation in time of this constraint gives us

$$-\alpha\ddot{q}_1 - \beta\,(\dot{q}_1 - \dot{q}_2) = 0;$$

substituting the equation of motion for $\ddot{q}_1$:

$$-\alpha\,[\alpha\dot{q}_2 - \beta\,(q_1 - q_2)] - \beta\,(\dot{q}_1 - \dot{q}_2) = 0.$$

On the other hand, the constraint we found implies that $-\beta\,(q_1 - q_2) = -\alpha\dot{q}_1$. So the equation above reduces to

$$-\alpha^2\,(\dot{q}_2 - \dot{q}_1) - \beta\,(\dot{q}_1 - \dot{q}_2) = \left(\alpha^2 - \beta\right)(\dot{q}_1 - \dot{q}_2) = 0.$$

Now, we look at two cases. If $\alpha^2 \neq \beta$, the equation above forces another constraint, $\dot{q}_1 = \dot{q}_2$. The first constraint is enough to determine $\dot{q}_1$ in terms of q_1 and q_2, and this new constraint also lets us solve $\dot{q}_2$, so the dynamics of the system are

$$\ddot{q}_1 = \ddot{q}_2 = \frac{\beta}{\alpha}\,(\dot{q}_1 - \dot{q}_2)\,.$$

If $\alpha^2 = \beta$, then the second constraint is trivially satisfied. There are no further equations, and $\dot{q}_2$ can not be exactly solved. We simply take q_2 as an arbitrary function of time, and the accelerations are

$$\ddot{q}_1 = \frac{1}{\alpha}\left(\dot{q}_1 - \dot{q}_2\right),$$
$$\ddot{q}_2 = \lambda(t).$$

2.2.2 Gauge invariance and constraints

When applying Hamilton's principle, we typically only consider paths that the action functional can have that leave the end points (the times t_1 and t_2) unchanged. However, we can also consider a more general case, where we have a curve C going from a point Q_1 at time t_1 to a point Q_2 at time t_2, and a infinitesimally displaced curve C' going from point Q_1' at t_1' to Q_2' at t_2'. The variations in time end points are

$$t_1' - t_1 = \Delta t_1 \quad t_2' - t_2 = \Delta t_2,$$

(we assume them to be small, i.e. can be approximated by a first order function) and the variation in a coordinate along the curve

$$q_k'(t) = q_k(t) + \delta q_k(t) \implies \dot{q}_k'(t) = \dot{q}_k(t) + \frac{d}{dt}\delta q_k(t).$$

If the action functional for the curve C is

$$S[C] = \int_{t_1}^{t_2} L(q, \dot{q}, t)dt,$$

then the functional variation is

$$\Delta S = S[C] - S[C'] = \int_{t_1}^{t_2} L(q_k, \dot{q}_k, t)dt - \int_{t_1'}^{t_2'} L(q_k', \dot{q}_k', t)dt$$
$$= \int_{t_1}^{t_2} \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k}\right]\delta q_k dt + \left[L\Delta t + \frac{\partial L}{\partial \dot{q}_k}\delta q_k(t)\right]_{t_1}^{t_2} \tag{2.10}$$

where $\Delta t = \Delta t_1 + \Delta t_2$.

Now, writing

$$\Delta q_k(t_j) = q_k'(t_j) - q_k(t_j) = \delta q_k(t_j) + \dot{q}_k(t_j)\Delta t_j, \quad j = 1, 2,$$

we can then write

$$\Delta S = \int_{t_1}^{t_2} \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q_k}} \right) - \frac{\partial L}{\partial q_k} \right] \delta q_k dt + \left[p_k \Delta q_k - H \Delta t \right]_{t_1}^{t_2}.$$

A Lagrangian L is said to be *quasi-invariant* under a transformation ϕ if the Lagrangian's variation can be written as a total time derivative, i.e.

$$\delta q_k = \epsilon \phi_k(q, \dot{q}),$$

$$\delta L = L(q + \delta q, \dot{q} + \delta \dot{q}) - L(q, \dot{q}) = \epsilon \frac{d}{dt} F(q, \dot{q}, t).$$

(The Lagrangian being fully invariant is a special case when $F = 0$) The variation in the action functional for a path C is then

$$\Delta S[C] = \int_{t_1}^{t_2} \delta L(q, \dot{q}, t) dt = \epsilon \; F(q, \dot{q}, t)|_{t_1}^{t_2}.$$

If we take our previous expression for the action functional (2.10) with fixed t_1 and t_2 (such that $\Delta t = 0$), we get

$$\Delta S[C] = \int_{t_1}^{t_2} \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q_k}} \right) - \frac{\partial L}{\partial q_k} \right] \delta q_k(t) dt + \left[p_k \delta q_k(t) \right]_{t_1}^{t_2},$$

and equating both expressions for ΔS,

$$- \int_{t_1}^{t_2} [EL]_k \, \delta q_k(t) dt = \left[p_k \delta q_k - \epsilon F(q_k, \dot{q_k}, t) \right]_{t_1}^{t_2} = \epsilon \left[p_k \phi(q, \dot{q}) - F(q_k, \dot{q_k}, t) \right]_{t_1}^{t_2}, \qquad (2.11)$$

where $[EL]_k := \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q_k}} \right) - \frac{\partial L}{\partial q_k}$.

If Lagrange's equations of motion are obeyed in the curve C, the left hand side integral vanishes, and the right hand gives us a constant of motion

$$p_k \phi(q, \dot{q}) - F(q_k, \dot{q_k}, t) = \text{constant}.$$

Let $L(q, \dot{q})$ be a Lagrangian with no explicit time dependence, such that there is a class of infinitesimal transformations in configuration space involving a finite number of arbitrary functions of time $f_a(t)$, under which L is quasi-invariant. The transformation would then be

$$\delta q_k = \epsilon f_a(t) \phi_{ak}(q, \dot{q}),$$

where ϵ is a small parameter, and $\phi_{ak}(q, \dot{q})$ functions of $q, \dot{q}$. The variation of the Lagrangian would be

$$\delta L = \epsilon \frac{d}{dt}\left(f_a(t) F_a(q,\dot{q})\right).$$

Substituting these into 2.11 we have

$$-\int_{t_1}^{t_2} [EL]_k \, \epsilon f_a(t)\phi_{ak}(q,\dot{q})dt = \epsilon \left[f_a(t)\left(p_k\phi_{ak}(q,\dot{q}) - F_a(q,\dot{q})\right)\right]_{t_1}^{t_2}.$$

The left hand expression is an integral and thus depends on the values of $f_a(t)$ between t_1 and t_2. However, the right hand side only depends on the values of $f_a(t)$ at t_1 and t_2. Since the $f_a(t)$ are arbitrary functions of time, this means that both sides must be zero. This means that

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k}\right]\phi_{ak}(q,\dot{q}) = 0; \tag{2.12}$$

$$p_k\phi_{ak}(q,\dot{q}) - F_a(q,\dot{q}) = 0. \tag{2.13}$$

The term in brackets in the first equation is the same as in the equations of motion (although we are not assuming it is equal to 0 itself). Thus we may rewrite this equation in terms of the Hessian matrix elements M_{jk} as

$$\phi_{ak}(q,\dot{q})M_{kl}(q,\dot{q})\ddot{q}_l = \phi_{ak}(q,\dot{q})w_k(q,\dot{q}).$$

We have not imposed the equations of motion on the system, and the acceleration only appears in the left-hand side. This means both sides must vanish for the equality to hold:

$$\phi_{ak}M_{kl} = \phi_{ak}w_k = 0.$$

Thus, $\mathbf{M}$ is singular and has null eigenvectors. And (2.12) shows the equations of motion are *not* independent of each other. The constant of motion associated with this quasi-invariant would be

$$f_a(t)\left(p_k\phi_{ak}(q,\dot{q}) - F_a(q,\dot{q})\right),$$

however this depends on time, and thus must be 0, which indeed it is due to (2.13).

If we have a solution for the equations of motion obeying boundary conditions at $t = t_0$, and alter it with such a δq transformation, we have a new solution and we can choose the arbitrary $f_a(t)$ such that the same boundary conditions are satisfied at $t = t_0$. Since the general solution depends on arbitrary functions of time, not all accelerations will be solved for.

We may also generalize this process as such: consider an infinitesimal transformation

$$\delta q_k = \epsilon \left(f_a(t)\phi_{ak}(q,\dot{q}) + \dot{f}_a(t)\psi_{ak}(q,\dot{q}) \right),$$

with $\phi_{ak}(q,\dot{q})$ and $\psi_{ak}(q,\dot{q})$ specific functions of $q, \dot{q}$. If L is quasi-invariant then

$$\begin{aligned} \delta L &= \epsilon \frac{d}{dt} \left(f_a(t)F_a(q,\dot{q}) + \dot{f}_a(t)G_a(q,\dot{q}) \right); \\ &- \int_{t_1}^{t_2} [EL]_k \, \epsilon \left(f_a(t)\phi_{ak}(q,\dot{q}) + \dot{f}_a(t)\psi_{ak}(q,\dot{q}) \right) dt \\ &= \epsilon \left[f_a(t) \left(p_k\phi_{ak}(q,\dot{q}) - F_a(q,\dot{q}) \right) + \dot{f}_a(t) \left(p_k\psi_{ak}(q,\dot{q}) - G_a(q,\dot{q}) \right) \right]_{t_1}^{t_2}. \end{aligned}$$

Integrating the left-hand side by parts, we have

$$\begin{aligned} &- \int_{t_1}^{t_2} f_a(t) \left\{ [EL]_k \, \phi_{ak}(q,\dot{q}) - \frac{d}{dt} \left(\psi_{ak}(q,\dot{q}) \, [EL]_k \right) \right\} dt \\ = &\left[f_a(t) \left(p_k\phi_{ak}(q,\dot{q}) + \psi_{ak}(q,\dot{q}) \, [EL]_k - F_a(q,\dot{q}) \right) + \dot{f}_a(t) \left(p_k\psi_{ak}(q,\dot{q}) - G_a(q,\dot{q}) \right) \right]_{t_1}^{t_2}. \end{aligned}$$

Once again, both sides must vanish, leading us to three identities

$$\begin{aligned} [EL]_k \, \phi_{ak}(q,\dot{q}) - \frac{d}{dt} \left\{ \psi_{ak}(q,\dot{q}) \, [EL]_k \right\} &= 0; \\ p_k\phi_{ak}(q,\dot{q}) + \psi_{ak}(q,\dot{q}) \, [EL]_k - F_a(q,\dot{q}) &= 0; \\ p_k\psi_{ak}(q,\dot{q}) - G_a(q,\dot{q}) &= 0. \end{aligned}$$

And in similar fashion to before, we notice that the first equation has $\dddot{q}$ terms, and the second equation has $\ddot{q}$ terms, that without the imposition of equations of motion, ultimately means that the $\mathbf{W}$ matrix has null eigenvectors

$$\psi_{ak}(q,\dot{q})W_{kl}(q,\dot{q}) = 0.$$

However, we cannot say for this set of eigenvectors that $\psi_{ak}\alpha_k = 0$ as we did before for ϕ. This means that we may have some constraints for the system, which was not necessarily the case for the previous transformation. In fact, the second equation itself may turn into a constraint equation when the equations of motion are imposed $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0$:

$$p_k\phi_{ak}(q,\dot{q}) - F_a(q,\dot{q}) = 0.$$

The conserved quantity for the quasi-invariance in this case would be

$$f_a(t)\left(p_k\phi_{ak}(q,\dot{q}) - F_a(q,\dot{q})\right) + \dot{f}_a(t)\left(p_k\psi_{ak}(q,\dot{q}) - G_a(q,\dot{q})\right),$$

which once again depends on time and thus must vanish. In fact, the last of the three identities tells us the coefficient on $\dot{f}_a(t)$ vanishes, and the constraint equation above tells us that imposing the equations of motion makes the coefficient on $f_a(t)$ vanish as well.

2.2.3 Hamiltonian theory

Let us look back on the definition of the canonical momenta when we transition from the Lagrangian to the Hamiltonian formalism:

$$p_j \coloneqq \frac{\partial L\left(q,\dot{q}\right)}{\partial \dot{q}^j}. \tag{2.14}$$

The parameters upon which the Lagrangian may depend have been left explicit to show that, taking the derivative, this may give us an expression that depends on the q and $\dot{q}$. For a regular system, this provides us with a system of equations relating the set of momenta p_j to the generalized velocities $\dot{q}^j$.

As we have seen in the Lagrangian formalism, singularity occurs when the Hessian matrix with respect to the generalized velocities (2.8) is singular. In the Hamiltonian formalism, this implies that at least one of the generalized velocities $\dot{q}^j$ can not be written as a combination of q and p components. This inability to relate all the canonical momenta to generalized velocities means also that some generalized velocities $\dot{q}$ will remain independent of all momenta p. In a regular system, the relation between momenta and velocities is bijective and therefore invertible. In a singular system, some momenta cannot be written as functions of velocities and vice-versa.

Let us consider a system with n degrees of freedom. At the same time, let us assume it is singular with only r momenta that can be written as functions of velocities. Nonetheless, the equations (2.14) will still give us important information about the dynamics of these momenta by relating them with coordinates and other momenta. Physically, these relations will give us *constraints* that the system is subject to. To remaining momenta will be written:

$$p_\rho = \gamma_\rho\left(q^j, p_k\right), \quad \rho = r+1,\dots,n; \quad j=1,\dots,n; \quad k=1,\dots,r. \tag{2.15}$$

We can now look at an interesting property of the Hamiltonian. Writing it as a function of the Lagrangian $H = \dot{q}^j p_j - L$, If we were to take the derivative with respect to $\dot{q}^\rho$ (where ρ refers to one of the constrained indices above, so that $\dot{q}_\rho$ is a velocity that cannot be written as a combination of q_j and p_j), we would have

$$\frac{\partial H}{\partial \dot{q}^k} = p_k - \frac{\partial L}{\partial \dot{q}_k} = 0.$$

This holds even for a singular system, as p_k is still equal to $\frac{\partial L}{\partial \dot{q}_k}$. Therefore, the Hamiltonian will not depend explicitly on the velocities, even those that are completely independent of the momenta. This may be a surprising result, given that we could expect that we would be unable to eliminate such velocities from the Hamiltonian in transitioning from the Lagrangian.

From this point, the Dirac-Bergmann and Faddeev-Jackiw methods have different approaches to obtaining the equations of motion for the system. We start by examining the first method.

2.2.3.1 Dirac-Bergmann method

The Dirac-Bergmann method was first introduced by Paul Dirac as a way to tackle the problem of quantizing singular systems (DIRAC, 1950; DIRAC, 1964). In quantum physics, Dirac's canonical quantization of a mechanical system (also called first quantization, in contrast with the second quantization of quantum field theory) is performed by elevating the position and momentum to Hermitian operators $\hat{x}$ and $\hat{p}$, and equating the canonical Poisson brackets to a commutator

$$\{\mathbf{x}, \mathbf{p}\} = \frac{1}{i\hbar}\left[\hat{x}, \hat{p}\right].$$

The main goal in doing this is transitioning to a quantum mechanical, Hermitian operator mathematical framework while preserving aspects of the classical theory encoded by its Poisson bracket algebra and the canonical bracket relations

$$\begin{aligned} \left\{q^j, q^k\right\} = \{p_j, p_k\} &= 0 \\ \left\{q^j, p_k\right\} &= \delta^j_k. \end{aligned}$$

For a singular system, however, these relations will not hold. Since not all p^k are independent of q^j, we will have $\{q^j, p_k\} = 0$ for all j in these cases. Dirac's idea to circumvent this problem was to introduce a new generalized Poisson bracket that accounted for some of the constraint's structure while preserving most important properties of the usual Poisson bracket. This came to be known as the Dirac bracket. We will define this bracket shortly; first, let us analyze further the constraints we have.

In (2.15), we obtained r constraints γ_ρ. These constraints arose from the very structure of the Lagrangian, during the transition to the Hamiltonian. We call them *primary constraints*. If we take the derivative of the Hamiltonian $H = p_j \dot{q}_j - L$ with respect to q_j and p_k where $k = 1, \ldots, r$ (so that p_k are the moment that are written as functions of q and $\dot{q}$), we have

$$\begin{cases} \frac{\partial H}{\partial q_j} &= \frac{\partial \gamma_\rho}{\partial q_j} \dot{q}_\rho - \frac{\partial L}{\partial q_j} \quad j = 1, \ldots, n \\ \frac{\partial H}{\partial p_k} &= \dot{q}_k + \frac{\partial \gamma_\rho}{\partial p_k} \dot{q}_\rho, \quad k = 1, \ldots, r \end{cases}$$

which we can rewrite using the Lagrangian equations of motion on the $\frac{\partial L}{\partial q_j}$ term as

$$\begin{cases} \dot{q}_k &= \frac{\partial H}{\partial p_k} - \frac{\partial \gamma_\rho}{\partial p_k} \dot{q}_\rho, \quad k = 1, \ldots, r \\ \dot{p}_j &= -\frac{\partial H}{\partial q_j} + \frac{\partial \gamma_\rho}{\partial q_j} \dot{q}_\rho, \quad j = 1, \ldots, n \end{cases} \tag{2.16}$$

As is known, Hamiltonian phase space has dimension $2n$. Constraints act on this phase space, reducing the degrees of freedom of the system to a surface of smaller dimension. This surface is defined by the constraints γ_ρ. Let $f\left(q^j, p_j\right)$ and $g\left(q^j, p_j\right)$ be two functions on the complete, unconstrained phase space. We call these functions *weakly equal*, denoted by

$$f \approx g$$

if they are equal under the constraints (i.e. setting $p_\rho = \gamma_\rho$). Furthermore. if $f \approx g$ and the phase space gradients of f and g (including derivatives with respect to all q and p) are also equal under the constraints, we say f and g are *strongly equal*, denoted by

$$f \equiv g.$$

We note that we may define the constrained hypersurface itself using a set of $n - r$ weak equalities

$$p_\rho - \gamma_\rho \approx 0, \;\; \rho = r + 1, \ldots, n.$$

This works naturally because each γ_ρ does not depend on the p_ρ. If we define a set of functions $\phi_\rho = p_\rho - \gamma_\rho$, then, the weak equality of two phase space functions f and g is equivalent to the strong equality

$$f - \frac{\partial f}{\partial p_\rho} \phi_\rho \equiv g - \frac{\partial g}{\partial p_\rho} \phi_\rho,$$

which also means that any function that vanishes weakly can be strongly identified as a sum $\frac{\partial f}{\partial p_\rho} \phi_\rho$

$$f \approx 0 \implies f \equiv \frac{\partial f}{\partial p_\rho} \phi_\rho.$$

With the tools of weak and strong equalities, we can rewrite the equations of motion in

(2.16), with constraints ϕ_ρ and unsolved velocities $\dot{q}_\rho$ as such[2]:

$$\begin{cases} \dot{q}^j & \approx \frac{\partial H}{\partial p_j} - \frac{\partial \phi_\rho}{\partial p_j}\frac{\partial H}{\partial p_\rho} + \dot{q}^\rho \frac{\partial \phi_\rho}{\partial p_j} \\ \dot{p}_j & \approx \frac{\partial \phi_\rho}{\partial q^j}\frac{\partial H}{\partial p_\rho} - \frac{\partial H}{\partial q^j} - \dot{q}^\rho \frac{\partial \phi_\rho}{\partial q^j} \end{cases}, \quad j = 1, \ldots, n; \quad \rho = r+1, \ldots, n.$$

If j happens to be on the set of unsolved indices ρ, then we simply obtain $\dot{q}_j = \dot{q}_\rho$ from the first equation. We can rewrite the equations above with Poisson brackets.

$$\begin{cases} \dot{q}^j & = \{q^j, H\} + \{q^j, \phi_\rho\}\,\dot{q}_\rho \\ \dot{p}_j & = \{p_j, H\} + \{p_j, \phi_\rho\}\,\dot{q}^\rho \end{cases}, \quad j = 1, \ldots, n; \quad \rho = r+1, \ldots, n.$$

Now, let us consider that for any function f in this phase space, the time derivative could be written as

$$\dot{f} = \{f, H\} + \{f, \phi_\rho\}\,\dot{q}^\rho. \tag{2.17}$$

The primary constraints given by ϕ_ρ are preserved in time. Therefore,

$$\dot{\phi}_\rho \approx \{\phi_\rho, H\} + \{\phi_\rho, \phi_\sigma\}\,\dot{q}^\sigma \approx 0. \tag{2.18}$$

Consider the matrix with elements $\{\phi_\rho, \phi_\sigma\}$. If it is nonsingular (in the constrained hypersurface), then the equations above will give relations for each unsolved velocity $\dot{q}_\rho$. In this case, we can find an inverse matrix $\mathbf{C}$ in the hypersurface, such that

$$C^{\rho\sigma}\{\phi_\sigma, \phi_\tau\} = \delta^\rho_\tau,$$

and then from (2.18) we can write

$$\dot{q}^\rho \approx -C^{\rho\sigma}\{\phi_\sigma, H\}.$$

Substituting in (2.17):

$$\dot{f} \approx \{f, H\}^* = \{f, H\} - \{f, \phi_\rho\}\,C^{\rho\sigma}\{\phi_\sigma, H\}.$$

This generalized Poisson bracket $\{f, H\}^*$ is the *Dirac bracket* of f and H. Combined with the equations $\varphi_\rho \approx 0$, it serves as a constrained substitute for the Poisson bracket.

However, the matrix of constraints $\{\phi_\rho, \phi_\sigma\}$ may also be singular, with rank $m \leq n-r$. In this case, we proceed in a similar manner to how we dealt with the Hessian matrix in the Lagrangian formalism. Since the matrix is singular, there will be a nontrivial set of a null eigenvectors:

[2]we may also notice, at this point, that we could use any function K that is strongly equal to H and obtain the same equations of motion.

$$v_j^\rho \{\phi_\rho, \phi_\sigma\} \approx 0, \quad j = 1, \ldots, a.$$

Then, contracting v_j^ρ with (2.18) we get

$$v_j^\rho \{\phi_\rho, H\} \approx 0, \quad j = 1, \ldots, a.$$

From this set of equations, it is possible we obtain even more constraints that are independent of ϕ_ρ. Let us denote these new constraints as χ_μ, and suppose there are a number a of them. Now we have a constrained hypersurface of dimension $n+r-m$ with constraints given by

$$\begin{cases} \phi_\rho \approx 0, \rho = r+1, \ldots, n, \\ \chi_\mu \approx 0, \mu = 1, \ldots, a. \end{cases}.$$

While the constraints ϕ_ρ, which were derived directly from the Hamiltonian, are called primary constraints, the χ_μ are *secondary constraints*.

With the new constraints, we restrict the matrix of primary constraint Poisson brackets $\{\phi_\rho, \phi_\sigma\}$ to the constrained hypersurface and check if it produces any new eigenvectors for 0 that lead to new independent constraints. In this case, these are also secondary constraints that are to be added to the χ_μ set. Note that with each iteration, our definition of the weak and strong equalities changes to refer to the constrained hypersurface we have obtained so far.

If we require that the secondary constraints also be preserved in time, the time derivatives of all constraints form a system of $n-r+m$ equations

$$\begin{cases} \dot{\phi}_\rho & \approx \{\phi_\rho, H\} + \{\phi_\rho, \phi_\sigma\} \dot{q}^\sigma \approx 0 \quad \rho = r+1, \ldots, n, \\ \dot{\chi}_\mu & \approx \{\chi_\mu, H\} + \{\chi_\mu, \phi_\sigma\} \dot{q}^\sigma \approx 0, \quad \mu = 1, \ldots, a. \end{cases} \tag{2.19}$$

which leads us to examine the $(n-r+a) \times (n-r)$ matrix

$$\mathcal{B} = \begin{bmatrix} \{\phi_\rho, \phi_\sigma\} \\ \{\chi_\mu, \phi_\sigma\} \end{bmatrix}, \quad \begin{matrix} \rho, \sigma = r+1, \ldots, n, \\ \mu = 1, \ldots, a. \end{matrix}$$

If $\mathcal{B}$ has maximal rank (which would be $n-r$), then there are exactly $n-r$ independent equations for the $\dot{q}_\rho$ among the system of constraint time derivatives (2.19), allowing us to write all $\dot{q}_\rho$ in terms of q and p. Otherwise, if this matrix has m null left-hand eigenvectors, which give us more equations

$$v_\rho \{\phi_\rho, \phi_\sigma\} + v_\mu \{\chi_\mu, \phi_\sigma\} \approx 0,$$

which, as with the primary constraints, can be used to obtain relations

$$v_\rho \{\phi_\rho, H\} + v_\mu \{\chi_\mu, H\} \approx 0.$$

Once again, these equations may give us additional secondary constraints, and we obtain a new matrix and repeat the process. If not, then we have a complete set of constraints, and the remaining unsolved velocities remain free as arbitrary functions of time. Now, since $\mathcal{B}$ has rank $m' \leq n - r$, there are only m' linearly independent columns and rows. Then we can find $n - r - m'$ relations

$$\begin{aligned} &\{\phi_\rho, \phi_\sigma\} \xi^\beta_\sigma \approx 0; \\ &\{\chi_\alpha, \phi_\sigma\} \xi^\beta_\sigma \approx 0; \ \beta = 1, \ldots, n - r - m'. \end{aligned}$$

We can replace the primary constraints $\phi_\rho \approx 0$ in our theory by any linear combination of themselves. In particular, we may choose to replace $n - r - m'$ constraints by linear combinations defined by

$$\phi_\beta = \xi^\beta_\sigma \phi_\sigma, \ \beta = 1, \ldots, n - r - m',$$

with ξ^β_σ being the components of the relations above. The remaining m' primary constraints will be denoted ϕ_γ. These new constraints ϕ_β have the property that their Poisson bracket with any other constraint vanishes

$$\{\phi_\beta, \phi_{\beta'}\} \approx \{\phi_\beta, \phi_\gamma\} \approx \{\phi_\beta, \chi_\alpha\} \approx 0.$$

The ϕ_γ and their linear combinations do not share this property, otherwise we would have more than $n - r - m'$ relations between the columns of the $\mathcal{B}$ matrix.

The ϕ_β are called *first-class constraints*, and the ϕ_γ are called *second-class constraints*. This division is useful because it tells us exactly which combinations of velocities will be solved for. We can also replace the χ_α by a new set χ'_α defined by

$$\chi'_\alpha = S_{\alpha\alpha'}\chi_{\alpha'} + S_{\alpha\beta}\phi_\beta + S_{\alpha\gamma}\phi_\gamma,$$

as long as the S_{jk} elements define a non-singular matrix. Using this, we may also divide the χ'_α into first-class χ_A and second-class χ_B constraints, with the χ_A obeying

$$\{\chi_A, \chi_{A'}\} \approx \{\chi_A, \chi_B\} \approx \{\chi_A, \phi_\beta\} \approx \{\chi_A, \phi_\gamma\} \approx 0$$

and no linear combination of the χ_B and ϕ_γ is first class.

Now, let us write the full system of velocities using our new convention for writing constraints. There will be a linear combination

$$\phi_\sigma \dot{q}_\sigma = \phi_\beta v_\beta + \phi_\gamma v_\gamma,$$

where v_β and v_γ are linearly independent combinations of $\dot{q}_\rho$. Our full system is defined by

$$\begin{cases} \{\phi_\beta, H\} & \approx 0, \\ \{\chi_A, H\} & \approx 0, \\ \{\phi_\gamma, H\} + \{\phi_\gamma, \phi_{\gamma'}\} v_{\gamma'} & \approx 0, \\ \{\chi_B, H\} + \{\chi_B, \phi_{\gamma'}\} v_{\gamma'} & \approx 0. \end{cases} \qquad \beta = 1, \ldots, n - r - m'. \tag{2.20}$$

The v_β vanish from these relations entirely as they are multiplied by primary first-class constraints ϕ_β. This allows us to conclude that there will be as many undetermined combinations of the velocities v_β as primary first-class constraints, while each v_γ will appear as an arbitrary function of time in the equations of motion.

In particular, each v_γ is defined by the last two equations in the (2.20) set above. Take the matrix of all constraint Poisson brackets

$$\mathbf{\Delta} = \begin{bmatrix} \{\phi_\gamma, \phi_{\gamma'}\} & \{\phi_\gamma, \chi_{B'}\} \\ \{\chi_B, \phi_{\gamma'}\} & \{\chi_B, \chi_{B'}\} \end{bmatrix}.$$

This matrix is skew-symmetric and non-singular; otherwise, there would be linearly dependent rows, which would imply that some of the second-class constraints are in fact first-class.

We can then find its inverse

$$\mathbf{\Delta}^{-1} = \mathbf{C} = \begin{bmatrix} C_{\gamma\gamma'} & C_{\gamma B} \\ C_{B\gamma'} & C_{BB'} \end{bmatrix}$$

Using $\mathbf{\Delta}$ and $\mathbf{C}$ and the equations for v_γ in (2.20) we get

$$v_\gamma \approx -C_{\gamma\gamma'} \{\phi_{\gamma'}, H\} - C_{\gamma B'} \{\chi_{B'}, H\} \tag{2.21}$$

$$0 \approx C_{B\gamma'} \{\phi_{\gamma'}, H\} + C_{BB'} \{\chi_{B'}, H\}. \tag{2.22}$$

Now, recalling the equations of motion

$$\frac{d}{dt} g(q, p) \approx \{g, H\} + \{g, \phi_\rho\} \dot{q}_\rho \approx \{g, H\} + \{g, \phi_\beta\} v_\beta + \{g, \phi_\gamma\} v_\gamma.$$

Substituting (2.21),

$$\frac{d}{dt}g(q,p) \approx \{g,H\} + \{g,\phi_\beta\}\, v_\beta - \{g,\phi_\gamma\}\, C_{\gamma\gamma'}\, \{\phi_{\gamma'},H\} - \{g,\phi_\gamma\}\, C_{\gamma B'}\, \{\chi_{B'},H\}\,,$$

and finally, we can make this expression symmetric in the second-class constraints ϕ_γ and χ_B by adding $-\{g,\chi_B\}\, C_{B\beta'}\, \{\phi_{\beta'},H\} - \{g,\chi_B\}\, C_{BB'}\, \{\chi_{B'},H\}$ to the right side. This does not change the weak equation, as (2.22) implies it vanishes weakly. Now if we denote the set of all second-class constraints by $\zeta_\mu = (\phi_\gamma, \chi_B)$, the equations of motion take the simpler for

$$\frac{d}{dt}g(q,p) \approx \{g,H\} + \{g,\phi_\beta\}\, v_\beta - \{g,\zeta_\mu\}\, C_{\mu\mu'}\, \{\zeta_{\mu'},H\}\,.$$

This is the final form of the general equations of motion, and will constitute the basic equations of the Dirac-Bergmann Hamiltonian theory together with the set of first-class constraints

$$\phi_\beta \approx 0, \quad \chi_A \approx 0$$

and the set of second-class constraints

$$\zeta_\mu \approx 0.$$

The Hamiltonian H is not arbitrary, as we originally found it to be strongly equal to the function $F(q,p)$. Also, the first-class constraints' Poisson brackets with H vanish weakly; this, coupled with the equation of motion and the way ζ_μ's appear in them, means that all constraints are preserved in time. Thus, if we choose initial values of q_k and p_k at $t = 0$ such that they are on the constrained hypersurface U, and the v_α are defined as arbitrary functions of time, we can solve the equations of motion and obtain the q_k and p_k for all later times. Similarly to the Lagrangian case, the constraints will be obeyed for all time, and the motion in phase space will always stay on U. Thus the general solution will also depend on arbitrary functions of time.

The separation of first-class and second-class constraints is more important than primary and secondary constraints. Sometimes, a given system of Lagrangian equations of motion can be derived from more than one Lagrangian, but classifying the constraints into primary and secondary depends on the functional form of the Lagrangian.

In our final equation of motion the second-class constraints all appear, but the first-class constraints do not. We could change this by adding the sum

$$\{g,\chi_A\}\, v_A$$

and allowing v_A to be arbitrary functions of time just like v_α. This makes the equations of motion have complete symmetry with all constraints ϕ and χ, but only the v_α correspond to the unsolved accelerations we originally found in the Lagrangian formalism.

The first-class and second-class constraints are very different mathematically. The matrix of second-class constraint Poisson brackets $[\{\zeta_\mu, \zeta_{\mu'}\}]$ is non-singular on the surface U, and we can assume that it is non-singular in a finite neighborhood of U. This means, of course, that the inverse matrix with elements $C_{\mu\mu'}$ also exists in this neighborhood. This means that the inverse matrix relation between $\mathbf{\Delta}$ and $\mathbf{C}$ is a strong equality $\mathbf{\Delta C} = \mathbf{I}$.

We may also rewrite the equations of motion using the Dirac bracket with respect to the second-class constraints ζ_μ:

$$\frac{dg}{dt} \approx \{g, H\}^* + \{g, \phi_\beta\} v_\beta.$$

Actually, we can include the second term in the Dirac bracket too, by again looking at a *total Hamiltonian* that includes the contributions from the arbitrary v_β:

$$\begin{aligned} H_T &= H + \phi_\beta v_\beta; \\ \Longrightarrow \frac{dg}{dt} &\approx \{g, H_T\}^*. \end{aligned}$$

When we first defined weak equations, we made sure that the differentiation of q's and p's should be carried out considering all variables independent and only afterwards replace the relations that defined U, This is important for computing Poisson brackets, because if a function vanishes weakly, its Poisson bracket with another arbitrary function will not in general vanish weakly. For Dirac brackets, this is different, because if a function is a second-class constraint ζ_μ, then its Dirac bracket vanishes weakly with *any* other function. Thus if we are working with Dirac brackets rather than Poisson brackets, it does not matter whether we constrain the variables before or after differentiation. This implies that the weak second-class vanishing equations

$$\zeta_\mu \approx 0$$

can be converted into strong equations

$$\zeta_\mu \equiv 0$$

and these equations can eliminate the corresponding number of q's and p's from the theory, expressing them as functions of the remaining q's and p's. This will lead us to the equations of motion in Dirac bracket form, and we will be dealing with a reduced number of q, p variables ($2n$ minus the number of second-class constraints). The only constraints left will be the first-class constraints. This is very similar to how we used the

A-type constraints to eliminate variables from the Lagrangian formalism while keeping the B-type constraints.

There is a further interesting property of first-class constraints to note. Let us take a dynamic function $f(t)$ in the phase space of a system with total Hamiltonian H_T. Consider an arbitrarily small time variation δt; The variation of f under this δt can be written

$$\begin{aligned} f(\delta t) &= f(0) + \dot{f}\delta t \\ &\approx f(0) + \{f, H_T\}^* \delta t \\ &\approx f(0) + \left(\{f, H\}^* + \{f, \phi_\beta\} v_\beta\right) \delta t. \end{aligned}$$

Since v_β is arbitrary at this point, we could have picked a different value, v^*_β, and obtained a mathematically different but physically equivalent $f(\delta t)$, and the difference between the two would be

$$\Delta f(\delta t) = \left(v_\beta - v^*_\beta\right) \delta t \{f, \phi_\beta\}$$

Taking $\epsilon_\beta = \left(v_\beta - v^*_\beta\right) \delta t$, we can write

$$\Delta f(\delta t) = \epsilon_\beta \{f, \phi_\beta\},$$

so, in this sense, the first-class constraints ϕ_β act as generator functions of transformations that do not change the phase space quantities physically. The gauge symmetries, of course, are known to induce such situations.

2.2.3.1.1 Example Let us analyze the system represented by the Lagrangian (2.9) using the Dirac-Bergmann method. This example is also found in (LEMOS, 2018). The canonical momenta are

$$\begin{aligned} p_1 &= \frac{\partial L}{\partial \dot{q}_1} = \dot{q}_1 + q_2, \\ p_2 &= \frac{\partial L}{\partial \dot{q}_2} = (1-\alpha) q_1. \end{aligned}$$

We already have a primary constraint, since p_2 does not depend on any $\dot{q}$.

$$\phi_1 = p_2 - (1-\alpha) q_1 \approx 0.$$

The Hamiltonian, obtained by the Legendre transformation of the Lagrangian, is

$$H = \frac{1}{2}\left(p_1 - q_2\right)^2 - \frac{\beta}{2}(q_1 - q_2)^2, \tag{2.23}$$

so the conservation of ϕ_1 leads us to

$$\phi_2 = \{\phi_1, H\} = \alpha\left(p_1 - q_2\right) - \beta\left(q_1 - q_2\right) \approx 0.$$

The conservation of ϕ_2 itself is then

$$\phi_3 = \{\phi_2, H\} + \lambda\{\phi_2, \phi_1\} = \alpha\beta\left(q_1 - q_2\right) - \beta\left(p_1 - q_2\right) + \lambda\left(\alpha^2 - \beta\right) \approx 0.$$

Now we analyze two cases. If $\alpha^2 = \beta$, then ϕ_3 turns out to be not independent of ϕ_2 and is thus not a new constraint. The total Hamiltonian is then

$$H_T = H + \phi_1\lambda = \frac{1}{2}\left(p_1 - q_2\right)^2 - \frac{\alpha^2}{2}(q_1 - q_2)^2 + \lambda\left(p_2 - (1-\alpha)\,q_1\right).$$

The generalized equations of motion are

$$\begin{aligned}
\dot{q}_1 &= \frac{\partial H_T}{\partial p_1} = p_1 - q_2;\\
\dot{q}_2 &= \frac{\partial H_T}{\partial p_2} = \lambda;\\
\dot{p}_1 &= -\frac{\partial H_T}{\partial q_1} = \alpha^2\left(q_1 - q_2\right) + (1-\alpha)\,\lambda;\\
\dot{p}_2 &= -\frac{\partial H_T}{\partial q_2} = p_1 - q_2 - \alpha^2\left(q_1 - q_2\right),
\end{aligned}$$

with λ an arbitrary function of time.

If $\alpha^2 \neq \beta$, then ϕ_1 and ϕ_2 are second-class constraints with $\{\phi_1, \phi_2\} = \beta - \alpha^2$, and the matrix $\mathbf{C}$ is given by

$$\mathbf{C} = \frac{1}{\alpha^2 - \beta}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

so the Dirac bracket is

$$\{f, g\}^* = \{f, g\} + \frac{\{f, \phi_1\}\{\phi_2, g\} - \{f, \phi_2\}\{\phi_1, g\}}{\alpha^2 - \beta},$$

from which we obtain the relevant Dirac brackets in phase space are

$$\begin{aligned}
\{q_1, q_2\}^* &= \frac{\alpha}{\alpha^2 - \beta};\\
\{q_1, p_1\}^* &= \frac{\alpha - \beta}{\alpha^2 - \beta};\\
\{q_1, p_2\}^* &= 0;\\
\{q_2, p_1\}^* &= -\frac{\beta}{\alpha^2 - \beta};\\
\{q_2, p_2\}^* &= \frac{\alpha\left(\alpha - 1\right)}{\alpha^2 - \beta};\\
\{p_1, p_2\}^* &= \frac{\left(\alpha - 1\right)\left(\alpha - \beta\right)}{\alpha^2 - \beta},
\end{aligned} \tag{2.24}$$

which, combined with the Hamiltonian (2.23), gives us the equations of motion as

$$\begin{aligned}
\dot{q}_1 &= p_1 - q_2;\\
\dot{q}_2 &= \frac{\beta}{\alpha}\left(q_1 - q_2\right);\\
\dot{p}_1 &= \frac{\beta}{\alpha}\left(q_1 - q_2\right);\\
\dot{p}_2 &= \frac{\beta\left(1 - \alpha\right)}{\alpha}\left(q_1 - q_2\right).
\end{aligned}$$

In this case, the equation $\phi_3 \approx 0$ allowed us to determine a fixed value for λ, which was eliminated from the theory.

2.2.3.2 Faddeev-Jackiw method

There is an alternate method introduced by L. Faddeev and R. Jackiw (FADDEEV; JACKIW, 1988) and expanded by L. Barcelos-Neto and C. Wotzasek in the case of singular systems (BARCELOS-NETO; WOTZASEK, 1992a). This method relies on concepts of symplectic geometry, mainly Darboux' theorem for singular 2-forms (for details see B.2).

With this in mind, a convenient way to describe the Faddeev-Jackiw method is using the tautological (or canonical) 1-form of the phase space manifold. The tautological 1-form can be thought of as a mathematical object that intertwines each coordinate and its canonical momenta. This is done by constructing a covector dual to the gradient of the generalized coordinates and having each basis element (represented by dq^n) contain their respective momentum as a coefficient: $\theta = p_j dq^j$. We already have written an expression for the Lagrangian in terms of the tautological one-form in (2.5)

$$L\left(\eta\right) = \theta_\mu\left(\eta\right)\dot{\eta}^\mu - H,$$

where $\theta_\mu(\eta)$ is the μ-th coefficient of the tautological 1-form.

This will be the basis for the Faddeev-Jackiw formalism. Notice that the Lagrangian written this way is linear in the velocities $\dot{\eta}^\mu$. For this reason, this is also known as a first-order formalism.

We already know the equations of motion (2.6)

$$\frac{\delta S}{\delta \eta^\mu} = \left(\frac{\partial \theta_\mu}{\partial \eta^\nu} - \frac{\partial \theta_\nu}{\partial \eta^\mu}\right)\dot{\eta}^\nu - \frac{\partial H}{\partial \eta^\mu} = 0.$$

The term in parentheses can be compared to the coefficients of the canonical 2-form, defined as

$$\omega = \frac{1}{2}\underbrace{\left(\frac{\partial \theta_\mu}{\partial \eta^\nu} - \frac{\partial \theta_\nu}{\partial \eta^\mu}\right)}_{\omega_{\mu\nu}} d\eta^\nu \wedge d\eta^\mu,$$

where the $\wedge$ symbol is the exterior product, which is an alternating operator on the space of differential forms.

Using this notation, the equations of motion are just

$$\omega_{\mu\nu}\dot{\eta}^\nu = \frac{\partial H}{\partial \eta^\mu} \tag{2.25}$$

which can be read as a matricial equation: the terms of $\omega_{\mu\nu}$ form a matrix, multiplying a vector $\dot{\eta}^\nu$ to obtain another vector $\frac{\partial H}{\partial \eta^\mu}$. Alternatively, we can express the equations of motion with Poisson brackets, using

$$\{\eta^\mu, \eta^\nu\} = \omega^{\mu\nu}, \tag{2.26}$$

where $\omega^{\mu\nu}\omega_{\mu\nu} = \delta^\nu_\mu$.

For singular systems, an issue arises: the tautological one-form is degenerate (i.e. it is not symplectic). This has the implication that the matrix of elements $\omega_{\mu\nu}$ is singular, and thus we will not find solutions for all $\dot{\eta}^\nu$ directly from (2.25), while the inverse matrix $\omega^{\mu\nu}$ in (2.26) will not exist.

In this case, then there exists at least one eigenvector v such that

$$\omega_{\mu\nu}v^\nu = 0,$$

or, contracting with the equations of motion (2.25),

$$v^\mu \frac{\partial H}{\partial \eta^\mu} = 0.$$

The functions $\gamma_r = v_r^\mu \frac{\partial H}{\partial \eta^\mu}$ are constraints of this system. We then have the set constraint equations,

$$\gamma_r = 0.$$

The next step is to rewrite the Hamiltonian by enforcing these constraint equations

$$H_1 = H|_{\gamma_r=0}\,.$$

Now, we proceed by adding each γ_r to the Lagrangian with a multiplier

$$L_1 = \theta_\mu \dot{\eta}^\mu - H_1 - \lambda^r \dot{\gamma}_r.$$

The action associated with this Lagrangian is

$$S = \int \left[\theta_\mu \dot{\eta}^\mu - H_1 - \lambda^r \dot{\gamma}_r\right] dt,$$

which, after an integration by parts disregarding the resulting surface term, is seen to be the same as

$$S = \int \left[\theta_\mu \dot{\eta}^\mu - H_1 + \gamma_r \dot{\lambda}^r\right] dt,$$

and from this, the variational principle gives us the equations of motion

$$\begin{cases} \omega_{\mu\nu}\dot{\eta}^\nu + \frac{\partial \gamma_r}{\partial \eta^\mu}\dot{\lambda}^r & = \frac{\partial H_1}{\partial \eta^\mu}, \\ \dot{\gamma}_r & = 0. \end{cases} \tag{2.27}$$

If at this stage any of the canonical coordinates does not appear in this system of equations, we can simply ignore it moving forward.

With the same integration by parts we performed in the action above, we may also rewrite the Lagrangian as

$$L_1 = \theta_\mu \dot{\eta}^\mu + \gamma_r \dot{\lambda}^r - H_1.$$

This Lagrangian hints that we can define a new tautological 1-form containing the multipliers as additional "coordinates":

$$\theta_1 = \theta_\nu d\eta^\nu + \gamma_r d\lambda^r,$$

from which, of course, we may define a new sympletic form by taking its exterior derivative.

$$\omega_1 = d\theta_1 = \omega + \frac{\partial \gamma_r}{\partial \eta^\mu} d\eta^\mu \wedge d\lambda^r,$$

and we obtain a new matrix representation of the coefficients $\omega_{1\mu\nu}$:

$$\omega_{1\mu\nu} = \begin{bmatrix} 0 & \cdots & \omega_{1n} & \frac{\partial\gamma_1}{\partial\eta^1} & \cdots & \frac{\partial\gamma_r}{\partial\eta^1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \omega_{n1} & \cdots & 0 & \frac{\partial\gamma_1}{\partial\eta^n} & \cdots & \frac{\partial\gamma_r}{\partial\eta^n} \\ -\frac{\partial\gamma_1}{\partial\eta^1} & \cdots & -\frac{\partial\gamma_r}{\partial\eta^1} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial\gamma_1}{\partial\eta^n} & \cdots & -\frac{\partial\gamma_r}{\partial\eta^n} & 0 & \cdots & 0 \end{bmatrix},$$

where $\omega_{\mu\nu}$ is the original symplectic form. From this point, we check if the matrix above is singular; if it is, we use its null eigenvectors to obtain more constraints γ_{1r}, and so on. Eventually, after a finite number of iterations, we will obtain a nonsingular matrix, from which we can invert to get expressions for the fundamental Poisson brackets, being

$$\{\eta^\mu, \eta^\nu\} = \omega^{\mu\nu},$$

which we then use to compute the final complete set of equations of motion.

2.2.3.2.1 Example Finally, let us work the example Lagrangian (2.9) in the Faddeev-Jackiw method. This example, in the $\alpha^2 \neq \beta$ case, is also found in (CARO *et al.*, 2021). We already computed the Hamiltonian and canonical momenta definitions in section 2.2.3.1.1, which gave us a primary constraint $p_2 = \frac{\partial L}{\partial \dot{q}_2} = (1-\alpha)q_1$. We now write the Lagrangian in first-order form:

$$\begin{aligned} L &= p_1\dot{q}_1 + p_2\dot{q}_2 - H \\ &= p_1\dot{q}_1 + (1-\alpha)\, q_1\dot{q}_2 - H, \end{aligned}$$

so that the tautological 1-form is

$$\theta = p_1 dq_1 + (1-\alpha)\, q_1 dq_2.$$

The canonical symplectic form is

$$\omega = dp_1 \wedge dq_1 + (1-\alpha)\, dq_1 \wedge dq_2,$$

and its matrix representation is

$$\mathbf{\Omega} = \begin{bmatrix} 0 & 1-\alpha & -1 \\ \alpha-1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

This is a 3×3 matrix since our system only has 3 independent canonical coordinates (q_1, q_2 and p_1). Clearly this matrix is singular: the second and third rows (or columns) are linearly dependent. It has a null eigenvector $\mathbf{v} = \begin{bmatrix} 0 & 1 & 1-\alpha \end{bmatrix}$, thus the equations of motion give us the constraint

$$\gamma = \frac{\partial H}{\partial q_2} + (1-\alpha)\frac{\partial H}{\partial p_1} = 0,$$

with

$$\begin{aligned} \frac{\partial H}{\partial q_2} &= q_2 - p_1 + \beta(q_2 - q_1); \\ \frac{\partial H}{\partial p_1} &= p_1 - q_2, \end{aligned}$$

such that

$$\begin{aligned} \gamma &= q_2 - p_1 + \beta(q_2 - q_1) + (1-\alpha)(p_1 - q_2) \\ &= -\alpha(p_1 - q_2) + \beta(q_2 - q_1) = 0, \\ \therefore \gamma = 0 \implies & \frac{\alpha}{\beta}(p_1 - q_2) = q_2 - q_1. \end{aligned}$$

The constrained Hamiltonian and Lagrangian are

$$\begin{aligned} H_1 = H|_{\gamma=0} &= \frac{1}{2}(p_1 - q_2)^2 - \frac{\alpha^2}{2\beta}(p_1 - q_2)^2 \\ &= \frac{1}{2}\left(1 - \frac{\alpha^2}{\beta}\right)(p_1 - q_2)^2, \\ L_1 &= p_1\dot{q}_1 + (1-\alpha)q_1\dot{q}_2 - [\alpha(p_1 - q_2) - \beta(q_2 - q_1)]\dot{\lambda}_1 - H_1. \end{aligned}$$

Note that if $\alpha^2 = \beta$, H_1 vanishes. Let us assume $\alpha^2 \neq \beta$ for now. The constrained 1-form is

$$\theta_1 = p_1 dq_1 + (1-\alpha)q_1 dq_2 + [-\alpha(p_1 - q_2) + \beta(q_2 - q_1)]\,d\lambda_1,$$

and the 2-form

$$\begin{aligned} \omega_1 &= d\theta_1 \\ &= dp_1 \wedge dq_1 + (1-\alpha)\,dq_1 \wedge dq_2 - \alpha dp_1 \wedge d\lambda_1 + \alpha dq_2 \wedge d\lambda_1 + \beta dq_1 \wedge d\lambda_1 - \beta dq_2 \wedge d\lambda_1. \end{aligned}$$

At this point, the Hamiltonian equations of motion (2.27) are

$$\begin{cases} -\dot{p}_1 + (1-\alpha)\,\dot{q}_2 + \beta\dot{\lambda}_1 & = 0; \\ (\alpha-1)\,\dot{q}_1 + (\alpha-\beta)\,\dot{\lambda}_1 & = \left(1-\frac{\alpha^2}{\beta}\right)(q_2-p_1); \\ \dot{q}_1 - \alpha\dot{\lambda}_1 & = \left(1-\frac{\alpha^2}{\beta}\right)(p_1-q_2); \\ -\alpha\,(\dot{p}_1-\dot{q}_2) + \beta\,(\dot{q}_2-\dot{q}_1) & = 0, \end{cases}$$

which can be simplified to

$$\begin{cases} \dot{\lambda}_1 & = \frac{\alpha}{\beta}\dot{q}_1, \\ \dot{q}_1 & = \dot{q}_2 = \dot{p}_1, \\ \left(1-\frac{\alpha^2}{\beta}\right)\dot{q}_2 & = \left(1-\frac{\alpha^2}{\beta}\right)(p_1-q_2). \end{cases} \tag{2.28}$$

If we continue the F-J analysis, we get the matrix

$$\mathbf{\Omega}_1 = \begin{bmatrix} 0 & 1-\alpha & -1 & \beta \\ \alpha-1 & 0 & 0 & \alpha-\beta \\ 1 & 0 & 0 & -\alpha \\ -\beta & -\alpha+\beta & \alpha & 0 \end{bmatrix}.$$

This matrix has determinant $(\alpha^2-\beta)^2$. Since we assumed $\alpha^2 \neq \beta$, then we can find its inverse as

$$\mathbf{\Omega}_1^{-1} = \begin{bmatrix} 0 & \frac{\alpha}{\alpha^2-\beta} & \frac{\alpha-\beta}{\alpha^2-\beta} & 0 \\ -\frac{\alpha}{\alpha^2-\beta} & 0 & -\frac{\beta}{\alpha^2-\beta} & -\frac{1}{\alpha^2-\beta} \\ \frac{\beta-\alpha}{\alpha^2-\beta} & \frac{\beta}{\alpha^2-\beta} & 0 & \frac{\alpha-1}{\alpha^2-\beta} \\ 0 & \frac{1}{\alpha^2-\beta} & \frac{1-\alpha}{\alpha^2-\beta} & 0 \end{bmatrix}.$$

Notice how the first three rows and columns, reflecting the phase space coordinates q_1, q_2, p_1 are exactly the same as the Dirac bracket values obtained in (2.24), and give us the following relations:

$$\begin{aligned} \left\{\eta_1^1, \eta_1^2\right\} &= \left(\mathbf{\Omega}_1^{-1}\right)_{12} = \{q_1, q_2\} = \frac{\alpha}{\alpha^2-\beta}; \\ \left\{\eta_1^1, \eta_1^3\right\} &= \left(\mathbf{\Omega}_1^{-1}\right)_{13} = \{q_1, p_1\} = \frac{\alpha-\beta}{\alpha^2-\beta}; \\ \left\{\eta_1^2, \eta_1^3\right\} &= \left(\mathbf{\Omega}_1^{-1}\right)_{23} = \{q_2, p_1\} = \frac{-\beta}{\alpha^2-\beta}; \end{aligned}$$

From this we obtain the equations of motion

$$\dot{q}_1 = \{q_1, H_1\} = p_1 - q_2;$$

$$\dot{q}_2 = \{q_2, H_1\} = p_1 - q_2;$$
$$\dot{p}_1 = \{p_1, H_1\} = p_1 - q_2.$$

and imposing the primary and secondary constraints, we obtain the same equations as the Dirac-Bergmann method gave us.

$$\begin{aligned}
\dot{q}_1 &= p_1 - q_2; \\
\dot{q}_2 &= \frac{\beta}{\alpha}\left(q_1 - q_2\right); \\
\dot{p}_1 &= \frac{\beta}{\alpha}\left(q_1 - q_2\right); \\
\dot{p}_2 &= \frac{\beta\left(1-\alpha\right)}{\alpha}\left(q_1 - q_2\right).
\end{aligned}$$

Now, if $\beta = \alpha^2$ then the matrix $\mathbf{\Omega}_1$ is still singular.

$$\mathbf{\Omega} = \begin{bmatrix} 0 & 1-\alpha & -1 & \alpha^2 \\ \alpha - 1 & 0 & 0 & -\alpha\left(\alpha-1\right) \\ 1 & 0 & 0 & -\alpha \\ -\alpha^2 & \alpha\left(\alpha-1\right) & \alpha & 0 \end{bmatrix}.$$

It has 2 null eigenvectors:

$$\ker \mathbf{\Omega}_1 = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 1-\alpha \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha \\ 0 \\ \alpha^2 \\ 1 \end{bmatrix} \right\}.$$

However, since the constraint makes $H_1 = 0$, there are no new constraints to be obtained. The equations of motion are

$$\begin{cases} \left(1-\alpha\right)\dot{q}_2 - \dot{p}_1 + \alpha^2\dot{\lambda}_1 &= 0, \\ \left(\alpha-1\right)\dot{q}_1 - \alpha\left(\alpha-1\right)\dot{\lambda}_1 &= 0, \\ \dot{q}_1 - \alpha\dot{\lambda}_1 &= 0, \\ -\alpha^2\dot{q}_1 + \alpha\left(\alpha-1\right)\dot{q}_2 + \alpha\dot{p}_1 &= 0. \end{cases} \implies \begin{cases} \left(1-\alpha\right)\dot{q}_2 - \dot{p}_1 &= -\alpha^2\dot{\lambda}_1, \\ \dot{q}_1 &= \alpha\dot{\lambda}_1. \end{cases}$$

Since $\dot{p}_1$ and $\dot{q}_2$ also do not get independent equations in terms of ps and qs, we may take either of them as an arbitrary function of time. We can recover the Dirac-Bergmann equations for this case by setting the arbitrary function $\lambda_1 = \frac{1}{\alpha}\left(q_1 - q_2\right)$ and $\dot{q}_2 = \lambda$ (the Dirac-Bergmann Lagrange multiplier).

3 The electromagnetic field as a singular system

In this chapter, we will use the methods described in the previous chapter, together with the field theoretic extensions of Lagrangian and Hamiltonian mechanics, using the free electromagnetic field as an example.

3.1 Lagrangian field theory and the free electromagnetic field

In 4-dimensional spacetime Lagrangian field theory, the action, where a set of fields $\varphi_\mu(\mathbf{x}, t)$ replace the general coordinates $q_j(t)$ may be written as

$$S = \int_\Sigma L(\varphi, \partial\varphi, \mathbf{x}, t)\, dt,$$

where $\partial\varphi$ is the 4-gradient of φ. More importantly, the Lagrangian itself may be taken as the 3-space integral of another function, the Lagrangian density $\mathcal{L}$

$$L = \int_V \mathcal{L}(\varphi, \partial\varphi, \mathbf{x}, t)\, d^3x,$$

so that we have separated the original action's integration surface Σ as a 3-space volume V and a time interval $[t_1, t_2]$.

By the variational principle we get the Lagrange field equations in terms of $\mathcal{L}$

$$\frac{\delta S}{\delta\varphi_\mu} = \frac{\partial\mathcal{L}}{\partial\varphi_\mu} - \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu\varphi_\mu)}\right) = 0,$$

or, expanding the second term

$$\frac{\partial\mathcal{L}}{\partial\varphi_\mu} - \frac{\partial^2\mathcal{L}}{\partial x^\nu\partial(\partial_\nu\varphi_\mu)} - \frac{\partial^2\mathcal{L}}{\partial(\partial_\mu\varphi_\nu)\partial\varphi_\rho}\partial_\nu\varphi_\rho - \frac{\partial^2\mathcal{L}}{\partial(\partial_\nu\varphi_\mu)\partial(\partial_\rho\varphi_\sigma)}\partial_\nu\partial_\rho\varphi_\sigma = 0. \tag{3.1}$$

This time, to check for singularity and constraints, since we are looking to solve for the last term with the second time derivatives of φ in Minkowski spacetime with signature $(+---)$ (we will maintain this convention for the rest of the work), we need the Hessian with respect to the x_0 derivatives of fields

$$M_{\mu\nu} = \frac{\partial^2 \mathcal{L}}{\partial\left(\partial_0\varphi_\mu\right)\partial\left(\partial_0\varphi_\nu\right)},$$

from which we can verify the system's singularity by a similar process as used in the previous chapter.

One of the most important fields from classical field theory is the free electromagnetic field. Its Lagrangian density is given by

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{3.2}$$

where F is the electromagnetic tensor, a rank 2 antisymmetric tensor with elements given by

$$F_{\mu\nu} = \begin{pmatrix} 0 & \frac{1}{c}E_x & \frac{1}{c}E_y & \frac{1}{c}E_z \\ -\frac{1}{c}E_x & 0 & -B_z & B_y \\ -\frac{1}{c}E_y & B_z & 0 & -B_x \\ -\frac{1}{c}E_z & -B_y & B_x & 0 \end{pmatrix},$$

or in contravariant form

$$F^{\mu\nu} = g^{\mu\rho}F_{\rho\sigma}g^{\sigma\nu} = \begin{pmatrix} 0 & -\frac{1}{c}E_x & -\frac{1}{c}E_y & -\frac{1}{c}E_z \\ \frac{1}{c}E_x & 0 & -B_z & B_y \\ \frac{1}{c}E_y & B_z & 0 & -B_x \\ \frac{1}{c}E_z & -B_y & B_x & 0 \end{pmatrix},$$

where $g^{\mu\nu}$ is the Minkowski metric tensor, with signature $(+---)$. Alternatively,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

where $A_\mu = \left(\frac{1}{c}\phi, \mathbf{A}\right)$ is the 4-potential. We will use these A^μ elements as our fields in the Lagrangian field equations. The Lagrangian does not depend on A_μ terms themselves, only on derivatives, so the Lagrangian equations of motion give us the divergence equation

$$\partial_\mu\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = \partial_\mu F^{\mu\nu} = 0.$$

The Hessian matrix elements are then

$$\frac{\partial^2 \mathcal{L}}{\partial\left(\partial_0 A_\mu\right)\partial\left(\partial_0 A_\nu\right)} = \delta_{0\mu}\delta_{0\nu} - \delta_{\mu\nu}.$$

This gives us the matrix

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

which is clearly singular with a kernel generated by $\begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$. We have the constraint

$$\begin{aligned} \partial_j \partial^j A^0 - \partial_j \partial^0 A^j &= 0 \\ = \partial_j F^{j0} = 0 &\implies \nabla \cdot \mathbf{E} = 0, \end{aligned}$$

which is Gauss' law from electrostatics (here and for the rest of this work we employ the convention that Greek letters represent 4-dimensional spacetime indices, while Latin letters represent 3-dimensional spatial indices only). Of course, Gauss' law is not a dynamic equation; it is obeyed by the field at all time. Its consistency condition

$$\partial_0 \partial_j F^{j0} = 0$$

is automatically satisfied, since the derivatives $\partial_0 \partial_j$ are symmetric in the indices $0j$, while F^{j0} is skew-symmetric.

The situation is then as follows: for spatial indices we have well-defined equations for the field's second derivatives

$$\partial_0 \partial^0 A^j = \partial_k \partial^k A^j + \partial_0 \partial^j A^0 + \partial_k \partial^j A^k,$$

or, using some vector calculus notation and the definitions of the 4-potential,

$$\ddot{\mathbf{A}} = \nabla^2 \mathbf{A} - \nabla \left(\dot{\phi} + \nabla \cdot \mathbf{A} \right).$$

We cannot, though, determine A_0 exactly. Instead, we take it as an arbitrary function of time on the provision that it satisfies our constraint, which is Gauss' law. This is the familiar gauge symmetry of electrodynamics.

3.2 Hamiltonian approach

Now let us move to field theory from the Hamiltonian formalism. For this, the first step is defining the canonical momenta of the fields, which we obtain from the Lagrangian density $\pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 \varphi_\mu)}$. Much like the Lagrangian, in field theory we can write the Hamiltonian as the spacetime integral of a Hamiltonian density

$$H = \int_V \mathcal{H}\left(\varphi, \nabla\varphi, \pi, \nabla\pi\right) d^3x.$$

Naturally for the Hamiltonian formalism, $\mathcal{H}$ can not depend on time derivatives, and only the 3-space gradient may appear, unlike the Lagrangian density. We can write the action in terms of this density via the Legendre transform

$$S = \int_\Sigma \left(\pi^\mu \partial_0 \varphi_\mu - \mathcal{H}\right) d^4x,$$

i.e. $\mathcal{L} = \pi^\mu \partial_0 \varphi_\mu - \mathcal{H}$.

The variational principle gives the set of Hamiltonian field equations as

$$\begin{cases} \dot{\varphi}_\mu & = \frac{\partial \mathcal{H}}{\partial \pi^\mu} - \nabla \cdot \frac{\partial \mathcal{H}}{\partial(\nabla \pi^\mu)}; \\ \dot{\pi}^\mu & = -\frac{\partial \mathcal{H}}{\partial \varphi_\mu} + \nabla \cdot \frac{\partial \mathcal{H}}{\partial(\nabla \varphi_\mu)}, \end{cases}$$

or, more concisely in terms of the functional/variational derivative

$$\begin{cases} \dot{\varphi}_\mu & = \frac{\delta H}{\delta \pi^\mu}, \\ \dot{\pi}^\mu & = -\frac{\delta H}{\delta \varphi_\mu}. \end{cases}$$

With this, we also need to take care in the definition of field theoretical Poisson brackets, which we may build by using the space integral of functional derivatives

$$\{F, G\} = \int \left(\frac{\delta F}{\delta \varphi_\mu}\frac{\delta G}{\delta \pi^\mu} - \frac{\delta G}{\delta \varphi_\mu}\frac{\delta F}{\delta \pi^\mu}\right) d^3x.$$

Due to the nature of the functional derivative, expressions like this Poisson bracket are not necessarily functions but rather distributions. The canonical Poisson bracket relations, for example, take the form

$$\begin{aligned} \{\varphi_\mu(\mathbf{x}, t), \varphi_\nu(\mathbf{y}, t)\} &= \{\pi^\mu(\mathbf{x}, t), \pi^\nu(\mathbf{y}, t)\} = 0; \\ \{\varphi_\mu(\mathbf{x}, t), \pi^\nu(\mathbf{y}, t)\} &= \delta_\mu^\nu \delta(\mathbf{x} - \mathbf{y}), \end{aligned}$$

where the first delta is the Kronecker delta and the second delta represents the distributional Dirac delta.

With the tools of Hamiltonian field theory now in hand we are ready to deal with the electromagnetic field action, defined by the Lagrangian in (3.2). The canonical momenta are

$$\pi^{\mu} = \frac{\partial \mathcal{L}}{\partial\left(\partial_0 A_{\mu}\right)} = \partial^{\mu} A^0 - \partial^0 A^{\mu}$$
$$= F^{\mu 0}.$$

Or, explicitly,

$$\begin{cases} \pi^0 = 0; \\ \pi^1 = E_x; \\ \pi^2 = E_y; \\ \pi^3 = E_z; \end{cases}$$

This is a constrained field with a primary constraint given promptly by $\pi^0 = 0$.

The Hamiltonian density is

$$\begin{aligned} \mathcal{H} &= \pi^{\mu}\partial_0 A_{\mu} - \mathcal{L} \\ &= \pi^j\left(F_{0j} + \partial_j A_0\right) + \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} \\ &= \pi^j \partial_j A_0 + \frac{1}{2}\pi_j\pi^j + \frac{1}{4}F_{jk}F^{jk}. \end{aligned}$$

Now, let us analyze how the Dirac-Bergmann and Faddeev-Jackiw methods work to obtain the field equations.

3.2.1 Dirac-Bergmann method

For the Dirac-Bergmann method, the next step in obtaining the solution is to enforce the consistency condition on our primary constraint $\gamma_0 = \pi_0$:

$$\{\gamma_0, H\} = \int \left\{\pi_0, \pi^j\partial_j A_0 + \frac{1}{2}\pi_j\pi^j + \frac{1}{4}F_{jk}F^{jk}\right\} d^3x.$$

Recalling that the canonical Poisson bracket relations give us

$$\left\{\pi^0(x), A_{\mu}(y)\right\} = -\delta^0_{\mu}\delta^3(x-y)$$

and

$$\left\{\pi^0, \pi^{\mu}\right\} = 0,$$

the relevant term in the consistency condition is

$$\left\{\pi_0, \pi^j\partial_j A_0\right\} = \pi^j\left\{\pi_0, \partial_j A_0\right\} = -\pi^j\partial_j\delta^3\left(x-y\right).$$

Thus

$$\{\gamma_0, H\} = \int \left[-\pi^j \partial_j \delta(x-y)\right] d^3x = \partial_j \pi^j = 0.$$

This consistency condition is, of course, merely Gauss' law:

$$\nabla \cdot \mathbf{E} = 0.$$

We can rewrite the total Hamiltonian as

$$H_T = \int \left[\frac{1}{2}\pi_j \pi^j + \frac{1}{4} F_{jk} F^{jk} + A_0 \partial_j \pi^j + \lambda \pi^0\right] d^3x,$$

where we have added a Lagrange multiplier, λ, and performed integration by parts in the term $\pi^j \partial_j A_0$, discarding the resulting surface terms (they do not affect the field equations). The field A_0 is then seen to also act as a Lagrange multiplier of the constraint $\gamma_1 = \partial_j \pi^j$.

Now, let us verify the Poisson brackets of γ_1:

$$\{\gamma_1, H_T\} = \int \left\{\partial_j \pi^j, \frac{1}{2}\pi_j \pi^j + \frac{1}{4} F_{jk} F^{jk} + A_0 \partial_j \pi^j + \lambda \pi^0\right\} d^3x.$$

The only possibly non-null term is

$$\begin{aligned}
\left\{\partial_j \pi^j(\mathbf{x}), F_{jk} F^{jk}(\mathbf{y})\right\} &= F^{jk}\left\{\partial_l \pi^l, F_{jk}\right\} + F_{jk}\left\{\partial_l \pi^l, F^{jk}\right\} \\
&= F^{jk} \cancel{(\partial_k \partial_j - \partial_j \partial_k)} \delta(\mathbf{x}-\mathbf{y}) + F^{jk} \cancel{\left(\partial^k \partial^j - \partial^j \partial^k\right)} \delta(\mathbf{x}-\mathbf{y}) \\
&= 0.
\end{aligned}$$

There are no further constraints to be obtained. Furthermore, both constraints are first-class, as they both only depend on π_j and thus their Poisson bracket is zero.

What remains is obtaining the equations of motion. This can easily be done by taking the Poisson brackets with the total Hamiltonian (including the Lagrange multiplier). We obtain

$$\begin{aligned}
\dot{A}_0 &= \{A_0, H_T\} = \lambda; \\
\dot{A}_j &= \{A_j, H_T\} = \pi^j + \partial_j A_0; \\
\dot{\pi}^0 &= \left\{\pi^0, H_T\right\} = -\partial_j \pi^j; \\
\dot{\pi}^j &= \left\{\pi^j, H_T\right\} = \partial_k F^{kj}.
\end{aligned}$$

3.2.2 Faddeev-Jackiw method

Now, we can attempt to solve this system using the Faddeev-Jackiw method. Writing the Lagrangian density in standard form:

$$\begin{aligned}\mathcal{L} &= \pi^{\mu}\dot{A}_{\mu} - \mathcal{H} \\ &= E^{j}\dot{A}_{j} - \mathcal{H}.\end{aligned}$$

with tautological 1-form

$$\theta = E_x dA_x + E_y dA_y + E_z dA_z,$$

and the symplectic 2-form, whose elements may be written $\frac{\delta\theta_{\mu}}{\delta\eta^{\nu}} - \frac{\delta\theta_{\nu}}{\delta\eta^{\mu}}$, is then (up to a Dirac delta)

$$\omega = d\theta = dE_x \wedge dA_x + dE_y \wedge dA_y + dE_z \wedge dA_z$$

Now, we notice that $\pi^0 = 0$ by the known primary constraint, so we can take a coordinate system that consists only of (A_{μ}, π^j). The matrix of coefficients of the 2-form above is

$$\mathbf{\Omega} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} \delta\left(\mathbf{x} - \mathbf{x}'\right).$$

This matrix clearly is singular, and it has the null eigenvector

$$\mathbf{v}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

which gives us the constraint

$$\gamma_1 = \frac{\delta H}{\delta A_0} = 0.$$

This functional derivative is

$$\begin{aligned}\frac{\delta H}{\delta A_0} &= \frac{\delta}{\delta A_0}\left[\int\left(\partial_j A_0 \pi^j + \frac{1}{2}\pi_j\pi^j + \frac{1}{4}F_{jk}F^{jk}\right) d^3x\right] \\ &= \int \pi^J \partial_j \delta\left(\mathbf{x} - \mathbf{x}'\right) d^3x \\ &= -\partial_j \pi^j \\ \implies \gamma_1 &= -\partial_j \pi^j = -\nabla \cdot \mathbf{E}.\end{aligned}$$

Therefore, this constraint γ_2 is Gauss' law. Just as in the Dirac-Bergmann method, instead of adding it as a new Lagrangian multiplier, it is convenient to simply perform integration by parts on the $\pi^j \partial_j A_0$ term and notice that we can replace A_0 with the Lagrangian multiplier for this constraint, λ_1, where $\partial_0 \lambda_1 = A_0$.

$$\begin{aligned}\mathcal{H}_1 &= \frac{1}{2}\pi_j \pi^j + \frac{1}{4} F_{jk} F^{jk} \\ \mathcal{L}_1 &= \pi^j \dot{A}_j - \frac{1}{2}\pi_j \pi^j - \frac{1}{4} F_{jk} F^{jk} - A_0 \partial_j \pi^j .\end{aligned}$$

So the new differential forms are

$$\begin{aligned}\theta_1 &= E_x dA_x + E_y dA_y + E_z dA_z + \left(\partial_j E^j\right) dA_0; \\ \Longrightarrow \omega_1 &= dE_x \wedge dA_x + dE_y \wedge dA_y + dE_z \wedge dA_z \\ &- \partial_j dE^j \wedge d\lambda_1 .\end{aligned}$$

The equations of motion (2.27), obtained by the variational principle, are given by

$$\begin{cases} \int \omega_{\mu\nu}\left(\mathbf{x}, \mathbf{x}'\right) \dot{\eta}^{\nu}\left(\mathbf{x}'\right) d^3\mathbf{x}' + \int \frac{\delta \gamma_r(\mathbf{x}')}{\delta \eta^{\mu}(\mathbf{x})} \dot{\lambda}^r \left(\mathbf{x}'\right) d^3\mathbf{x}' & = \frac{\delta H_1}{\delta \eta^{\mu}(\mathbf{x})}; \\ \dot{\gamma}_r & = 0. \end{cases}$$

For our system,

$$\begin{cases} -\dot{E}_j & = -\partial_k F^{kj}; \\ \dot{A}_j - \partial_j A_0 & = E_j; \\ \partial_j \dot{E}^j & = 0. \end{cases}$$

These are equivalent to the equations we obtained for the Dirac-Bergmann method. If we continue the Faddeev-Jackiw analysis, we find that the matrix representation of ω_1 is

$$\boldsymbol{\Omega}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & \partial_x & \partial_y & \partial_z \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -\partial_x & -1 & 0 & 0 & 0 & 0 & 0 \\ -\partial_y & 0 & -1 & 0 & 0 & 0 & 0 \\ -\partial_z & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} \delta\left(\mathbf{x} - \mathbf{x}'\right).$$

Unlike the mechanical version of the Faddeev-Jackiw method, the matrix has non-numeric entries, and we can interpret it as a linear differential operator with constant coefficients. Inverting the matrix does not become a question of finding $\boldsymbol{\Omega}_1^{-1}$ such that $\boldsymbol{\Omega}_1 \boldsymbol{\Omega}_1^{-1} = \mathbf{I}$ but rather a distributional inverse matrix

$$\mathbf{\Omega}_1\mathbf{\Omega}_1^{-1}\left(\mathbf{x},\mathbf{x}'\right)=\delta\left(\mathbf{x}-\mathbf{x}'\right)\mathbf{I},$$

which is essentially Green's function for the linear differential operator, also represented as a matrix. This can be found directly by finding Green's function for each basis vector

$$\mathbf{\Omega}_1 G_j\left(\mathbf{x},\mathbf{x}'\right)=\delta(\mathbf{x}-\mathbf{x}')\mathbf{e}_j,$$

where j takes the values of the matrix dimensions (in the case of our matrix above, $j=1,\ldots,8$). Then, we build the inverse matrix as

$$\mathbf{\Omega}_1^{-1}\left(\mathbf{x},\mathbf{x}'\right)=\begin{bmatrix}G_1 & \cdots & G_j\end{bmatrix}.$$

An alternative, usually more straightforward method for an operator with constant coefficients is simply to take the Fourier transform of each matrix coefficient, from which we obtain a numerical matrix that we can manipulate and then bring back to differential operator coefficients by performing an inverse Fourier transform.

Whatever option we choose, we will notice the matrix above is still singular. This is obvious since it is skew-symmetric but has an odd number of columns and rows. It still has a null eigenvector, given by $\mathbf{v}=\begin{bmatrix}1 & -\partial_x & -\partial_y & -\partial_z & 0 & 0 & 0\end{bmatrix}^T$. The resulting constraint would be

$$\gamma_2=\frac{\delta H_1}{\delta A_0}-\partial_j\frac{\delta H_1}{\delta A_j}=-\partial_j\pi^j=\gamma_1,$$

so we have no new constraints. We need to fix a gauge to proceed. Take the Coulomb gauge for example, $\partial_j A^j=0$. Adding it as a new Lagrange multiplier, we have a new 2-form and symplectic matrix

$$\begin{aligned}\omega_1 &= dE_x\wedge dA_x+dE_y\wedge dA_y+dE_z\wedge dA_z\\ &\quad-\partial_j dE^j\wedge d\lambda_1+\partial_j dA_j\wedge d\lambda_2,\end{aligned}$$

$$\mathbf{\Omega}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & \partial_x & \partial_y & \partial_z & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \partial_x \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \partial_y \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \partial_z \\ -\partial_x & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\partial_y & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -\partial_z & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -\partial_x & -\partial_y & -\partial_z & 0 & 0 & 0 & 0 \end{bmatrix} \delta\left(\mathbf{x}-\mathbf{x}'\right).$$

This matrix is nonsingular. Its inverse, obtained by the methods described above, is

$$\Omega_2^{-1} = \begin{bmatrix} 0 & \frac{\partial_x}{\nabla^2} & \frac{\partial_y}{\nabla^2} & \frac{\partial_z}{\nabla^2} & 0 & 0 & 0 & -\frac{1}{\nabla^2} \\ -\frac{\partial_x}{\nabla^2} & 0 & 0 & 0 & -1+\frac{\partial_x^2}{\nabla^2} & \frac{\partial_x\partial_y}{\nabla^2} & \frac{\partial_x\partial_z}{\nabla^2} & \frac{\partial_x^2}{\nabla^2} \\ \frac{\partial_y}{\nabla^2} & 0 & 0 & 0 & \frac{\partial_y\partial_x}{\nabla^2} & -1+\frac{\partial_y^2}{\nabla^2} & -\frac{\partial_y\partial_z}{\nabla^2} & \frac{\partial_y^2}{\nabla^2} \\ -\frac{\partial_z}{\nabla^2} & 0 & 0 & 0 & \frac{\partial_z\partial_x}{\nabla^2} & \frac{\partial_z\partial_y}{\nabla^2} & -1+\frac{\partial_z^2}{\nabla^2} & \frac{\partial_z^2}{\nabla^2} \\ 0 & 1-\frac{\partial_x^2}{\nabla^2} & -\frac{\partial_x\partial_y}{\nabla^2} & -\frac{\partial_x\partial_z}{\nabla^2} & 0 & 0 & 0 & 0 \\ 0 & -\frac{\partial_y\partial_x}{\nabla^2} & 1-\frac{\partial_y^2}{\nabla^2} & \frac{\partial_y\partial_z}{\nabla^2} & 0 & 0 & 0 & 0 \\ 0 & -\frac{\partial_z\partial_x}{\nabla^2} & -\frac{\partial_z\partial_y}{\nabla^2} & 1-\frac{\partial_z^2}{\nabla^2} & 0 & 0 & 0 & 0 \\ \frac{1}{\nabla^2} & -\frac{\partial_x^2}{\nabla^2} & -\frac{\partial_y^2}{\nabla^2} & -\frac{\partial_z^2}{\nabla^2} & 0 & 0 & 0 & 0 \end{bmatrix},$$

from which we obtain the (Coulomb) *gauge-fixed* Poisson brackets of A_j and E_j

$$\left\{A_j\left(\mathbf{x}\right), E_k\left(\mathbf{x}'\right)\right\} = \int \left(\delta_{jk} - \frac{\partial_j\partial_k}{\nabla^2}\right)\delta\left(\mathbf{x}-\mathbf{x}'\right)d^3x.$$

So, with the Hamiltonian H_1, we obtain the equations of motion as

$$\begin{aligned} \dot{A}_j &= \int \left\{A_j, \mathcal{H}_1\right\} d^3x = \left(\delta_{jk} - \frac{\partial_j\partial_k}{\nabla^2}\right) E^k, \\ \dot{E}_j &= \int \left\{E_j, \mathcal{H}_1\right\} d^3x = \left(\delta_{jk} - \frac{\partial_j\partial_k}{\nabla^2}\right) \partial_l F^{lj} = \partial_k F^{kj}. \end{aligned}$$

We see that the field equations for $\dot{A}_j$ are now no longer the complete electric fields, but only its transverse parts. Thus, by forcing the vector potential to be transverse with the Coulomb gauge, we see that the only the transverse parts of A_j and E_j are dynamically meaningful; the longitudinal part is related to the gauge freedom.

We also have the gauge consistency condition, which for the Lorentz gauge is

$$\partial_j \dot{A}^j = 0.$$

Since $\dot{A}^j = \partial^j A^0 - F^{0j}$, and Gauss' law says $\partial_j F^{0j} = 0$, this means that A_0 is an harmonic function in the space coordinates

$$\partial_j \partial^j A^0 = 0.$$

Thus, the gauge-fixing condition allows us to obtain an equation for A_0.

4 Applications to nonlinear electrodynamics

In the previous chapter, we introduced how singular system formalisms may be applied to field theory using classical, Maxwell electrodynamics. This classical theory of electrodynamics has many mathematical properties that make it easier to work with. However, by the 20th century, advancements in our understanding of theoretical physics led to the development of several models of electromagnetic fields with different properties that, nonetheless, remained physically interesting. In particular, Maxwell's electrodynamics are linear, at least in their microscopic formulations (in macroscopic settings, nonlinear effects may occur at different materials). Still, many generalized *nonlinear electrodynamics* formalisms have proposed by the likes of Born and Infeld (Born-Infeld electrodynamics) and Heisenberg and Euler (Heisenberg-Euler electrodynamics). These types of theories remain of interest in theoretical physics to this day.

(SOROKIN, 2022) and (BIALYNICKI-BIRULA, 1983) have provided in-depth introductions to the theory of nonlinear electrodynamics (NLED), which we shall briefly summarize here.

The Maxwell Lagrangian density we know (3.2) is composed of a scalar obtained by contracting the electromagnetic tensor with itself. This scalar is Lorentz invariant. There is another notable Lorentz invariant in electromagnetism, which can be obtained contracting the electromagnetic tensor with its Hodge dual F^*, explicitly written as

$$F^{*\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} = \begin{bmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & \frac{1}{c}E_x & -\frac{1}{c}E_y \\ B_y & -\frac{1}{c}E_x & 0 & \frac{1}{c}E_z \\ B_z & \frac{1}{c}E_y & -\frac{1}{c}E_z & 0 \end{bmatrix}.$$

Let us denote these two Lorentz invariants as

$$\begin{cases} \mathcal{S} & := F_{\mu\nu}F^{\mu\nu}, \\ \mathcal{P} & := F_{\mu\nu}F^{*\mu\nu}. \end{cases} \tag{4.1}$$

Then, any Lagrangian density that is a function of only $\mathcal{S}$ and $\mathcal{P}$ will also give rise to a Lorentz invariant action by transitivity

$$S = \int \mathcal{L}\left(\mathcal{S}, \mathcal{P}\right) d^4x.$$

The Euler-Lagrange field equations for such an action will give us

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial F_{\mu\nu}} \right) = 0,$$

or equivalently

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \mathcal{S}} F^{\mu\nu} + \frac{\partial \mathcal{L}}{\partial \mathcal{P}} F^{*\mu\nu} \right) = 0. \tag{4.2}$$

This action is the most general starting point for NLED theories. If we want to impose certain conditions other than Lorentz invariance, we may need other relations regarding the Lagrangian. For example, Maxwell electrodynamics is symmetric under scale transformations (dilations), but the general NLED action above need not be. There is a further condition needed to guarantee this additional symmetry, being an homogeneity

$$\mathcal{L}\left(\alpha\mathcal{S}, \alpha\mathcal{P}\right) = \alpha\mathcal{L}\left(\mathcal{S}, \mathcal{P}\right).$$

This condition guarantees that, if we transform spacetime via a scale transformation $x^\mu \to \sigma x^\mu$, the Lagrangian density transforms as $\sigma^{-4}\mathcal{L}\left(\mathcal{S}, \mathcal{P}\right)$. This type of dilation symmetry is important since in our construction of nonlinear electrodynamics, it is the only non-trivial condition for our action to be conformal invariant, much like Maxwell electrodynamics is.

Another celebrated property of Maxwell electrodynamics is the duality rotation symmetry. This is a stronger, continuous formulation of the well-known electric-magnetic duality. This duality rotation is a transformation acting on the electromagnetic tensor $F_{\mu\nu}$ and its dual via a rotation matrix

$$\begin{bmatrix} F'_{\mu\nu} \\ F'^{*}_{\mu\nu} \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} F_{\mu\nu} \\ F^*_{\mu\nu} \end{bmatrix}.$$

Taking $\alpha = \frac{\pi}{2}$ rad is equivalent to swapping the electric and magnetic fields.

For the NLED action, this transformation takes the form

$$\begin{bmatrix} -2\frac{\partial\mathcal{L}(\mathcal{S}',\mathcal{P}')}{\partial F'_{\mu\nu}} \\ F'^{*}_{\mu\nu} \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} -2\frac{\partial\mathcal{L}(\mathcal{S},\mathcal{P})}{\partial F_{\mu\nu}} \\ F^*_{\mu\nu} \end{bmatrix}.$$

and a condition for this invariance is that the following equation holds

$$\mathcal{P} = \mathcal{P}\left[\left(\frac{\partial \mathcal{L}}{\partial \mathcal{S}}\right)^2 - \left(\frac{\partial \mathcal{L}}{\partial \mathcal{P}}\right)^2\right] - 2\mathcal{S}\frac{\partial \mathcal{L}}{\partial \mathcal{S}}\frac{\partial \mathcal{L}}{\partial \mathcal{P}}.$$

One of the most notable theories of nonlinear electrodynamics was proposed by Born and Infeld (BORN; INFELD, 1934). During a moment when quantum mechanics was gaining traction and physicists were still beginning to work out how to possibly unify it with relativity, their proposal was a theory where the electric field self-energy was limited. The classical Maxwell Lagrangian (3.2) depends only on $\mathcal{S}$, and can be written in terms of the electric and magnetic field magnitudes as

$$\mathcal{L} = \frac{1}{2}\left(\frac{1}{c^2}\mathbf{E}^2 - \mathbf{B}^2\right)$$

and the Hamiltonian density (which is also the energy density)

$$\mathcal{H} = \frac{1}{8\pi}\left(\frac{1}{c^2}\mathbf{E}^2 + \mathbf{B}^2\right).$$

However, this may lead to undefined behavior near the origin. Consider the electric field generated by a static charge by Coulomb's law

$$E\left(r\right) = \frac{e}{r^2}.$$

Clearly, the energy density for such a field diverges as $r \to 0$ (the self-energy).

The Born-Infeld Lagrangian incorporates a parameter τ, which we will call the *tension*, as follows:

$$\mathcal{L} = \tau^2\left(1 - \sqrt{1 + \frac{\mathcal{S}}{2\tau^2} - \frac{\mathcal{P}^2}{16\tau^4}}\right).$$

We can recover the Maxwell Lagrangian from the weak field limit $F_{\mu\nu} \ll \tau$.

More recently, a new family of related theories have been suggested (KRUGLOV, 2017). They are built from the Born-Infeld Lagrangian introducing two new parameters γ and σ

$$\mathcal{L} = \tau^2\left[1 - \left(1 + \frac{\mathcal{S}}{2\tau^2\sigma} - \frac{\gamma\mathcal{P}^2}{32\sigma\tau^4}\right)^\sigma\right]. \tag{4.3}$$

This reduces to the Born-Infeld Lagrangian when $\sigma = \frac{1}{2}$ and $\gamma = 1$. Using (4.2), we get the field equations

$$\partial_\mu \left[\left(1 + \frac{\mathcal{S}}{4\tau^2\sigma} - \frac{\gamma \mathcal{P}^2}{32\sigma\tau^4} \right)^{\sigma-1} \left(-\frac{1}{2} F^{\mu\nu} + \frac{\gamma \mathcal{P}}{16\tau^2} F^{*\mu\nu} \right) \right] = 0.$$

Kruglov's motivation for generalizing the Born-Infeld model in this way was that the Born-Infeld Lagrangian did not account for vacuum birefringence in external magnetic fields, which is an effect predicted by quantum electrodynamics. It can, however, be modeled by this type of generalization.

Now, we will show how to obtain the field equations and constraints for a general NLED action using a "Jordan-Palatini" frame and the Fadeev-Jackiw method. After that, we will use the Lagrangian (4.3) as an application.

4.1 Jordan-Palatini frame for non-linear electrodynamics

We want to work with theories such as those described by action (4.3) using the Faddeev-Jackiw formalism. To this end, we will first introduce a way to rewrite NLED actions with auxiliary fields. This is inspired by the use of Jordan-Palatini frames in $f(R)$ theories in general relativity (SOTIRIOU; FARAONI, 2010). Such theories are generalizations of the Einstein-Hilbert action, which depends on the Ricci scalar R, to actions that depend on *nonlinear functions* of this scalar, i.e. $f(R)$. It is not hard to see a parallel with how nonlinear electrodynamics is a generalization of the Maxwell action; rather than depending just on the scalar $\mathcal{S}$, we now have a function $f(\mathcal{S}, \mathcal{P})$. There are two steps to the process of building a Jordan-Palatini formalism: first, we perform a new Legendre transform with auxiliary fields, which is the Jordan frame, and then we perform the Palatini-style variation, which is a different approach to obtain the field equations.

4.1.1 Jordan frame

Take a general NLED action with Lagrangian $\mathcal{L}(\mathcal{S}, \mathcal{P})$. if we introduce two auxiliary fields χ and ξ so that

$$S = \int \left[\mathcal{L}(\chi, \xi) + \frac{\partial \mathcal{L}}{\partial \chi} (\mathcal{S} - \chi) + \frac{\partial \mathcal{L}}{\partial \xi} (\mathcal{P} - \xi) \right] d^4x.$$

The variational principle applied to χ and ξ then gives

$$\begin{cases} \frac{\partial^2 \mathcal{L}}{\partial \chi^2} (\mathcal{S} - \chi) + \frac{\partial^2 \mathcal{L}}{\partial \chi \partial \xi} (\mathcal{P} - \xi) & = 0, \\ \frac{\partial^2 \mathcal{L}}{\partial \xi \partial \chi} (\mathcal{S} - \chi) + \frac{\partial^2 \mathcal{L}}{\partial \xi^2} (\mathcal{P} - \xi) & = 0, \end{cases}$$

or in matrix form

$$\begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \chi^2} & \frac{\partial^2 \mathcal{L}}{\partial \chi \partial \xi} \\ \frac{\partial^2 \mathcal{L}}{\partial \xi \partial \chi} & \frac{\partial^2 \mathcal{L}}{\partial \xi^2} \end{bmatrix} \begin{bmatrix} (\mathcal{S} - \chi) \\ (\mathcal{P} - \xi) \end{bmatrix} = \mathbf{0}.$$

If the Hessian matrix of second derivatives of $\mathcal{L}$ with respect to χ, ξ is nonsingular, this gives us two constraints

$$\begin{cases} \mathcal{S} & = \chi, \\ \mathcal{P} & = \xi, \end{cases}$$

which, substituted back in the Lagrangian, lead us back to the original NLED Lagrangian.

Defining $\phi = \frac{\partial \mathcal{L}}{\partial \chi}$ and $\psi = \frac{\partial \mathcal{L}}{\partial \xi}$, taking χ and ξ as functions of ϕ and ψ, and defining another function

$$V(\phi, \psi) = \chi(\phi, \psi)\,\phi + \xi(\phi, \psi)\,\psi - \mathcal{L}\left(\chi(\phi, \psi), \xi(\phi, \psi)\right),$$

we can write the action as

$$S = \int \left[\mathcal{S}\phi + \mathcal{P}\psi - V(\phi, \psi)\right] d^4x. \tag{4.4}$$

The field equations obtained from this are

$$\begin{cases} \mathcal{S} & = \frac{\partial V}{\partial \phi}, \\ \mathcal{P} & = \frac{\partial V}{\partial \psi}, \\ \partial_\mu \left(F^{\mu\nu}\phi + F^{*\mu\nu}\psi\right) & = 0. \end{cases} \tag{4.5}$$

This formulation of NLED by means of a Legendre transform with $\mathcal{S}$ and $\mathcal{P}$ is what we will call the *Jordan frame* for nonlinear electrodynamics. Notice how it is in some ways analogue to the Legendre transform that we use in transitioning from the Lagrangian to the Hamiltonian formalism, in the sense that we inverted the definitions of ϕ and ψ to rewrite χ and xi in the final Lagrangian. The result was an action where ϕ and ψ appear as conjugates to $\mathcal{S}$ and $\mathcal{P}$, respectively, and $V(\phi, \psi)$ acts as a sort of potential or Hamiltonian.

4.1.2 Palatini variation

Next, we will perform the "Palatini" variation. In relativity, this means that relations i.e. between the metric and connection, which are normally taken as *a priori* definitions, are not imposed as such, but rather obtained via the variational principle (TSAMPARLIS, 1978). For our Palatini-type variation in NLED, we *do not* immediately write the $F^{\mu\nu}$

and $F^{*\mu\nu}$ tensors as functions of each other and A_μ, but rather assume them to be fully independent of each other, and the relations between them will emerge when analyzing the system, either as constraints (if the relations have no dynamics) or as additional field equations. This, of course, also requires some modifications to the Lagrangian density to ensure that these constraints and field equations can be derived properly.

In the case of our Jordan frame action (4.4), we have to rewrite $\mathcal{S}$ and $\mathcal{P}$ as

$$\begin{cases} \mathcal{S} &= 2F^{\mu\nu}\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) - F_{\mu\nu}F^{\mu\nu}, \\ \mathcal{P} &= 2F^{*\mu\nu}\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) + \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{*\alpha\beta}F^{*\mu\nu}. \end{cases} . \tag{4.6}$$

If we were to assume both tensors $F^{\mu\nu}$ and $F^{*}\mu\nu$ as being written in terms of A_μ as we did before, this is exactly the same as our previous definition (4.1). However, the point of the Palatini formalism is to consider them independent. Our action is still

$$S = \int \left[\mathcal{S}\phi + \mathcal{P}\psi - V\left(\phi,\psi\right)\right] d^4x,$$

but now, rather than just 6 fields to work with (A_μ, ϕ and ψ), we have a total of 38, which includes the 16 independent components of $F^{\mu\nu}$ and 16 independent components of $F^{*\mu\nu}$.

Taking the variation of each term in S, while performing integration by parts on the resulting $\partial_\mu \delta A_\nu$ terms, we get

$$\begin{aligned} \delta\int\left[\mathcal{S}\phi\right] d^4x = &\int \{2\partial_\mu\left(\phi F^{\nu\mu} - \phi F^{\mu\nu}\right)\delta A_\nu \\ &+ \left[2\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\delta F^{\mu\nu} - 2F^{\mu\nu}\delta F_{\mu\nu}\right]\phi \\ &+ \left[2F^{\mu\nu}\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) - F_{\mu\nu}F^{\mu\nu}\right]\delta\phi\} \, d^4x; \\ \delta\int\left[\mathcal{P}\psi\right] d^4x = &\int \{2\partial_\mu\left(\psi F^{*\nu\mu} - \psi F^{*\mu\nu}\right)\delta A_\nu \\ &+ \left[2\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\delta F^{*\mu\nu} + \epsilon_{\mu\nu\alpha\beta}F^{*\alpha\beta}\delta F^{*\mu\nu}\right]\psi \\ &+ \left[2F^{*\mu\nu}\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) + \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{*\alpha\beta}F^{*\mu\nu}\right]\delta\psi\Bigg\} \, d^4x; \\ \delta\int\left[V\left(\phi,\psi\right)\right] d^4x = &\int\left[\frac{\partial V}{\partial\phi}\delta\phi + \frac{\partial V}{\partial\psi}\delta\psi\right] d^4x. \end{aligned}$$

The variational principle gives us the field equations

$$\begin{cases} (\partial_\mu A_\nu - \partial_\nu A_\mu) & = F_{\mu\nu}; \\ 2\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) & = -\epsilon_{\mu\nu\alpha\beta} F^{*\alpha\beta}; \\ \partial_\mu\left(\phi F^{\nu\mu} - \phi F^{\mu\nu} + \psi F^{*\nu\mu} - \psi F^{*\mu\nu}\right) & = 0; \\ \mathcal{S} & = \frac{\partial V}{\partial \phi}; \\ \mathcal{P} & = \frac{\partial V}{\partial \psi}. \end{cases}$$

As expected, the first two equations retrieve the relations between $F^{\mu\nu}$, $F^{*\mu\nu}$ and A_μ, which also implies the antisymmetry of both $F^{\mu\nu}$ and $F^{*\mu\nu}$ tensors. Applying this antisymmetry to the third equation gives us the Jordan frame divergence equation from (4.5). Thus, this set of 5 equations is ultimately equivalent to our original Jordan frame NLED field equations (4.5).

4.2 Jordan-Palatini NLED with the Faddeev-Jackiw method

Finally, in this section, we will use the Faddeev-Jackiw method to analyze the Jordan-frame action (4.4) in the Palatini approach, which will allow us to not only obtain the field equations but also classify the constraint equations. Recall that in the Palatini approach, all components in the Lagrangian density are being taken as independent, and $\mathcal{S}$ and $\mathcal{P}$ are defined by (4.6). Taking the canonical momenta for each component, we have

$$\begin{cases} \Pi_{A_\mu} & = 2\left(F^{0\mu} - F^{\mu 0}\right)\phi + 2\left(F^{*0\mu} - F^{*\mu 0}\right)\psi; \\ \Pi_{F^{\mu\nu}} & = 0; \\ \Pi_{F^{*\mu\nu}} & = 0; \\ \Pi_\phi & = 0; \\ \Pi_\psi & = 0. \end{cases}$$

The resulting Hamiltonian density is

$$\begin{aligned} \mathcal{H} = & -2\left[\phi\left(F^{j\nu} - F^{\nu j}\right) + \psi\left(F^{*j\nu} - F^{*\nu j}\right)\right]\partial_j A_\nu \\ & + F_{\mu\nu}F^{\mu\nu}\phi - \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{*\alpha\beta}F^{*\mu\nu}\psi + V(\phi,\psi), \end{aligned}$$

and Lagrangian

$$\mathcal{L} = 2\left[\left(F^{0j} - F^{j0}\right)\phi + \left(F^{*0j} - F^{*j0}\right)\psi\right]\partial_0 A_j - \mathcal{H}.$$

So the tautological form is

$$\theta = 2\left[\left(F^{0j} - F^{j0}\right)\phi + \left(F^{*0j} - F^{*j0}\right)\psi\right] dA_j,$$

and symplectic 2-form

$$\begin{aligned} \omega = & 2\phi dF^{0j} \wedge dA_j - 2\phi dF^{j0} \wedge dA_j \\ & + 2\psi dF^{*0j} \wedge dA_j - 2\psi dF^{*j0} \wedge dA_j \\ & + 2\left(F^{0j} - F^{j0}\right) d\phi \wedge dA_j + 2\left(F^{*0j} - F^{*j0}\right) d\psi \wedge dA_j. \end{aligned}$$

The resulting matrix representation has dimensions 38×38.

$$\mathbf{\Omega} = \begin{bmatrix} \mathbf{0}_4 & -2\phi\mathbf{\Omega}_F^T & -2\psi\mathbf{\Omega}_F^T & -\mathbf{\Omega}_\phi & -\mathbf{B} \\ 2\phi\mathbf{\Omega}_F & \mathbf{0}_{16} & \mathbf{0}_{16} & \mathbf{0}_{16\times1} & \mathbf{0}_{16\times1} \\ 2\psi\mathbf{\Omega}_F & \mathbf{0}_{16} & \mathbf{0}_{16} & \mathbf{0}_{16\times1} & \mathbf{0}_{16\times1} \\ \mathbf{\Omega}_\phi^T & \mathbf{0}_{1\times16} & \mathbf{0}_{1\times16} & 0 & 0 \\ \mathbf{\Omega}_\psi^T & \mathbf{0}_{1\times16} & \mathbf{0}_{1\times16} & 0 & 0 \end{bmatrix},$$

where

$$\mathbf{\Omega}_F = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{\Omega}_\phi = \begin{bmatrix} 0 \\ 2\left(F^{01} - F^{10}\right) \\ 2\left(F^{02} - F^{20}\right) \\ 2\left(F^{03} - F^{30}\right) \end{bmatrix}, \quad \mathbf{\Omega}_\psi = \begin{bmatrix} 0 \\ 2\left(F^{*01} - F^{*10}\right) \\ 2\left(F^{*02} - F^{*20}\right) \\ 2\left(F^{*03} - F^{*30}\right) \end{bmatrix}.$$

This matrix is singular. We obtain the following constraints from its null eigenvectors:

$$\begin{aligned} \gamma_{F^{jk}} &= 2\phi\left(F^{jk} - \partial_j A_k + \partial_k A_j\right); \\ \gamma_{F^{*jk}} &= 2\psi\left(\partial_k A_j - \partial_j A_k\right) - \psi\epsilon_{jk\mu\nu}F^{*\mu\nu}; \\ \gamma_{F^{0j}} &= 2\phi\left(F_{0j} + F_{j0}\right) \end{aligned}$$

$$\begin{aligned}
\gamma_{F^{*0j}}; &= \psi\left(2F_{j0} - \epsilon_{0j\mu\nu}F^{*\mu\nu}\right) \\
\gamma_{F^{00}} &= 2\phi F_{00}; \\
\gamma_{A_0} &= 2\partial_j\left[\phi\left(F^{j0} - F^{0j}\right) + \psi\left(F^{*j0} - F^{*0j}\right)\right]; \\
\gamma_\phi &= -2\left(F^{jk} - F^{kj}\right)\partial_j A_k + F_{\mu\nu}F^{\mu\nu} + \frac{\partial V}{\partial\phi} + 2\left(F^{0j} - F^{j0}\right)F_{j0}; \\
\gamma_\psi &= -2\left(F^{*jk} - F^{*kj}\right)\partial_j A_k - \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{*\alpha\beta}F^{*\mu\nu} + \frac{\partial V}{\partial\psi} + 2\left(F^{*0j} - F^{*j0}\right)F_{j0}.
\end{aligned}$$

Imposing these constraints $\gamma = 0$ gives us

$$\begin{cases}
F_{jk} & = \partial_j A_k - \partial_k A_j; \\
\epsilon_{jk\mu\nu}F^{*\mu\nu} & = 2\left(\partial_k A_j - \partial_j A_k\right); \\
F_{0j} & = -F_{j0}; \\
F_{j0} & = \frac{1}{2}\epsilon_{0j\mu\nu}F^{*\mu\nu}; \\
F_{00} & = 0; \\
0 & = \partial_j\left[\phi\left(F^{j0} - F^{0j}\right) + \psi\left(F^{*j0} - F^{*0j}\right)\right]; \\
\frac{\partial V}{\partial\phi} & = 2\left(F^{jk} - F^{kj}\right)\partial_j A_k - F_{\mu\nu}F^{\mu\nu} - 2\left(F^{0j} - F^{j0}\right)F_{j0}; \\
\frac{\partial V}{\partial\psi} & = 2\left(F^{*jk} - F^{*kj}\right)\partial_j A_k + \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{*\alpha\beta}F^{*\mu\nu} - 2\left(F^{*0j} - F^{*j0}\right)F_{j0};
\end{cases}$$

The last two equations can be combined with the other relations involving $F^{\mu\nu}$ and $F^{*\mu\nu}$ to obtain

$$\begin{cases}
\frac{\partial V}{\partial\phi} & = F_{\mu\nu}F^{\mu\nu}, \\
\frac{\partial V}{\partial\psi} & = -\frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{*\alpha\beta}F^{\mu\nu} = F_{\mu\nu}F^{*\mu\nu}.
\end{cases}$$

So in total the constraints already gives us most of the kinematical relations we expected for $F^{\mu\nu}$ and $F^{*\mu\nu}$. These constraints also show that F^{jk} and $F^{*\mu\nu}$ have no independent dynamics, and that both are skew symmetric. Furthermore, an integration by parts on the $\partial_j A_0$ term of the Hamiltonian shows that A_0 itself acts as a Lagrange multiplier for the γ_{A_0} constraint; this is similar to the Maxwell electrodynamics case, where A_0 could be seen to be a multiplier for the Gauss' law constraint.

We conclude that we can choose A_j and F^{j0} as the only dynamical variables in this system. The remaining F^{jk} and $F^{*\mu\nu}$ variables are not independent of them, while ϕ and ψ are auxiliary fields with no dynamics, as their time derivatives do not appear explicitly in the Lagrangian density.

Let $\mathbf{\Omega}_1$ be the 6×6 matrix containing only columns and rows of $\mathbf{\Omega}$ corresponding to the A_j and F^{j0} coefficients in the 2-form ω, i.e.

$$\mathbf{\Omega}_1 = 2\phi \begin{bmatrix} \mathbf{0}_3 & -\delta_j^k \\ \delta_j^k & \mathbf{0}_3 \end{bmatrix}.$$

This matrix is nonsingular and has an inverse:

$$(\mathbf{\Omega}_1)^{-1} = \frac{1}{2\phi} \begin{bmatrix} \mathbf{0}_3 & \delta_j^k \\ -\delta_j^k & \mathbf{0}_3 \end{bmatrix}.$$

We can use this to calculate the field equations in the following manner. Let $\xi^a = \{A^j, F^{j0}\}$ the set of phase space coordinates encompassed by the invertible matrix above, and $\xi^A = \left\{A^0, F^{j0}, F^{jk}, F^{*\mu\nu}, \phi, \psi\right\}$ (the remaining coordinates). From the field equations $\frac{\delta H}{\delta \xi^\mu} = \omega_{\mu\nu}\dot{\xi}^\nu$, we have

$$\frac{\delta H}{\delta \xi^\mu} = \omega_{\mu a}\dot{\xi}^a + \omega_{\mu A}\dot{\xi}^A,$$

which, since $\omega_{\mu a}$ is invertible, can be rearranged into

$$\dot{\xi}^a = \left(\omega^{-1}\right)^{\mu a} \frac{\delta H}{\delta \xi^\mu} - \left(\omega^{-1}\right)^{\mu a} \omega_{\mu A}\dot{\xi}^A. \tag{4.7}$$

The components of $(\omega^{-1})^{\mu a}$ are the entries in the $(\mathbf{\Omega}_1)^{-1}$ matrix, while the components of $\omega_{\mu A}$ are the remaining elements of the $\mathbf{\Omega}$ matrix corresponding to $d\xi^A$ elements in the symplectic 2-form. The relevant functional derivatives are

$$\begin{aligned} \frac{\delta H}{\delta A^j} &= 2\partial_k \left[\phi\left(F^{kj} - F^{jk}\right) + \psi\left(F^{*kj} - F^{*jk}\right)\right]; \\ \frac{\delta H}{\delta F^{j0}} &= -2\phi\partial^j A^0 + 2\phi F^{j0}. \end{aligned}$$

so

$$\left(\omega^{-1}\right)^{\mu a} \frac{\delta H}{\delta \xi^\mu} = \begin{cases} \partial^j A^0 - F^{j0} & \text{if } a = A_j, \\ \frac{1}{\phi}\partial_k \left[\phi\left(F^{kj} - F^{jk}\right) + \psi\left(F^{*kj} - F^{*jk}\right)\right] & \text{if } a = F^{j0}. \end{cases}$$

For the remaining term $(\omega^{-1})^{\mu a}\,\omega_{\mu A}\dot{\xi}^A$, we can write the matrix representation of $\omega_{\mu A}$ as

$$\mathbf{\Omega}_A = \begin{bmatrix} \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 3} \\ -2\phi\delta_k^j & \mathbf{0}_3 \\ \mathbf{0}_{9\times 3} & \mathbf{0}_{9\times 3} \\ \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 3} \\ 2\psi\delta_k^j & \mathbf{0}_{3\times 9} \\ -2\psi\delta_k^j & \mathbf{0}_{1\times 9} \\ \mathbf{0}_{9\times 3} & \mathbf{0}_{3\times 9} \\ \mathbf{0}_{1\times 3} & \mathbf{0}_{3\times 9} \\ -2\left(F^{0j}-F^{j0}\right) & \mathbf{0}_9 \\ -2\left(F^{*0j}-F^{*j0}\right) & \mathbf{0}_{1\times 9} \end{bmatrix}^T,$$

from where we obtain

$$\left(\omega^{-1}\right)^{\mu a}\omega_{\mu A}\dot{\xi}^A = \begin{cases} 0 & \text{if } a = A^j, \\ -\dot{F}^{0j} + \frac{\psi}{\phi}\left(\dot{F}^{*0j} - \dot{F}^{*j0}\right) & \\ -\frac{1}{\phi}\left(F^{0j}-F^{j0}\right)\dot{\phi} - \frac{1}{\phi}\left(F^{*0j}-F^{*j0}\right)\dot{\psi} & \text{if } a = F^{j0}. \end{cases}$$

So, putting it all together, (4.7) gives us

$$\begin{cases} \dot{A}^j & = \partial^j A^0 - F^{j0}, \\ \phi\dot{F}^{j0} & = \partial_k\left[\phi\left(F^{kj}-F^{jk}\right) + \psi\left(F^{*kj}-F^{*jk}\right)\right] \\ & \quad -\phi\dot{F}^{0j} + \psi\left(\dot{F}^{*0j}-\dot{F}^{*j0}\right) - \left(F^{0j}-F^{j0}\right)\dot{\phi} - \left(F^{*0j}-F^{*j0}\right)\dot{\psi}. \end{cases}$$

The second equation can be rearranged into

$$\begin{aligned} & 0 = \partial_0\left[\phi\left(F^{j0}-F^{0j}\right) + \psi\left(F^{*j0}-F^{*0j}\right)\right] + \partial_k\left[\phi\left(F^{jk}-F^{kj}\right) + \psi\left(F^{*jk}-F^{*kj}\right)\right] \\ \implies & 0 = \partial_\mu\left[\phi\left(F^{j\mu}-F^{\mu j}\right) + \psi\left(F^{*j\mu}-F^{*\mu j}\right)\right], \end{aligned}$$

which, combined with the Gauss' law-like constraint $\gamma_{A_0} = 0$, gives

$$0 = \partial_\mu\left[\phi\left(F^{\nu\mu}-F^{\mu\nu}\right) + \psi\left(F^{*\nu\mu}-F^{*\mu\nu}\right)\right],$$

which can be further rearranged by noting that the constraints $\gamma_{F^{jk}} = 0, \gamma_{F^{*jk}} = 0, \gamma_{F^{0j}} = 0$ and $\gamma_{F^{*0j}} = 0$ impose skew-symmetry on F and F^*, resulting in

$$0 = \partial_\mu\left(\phi F^{\nu\mu} + \psi F^{*\nu\mu}\right).$$

This equation and the constraint equations $\gamma_\phi = 0$ and $\gamma_\psi = 0$, are equivalent to the set of field equations obtained before for a Jordan frame NLED in (4.5).

4.3 Example: the generalized Born-Infeld action

Let us look back at the action defined in (4.3). Replacing $\mathcal{S}$ and $\mathcal{P}$ by auxiliary fields ξ and χ, the Lagrangian is written

$$\mathcal{L}(\xi,\chi) \coloneqq f(\xi,\chi) = \tau^2\left[1-\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{\sigma}\right]. \tag{4.8}$$

The derivatives are

$$\begin{cases} \frac{\partial f}{\partial \xi} &= -\frac{1}{4}\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{\sigma-1}, \\ \frac{\partial f}{\partial \chi} &= \frac{\gamma\chi}{16\sigma\tau^4}\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{\sigma-1}. \end{cases} \tag{4.9}$$

So the elements of the Hessian matrix are

$$\begin{cases} \frac{\partial^2 f}{\partial \xi^2} &= -\frac{(\sigma-1)}{16\tau^2\sigma}\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{\sigma-2}, \\ \frac{\partial^2 f}{\partial \xi\partial\chi} &= \frac{\partial^2 f}{\partial \chi\partial\xi} = -\frac{\gamma\chi(\sigma-1)}{64\tau^2\sigma}\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{\sigma-2}, \\ \frac{\partial^2 f}{\partial \chi^2} &= \frac{\gamma}{16\tau^2\sigma}\left[1+\frac{\xi}{4\tau^2\sigma}+\frac{\gamma\chi^2(1-2\sigma)}{32\tau^2\sigma}\right]\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{\sigma-2}. \end{cases}$$

And the determinant is

$$\begin{aligned} \det H &= \frac{1}{256\tau^4\sigma^2}\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{2\sigma-4}\det\begin{bmatrix}(1-\sigma) & \frac{1\xi(\sigma-1)}{4\tau^2} \\ \frac{\gamma\xi(\sigma-1)}{4\tau^2} & \gamma\sigma\left[1+\frac{\xi}{4\tau^2\sigma}+\frac{\gamma\chi^2(1-2\sigma)}{32\tau^2\sigma}\right]\end{bmatrix} \\ &= \frac{\gamma(1-\sigma)}{256\tau^4\sigma^2}\left(1+\frac{\xi}{4\tau^2}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{2\sigma-3}. \end{aligned}$$

we can conclude that this theory is regular for $\sigma \neq 1$ and $\gamma \neq 0$. The auxiliary fields ϕ and ψ are defined by the derivatives (4.9)

$$\begin{cases} \phi = \frac{\partial f}{\partial \xi} &= -\frac{1}{4}\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{\sigma-1}, \\ \psi = \frac{\partial f}{\partial \chi} &= \frac{\gamma\chi}{16\sigma\tau^4}\left(1+\frac{\xi}{4\tau^2\sigma}-\frac{\gamma\chi^2}{32\sigma\tau^4}\right)^{\sigma-1} = -\frac{\chi\gamma\phi}{4\tau^2}, \end{cases}$$

which can be inverted

$$\begin{cases} \chi &= -\frac{4\tau^2\psi}{\phi\gamma}, \\ \xi &= 4\tau^2\sigma\left[(-4\phi)^{\frac{1}{\sigma-1}}-1+\frac{\psi^2}{2\phi^2\gamma\sigma}\right]. \end{cases}$$

Substituting into (4.8), we obtain

$$f\left(\xi\left(\phi,\psi\right),\chi\left(\phi,\psi\right)\right)=\tau^{2}\left[1-\left(-4\phi\right)^{\frac{\sigma}{\sigma-1}}\right],$$

and the Jordan frame action (4.4) is

$$S=\int\left[\phi\mathcal{S}+\psi\mathcal{P}-V\left(\phi,\psi\right)\right]d^{4}x,$$

where

$$\begin{aligned}V\left(\phi,\psi\right)&=\chi\left(\phi,\psi\right)\phi+\xi\left(\phi,\psi\right)\psi-f\left(\chi\left(\phi,\psi\right),\xi\left(\phi,\psi\right)\right)\\&=-4\tau^{2}\sigma\left[\phi\left(1-\frac{\left(\sigma-1\right)}{\sigma}\left(-4\phi\right)^{\frac{1}{\sigma-1}}\right)+\frac{1}{4\sigma}+\frac{\psi^{2}}{2\gamma\sigma\phi}\right],\end{aligned}$$

with derivatives

$$\begin{cases}\frac{\partial V}{\partial\phi}&=-4\tau^{2}\sigma\left(1-\left(-4\phi\right)^{\frac{1}{\sigma-1}}\right)-\frac{\psi^{2}}{2\gamma\sigma\phi^{2}},\\\frac{\partial V}{\partial\psi}&=-\frac{4\tau^{2}\psi}{\gamma\phi}.\end{cases}$$

So, the field equations for this theory are

$$\begin{cases}\mathcal{S}&=\frac{\partial V}{\partial\phi}=-4\tau^{2}\sigma\left(1-\left(-4\phi\right)^{\frac{1}{\sigma-1}}\right)-\frac{\psi^{2}}{2\gamma\sigma\phi^{2}},\\\mathcal{P}&=\frac{\partial V}{\partial\psi}=-\frac{4\tau^{2}\psi}{\gamma\phi},\\0&=\partial_{\mu}\left(\phi F^{\nu\mu}+\psi F^{*\nu\mu}\right).\end{cases}.$$

With this, we conclude our brief investigation into how NLED theories may be analyzed as singular systems with the help of the Jordan-Palatini frame.

5 Conclusion

In this work, we studied many aspects of singular systems, field theories and their connections with physical constraints and gauges. The Lagrangian and Hamiltonian formalisms allow us to inspect theories in different mathematical schemes, each with their own traits. The Lagrangian method usually provides a simpler, more direct calculation, however, the Hamiltonian is of foremost importance for modern physics due to being the basis for the quantization of mechanical systems. In this sense, both the Dirac-Bergmann and Faddeev-Jackiw methods provide a natural path for quantization by constructing generalized Poisson brackets, though the Dirac-Bergmann method relies more on manipulating the system with its Poisson brackets as usually defined, while Faddeev-Jackiw depends on using differential forms defined by the system to uncover its symplectic structure, from which we may obtain the generalized bracket relations.

Our applications of each method to the free electromagnetic field and nonlinear electrodynamics are examples of how these procedures may be used in practice. Furthermore, the prevalence of electromagnetism as a field theory by itself already demonstrates the usefulness of constraint analysis and the study of singular systems. And while electrodynamics is not the only singular field of profound interest to physicists, its status as a fundamental element of classical field theory makes it a important didactic example to understand concepts while also directly being the basis for other theories, including newer proposals like the generalized Born-Infeld electrodynamics.

The Faddeev-Jackiw approach used in this work for the Jordan-Palatini formulation of nonlinear electrodynamics involved reducing the phase space after the first iteration and obtaining the field equations solely for our desired fields (which we knew were the only fields with dynamics in the system due to the constraints). This greatly simplified the resulting calculations. We leave the full Faddeev-Jackiw iteration with Lagrange multipliers, as well as further analyses of theoretical aspects of NLED models, as possible future extensions to this work. Another possible branching of this work is understanding how the Faddeev-Jackiw method interfaces with multisymplectic and polysymplectic formulations of classical field theory (e.g. (GIACHETTA *et al.*, 2009; FORGER; GOMES, 2013)).

It should also be noted that the methods described in this work are important tools for the study of singular systems, however, they are by no means the only possible ways to analyze such systems. For example, the Hamilton-Jacobi formalism from analytical mechanics has also been adapted and applied to singular systems (BERTIN *et al.*, 2005).

Bibliography

ABRAHAM, R.; MARSDEN, J. E. **Foundations of mechanics**. [*S.l.*]: American Mathematical Soc., 2008.

BARCELOS-NETO, J.; WOTZASEK, C. Faddeev-Jackiw quantization and constraints. **International Journal of Modern Physics A**, World Scientific, v. 7, n. 20, p. 4981–5003, 1992.

BARCELOS-NETO, J.; WOTZASEK, C. Symplectic quantization of constrained systems. **Modern Physics Letters A**, World Scientific, v. 7, n. 19, p. 1737–1747, 1992.

BERTIN, M.; PIMENTEL, B.; POMPEIA, P. First-order actions: a new view. **Modern Physics Letters A**, World Scientific, v. 20, n. 37, p. 2873–2889, 2005.

BIALYNICKI-BIRULA, I. Nonlinear electrodynamics: Variations on a theme by Born and Infeld. **Quantum theory of particles and fields**, World Scientific Singapore, p. 31–48, 1983.

BORN, M.; INFELD, L. Foundations of the new field theory. **Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character**, The Royal Society London, v. 144, n. 852, p. 425–451, 1934.

CARMO, M. P. D. **Differential forms and applications**. [*S.l.*]: Springer, 2012.

CARO, L.; PIMENTEL, B.; ZAMBRANO, G. Método de Faddeev-Jackiw na mecânica clássica. **Revista Brasileira de Ensino de Física**, SciELO Brasil, v. 43, p. e20210273, 2021.

CARO, L. G. Da teoria clássica à quântica de campos através da abordagem de Faddeev-Jackiw. Universidade Estadual Paulista, 2024.

CHERN, S.-s.; WOLFSON, J. G. A simple proof of Frobenius theorem. *In*: **Manifolds and Lie Groups: Papers in Honor of Yozô Matsushima**. [*S.l.*]: Springer, 1981. p. 67–69.

DIRAC, P. A. M. Generalized hamiltonian dynamics. **Canadian journal of mathematics**, Cambridge University Press, v. 2, p. 129–148, 1950.

DIRAC, P. A. M. **Lectures on Quantum Mechanics**. [*S.l.*]: Yeshiva University, 1964.

FADDEEV, L.; JACKIW, R. Hamiltonian reduction of unconstrained and constrained systems. **Physical Review Letters**, APS, v. 60, n. 17, p. 1692, 1988.

FLANDERS, H. **Differential forms with applications to the physical sciences**. [*S.l.*]: Courier Corporation, 1963.

FORGER, M.; GOMES, L. G. Multisymplectic and polysymplectic structures on fiber bundles. **Reviews in Mathematical Physics**, World Scientific, v. 25, n. 09, p. 1350018, 2013.

GIACHETTA, G.; SARDANASHVILY, G. A.; MANGIAROTTI, L. **Advanced classical field theory**. [*S.l.*]: World Scientific, 2009.

GOLDSTEIN, H.; POOLE, C.; SAFKO, J.; ADDISON, S. R. **Classical Mechanics, Third Edition**. [*S.l.*]: Pearson, 2002.

KRUGLOV, S. Notes on Born–Infeld-type electrodynamics. **Modern Physics Letters A**, World Scientific, v. 32, n. 36, p. 1750201, 2017.

LEMOS, N. A. **Analytical mechanics**. [*S.l.*]: Cambridge University Press, 2018.

SOROKIN, D. P. Introductory notes on non-linear electrodynamics and its applications. **Fortschritte der Physik**, Wiley Online Library, v. 70, n. 7-8, p. 2200092, 2022.

SOTIRIOU, T. P.; FARAONI, V. f (r) theories of gravity. **Reviews of Modern Physics**, APS, v. 82, n. 1, p. 451–497, 2010.

SUDARSHAN, E. C. G.; MUKUNDA, N. **Classical dynamics: a modern perspective**. [*S.l.*]: World Scientific, 1974.

SUNDERMEYER, K. **Constrained dynamics with applications to Yang-Mills theory, general relativity, classical spin, dual string model**. [*S.l.*]: Springer, 1982.

TSAMPARLIS, M. On the Palatini method of variation. **Journal of Mathematical Physics**, American Institute of Physics, v. 19, n. 3, p. 555–557, 1978.

TU, L. W. **Introduction to Manifolds, Second Edition**. [*S.l.*]: Springer, 2011.

YANO, K. **The theory of Lie derivatives and its applications**. [*S.l.*]: Courier Dover Publications, 2020.

Appendix A - Differential forms

In this appendix, we will give a brief summary of the mathematical theory of differential forms. which was used in this work to define the generalized Poisson bracket, as well as the tautological 1-form and symplectic forms used in the Faddeev-Jackiw method. For more complete discussions on this topic, see (CARMO, 2012) and (FLANDERS, 1963).

A.1 Differential forms on a manifold

We start by defining a general k-form in a manifold as an element of the space of alternating multilinear k-forms on cotangent space, denoted $\bigwedge^k \left(T_p^* M\right)$.

Definition A.1.1. A *k-form* on a differentiable manifold M of dimension n is a set of elements $\omega(p) \in \bigwedge^k \left(T_p^* M\right)$ for each point $p \in M$.

Definition A.1.2. The exterior product, $\wedge$, of a k-form and a l-form, is a bilinear map

$$\wedge : \bigwedge^k \left(T_p^* M\right) \times \bigwedge^l \left(T_p^* M\right) \to \bigwedge^{k+l} \left(T_p^* M\right)$$

that is distributive, associative and satisfies the following identity, for a k-form α and l-form β

$$\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$$

Locally, a coordinate chart $x_1, \ldots, x_n$ allows us to identify such a k-form with a real k-form on $\bigwedge^k \left(T_p^* \mathbb{R}^n\right)$. The elements of $T_p^* \mathbb{R}^n$ themselves are 1-forms, with basis 1-forms written as $dx_1, \ldots, dx_n$. k-forms can be written as linear combinations of exterior products of k basis 1-forms. For example, a k-form in general is

$$\omega_p = a_{j_1, \ldots, j_k} (p) \, dx^{j_1} \wedge \cdots \wedge dx^{j_k}.$$

A k-form is called a *differential k-form* if the coefficients of its local representation (such as the $a_{j_1, \ldots, j_k} (p)$ above) are differentiable functions at each point p. The set of k-forms on a manifold M is denoted $\Omega^k (M)$. A differential 0-form is merely a differentiable function on M.

Since a k-form is a multilinear map of cotangent vectors, which are linear functionals of tangent vectors, we can also interpret a k-form as taking k tangent vector fields (defined for each point p) as parameters. For a k-form ω acting on vector fields $X_1, \ldots, X_k$, we can denote this as $\omega\left(X_1, \ldots, X_k\right)$. This allows us to easily define an operation that decreases the rank of a k-form:

Definition A.1.3. The *interior product* $\iota_{\mathbf{V}}\omega$ of a k-form $\omega\left(X_1, \ldots, X_k\right)$ by a vector field $\mathbf{V}$ is the $(k-1)$-form given by

$$\left(\iota_{\mathbf{V}}\omega\right)\left(X_1, \ldots, X_{k-1}\right) = \omega\left(V, X_1, \ldots, X_{k-1}\right)$$

This interior product is clearly nilpotent when applied with the same vector field and k-form multiple times, by the alternating nature of k-forms (which implies that it is null if applied on any repeated vector fields arguments):

$$\iota_{\mathbf{V}}\left(\iota_{\mathbf{V}}\omega\right) = \omega\left(V, V, X_1, \ldots, X_{k-2}\right) = 0.$$

The last operation we will define here, acting specifically on differential forms, is a derivative.

Definition A.1.4. Let ω be a k-form. The *exterior derivative* is a linear map

$$\begin{aligned} d : \Omega^k(M) &\longrightarrow \Omega^{k+1}(M) \\ \omega\left(X_1, \ldots, X_k\right) &\longmapsto d\omega\left(X_1, \ldots, X_{k+1}\right) \end{aligned}$$

satisfying, for any k-form α and l-form β

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta$$

and

$$d(d\alpha) = 0$$

.

Locally, we write the exterior derivative of a k-form ω as

$$d\omega = \frac{\partial a_{j_1,\ldots,j_k}(p)}{\partial x_l} dx^l \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k}$$

such that, in the case of a 2-form, considering the skew-symmetry of the exterior product, the coefficients are

$$(d\omega)_{ij} = \partial_i \omega_j - \partial_j \omega_i.$$

A.1.1 Poincaré's lemma

We now classify differential k-forms in two categories:

Definition A.1.5. Let ω be a differential k-form. We say that

- ω is a *closed form* if $d\omega = 0$.
- ω is an *exact form* if there exists a $(k-1)$-form θ such that $\omega = d\theta$.

By the properties of the exterior derivative, every exact form in any manifold is closed. The converse is not necessarily true for a differential form defined in any manifold. Poincaré's lemma gives us a condition to guarantee that it holds. Before we get to this result, we define a few further operations on differential forms that will be useful for the proof.

Definition A.1.6. Let M, N be two differentiable manifolds and f a diffeomorphism (i.e. a differentiable bijective map with differentiable inverse) between M and N. The *push-forward* of f at a point $p \in M$ is a map

$$\begin{aligned} T_p f: \ T_pM &\longrightarrow T_{f(p)}N \\ \gamma_p(g) &\longmapsto \gamma_p(g \circ f) \end{aligned}$$

(where γ_p is a tangent vector at p, and g is a smooth function in M).

The *pullback* f^* is the function on k-forms ω

$$f^*\omega_p\left(v_1, \ldots, v_k\right) \coloneqq \omega_{f(p)}\left(T_p f\left(v_1\right), \ldots, T_p f\left(v_k\right)\right)$$

We recall that any smooth vector field $\mathbf{V}$ on a manifold M defines a function

$$\Phi_{\mathbf{V}}(p,t) : M \times \mathbb{R} \to M$$

called the *flow*, which satisfies, for all $p \in M$

- $\Phi_{\mathbf{V}}(p, 0) = p$
- $\frac{d}{dt}\Phi_{\mathbf{V}}(p,t) = \mathbf{V}(\Phi_V(p,t))$.

The existence and uniqueness theorem (Picard-Lindelöf theorem) guarantees that such a function $\Phi_{\mathbf{V}}$ exists and is unique provided that $\mathbf{V}$ is smooth. From now on, we will denote $\Phi^t_{\mathbf{V}}(p) \coloneqq \Phi_{\mathbf{V}}(p,t)$ for simplicity.

We then define the *Lie derivative* of a differential form.

Definition A.1.7. Let $\mathbf{V}$ be a smooth vector field, and f a smooth function. The *Lie derivative* of f with respect to $\mathbf{V}$ is

$$\mathcal{L}_{\mathbf{V}} f(p) := \lim_{t \to 0} \frac{f\left(\Phi^t_{\mathbf{V}}(p)\right) - f(p)}{t}.$$

For a differential k-form $\omega \in \Omega^k(M)$ we define

$$\mathcal{L}_{\mathbf{V}}\omega := \lim_{t \to 0} \frac{\left(\Phi^t_{\mathbf{V}}\right)^* \omega - \omega}{t} = \left.\frac{d}{dt}\right|_{t=0} \left(\Phi^t_{\mathbf{V}}\right)^* \omega$$

or, in general for any t,

$$\frac{d}{dt}\left(\Phi^t_{\mathbf{V}}\right)^* \omega = \left(\Phi^t_{\mathbf{V}}\right)^* \mathcal{L}_{\mathbf{V}}\omega$$

The Lie derivative satisfies many properties associated with the usual derivation operation, such as linearity and the Leibniz rule. We are only introducing it as a tool to prove Poincaré's lemma, so we will not go into much detail regarding these properties, There are many in-depth texts in the literature expanding upon the theory of Lie derivatives, including applications to physics (YANO, 2020; ABRAHAM; MARSDEN, 2008).

For the proof of the following theorem, the only other important property of the Lie derivative of a k-form is that it commutes with the exterior derivative.

$$\mathcal{L}_{\mathbf{V}} d\omega = d\mathcal{L}_{\mathbf{V}}\omega.$$

This can be seen from the

Theorem A.1.1. *The Lie derivative of a k-form ω can be written in terms of the interior product and exterior derivative as*

$$\mathcal{L}_{\mathbf{V}}\omega = d\left(\iota_{\mathbf{V}}\omega\right) + \iota_{\mathbf{V}} d\omega.$$

This is known as the Cartan homotopy formula *(TU, 2011).*

Proof. We prove this by induction on the rank of the k-forms.

For a 0-form, that is, a smooth function f, the interior product $\iota_{\mathbf{V}} f$ will always be zero. The Lie derivative would be

$$\left(\mathcal{L}_{\mathbf{V}} f\right)(p) = \left.\frac{d}{dt}\right|_{t=0} f\left(\Phi^t_{\mathbf{V}}(p)\right) = (df)_p \left(\left.\frac{d}{dt}\right|_{t=0} \Phi^t_{\mathbf{V}}(p)\right) = (df)_p(\mathbf{V}) = \iota_{\mathbf{V}} df$$

Now, for the induction step, we assume the formula is valid for any $(k-1)$-form θ. Locally, we may then write any k-form as $\omega = \sum_j df_j \wedge \theta_j$ where f is a smooth function.

The Lie derivative is linear such that

$$\mathcal{L}_{\mathbf{V}}\omega = \sum_j \left(\mathcal{L}_{\mathbf{V}} df_j\right) \wedge \theta_j + df_j \wedge \left(\mathcal{L}_{\mathbf{V}}\theta_j\right).$$

On the other hand,

$$\begin{aligned}
\iota_{\mathbf{V}} d\left(df \wedge \theta\right) + d\iota_{\mathbf{V}}\left(df \wedge \omega\right) &= \cancel{\iota_{\mathbf{V}}\left(d(df) \wedge \theta\right)} - \iota_{\mathbf{V}}\left(df \wedge d\theta\right) + d\left(\left(\iota_{\mathbf{V}} df\right) \wedge \theta - df \wedge \left(\iota_{\mathbf{v}}\theta\right)\right) \\
&= \cancel{-\iota_{\mathbf{V}} df \wedge d\theta} + df \wedge \left(\iota_{\mathbf{V}} d\theta\right) + \left(d\iota_{\mathbf{V}} df\right) \wedge \theta + \cancel{\left(\iota_{\mathbf{V}} df\right) \wedge \left(d\theta\right)} \\
&\quad + \cancel{\left(d\left(df\right)\right) \wedge \left(\iota_{\mathbf{V}}\theta\right)} + df \wedge \left(d\iota_{\mathbf{V}}\theta\right) \\
&= df \wedge \left(\mathcal{L}_{\mathbf{V}}\theta\right) + \left(d\mathcal{L}_{\mathbf{V}} f\right) \wedge \omega \\
&= df \wedge \left(\mathcal{L}_{\mathbf{V}}\theta\right) + \left(\mathcal{L}_{\mathbf{V}} df\right) \wedge \theta
\end{aligned}$$

so that, summing this expression over all the df_j and θ_j, we have exactly the formula as stated. □

We can now state and prove the last and most important theorem regarding differential forms.

Theorem A.1.2 (Poincaré's lemma)**.** *Let ω be a differential k-form on an open ball contained in $\mathbb{R}^n$. Then, if ω is closed, it is also exact.*

Proof. Let U be an open ball centered at the origin of $\mathbb{R}^n$. For every point $p \in U$, we can define a vector field $\mathbf{V}$ with flow $\Phi^t_{\mathbf{V}}(p) = (1-t)p$ and note that, for $t \in [0,1]$, this flow is entirely contained in U.[1] Then, by the definition of the Lie derivative,

$$\frac{d}{dt}\left(\Phi^t_{\mathbf{V}}\right)^* \omega = \mathcal{L}_{\mathbf{V}}\omega$$

Integrating from 0 to 1 and applying Cartan's homotopy formula on the right side,

$$\begin{aligned}
\left(\Phi^0_{\mathbf{V}}\right)^* \omega - \left(\Phi^1_{\mathbf{V}}\right)^* \omega &= \int_0^1 \left(d\left(\iota_{\mathbf{V}}\omega\right) + \iota_{\mathbf{V}} d\omega\right) dt \\
\left(\Phi^0_{\mathbf{V}}\right)^* \omega - \cancel{\left(\Phi^1_{\mathbf{V}}\right)^* \omega} &= \int_0^1 \left(d\left(\iota_{\mathbf{V}}\omega\right) + \cancel{\iota_{\mathbf{V}} d\omega}\right) dt \\
\omega &= d\int_0^1 \iota_{\mathbf{V}}\omega dt.
\end{aligned}$$

[1]This is, in fact, the *only* real property of the open ball that we need for this proof. Since we are generally working with manifolds, proving this theorem for an open ball im $\mathbb{R}^n$ is a powerful enough result to allow us to apply this theorem to an open neighborhood of any point in a manifold. We could, however, be more abstract and generalize Poincaré's lemma to every set that satisfies the property that this flow, which is an homotopy with the origin, be contained in the set. A set that satisfies this property is called *contractible*.

The integral is well defined, as $\iota_{\mathbf{V}}\omega$ is a $(k-1)$-form in $\mathbb{R}^n$, and the integration over t only affects its coefficients. Since these coefficients are differentiable (by the definition of a differential form), they are also integrable, and $\int_0^1 \iota_{\mathbf{V}}\omega dt$ is a well-defined $(k-1)$- form; then, ω is its exterior derivative, which means ω is exact, proving our hypothesis. □

Appendix B - Symplectic geometry and Darboux's theorem

In this appendix, we take note of some definitions and theorems from symplectic geometry, which set a foundation for the mathematical theory of Hamiltonian systems. In particular our final objective will be to give a solid footing needed to state and prove the generalized Darboux's theorem, which was the basis for the Faddeev-Jackiw method.

Several concepts from appendix B.1.1 are also to be used throughout this appendix.

B.1 Symplectic vector spaces and manifolds

Let V be a vector space over $\mathbb{R}$ (the field of real numbers) with dimension m, and $f : V \times V \to \mathbb{R}$ a bilinear skew-symmetric map.

Theorem B.1.1. *There is a basis of* V $\{u_1, \ldots, u_r, v_1, \ldots, v_n, w_1, \ldots, w_n\}$ *satisfying, for all indices* $j \in 1, \ldots, r; k, l \in 1, \ldots, n$

$$\begin{cases} f\left(u_j, v_k\right) = 0 \\ f\left(v_k, v_l\right) = f\left(w_k, w_l\right) = 0 \\ f\left(v_k, w_l\right) = \delta_{kl} \end{cases}$$

Proof. First we take $u_1, \ldots, u_r$ as a basis of the subspace U defined by

$$U = \{y \in V : f\left(x, y\right) = 0, \forall x \in V\}$$

then, take v_1 as any (non-zero) vector in the complement $W = V - U$. There also exists a $w_1 \in V - U$ such that $f\left(v_1, w_1\right) \neq 0$, otherwise we would have $v_1 \in U$. By the bilinearity of f and the properties of V as a vector space we can take w_1 so that $f\left(v_1, w_1\right) = 1$.

Now, consider the subspace $W_1 \subseteq W$ spanned by v_1, w_1. We repeat the same split as before, considering also the subspace

$$W_1^f = \{y \in W : f\left(x, y\right) = 0, \forall x \in W_1\}.$$

Let $z \in W_1$, i.e. $z = c_1 v_1 + c_2 w_1$. Suppose also that $z \in W_1^f$. Thus

$$\begin{cases} f(z, v_1) = 0 \\ f(z, w_1) = 0 \end{cases}$$

however, exploiting the bilinearity of f, we have

$$\begin{cases} f(z, v_1) = f(c_1 v_1 + c_2 w_1, v_1) = -c_2 \\ f(z, w_1) = f(c_1 v_1 + c_2 w_1, w_1) = c_1 \end{cases}$$

and thus $c_1 = c_2 = 0 \implies W_1 \cap W_1^f = \{0\}$.

Now, let $z \in W$. Writing $f(z, v_1) = c_1$ and $f(z, w_1) = c_2$, then we may write

$$z = (c_2 w_1 - c_1 v_1) + (z - c_2 w_1 + c_1 v_1).$$

The term in the first set of parentheses is in W_1, being a linear combination of v_1 and w_1. The term on the second set of parentheses is in W_1^f, as applying f to it and either v_1 or w_1 gives us zero. Then, we conclude that any vector $z \in W$ can be written as a sum of a vector in W_1 and a vector in W_1^f.

We continue the process by taking a non-zero vector $v_2 \in W_1^f$ and once again choosing w_2 such that $f(v_2, w_2) = 1$. Then, we set W_2 as the subspace spanned by v_2, w_2.

We repeat this, generating new subspaces spanned by vectors $v_3, w_3, \ldots$ until eventually, since V is finite dimensional, we finally have a complete basis given by

$$\{u_1, \ldots, u_r, v_1, \ldots, v_n, w_1, \ldots, w_n\}$$

where the vectors satisfy the three properties we stated in the theorem. □

Definition B.1.1. If the basis encountered by the process in the theorem above is composed solely of $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$, that is, if there are no non-zero u_r vectors such that $f(u_r, x) = 0$ for all $x \in V$, then f is called a *symplectic map* and (V, f) is a *symplectic vector space.*

Remark. Since a symplectic vector space has a basis given by pairs of vectors v_n, w_n, its dimension has to be an even number.

Now, we extend this notion of a symplectic structure to more general manifolds. Differential forms give us a natural way to achieve this. We note that any 2-form ω in a manifold M gives rise to a bilinear skew-symmetric map ω_x from $T_xM \times T_xM$ to $\mathbb{R}$ at each point $x \in M$. This can be seen, for example, if we write a matrix representation of ω as $\mathbf{\Omega}$, then

$$\omega_x(\mathbf{v}, \mathbf{w}) = \mathbf{v}^T \mathbf{\Omega} \mathbf{w}$$

where $\mathbf{v}$ and $\mathbf{w}$ are two covectors in T_x^*M. Thus, we can define a symplectic form as simply:

Definition B.1.2. Let ω be a 2-form defined on a manifold M. If ω is closed (that is, $d\omega = 0$), and ω_x is a symplectic map for every point $x \in M$, then we call ω a *symplectic form.* In this case, we say (M, ω) is a *symplectic manifold*,

Remark. Let ω be a closed 2-form defined on a manifold M represented by a matrix $\mathbf{\Omega}$. Then ω is a symplectic form if and only if $\det \mathbf{\Omega} \neq 0$. This can be seen quite directly from the definitions.

Now, for the mathematical theory of analytical mechanics, we work not only in manifolds, but on structures called tangent and cotangent bundles. Bundles also have symplectic structures. Firstly, formally, we define a bundle:

Definition B.1.3. Let B and T be two real topological spaces and $\pi : T \to B$ a surjective and continuous projection. If there exists some natural number n such that for every point $p \in B$, the set $\pi^{-1}[\{p\}]$ is homeomorphic to $\mathbb{R}^n$, we say that the B, T and π form a *fiber bundle.*

Furthermore if the set $\pi^{-1}[\{p\}]$ is a vector space, then this structure is called a *vector bundle.*

Definition B.1.4. Let M be a n-dimensional smooth manifold. Define the set

$$TM = \bigcup_{p \in M} T_pM$$

where T_pM is the tangent space at p. The projection

$$\begin{aligned} \pi : \quad TM &\longrightarrow M \\ (p, v) &\longmapsto p \end{aligned}$$

gives a fiber structure to (TM, M, π). This $2n$-dimensional fiber bundle is called *tangent bundle.*

In similar fashion, we can define the set

$$T^*M = \bigcup_{p \in M} T_p^*M$$

where T_p^*M is the cotangent space at p (dual of the tangent space), and a projection

$$\begin{aligned} \rho: \quad T^*M &\longrightarrow M \\ (p,\theta) &\longmapsto p \end{aligned}$$

forms a 2n-dimensional vector bundle (T^*M, M, ρ) called *cotangent bundle*.

For a given n-dimensional manifold M, the cotangent bundle T^*M naturally forms a $2n$-dimensional manifold. The manifold structure of T^*M can be constructed as follows: given a coordinate chart (U, x_j) for M, a basis for the *tangent space* T_xM at a point $x \in U$ is given by $\frac{\partial}{\partial x_j}$. A basis for the cotangent space T_x^*M at this same point is given by differentials dx_j, so that any covector $\xi \in T_x^*M$ may be written $\xi = \xi_j dx_j$. We can then define a coordinate chart on the cotangent bundle T^*M as (T^*U, x_j, ξ_j). The cotangent bundle is then a $2n$-dimensional manifold.

Definition B.1.5 (Tautological one-form)**.** Let M be a manifold, with (U, x_j) a coordinate chart for M and (T^*U, x_j, ξ_j) a coordinate chart for the cotangent bundle T^*M. The 1-form defined by

$$\theta = \xi_j dx_j$$

is called the *tautological 1-form.* The exterior derivative $\omega = -d\theta = dx_j \wedge d\xi_j$ is the *canonical symplectic form.*

While it may not be obvious from the above definition, the tautological one form is coordinate independent. If (U, x_j) and (V, y_j) are two coordinate charts for M such that $U \cap V \neq \emptyset$. Then, taking the cotangent space charts (T^*U, x_j, ξ_j) and (T^*V, y_j, o_j), we can equate the coordinates in $U \cap V$ through

$$o_k = \xi_j \frac{\partial x_j}{\partial y_k}$$

but $dy_k = \frac{\partial y_k}{\partial x_j} dx_j$. So we write the tautological one-form in $U \cap V$ as

$$\theta = \xi_j dx_j = o_j dy_j,$$

Therefore θ, and by extension its exterior derivative ω, are coordinate independent.

Definition B.1.6. If (M, ω_1) and (N, ω_2) are symplectic manifolds, and $f^*\omega_2 = \omega_1$ (the pullback of f acting on ω_2 is equal to ω_1), then f is called a *symplectomorphism* between the two symplectic manifolds.

A symplectomorphism is a formal mathematical definition of what we call a "canonical transformation" in physics.

B.2 Darboux's theorem

We finish by stating and proving a central theorem in symplectic geometry. Several of the concepts laid in appendix A are important for this.

Theorem B.2.1. *Let* (M, ω) *be a symplectic manifold of dimension* $2n$*. For each point* $p \in M$*, there exists a coordinate chart* $\xi = (q_1, \ldots, q_n, p_1, \ldots, p_n)$ *such that in some neighborhood* U *of* p

$$\omega|_U = dq_j \wedge dp_j$$

and the coordinates ξ *are called Darboux coordinates or canonical coordinates.*

Proof. Let $\omega_0 = dq_j \wedge dp_j$, where q_j, p_j is a symplectic basis of the coordinate chart near p. Our goal is to find a diffeomorphism φ such that $\varphi^*\omega = \omega_0$.

Take the family of 2-forms $\omega_t = (1-t)\omega + t\omega_0$. Then

$$\frac{d}{dt}\omega_t = \omega_0 - \omega.$$

Due to Poincaré's lemma (theorem A.1.2), in some neighborhood of p, there is a 1-form θ such that $d\theta = \omega_0 - \omega$.

Now, we may try to find a family of vector fields V_t such that their flows satisfy

$$\left(\phi^t_{\mathbf{V}_t}\right)^* \omega_t = \omega_0$$

Taking the derivative of both sides with respect to the parameter t:

$$\left[\frac{d}{dt}\left(\phi^t_{\mathbf{V}_t}\right)^*\right]\omega_t + \left(\phi^t_{\mathbf{V}_t}\right)^*\left(\frac{d}{dt}\omega_t\right) = 0$$

The first time derivative is the Lie derivative (definition A.1.2). The second is $\omega_0 - \omega = d\theta$, by the definition of ω_t and θ. So

$$\begin{aligned} \left(\phi^t_{\mathbf{V}_t}\right)^*\left(\mathcal{L}_{\mathbf{V}_t}\omega_t + d\theta\right) &= 0 \\ \implies \mathcal{L}_{\mathbf{V}_t}\omega_t + d\theta &= 0 \end{aligned}$$

If we use theorem A.1.1 to expand the Lie derivative,

$$\mathcal{L}_{\mathbf{V}_t}\omega_t = d\iota_{\mathbf{V}_t}\omega_t + \iota_{\mathbf{V}_t}d\omega_t$$

but ω_t is a closed form, so the second term on the right hand side vanishes. We then conclude

$$\begin{aligned} d\iota_{\mathbf{V}_t}\omega_t &= -d\theta \\ \implies \iota_{\mathbf{V}_t}\omega_t &= -\theta \end{aligned}$$

Since ω is a non-degenerate 2-form, there exists a unique vector field family $\mathbf{V}_t$ satisfying this relation. From there, we could solve the ODE in the parameter t

$$\frac{d}{dt}\left(\phi^t_{\mathbf{V}_t}\right) = \mathbf{V}_t\left(\phi^t_{\mathbf{V}_t}\right)$$

with the initial value $\phi^0_{\mathbf{V}_0} = 1$. From this, we get

$$\left(\phi^1_{\mathbf{V}_1}\right)^* \omega = \omega_0.$$

□

For singular systems, we would like to generalize this result to presymplectic (degenerate) forms. This is not trivial. We will give an outline of the theory needed to prove it.

Definition B.2.1. Let M be a smooth n-dimensional manifold. A *smooth k-differential system D* is a family of vector spaces $\{D_p \subset T_pM\}$ for any point $p \in M$ such that there is a neighborhood U of p and smooth vector fields $\mathbf{V}_1, \ldots, \mathbf{V}_k$ on U (called the *basis*) with $D_p = \operatorname{span}\{\mathbf{V}_1(p), \ldots, \mathbf{V}_k(p)\}$

A differential system is called *completely integrable* if for every point $p \in M$ there is a local coordinate chart $x_1, \ldots, x_n$ such that $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k}$ is a basis of D.

The following is an important theorem regarding smooth differential systems. The proof presented here is due to (CHERN; WOLFSON, 1981).

Theorem B.2.2 (Frobenius theorem)**.** *Let D be a smooth k-differential system in a smooth manifold M. Then D is completely integrable if, and only if, for every basis $\mathbf{V}_1, \ldots, \mathbf{V}_k$ in an open set $U \subset M$, the Lie bracket $[\mathbf{V}_i, \mathbf{V}_j] := \mathbf{V_i}(\mathbf{V_j}) - \mathbf{V_j}(\mathbf{V_i})$ for any $1 \leq i, j \leq k; i \neq j$ is also contained in D_p.*

Proof. Firstly, we note that if D is completely integrable, by definition, it has local bases of form $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k}$. The Lie bracket $\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right]$ is always 0 since these vector fields commute. Then, trivially, the Lie bracket is also in D_p.

For the converse, we use induction on k.

For $k = 1$, our differential system D would be (locally) spanned by a single vector field $\mathbf{V}$. For a given point $p \in M$, obviously we need $\mathbf{V}(p) \neq 0$. There exist coordinates

$x^1, \ldots, x^k$ such that $X = \frac{\partial}{\partial x_1}$; we can construct them from any arbitrary coordinate chart by composing the linear transformation that takes the representation of $\mathbf{V}$ in said chart to $\{1, 0, \ldots, 0\}$.

Now, for $k \geq 2$, assuming the result is valid for $k-1$, we know that we can choose a coordinate system centered on p such that

$$\mathbf{V}_k = \frac{\partial}{\partial x^k}$$

then, for each other vector field in the basis, we define a new field

$$\mathbf{V}_j' := \mathbf{V}_j - \mathbf{V}_j \left(x^k\right) \mathbf{V}_k, \quad j = 1, \ldots, k-1$$

We now define a $k-1$-differential system D' with basis $\mathbf{V}_1', \ldots, \mathbf{V}_{k-1}'$. With this basis, we have

$$\mathbf{V}_j' \left(y^k\right) = 0$$

for any j. And thus, the Lie bracket of any two such fields will also vanish at y^k

$$\left[\mathbf{V}_j', \mathbf{V}_{j'}'\right] \left(y^k\right) = 0$$

This means that D' satisfies our hypothesis. From the induction, there is another coordinate system centered on p, call it $z^1, \ldots, z^n$, where up to a linear isomorphism,

$$\mathbf{V}_j' = \frac{\partial}{\partial z^j}, \quad j = 1, \ldots, n$$

Then, the k-differential system with basis $\mathbf{V}_1', \ldots, \mathbf{V}_{k-1}', \mathbf{V}_k$, which is D, still satisfies the induction hypothesis. Now, write

$$\mathbf{V}_k = \alpha_j \frac{\partial}{\partial z^j}. \quad j = 1, \ldots, k-1$$

and the Lie bracket

$$\left[\mathbf{V}_j', \mathbf{V}_k\right] = \frac{\partial \alpha_l}{\partial z^j} \frac{\partial}{\partial z^l}$$

which can be seen to obligatorily be a linear combination of $\mathbf{V}_j'$ (i.e. $\frac{\partial}{\partial z^j}$) by noting that $\left[\mathbf{V}_j', \mathbf{V}_k\right] = 0$ at y^k, and thus the bracket has no $\mathbf{V}_k$-component. This also means that

$$\frac{\partial \alpha_l}{\partial z^j} = 0, \quad j = 1, \ldots, k-1, \quad l = k, \ldots, n$$

so that the last $n-k$ coefficients α_l do not depend on the first $k-1$ coordinates z^j. On the other hand, we already have the previous V_j' generating the $\frac{\partial}{\partial z^j}$ components for

$j = 1, \ldots, k-1$, so there exists a vector field

$$\mathbf{V}_k^{'} = \alpha_l \frac{\partial}{\partial z^l}, \quad l = k, \ldots, n$$

such that $\mathbf{V}_1^{'}, \ldots, \mathbf{V}_{k-1}^{'}, \mathbf{V}_k^{'}$ is also a basis of D. We now apply the base case $k = 1$ to $\mathbf{V}_k^{'}$, and we find a coordinate system $w^k, \ldots, w^n$ such that

$$\mathbf{V}_k^{'} = \frac{\partial}{\partial w^k}.$$

Then, the differential system D is generated by the set $\frac{\partial}{\partial z^1}, \ldots, \frac{\partial}{\partial z^{k-1}}, \frac{\partial}{\partial w^k}$ and is completely integrable. □

Lemma B.2.3. *Let ω be a degenerate differential 2-form with constant rank $2k$. The kernel of ω induces a differential system.*

Proof. This follows from defining the family $D = \{\ker \omega_p, p \in M\}$, the kernel of 2-form at each point, and the basis of the kernel at p is the basis of $\ker \omega_p$, giving D the structure of a differential system. □

Definition B.2.2. Let M be a n-dimensional manifold. A *foliation* is a partition of M into k-dimensional submanifolds K_j. Each K_j is called a *leaf.*

Any completely integrable k-differential system induces a k-foliation, where each leaf is generated by each basis vector field.

Theorem B.2.4 (Generalized Darboux theorem)**.** *Let M be a manifold of dimension n, and ω a closed differential 2-form on M of constant rank $2k$ in a neighborhood of a point p. There exists a coordinate chart $\xi = (q_1, \ldots, q_k, p_1, \ldots, p_k, z_1, \ldots, z_r)$, $\quad r = n - k$ such that in some neighborhood U of p*

$$\omega|_U = dq_j \wedge dp_j$$

Proof. The kernel of ω is composed of vector fields $\mathbf{V}$ such that

$$\iota_{\mathbf{V}} \omega = 0 \iff \omega(\mathbf{V}, \mathbf{W}) = 0, \ \forall \ \mathbf{W} \in TM$$

from Cartan's homotopy formula (A.1.1), the Lie derivative is also null $\mathcal{L}_{\mathbf{V}} \omega = 0$. We note that the Lie derivative is directly related to the Lie bracket of vector fields as

$$\mathcal{L}_{\mathbf{V}} \mathbf{W} = [\mathbf{V}, \mathbf{W}].$$

Then, from the Leibniz rule for a Lie derivative, we obtain

$$\mathcal{L}_{\mathbf{V}}\left[\omega\left(\mathbf{W},\mathbf{X}\right)\right]=\mathbf{V}\left(\omega\left(\mathbf{W},\mathbf{X}\right)\right)-\omega\left(\left[\mathbf{V},\mathbf{W}\right],\mathbf{X}\right)-\omega\left(\mathbf{W},\left[\mathbf{V},\mathbf{X}\right]\right)$$

if $\mathbf{V},\mathbf{W}\in\ker\omega$, then for any field $\mathbf{Z}$ this gives us

$$\omega\left(\left[\mathbf{V},\mathbf{W}\right],\mathbf{X}\right)=0$$

for any $\mathbf{Z}$. This implies that $[\mathbf{V},\mathbf{W}]\in\ker\omega$. So, the differential system in lemma (B.2.3) is also completely integrable due to Frobenius' theorem. There is then a coordinate system near any point p

$$\left(x^1,\ldots,x^{2k},z^1,\ldots,z^r\right)$$

where $\frac{\partial}{\partial z^j}$ is a basis for $\ker\omega$, and x^j is a basis for the foliation. Locally, the 2-form is written

$$\omega=\frac{1}{2}\omega_{jk}\left(x,z\right)dx^j\wedge dx^k \tag{B.1}$$

Now we use a modified version of Poincaré's lemma (A.1.2). We have not just a differential form, but rather a family of differential forms that depend smoothly on an additional parameter (in our case, the z coordinates).

Lemma B.2.5. *Let $\omega\left(y,z\right)$ be a family of differential k-forms on any open ball u in $\mathbb{R}^n$ depending smoothly on a parameter z. Then, if ω is closed for every z, and there is a z_0 such that $\omega\left(y,z_0\right)=0,\ \forall y\in U$, then there also exists a family of $(k-1)$-forms $\theta\left(y,z\right)$ depending smoothly on z with*

$$d_y\theta\left(y,z\right)=\omega\left(y,z\right)$$

and

$$\theta\left(y,z_0\right)=0,\forall y\in U.$$

We will not write out the proof of this lemma in full, but it is not too different from the proof of the regular Poincaré's lemma we used: if we fix a value of z, the regular theorem is valid, and the procedures used in the proof do not violate smoothness. The additional condition that the k-form vanishes on some value z_0 guarantees that θ is unique.

With this lemma, we can also apply the same strategy of the proof of the standard Darboux's theorem to the 2-form ω in (B.1). If we fix a value of z, then this ω is a symplectic.$2k$-form. This means that it has Darboux coordinates (q_j,p_j) such that

$$\omega=dq_j\wedge dp_j$$

once again, none of the procedures used in proving Darboux's theorem break smoothness on the parameters z^j, so ω still depends smoothly on them. We conclude that there exists a coordinate chart $\xi = (q_1, \ldots, q_k, p_1, \ldots, p_k, z_1, \ldots, z_r)$, $\quad r = n - k$, where locally $\omega = dq_j \wedge dp_j$ and $z_1, \ldots, z_r$ is a basis of $\ker \omega$. $\square$

FOLHA DE REGISTRO DO DOCUMENTO

1. CLASSIFICAÇÃO/TIPO	2. DATA	3. DOCUMENTO Nº	4. Nº DE PÁGINAS
DM	05 de março de 2026	DCTA/ITA/DM-016/2026	87

5. TÍTULO E SUBTÍTULO:

Singular systems and first-order actions: a comparative study

6. AUTOR(ES):

Gabriel Ronconi Prada

7. INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES):

Instituto Tecnológico de Aeronáutica – ITA

8. PALAVRAS-CHAVE SUGERIDAS PELO AUTOR:

Sistemas singulares; Teoria de campos; Eletrodinâmica; Física nuclear; Física

9. PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO:

Teoria quântica de campos; Campos eletromagnéticos; Eletrodinâmica; Equações de Faddeew; Física nuclear; Física.

10. APRESENTAÇÃO: (**X**) **Nacional** ( ) **Internacional**

ITA, São José dos Campos. Curso de Mestrado. Programa de Pós-Graduação em Física. Área de Física Nuclear. Orientador: Prof. Dr. Pedro José Pompeia. Defesa em 20/02/2026. Publicada em 2026.

11. RESUMO:

Singular systems are a vast and important category of physical systems, both in mechanics with finite degrees of freedom and in field theories. Among them, we can consider not only systems with physically evident kinematical constraints, but also theories with gauge symmetries, which includes many field theories such as electromagnetism. The development of different formalisms in theoretical physics has also provided various ways to understand and deal with the intricacies of this sort of system. In this work, we will compare some of these methods, studying singular systems in the Lagrangian formalism, and in the Hamiltonian formalism using two different procedures: the Dirac-Bergmann method and the Faddeev-Jackiw method. We begin by showing how each method is applied to a mechanical system, before making the transition to a field, which requires some changes to properly handle the infinite degrees of freedom. To this end, we will use the free electromagnetic field as a motivating example. Lastly, we will apply some of the concepts we acquired in these analyses to nonlinear electrodynamics (NLED), a family of various field theories with a broad range of applications, from which we will give special attention to the specific cases of Born-Infeld electrodynamics and a generalization recently proposed by Kruglov. To do this, we will also use a formulation of NLED based on Jordan frames and Palatini variations, which also allow us to extract more mathematical relations as system constraints, further reaffirming the importance of analyzing it as a singular system.

12. GRAU DE SIGILO:

(**X**) **OSTENSIVO** ( ) **RESERVADO** ( ) **SECRETO**