Characteristic Hypersurfaces and Constraint Theory

Thesis submitted for the Degree of MSc by
Patrik Omland

Under the supervision of
Prof. Stefan Hofmann

Arnold Sommerfeld Center for Theoretical Physics
Chair on Cosmology

Abstract

In this thesis I investigate the occurrence of additional constraints in a field theory, when formulated in characteristic coordinates. More specifically, the following setup is considered: Given the Lagrangian of a field theory, I formulate the associated (instantaneous) Hamiltonian problem on a characteristic hypersurface (w.r.t. the Euler-Lagrange equations) and find that there exist new constraints. I then present conditions under which these constraints lead to *symplectic* submanifolds of phase space. Symplecticity is desirable, because it renders Hamiltonian vector fields well-defined. The upshot is that symplecticity comes down to analytic rather than algebraic conditions.

Acknowledgements

After five years of study, there are many people I feel very much indebted to. Foremost, without the continuous support of my parents and grandparents, sister and aunt, for what is by now a quarter of a century I would not be writing these lines at all.

Had Anja not been there to show me how to get to my first lecture (and in fact all subsequent ones), lord knows where I would have ended up. It was through Prof. Cieliebaks inspiring lectures and help that I did end up in the TMP program. Thank you, Robert, for providing peculiar students with an environment, where they could forget about life for a while and collectively find their limits, respectively. In particular, I would like to thank the TMP lonely island faction.

It is thanks to Stefan Hofmann's open mindedness and enthusiasm that I will attempt a part-time PhD. It seems truly special in the theoretical sciences that professors directly respond to and communicate with students. And so it was also, when Mark J. Gotay wrote a detailed list of master thesis proposals to an unacquainted student that had contacted him via email.

My very special thanks goes out to Tine, for sharing all bitter times, albeit not being able to share the sweet times, of my bittersweet relationship with mathematical physics.

Thank You.

This thesis was written in LaTeX using MiKTeX and Emacs. Heavy use was made of the AMS-LaTeX packages and one or the other TikZ-picture shows up in this work. Thanks to all developers who make this available through some public license.

Contents

1. Introduction

In classical mechanics, at some point every physicist-to-be is asked to calculate the motion of a pendulum in three dimensions. The motion of the pendulum's point-mass is easily seen to be *constrained* to lie within a sphere of radius the length of the string. Usually, one then switches to a spherical coordinate system with fixed radius and has the constraint thus implemented from the start. We will briefly recall what happens when continuing in Cartesian coordinates.

[ST95] Take a free particle confined to the two-dimensional unit sphere in three-dimensional ambient space. Using a Lagrange multiplier λ and coordinates q^i, $i = 1, 2, 3$, the Lagrangian reads

$$L(q^i, \dot{q}^i, \lambda, \dot{\lambda}) = \frac{m}{2}\dot{q}^2 + \lambda(q^2 - 1)$$

The Euler-Lagrange equations are of course

$$\begin{aligned} m\ddot{q}_i - 2\lambda q_i &= 0 \quad \text{(evolution equation)} \\ q^2 - 1 &= 0 \quad \text{(constraint)} \end{aligned}$$

Trivially, the constraint equation has consequences on the initial position we may choose: It has to lie on the sphere! But this is not the whole story. Quite formally, the whole system of equations reads

$$\begin{aligned} m\ddot{q}_i - 2\lambda q_i &= 0 & \ddot{\lambda} &= 0 \\ q\dot{q} &= 0 & \dot{\lambda} &= 0 \\ q^2 - 1 &= 0 & m\dot{q}^2 + 2\lambda &= 0 \end{aligned}$$

In particular, the initial velocity needs to be tangent to the sphere. The Hamiltonian formalism lives on the space of initial conditions - so let's see what happens, when we switch to phase space. For the canonically conjugate momenta p_i, π, we calculate $p_i = m\dot{q}_i$, $\pi = 0$. Hence, from the start phase space is constrained to

$$\mathcal{P} = \{(q_i, \lambda, p_i, \pi) \in \mathbb{R}^8 \mid \pi = 0\} \cong \mathbb{R}^7$$

$\pi = 0$ is called a *primary constraint* and $\mathcal{P}$ the *primary constraint submanifold* of phase space. Primary constraints arise, when the Legendre transformation is not onto. The Hamiltonian becomes

$$\begin{aligned} H : \mathcal{P} &\longrightarrow \mathbb{R} \\ (q_i, \lambda, p_i, \pi) &\longmapsto \frac{1}{2m}p^2 - \lambda(q^2 - 1) \end{aligned}$$

Note, that being odd-dimensional, $\mathcal{P}$ is not a symplectic submanifold w.r.t. the canonical symplectic form. The Hamiltonian is defined only on $\mathcal{P}$, so we better restrict to initial values

such that we do not flow off $\mathcal{P}$ by introducing new constraints: In view of more intricate systems, Dirac[1] has developed an algorithm based on Poisson brackets to tackle this problem. In subsection 5.4 we introduce an algorithm due to Gotay, Nester and Hinds [GNH78] that is based on the presymplectic structure of $\mathcal{P}$.
In this case, also by simply using the Hamilton equations, we find the *secondary constraints*

$$q^2 = 1 \ , \ pq = 0 \ , \ p^2 = -2m\lambda$$

Secondary constraints are those that come from Hamilton's equation by these considerations. We call the submanifold satisfying all constraints the *final constraint submanifold* and denote it by C. Finally, by a classification due to Dirac, all these constraints are *second class* and this classification allows us to calculate the dynamical degrees of freedom via

$$\begin{aligned}\mathrm{DoF} &= D_{\mathrm{Dim.conf.space}} - \#_{\mathrm{1stClass}} - \frac{1}{2} \cdot \#_{\mathrm{2ndClass}} \\ &= 4 - 0 - \frac{1}{2} \cdot 4\end{aligned}$$

Which is equal to two - the dimension of the sphere. The classification will be explained in section 5.4.

Constraints are ubiquitous in field theory. In fact, if a field theory features gauge invariance, it necessarily features constraints (Proposition 5.5.3). The furthest reaching attempt to characterize the relationship between constraints and gauge invariance that the author is aware of, has been made in [GM06]. From there we take the Hamiltonian formulation of a field theory, recapitulated in section 6, and point out where the "gauge ↔ constraints" correlation starts off. In particular, we give an example in section 7.

So where do characteristics come in? Now, hyperbolic partial differential equations are primarily defined through the existence of associated characteristic hypersurfaces. For physicists, hyperbolic PDEs are formally defined as those PDEs, which feature unique solutions for an initial-value problem and whose solutions propagate with finite speed. In section 4 we will make the connection.

What makes characteristic hypersurfaces special in the theory of constraints is that they tend to bring about new constraints. This was already noted by [Ste80], who investigated quantization on characteristic hypersurfaces. We shall skip all the functional analysis involved for now (see section 5) and take as an example the massless Klein-Gordon field on a 4-Minkowski background:

$$\mathcal{L}(\phi, \partial_\mu \phi) = \int \frac{1}{2} \eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi \ , \ \eta \text{ Minkowski Metric}$$

[1] P.A.M. Dirac. Generalized Hamiltonian dynamics. Can. J. Math., 2:129-148, 1950

$$= \int \frac{1}{2} \left(\partial_+\phi\partial_-\phi - \partial_n\phi\partial_n\phi\right) \ , \ x_n = x_n \ , n = 2, 3 \ , x_\pm := \frac{1}{2}(x_0 \pm x_1)$$

Now, in the first coordinate system we choose x_0 as the evolution direction and $\{x_0 = 0\}$ as initial value hypersurface. In the second, we choose x_+ as the evolution direction and correspondingly $\{x_+ = 0\}$ as initial value hypersurface. Latter, as we will see in section 3, is characteristic w.r.t. the Euler-Lagrange equations. We calculate the canonical momenta π_0 and π_+, respectively

$$\pi_0 := \frac{\partial\mathcal{L}}{\partial(\partial_0\phi)}\Big|_{x_0=0} = \partial_0\phi\Big|_{x_0=0}$$
$$\pi_+ := \frac{\partial\mathcal{L}}{\partial(\partial_+\phi)}\Big|_{x_+=0} = \partial_-\phi\Big|_{x_+=0}$$

In the first case, it is clearly possible to substitute the canonical momentum π_0 for the velocity $\partial_0\phi$ on the hypersurface and obtain the *unconstrained* phase space $\mathcal{P}_0 = \{(\phi|_{x_0=0}, \pi_0)\}$. In the second case, this is not possible. Rather, we obtain the *constrained* subspace $\mathcal{P}_+ = \{(\phi|_{x_+=0}, \pi_+) \mid \pi_+ - \partial_-\phi|_{x_+=0} = 0\}$, where again, $\pi_+ - \partial_-\phi|_{x_+=0} = 0$ is referred to as a *primary constraint*.

This observation marks the starting point of this thesis: Is there a general property to this sort of new constraint? As Dirac's classification was introduced above, one could for instance wonder whether the constraint is first class. It turns out to be second class. Dirac conjectured that first class secondary constraints correspond to gauge freedom. Will this primary constraint give rise to such a secondary constraint? Does gauge invariance then depend on the initial value hypersurface? If, instead, it gives rise to a second class constraint, does the Klein Gordon field then have no degree of freedom (by the DoF-formula)? These questions are, of course, oversimplified, but not as much as one might think.

In fact, we found that the associated constraint submanifold of this particular setup is *symplectic* w.r.t. the canonical symplectic structure. In particular this means (as explained in section 5.4) that it will not give rise to a secondary constraint submanifold, whence all these questions are answered at once! Thus, this is the situation one would hope for. It then seems possible to ignore the fact that one is not working in the symplectic phase space, but rather in the symplectic characteristic constraint submanifold of phase space and continue with the usual analysis. Finally, we have set the stage for the question investigated in this thesis

Under which conditions is the characteristic constraint submanifold symplectic?

Part I. Mathematical Preliminaries

2. Implicit Systems of Partial Differential Equations

In this section we will give a short overview of the more general obstacles that arise when analyzing general differential equations. What is described here is essentially the reason for most of the machinery developed in later sections, just that the origin will be somewhat disguised. This section is based on the introduction in [Sei12].

As customary, we will be using subscripts to denote derivatives and split coordinates $(t, x) \in \mathbb{R}\times\mathbb{R}^n$, where t is the "distinguished"coordinate. The way in which this coordinate is distinguished becomes clear when one considers a **normal** system for an unknown function $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^m$

$$u_t = \phi(t, x, u, u_x)$$

Such a system is also said to be in **Cauchy-Kovalevskaya form**, because a well known theorem by the same name ensures local existence and uniqueness in case of analytic ϕ and initial conditions. Formally, this is readily seen, as one may consider the partial differential equation in normal form as an *ordinary* differential equation, in t, on an infinite dimensional space.

Physics is abound with *non*-normal systems, examples of which will be discussed in the course of this work. Hence, we consider the **implicit** first order system

$$\Phi(t, x, u, u_t, u_x) = 0$$

By the inverse function theorem, this system is equivalent to a normal system if and only if - possibly after a transformation of the independent variables (t, x) - the Jacobian $\partial\Phi/\partial u_t$ is regular. If this is not the case, one can separate the unknown function $u = (v, w)$ such that already the partial Jacobian $\partial\Phi/\partial v_t$ has the same rank as the full one $\partial\Phi/\partial u_t$. We may then make use of the implicit function theorem to obtain the **semi-explicit form**

$$v_t = \phi(t, x, v, w, v_x, w_x, w_t) \tag{1}$$

$$0 = \psi(t, x, v, w, v_x, w_x) \tag{2}$$

This system yields so-called **integrability conditions**: Taking a derivative w.r.t. t of Equation 2, we obtain

$$0 = \frac{\partial\psi}{\partial t} + \frac{\partial\psi}{\partial v} v_t + \frac{\partial\psi}{\partial w} w_t + \frac{\partial\psi}{\partial v_x} v_{xt} + \frac{\partial\psi}{\partial w_x} w_{xt}$$

we can now substitute v_t from Equation 1 to check consistency of the non-normal system: This will either yield an identity or a new equation. Repetition of this process should lead to an identity if the system is consistent.

Of course the process does not stop at this point, for Equation 2 is actually just another implicit first order system, whence the above process has to be repeated until one is left with an ordinary differential equation. The overall system is then said to be in **Cartan normal form**.[2]

In particular, it should be clear that one can not choose arbitrary initial data for the system. Choosing $v(0, x) = v_0$ and $w(0, x) = w_0$, Equation 2 requires $\psi(0, x, v_x, w_x) = 0$ and this condition is not even sufficient if integrability conditions are hidden in the system.

The most prominent examples of such systems are so called characteristic initial value problems.

3. The Characteristic Initial Value Problem

All of this section is taken from [zHS79]. In the first part of the section we will briefly review some basic notions in the theory of partial differential equations.Then we give an example, which illustrates the importance of the choice of hypersurface on which to place the initial data. We conclude this section with an existence theorem for compact domains, to emphasize that characteristic initial value problems can be well-posed problems and arise naturally in the realm of general relativity.
For further study, [Hoe83] features a whole section on the characteristic Cauchy Problem. He focuses on existence of a solution with values prescribed on a characteristic hyperplane, but does not treat uniqueness. [Lun78] gives an analysis of the Klein-Gordon equation with Light-Cone data, showing that this initial value problem is well-posed.

3.1. Basic Notions

Let $\mathcal{S}$ be the set of solutions $u : M \to \mathbb{R}^n$ of the partial differential equation

$$L[u] = L(D^2u, Du, u) = F(Du, u)$$

on a domain M with L a linear and F an arbitrary function. Let $\mathcal{D}$ be the **data** that is the set of certain values of u on a hypersurface H. The correspondence

$$\begin{aligned} \varphi : \mathcal{S} &\to \mathcal{D} \\ u &\mapsto \text{data}[u] \text{ (the } \textbf{datum} \text{ of } u \text{ on } H) \end{aligned}$$

[2]When there is gauge invariance, one does not have enough equations to obtain the Cartan normal form - thus one has to fix the gauges before the PDE can be solved.

is (normally) a well-defined function and continuous in any natural topology. If the inverse, φ^{-1}, is

- Defined on the whole of $\mathcal{D}$, we have an **existence property**
- A function on $\varphi(\mathcal{S}) \subset \mathcal{D}$, we have a **uniqueness property**
- A continuous function, we have a **stability property**

Let $\mathcal{B}$, $\mathcal{B}'$ be topological function spaces. If

$$\varphi^{-1} : \mathcal{D} \cap \mathcal{B}(H) \to \mathcal{S} \cap \mathcal{B}'(M)$$

is a continuous function, the search for the values of φ^{-1} is called a **well-posed** or **properly posed problem**.

3.2. The Wave Equation

For $(t, x, y) \in \mathbb{R}^3$, consider the partial differential equation

$$\partial_t^2 u = \partial_x^2 u + \partial_y^2 u$$

- Boundary-Value Problem in two Dimensions: Let $M \in \mathbb{R}^2 \ni (t, x)$ be the unit disk, $H := \partial M$. This problem is improperly posed, consider e.g. $u_a := a(x^2 + t^2) - a$. u_a solves the 2-dim. wave equation for all $a \in \mathbb{R}$, but $u_a|_H = 0$.[3]
- The analytic Cauchy problem. $H = \{t = 0\}$: $\varphi(x, y) = u|_h$, $\psi(x, y) = \partial_t u|_H$. Using the PDE, all derivatives can be calculated algebraically from the data at any point of H:

  $$\begin{aligned} \partial_t^{2k} u\Big|_H &= (\partial_x^2 + \partial_y^2)^k \varphi \\ \partial_t^{2k+1} u\Big|_H &= (\partial_x^2 + \partial_y^2)^k \psi \end{aligned}$$

  This problem is properly posed on suitable function spaces.
- The characteristic problem. $H = \{x = t\}$, introducing new coordinates, $\xi := t + x$, $\eta = t - x$, the wave equation reads

  $$4\partial_\eta \partial_\xi u = \partial_y^2 u$$

  Then, with $\varphi(\xi, y) = u|_H$ we obtain

  $$4^k \partial_\xi^k (\partial_\eta^k u) = \partial_y^{2k} u \tag{3}$$

  $$\partial_\eta u\Big|_H (\xi, y) = \frac{1}{4} \int^\xi \partial_y^2 \varphi(\bar{\xi}, y) d\bar{\xi} \tag{4}$$

[3] On the other hand, for e.g. Poisson's equation, this is properly posed.

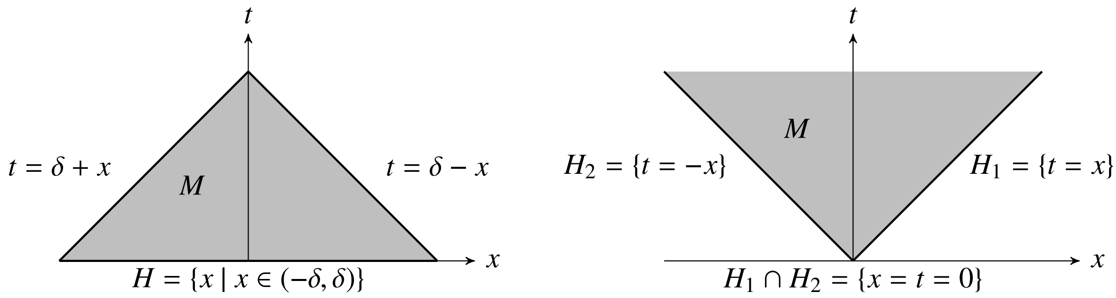

Figure 1: Two plots, the first of which picturing the analytic Cauchy problem (for bounded H) and the second one the characteristic problem for the wave equation, both in two dimensions. Compare this to Corollary 4.0.6ff.

$$\partial_\eta^k u\Big|_H = \left(\frac{1}{4}\right)^k \int^{\xi} (...) \left(\int^{\xi'} \partial_y^{2k} \varphi(\bar{\xi}, y) d\bar{\xi} \right) d\xi^1 ... d\xi^n \tag{5}$$

Hence these derivatives neither can be arbitrarily given as another datum on H, nor can be determined, as the constant of integration is not fixed.

Definition 3.2.1. A hypersurface H on which the kth outgoing derivative of a solution of Equation 3 is not determined by the derivatives up to the $(k-1)$-st order given on H is called **characteristic**. The lines along which the outgoing derivatives can be calculated by propagation equations of the type of Equation 3 are called **bicharacteristics**.

Say, we define $H = H_1 \cup H_2$, $H_1 = \{0 \le t = x\}$, $H_2 = \{0 \le t = -x\}$ with data $\varphi_1(\xi, y) = u|_{H_1}$, $\varphi(\eta, y) = u|_{H_2}$ s.t. $\varphi_1(0, y) = \varphi_2(0, y)$ (cpw. Figure 1). Then the constant of integration in Equation 4 is fixed on $H_\cap := H_1 \cap H_2$ by $\partial_\eta u|_{H_1}(0, y) = \partial_\eta \varphi_2(0, y)$.
Still, if we want to calculate the derivatives at points of $H \backslash H_\cap$, then, in contrast to the non-char. problem, we can not use an algebraic condition, but have to solve the propagation equations, Equation 3. One can show (e.g. [Lun78]) that this problem *is well-posed.*

3.3. The more general Problem

[zHS79] is really based on more general considerations, which can be found in [zH90]. We chose to stick with Lorentzian metrics as in [zHS79].
Let M be a compact C^∞ 4-manifold with boundary ∂M, $\hat{g}$ a metric of signature $(- + ++)$, defining an affine connection with covariant derivative D. g a metric of the same signature. For raising/lowering indices, g will be used. Let H_1 and H_2 be two non-parallel g-null-hypersurfaces that intersect along a 2-submanifold and define $H = H_1 \cup H_2$ (see Figure 1). These will be characteristic for the following PDE.[4]

[4]See 8.6 for the explicit calculation

Assume M is a hyperbolic set based on H (cp. [HE75]). We consider the 2nd order hyperbolic equation

$$L[u] = g^{ab} D_a D_b u + B^a D_a u + Cu = F$$

With an unknown tensor field u.[5] This equation is called

- **linear** if the coefficients B, C, F and the metrics g, $\hat{g}$ are given tensor fields[6]
- **weakly coupled quasi-linear** if g, $\hat{g}$ depend on u and the coefficients depend on u, Du.

Definition 3.3.1. Let $\tilde{W}^m(M)$ denote the set of functions having finite $\|\cdot\|_m^M$-norm, where

$$\left(\|q\|_m^M\right)^2 := \|q\|_{H^m(M)} + \sum_{i=1,2} \sum_{k=0}^{m-1} \left(\|(n_i^a D_a)^k q\|_{H^{2(m-k)-1}(H_i)}\right)^2$$

with n_i^a the g-null vector field on H_i, $i = 1, 2$. Note that $C^\infty(M)$ is dense in $\tilde{W}^m(M)$.

Theorem 3.3.2. *Assume the following norms exist:* $\|\hat{g}\|_{m+1}^M$, $\|\cdot\|_m^M$ *for* g *and* n_i^a, $\|\cdot\|_{m-1}^M$ *for* B, C, F. *For the data* $u|_{H_1}$ *and* $u|_{H_2}$ *we require that* $\|\cdot\|_{H^m(H_i)}$ *is finite and* $u|_{H_1} = u|_{H_2}$ *on* $H_\cap$. *Then*

1. $\exists$ *uniquely determined solution* u
2. $\|u\|_m^M < \infty$.

3.3.0.1. Remark:

- Note the high differentiability requirements in the second term of $\|\cdot\|_m^M$.
- There are local version of this theorem for the quasi-linear case
- There are existence and uniqueness theorems for the characteristic initial-value problem for Einstein's vacuum field equations

In the next section we will see that characteristics need not only be considered in characteristic initial value problems. More generally, they lead to the most global statement one can make on (hyperbolic) PDEs that at the same time goes to the very heart of physics.

[5] Note, that assuming D to be a covariant derivative of another metric is nothing fancy, for we could just as well truncate the non-derivative parts of D into B and C.

[6] We could then write in index notation: $L[u] = g^{ab} D_a D_b u^A + b^a{}_B{}^A D_a u^B + c^A{}_B u^B = F^A$

4. Characteristics

This section is based on [Rau97]. We give a version of one of the most prominent theorems of the theory of partial differential equations, John's Global Holmgren Theorem. For mathematicians this theorem establishes uniqueness of solutions to a partial differential equation in its scope. For physicists it establishes the mathematical formulation of causality.
We briefly mention another property of characteristics concerning uncontinuities (e.g. shock waves) that will not be elaborated further: Suppose a solution, u, to a linear partial differential equation, $Pu = 0$, is piecewise smooth along a hypersurface Σ. This basically means that for every point of the hypersurface, the solution is smooth on either side of the hypersurface and can be C^m- continued to the hypersurface from each side, respectively. Then, if u is C^m, $Pu \in C^\infty$, and u is not C^∞ on a neighborhood of $x \in \Sigma$, then Σ must be characteristic at x.

Let $\Omega \subset \mathbb{R}^d$ be an open subset and $P(x, D)$ an mth order linear partial differential operator on Ω with coefficients in $C^\omega(\Omega)$.

Definition 4.0.3. The **principal symbol** of $P = \sum a_\alpha(x)D^\alpha$ is the function

$$P_m(x, \xi) := \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha$$

Where we used multi-index notation. Note that P_m is a homogeneous polynomial of degree m in ξ.

Definition 4.0.4. A smooth hypersurface, $\Sigma \subset \Omega$, is **characteristic** at $x \in \Sigma$ if $P_m(x, \xi) = 0$, for all $\xi \in N_x^*(\Sigma)$. Here, $N_x^*(\Sigma)$ denotes the annihilator of $T_x\Sigma$. A hypersurface which is characteristic at all points is called a **characteristic hypersurface**.

Suppose that $\Sigma \subset \Omega$ is a non-characteristic immersed C^m hypersurface. We will now define a continuous deformation of Σ through a one-parameter family of hypersurfaces Σ_λ, whose ends lie in Σ.
We require that the hypersurfaces Σ_λ be images of a fixed set $\theta \subset \mathbb{R}^{d-1}$ by a map σ depending on $\lambda \in [0, 1]$. We suppose that

1. $\theta \subset \mathbb{R}^{d-1}$ is open and $\sigma : [0, 1] \times \text{closure}(\theta) \to \Omega \subset \mathbb{R}^d$ is continuous.
2. For each $\lambda \in [0, 1]$, $\sigma(\lambda, \cdot) : \theta \to \mathbb{R}^d$ is a C^m immersion of a non-characteristic hypersurface, Σ_λ.
3. The initial surface, Σ_0 is a subset of Σ.
4. $\sigma([0, 1] \times \partial\theta) \subset \Sigma$, which expresses the fact that the edge $\sigma(\lambda, \partial\theta)$ of Σ_λ lies in Σ.

Then we have:

Theorem 4.0.5. *(John's Global Holmgren Theorem). If $u \in C^m(\Omega)$, $Pu = 0$ in Ω and $\partial^\alpha u|_\Sigma = 0$ for $|\alpha| \leq m-1$, then for $|\alpha| \leq m-1$, $\partial^\alpha u = 0$ on $\sigma([0,1] \times \mathrm{closure}(\theta))$.*

As an application we take the d+1-dimensional wave equation, $P(\partial_t, \partial_x) = \partial_t^2 - c^2\Delta_x := \Box$. As an immediate consequence we obtain

Corollary 4.0.6. *If $u \in C^2(\mathbb{R}_t \times \mathbb{R}_x^d)$ satisfies $\Box u = 0$ and $u|_{t=0} = \partial_t u|_{t=0} = 0$ on $|x| < R$, then*

$$u = 0 \text{ in } \Big\{(t,x) : |x| < R - c|t|\Big\}$$

And in the same spirit we can deduce

Corollary 4.0.7. *Suppose that $u \in C^2(\mathbb{R}^{1+d})$ satisfies $\Box u = 0$, and that K is the support of the initial data*

$$K = \operatorname{supp} \partial_t u|_{t=0} \cup \operatorname{supp} u|_{t=0} \subset \mathbb{R}_x^d$$

Then

$$\operatorname{supp} u \subset \Big\{(t,x) \mid \text{ distance}(x,K) \leq c|t|\Big\}$$

I.e. the waves (solutions u), propagate at speeds less than or equal to c. Now, the reader may go back to Figure 1 to make a connection between this and the preceding section.

5. Banach Spaces and Differential Calculus

The rest of this work will predominantly be concerned with vector space structures. In particular the vector space structure certain function spaces are equipped with. While this might be familiar to most readers, the generalization of calculus to functions on such vector spaces, let alone infinite dimensional manifolds based on those, might be less so. This section gives a brief introduction that turns quickly to symplectic structures. Here, readers acquainted with symplectic geometry will find some basic results generalized to the infinite-dimensional realm. The following section is based on [AMR88] and we will pick up citations again, once different sources come in.

5.1. Basics

Definition 5.1.1. Let $(E, ||\cdot||)$ be a normed space. If the associated metric d is complete, we say $(E, ||\cdot||)$ is a **Banach Space**. If $(E, \langle\cdot,\cdot\rangle)$ is an inner product space, whose associated metric is complete, we say $(E, \langle\cdot,\cdot\rangle)$ is a **Hilbert Space**.

But then, in a Hilbert space, we would of course like to know what orthogonal means

Proposition 5.1.2. *If E is a Hilbert space and F a closed subspace, then $E = F \oplus F^{\perp}$. Thus every closed subspace of a Hilbert space splits.*

In finite dimensional vector spaces, one is used to identifying covectors with vectors. It turns out that this is also possible in Hilbert spaces:

Theorem 5.1.3. *(Riesz Representation Theorem) Let E be a real (resp. complex) Hilbert space. The map $e \mapsto \langle\cdot, e\rangle$ is a linear (resp. anti-linear) norm-preserving isomorphism of E with E^*; for short $E \cong E^*$.*

This also implies that in a real Hilbert space E, every continuous linear function $l : E \to \mathbb{R}$ can be written $l(e) = \langle e, e_0\rangle$ for some $e_0 \in E$ and $||l|| = ||e_0||$.[7]

In a general Banach space E, we do not have such an identification of E and E^*. We do, however have a canonical map $i : E \to E^{**}$ defined by $i(e)(l) := l(e)$. One calls E **reflexive** if i is onto. Every finite dimensional or Hilbert space is reflexive, as are the L^p-spaces for $1 < p < \infty$.

Now, we are already set to start with the calculus part that really is 'only' a very obvious generalization of known concepts.

Definition 5.1.4. Let E, F be normed vector spaces, with maps $f, g : U \subset E \to F$ where U is open in E. We say f and g are **tangent** at the point $u_0 \in U$ if

$$\lim_{u\to u_0} \frac{||f(u) - g(u)||}{||u - u_0||} = 0$$

where $||\cdot||$ represents the norm on the appropriate space.

[7] $||\cdot||$ also denoting the induced operator norm, $||l|| := \sup\{||le|| \mid ||e|| = 1\}$

Proposition 5.1.5. *For $f : U \subset E \to F$ and $u_0 \in U$ there is at most one $L \in L(E, F)$* [8] *such that the map $g_L : U \subset E \to F$ given by $g_L(u) = f(u_0) + L(u - u_0)$ is tangent to f at u_0.*

Definition 5.1.6. If, in Proposition 5.1.5, there is such an $L \in L(E, F)$, we say f is **differentiable at** u_0 and define the **derivative** of f at u_0 to be $Df(u_0) = L$. If f is differentiable at each $u_0 \in U$, the map

$$Df : U \to L(E, F);\ u \mapsto Df(u)$$

is called the **derivative** of f.

As we will be working a lot with sections of fiber bundles, it is common to define

Definition 5.1.7. Suppose $f : U \subset E \to F$ is of class C^1. Define the **tangent** of f to be the map

$$Tf : U \times E \to F \times F;\ Tf(u, e) = (f(u), Df(u)e)$$

It turns out that we can carry over many properties of the derivative that we are used to from finite-dimensional calculus. To explicitly list two non-trivial ones

Proposition 5.1.8.

- $T(g \circ f) = Tg \circ Tf$ *(Chain Rule)*
- *Let $f_i : U \subset E \to F_i$, $i = 1, 2$ be differentiable maps and $B \in L(F_1, F_2; G)$. Then the mapping $B(f_1, f_2) = B \circ (f_1 \times f_2) : U \subset E \to G$ is differentiable and*

$$D\left(B(f_1, f_2)\right)(u)e = B\left(Df_1(u)e, f_2(u)\right) + B\left(f_1(u), Df_2(u)e\right)\ \textit{(Leibniz Rule)}$$

We can also carry over the differentiation along a path

Definition 5.1.9. Let $f : U \subset E \to F$ and let $u \in U$. We say that f **has a derivative in the direction** $e \in E$ at u if

$$\left.\frac{d}{dt} f(u + te)\right|_{t=0}$$

exists. We call this element of F the **directional derivative** of f in the direction e at u.

Proposition 5.1.10. *If f is differentiable at u, then the directional derivatives of f exist at u and are given by*

$$\left.\frac{d}{dt} f(u + te)\right|_{t=0} = Df(u)e$$

Finally, we state the implicit function theorem. This will become important later, when we look at the Legendre transformation.

[8] $L(E, F)$ denoting the normed space of all continuous linear maps of E to F

Theorem 5.1.11. *(Implicit Function Theorem) Let $U \subset E$, $V \subset F$ be open and $f : U \times V \to G$ be C^r, $r \geq 1$. For some $x_0 \in U$, $y_0 \in V$ assume $D_2 f(x_0, y_0) : F \to G$ is an isomorphism.*[9] *Then there are neighborhoods U_0 of x_0 and W_0 of $f(x_0, y_0)$ and a unique C^r map $g : U_0 \times W_0 \to V$ such that for all $(x, w) \in U_0 \times W_0$,*

$$f(x, g(x, w)) = w$$

5.2. The Functional Derivative and the Calculus of Variations

In view of the Euler-Lagrange equations the importance of the functional derivative in this work can not be overstressed.

Definition 5.2.1. Let E and F be Banach spaces. A continuous bilinear functional $\langle \cdot, \cdot \rangle : E \times F \to \mathbb{R}$ is called E-**non-degenerate** if $\langle x, y \rangle = 0$ for all $y \in F$ implies $x = 0$. Similarly, it is F-non-degenerate if (...). If it is both, we say it is **non-degenerate**. Equivalently, the maps $x \mapsto \langle x, \cdot \rangle$ and $y \mapsto \langle \cdot, y \rangle$ are one-to-one. If they are isomorphisms, $\langle \cdot, \cdot \rangle$ is called $E-$ or $F-$ **strongly non-degenerate**.
We say E and F are in **duality** if there is a non-degenerate bilinear functional $\langle \cdot, \cdot \rangle : E \times F \to \mathbb{R}$, also called a **pairing** of E and F.

Definition 5.2.2. Let E and F be normed spaces and $\langle \cdot, \cdot \rangle : E \times F \to \mathbb{R}$ be an E-weakly non-degenerate pairing. Let $f : F \to \mathbb{R}$ be differentiable at the point $\alpha \in F$. The **functional derivative** $\delta f / \delta \alpha$ of f w.r.t. α is the unique element in E, if it exists, such that

$$Df(\alpha)\beta = \left\langle \frac{\delta f}{\delta \alpha}, \beta \right\rangle \ , \ \forall \beta \in F$$

Suppose E is a Banach space of functions φ on a region $\Omega \subset \mathbb{R}^n$ and $f : E \to \mathbb{R}$. The functional derivative $\frac{\delta f}{\delta \varphi}$ of f w.r.t. φ is the unique element $\frac{\delta f}{\delta \varphi} \in E$, if it exists, s.t.

$$Df(\varphi)\psi = \left\langle \frac{\delta f}{\delta \varphi}, \psi \right\rangle = \int_\Omega \left(\frac{\delta f}{\delta \varphi} \right)(x) \psi(x) d^n x \ , \ \forall \psi \in E$$

Proposition 5.2.3. *Let E be a space of functions, as above. A necessary condition for a differentiable function $f : E \to \mathbb{R}$ to have an extremum at φ is that*

$$\frac{\delta f}{\delta \varphi} = 0$$

[9] D_2 denoting the partial derivative w.r.t. the second argument

5.3. Symplectic Geometry

Analogously to finite dimensional manifolds, we define a possibly *infinite* dimensional **Banach manifold** M, to be a a pair $(S, \mathcal{D})$, where S is a set and $\mathcal{D}$ is a differential structure on S, the charts of which take their values in a Banach space E. E is then called the **model space** (cp. [AMR88]).[10]

While the definition of differential forms directly carries over to the infinite dimensional case, their exterior derivative does not. Rather than defining the exterior derivative directly, we will define it through an operation whose definition does carry over directly to the infinite dimensional case

Definition 5.3.1. If $X \in \mathfrak{X}(M)$, we let $\pounds_X$ be the unique *differential operator* on $\mathcal{T}(M)$ [11], called the **Lie derivative** w.r.t. X, such that

1. For $f : M \to \mathbb{R}$ differentiable, we have $\pounds_X f(m) = df(m) \cdot X(m)$ for any $m \in M$.
2. If $X, Y \in \mathfrak{X}^r(M)$, $r \geq 1$, and X has flow F_t, $\pounds_X Y = [X, Y] := \frac{d}{dt}\big|_{t=0}(F_t^* Y)$

Note, how this defines the operation of $\pounds_X$ on covector fields α: Let $\alpha(Y) \in C^r(M)$, $r \geq 1$, $Y \in \mathfrak{X}(M)$. Then $\pounds_X(\alpha(Y)) = (\pounds_X\alpha)(Y) + \alpha(\pounds_X(Y))$, because $\pounds_X$ is a differential operator. This equation is readily solved for $\pounds_X\alpha$. Now, we can formulate a global definition for the exterior derivative:

Definition 5.3.2. The **exterior derivative** of a differential k-form ω is the unique k+1-form $d\omega$ defined by

$$
\begin{aligned}
d\omega(X_0, X_1, ..., X_k) := &\sum_{l=0}^{k}(-1)^l \pounds_{X_l}\left(\omega(X_0, ..., \hat{X}_l, ..., X_k)\right) \\
&+ \sum_{0\leq i\leq j\leq k}(-1)^{i+j}\omega\left(\pounds_{X_i}(X_j), X_0, ..., \hat{X}_i, ..., \hat{X}_j, ..., X_k\right)
\end{aligned}
$$

where $\hat{X}_l$ denotes that X_l is deleted.

In a coordinate neighborhood U, denoting elements of TU by (u, v), the definition reduces to the more familiar expression

$$
d\omega(u)(v_0, ..., v_k) = \sum_{i=0}^{k}(-1)^i(D\omega(u) \cdot v_i)(v_0, ..., \hat{v}_i, ..., v_k)\ ,\ \ v_i \in T_uU
$$

In particular one can check that Cartan's magic formula holds in the infinite dimensional case, namely

[10] For a short introduction to infinite dim. manifolds, see `http://johncarlosbaez.wordpress.com/2012/03/12/fluid-flows-and-infinite-dimensional-manifolds/`

[11] $\mathcal{T}(M)$ denotes the full tensor algebra of M

$$\pounds_X\alpha = i_X d\alpha + d i_X\alpha$$

Definition 5.3.3. [GNH78] Let M be a manifold modeled on a Banach space E and suppose that ω is a closed 2-form on M. Then (M,ω) is said to be a **strong symplectic manifold** if the linear map $\flat : TM \to T^*M$ defined by $\flat(X) =: X^\flat := X \lrcorner \omega$ is an isomorphism.
If $\flat$ is injective but not surjective, (M,ω) will be called a **weak symplectic manifold**, ω being *weakly* nondegenerate.

Note that in the finite dimensional case, there is no distinction between weak and strong symplectic forms. For brevity we will call strong symplectic manifolds simply **symplectic manifolds**, while weakly nondegenerate and degenerate forms will be referred to as **presymplectic**. Finally, we say that ω is **topologically closed** provided the map $\flat$ is a closed map.

In Physics, phase space is usually defined to be the cotangent bundle of configuration space (a manifold). On a cotangent bundle, there exists a natural *weakly* symplectic structure:

Definition 5.3.4. [CM70] Let M be a manifold modeled on a Banach space E. Let T^*M be its cotangent bundle and $\tau : T^*M \to M$ the projection. The **canonical 1-form** on T^*M is defined by

$$\theta(\alpha_m)\cdot w = \alpha_m \cdot T\tau(w)\ ,\ \alpha_m \in T^*_m M\ ,\ w \in T_{\alpha_m}(T^*M)$$

In a chart $U \subset E$, we have $\theta(x,\alpha)\cdot(e,\beta) = \alpha(e)$, with $(x,\alpha) \in U \times E^*$, $(e,\beta) \in E \times E^*$.
The **canonical 2-form** is defined by

$$\begin{aligned} \omega &: = -d\theta\ ,\ \text{ locally we have} \\ \omega(x,\alpha)\cdot((e_1,\alpha_1),(e_2,\alpha_2)) &= d(\theta(e_1,\alpha_1))(e_2,\alpha_2) - (1 \leftrightarrow 2) \\ &= d(\alpha(e_1))(e_2,\alpha_2) - (1 \leftrightarrow 2) \\ &= \alpha_2(e_1) - \alpha_1(e_2) \end{aligned}$$

Of course, we will now have to justify the name given to the two-form defined above:

Theorem 5.3.5. *[CM70] The canonical symplectic form is*

1. *Weakly symplectic on T^*M*
2. *Strongly symplectic iff E is reflexive*

Proof. For the proof we need a well-known corollary to the Hahn-Banach theorem, which we will state as a lemma

Lemma 5.3.6. *[AMR88] Let E be a normed vector space and $e \neq 0$. Then there exists $f \in E^*$ such that $f(e) \neq 0$. In other words, if $f(e) = 0$ for all $f \in E^*$, then $e = 0$; that is, E^* separates points of E.*

Using the lemma above for the expression of the canonical symplectic form in local coordinates, the first part of the theorem holds trivially. For the second part we must show that the map

$$\omega^\flat : E \times E^* \to (E \times E^*)^* = E^* \times E^{**} \ , \ \omega^\flat(e_1, \alpha_1) \cdot (e_2, \alpha_2) = \alpha_2(e_1) - \alpha_1(e_2)$$

is onto, i.e. we assume that E is reflexive, then $E^* \times E^{**} \cong E^* \times E$ and the map is obviously onto. Conversely, if $\omega^\flat$ is onto, then for $(\beta, f) \in E^* \times E^{**}$, there is (e_1, α_1) s.t. $f(\alpha_2) + \beta(e_2) = \alpha_2(e_1) - \alpha_1(e_2)$, $\forall e_2, \alpha_2$. Setting $e_2 = 0$, we see $f(\alpha_2) = \alpha_2(e_1)$, so $E \to E^{**}$ is onto. □

Of course, one is always interested in structure-preserving maps. In symplectic geometry, these are called symplectomorphisms and it turns out that they describe the evolution of a physical system. Without further citation, the rest of the section is based on [CM70].

Let (P, ω), (P', ω'), be weak symplectic manifolds (ω, ω' are weak symplectic). A (smooth) map $f : P \to P'$ is called **canonical** (usually in physics) or **symplectic** or a **symplectomorphism** when $f^*\omega' = \omega$.
As an example that we will encounter again later, we show how to lift a function on a manifold to a function on the associated cotangent bundle of that manifold, s.t. the latter is a symplectomorphism w.r.t to the canonical structure.

Proposition 5.3.7. *Let M be a manifold and $f : M \to M$ a diffeomorphism; define the* ***canonical lift*** *of f by*

$$T^*f : T^*M \to T^*M$$
$$T^*f(\alpha_m) \cdot v := \alpha_m \cdot (Tf \cdot v) \ , \ v \in T^*_{f^{-1}(m)}M$$

Then T^*f is symplectic w.r.t. the canonical symplectic form.

Definition 5.3.8. Let (P, ω) be a strongly symplectic manifold and $H : P \to \mathbb{R}$ a given smooth function. We define the **Hamiltonian vector field** of H, X_H, as the one satisfying

$$\omega_x \left(X_H(x), v\right) = dH_x \cdot v \ , \ \forall x \in P \ , \ v \in T_xP$$

It is easily seen that $X_H[H] = 0$. In particular this means that along an integral curve of X_H, H will be constant (conservation of energy). But now to the familiar form of **Hamilton's equations**:

Proposition 5.3.9. *Suppose M is modeled on a reflexive space E. If $P = T^*M$ and H is a smooth function on P, then locally*

$$X_H = (D_2H, -D_1H)$$

In a finite-dimensional symplectic space (M, ω) of $2n$ dimensions, say, one can proof that ω^n is a volume form. The following (valid also in infinite dimensions) will then directly lead to the well known Liouville theorem

Proposition 5.3.10. *Let (P, ω) be a symplectic manifold, $H : P \to \mathbb{R}$, and let F_t be the flow of X_H. Then for each t, F_t is a symplectomorphism.*

Proof.

$$0 = d\left(i_{X_H}\omega\right) = \pounds_{X_H}\omega = \frac{d}{dt}\Big|_{t=0} F_t^*\omega$$

Hence $F_t^*\omega$ is constant. But $F_{t=0} = \mathrm{Id}$. □

5.3.0.2. Example: Klein-Gordon on flat background

We consider the Klein-Gordon field on $\mathbb{R}^n$ that is some function space E over $\mathbb{R}^n$. On the tangent space $TE = E \oplus E$, we have the Lagrangian

$$\begin{aligned} L(\phi, \dot\phi) &= \frac{1}{2}\int_{\mathbb{R}^n}\left(\dot\phi^2 - (\nabla\phi)^2 + m^2\phi^2 + 2F(\phi)\right)dx \\ &= \frac{1}{2}\left(\langle\dot\phi, \dot\phi\rangle_{L^2(\mathbb{R}^n)} - \langle\nabla\phi, \nabla\phi\rangle_{L^2(\mathbb{R}^n)} + m^2\langle\phi, \phi\rangle_{L^2(\mathbb{R}^n)}\right) + \int_{\mathbb{R}^n} F(\phi)dx \end{aligned}$$

Judging from the Lagrangian, a sensible model space for the tangent space of this theory might be $TE = H^1(\mathbb{R}^n) \times L^2(\mathbb{R}^n)$.[12] We do not need to switch to T^*E, because in this case we can use the metric associated to the L^2-inner product[13] to pull back the canonical symplectic structure from $T^*E = E \oplus E^*$ (this is the reason why we will introduce L-subscripts to the canonical forms).[14] The canonical forms *on* TE read

$$\begin{aligned} \theta_L(x, e) \cdot (\alpha, \beta) &= -\langle e, \alpha\rangle_{L^2} \\ \omega_L(x, e) \cdot (\alpha, \beta; \alpha', \beta') &= \langle\alpha, \beta'\rangle_{L^2} - \langle\alpha', \beta\rangle_{L^2} \end{aligned}$$

Defining the Hamiltonian of the theory in the usual way,

$$H(\phi, \dot\phi) = \frac{1}{2}\int_{\mathbb{R}^n}\left(\dot\phi^2 + (\nabla\phi)^2 + m^2\phi^2 + 2F(\phi)\right)dx$$

We seek the Hamiltonian vector field on $E \oplus E$. We compute

$$dH(\phi, \dot\phi) \cdot (\alpha, \beta) = \int_{\mathbb{R}^n}\left(\dot\phi\beta + \nabla\phi \cdot \nabla\alpha + m^2\phi\alpha + F'(\phi)\alpha\right)dx$$

Writing $X_H(\phi, \dot\phi) = (Y(\phi, \dot\phi), Z(\phi, \dot\phi))$, we set

$$dH(\phi, \dot\phi) \cdot (\alpha, \beta) = \omega_L(\phi, \dot\phi) \cdot (Y, Z; \alpha, \beta)$$

[12] For a more complicated example, see [GNH78], where a well defined Lagrangian does not lead to a well defined Hamiltonian theory.

[13] As the Lagrangian has to be well defined, we will in any case be inside some L^2 subspace

[14] Note that this is what is usually done in introductory courses to mechanics. Though there of course in a finite-dimensional setting.

and, using some partial integration, find

$$
\begin{aligned}
Y(\phi, \dot{\phi}) &= \dot{\phi} \\
Z(\phi, \dot{\phi}) &= \Delta\phi - m^2\phi - F'(\phi)
\end{aligned}
$$

Now, if we take a 1-parameter family $\phi_t(x) = \phi(t, x)$ and lift it to $(\phi, \frac{d}{dt}\phi) \in TE$, *formally*, it will be the flow of X_H, if

$$\frac{\partial^2 \phi}{\partial t^2} = \Delta\phi - m^2\phi - F'(\phi)$$

We emphasize 'formally', because Δ, etc. are *not* continuous operations on most function spaces.

5.4. Geometric Constraint Theory

The following algorithm is due to Gotay, Nester and Hinds [GNH78]. As mentioned in the introduction, Dirac came up with a first procedure to treat constrained dynamics. However, Dirac's method is based on ideas from (finite-dimensional) classical mechanics. Contrary to common believe, it does not simply generalize to the infinite-dimensional realm. For discussion, please refer to [Ste80], [Sun82], [ST95] and [Sei12].

We have seen in the last section that the Hamilton equations lead to the "equations of motion" for a field theory - but what do we do if the underlying space is only presymplectic?

In this section we will investigate the question of what is meant by "consistent equations of motion", given a *pre*symplectic phase space (M, ω) and a Hamiltonian H. This section is based on [GNH78] and everything that is not referenced otherwise will be assumed to be taken from there, possibly with modifications.

Specifically, we will investigate the following system. Let N be a submanifold of the presymplectic manifold (M, ω) with inclusion j. The manifold N is called **constraint submanifold**. We define the **symplectic polar**

$$
\begin{aligned}
TN^{\perp} :&= \{Z \in TM \mid \omega|_N(X, Z) = 0\,,\ \forall X \in TN\} \\
&= \{Z \in TM \mid j^*(Z \lrcorner \omega) = 0\}
\end{aligned}
$$

Furthermore, given a subspace S of a Banach space E, we define the **annihilator** of S as

$$S^{\vdash} := \{\beta \in E^* \mid \beta(v) = 0\,,\ \forall v \in S\} \subset E^*$$

Proposition 5.4.1. *If M is reflexive and ω is topologically closed, then*

$$(TN^{\perp})^{\vdash} = TN^{\flat}$$

where we denoted $TN^{\flat} := \mathrm{Im}(\flat|_{TN})$, with $\flat : TM \to T^*M,\ Z \mapsto \omega(Z, \cdot)$, as usual.

Definition 5.4.2. The constraint submanifold N is said to be

1. **isotropic** if $TN \subset TN^\perp$
2. **coisotropic** or **first-class** if $TN^\perp \subset TN$
3. **weakly symplectic** or **second-class** if $TN \cap TN^\perp = \{0\}$
4. **Lagrangian** if $TN = TN^\perp$

Locally, a fist-class constraint submanifold N can be described by the vanishing of functions $f_i : U \subset N \to \mathbb{R}$, s.t. $df_i|_N \cdot W = 0$ for all $W \in TN^\perp$. We call any function f (resp. 1-form γ) on M such that $f|_N = 0$ (resp. $j^*\gamma = 0$) a **constraint function** (resp. **constraint form**). And any function g (resp. 1-form σ) on M such that $dg|_N(W) = 0$ (resp. $\sigma(W)|_N = 0$) for all $W \in TN^\perp$ is called **first class**. Functions, which are not first-class are called **second-class** and a second-class constraint submanifold can locally be described by such.

5.5. The Constraint Algorithm

Let $(\mathcal{P}, \omega)$ be a presymplectic manifold and $H : \mathcal{P} \to \mathbb{R}$ a differentiable function (the Hamiltonian). We investigate when it is possible to solve the canonical equations of motion,

$$i_X\omega = dH$$

or, more generally, when we can solve

$$i_X\omega = \alpha$$

for some closed 1-form α on $\mathcal{P}$ such that the vector field X thus defined gives rise to a well behaved flow. Now, in order for this equation to possess any solution at all, we must of course restrict to the subspace

$$\mathcal{P}_2 := \{m \in \mathcal{P} \mid \alpha \in T_m\mathcal{P}^\flat\}$$

consequently we are led to solve the equation $(i_X\omega - \alpha) \circ j_2 = 0$, where $j_2 : \mathcal{P}_2 \to \mathcal{P}$ is the inclusion. However, the solution to that equation needs to be tangent to $\mathcal{P}_2$, otherwise we will not be able to define the flow of the resulting vector field X. Hence we are led to consider the subspace $\mathcal{P}_3 \subset \mathcal{P}_2$ defined by

$$\mathcal{P}_3 := \{m \in \mathcal{P}_2 \mid \alpha(m) \in T\mathcal{P}_2^\flat\}$$

Finally, it is obvious how this algorithm is to proceed and we generate a string of submanifolds

$$C = \mathcal{P}_k \xrightarrow{j_k} \dots \xrightarrow{j_4} \mathcal{P}_3 \xrightarrow{j_3} \mathcal{P}_2 \xrightarrow{j_2} \mathcal{P} \text{ with } \mathcal{P}_{l+1} := \{m \in \mathcal{P}_l \mid \alpha(m) \in T\mathcal{P}_l^\flat\}$$

Eventually this algorithm terminates in a (possibly zero dimensional) submanifold C, which we will call the **final constraint manifold**.

Starting from this intuitive algorithm, we will now present a computationally more viable method to arrive at C. Using Proposition 5.4.1 one can show

Proposition 5.5.1. *If ω is topologically closed and if $\mathcal{P}$ is reflexive we have*

$$\mathcal{P}_{l+1} := \{m \in \mathcal{P}_l \mid \langle T\mathcal{P}_l^{\perp}, \alpha\rangle(m) = 0\} \text{ with}$$
$$T\mathcal{P}_l^{\perp} := \{W \in T\mathcal{P} \mid k_l^*(i_W\omega) = 0\}$$

with $k_l := j_2 \circ j_3 \circ ... \circ j_l$ and $k_1 := \mathrm{id}|_{\mathcal{P}}$.

The conditions $\langle T\mathcal{P}_l^{\perp}, \alpha\rangle(m) = 0$ are called **secondary-, tertiary-,...., constraints**.

Theorem 5.5.2. *Assume ω is topologically closed and $\mathcal{P}$ reflexive. The equations $(X \lrcorner \omega = \alpha)|_{\mathcal{P}_k}$ possess solutions tangent to $\mathcal{P}_k$ iff $\langle T\mathcal{P}_k^{\perp}, \alpha\rangle = 0$.*

Denote by X_f the **Hamiltonian vector field** associated to $f : \mathcal{P} \to \mathbb{R}$, i.e. satisfying $X_f \lrcorner \omega = df$.

Proposition 5.5.3. *[GM06]*

1. *Let f be a constraint. Then X_f exists along $C \Leftrightarrow T\mathcal{P}^{\perp}[f]|_C = 0$. If it exists, then $X_f \in \mathfrak{X}(C)^{\perp}$.*
2. *Conversely, suppose $(\mathcal{P}, \omega)$ is symplectic. Then $\forall p \in C$, $T_pC^{\perp}$ is spanned by the Hamiltonian vector fields of constraints.*
3. *Let f be a first class constraint. Then X_f exists along C and $X_f \in \mathfrak{X}(C) \cap \mathfrak{X}(C)^{\perp}$.*
4. *Conversely, suppose $(\mathcal{P}, \omega)$ is symplectic. Then $\forall p \in C$, $T_pC \cap T_pC^{\perp}$ is spanned by the Hamiltonian vector fields of first class constraints.*
5. *C first class $\Leftrightarrow$ Every constraint is first class.*
6. *C second class $\Leftrightarrow$ Every **effective**[15] constraint is second class.*

Note in particular point 3.: This led Dirac to conjecture that secondary first class constraints generate all gauge transformations. See [GM06] or [Got83] for a extensive discussion - we will give a brief outline taken from the former.

By **gauge freedom** is meant that a given a set of initial data $(\varphi, \pi) \in C_\lambda$ does not uniquely determine a dynamical trajectory. Now, take the Hamiltonian vector field $X_f \in \mathfrak{X}(C_\lambda) \cap \mathfrak{X}(C_\lambda)^{\perp}$ as predicted by 3. and set $X'_\lambda = X_\lambda + X_f$, where X_λ solves Hamilton's equations. We see that

$$\left(X'_\lambda \lrcorner \omega_\lambda - d(H_{\lambda,\zeta} + f)\right)\Big|_{C_\lambda} = 0$$

Whence, if X_λ is a tangential solution of Hamilton's equations along C_λ with Hamiltonian $H_{\lambda,\zeta}$, then X'_λ is a tangential solution of Hamilton's equations along C_λ with Hamiltonian $H'_{\lambda,\zeta} + f$. But

[15] A constraint, f, is **effective**, provided $df|_C \not\equiv 0$.
(If f is any constraint, then f^2 is first class)

$H'_{\lambda,\zeta}|_{C_\lambda} = H_{\lambda,\zeta}|_{C_\lambda}$. Thus physically, both solutions are indistinguishable, as physically, the region outside of C_λ is dynamically inaccessible.

Though we will not go into constraint theory as presented here, the first step of the algorithm is one of the main motivations for the problem investigated in this work - see section 8.1. The next section mainly serves to introduce the Legendre transformation generalized to the infinite-dim. setting of field theory. The image of this Legendre transformation will turn out to be a presymplectic manifold on which we define the usual Hamiltonian, which in turn brings us back to the tools developed in this section and leads straight to the original work of this thesis.

6. A Hamiltonian formulation of Classical Fields

The following is basically a recapitulation of [GM06], where we felt free to elucidate on parts that were not treated explicitly there. Consequently we only reference third sources and will not highlight small additions by the author. Before introducing the Lagrangian, we will first introduce the Jet bundle - the space the Lagrangian will be defined on.

6.1. The Jet Bundle

See also [Sei12] for a very hands-on introduction to Jet bundles.
Let

$$\pi_{XY} : Y \to X \ , \ X \text{ oriented manifold}$$

be a finite dimensional fiber bundle, the **covariant configuration bundle** and $\dim X = n + 1$, $\dim \pi_{XY}^{-1}(x) = N \ , \ x \in X$. Denote local coordinates

$$\begin{aligned} X \supset U &\to \mathbb{R}^{n+1} & \pi_{XY}^{-1}(x) \supset F &\to \mathbb{R}^N \\ x &\mapsto (x^\mu)_{\mu=0,1,\dots,n} & y &\mapsto (y^A)_{A=1,\dots,N} \end{aligned}$$

The **first jet bundle**, J^1Y, of Y is defined as: Let $\phi_i \ , \ i = 1, 2$, be local sections of Y.

$$\begin{aligned} \phi_1 \sim \phi_2 :\Leftrightarrow & \phi_1(x) = \phi_2(x) \\ & T_x\phi_1 = T_x\phi_2 \text{ (as maps } T_xX \to T_{\phi(x)}Y) \\ J_1Y := \{\text{Local sections of } Y \to X\}/\sim & \end{aligned}$$

J^1Y is naturally identified with the affine bundle over Y defined by

$$\begin{aligned} &\pi_{YJ^1Y} : J^1Y \to Y \\ &\pi_{YJ^1Y}^{-1}(y) = \left\{ \gamma : T_xX \to T_yY \text{ linear} \,\middle|\, T\pi_{XY} \circ \gamma = \mathrm{Id}_{T_xX} \right\} \end{aligned}$$

To see this, let $j^1(\phi)(x)$ be the jet of ϕ at $x \in X$. Clearly, $j^1(\phi)(x)$ depends only on the values of ϕ in some neighborhood of x. Thus, if $\phi(x) = (x, y)$, we can find charts (U, φ) and (F, ψ) around x and y, respectively, and easily observe that $\tilde{\phi} \in j^1(\phi)(x)$ iff

$$T(\psi \circ \phi \circ \varphi^{-1})(\varphi(p)) = T(\psi \circ \tilde{\phi} \circ \varphi^{-1})(\varphi(x))$$

where we dropped the identity 'x-component' of $\phi(x)$. A jet $j^1(\phi)(x)$ thus identified with $T(\phi)(x)$, in local coordinates is an element of $\mathbb{R}^{n+1} \oplus \mathbb{R}^N \oplus L(\mathbb{R}^{n+1}, \mathbb{R}^N)$, where $L(\mathbb{R}^{n+1}, \mathbb{R}^N)$ denotes the space of linear maps between the respective spaces. Consequently, J^1Y itself is locally diffeomorphic to an open set in $\mathbb{R}^{n+1} \oplus \mathbb{R}^N \oplus L(\mathbb{R}^{n+1}, \mathbb{R}^N)$. This construction immediately yields induced coordinates on the fibers of J^1Y, which we denote by $v^A{}_\mu$.

Definition 6.1.1. [Lee12] Let $F : M \to N$ be a smooth submersion, where M and N are smooth manifolds. A vector field V on M is said to be **vertical** if V is everywhere tangent to the fibers of F.

We define the **vertical subbundle**, $VY \subset TY$,

$$\begin{aligned} \pi_{YVY} &: VY \to Y \\ \pi_{YVY}^{-1}(y) &=: V_y Y = \ker(T\pi_{XY}) \end{aligned}$$

For later use, we also point out

$$\gamma \in J^1_y Y \Rightarrow T_y Y = \operatorname{im}\gamma \oplus V_y Y \quad (\gamma \text{ induces a connection on } Y)$$

which can be seen immediately from the local expression above upon keeping the identity-component that was dropped.

Concluding, if we let ϕ be a (local) section of Y, then $T_x\phi \in J^1_{\phi(x)}Y$ and hence we can use local coordinates and define

$$\begin{aligned} j^1\phi : X &\to J^1_{\phi(x)}Y \\ x &\mapsto T_x\phi \\ x^\mu &\mapsto \left(x^\mu, \phi^A(x^\mu), \partial_\nu\phi^A(x^\mu)\right) \end{aligned}$$

which is clearly a section of J^1Y. We call $j^1\phi$ the **first jet prolongation**. Finally, let $\hat{\phi}$ be a section of J^1Y. If $\hat{\phi} = j^1\phi$ is a first jet prolongation of some section ϕ of Y, it is called **holonomic**.

6.2. The Dual Jet Bundle

Definition 6.2.1. [Wik13]: An **affine map** $f : A \to B$ between two affine spaces, A and B, is a map that determines a linear transformation, φ, s.t.

$$\overrightarrow{f(P)f(Q)} = \varphi(\overrightarrow{PQ})$$

E.g. $f : \mathbb{R}^n \to \mathbb{R}^m$ is affine iff it is of the form $f(x) = Ax + c$, A linear and $c \in \mathbb{R}^m$ constant.

The **dual jet bundle**, J^1Y^*, is the vector bundle

$$\begin{aligned} \pi_{YJ^1Y^*} &: J^1Y^* \to Y \\ \pi_{YJ^1Y^*}^{-1}(y_x) &= \left\{\gamma : J^1_yY \to \Lambda^{n+1}_x X \;\middle|\; \gamma \text{ affine mapping}\right\} \end{aligned}$$

where $\Lambda^{n+1}(X)$ denotes the bundle of $(n+1)$-forms on X.

$\gamma \in \pi^{-1}_{YJ^1Y^*}(y_x)$ can be represented in coordinates by the affine mapping $\gamma : v^A{}_\mu \mapsto \left(p + p^\mu{}_A v^A{}_\mu\right) d^{n+1}x$, i.e. by the coordinates $(p, p^\mu{}_A)$.

Let $\Lambda := \Lambda^{n+1} Y$ and denote the fiber of Λ over $y \in Y$ by Λ_y.

$$Z_y := \left\{ z \in \Lambda_y \mid i_v i_w z = 0 \ , \ \forall v, w \in V_y Y \right\}$$

In coordinates

$$\begin{aligned} &Z \ni z = p d^{n+1}x + p^\mu{}_A dy^A \wedge d^n x_\mu \\ &d^n x_\mu := \partial_\mu \lrcorner\, d^{n+1}x \end{aligned}$$

i.e. z can be represented by $(p, p^\mu{}_A)$.

Proposition 6.2.2. *$J^1Y^* \cong Z$,* canonically*, as vector bundles over Y.*

Proof. We will construct the canonical isomorphism. Let $\langle \cdot, \cdot \rangle$ denote the dual pairing and define

$$\begin{aligned} &\Phi : Z \to J^1Y^* \text{ such that} \\ &\langle \Phi(z), \gamma \rangle = \gamma^* z \in \Lambda^{n+1}_x X \end{aligned}$$

where $z \in Z_y$, $\gamma \in J^1_y Y$, $x = \pi_{XY}(y)$. By the rank-nullity theorem we need only show that if $\gamma^* z = 0\ \forall z$, then $\gamma = 0$. $\gamma^* : T^*_y Y \to T^*_x X$ and we are free to choose $z \in \Lambda^{n+1}_y Y$ s.t. $i_{e_i} z \neq 0$ for some basis $\{e_i\}$ of TY_x. $\square$

We can of course do this calculation explicitly in coordinates: Let $\gamma = v^A{}_\mu$, then

$$\begin{aligned} \gamma^* dx^\mu &= dx^\mu \\ \gamma^* dy^A &= v^A{}_\mu dx^\mu \end{aligned}$$

and thus

$$\gamma^* \left(p d^{n+1}x + p^\mu{}_A dy^A \wedge d^n x_\mu \right) = \left(p + p^\mu{}_A v^A{}_\mu \right) d^{n+1}x$$

So we are just equating the coordinates $(x^\mu, y^A, p, p^\mu{}_A)$ of Z and J^1Y^*.

We define the **canonical** $(n+1)$**-form**, Θ_Λ, on Λ $(= \Lambda^{n+1}Y)$, by

$$\Theta_\Lambda(z)(u_1, ..., u_{n+1}) := \left(\pi^*_{Y\Lambda} z \right)(u_1, ..., u_{n+1})$$

where $z \in \Lambda$, $u_1, ..., u_{n+1} \in T_z \Lambda$.

Define the **canonical** $(n+2)$**-form**, Ω_Λ, on Λ by

$$\Omega_\Lambda = -d\Theta_\Lambda$$

Let $i_{\Lambda Z} : Z \to \Lambda$ be the inclusion map. The **canonical** $(n+1)$**-form**, Θ, on Z is defined by

$$\Theta := i_{\Lambda Z}^{*}\Theta_{\Lambda} \tag{6}$$

and the **canonical** $(n+2)$**-form**, Ω, on Z by

$$\Omega := -d\Theta = i_{\Lambda Z}^{*}\Omega_{\Lambda} \tag{7}$$

Or, in coordinates:

$$\Theta = p^{\mu}{}_{A} dy^{A} \wedge d^{n}x_{\mu} + p d^{n+1}x$$
$$\Omega = dy^{A} \wedge dp^{\mu}{}_{A} \wedge d^{n}x_{\mu} - dp \wedge d^{n+1}x$$

Definition 6.2.3. (Z, Ω) is called **multiphase space** or **covariant phase space**. It is an example of a multisymplectic manifold.

Definition 6.2.4. A **multisymplectic manifold**, (M, Ω), is a manifold, M, endowed with a closed k-form, Ω, which is nondegenerate i.t.s.t. $i_v\Omega \neq 0$ for $0 \neq v \in TM$.[16]

Finally, through direct calculation, we observe that similar to the canonical (Liouville-) one-form in symplectic geometry, we have

Proposition 6.2.5. *If σ is a section of π_{XZ} and $\phi = \pi_{YZ} \circ \sigma$, then*

$$\sigma^{*}\Theta = \phi^{*}\sigma$$

where $\phi^{*}\sigma$ means the pull-back by ϕ to X of σ regarded as an $(n+1)$-form on Y along ϕ.

6.3. Lagrangian Dynamics

Let the **Lagrangian density** $\mathcal{L} : J^{1}Y \to \Lambda^{n+1}X$ be given in coordinates by a smooth bundle map $\mathcal{L} = L(x^{\mu}, y^{A}, v^{A}{}_{\mu})d^{n+1}x$ over X.[17]
In classical mechanics, the Legendre transform is defined through the fiber derivative of the tangent space of the configuration space:

Definition 6.3.1. [BSF06]
Let M be a manifold (configuration space) and $L : TM \to \mathbb{R}$ a smooth function. The **fiber derivative** of L is a strong bundle map $\mathbb{F}L : TM \to T^{*}M$ defined by $\mathbb{F}L(v_p)(w_p) := dL(v_p)(w_p)$ for all $v_p, w_p \in T_pM$ and $p \in M$, where w_p on the right hand side is considered as an element in $T_{v_p}(T_pM) \subset T^{2}M$.

[16] See http://ncatlab.org/nlab/show/multisymplectic+geometry for more information on multisymplectic manifolds - there seems to be no general agreement on a proper abstract definition of a multisymplectic manifold yet.
[17] Recall that we assume X to be oriented

Of course we can express this as $\mathbb{F}L(v_p)(w_p) = \frac{d}{dt}\Big|_{t=0} L(p, v_p + tw_p)$. Now, as J^1Y is a vector bundle and in the above we haven't made use of the 'tangent-structure' of the tangent bundle TM, we may naturally generalize:

The **covariant Legendre transformation** for $\mathcal{L}$ is the fiber derivative

$$\mathbb{F}\mathcal{L} : J^1Y \to J^1Y^* \cong Z \tag{8}$$

$$\langle \mathbb{F}\mathcal{L}(\gamma), \gamma' \rangle := \mathcal{L}(\gamma) + \frac{d}{d\epsilon}\mathcal{L}(\gamma + \epsilon(\gamma' - \gamma))\Big|_{\epsilon=0} \tag{9}$$

where $\gamma, \gamma' \in J^1_y Y$. In coordinates:

$$p^\mu{}_A = \frac{\partial L}{\partial v^A{}_\mu} \ , \ p = L - \frac{\partial L}{\partial v^A{}_\mu} v^A{}_\mu \tag{10}$$

where we call $p^\mu{}_A$ the **multimomenta** and p the **covariant Hamiltonian**. To verify the coordinate expression, suppose $\gamma = v^A{}_\mu$ and $\gamma' = w^A{}_\mu$. Then the right hand side of Equation 9 reads

$$\left(L(\gamma) + \frac{\partial L}{\partial v^A{}_\mu}(w^A{}_\mu - v^A{}_\mu) \right) d^{n+1}x = \left(p + p^\mu{}_A w^A{}_\mu \right) d^{n+1}x$$

Definition 6.3.2. The **Cartan form** is the $(n+1)$-form, $\Theta_\mathcal{L}$, on J^1Y

$$\Theta_\mathcal{L} : = \mathbb{F}\mathcal{L}^*\Theta \tag{11}$$

$$= \frac{\partial L}{\partial v^A_\mu} dy^A \wedge d^n x_\mu + \left(L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu \right) d^{n+1}x \tag{12}$$

We also define:

$$\begin{aligned} \Omega_\mathcal{L} &: = -d\Theta_\mathcal{L} = \mathbb{F}\mathcal{L}^*\Omega \\ &= dy^A \wedge d\left(\frac{\partial L}{\partial v^A_\mu} \right) \wedge d^n x_\mu - d\left(L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu \right) \wedge d^{n+1}x \end{aligned}$$

The construction of $\Theta_\mathcal{L}$ might seem artificial at this point. There is, however, a more natural way to arrive at $\Theta_\mathcal{L}$: A direct calculation shows that

Corollary 6.3.3.

$$\mathcal{L}(j^1\phi) = (j^1\phi)^*\Theta_\mathcal{L}$$

When we vary the Lagrangian in the following, $\Theta_\mathcal{L}$ appears naturally in the boundary term. [MPS98] show that $\Theta_\mathcal{L}$ is the unique one-form 'generating' a more general class of boundary terms.

Definition 6.3.4. Let ϕ be a section of Y. A **variation** of ϕ is a curve $\phi_\lambda = \eta_\lambda \circ \phi$, where η_λ is the flow of a vertical vector field V on Y which is compactly supported in X. One says that ϕ is a **stationary point** of $L = \int \mathcal{L}$, if

$$\frac{d}{d\lambda}\left[\int_X \mathcal{L}(j^1\phi_\lambda)\right]\bigg|_{\lambda=0} = 0$$

for all variations ϕ_λ of ϕ.

The **Euler-Lagrange** equations in coordinates read

$$\frac{\delta L}{\delta \phi^A} = \frac{\partial L}{\partial y^A}(j^1\phi) - \frac{\partial}{\partial x^\mu}\left(\frac{\partial L}{\partial {v^A}_\mu}(j^1\phi)\right) = 0 \tag{13}$$

for a (local) section, ϕ, of Y.

Theorem 6.3.5. *Let ϕ be a section of the fiber bundle $\pi_{XY} : Y \to X$. The following are equivalent:*

1. *ϕ is a stationary point of the* ***action integral***

$$\int_X \mathcal{L}(j^1\phi)$$

2. *The Euler-Lagr. equ. hold in coordinates*

3. *For any vector field W on J^1Y:*

$$(j^1\phi)^*(W \lrcorner \Omega_{\mathcal{L}}) = 0$$

6.3.0.3. Remark:

1. Since the E-L-Equ. is a PDE, this theorem holds if the section ϕ is only a local section, $\phi : U \to Y_U$, where $U \subset X$, open. The integral is then only taken over U.

2. Because the Euler-Lagr. equations are equivalent to the *intrinsic* conditions, 1. and 3., above, they, too, must be intrinsic. One can, in fact, write the Euler-Lagr. derivative, $\frac{\delta L}{\delta \phi^A}$, intrinsically in a direct way - see [MPS98] .

Corollary 6.3.6. *For any vector field, W, on J^1Y we have*

$$(j^1\phi)^*(W \lrcorner \Omega_{\mathcal{L}}) = -W^A \frac{\delta L}{\delta \phi^A} d^{n+1}x$$

6.4. Cauchy surfaces and Space of Sections

Now, that we have covered the Lagrangian formalism, we will move towards the Hamiltonian formalism. For this, though, we need to define the space of initial conditions on which the Hamiltonian formalism takes place. Throughout, let Σ denote a compact, oriented, connected, boundaryless n-manifold. For a detailed exposition of spaces of sections the reader is referred to [Pal68].

Definition 6.4.1. For $\tau \in \mathrm{Emb}(\Sigma, X) \cap \{\text{fixed char. class}\}$ [18], let $\Sigma_\tau = \tau(\Sigma)$. We view Σ as a **reference** or **model Cauchy surface**. (Σ_τ will eventually be a Cauchy surface for the dynamics)

Let $\pi_{XK} : K \to X$ be a fiber bundle over X. Then we denote the space of sections of K by $\mathcal{K}$ and $K_\tau := K|_{\Sigma_\tau \subset X}$, with $\mathcal{K}_\tau$ denoting the space of sections of K_τ. Furthermore, we define the bundle

$$\mathcal{K}^\Sigma := \{ \mathcal{K}_\tau \mid \tau \in \mathrm{Emb}(\Sigma, X)\} \to \mathrm{Emb}(\Sigma, X) \tag{14}$$

The **tangent space** to $\mathcal{K}$ at a point σ is defined by[19]

$$T_\sigma \mathcal{K} = \{W : X \to VK \mid W \text{ covers } \sigma\}$$

where VK denotes the vertical tangent bundle of K. Similarly, the **smooth cotangent space**

$$T^*_\sigma \mathcal{K} = \{\pi : X \to L(VK, \Lambda^{n+1}X) \mid \pi \text{ covers } \sigma\}$$

where $L(VK, \Lambda^{n+1}X)$ is the vector bundle over K defined by

$$\begin{aligned}
&\pi_{KL} : L(VK, \Lambda^{n+1}X) \to K \\
&\pi^{-1}_{KL}(k) := \{F : V_k K \to \Lambda^{n+1}_x X \mid F \text{ linear map}\},\ k \in K_x
\end{aligned}$$

For $\pi \in T^*_\sigma \mathcal{K}$ and $V \in T_\sigma \mathcal{K}$, we naturally define

$$\langle \pi, V \rangle = \int_X \pi(V)$$

Similarly we define the (co-) tangent space of $\mathcal{K}_\tau$ by replacing X with Σ_τ and K with K_τ.

[18] $\mathrm{Emb}(\Sigma, X)$...Smooth embeddings $\Sigma \to X$. "Fixed char. class" meaning either char. or non-char. w.r.t. the Euler-Lagrange equations. [GM06] require the surfaces to be spacelike, i.e. non-char. in physical applications. Refer to section 8.1 for an explanation why the author chose to change this.

[19] The definition of a tangent bundle for an infinite-dim. manifold is not unambiguous. Following [AMR88]; the definition of the tangent bundle to an (infinite) dimensional manifold: Let M be a smooth manifold modeled on a Banach space E with smooth atlas $\{(U_\alpha, \phi_\alpha) \mid \alpha \in A\}$. Consider the collection of triples (α, p, v), where $\alpha \in A$, $p \in U_\alpha$ and $v \in E$ and define

$$(\alpha, p, v) \sim (\beta, q, w) :\Leftrightarrow p = q \text{ and } w = D(\phi_\beta \circ \phi_\alpha^{-1})(\phi_\alpha(p))v$$

the tangent bundle of M is the set of equivalence classes.

Definition 6.4.2. If ξ_K is any π_{XK}-projectable vector field on K, define the **Lie-derivative** of $\sigma \in \mathcal{K}$ along ξ_K by

$$\pounds_{\xi_K}\sigma := T\sigma \circ \xi_X - \xi_K \circ \sigma \ \in T_\sigma \mathcal{K}$$
$$\left(\pounds_{\xi_K}\sigma\right)^A = \sigma^A{}_{,\mu}\xi^\mu - \xi^A \circ \sigma$$

where, as usual, we denote $\xi_K = (\xi^\mu, \xi^A)$.

Note that $-\pounds_{\xi_K}\sigma$ is exactly the vertical component of $\xi_K \circ \sigma$ ($T\pi_{XK}(\pounds_{\zeta_K}\sigma) = \zeta_X - \zeta_X = 0$).

Definition 6.4.3. If $f : \mathcal{K} \to C^\infty(X)$, then define the "formal" partial derivative

$$D_\mu f : \mathcal{K} \to C^\infty(X)$$
$$D_\mu f(\sigma) := f(\sigma)_{,\mu}$$

6.5. Canonical Forms on $T^*\mathcal{Y}_\tau$ and $\mathcal{Z}_\tau$

The following construction should be compared to the construction of the canonical symplectic form in section 5.3.4. In the notation introduced above:

Definition 6.5.1. For $\tau \in \mathrm{Emb}(\Sigma, X)$, we call $\mathcal{Y}_\tau$ the **τ-configuration space** and $T^*\mathcal{Y}_\tau$ the **τ-phase space**.

Definition 6.5.2.

$$\theta_\tau(\varphi,\pi)(V) := \int_{\Sigma_\tau} \pi\left(T\pi_{\mathcal{Y}_\tau,T^*\mathcal{Y}_\tau} V\right) \ \left(= \int_\Sigma \tau^*(\ldots)\right)$$

where $(\varphi,\pi) \in T^*\mathcal{Y}_\tau$, $V \in T_{(\varphi,\pi)}T^*\mathcal{Y}_\tau$ and $\pi_{\mathcal{Y}_\tau,T^*\mathcal{Y}_\tau} : T^*\mathcal{Y}_\tau \to \mathcal{Y}_\tau$. Moreover, we define the symplectic form

$$\omega_\tau := -d\theta_\tau$$

Making $(T^*\mathcal{Y}_\tau, \omega_\tau)$ a symplectic manifold.

Next, we will give a coordinate representation. The forms depend on the hypersurface Σ_τ. So naturally, we will start by choosing special coordinates of X that already conveniently describe the hypersurface.[20]

Definition 6.5.3. A chart $(x^0, x^1, \ldots)$ on X is **adapted** to τ if Σ_τ is locally a level set of x^0.

[20]Note: Easily choosing adapted coordinates works only if the Lagrangian is somewhat parametrization invariant (ref. to section 8.1). Practically, one works in adapted coordinates from the start.

By definition of a submanifold, such coordinates always exist. For an adapted chart, $(x^0, x^1, ...)$, on X, $\pi \in T^*_\varphi \mathcal{Y}_\tau$, regarded as a map $\pi : \Sigma_\tau \to L(VY_\tau, \Lambda^n \Sigma_\tau)$, is expressible as

$$\pi = \pi_A dy^A \otimes d^n x_0 \tag{15}$$

Thus, for an adapted chart on X:

$$\theta_\tau(\varphi, \pi) = \int_{\Sigma_\tau} \pi_A d\varphi^A \otimes d^n x_0$$
$$\omega_\tau(\varphi, \pi) = \int_{\Sigma_\tau} \left(d\varphi^A \wedge d\pi_A\right) \otimes d^n x_0$$

e.g. for $T_{(\varphi,\pi)}(T^*\mathcal{Y}_\tau) \ni V = (V^A, W_A)$ (in adapted coordinates):

$$\theta_\tau(\varphi, \pi)(V) = \int_{\Sigma_\tau} \pi_A V^A d^n x_0$$

The rest of the preliminary part of this work will be concerned with finding a Hamiltonian system $(\mathcal{P}_\tau, \omega_\tau, H) \subset (T^*\mathcal{Y}_\tau, \omega_\tau)$, such that its solutions correspond to solutions of the Euler-Lagrange equations Equation 13. For this, we will need to relate $T^*\mathcal{Y}_\tau$ to $J^1 Y$. As a first step, we relate $T^*\mathcal{Y}_\tau$ to $\mathcal{Z}_\tau$.

Definition 6.5.4. We define the **canonical one-form**, Θ_τ, on $\mathcal{Z}_\tau$, by

$$\Theta_\tau(\sigma)(V) := \int_{\Sigma_\tau} \sigma^* \left(i_V \Theta\right)$$

where $\sigma \in \mathcal{Z}_\tau$, $V \in T_\sigma \mathcal{Z}_\tau$ and Θ the canonical one-form on Z (cp. Equation 6).
The **canonical two-form**, Ω_τ, on $\mathcal{Z}_\tau$, is

$$\Omega_\tau := -d\Theta_\tau$$

Lemma 6.5.5. *At $\sigma \in \mathcal{Z}_\tau$ and with Ω the canonical $(n+2)$-form on Z (Equation 7)*

$$\Omega_\tau(\sigma)(V, W) = \int_{\Sigma_\tau} \sigma^*(i_W i_V \Omega)$$

6.6. Reduction of $\mathcal{Z}_\tau$ to $T^*\mathcal{Y}_\tau$

We define

$$R_\tau : \mathcal{Z}_\tau \to T^*\mathcal{Y}_\tau \tag{16}$$
$$\langle R_\tau(\sigma), V \rangle := \int_{\Sigma_\tau} \varphi^*(i_V \sigma) \tag{17}$$

where $\varphi = \pi_{YZ} \circ \sigma$ and $V \in T_\varphi \mathcal{Y}_\tau$. Fixing σ, we may also interpret $R_\tau(\sigma)$ as a map $\Sigma_\tau \to L(VY_\tau, \Lambda^n \Sigma_\tau)$ which covers φ, given by

$$\langle R_\tau(\sigma)(x), v\rangle = \varphi^* i_v \sigma(x)$$

where $v \in V_{\varphi(x)} Y_\tau$.
In adapted coordinates:

$$\mathcal{Z}_\tau \ni \sigma = (p_A{}^\mu \circ \sigma)\, dy^A \wedge d^n x_\mu + (p \circ \sigma) d^{n+1} x \tag{18}$$

$$\rightsquigarrow R_\tau(\sigma) = \left(p_A{}^0 \circ \sigma\right) dy^A \otimes d^n x_0 \tag{19}$$

Since one can always choose adapted coordinates, we have

Corollary 6.6.1. *R_τ is a surjective submersion with*

$$\ker R_\tau = \left\{\sigma \in \mathcal{Z}_\tau \mid p_A{}^0 \circ \sigma = 0\right\}$$

Comparing Equation 15 with Equation 19, we see that the **instantaneous momenta**, π_A, correspond to the temporal components of the multimomenta $p_A{}^\mu$.

Proposition 6.6.2.

$$R_\tau^* \theta_\tau = \Theta_\tau$$

Proof. Let $V \in T_\sigma \mathcal{Z}_\tau$, then

$$\begin{aligned}\langle (R_\tau^* \theta_\tau)(\sigma), V\rangle &= \langle \theta_\tau \left(R_\tau(\sigma)\right), TR_\tau V\rangle \\ &= \langle R_\tau(\sigma), T\pi_{\mathcal{Y}_\tau, T^*\mathcal{Y}_\tau} V\rangle\end{aligned}$$

Using the definition of R_τ from above and noting that $\pi_{YZ}^* \sigma = \Theta \circ \sigma$ yields the claim. □

Corollary 6.6.3. *1.* $R_\tau^* \omega_\tau = \Omega_\tau$

2. $\ker T_\sigma R_\tau = \ker \Omega_\tau(\sigma)$

3. The induced quotient map, $\mathcal{Z}_\tau / \ker R_\tau = \mathcal{Z} / \ker \Omega_\tau \to T^* \mathcal{Y}_\tau$ is a symplectic diffeomorphism.

6.7. Initial Value Analysis of Field Theories

Now, that we have discussed the "space of initial values", we are of course interested in "time evolution". For this, we need to choose a direction of time and equal-time-hypersurfaces.

Definition 6.7.1. A **slicing** of an $(n+1)$-dim. spacetime, X, consists of an n-dim. mfd. Σ (the **reference Cauchy surface**) and a diffeomorphism

$$\mathfrak{s}_X : \Sigma \times \mathbb{R} \to X$$

For $\lambda \in \mathbb{R}$,

$$\begin{aligned} \Sigma_\lambda &:= \mathfrak{s}_X(\Sigma \times \{\lambda\}) \\ \tau_\lambda &: \Sigma \to \Sigma_\lambda \subset X \\ \tau_\lambda(x) &:= \mathfrak{s}_X(x, \lambda) \end{aligned}$$

Definition 6.7.2. The **generator**, ζ_X, of $\mathfrak{s}_X$ is the vector field on X defined by

$$\frac{\partial}{\partial \lambda}\mathfrak{s}_X(x, \lambda) =: \zeta_X\left(\mathfrak{s}_X(x, \lambda)\right)$$

Definition 6.7.3. Given a bundle $K \to X$ and a slicing $\mathfrak{s}_X$ of X, a **compatible slicing** of K is a bundle $K_\Sigma \to \Sigma$ and a bundle diffeomorphism $\mathfrak{s}_K : K_\Sigma \times \mathbb{R} \to K$, s.t.

$$\begin{tikzcd} K_\Sigma \times \mathbb{R} \arrow[r, "\mathfrak{s}_K"] \arrow[d, "\mathrm{nat}"'] \arrow[dr, dashed] & K \arrow[d, "\mathrm{nat}"] \\ \Sigma \times \mathbb{R} \arrow[r, "\mathfrak{s}_X"'] & X \end{tikzcd}$$

For $\lambda \in \mathbb{R}$,

$$\begin{aligned} K_\lambda &:= \mathfrak{s}_K(K_\Sigma \times \{\lambda\}) \\ \mathfrak{s}_\lambda &: K_\Sigma \to K_\lambda \subset K \\ \mathfrak{s}_\lambda(k) &:= \mathfrak{s}_K(k, \lambda) \\ \zeta_K &:= T\mathfrak{s}_K \frac{\partial}{\partial \lambda} \end{aligned}$$

Corollary 6.7.4. *Every compatible slicing, $(\mathfrak{s}_K, \mathfrak{s}_X)$ of $K \to X$ defines a 1-param. group of bundle automorphisms: The flow, f_λ, of the generating vector field ζ_K, which is given by*

$$f_\lambda(k) = \mathfrak{s}_K\left(\mathfrak{s}_K^{-1}(k) + \lambda\right)$$

"$+\lambda$"...addition of λ to second factor of $K_\Sigma \times \mathbb{R}$.

Recall the definition of $\mathcal{K}^\Sigma$ from *Equation* 14. We are of course mainly interested in a curve of embeddings $\lambda \mapsto \tau_\lambda$, for $\lambda \in \mathbb{R}$ resembling time (and not the space of all embeddings). We define

$$\mathcal{K}^\tau := \bigcup_{\lambda \in \mathbb{R}} \mathcal{K}_\lambda \subset \mathcal{K}^\Sigma$$

where we have dropped τ from the notation on the r.h.s..

Corollary 6.7.5. *The slicing* $\mathfrak{s}_K : K_\Sigma \times \mathbb{R} \to K$ *induces a trivialization*

$$\mathfrak{s}_{\mathcal{K}} : \mathcal{K}_\Sigma \times \mathbb{R} \to \mathcal{K}^\tau$$
$$\mathfrak{s}_{\mathcal{K}}(\sigma_\Sigma, \lambda) := \mathfrak{s}_\lambda \circ \sigma_\Sigma \circ \tau_\lambda^{-1}$$

And, as above, we again define the generator $\zeta_{\mathcal{K}} := T\mathfrak{s}_{\mathcal{K}} \frac{\partial}{\partial \lambda}$. Using the definition of $s_{\mathcal{K}}$, we observe

$$\zeta_{\mathcal{K}}(\sigma) = \zeta_K \circ \sigma$$

6.7.0.4. Remarks

1. Slicings of the configuration bundle $Y \to X$ naturally induce slicings of certain bundles over it. E.g.:
   a) If ζ_Y generates $\mathfrak{s}_Y$, then $\mathfrak{s}_Z$ is generated by the canonical lift, ζ_Z, of ζ_Y to Z. That is if $\eta_Y : Y \to Y$ is a π_{XY}-bundle automorphism, its **canonical lift** $\eta_Z : Z \to Z$ is defined by $\eta_Z(z) = (\eta_Y^{-1})^* z$. If V is a vector field on Y whose flow η_λ maps fibers of π_{XY} to fibers, its canonical lift to Z is the vector field V_Z that generates the canonical lift of this flow to Z ($\rightsquigarrow \pounds_{\zeta_Z}\Theta = 0$).
   b) A slicing of J^1Y is generated by the jet prolongation $\zeta_{J^1Y} = j^1\zeta_Y$ of ζ_Y to J^1Y.[21]

2. By a theorem due to Geroch:[22]
   If X is a globally hyperbolic spacetime, then $X \cong \Sigma \times \mathbb{R}$.

3. Sometimes one can also allow for more general curves of embeddings that are not slicings. This need arises for instance when working with characteristic initial value problems, where hypersurfaces need to intersect in order to have a well-posed initial value problem ($\mathfrak{s}$ is not a diffeomorphism in this case)- see [GM06] for references.

For future reference we note that there are two special classes of slicings:

Definition 6.7.6. If a slicing is induced by a 1-param. subgroup of the gauge group G of the theory, it is called a **G-slicing**.
For a given field theory, a slicing, $\mathfrak{s}_Y$, is **Lagrangian** if the Lagrangian density $\mathcal{L}$ is equivariant w.r.t. the 1-param. groups of automorphisms associated to the induced slicings of J^1Y and $\Lambda^{n+1}X$. I.e., let f_λ be the flow of ζ_Y, so that $j^1 f_\lambda$ is the flow of ζ_{J^1Y}, then **equivariance** means

$$\mathcal{L}\left(j^1 f_\lambda(\gamma)\right) = \left(h_\lambda^{-1}\right)^* \mathcal{L}(\gamma)$$

where h_λ is the flow of ζ_X.

[21] To prolong ζ_Y to J^1Y, let f_λ be the flow of ζ_Y. We know what $j^1 f_\lambda$ is and define ζ_{J^1Y} to be its generator.
[22] Smoothness, i.e. that the spaces are actually diffeomorphic, was proven rather recently by Bernal and Sanchez (2005, *arXiv :gr-gc/0404084*)

Now, in principle slicings can be chosen arbitrarily, not necessarily according to a given *a priori* rule. Thus they can be chosen as to implement certain "gauge conditions" on the fields. Most prominently, one can impose the Coulomb gauge condition $\nabla \cdot A = 0$ in Maxwell's theory by adding the corresponding equations (restrictions) for the slicing to equations for the fields.
One may also use 'adaptive slicings' for numerical calculations to improve accuracy.

Thus one is lead to consider the initial value problem for slicings and its interplay with the dynamics of the field. While we will not investigate the former, it turns out that the instantaneous formalism only needs the Cauchy data of the slicing, whence we make the following definition

Definition 6.7.7. An **infinitesimal slicing** of a spacetime X consists of a Cauchy surface, Σ_τ, along with a spacetime vector field, ζ_X, defined over Σ_τ, which is everywhere transverse to Σ_τ.
An **infinitesimal slicing** of a bundle $K \to X$ consists of K_τ along with a vector field ζ_K,...
The infinitesimal slicings (Σ_τ, ζ_X) and (K_τ, ζ_K) are called **compatible** if ζ_K projects to ζ_X.

Fix an infinitesimal slicing $(Y_\tau, \zeta := \zeta_Y)$ of Y and denote [23]

$$\varphi := \phi|_{\Sigma_\tau} \qquad\qquad \dot\varphi := \pounds_\zeta \phi|_{\Sigma_\tau} \tag{20}$$
$$\varphi^A = \phi^A|_{\Sigma_\tau} \qquad\qquad \dot\varphi^A = \left(\zeta^\mu \phi^A{}_{,\mu} - \zeta^A \circ \phi\right)\Big|_{\Sigma_\tau} \tag{21}$$

We define an affine bundle map

$$\beta_\zeta : (J^1 Y)_\tau \to J^1(Y_\tau) \times VY_\tau$$
$$\beta_\zeta\left(j^1\phi(x)\right) := \left(j^1\varphi(x), \dot\varphi(x)\right)\ ,\ x \in \Sigma_\tau$$
$$\beta_\zeta\left(x^i, y^A, v^A{}_\mu\right) = \left(x^i, y^A, v^A{}_j, \dot y^A\right) \ \text{ (in adapted coordinates with } \zeta = \zeta_0\partial_0)$$

If $\zeta|_{Y_\tau} = \frac{\partial}{\partial x_0}$, then $\dot y^A = v^A{}_0$. Thus (since one can always choose adapted coordinates)

Proposition 6.7.8. *If ζ_X is transverse to Σ_τ, then β_ζ is an isomorphism.*

If ζ_X is transverse to Σ_τ, the bundle isomorphism β_ζ is called **jet decomposition map** and its inverse the **jet reconstruction map**.

Corollary 6.7.9. *β_ζ induces an isomorphism*

$$\left(j^1\mathcal{Y}\right)_\tau := \left\{ j^1\phi\Big|_{\Sigma_\tau} \,\middle|\, \phi \in \Gamma_{\text{loc}}(Y)\right\} \cong T\mathcal{Y}_\tau$$

Proof. Denoting by $i_\tau : \Sigma_\tau \to X$ the inclusion, $\beta_\zeta(j^1\phi \circ i_\tau) = (j^1\varphi, \dot\varphi) \in J^1(Y_\tau) \times T_\varphi\mathcal{Y}_\tau$ as $\dot\varphi : \Sigma_\tau \to VY_\tau$ covers φ. The claim follows from the previous proposition. □

[23] cp. Definition 6.4.2

6.8. The Instantaneous Legendre Transform

The Lagrangian is a top-degree form on the oriented manifold X. From differential geometry we know that an oriented manifold features a special class of top-degree forms, which induce the orientation. To show that the definition of the instantaneous Lagrangian hereafter is natural, we quote the following proposition

Proposition 6.8.1. *[Lee12] Suppose M is an oriented smooth n-manifold with or without boundary, S is an immersed hypersurface with or without boundary in M, and N is a vector field along S that is nowhere tangent to S. Then S has a unique orientation such that for each $p \in S$, $(E_1, ..., E_{n-1})$ is an oriented basis for T_pS if and only if $(N_p, E_1, ..., E_{n-1})$ is an oriented basis for T_pM. If* dVol *is an orientation form for M, then $i_S^*(N \lrcorner \, \mathrm{dVol})$ is an orientation form for S w.r.t. this orientation, where $i_S : S \to M$ is inclusion.*

Using the jet reconstruction map, we may decompose the Lagrangian as follows: Define

$$
\begin{aligned}
&\mathcal{L}_{\tau,\zeta} : J^1(Y_\tau) \times VY_\tau \to \Lambda^n\Sigma_\tau \\
&\mathcal{L}_{\tau,\zeta}\left(j^1\varphi(x), \dot{\varphi}(x)\right) := i_\tau^*\left(\zeta_X \lrcorner \mathcal{L}\left(j^1\phi(x)\right)\right)
\end{aligned}
$$

where $j^1\phi \circ i_\tau$ is the reconstruction of $(j^1\varphi, \dot{\varphi})$.

Definition 6.8.2. The **instantaneous Lagrangian**, $L_{\tau,\zeta}$, is defined as

$$
\begin{aligned}
L_{\tau,\zeta} &: T\mathcal{Y}_\tau \to \mathbb{R} \\
L_{\tau,\zeta}(\varphi, \dot{\varphi}) &= \int_{\Sigma_\tau} \mathcal{L}_{\tau,\zeta}\left(j^1\varphi, \dot{\varphi}\right) \\
&= \int_{\Sigma_\tau} L\left(j^1\varphi, \dot{\varphi}\right)\zeta^0 d^n x_0 \text{ (in adapted coordinates with } \zeta = \zeta_0\partial_0)
\end{aligned}
$$

and the **instantaneous Legendre transform** as usual through the fiber derivative

$$
\begin{aligned}
&\mathbb{F}L_{\tau,\zeta} : T\mathcal{Y}_\tau \to T^*\mathcal{Y}_\tau \\
&\mathbb{F}L_{\tau,\zeta}(v)w := \left.\frac{d}{dt}\right|_{t=0} L_{\tau,\zeta}(v + tw)\ ,\ v, w \in T_\varphi\mathcal{Y}_\tau
\end{aligned}
$$

Note that $\mathbb{F}L_{\tau,\zeta}$ is fiber-preserving. In a local chart, the fiber derivative is given by

$$
\mathbb{F}L_{\tau,\zeta}(\varphi, \dot{\varphi}) = \left(\varphi, D_2L_{\tau,\zeta}(\varphi, \dot{\varphi})\right)
$$

Where D_2 denotes the partial derivative w.r.t. the second argument. In adapted coordinates: $\pi = \pi_A dy^A \otimes d^n x_0$ with $\pi_A = \frac{\partial \mathcal{L}_{\tau,\zeta}}{\partial \dot{y}^A}$.[24] We call

$$
\mathcal{P}_{\tau,\zeta} := \operatorname{im} \mathbb{F}L_{\tau,\zeta} \subset T^*\mathcal{Y}_\tau
$$

the **instantaneous** or **τ-primary constraint set**.

[24]Here it was used that $\mathcal{L}$ is of first order

Assumption 6.8.3. ***(Almost regularity)***
Assume that $\mathcal{P}_{\tau,\zeta}$ is a smooth, closed, submanifold of $T^*\mathcal{Y}_\tau$ and that $\mathbb{F}L_{\tau,\zeta}$ is a submersion onto its images with connected fibers.

Notice that we already defined a *covariant* Legendre transformation on J^1Y in Equation 9. It is natural to ask in which sense it 'agrees' with the one operating on $T\mathcal{Y}_\tau$. A direct calculation in adapted coordinates shows

Proposition 6.8.4. *Assume ζ_X is transverse to Σ_τ. Then*

$$\begin{tikzcd} \left(j^1\mathcal{Y}\right)_\tau \arrow[r, "\mathbb{F}\mathcal{L}"] \arrow[d, "\beta_\zeta"'] \arrow[dr, dashed] & \mathcal{Z}_\tau \arrow[d, "R_\tau"] \\ T\mathcal{Y}_\tau \arrow[r, "\mathbb{F}L_{\tau,\zeta}"'] & T^*\mathcal{Y}_\tau \end{tikzcd}$$

Where $R_\tau : (\mathcal{Z}_\tau, \Omega_\tau) \to (T^*\mathcal{Y}_\tau, \omega_\tau)$ was defined in Equation 16.

Hence it makes sense to call the following the **covariant primary constraint set**

$$N := \mathbb{F}\mathcal{L}\left(J^1Y\right) \subset Z$$

and, by abuse of notation,

$$\mathcal{N}_\tau := \mathbb{F}\mathcal{L}\left((j^1\mathcal{Y})_\tau\right) \subset \mathcal{Z}_\tau \tag{22}$$

As direct consequence of Corollary 6.7.9 and Proposition 6.8.4 we obtain

Corollary 6.8.5. *If ζ_X is transverse to Σ_τ, then*

$$R_\tau(\mathcal{N}_\tau) = \mathcal{P}_{\tau,\zeta}$$

In particular, $\mathcal{P}_{\tau,\zeta} =: \mathcal{P}_\tau$ is independent of ζ.

Denote $\omega_{\mathcal{P}_\tau} := i^*\omega_\tau$, $i : \mathcal{P}_\tau \to T^*\mathcal{Y}_\tau$ inclusion. $(\mathcal{P}_\tau, \omega_{\mathcal{P}_\tau})$ will in general be only presymplectic. However, the fact that $\mathbb{F}L_{\tau,\zeta}$ is fiber preserving together with Assumption 6.8.3 imply that

Corollary 6.8.6. $\ker \omega_{\mathcal{P}_\tau}$ *is a regular distribution on $\mathcal{P}_\tau$ (i.t.s.t. it defines a subbundle of $T\mathcal{P}_\tau$).*

For this work, all concepts needed to understand the results have been covered at this point. However, without having defined the Hamiltonian system, there is of course no motivation for the results.

6.9. Hamiltonian Dynamics

The **instantaneous Hamiltonian** is defined

$$
\begin{aligned}
H_{\tau,\zeta} &: \mathcal{P}_\tau \to \mathbb{R} \\
H_{\tau,\zeta}(\varphi,\pi) &:= \langle \pi, \dot{\varphi} \rangle - L_{\tau,\zeta}(\varphi,\dot{\varphi})
\end{aligned}
$$

Of course we want to investigate the relationship between the covariant and instantaneous formalism, with ultimate goal of showing in which way solutions of the Euler-Lagrange equations relate to solutions of the Hamilton equations by the end of this section. Obviously, R_τ will be of central importance again:

Definition 6.9.1. A **holonomic lift**, σ, of $(\varphi,\pi) \in \mathcal{P}_\tau$ is any element $\sigma \in R_\tau^{-1}\{(\varphi,\pi)\} \cap \mathcal{N}_\tau$. They always exist by virtue of Proposition 6.8.4.

A straight forward calculation in adapted coordinates shows

Proposition 6.9.2. *Let $(\varphi,\pi) \in \mathcal{P}_\tau$. Then for any holonomic lift, σ, of (φ,π),*

$$
H_{\tau,\zeta}(\varphi,\pi) = -\int_{\Sigma_\tau} \sigma^* \left(i_{\zeta_Z} \Theta \right) =: \int_{\Sigma_\tau} \mathfrak{H}_{\tau,\zeta}(\varphi,\pi)
$$

where ζ_Z is the canonical lift of ζ to Z (cp. Remark 6.7.0.4).

Now, fix compatible (Lagrangian) slicings, $\mathfrak{s}_Y$ and $\mathfrak{s}_X$ of Y and X with generating vector fields ζ and ζ_X, respectively. As before, let τ be the curve of embeddings, i.e.

$$
\begin{aligned}
\tau &: \mathbb{R} \to \mathrm{Emb}(\Sigma, X) \\
\tau(\lambda)(x) &:= \mathfrak{s}_X(x,\lambda)
\end{aligned}
$$

and denote

$$
\mathcal{P}^\tau := \cup_{\lambda \in \mathbb{R}} \mathcal{P}_{\tau(\lambda)}
$$

We view the $(n+1)$-evolution of the fields as being given by a curve, c, such that[25]

$$
\begin{tikzcd}
\mathbb{R} \arrow[r, "c\,:\,\lambda \mapsto (\varphi(\lambda),\pi(\lambda))"] \arrow[dr, "\tau"'] & \mathcal{P}^\tau \arrow[d, dashed, "\mathrm{proj}"] \\
& \mathrm{Emb}(\Sigma, X)
\end{tikzcd}
$$

[25] Dashed arrows mean that the diagram commutes.

But this leads us to the question on how exactly we can take the structures we have defined on $T^*\mathcal{Y}_{\tau(\lambda)}$ and $\mathcal{P}_{\tau(\lambda)}$ and extend them to the spaces $T^*\mathcal{Y}^\tau$ and $\mathcal{P}^\tau$, respectively. In more general terms, the problem is stated in [GLSW83] as follows:

A **bundle of symplectic manifolds** is a differentiable fibre bundle $F \to E \xrightarrow{\pi} B$ whose structure group[26] preserves a symplectic structure on F. In particular this means that the vertical subbundle $\mathcal{V} = \mathrm{Ker}(T\pi) \subset TE$ carries a field of bilinear forms which we call the **symplectic structure along the fibers** and denote by ω. Of course we can restrict any 2-form Ω on E to $\mathcal{V}$; if this restriction is ω, we call Ω an **extension** of ω.

In [GLSW83] it is investigated when such a closed extension exists, which eventually comes down to a cohomology condition. In the following we explicitly construct such a form in a very natural way.

Definition 6.9.3. A trivialization is **(pre-) symplectic** if the associated flow restricts to a (pre-) symplectic isomorphism on the fiber.

Recall from Corollary 6.7.5 and the remark thereafter that the slicing $\mathfrak{s}_Y$ of Y gives rise to a trivialization, $\mathfrak{s}_{\mathcal{Y}}$ of $\mathcal{Y}^\tau$ and hence induces trivializations

- $\mathfrak{s}_{j^1\mathcal{Y}}$ of $(j^1\mathcal{Y})^\tau$ by jet prolongation
- $\mathfrak{s}_{\mathcal{Z}}$ of $\mathcal{Z}^\tau$ (presymplectic) and $\mathfrak{s}_{T^*\mathcal{Y}}$ of $T^*\mathcal{Y}^\tau$ (symplectic) by pullback.

To picture the latter, suppose $Y = X \times F$ is a trivial bundle. Define a slicing on Y by $\mathbb{R} \times \Sigma \times F \to X \times F$, i.e. $Y_\Sigma = \Sigma \times F$. We have $\mathcal{Y}_\Sigma = \{\varphi_\Sigma : \Sigma \to F\}$ and $\mathcal{Y}_\tau = \{\varphi : \Sigma_\tau \to F\}$. Then

$$\mathfrak{s}_{\mathcal{Y}}(\varphi_\Sigma, \lambda) = \mathfrak{s}_Y \circ \varphi_\Sigma \circ \tau_\lambda^{-1} =: \mathfrak{s}_Y \circ \varphi_\lambda = \varphi_\lambda : \Sigma_{\tau(\lambda)} \to F$$

The fiber of $T^*\mathcal{Y}^\tau$ over $\tau(\lambda)$ is simply $T^*\mathcal{Y}_{\tau(\lambda)}$. $\mathfrak{s}_{T^*\mathcal{Y}^\tau} := (\mathfrak{s}_{\mathcal{Y}}^{-1})^* : T^*\mathcal{Y}_\tau \times T^*\mathbb{R} \to T^*\mathcal{Y}^\tau$. We have

$$\mathfrak{s}^*_{T^*\mathcal{Y}^\tau}\omega_{T^*\mathcal{Y}_\tau} = \mathfrak{s}^*_{T^*\mathcal{Y}^\tau} d\pi^*_{\mathcal{Y}_\tau, T^*\mathcal{Y}_\tau}\pi_\lambda = d(\pi_{\mathcal{Y}_\tau, T^*\mathcal{Y}_\tau} \circ \mathfrak{s}_{T^*\mathcal{Y}^\tau})^*\pi_\lambda = d(\pi_{\mathcal{Y}_\Sigma, T^*\mathcal{Y}_\Sigma})^*\pi_\Sigma = \omega_{T^*\mathcal{Y}_\Sigma}$$

with π_λ and π_Σ denoting elements of $T^*_\varphi\mathcal{Y}_{\tau(\lambda)}$ and $T^*_{\varphi_\Sigma}\mathcal{Y}_\Sigma$, respectively.

Proposition 6.9.4. *If $\mathfrak{s}_Y$ is Lagrangian and Assumption 6.8.3 holds, then $\mathcal{P}^\tau$ is a subbundle[27] of $T^*\mathcal{Y}^\tau$ and the symplectic trivialization $\mathfrak{s}_{T^*\mathcal{Y}}$ on $T^*\mathcal{Y}^\tau$ restricts to a presymplectic trivialization, $\mathfrak{s}_{\mathcal{P}}$, of $\mathcal{P}^\tau$.*

[26] Let (ϕ_i, U_i) be a trivializations of the fiber bundle. Consider $\phi_i\phi_j^{-1} : (U_i \cap U_j) \times F \to (U_i \cap U_j) \times F$. If $\phi_i\phi_j^{-1}(x,\xi) = (x, t_{ij}(x)\cdot\xi)$, with $t_{ij} : U_i \cap U_j \to G$ and G a group (with action "$\cdot$"), then G is called the structure group.

[27] Though not always as a vector bundle. For the relativistic particle, for instance, we have $\pi^\mu\pi_\nu = -m^2$. However, [GM06] give conditions under which this holds as a vector bundle.

The examples throughout this text will convince you that in the usual cases $\mathcal{P}^\tau$ is a subbundle of $T^*\mathcal{Y}^\tau$. The second part then follows directly by including the inclusion $i : \mathcal{P}_\tau \to T^*\mathcal{Y}_\tau$ in the calculation above: Since the inclusion is not even a submersion in general, we can not expect the pulled back form $\omega_{\mathcal{P}_\tau} := i^*\omega_{T^*\mathcal{Y}_\tau}$ to be still symplectic.

We use $\mathfrak{s}_\mathcal{P}$ to coordinatize $\mathcal{P}^\tau$ by (φ, π, λ) and in subscripts abbreviate $\tau(\lambda)$ with λ. Using $\zeta_\mathcal{P}$ ("$\partial/\partial\lambda$ on $\mathcal{P}^{\tau(\lambda)}$"), we can uniquely extend the forms ω_λ along the fibers $\mathcal{P}_\lambda$ to a (degenerate) closed 2-form, ω, on $\mathcal{P}^\tau$ by

$$\omega(\mathcal{V}, \mathcal{W}) := \omega_\lambda(\mathcal{V}, \mathcal{W})$$
$$\omega(\zeta_\mathcal{P}, \cdot) := 0$$

$\forall(\varphi, \pi) \in \mathcal{P}_\lambda$, where $\mathcal{V}$ and $\mathcal{W}$ are vertical vectors on $\mathcal{P}^\tau$ (i.e. tangent to $\mathcal{P}_\lambda$) at (φ, π).

Similarly,
$$H_\zeta : \mathcal{P}^\tau \to \mathbb{R}$$
$$H_\zeta(\varphi, \pi, \lambda) := H_{\lambda,\zeta}(\varphi, \pi)$$
and
$$L_\zeta : T\mathcal{Y}^\tau \to \mathbb{R}$$
$$L_\zeta(\varphi, \dot\varphi, \lambda) := L_{\lambda,\zeta}(\varphi, \dot\varphi)$$

If we take the slicing to be Lagrangian, from the definition of a Lagrangian slicing, it will be obvious that $\mathcal{L}$ can not explicitly depend on the slicing parameter, λ. From the definition of $L_{\lambda,\zeta}$ we see that $L_{\lambda,\zeta}$ then only depends on the infinitesimal slicing, meaning the embedding $\tau(\lambda)$ and the slicing generator $\zeta_X(x, \lambda)$. This yields

Corollary 6.9.5. *Assume ζ_Y is associated to a Lagrangian slicing. Then L_ζ is independent of λ. From the definition of $H_{\lambda,\zeta}$, we then have $\zeta_\mathcal{P}\left[H_\zeta\right] = 0$.*

Consider the 2-form $\omega + dH_\zeta \wedge d\lambda$ on $\mathcal{P}^\tau$. By construction:

$$£_{\zeta_\mathcal{P}}\left(\omega + dH_\zeta \wedge d\lambda\right) = 0$$

Definition 6.9.6. $c : \mathbb{R} \to \mathcal{P}^\tau$, as above, is a **dynamical trajectory** if $c(\lambda)$ covers $\tau(\lambda)$ and

$$\dot c(\lambda) \lrcorner \left(\omega + dH_\zeta \wedge d\lambda\right) = 0$$

Note that the tangent, $\dot c$, to any curve, c, in $\mathcal{P}^\tau$ covering τ can uniquely be split as

$$\dot c =: X + \zeta_\mathcal{P}$$

where X is vertical in $\mathcal{P}^\tau$. Set $X_\lambda := X|_{\mathcal{P}_\lambda}$. Plugging this into the respective equations, we verify:

Proposition 6.9.7. *A curve, c, in $\mathcal{P}^\tau$ is a dynamical trajectory iff* ***Hamilton's equations***

$$X_\lambda \lrcorner \omega_\lambda = dH_{\lambda,\zeta}$$

hold at $c(\lambda)$ for every $\lambda \in \mathbb{R}$.

Corollary 6.9.8. *Let $\mathcal{V}$ be a vector field on $\mathcal{P}^\tau$, whose integral curves are dynamical trajectories. Let $\mathcal{F}_{\lambda_1,\lambda_2} : \mathcal{P}_{\lambda_1} \to \mathcal{P}_{\lambda_2}$ be its flow. Then $\mathcal{F}_{\lambda_1,\lambda_2}$ is symplectic, i.e.*

$$\mathcal{F}^*_{\lambda_1,\lambda_2} \omega_{\lambda_2} = \omega_{\lambda_1}$$

Proof. Using the standard formula $\frac{d}{d\lambda}\psi^*_\lambda \alpha = \psi^*_\lambda \pounds_{\mathcal{V}} \alpha$ for some form α and ψ_λ the flow of $\mathcal{V}$, we have

$$0 = \psi^*_\lambda \pounds_{\mathcal{V}}(\omega + dH \wedge d\lambda) = \frac{d}{d\lambda}\psi^*_\lambda(\omega + dH \wedge d\lambda)$$

by integration over $[\lambda_1, \lambda_2]$, we find

$$\psi^*_{\lambda_1}\omega = \psi^*_{\lambda_2}\omega + (\psi^*_{\lambda_2} - \psi^*_{\lambda_1})dH \wedge d\lambda$$

Applying the trivialization $s_{\mathcal{P}}$ yields the claim. □

At this point we are ready to formulate the relation between solutions to the Euler-Lagrange-equations and Hamilton's equations from above. First, we need to relate elements of $\mathcal{Y}$ (e.g. solutions to the E-L-equ.), to elements of $\mathcal{P}_\lambda$.

Definition 6.9.9. Given $\phi \in \mathcal{Y}$, set $\sigma := \mathbb{F}\mathcal{L}(j^1\phi)$ and define the **canonical decomposition** of ϕ w.r.t. a given slicing[28]

$$c_\phi(\lambda) := R_\lambda(\sigma_\lambda)$$

where $\sigma_\lambda := \sigma \circ i_\lambda$, $i_\lambda : \Sigma_\lambda \to X$ the inclusion.

Corollary 6.9.10. *By Proposition 6.8.4: $c_\phi(\lambda) \in \mathcal{P}_\lambda$*

Finally, we state the theorem giving the relation of solutions to the Euler-Lagrange-equation and Hamilton's equation. Unfortunately, the proof is rather involved and in any case not relevant to this work.

Theorem 6.9.11.
1. *$\phi \in \mathcal{Y}$ solution of Euler-Lagrange equ.*
   $\Rightarrow c_\phi$ satisfies Hamilton's equ. w.r.t. any Lagrangian slicing.
2. *c solution of Hamilton's equ.*
   $\Rightarrow c = c_\phi$ with ϕ a solution of the Euler-Lagrange equ.

[28] Recall: $J^1Y^* \cong Z$.

6.10. Constraint Theory

We have now arrived in the setting of subsection 5.5, searching in the presymplectic manifold $(\mathcal{P}_\lambda, \omega_\lambda)$ for solutions of Hamilton's equations

$$X_\lambda \lrcorner \omega_\lambda = dH_{\lambda,\zeta} \tag{23}$$

Denote the final constraint manifold by C that is

$$C := \mathcal{P}^l_\lambda \text{ , with } \mathcal{P}^l_\lambda \text{ s.t. } \left(T_p \mathcal{P}^l_\lambda\right)^\perp [H] = 0 \; \forall p \in \mathcal{P}^l_\lambda$$

With the sequence of constraint submanifolds

$$C_{\tau,\zeta} = \mathcal{P}^l_{\tau,\zeta} \subset \ldots \subset \mathcal{P}^2_{\tau,\zeta} \subset \mathcal{P}_\tau \subset T^* \mathcal{Y}_\tau$$

in the notation of the constraint algorithm in subsection 5.5.

6.10.0.5. Remark

1. While in finite dimensions, $(i_X \omega_\lambda - dH_{\lambda,\zeta})|_C = 0$ is a system of ODEs and integrability follows automatically, the previous results do not imply integrability in the infinite dimensional case. That X can actually be integrated to a flow is then a difficult analytic problem.

2. We assume that each $\mathcal{P}^l$ as well as C are smooth submanifold of $\mathcal{P}$. In practice, $\mathcal{P}^{l>1}$ and hence C typically have quadratic singularities.
While singularities remain important for questions of linearization stability, quantization, etc., they present no problem in calculating the constraint set.

3. Solutions X, when they exist, are usually *not unique*:
If X is a solution, then so is $X + V$ for any $V \in \ker \omega_\lambda \cap \mathfrak{X}(C)$. Thus, besides being **overdetermined**, signaled by $C \subsetneq \mathcal{P}$, Equation 23 is in general also **underdetermined**, signaling the presence of gauge freedom in the theory as was already mentioned following Proposition 5.5.3.

Finally, we need another definition of well-posedness that will not yield any system with gauge freedom ill posed, as does the definition in section 3. One possibility is to call a system well-posed if it is well-posed after fixing gauges, i.e. introducing new constraints until

$$\ker \omega_\lambda \cap \mathfrak{X}(C \cap \{\text{new constraints}\}) = \emptyset$$

[GM06] chose a more straight forward definition that is easily seen to include this class of systems.

Definition 6.10.1. The Euler-Lagrange equations are **well-posed** relative to a slicing $\mathfrak{s}_Y$ if every $(\varphi, \pi) \in C_{\lambda,\zeta}$ can be extended to a dynamical trajectory $c : \mathbb{R} \supset (\lambda - \epsilon, \lambda + \epsilon) \to \mathcal{P}^\tau$ with $c(\lambda) = (\varphi, \pi)$ and that this solution trajectory depends continuously (in a chosen function space topology) on (φ, π).

Part II. Thesis

7. Warm up: Parametrization- Invariant Theories

Constraints in field theories are not mere mathematical peculiarities. In fact, there is a close relationship between local gauge transformations (a.k.a. gauge transformation of the second kind) and constraints, the starting point of which is the following theorem

Theorem 7.0.2. *[Sun82, Noether's 2. theorem] If the action is invariant under infinitesimal transformations of an infinite continuous group parametrized by r arbitrary functions there exist r independent algebraic or differential identities for the Euler derivatives of the Lagrange function.*

By Euler derivatives are meant the Euler-Lagrange equations, denoted by L_A below. We will not show here, how the relation may be facilitated, but rather take the specific example of parametrization-invariant theories and do an explicit calculation. For a more general treatment, please refer to [Sun82], where also the definition of a param.-inv. theory was taken from, following below. For coordinate free and general results on the correspondence between Gauge groups and constraint submanifolds, please refer to [GM06].

Take a Lagrangian of the form

$$L = \int_X d^{n+1}x \, \mathcal{L}\left(\phi^A, \frac{\partial \phi^A}{\partial x_\mu}\right)$$

note in particular that $\mathcal{L}$ has no explicit x-dependence. Now, if we reparametrize to new coordinates, $y(x)$, say, the Lagrangian changes to

$$L = \int_Y d^{n+1}y \ \det\left(\frac{\partial x_\mu}{\partial y_\nu}\right) \mathcal{L}\left(\phi^A, \frac{\partial y_\mu}{\partial x_\nu}\frac{\partial \phi^A}{\partial y_\mu}\right)$$

We say that the corresponding field theory is **parametrization invariant**, if

$$\mathcal{L}\left(\phi^A, \lambda_\mu{}^\nu \partial_\nu \phi^A\right) = \det\left(\lambda_{\mu\nu}\right) \mathcal{L}\left(\phi^A, \partial_\mu \phi^A\right)$$

with the consequence, that the Lagrangian above actually look the same, just with x and y switched.[29] One can show that the condition is equivalent to[30]

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)} \partial_\nu \phi^A = \delta^\mu{}_\nu \mathcal{L} \tag{24}$$

[29] Note that in physics something similar is referred to as "generally covariant". That notion does not always agree with parametrization invariant as defined here.

[30] Look up homogeneous function, e.g. on Wikipedia. In one dimension one can show for a function $L(\dot{x})$

$$L(\lambda \dot{x}) = \lambda L(\dot{x}) \Leftrightarrow \frac{\partial L}{\partial \dot{x}}\dot{x} - L = 0$$

There are only a few theories which are parametrization-invariant. Notably, Einstein's gravity, the relativistic point particle and the relativistic string. In this section we will apply some of the tools developed in the preceding part of this thesis to a general param.-inv. theory, so that the subsequent sections become more clear.

The Euler-Lagrange equations are as usual

$$L_A := \frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)} = 0$$

However, a straight forward calculation reveals that the equations are not independent for a param.-inv. theory. We have for *any* ϕ the so called generalized Bianchi identities (the identities for the Euler-derivatives as predicted by Theorem 7.0.2)

$$L_A \partial_\mu \phi^A = 0$$

These $n+1$ equations already suggest that there will be $n+1$ primary constraints (cp. [Sun82]). Assuming that $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi)$ has no explicit coordinate dependence we calculate

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)} = \frac{\partial \mathcal{L}}{\partial \phi^B \partial \phi^A_{,\mu}} \partial_\mu \phi^B + \frac{\partial \mathcal{L}}{\partial \phi^B_{,\nu} \partial \phi^A_{,\mu}} \partial_\mu \partial_\nu \phi^B$$

And read off the symbol $F^{\mu\nu}_{AB}$ of the Euler-Lagrange equation by assuming the following expression to be nonvanishing for some indices

$$\frac{\partial \mathcal{L}}{\partial \phi^B_{,\nu} \partial \phi^A_{,\mu}} := F^{\mu\nu}_{AB}$$

The reader may check that the following definition of a characteristic hypersurface in fact coincides with Definition 3.2.1.[31] A hypersurface (locally) given by $z(x) = 0$ is characteristic w.r.t. the Euler-Lagrange equations if we have for the symbol

$$\det\left(\frac{\partial z}{\partial x_\mu} \frac{\partial z}{\partial x_\nu} F^{\mu\nu}_{AB}\right)\Big|_{z(x)=0} = 0 \tag{25}$$

Continuing to the instantaneous formalism we choose local coordinates (z, z_i) of X, s.t. the hypersurface Σ_z is (locally) parametrized by $(0, z_i) : X \supset \Sigma_z \to \Sigma$, and verify

$$L_{z,\zeta}(\varphi, \dot\varphi) = \int_\Sigma i_z^* i_\zeta \left(\mathcal{L}(\phi^A, \partial_\mu \phi^A) d^{n+1} x\right)$$

Note, if $L = L(x, \dot x)$ was a Lagrangian this implies for the associated Hamiltonian

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

That the associated Hamiltonian vanishes "weakly" (i.e. on the constr. submfd.) is a general feature of param.-inv. (field) theories.

[31] Use the inverse function theorem.

$$
\begin{aligned}
&= \int_\Sigma \mathcal{L}\left(\phi^A, \frac{\partial z_\nu}{\partial x_\mu}\frac{\partial \phi^A}{\partial z_\nu}\right)\Bigg|_{z(x)=0} \zeta^\sigma \Phi_\sigma d^n z \\
&= \int_\Sigma \mathcal{L}\left(\varphi^A, \frac{\partial \phi^A}{\partial z}\Bigg|_{z(x)=0}, \frac{\partial \varphi^A}{\partial z_i}\right) \det\left(\frac{\partial z_\nu}{\partial x_\mu}\right) \zeta^\sigma \Phi_\sigma d^n z \\
&= \int_\Sigma \mathcal{L}\left(\varphi^A, \frac{\left(\dot\varphi^A - S^A\right)}{\zeta^\epsilon \frac{\partial z}{\partial x_\epsilon}}, \frac{\partial \varphi^A}{\partial z_i}\right) \det\left(\frac{\partial z_\nu}{\partial x_\mu}\right) \zeta^\sigma \Phi_\sigma d^n z
\end{aligned}
$$

Where we have truncated some parts. In particular, the definition of S^A becomes clear upon revisiting Equation 21. Denoting the integrand by $\mathcal{L}_{z,\zeta}$, we calculate the instantaneous momenta by

$$
\pi_A := \frac{\partial \mathcal{L}_{z,\zeta}}{\partial \dot\varphi^A} = \frac{1}{\zeta^\mu \frac{\partial z}{\partial x_\mu}} \frac{\partial \mathcal{L}}{\partial(\partial_z \phi^A)}(\varphi^A, \dot\varphi^A) \tag{26}
$$

But then, because the theory is param.-inv., by Equation 24, we have

$$
\begin{aligned}
\pi_A \partial_i \varphi^A &= 0 \text{ (n primary constraints)} \\
\pi_A \dot\varphi^A - \zeta^\mu \frac{\partial z}{\partial x_\mu}\mathcal{L} - \pi_A S^A &= 0 \text{ (1 primary constraint)}
\end{aligned}
$$

Setting $\zeta = \partial_z$ this last constraint becomes the vanishing of the Hamiltonian

$$
0 = \pi_A \dot\varphi^A - \mathcal{L} = H
$$

Hence any param.-inv. theory has $n+1$ primary constraints. But how can one investigate with this general expression whether a characteristic hypersurface yields new primary constraints, i.e. connect Equation 25 with Equation 26? From Equation 25 we know that there exist functions K^A, such that

$$
\frac{\partial z}{\partial x_\mu}\frac{\partial z}{\partial x_\nu}\frac{\partial^2 \mathcal{L}}{\partial \phi^A_{,\mu}\partial \phi^B_{,\nu}} K^A = 0
$$

From this, we would like to derive an identity for the π_A of the form

$$
\left(\frac{\partial z}{\partial x_\mu}\frac{\partial \mathcal{L}}{\partial \phi^A_{,\mu}} N^A =\right) \pi_A N^A = \text{something}
$$

for some functions N^A. We leave this investigation open for future research and proceed with a different class of Lagrangians in the next section.

8. Characteristic Hypersurfaces and Constraint Theory

As announced in the main introduction, in the following, we will investigate under which conditions the additional constraint due to a characteristic hypersurface yields a symplectic submanifold.

8.1. Introduction

We start off with a technical remark that goes back to Definition 6.4.1: [GM06] require the Cauchy surfaces to be *spacelike* that is, they require them to be non-characteristic. We believe that this is not used explicitly in the first two parts of [GM06] - the parts we are essentially covering. Of course, if the initial value problem is not well-posed[32], we potentially run into trouble. So we assume that the initial-value hypersurfaces were chosen w.r.t. to the Euler-Lagrange equations, such that the characteristic initial value problem is well-posed (ref. to section 3).

Furthermore, we need the primary constraint manifolds for different times to be diffeomorphic (see Proposition 6.9.4). For this, it should suffice to require that the Cauchy surfaces be of the same class (spacelike, null; i.e. non-char. or characteristic) as the characteristic hypersurfaces vary just as smoothly as the non-characteristic ones do and thus do the constraints under the same reasonable (physical) assumptions made in the non-char. case. These assumptions become clear in section 8.3. In conclusion, we will use the procedure of section 6 on characteristic Cauchy surfaces in the following.

We consider a field theory consisting of an appropriate function space $\mathcal{F}(M)$ over a $n+1$ - dim. manifold M. We choose a coordinate neighborhood U of M with coordinates $x : U \to X \subset \mathbb{R}^{n+1}$ and assume a Lagrangian of the form

$$L = \int_X \left(F^{\mu\nu}_{AB}(x)\partial_\mu \phi^A \partial_\nu \phi^B + G(\phi, x) \right) \, d^{n+1}x \tag{27}$$

with $\phi \in \mathcal{F}(X)$, F, G at least continuous and s.t. the integral is defined and not divergent. When it comes to choosing neighborhoods we chose to be less rigorous in keeping count favoring shorter notation instead. As in the integral above, we will ignore the fact that we are actually considering many neighborhoods, U_i, and should therefore sum over the X_i - this we hide by simply writing X.

Most field theories have an associated *constrained* Hamiltonian system: For non-constrained systems, the Hamiltonian dynamics take place on the whole cotangent bundle over a configuration manifold and are defined through the canonical symplectic form of the cotangent bundle. The dynamics of constrained Hamiltonian systems is restricted to a submanifold which is, in general, only *pre*symplectic.

Using characteristic hypersurfaces for the space of Cauchy data introduces new constraints for field theories of the form above, as was already shown in the main introduction for the Klein-Gordon field. It seems natural to investigate, under which conditions these constraints yield symplectic submanifolds: They would then not give rise to secondary constraint manifolds

[32] I.t.s.o Definition 6.10.1. In most cases this should be equivalent to well-posedness after gauge-fixing (i.e. adding constraints by hand) , see discussion following Proposition 5.5.3 and section 9.

(see Proposition 5.5.1) and this means one could possibly (see section 9) carry over any result from the non-char. hypersurface case that only makes use of the symplectic structure of the cotangent bundle. In more physical terms, they would not be introducing additional gauges (cpw. Proposition 5.5.3). Whence, we will investigate under which conditions the additional constraint yields a symplectic submanifold in the following.

On the other hand, one might be able to turn this around in the sense that a reasonable physical theory should, when restricted to char. hypersurfaces, only introduce constraints that lead to symplectic submanifolds - but we will not pursue this route.

In a finite dimensional vector space setting, it is easy to see that a constraint, which is linear in the momenta, can not yield a symplectic submanifold. That this holds for the case at hand is shown in Corollary 8.4.2. Under certain circumstances on a non-char. hypersurface, it is shown in [GM06] that all primary constraints are linear in the instantaneous momenta. Proposition 8.3.6 gives a tangible condition under which the new constraints due to a char. hypersurface are always non-linear in the inst. momenta.

As the example of a Klein-Gordon-type field on a general background shows (cp. section 8.6) there can be no easy algebraic condition on whether a theory will feature symplectic constraints due to a char. hypersurface. Proposition 8.4.1 gives the sufficient condition for a constraint to yield a symplectic manifold. Finally, since there may be several new constraints with a char. hypersurface (cp. EM, section 8.7), we show in Lemma 8.4.3 that the intersection of symplectic constraints is itself symplectic and combine the last two statements in Proposition 8.4.4.

8.2. The Characteristic Hypersurface(s)

In section 3 we stated that *a* condition for well-posedness of the characteristic initial value problem associated to the Euler-Lagrange equations corresponding to the Lagrangian in Equation 27 could be[33] that data be given on two intersecting transverse characteristic hypersurfaces. Let these hypersurfaces (locally) be described by

$$\begin{aligned} &H_1 = \{z_1(x) = 0\} \text{ and } H_2 = \{z_2(x) = 0\} \\ &\text{s.t. } H_1 \cap H_2 \text{ is a 2-dim. submfd.} \end{aligned}$$

Of course, the instantaneous Lagrangian will now split in two parts:

$$\begin{aligned} L_{z_{1/2}\zeta}(\varphi, \dot{\varphi}) &= \int_{H_1 \cup H_2} i^*_{H_1 \cup H_2} i_\zeta \mathcal{L} \\ &= \int_{H_1} i^*_{z_1} i_\zeta \mathcal{L} + \int_{H_2} i^*_{z_2} i_\zeta \mathcal{L} \end{aligned}$$

[33] Note, that in section 3 the Lagrangian is assumed such that all Euler-Lagrange equations are lin. independent equations of second order. As we have shown, this will not be the case in most applications we consider.

Hence, calculating the instantaneous momenta is not as straight forward as in the preceding example from the last section. Recalling that the Legendre transform is defined as a map

$$\begin{aligned} \mathbb{F}L_{z_{1/2},\zeta} : T\mathcal{Y}_{z_{1/2}} &\to T^*\mathcal{Y}_{z_{1/2}} \\ \text{with } T^*_\varphi\mathcal{Y}_{z_{1/2}} &:= \Big\{\pi : H_1 \cup H_2 \to L\Big(VY|_{H_1\cup H_2}, \Lambda^n(H_1 \cup H_2)\Big)\Big\} \end{aligned}$$

The way to go is

$$\pi_A = \pi^1_A\delta\left(z_1(x)\right) + \pi^2_A\delta\left(1 - z_1(x)\right)$$

where π^i_A, $i = 1, 2$, are the "usual" instantaneous momenta for the respective integral. The first question that comes into mind now, is, whether π^i_A agree on $H_1 \cap H_2$. This is obviously the case as they are defined through restrictions of functions defined on $H_1 \cup H_2$. The more important question concerns derivatives and continuity at the intersection. For instance, in Figure 1, $H_1 \cup H_2$ is not a differentiable manifold, but a manifold with corner! We chose to dodge this question and refer to e.g. D. Joyce, *arXiv:0910.3518* for more information. Since all functions that go into the Legendre transformation are assumed smooth on X, we should be allowed to "smoothen out" the corner when handling the space of initial data on $H_1 \cup H_2$. The instantaneous formalism is mostly concerned with functionals that are defined through some integration over $H_1 \cup H_2$. In this case, $H_1 \cap H_2$ is of measure zero and can be ignored (see also 3. of Remark 6.7.0.4).

In the following section we will be concerned with local questions and consequently assume that we are on either null-hypersurface. In fact, for readability, we will be ignoring the other null-hypersurface right from the start, in the understanding that π_A has "property" should be substituted by π^i_A, $i = 1, 2$, have "property", respectively.

8.3. The Constraints

From the Euler-Lagrange equations associated to Equation 27, we can read off the principal symbol of the differential operator by neglecting lower derivative terms:

$$\begin{aligned} \frac{\partial L}{\partial {\phi^A}_{,\mu}} &= F^{\mu\nu}_{AB}{\phi^B}_{,\nu} + F^{\nu\mu}_{BA}{\phi^B}_{,\nu} \\ \frac{\partial}{\partial x^\mu}\left(\frac{\partial L}{\partial {\phi^A}_{,\mu}}\right) &= \left(F^{\mu\nu}_{AB} + F^{\mu\nu}_{BA}\right){\phi^B}_{,\mu\nu} + \ldots \\ &= 2F^{\mu\nu}_{AB}{\phi^B}_{,\mu\nu} + \ldots \end{aligned}$$

Where in the last step we realized that due to the form of the Lagrangian, F can be assumed to feature this symmetry. Now, let a hypersurface $\Sigma_z \subset X$ (locally) be described by $z(x) = 0$. Usually, Σ_z is said to be characteristic for the E-L-Equ. if

$$\det\left(\frac{\partial z}{\partial x_\mu}\frac{\partial z}{\partial x_\nu}F^{\mu\nu}_{AB}\right)\Bigg|_{z(x)=0} = 0$$

whence $d_x z \in T_x^* X$ belongs to the characteristic variety. But for most field theories, this holds independent of the choice of hypersurface (cp. Vacuum Electrodynamics, section 8.5). Thus we will make the following definition

Definition 8.3.1. We refer to a hypersurface Σ_z as a **characteristic hypersurface**[34] only if this choice of hypersurface further lowers the rank of the symbol, i.e. there exists through every point of Σ_z another hypersurface $\Sigma_{\tilde{z}}$, s.t.

$$\text{Rank}\ \frac{\partial z}{\partial x_\mu}\frac{\partial z}{\partial x_\nu}F^{\mu\nu}_{AB} < \text{Rank}\ \frac{\partial \tilde{z}}{\partial x_\mu}\frac{\partial \tilde{z}}{\partial x_\nu}F^{\mu\nu}_{AB}$$

Unfortunately, we need to restrict to special coordinates[35] for reasons that will become clear when we calculate the instantaneous momenta. For a hypersurface, one can always choose (local) coordinates $(z, z_i) : X \supset V \to I \times \Sigma \subset \mathbb{R}^{n+1}$ $(i = 1, ..., n)$ such that $V \cap \Sigma_z = z^{-1}(0)$, as above. We will again drop the neighborhoods V from the notation.

For an infinitesimal slicing (Σ_z, ζ) and denoting by $i_z : \Sigma \to \Sigma_z \subset X$ the inclusion, the instantaneous Lagrangian reads (neglecting non-derivative terms)

$$\begin{aligned}
L_{z,\zeta}(\varphi, \dot{\varphi}) &= \int_\Sigma i_z^* i_\zeta \mathcal{L} \\
&= \int_\Sigma \Bigg(\frac{\partial z}{\partial x^\mu}\frac{\partial z}{\partial x^\nu}F^{\mu\nu}_{AB}\frac{\partial \phi^A}{\partial z}\frac{\partial \phi^B}{\partial z} + \frac{\partial z}{\partial x^\mu}\frac{\partial z_i}{\partial x^\nu}F^{\mu\nu}_{AB}\frac{\partial \phi^A}{\partial z}\frac{\partial \phi^B}{\partial z_i} + \frac{\partial z_i}{\partial x^\mu}\frac{\partial z}{\partial x^\nu}F^{\mu\nu}_{AB}\frac{\partial \phi^A}{\partial z_i}\frac{\partial \phi^B}{\partial z} \\
&\qquad + \frac{\partial z_i}{\partial x^\mu}\frac{\partial z_j}{\partial x^\nu}F^{\mu\nu}_{AB}\frac{\partial \phi^A}{\partial z_i}\frac{\partial \phi^B}{\partial z_j} + (...)\Bigg)\zeta^\sigma \Phi_\sigma\ d^n z\ ,\ i, j = 1, ..., n \\
&\qquad \text{where } \Phi_\mu := (-1)^\mu \det\left(\frac{\partial x_\nu}{\partial z_i}\bigg|_{\nu\neq\mu}\right)\ ,\ d^n z := dz_1 \wedge dz_2 \wedge ... \wedge dz_n \\
&= \int_\Sigma \Bigg(\frac{1}{\zeta_z^2}\frac{\partial z}{\partial x^\mu}\frac{\partial z}{\partial x^\nu}F^{\mu\nu}_{AB}\left(\dot{\varphi}^A - S^A\right)\left(\dot{\varphi}^B - S^B\right) \\
&\qquad + \frac{1}{\zeta_z}\frac{\partial z}{\partial x^\mu}\frac{\partial z_i}{\partial x^\nu}F^{\mu\nu}_{AB}\left(\dot{\varphi}^A - S^A\right)\left(D_i\varphi^B\right) \\
&\qquad + \frac{1}{\zeta_z}\frac{\partial z_i}{\partial x^\mu}\frac{\partial z}{\partial x^\nu}F^{\mu\nu}_{AB}\left(D_i\varphi^A\right)\left(\dot{\varphi}^B - S^B\right) + (...)\Bigg)\zeta^\sigma \Phi_\sigma\ d^n z\ ,\ \text{with } \zeta_z := \zeta^\mu \frac{\partial z}{\partial x_\mu}
\end{aligned}$$

Where S^A denotes the "rest terms" upon solving for $\partial\phi^A/\partial z$ (cpw. Equation 21) and D_i denoting formal partial derivatives w.r.t. z_i (cpw. Definition 6.4.3). Taking the Frechet derivative of the second argument of $L_{z,\zeta}$, i.e. w.r.t $\dot{\varphi}$, the instantaneous momenta read

$$\frac{1}{\zeta^\mu \Phi_\mu}\pi_A = \left(\frac{1}{\zeta_z}\right)^2 \frac{\partial z}{\partial x^\mu}\frac{\partial z}{\partial x^\nu}\left(F^{\mu\nu}_{AB} + F^{\mu\nu}_{BA}\right)\left(\dot{\varphi}^B - S^B\right)$$

[34]This is this papers definition and does *not* agree with the usual one.

[35]We are not working in adapted coordinates from the start. In our case the Lagrangian is not form-invariant for a general change of coordinates. We will choose an "adapted embedding" instead, which of course amounts to a coordinate change of the Lagrangian to adapted coordinates.

$$+\frac{1}{\zeta_z}\frac{\partial z}{\partial x^\mu}\frac{\partial z_i}{\partial x^\nu}F^{\mu\nu}_{AB}\left(D_i\varphi^B\right)+\frac{1}{\zeta_z}\frac{\partial z_i}{\partial x^\mu}\frac{\partial z}{\partial x^\nu}F^{\mu\nu}_{BA}\left(D_i\varphi^B\right)$$

Because Σ_z is characteristic, there exists a **constraint vector** $K^A(z_1,...)$ s.t.[36]

$$\frac{1}{\zeta^\mu\Phi_\mu}\pi_A K^A=\frac{2}{\zeta_z}\frac{\partial z}{\partial x_\mu}\frac{\partial z_i}{\partial x_\nu}F^{\mu\nu}_{AB}K^A\left(D_i\varphi^B\right)$$

Hence we can choose $\overline{E}$ s.t. $K^{\overline{E}}\neq 0$ and substitute

$$\pi_{\overline{E}}=\frac{1}{K^{\overline{E}}}\left(-\pi_E K^E+\underbrace{\zeta^\sigma\Phi_\sigma\frac{2}{\zeta_z}\frac{\partial z}{\partial x_\mu}\frac{\partial z_i}{\partial x_\nu}F^{\mu\nu}_{AB}}_{=:\Xi^i_{AB}}K^A\left(D_i\varphi^B\right)\right),\ E\neq\overline{E} \tag{28}$$

which is a primary constraint. We can readily read off the following result:

Proposition 8.3.2. *The primary constraint associated to K^A is linear in the instantaneous momenta iff $\Xi^i_{AB}K^A=0$.*

One might be tempted at this point to define a class of special, hypersurface independent constraint vectors, as those satisfying $F^{\mu\nu}_{AB_p}K^A\neq 0$ (for some B_P). Unfortunately, this class seems to be of little interest, as already for electromagnetism on a general background the constraint vectors depend on the hypersurface chosen (cp. section 8.7).

For future reference and for a better understanding of the following, we show

Lemma 8.3.3. *Let K^A be a constraint vector. If the associated constraint is linear in the instantaneous momenta, then the mappings $\flat_B : v_\mu\mapsto F^{\mu\nu}_{AB}K^A v_\mu$ are degenerate, with $\partial_\mu z$ an element of the kernel.*

Proof. We have $\Xi^i_{AB}K^A=\flat_B(\partial z/\partial x_\mu)\partial z_i/\partial x_\mu=0$ (i.e. that it is linear in the inst. momenta). $\{\partial z/\partial x_\mu,\ \partial z_i/\partial x_\mu\}$ span a basis and we have $\flat_B(\partial z/\partial x_\mu)v=0$ for any basis element v, as $\flat_B(\partial z/\partial x_\mu)\partial z/\partial x_\mu=0$ by assumption. Hence $\flat_B$ is degenerate for all B, with $\partial z/\partial x_\mu\in\mathrm{Ker}(\flat_B)$. □

Lemma 8.3.4. *Let $z_s(x)$ define a smooth family of hypersurfaces given by $z_s(x)=0$. Then the constraint vectors K_{i_s} vary smoothly with s.*

Proof. The K_{i_s} are defined by

$$\partial_\mu z_s\partial_\nu z_s F^{\mu\nu}_{AB}K^A_{i_s}=0$$

And thus depend smoothly on s.[37] □

[36]Recalling the preceding section: How does $K^A(z_1,...)$ behave when transitioning through the intersection to the other null-hypersurface?

[37]I.g. eigenvectors and alike need not vary smoothly with a smooth 1-param. family of matrices

Lemma 8.3.5. *For every characteristic hypersurface, z, there exist for every point a neighborhood and a family of non-characteristic hypersurfaces, z_s, s.t. $z_s \to z$ in that neighborhood.*

Proof. By the definition of characteristic hypersurface, there exist through every point of z at least one non-characteristic hypersurface, z_0, intersecting z transversely - otherwise the char. hypersurface would not be lowering the rank at this point.
Fix a point and using both hypersurface's (local) orthogonal vector fields, n_z and n_{z_0}, choose a neighborhood that is a tubular neighborhood for both hypersurfaces, respectively. By definition of tubular neighborhoods both orthogonal vector fields are now continued to non vanishing vector fields on the whole neighborhood, which we still denote by n_z and n_{z_0}. Define a vector field $a(x)(sn_z + (s-1)n_{z_0})$, with $a(x)$ s.t. points on z_0 will be mapped to points on z_1 by the time-one-flow and $a(x) = 0$ for x on the intersection. In particular, $a(x)$ will have to change sign there. The $s = \text{const.}$ hypersurfaces are the ones we were looking for. □

Proposition 8.3.6. *Assume* $\text{rank}(\partial_\mu f F^{\mu\nu}_{AB}) = \text{const.}$, $\forall \nu$ *and all functions f with $df \neq 0$ that define hypersurfaces in a half open neighborhood whose boundary coincides with the char. hypersurface. Then the new,* ***characteristic constraint vectors****, $\tilde{K}_i$, resulting from choosing a characteristic hypersurface, z, yield constraints non-linear in the momenta.*

Proof. Let z_s be a smooth family of non-characteristic surfaces, s.t. $z_s \to z$, which exists by virtue of Lemma 8.3.5. As shown in Lemma 8.3.4, the K_{i_s} depend smoothly on s. By assumption

$$\text{rank}\left(\partial_\mu z F^{\mu\nu}_{AB}\right) = \text{rank}\left(\partial_\mu z_s F^{\mu\nu}_{AB}\right)$$

Thus we can choose the K_{i_s} so that they define an orthonormal frame varying smoothly with s.
The only things that could go wrong w.r.t. to the conclusion of the prop. is that two constraint vectors, say K_{1_s} and K_{2_s} converge to the same element or one of them vanishes.[38] Because by assumption we would then get at least two new constraint vectors,[39] $\tilde{K}_{1/2}$, with one of which, $\tilde{K}_1$, say, satisfying $\partial_\mu z F^{\mu\nu}_{AB} \tilde{K}^A_1 = 0$.
By smoothness, however, K_{1_s} and K_{2_s} have to stay orthogonal (the eucl. scalar product is a smooth function). Since we furthermore chose them normalized (w.r.t. eucl. scalar product), they can not vanish. In conclusion a new constraint, (i.p. linear independent of the K_i), $\tilde{K}_1$, can not belong to the kernel. □

In the setting of [GM06], the primary constraint manifolds of a non-char. slicing are independent of the slicing chosen - in particular the number of constraint vectors is constant along non-char. hypersurfaces. [GM06] give conditions under which the constraints are linear in the inst. momenta (these are the theories they focus on). Via Lemma 8.3.3, the constant rank assumption in Proposition 8.3.6 is then satisfied for f defining non-char. hypersurfaces. It would be nice to show something along the lines of:

[38]"creating space for a new vector in the kernel"
[39]By definition of constraint hypersurface we need to end up with more constraint vectors than before

In the setting of [GM06]. If for non-char. hypersurfaces all constraints are linear in the inst. momenta and (...) then the new constraints due to choosing a char. hypersurface are non-linear in the instantaneous momenta.
To show this we 'only' need to show that the rank will still be the same in the limit to the char. hypersurface - possibly by introducing another assumption "(...)".

8.4. Symplecticity of the Constraint Submanifold

Now, we will take a closer look at the constraint submanifold $(\mathcal{P}_z^K, i^*_{\mathcal{P}_z^K}\omega_z) \subset (T^*\mathcal{Y}_z, \omega_z)$, created by a constraint vector K. We recall that $(T^*\mathcal{Y}_z, \omega_z)$ is in fact symplectic. The symplectic form ω_z is, in particular, a bilinear anti-symmetric form. We *assume* that for vector fields $(V, \tilde{V})$, $(W, \tilde{W}) \in T(T^*\mathcal{Y}_z)$, ω_z is of the form

$$\omega_z\Big((V, \tilde{V}), (W, \tilde{W})\Big) = \langle V^A, \tilde{W}_A\rangle - \langle W^A, \tilde{V}_A\rangle$$

for a dual pairing $\langle\cdot,\cdot\rangle$ satisfying

$$\langle D_i v, \tilde{w}\rangle = -\langle v, D_i\tilde{w}\rangle \text{ and } \langle v, \tilde{w}\rangle = \langle v\tilde{w}, 1\rangle$$

The most prominent example of such a pairing is the $L^2(\Sigma_z)$ inner product with vanishing boundary terms. The L^2 inner product is usually also the one encountered in field theories. In fact, going back to section 6.5, will convince the reader that we are usually in this environment. What is more, the symplectic form consists of an integral over the initial value surface. Going back to section 8.2, we see, that we can ignore the problem of continuity at the intersection, as the intersection is of measure zero in the integral. Furthermore, the symplectic form will split

$$\omega_z = \int_{H_1}(...) + \int_{H_2}(...) \tag{29}$$

Thus, we may again focus on one term and ignore the second.

In this section we will proof the following proposition and discuss its meaning using some examples.

Proposition 8.4.1. *In the setting above we have: $(\mathcal{P}_z^K, i^*_{\mathcal{P}_z^K}\omega)$ is symplectic iff the only solution to*

$$0 = \left[D_i\left(\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AB}K^AK^B\right)\right]L^{\overline{E}} + 2\left[\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AB}K^AK^B\right]D_iL^{\overline{E}} \tag{30}$$

in the appropriate function space[40] *is* $L^{\overline{E}} = 0$.

[40] $L^{\overline{E}}$ *should* be the component of a vector in $T\mathcal{P}_z^K$. In the proof, one can read off the remaining components (depending on $L^{\overline{E}}$). Then, one can check whether $(L, \tilde{L}) \in T\mathcal{P}_z^k$.

Proof. We need to show: $\forall (V,\tilde V) \in T\mathcal{P}^K_z \subset T(T^*\mathcal{Y}_z)$, there exists $(W,\tilde W) \in T\mathcal{P}^K_z$ s.t. $\omega_z((V,\tilde V),(W,\tilde W)) \neq 0$.[41] So first, we need to figure out what $T\mathcal{P}^k_z$ actually is. As taking variational derivatives before and after substituting $\pi_{\overline{E}} = 1/K^{\overline{E}}(\ldots)$ (Equation 28) needs to agree, it is readily seen

$$T\mathcal{P}^K_z = \left\{ L^A \frac{\delta}{\delta\varphi^A} + \tilde L_A \frac{\delta}{\delta\pi_A} \,\middle|\, \tilde L_{\overline{E}} = \frac{1}{K^{\overline{E}}}\left(-\tilde L_E K^E + \Xi^i_{AB} K^A \left(D_i L^B\right)\right) \;,\; E \neq \overline{E} \right\}$$

Thus for $(L,\tilde L), (M,\tilde M) \in T\mathcal{P}^K_z$ and denoting by $\langle\cdot,\cdot\rangle$ the $L^2(\Sigma)$ inner product, we get

$$\begin{aligned}
\omega\left((L,\tilde L),(M,\tilde M)\right) &= \langle L^A, \tilde M_A\rangle - \langle M^A, \tilde L_A\rangle \\
&= \langle L^E, \tilde M_E\rangle + \langle L^{\overline{E}}, \frac{1}{K^{\overline{E}}}\left(\Xi^i_{AB}K^A(D_iM^B) - K^E\tilde M_E\right)\rangle \\
&\quad - \langle M^E, \tilde L_E\rangle - \langle M^{\overline{E}}, \frac{1}{K^{\overline{E}}}\left(\Xi^i_{AB}K^A(D_iL^B) - K^E\tilde L_E\right)\rangle \\
&= \langle L^E - \frac{K^E}{K^{\overline{E}}}L^{\overline{E}}, \tilde M_E\rangle \\
&\quad - \langle D_i\left(\frac{1}{K^{\overline{E}}}\Xi^i_{AE}K^A L^{\overline{E}}\right) + \tilde L_E, M^E\rangle \\
&\quad - \langle D_i\left(\frac{1}{K^{\overline{E}}}\Xi^i_{A\overline{E}}K^A L^{\overline{E}}\right) + \frac{1}{K^{\overline{E}}}\left(\Xi^i_{AB}K^A(D_iL^B) - K^E\tilde L_E\right), M^{\overline{E}}\rangle \\
&\overset{!}{=} 0 \;,\; \forall (M,\tilde M) \in T\mathcal{P}^K_z
\end{aligned}$$

For this to hold, the left hand side of each inner product has to vanish, respectively. Plugging the first two resulting equations into the last yields

$$\begin{aligned}
0 &= D_i\left(\frac{1}{K^{\overline{E}}}\Xi^i_{A\overline{E}}K^A L^{\overline{E}}\right) + \frac{1}{K^{\overline{E}}}\left(\Xi^i_{AE}K^A\left(D_i\left[\frac{K^E}{K^{\overline{E}}}L^{\overline{E}}\right]\right) + \Xi^i_{A\overline{E}}K^A(D_iL^{\overline{E}})\right) \\
&\quad + \frac{1}{K^{\overline{E}}}K^E D_i\left(\frac{1}{K^{\overline{E}}}\Xi^i_{AE}K^A L^{\overline{E}}\right) \\
&= \left[D_i\left(\frac{1}{K^{\overline{E}}}\Xi^i_{A\overline{E}}K^A\right) + \frac{1}{K^{\overline{E}}}\Xi^i_{AE}K^A\left(D_i\left[\frac{K^E}{K^{\overline{E}}}\right]\right) + \frac{1}{K^{\overline{E}}}K^E D_i\left(\frac{1}{K^{\overline{E}}}\Xi^i_{AE}K^A\right)\right]L^{\overline{E}} \\
&\quad + 2\left[\frac{1}{K^{\overline{E}}}\Xi^i_{A\overline{E}}K^A + \left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AE}K^AK^E\right]D_iL^{\overline{E}} \\
&= \left[D_i\left(\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AB}K^AK^B\right)\right]L^{\overline{E}} + 2\left[\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AB}K^AK^B\right]D_iL^{\overline{E}}
\end{aligned}$$

$(\mathcal{P}^K_z, \omega)$ is symplectic iff the only solution is $L^{\overline{E}} = 0$. □

[41] We are only showing this for one term in Equation 29. So what if the two terms happened to cancel each other? In general theories this is extremely unlikely: For one thing, $F^{\mu\nu}_{AB}(x)$ would probably have to feature some $H_1 \leftrightarrow H_2$ symmetry. Then, still, one could choose a different parametrization of z_2 or change ζ.

The easiest example where the condition is satisfied, is for

$$\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AB} K^A K^B = c(z_1, ..., \hat{z}_j, ..., z_n) \cdot \delta^{ij} \ , \ \text{no sum over } j, z_j \text{ excluded}$$

Then the only solution in any differentiable function space will obviously be the null solution. This is for instance the case for the Klein-Gordon - and electromagnetic field on a Minkowski background (cp. 8.5). With more general backgrounds, however, one can not expect the coefficients to be independent of a coordinate.
We can rewrite the general condition to

$$D_{\bar{l}} L^{\overline{E}} + \frac{A^l}{A^{\bar{l}}} D_l L^{\overline{E}} + \frac{1}{2A^{\bar{l}}} (D_l A^l) L^{\overline{E}} = 0 \ , \ \text{sum over } l \neq \bar{l}$$

By the Cauchy-Kowalevski Theorem, we know that this equation has a unique solution if the coefficients are real analytic.[42] In general, we should expect to find a solution that is not the null solution. One then has to check whether the solution is actually in the function space (vector space) considered. In the case of a Klein-Gordon field on (half) de-Sitter background (cp. section 8.6), the solution can, for example, be readily discarded as failing this essential requirement.

For future reference we note that when all coefficients in the condition vanish, it will be satisfied for any function. As an immediate consequence of Proposition 8.3.2 we get

Corollary 8.4.2. *Constraints that are linear in the instantaneous momenta yield non-symplectic submanifolds.*

Of course, now that we have established criteria for the symplecticity of a single constraint manifold, we are interested in the nature of their intersections. After all, the field dynamics (i.e. the Hamiltonian formalism), are taking place in the intersection of all constraint manifolds.

Lemma 8.4.3. *Let $(V, \langle\cdot,\cdot\rangle)$ be a separable Hilbert space and consider the symplectic vector space $(V \times V, \omega)$ with the obvious symplectic form ω. Let $W_i \subset V \times V$ be symplectic subspaces. Then $W = \bigcap W_i$ is a symplectic subspace of $(V \times V, \omega)$.*

Proof. By assumption we can choose an orthonormal basis $\mathcal{B} = \{u_1, u_2, ...\}$ of V. Let $W_1 = \text{span}\{u_i\}$, $i \in \sigma_{W_1}$, where σ_{W_1} denotes the corresponding index set for W_1 and define σ_{W_2}, σ_W analogously.
Assume $\exists (k,l) \in \sigma_W$ s.t. $\forall (i,j) \in \sigma_W : 0 = \omega((u_k, u_l), (u_i, u_j))$. Then there exist $(m,n) \in \sigma_{W_1}/\sigma_{W_2}$ and $(o,p) \in \sigma_{W_2}/\sigma_{W_1}$ such that

$$0 \neq \langle u_k, u_n \rangle - \langle u_l, u_m \rangle$$
$$0 \neq \langle u_k, u_p \rangle - \langle u_l, u_o \rangle$$

[42] I.p. there are counter examples with coefficients that are merely smooth, e.g. "Lewys example"

Whence $k = n$, $l = o$ or $l = m$, $k = p$ (otherwise $(u_m = u_o, u_n = u_p) \in W$). Choose the first option, in particular: $(o, n) \notin \sigma_W$ (otherwise, we would be done already). But, since by assumption $(k, l) \in \sigma_W$, we have $(n, o) \in \sigma_W$. This is a contradiction, as it is easy to show that: $(u_i, 0) \in W \Leftrightarrow (0, u_i) \in W$ and thus $(n, o) \in \sigma_W \Leftrightarrow (o, n) \in \sigma_W$. □

Proposition 8.4.4. *If all characteristic constraints satisfy Proposition 8.4.1 and provided we are working in a separable Hilbert space, then the* ***characteristic constraint manifold*** *- the intersection of the respective constraint manifolds - is symplectic.*

8.5. Vacuum Electrodynamics on Minkowski $(-1, 1, 1, 1)$:

$$
\begin{aligned}
F^{\mu\sigma}_{\nu\epsilon} =& 2\left(\eta^{\mu\sigma}\eta^{\nu\epsilon} - \eta^{\mu\epsilon}\eta^{\nu\sigma}\right) \\
&\Rightarrow F^{\mu\sigma}_{\nu\epsilon} = F^{\sigma\mu}_{\epsilon\nu} = F^{\nu\epsilon}_{\mu\sigma} \ , \ F^{\mu\sigma}_{\nu\epsilon} = -F^{\mu\epsilon}_{\nu\sigma} \\
z =& \frac{1}{2}(x_0 + x_1) - x^+ \ , \ x^+ = \text{const.} \\
z_1 =& \frac{1}{2}(x_0 - x_1) \\
z_m =& x_m \ , \ m = 2, 3
\end{aligned}
$$

Using the formulas from above we calculate

$$
\begin{aligned}
\pi_\mu = \frac{1}{\zeta_z}\frac{1}{4}\Big[& \left(F^{00}_{\mu\nu} + F^{00}_{\nu\mu}\right)(\dot{\varphi}^\nu - S^\nu) \\
& + \left(F^{01}_{\mu\nu} + F^{10}_{\nu\mu}\right)(\dot{\varphi}^\nu - S^\nu) + \left(F^{10}_{\mu\nu} + F^{01}_{\nu\mu}\right)(\dot{\varphi}^\nu - S^\nu) \\
& + \left(F^{11}_{\mu\nu} + F^{11}_{\nu\mu}\right)(\dot{\varphi}^\nu - S^\nu) \Big] + (...)
\end{aligned}
$$

And by using the symmetries of $F^{\mu\sigma}_{\nu\epsilon}$ above,

$$
\pi_\mu = \frac{1}{\zeta_z}\frac{1}{2}\left(F^{00}_{\mu\nu} + F^{01}_{\mu\nu} + F^{10}_{\mu\nu} + F^{11}_{\mu\nu}\right)(\dot{\varphi}^\nu - S^\nu) + (...)_\mu \tag{31}
$$

We check that the usual constraint holds, namely for the "evolution direction", $K = (1, 1, 0, 0)$, we get $\pi_\mu K^\mu = 0$. This holds also for non-characteristic, spacelike hypersurfaces. For these this is the only constraint.

For the characteristic hypersurface $z = 0$, we verify that for $K^\mu = \delta^{\mu\bar{\mu}}$, $\bar{\mu} \neq 0, 1$, $\pi_\mu K^\mu = (...)_\mu K^\mu \neq 0$, i.e. the part we have written out in Equation 31 vanishes, while the rest, $(...)_\mu K^\mu$, does not. Continuing our algorithm with this K, we have[43]

$$
\left(\frac{1}{K^{\bar{\mu}}}\right)^2 \Xi^i_{\mu\nu} K^\mu K^\nu = \Xi^i_{\bar{\mu}\bar{\mu}}
$$

[43] We still sum $\mu, \nu = 0, .., 3$

$$= \frac{1}{2}\left(\delta_{\mu 0}+\delta_{\mu 1}\right)\delta_{i1}\left(\delta_{\nu 0}-\delta_{\nu 1}\right)F^{\mu\nu}_{\overline{\mu\mu}}+\left(\delta_{\mu 0}+\delta_{\mu 1}\right)\delta_{im}F^{\mu m}_{\overline{\mu\mu}}\ ,\ m\neq 1$$
$$= -\delta_{i1}$$

So as noted following the proof of Proposition 8.4.1, the constraint submanifold associated to the characteristic constraint is symplectic.

8.6. $F^{\mu\nu}_{AB} \propto \sqrt{-g}g^{\mu\nu}$ with Lorentz Metric g

We will first show that the symplecticity condition can in general be reduced to an ordinary differential equation whose solution is then given and explicitly calculated as an example for the 'half De Sitter' spacetime metric.

The characteristic hypersurfaces are exactly the null hypersurfaces for g (cp. [HE75, p.44]) that is, let $\Sigma_z \subset X$ be locally given by $z(x) = 0$ and denote $dz = n_a dx^a$. Then the induced metric on Σ, $i_z^* g$, will be degenerate if $g^{\mu\nu}n_\mu n_\nu = 0$, in which case Σ_z is called a **null hypersurface**. We can proceed directly with the algorithm and calculate

$$\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AB}K^A K^B = \Xi^i = \zeta^\mu \det\left(\frac{\partial x_\nu}{\partial z_j}\Big|_{\nu\neq\mu}\right)\frac{1}{\zeta^\sigma \frac{\partial z}{\partial x_\sigma}}\frac{\partial z}{\partial x_\mu}\frac{\partial z_i}{\partial x_\nu}g^{\mu\nu}\sqrt{-g}$$

Where of course we do not have any 'A, B-components'. Defining $L^\nu := \frac{\partial z}{\partial x_\mu}g^{\mu\nu}$, we have

$$\langle L, L\rangle_g = 0 \text{ (By construction)}$$
$$\langle L, V\rangle_g = dz(V) = 0, \forall V \in T\Sigma_z \ (\Sigma_z \text{ is given by } z(x) = \text{const.})$$

As $\langle\cdot,\cdot\rangle_g$ is non-degenerate, and $\operatorname{codim}\Sigma_z = 1$, we have

$$\langle L, N\rangle_g \neq 0 \Leftrightarrow N \in \operatorname{Span}\left(\frac{\partial z}{\partial x^\mu}\right) = T\Sigma_z^\perp$$

Whence $L \in T\Sigma_z$. Note that L defines a smooth, non-vanishing vector field on Σ_z. Define $\hat{L} := L/||L||^2_{\text{Eucl}}$. The flow of $\hat{L}$ induces a foliation $\psi : \Sigma_z \xrightarrow{\cong} I \times U \subset \mathbb{R}\times\mathbb{R}^2$. Choosing $z_i = \psi_i$, we verify: $\partial_\mu z_1 = \hat{L}^\mu$, $\hat{L}^\mu \partial_\mu z_i = 0$, $i = 2, 3$ (the leaves of the foliation are given by $z_1 = $ const.). In these coordinates we immediately obtain:

$$\Xi^i = \zeta^\delta \det\left(\frac{\partial x_\kappa}{\partial z_j}\Big|_{\kappa\neq\delta}\right)\frac{1}{\zeta^\sigma\frac{\partial z}{\partial x_\sigma}}\sqrt{-g}\delta^{i1}$$

Since this term is not constant, we actually need to look at Equation 30, which reads (changing $L^{\overline{E}}$ to f and using h in an obvious way)

$$0 = (D_1 h)f + 2h(D_1 f)$$

$$\Rightarrow f = c(z_2, z_3)h^{-\frac{1}{2}}$$

So that we can write down the solution in our special coordinates:

$$f = c(z_2, z_3)\left(\frac{\zeta^\sigma \frac{\partial z}{\partial x_\sigma}}{\zeta^\delta \det\left(\left.\frac{\partial x_\kappa}{\partial z_j}\right|_{\kappa\neq\delta}\right)\sqrt{-g}}\right)^{\frac{1}{2}} \tag{32}$$

Since we actually did find a solution, we need to check whether it is in the vector space we considered. Below we will do this explicitly for the de Sitter metric using special coordinates that only cover half of the spacetime.

8.6.0.6. "Half De Sitter"

Presumably our universe becomes a de Sitter universe for large times. In coordinates covering only half of de Sitter (cp. [HE75]):

$$g = -dx_0 \otimes dx_0 + a(x_0)^2 dx_i \otimes dx_i$$

$\Sigma_z := \{z(x) = \text{const.}\}$ is characteristic if

$$0 = \frac{\partial z}{\partial x_\mu}\frac{\partial z}{\partial x_\nu}g^{\mu\nu} = -(\partial_0 z)^2 + \frac{1}{a^2}(\partial_i z)^2$$

With the ansatz $z = z(x_0, x_1)$, this becomes $0 = \partial_0 z \pm \frac{1}{a}\partial_1 z$. For de Sitter, we take $a = e^{x_0}$, yielding $z_\pm = c(x_2, x_3)e^{e^{-x_0}\pm x_1}$. We define new coordinates:

$$\begin{aligned} z &= e^{e^{-x_0}+x_1} \\ z_1 &= e^{e^{-x_0}-x_1} \\ z_i &= x_i\ ,\ i = 2, 3 \end{aligned}$$

These are valid coordinates:

$$\det\left(D\begin{pmatrix} z \\ z_1 \end{pmatrix}\right) = \det\begin{pmatrix} -e^{-x_0}z & z \\ -e^{-x_0}z_1 & -z_1 \end{pmatrix} = 2e^{-(x_0+2e^{-x_0})} \neq 0$$

Or, more directly

$$\begin{aligned} x_0 &= -\ln\left(\ln\sqrt{zz_1}\right) \\ x_1 &= \ln\left(\sqrt{\frac{z}{z_1}}\right) \end{aligned}$$

Thus the coordinates are only defined for $\{z_1, z \in \mathbb{R} \mid zz_1 > 1\}$. We calculate

$$\frac{\partial x_0}{\partial z} = -\frac{1}{2z\ln\sqrt{zz_1}}$$

$$\frac{\partial x_1}{\partial z} = \frac{1}{2z}$$

and for the other derivatives

$$\left(\frac{\partial x_\nu}{\partial z_j}\right) = \begin{pmatrix} -\frac{1}{2z_1 \ln \sqrt{zz_1}} & -\frac{1}{2z_1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{aligned} \zeta^\mu \det\left(\frac{\partial x_\nu}{\partial z_j}\Big|_{\nu\neq\mu}\right) &= \zeta^0\left(-\frac{1}{2z_1}\right) + \zeta^1\left(-\frac{1}{2z_1 \ln \sqrt{zz_1}}\right) \\ \frac{1}{\zeta^\sigma \frac{\partial z}{\partial x^\sigma}} &= \left(-\zeta^0 z \ln \sqrt{zz_1} + \zeta^1 z\right)^{-1} \end{aligned}$$

and we can finally write down

$$\begin{aligned} \Xi^i &= \zeta^\mu \det\left(\frac{\partial x_\nu}{\partial z_j}\Big|_{\nu\neq\mu}\right) \frac{1}{\zeta^\sigma \frac{\partial z}{\partial x_\sigma}} \frac{\partial z}{\partial x_\mu} \frac{\partial z_i}{\partial x_\nu} g^{\mu\nu} \sqrt{-g} \\ &= \frac{\zeta^0 + \frac{1}{\ln \sqrt{zz_1}}\zeta^1}{\ln \sqrt{zz_1}\zeta^0 - \zeta^1} \left(\ln \sqrt{zz_1}\right)^{-1} := h \end{aligned}$$

Plugging these calculations into the more general formula Equation 32, we require $f \in L^2(\Sigma_z)$, i.e.

$$\infty > \int_{\Sigma_z} f^2 d^3 z = \int_{\mathbb{R}^2} dz_2 dz_3 c(z_2, z_3)^2 \int_{\{zz_1 > 1\}} dz_1 \frac{1}{|h|}$$

But it is easily verified that

$$\frac{1}{h} = \ln \sqrt{zz_1} \frac{\ln \sqrt{zz_1}\zeta^0 - \zeta^1}{\zeta^0 + \frac{1}{\ln \sqrt{zz_1}}\zeta^1} \xrightarrow{z_1 \to \infty} \infty$$

Whence our solution, f, can not be in e.g. $L^2(\Sigma_z)$ and the only solution in $L^2(\Sigma_z)$ is $f = 0$. Ad hoc, one would even expect the function space to be more restricted in order for the Lagrangian with diverging metric part to be well defined. Note that it is also that factor, $\sqrt{-g}$, which is responsible for the divergence above.

8.7. EM on curved spacetime

In this section we will see that the analysis from the last section carries over to electromagnetism on a curved background. Through the same procedure we will again be able to reduce the symplecticity condition to an ordinary differential equation and explicitly calculate the solution for the 'half De Sitter' spacetime metric.

$$L = -\int_X \frac{1}{4}\sqrt{-g}g^{\mu\nu}g^{\sigma\epsilon}\nabla_{[\mu}A_{\sigma]}\nabla_{[\nu}A_{\epsilon]} = -\int_X \frac{1}{4}\sqrt{-g}g^{\mu\nu}g^{\sigma\epsilon}\partial_{[\mu}A_{\sigma]}\partial_{[\nu}A_{\epsilon]}$$

Expressing the Lagrangian through covariant derivatives, the Christoffel symbols will cancel from the Lagrangian due to all the anti-symmetries involved. We can read off right away that

$$\begin{aligned} F^{\mu\nu}_{\sigma\epsilon} &= -\frac{1}{2}\sqrt{-g}(g^{\mu\nu}g^{\sigma\epsilon} - g^{\mu\epsilon}g^{\sigma\nu}) \\ &= F^{\nu\mu}_{\sigma\epsilon} = -F^{\mu\epsilon}_{\sigma\nu} = -F^{\sigma\nu}_{\mu\epsilon} \end{aligned}$$

Before calculating a characteristic hypersurface, we note that actually for any vector V_μ we have

$$\det\left(V_\mu V_\nu\left(F^{\mu\nu}_{\sigma\epsilon} + F^{\nu\mu}_{\epsilon\sigma}\right)\right) = \det\left(2V_\mu V_\nu F^{\mu\nu}_{\sigma\epsilon}\right) = 0$$

To see this, assume that g is diagonal, which, at a point, one is always free to do. Take a vector $K \neq 0$ and calculate

$$\left(V_\mu V_\nu F^{\mu\nu}_{\sigma\epsilon}\right)K^\sigma = \left((V_\mu)^2 g^{\mu\mu}g^{\sigma\epsilon} - V_\sigma V_\epsilon g^{\sigma\sigma}g^{\epsilon\epsilon}\right)K^\sigma \quad \text{(no sum over } \epsilon)$$

Setting $K = V$ the expression vanishes. For $V_\mu V_\nu g^{\mu\nu} \neq 0$, V itself spans the kernel. Note that $V_\mu F^{\mu\nu}_{\sigma\epsilon}V_\sigma = 0$, but $V_\mu F^{\mu\nu}_{\sigma\epsilon}V_\epsilon \neq 0$.

Since the principal symbol is degenerate for a generic choice of V, we want to find a special choice that further lowers the rank. As our findings above suggest, we see that for V a null-vector, does not only V itself lie in the kernel, but additionally any vector in $\{K \mid g^{\mu\nu}V_\mu K^\nu = 0\}$, which, since g is nondegenerate, is a subspace of codimension one. This, in turn, means that the coimage of the principal symbol is one-dimensional and seen to be spaned by $\tilde{V}_\nu = g_{\mu\nu}V_\mu$.

Consequently, the characteristic hypersurfaces coincide again with the g-null hypersurfaces. We have already found the 3-dim. null-subspace and can now choose three constraint vectors K_i to span this space. We choose $K_1^\mu = \partial_\mu z$ and some K_2, K_3 such that we end up with an orthogonal frame. Since K_1 is the 'old constraint'[44], we will pick e.g. $K_2^\mu =: V_\mu$ with $g^{\mu\nu}\partial_\mu z V_\nu = 0$, and get

$$\begin{aligned} \left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{\sigma\epsilon}K^\sigma K^\epsilon &= 2\zeta^\lambda\Phi_\lambda\frac{1}{\zeta_z}\frac{\partial z}{\partial x_\mu}\frac{\partial z_i}{\partial x_\nu}F^{\mu\nu}_{\sigma\epsilon}K_2^\sigma K_2^\epsilon \\ &= \zeta^\lambda\Phi_\lambda\frac{1}{\zeta_z}\frac{\partial z}{\partial x_\mu}\frac{\partial z_i}{\partial x_\nu}\sqrt{-g}(g^{\mu\nu}g^{\sigma\epsilon} - g^{\mu\epsilon}g^{\sigma\nu})V_\sigma V_\epsilon \end{aligned}$$

[44] Yielding, as is easily seen, a constraint that is linear in the inst. momenta

$$= \zeta^\lambda \Phi_\lambda \frac{1}{\zeta_z} \frac{\partial z}{\partial x_\mu} \frac{\partial z_i}{\partial x_\nu} \sqrt{-g} g^{\mu\nu} g^{\sigma\epsilon} V_\sigma V_\epsilon$$

Thus with the same special choice of coordinates z_i from section 8.6, we obtain

$$\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{\sigma\epsilon} K^\sigma K^\epsilon = \zeta^\lambda \Phi_\lambda \frac{1}{\zeta_z} \sqrt{-g} g^{\sigma\epsilon} V_\sigma V_\epsilon \delta^{i1}$$

Analogous to Equation 32, we can simply solve the now ordinary differential equation Equation 30 and obtain the solution

$$L^{\overline{E}} = c(z_2, z_3) \left(\zeta^\lambda \Phi_\lambda \frac{1}{\zeta_z} \sqrt{-g} g^{\sigma\epsilon} V_\sigma V_\epsilon\right)^{-\frac{1}{2}}$$

Whence, again, we can not make a general prediction but need to investigate the specific spacetime.

8.7.0.7. ”Half De Sitter”

Using the calculations from the end of section 8.6 and setting $V = \frac{\partial}{\partial z_2}$, we find

$$\begin{aligned} L^{\overline{E}} &= c(z_2, z_3) \left(\frac{\zeta^0 + \frac{1}{\ln \sqrt{z z_1}} \zeta^1}{\ln \sqrt{z z_1} \zeta^0 - \zeta^1} \left(\ln \sqrt{z z_1}\right)^{-1} \ln \sqrt{z z_1} \right)^{-\frac{1}{2}} \\ &= c(z_2, z_3) \left(\frac{\zeta^0 + \frac{1}{\ln \sqrt{z z_1}} \zeta^1}{\ln \sqrt{z z_1} \zeta^0 - \zeta^1} \right)^{-\frac{1}{2}} \xrightarrow{z_y \to \infty} \infty \end{aligned}$$

as in section 8.6.

8.8. More General

In order to also account for theories with covariant derivatives in the Lagrangian, where the Christoffel symbols do not cancel, or theories that are coupled, we need to add a new, single derivative term. The Lagrangian now reads:

$$L = \int_X \left(F^{\mu\nu}_{AB}(x)\partial_\mu\phi^A\partial_\nu\phi^B + G^\mu_{AB}(x)\partial_\mu\phi^A\phi^B + H(\phi, x) \right) d^{n+1}x$$

The change has no effect on the top-degree derivative term in the E-L-Equations. In fact, the procedure for the new Lagrangian is so similar to the 'old' procedure that we will be able to abbreviate calculations tremendously using obvious short-hand notation.

$$\begin{aligned} L_{z,\zeta}(\varphi, \dot\varphi) &= \int_\Sigma \left((...) + \frac{1}{\zeta_z}\frac{\partial z}{\partial x^\mu} G^\mu_{AB}\left(\frac{\partial\phi^A}{\partial z}\right)\phi^B + (...) \right) \zeta^\sigma\Phi_\sigma d^3z \\ &= \int_\Sigma \left((...) + \frac{1}{\zeta_z}\frac{\partial z}{\partial x^\mu} G^\mu_{AB}\left(\dot\varphi - S^A\right)\varphi^B + (...) \right) \zeta^\sigma\Phi_\sigma d^3z \end{aligned}$$

From which we read off the instantaneous momenta

$$\pi_A = \left(\frac{1}{\zeta_z}\right)^2 (...) + \frac{1}{\zeta_z}(...) + \underbrace{\zeta^\sigma\Phi_\sigma\frac{1}{\zeta_z}\frac{\partial z}{\partial x^\mu}G^\mu_{AB}}_{:=G_{AB}}\varphi^B$$

And choosing $\overline{E}$ and K^A as before,

$$\begin{aligned} \pi_{\overline{E}} &= \frac{1}{K^{\overline{E}}}\left(-\pi_E K^E + \Xi^i_{AB}K^A D_i\varphi^B + G_{AB}\varphi^B K^A \right) \\ \tilde L_{\overline{E}} &= \frac{1}{K^{\overline{E}}}\left((...) + G_{AB}L^B K^A \right) \\ \omega\left((L,\tilde L),(M,\tilde M)\right) &= \langle L^A, \tilde M_A\rangle - \langle M^A, \tilde L_A\rangle \\ &= (...) + \langle L^{\overline{E}}, \frac{1}{K^{\overline{E}}}G_{AB}K^A M^B\rangle - \langle M^{\overline{E}}, \frac{1}{K^{\overline{E}}}G_{AB}K^A L^B\rangle \\ &= (...) + \langle \frac{1}{K^{\overline{E}}}G_{AE}K^A L^{\overline{E}}, M^E\rangle + \langle \frac{1}{K^{\overline{E}}}G_{AE}K^A L^E, M^{\overline{E}}\rangle \\ &\overset{!}{=} 0 \ , \ \forall (M,\tilde M) \in T\mathcal{P}^0_z \end{aligned}$$

Hence the old condition from the $\tilde M^E$-term does not change, while the M^E- and $M^{\overline{E}}$-terms pick up additional terms. The final equations reads

$$0 = \left[D_i\left(\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AB}K^A K^B \right) + 2\left(\frac{1}{K^{\overline{E}}}\right)^2 G_{AE}K^A K^E \right] L^{\overline{E}} + 2\left[\left(\frac{1}{K^{\overline{E}}}\right)^2 \Xi^i_{AB}K^A K^B \right] D_i L^{\overline{E}}$$

8.8.1. Charged Scalar Field

$$
\begin{aligned}
\mathcal{L} &= \sqrt{-g}\left(g^{\mu\nu}(\mathcal{D}_\mu\phi)(\mathcal{D}_\nu\phi)^* - V(\phi)\right)\ ,\ \mathcal{D}_\mu = (\partial_\mu - ieA_\mu) \\
&= \sqrt{-g}g^{\mu\nu}\left(\delta_{AB}\partial_\mu\phi^A\partial_\nu\phi^B + 2\epsilon_{AB}\partial_\mu\phi^A A_\nu\phi^B\right) + (\ldots)\ ,\ A = 1,2
\end{aligned}
$$

where we substituted $\phi = \phi^1 + i\phi^2$ and neglected non-derivative terms. We immediately read off $F^{\mu\nu}_{AB} = \sqrt{-g}g^{\mu\nu}\delta_{AB}$ and $G^{\mu}_{AB} = 2\sqrt{-g}g^{\mu\nu}A_\nu\epsilon_{AB}$.

This section was just meant to show that a first generalization to more general Lagrangians is straight forward and to serve as a starting point for further investigation, which, however, will not be carried out in this work.

9. Conclusion

In the introduction to section 8, we have noted that "well-posedness" of a theory with gauge freedom of course means well-posedness of the associated Cauchy-problem *after* fixing gauges. Following the preceding discussion it does not seem all too obvious that the gauge freedom (i.e. the gauge groups) should stay the same on a characteristic hypersurface. Proposition 8.3.6 points in this direction and at least for the vacuum Maxwell field on a Minkowski background, this is seen to hold (cp. subsection A.2). However, if the characteristic constraint gives rise to a symplectic submanifold, then at least it does not contribute to gauge freedom. We have shown that there should be a large class of theories for which the char. constr.submfd. is symplectic. Thus, a next question that could be tackled may be formulated as:

Assuming the characteristic constraint submanifold to be symplectic, under which conditions do we recover the same gauge groups as on a non-characteristic hypersurface and does fixing these gauges alone give rise to a well-posed system as in the non-char. case?

A. Appendix

Here, we are including calculations done in the very beginning to get a feel for the way [GM06] behaves on characteristic initial value hypersurfaces. Essentially, we calculated every object defined in [GM06] for the example of a Klein Gordon field on a light cone and electromagnetism on a characteristic hypersurface, both on a Minkowski background.

A.1. Klein Gordon on a Light Cone

$$X = \mathbb{R}^2$$
$$Y = \pi_{\mathbb{R}^2,\mathbb{R}} \ , \ J^1(Y) = \pi_{\mathbb{R}^2,\mathbb{R}^3}$$

Coordinates on $J^1(Y)$ are denoted $(x^+, x^-, \phi, \phi_+, \phi_-)$.

$$\mathcal{L} = \left(\phi_+\phi_- - \frac{1}{2}m\phi^2\right) dx^+ \wedge dx^-$$

$$\theta_{\mathcal{L}} = \phi_- d\phi \wedge dx^+ + \phi_+ d\phi \wedge dx^- - \left(\phi_+\phi_- + \frac{1}{2}m\phi^2\right) dx^+ \wedge dx^-$$
$$\Omega_{\mathcal{L}} = d\phi \wedge (d\phi_- \wedge dx^+ + d\phi_+ \wedge dx^- + m\phi dx^+ \wedge dx^-) + (\phi_- d\phi_+ + \phi_+ d\phi_-) \wedge dx^+ \wedge dx^-$$

The Euler-Lagr. equ.: $m\phi + 2\partial_+\partial_-\phi = 0$.

It is easy to see that the Lagrangian is equivariant w.r.t. to the following **group actions**:

$(\mathbb{R}^2, +) \ni (a, b)$ acts on $Y = \pi_{\mathbb{R}^2,\mathbb{R}}$ with

$$\eta_Y((a, b), x^+, x^-, \phi) = (x^+ + a, x^+ + b, \phi)$$
$$\eta_{J^1Y}(...) = (x^+ + a, x^- + b, \phi, ...)$$
$$\xi_Y(x, \phi) = (\xi, 0)$$

and the corresp. Noether current reads:

$$J(\xi)(z) = p(\xi^+ dx_- - \xi^- dx_+) - (p^+\xi^- - p^-\xi^+)d\phi$$
$$(j^1\phi)^* J^{\mathcal{L}}(\xi) = (\partial_-\phi(-\partial_+\phi\xi^+ - \partial_-\phi\xi^-) + L\xi^+)dx^- + (\pm \text{ interchanged})dx^+$$

And, of course, SO(1, 1),[45] i.e.

$$\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \mapsto \frac{1}{2}\begin{pmatrix} \lambda + \frac{1}{\lambda} & \lambda - \frac{1}{\lambda} \\ \lambda - \frac{1}{\lambda} & \lambda + \frac{1}{\lambda} \end{pmatrix}\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$$

[45]The non-char. Lagrangian is equivariant w.r.t. this action.

which in light-cone coordinates diagonalizes:
$\mathrm{GL}(1) = (\mathbb{R}/\{0\}, \cdot) \ni \lambda$ acts on $Y = \pi_{\mathbb{R}^2,\mathbb{R}}$ with

$$\eta_Y(\lambda, x^+, x^-, \phi) = (\lambda x^+, \frac{1}{\lambda}x^-, \phi)$$
$$\eta_{J^1Y}(\lambda, x^+, x^-, \phi, \phi_+, \phi_-) = (\lambda x^+, \frac{1}{\lambda}x^-, \phi, \frac{1}{\lambda}\phi_+, \lambda\phi_-)$$

$$T_e(\mathrm{GL}(1)) \cong T_1(\mathbb{R}/\{0\}) \cong \mathbb{R} \ni \alpha$$
$$\exp(t\alpha) = e^{t\alpha} \Rightarrow \alpha_Y(x, \phi) = (\alpha x^+, -\alpha x^-, 0)$$

and the corresp. Noether current reads:

$$J(\alpha)(z) = \alpha p(x^+dx^- + x^-dx^+) + \alpha(p^+x^- + p^-x^+)d\phi$$
$$(j^1\phi)^* J^{\mathcal{L}}(\alpha) = \Big[\partial_-\phi - (\partial_-\phi x^-\alpha - \partial_+\phi x^+\alpha) + \alpha x^+ L\Big]dx^- - \Big[\partial_+\phi(\partial_-\phi x^-\alpha - \partial_+\phi x^+\alpha) - \alpha x^- L\Big]dx^+$$

We check: $0 = d(j^1\phi)^* J^{\mathcal{L}}(\alpha) = d(j^1\phi)^* J^{\mathcal{L}}(\xi)$.

Slicing: Σ is a 1-mfd., $\mathrm{Emb}(\Sigma, X) \ni \tau$ a parametrized curve.

$$\mathcal{Y}_\tau \ni \varphi\ (= \phi|_{\Sigma_\tau})$$
$$\mathcal{Z}_\tau \ni (\varphi, p, p^\mu)$$
$$T^*\mathcal{Y}_\tau \ni (\varphi, \pi)$$

We verify that a slicing generated by $\zeta^\mu \frac{\partial}{\partial x^\mu}$ is Lagrangian, i.e. equivariant w.r.t. its flow, iff

$$0 = \frac{1}{2}m\phi^2(\zeta^+_{,+} + \zeta^-_{,-}) + \phi_+\phi_+\zeta^+_{,-} + \phi_-\phi_-\zeta^-_{,+}$$

However, from hereon we will choose a slicing with $x^+ =$ const., which amounts to working in adapted coordinates.

$$\dot\varphi := \pounds_\zeta\phi|_{\Sigma_\tau} := \Big(\zeta^\mu\phi_{,\mu} - \zeta \circ \phi\Big)|_{\Sigma_\tau} = (\zeta^+\partial_+\phi + \zeta^-\partial_-\phi)\,|_{\Sigma_\tau}$$

Now, choosing $\Sigma_\tau = \Sigma_{x^+} := \{x^+ = \text{const.}\}$ we calculate:

$$L_{x^+,\zeta}(\varphi, \dot\varphi) = \int_{\Sigma_\tau} i^*_{x^+} i_{\zeta_X}\mathcal{L}(j^1\phi(x))\ ,\ i_{x^+} : \Sigma_{x^+} \to X \text{ inclusion}$$
$$= \int_{\Sigma_{x^+}} \left[D\varphi\left(\frac{\dot\varphi - \zeta^- D\varphi}{\zeta^+}\right) - \frac{1}{2}m\varphi^2\right]\zeta^+\,dx^-,\ \text{where } D\varphi := \partial_-\varphi$$

$$= \int_{\Sigma_{x^+}} \left[D\varphi(\dot{\varphi} - \zeta^- D\varphi) - \frac{1}{2}\zeta^+ m\varphi^2 \right] dx^-$$

$$\pi := \frac{\partial \mathcal{L}_{x^+,\zeta}(j^1\varphi, \dot{\varphi})}{\partial \dot{\varphi}} d\varphi \otimes dx^- = D\varphi \, d\varphi \otimes dx^-$$

$$\mathcal{P}_{x^+} := \operatorname{Im} \mathbb{F}\mathcal{L}_{x^+,\zeta} = \operatorname{Im}\Big((\varphi, \dot{\varphi}) \mapsto (\varphi, D\varphi)\Big) = \Big\{(\varphi, \pi) \in T^*\mathcal{Y}_{x^+} \;\Big|\; \pi - D\varphi = 0\Big\}$$

Where, by abuse of notation, we also denote by π its density function. We note that $\mathcal{P}_{x^+}$ is a linear subspace of $T^*\mathcal{Y}_{x^+}$. The same calculation shows that adding more spacetime dimensions or scalar fields to the Klein-Gordon Lagrangian has no effect on the primary constraint set on a null hyperplane.

$$\begin{aligned} H_{x^+,\zeta}(\varphi, \pi) &: = \langle \pi, \dot{\varphi} \rangle - L_{x^+,\zeta}(\varphi, \dot{\varphi}) \\ &= \int_{\Sigma_{x^+}} \left(\zeta^-(D\varphi)^2 + \frac{1}{2}\zeta^+ m\varphi^2 \right) dx^- \end{aligned}$$

The **Space of Cauchy data**: We choose $\mathcal{Y}_{x^+} = H^1(\Sigma_{x^+})$ as the space of initial data[46] and assume $\zeta^\pm$ to be bounded functions. Since $H^1(\Sigma_{x^+})$ is a complete vector space, we know that $\frac{d}{d\lambda} f_\lambda(x) \in H^1(\Sigma_{x^+})$ for any continuous path f_λ in $H^1(\Sigma_{x^+})$. Hence, $T_\varphi \mathcal{Y}_{x^+} = H^1(\Sigma_{x^+})$. By the Riesz representation theorem we may then identify $T^*_\varphi \mathcal{Y}_{x^+} \cong L^2(\Sigma_{x^+})$. Finally, $T_{(\varphi,\pi)}(T^*\mathcal{Y}_{x^+}) \cong T(T_\varphi \mathcal{Y}_{x^+}) \oplus T_\pi(T^*_\varphi \mathcal{Y}_{x^+}) \cong H^1(\Sigma_{x^+}) \oplus L^2(\Sigma_{x^+})$.
If we identify the tangent space with the space of variational derivatives, then, because taking variational derivatives before and after substituting $D\varphi = \pi$ needs to agree,

$$T\mathcal{P}_{x^+} = \left\{ X \;\Big|\; X = f\frac{\delta}{\delta\varphi} + Df\frac{\delta}{\delta\pi} \;,\; f : \mathcal{P}_{x^+} \to H^1(\Sigma_{x^+}) \right\}$$

Meaning $\delta\varphi = f$ and $\delta\pi = Df$.

$$\omega_{\mathcal{P}_{x^+}} := i^*\omega_{x^+} \;,\; i : \mathcal{P}_{x^+} \to T^*\mathcal{Y}_{x^+} \text{ inclusion}$$

That is, for $(f, Df), (g, Dg) \in T\mathcal{P}_{x^+}$,

$$\begin{aligned} \omega_{x^+}((f, Df), (g, Dg)) &= \langle f, Dg \rangle_{L^2} - \langle g, Df \rangle_{L^2} \\ &= -2\langle Df, g \rangle_{L^2} \\ &= 0 \;\forall g \in H^1(\Sigma_{x^+}) \Leftrightarrow Df = 0 \text{ a.e.} \end{aligned}$$

Whence $(\mathcal{P}_{x^+}, \omega_{x^+})$ (ω_{x^+} now denoting the pulled back form) is a symplectic subspace of $T^*\mathcal{Y}_{x^+}$ and the constraint algorithm terminates. Since the $\delta/\delta\phi$-part fixes the $\delta/\delta\pi$ part, we shall henceforth drop the latter.

[46]Cf. Lars-ErikLundberg. The Klein-Gordon Equation with light-cone data. Communications in Mathematical Physics, 62:107-118, 1978

Check: Upon choosing the usual evolution direction, $\zeta = \zeta^+ \frac{\partial}{\partial x^+}$, the Hamiltonian equations of motion yield for $X = \frac{d\varphi_{x^+}}{dx^+} \frac{\delta}{\delta\varphi}$:

$$\begin{aligned}
i_X \omega_{x^+} &= dH_{x^+,\zeta^+} \\
\Leftrightarrow \omega_{x^+}\left(\frac{d\varphi_{x^+}}{dx^+}, g\right) &= g\frac{\delta}{\delta\varphi}\big[H_{x^+,\zeta^+}\big],\ \forall g \in T\mathcal{P}_{x^+} \\
\Leftrightarrow -2\left\langle D\frac{d\varphi_{x^+}}{dx^+}, g\right\rangle_{L^2} &= m\langle\varphi, g\rangle_{L^2} \\
\Leftrightarrow -2D\frac{d\varphi_{x^+}}{dx^+} &= m\varphi \quad \text{a.e.}
\end{aligned}$$

Calculating the **instantaneous energy-momentum map**:

- $(\mathbb{R}^2, +)$:

$$\begin{aligned}
\langle E_{x^+}(\sigma), \xi\rangle :&= \int_{\Sigma_{x^+}} \sigma^*\langle J, \xi\rangle\ ,\ \sigma \in \mathcal{Z}_{x^+} \\
&= \int_{\Sigma_{x^+}} (p\xi^+ - (p^+\xi^- - p^-\xi^+)\, D\varphi)\, dx^-
\end{aligned}$$

$$\begin{aligned}
\langle \mathcal{E}_{x^+}(\varphi,\pi), \xi^-\frac{\partial}{\partial x^-}\rangle &= \langle \mathcal{J}_{x^+}(\varphi,\pi), \xi^-\frac{\partial}{\partial x^-}\rangle\ ,\ \text{as } \xi^-\frac{\partial}{\partial x^-} \text{ is everywhere tangent to } \Sigma_{x^+} \\
&= -\xi^- \int_{\Sigma_{x^+}} (D\varphi)^2 dx^- \\
\langle \mathcal{E}_{x^+}(\varphi,\pi), \xi^+\frac{\partial}{\partial x^+}\rangle &= -H_{x^+,\xi^+}\ ,\ \text{as } \xi^+\frac{\partial}{\partial x^+} \text{ is everywhere transverse to } \Sigma_{x^+} \\
&= -\frac{1}{2}m\xi^+ \int_{\Sigma_{x^+}} \varphi^2 dx^-
\end{aligned}$$

- GL(1):

$$\begin{aligned}
\langle E_{x^+}(\sigma), \alpha\rangle :&= \int_{\Sigma_{x^+}} \sigma^*\langle J, \alpha\rangle\ ,\ \sigma \in \mathcal{Z}_{x^+} \\
&= \alpha\int_{\Sigma_{x^+}} \Big(px^+ + (p^+x^- + p^-x^+)D\varphi\Big)\, dx^- \\
\langle \mathcal{E}_{x^+}(\varphi,\pi), \alpha\rangle &= \alpha\int_{\Sigma_{x^+}} \left(-\frac{1}{2}m\varphi^2 x^+ + (D\varphi)^2 x^-\right) dx^- \ \text{(diverges i.g.)}
\end{aligned}$$

No **Gauge Groups**:

- ξ_Y is constant, hence it can not be localizable

- $\alpha_Y(x, \phi) = (\alpha x^+, -\alpha x^-, 0)$ is obviously not localizable

$\Rightarrow$ No Gauge groups i.t.s.o. [GM06].
As $C_{x^+} = \mathcal{P}_{x^+}$ is symplectic, we have $\mathfrak{X}(\mathcal{P}_{x^+})^\perp = \emptyset$ (w.r.t. itself), so that the fullness assumption is trivially satisfied.

We can still calculate the **instantaneous generators**:

$$\xi_{\mathcal{P}^+_x} \lrcorner \omega_{x^+} - d\langle \mathcal{E}_{x^+}, \xi \rangle = 0$$

Yielding:

$$\begin{aligned}
\xi^-_{\mathcal{P}_{x^+}} &= -\xi^- D\varphi \frac{\delta}{\delta\varphi} \ \text{ a.e.} \\
\xi^+_{\mathcal{P}_{x^+}} &= f\frac{\delta}{\delta\varphi}\ ,\ Df = \frac{1}{2} m\xi^+\varphi \ \text{ a.e.} \\
\alpha_{\mathcal{P}_{x^+}} &= (f_1 + f_2)\frac{\delta}{\delta\varphi}\ ,\ Df_1 = -\frac{1}{2} m\alpha x^+\varphi\ ,\ f_2 = \alpha x^- D\varphi \ \text{ a.e.}
\end{aligned}$$

Vanishing Theorem (this is posed for any hypersurface, i.p. characteristic ones):

- $(\mathbb{R}^2, +)$:

$$\begin{aligned}
\int_{\Sigma_\tau} \tau^*(j^1\phi)^* J^{\mathcal{L}}(\xi) &= \int_{\Sigma_\tau} \tau^* \left[(\partial_-\phi(-\partial_+\phi\xi^+ - \partial_-\phi\xi^-) + L\xi^+)dx^- + (\pm \text{ interchanged })dx^+\right] \\
&= -\int_{x^+ = \text{ const.}} \left((D\varphi)^2\xi^- + \frac{1}{2}\xi^+ m\varphi^2\right) dx^- \ (\text{ taking } \Sigma_\tau = \{x^+ = \text{ const.}\}) \\
&= 0 \ \Leftrightarrow \varphi = 0 \ \text{ a.e.}
\end{aligned}$$

  As expected, for this group action is not localizable.

- GL(1)

$$\begin{aligned}
\int_{\Sigma_\tau} \tau^*(j^1\phi)^* J^{\mathcal{L}}(\xi) &= \alpha \int_{\Sigma_{x^+}} \left((D\varphi)^2 x^- - D\varphi\partial_+\phi x^+ + x^+ L\right) dx^- \\
&= \alpha \int_{\Sigma_{x^+}} \left((D\varphi)^2 x^- - \frac{1}{2} x^+ m\varphi^2\right) dx^- \\
&= 0 \ \Leftrightarrow \varphi = 0 \ \text{ a.e.}
\end{aligned}$$

  Again, this group action is not localizable and in this case the integral even diverges in general - as expected from the Poincare group.

Primary Constraints and the Momentum Map:
In this case:

$$\mathfrak{p}_{x^+} := \left\{ \xi \in \mathfrak{g}_{x^+} \,\middle|\, \xi_{Y_{x^+}} = 0 \text{ and } [\xi, \mathfrak{g}] \subset \mathfrak{g}_{x^+} \right\} = \emptyset$$

i.e. there exist no (non-triv.) infinitesimal generators ξ of the above symmetry groups satisfying:

$$\xi^\mu|_{Y_{x^+}} = 0 = \xi^A|_{Y_{x^+}} \text{ and } \xi^+_{,+}|_{Y_{x^+}} = 0$$

Hence $\dot{\mathfrak{g}}_{x^+} = \emptyset$, whence for $\dot{\mathcal{J}}_{x^+} : T^*\mathcal{Y}_{x^+} \to \dot{\mathfrak{g}}^*_{x^+}$ we can only define $\dot{\mathcal{J}}^{-1}_{x^+}(0) = T^*\mathcal{Y}_{x^+}$. Because our primary constraint is not linear in π, whereas $\dot{\mathcal{J}}_{x^+}$ is, the construction could not have worked to begin with.
As $\mathfrak{g}_{x^+} = \{\xi^- \frac{\partial}{\partial x^-}\}$, we get

$$\begin{aligned}\langle \mathcal{J}_{x^+}(\varphi, \pi), \xi^- \rangle = \langle \pi, \xi^-_{\mathcal{Y}_{x^+}}(\varphi) \rangle &= \int_{\Sigma_{x^+}} \pi\left(\xi^-_{\mathcal{Y}_{x^+}}(\varphi)\right) \\ &= \int_{\Sigma_+} \pi_A((\xi^-)^A \circ \varphi - \varphi^A_{,i}(\xi^-)^i)\, dx^- = \xi^- \int_{\Sigma_{x^+}} \pi D\varphi\, dx^-\end{aligned}$$

And as stated already in [GM06, p.188], the zero set of $\mathcal{J}_{x^+}$ need not vanish on $\mathcal{P}_{x^+}$. Actually, $\mathcal{J}^{-1}_{x^+}(0) \cap \mathcal{P}_{x^+} = \{D\varphi = 0\}$.
Since (p.197): $\mathcal{J}^{-1}_{x^+}(0) \subset \dot{\mathcal{J}}^{-1}_{x^+}(0)$, our results for both zero sets are in agreement and we could not have gotten $\mathcal{P}_{x^+}$ as a zero set of $\dot{\mathcal{J}}_{x^+}$.

A.2. Vacuum Maxwell on a null hypersurface

- We are considering the case in $X = \mathbb{R}^4$ with a Minkowski background metric, i.e. the theory is not parametrized.

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} d^4x$$

$\mathcal{G} = C^\infty(X) \ni f$ acts on $Y = \Lambda^1 X \ni A$ as $(f \cdot A)(x) = A(x) + df(x)$
For $\chi \in \mathfrak{g} \cong C^\infty(X)$,

$$\chi_Y = \chi_{,\nu} \frac{\partial}{\partial A_\nu}$$

The action is obviously vertically transitive. We have $J : Z \to \mathfrak{g}^*$,

$$\langle J(x^\mu, A_\nu, p, \mathcal{F}^{\mu\nu}), \chi \rangle = \mathcal{F}^{\mu\nu} \chi_{,\nu}\, d^3x_\mu$$

$$(j^1A)^* J^{\mathcal{L}}(\chi) = (j^1A)^* \left(F^{\mu\nu} \chi_{,\nu}\, d^3x_\mu\right) = (A^{\mu,\nu} - A^{\nu,\mu}) \chi_{,\nu}\, d^3x_\mu$$

$$d\left[(j^1A)^* J^{\mathcal{L}}(\chi)\right] = \partial_\mu (A^{\mu,\nu} - A^{\nu,\mu}) \chi_{,\nu}\, d^4x$$

$$= 0 \ \forall \chi \in \mathfrak{g} \Leftrightarrow \partial_\mu \left(A^{\mu,\nu} - A^{\nu,\mu} \right) = 0$$

which are simply the E-L-Equ. (cp. Vanishing Thm.).[47]
We take $\Sigma_{x^+} := \{x^+ = \text{const.}\}$ and denote $T^*\mathcal{Y}_{x^+} \ni (A_\nu, \mathfrak{E}^\nu)$ with the embedding

$$\begin{aligned} i_{x^+} :& \Sigma \to \Sigma_{x^+} \subset X \\ & \begin{pmatrix} x^- \\ x^i \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{2}(x^+ - x^-) \\ \frac{1}{2}(x^+ + x^-) \\ x^i \end{pmatrix} , \ i = 2,3 \\ s :& \Sigma \times \mathbb{R} \xrightarrow{\cong} X , \ ((x^-, x^i), x^+) \mapsto i_{x^+}(x^-, x^i) \end{aligned}$$

Any slicing of X induces a slicing of Y by 'push forward':
If $x = s(x^+, x^-, x^i)$ and $A_x \in \Lambda^1_x X$, then

$$s\left(s^{-1}(x) - (\epsilon, 0, 0)\right)^* A_x \in \Lambda^1_{\tilde{x}} X$$

where $\tilde{x} = s(x^+ + \epsilon, x^-, x^i)$.
Hence the most general generator of a slicing reads

$$\zeta_Y = \zeta^\mu \frac{\partial}{\partial x^\mu} + \left(\chi_{,\alpha} - A_\nu \zeta^\nu{}_{,\alpha}\right) \frac{\partial}{\partial A_\alpha}$$

where ζ_X must be a Killing vector field for the slicing to be Lagrangian. In our Minkowski case we thus require ζ_X to be a generator of the Poincare group[48].

$$\begin{aligned} \dot{A}_\mu &= \left[\zeta^0 D_0 A_\mu + \zeta^i D_i A_\mu - \chi_{,\mu} + A_\nu \zeta^\nu{}_{,\mu}\right]\Big|_{\Sigma_{x^+}} \\ &= \zeta^+ D_+ A_\mu + \underbrace{\zeta^- D_- A_\mu + \zeta^m D_m A_\mu - \chi_{,\mu} + A_\nu \zeta^\nu{}_{,\mu}}_{=:S_\mu} , \ m = 2,3 \end{aligned}$$

Note, because S_μ so defined is linear in A_μ, we can simply write $A_\pm$ with the obvious meaning.[49]

$$\begin{aligned} L_{x^+,\zeta}(A, \dot{A}) &= \int_{\Sigma_+} i^*_{x^+} i_{\zeta_X} \mathcal{L} , \ i_{x^+} : \Sigma_+ \to X \text{ inclusion} \\ &= \int_{\Sigma_+} \left[\frac{1}{2}(F_{+-})^2 - (F_{+m})(F_{-m}) - \frac{1}{4}(F_{mn})^2\right] \zeta^+ d^3 x_+ , \ \text{with } D_+ A_\mu := \frac{\dot{A}_\mu - S_\mu}{\zeta^+} \end{aligned}$$

Denoting the instantaneous momenta by $\mathfrak{E}$, we get

$$\mathfrak{E}^+ = 0$$

[47] The $\chi_{,\mu\nu}$-term vanishes due to (anti-) symmetry
[48] If we assume ζ_μ to be bounded, does this exclude all but translations?
[49] We have not been consistent with the sum convention, at some times we sum with η, at others not.

$$\begin{aligned}\mathfrak{E}^m &= -F_{-m}\\ \mathfrak{E}^- &= \left(\frac{\dot{A}_- - S_-}{\zeta^+} - D_- A_+\right)\end{aligned}$$

We verify that the first two equations are constraint equations, whence

$$\mathcal{P}_{x^+} = \{(A, \mathfrak{E}) \in T^+\mathcal{Y}_+ \mid \mathfrak{E}^+ = 0 \text{ and } \mathfrak{E}^m = -F_{-m}\}$$

Because variational derivatives need to agree and $\mathfrak{E}^m = D_m A_- - D_- A_m$, we have

$$T_{(A,\mathfrak{E})}\mathcal{P}_+ = \left\{ L_\mu \frac{\delta}{\delta A_\mu} + M_\mu \frac{\delta}{\delta \mathfrak{E}^\mu} \,\middle|\, M_+ = 0\ ,\ M_m = D_m L_- - D_- L_m \right\}$$

for $L_\mu, M_\mu : \mathcal{P}_+ \to C^\infty(\Sigma_+)$.

Let $(L, \tilde{L}), (G, \tilde{G}) \in \mathfrak{X}(\mathcal{P}_+)$. We calculate, denoting by $\langle \cdot, \cdot \rangle$ the L^2 inner product

$$\begin{aligned}\omega\left((L, \tilde{L}), (G, \tilde{G})\right) &= \langle L_-, \tilde{G}_- \rangle + \langle L_m, \tilde{G}_m \rangle - \langle G_-, \tilde{L}_- \rangle - \langle G_m, \tilde{L}_m \rangle\\ &= \langle L_-, \tilde{G}_- \rangle - \langle G_-, \tilde{L}_- \rangle - \langle D_m L_m, G_- \rangle + 2\langle D_- L_m, G_m \rangle - \langle D_m L_-, G_m \rangle\\ &\overset{!}{=} 0\ \forall (G, \tilde{G}) \in \mathfrak{X}(\mathcal{P}_+)\end{aligned}$$

$\Rightarrow L_- \overset{!}{=} 0$ for $\tilde{G}_-$ is independent of the other components. We have

$$\begin{aligned}&0 \overset{!}{=} -\langle D_m L_m + \tilde{L}_-, G_- \rangle + \langle 2D_- L_m, G_m \rangle\\ &\Rightarrow D_- L_m = 0 \text{ a.e.}\\ &\Rightarrow \tilde{L}_- = 0 \text{ a.e.}\\ &\Rightarrow \mathfrak{X}(\mathcal{P}_+)^\perp = \operatorname{Span}\left\{\frac{\delta}{\delta A_+}\right\}\end{aligned}$$

We note that the constraint due to the characteristic slicing does not lower the rank of $\omega_{\mathcal{P}_+} = \omega_+$ (by abuse of notation).

$$\begin{aligned}H_{x^+,\zeta} &= \int_{\Sigma_+} \langle \mathfrak{E}, \dot{A} \rangle - \mathcal{L}_{x^+,\zeta}\\ &= \int_{\Sigma_+} \mathfrak{E}^- \left((\mathfrak{E}^- + D_- A_+)\zeta^+ + S_-\right) - F_{-m}\dot{A}_m - \frac{1}{2}\zeta^+(F_{+-})^2 + (F_{+m})(F_{-m})\zeta^+ + \frac{1}{4}(F_{mn})^2\zeta^+\\ &= \int_{\Sigma_+} \frac{1}{2}\zeta^+(\mathfrak{E}^-)^2 + \mathfrak{E}^- \left(D_- A_+ \zeta^+ + S_-\right) - F_{-m}\left(S_m + \zeta^+ D_m A_+\right) + \frac{1}{4}(F_{mn})^2\zeta^+\end{aligned}$$

Since $\mathfrak{X}(\mathcal{P}_+)^\perp = \operatorname{Span}\{\frac{\delta}{\delta A_+}\}$, we require $\forall L : \mathcal{P}_+ \to C^\infty(\Sigma)$

$$0 \overset{!}{=} L \frac{\delta}{\delta A_+}\left[H_{x^+,\zeta}\right]$$

$$
\begin{aligned}
&= \int_{\Sigma_+} \mathfrak{E}^-\zeta^+(D_L) + \mathfrak{E}^-(D_-\zeta^+)L - F_{-m}(D_m\zeta^+)L - F_{-m}\zeta^+D_mL \\
&= \int_{\Sigma_+} \zeta^+ L(D_mF_{-m} - D_-\mathfrak{E}^-)d^3x_+ \\
&\Rightarrow D_mF_{-m} - D_-\mathfrak{E}^- \overset{!}{=} 0 \text{ a.e. (Secondary Constraint)} \\
&\Rightarrow \mathcal{P}^2_{x^+,\zeta} = \left\{(A,\mathfrak{E}) \in \mathcal{P}_{x^+,\zeta} \;\middle|\; D_mF_{-m} - D_-\mathfrak{E}^- = 0\right\}
\end{aligned}
$$

$$
T_{(A,\mathfrak{E})}\mathcal{P}^2_{x^+,\zeta} = \left\{(L,\tilde{L}) \in T_{(A,\mathfrak{E})}\mathcal{P}_{x^+,\zeta} \;\middle|\; D_mD_-L_m - D_m^2L_- = D_-\tilde{L}_-\right\}
$$

We are searching for $(G,\tilde{G}) \in T_{(A,\mathfrak{E})}\mathcal{P}_{x^+,\zeta}$, s.t.

$$
\begin{aligned}
0 &= \omega_+\left((L,\tilde{L}),(G,\tilde{G})\right) \ , \ \forall (L,\tilde{L}) \in T_{(A,\mathfrak{E})}\mathcal{P}^2_{x^+,\zeta} \\
&= \langle L_-, \tilde{G}_-\rangle - \langle G_-, \tilde{L}_-\rangle - \langle D_mL_m, G_-\rangle + 2\langle D_-L_m, G_m\rangle - \langle D_mL_-, G_m\rangle
\end{aligned}
$$

Setting $G_- = D_-g_-$, for any $g_- : \mathcal{P}_{x^+,\zeta} \to C^\infty(\Sigma_+)$, we have

$$
\begin{aligned}
0 &= \langle \tilde{G}_- - D_m^2g_- + D_mG_m, L_-\rangle + \langle 2D_-D_mg_- - 2D_-G_m, L_m\rangle \\
&\Rightarrow D_mg_- = G_m \text{ a.e.} \\
&\Rightarrow \tilde{G}_- = 0 \text{ a.e.} \\
&\Rightarrow \mathfrak{X}(\mathcal{P}^2_{x^+,\zeta})^\perp = \text{Span}\left\{\frac{\delta}{\delta A_+}\right\} \cup \left\{D_-g\frac{\delta}{\delta A_-} + D_mg\frac{\delta}{\delta A_m} \;\middle|\; g : \mathcal{P}^2_{x^+,\zeta} \to C^\infty(\Sigma_+)\right\} \subset \mathfrak{X}(\mathcal{P}^2_{x^+,\zeta})
\end{aligned}
$$

Again, we require

$$
\begin{aligned}
0 &\overset{!}{=} \left(D_-g\frac{\delta}{\delta A_-} + D_mg\frac{\delta}{\delta A_m}\right)\left[H_{x^+,\zeta}\right] \\
&= \int_{\Sigma_+} \mathfrak{E}^-\left(\zeta^-(D_-D_-g) + \zeta^m(D_mD_-g) + (D_-g)(D_-\zeta^-)\right) \\
&\quad +(D_mD_-g)(S_m + \zeta^+D_mA_+) - F_{-m}(D_g)(D_m\zeta^-) \\
&\quad +\mathfrak{E}^-(D_mg)(D_-\zeta^m) - (D_-D_mg)(S_m + \zeta^+D_mA_+) \\
&\quad -F_{-m}(\zeta^-(D_-D_mg) + \zeta^m(D_mD_mg) + (D_mg)(D_m\zeta^m) \\
&\quad + \underbrace{(D_lD_mg)(D_lA_m - D_mA_l)}_{=0} \\
&= \int_{\Sigma_+} \mathfrak{E}^-\left(\zeta^-(D_-D_-g) + \zeta^m(D_mD_-g) + (D_-g)(D_-\zeta^-) + (D_mg)(D_-\zeta^m)\right) \\
&\quad -F_{-m}\left((D_g)(D_m\zeta^-) + \zeta^-(D_-D_mg) + \zeta^m(D_mD_mg) + (D_mg)(D_m\zeta^m)\right) \\
&= \int_{\Sigma_+} (D_mF_{-m} - D_-\mathfrak{E}^-)(\zeta^-(D_-g) + \zeta^m(D_mg))
\end{aligned}
$$

$$= 0$$

because we are inside $\mathcal{P}^2_{x^+,\zeta}$. As a result we find for the final constraint set $C_{x^+} = \mathcal{P}^2_{x^+,\zeta}$.

$$\begin{aligned}
\langle R_{x^+}(\sigma), V\rangle := &\int_{\Sigma_+} \varphi^*(i_V\sigma)\ ,\ \sigma \in \mathcal{Z}_{x^+},\ \text{i.e. } \sigma : \Sigma_+ \to Z|_{\Sigma_+}\ ,\ \varphi := \pi_{YZ} \circ \sigma\ ,\ V \in T_\varphi \mathcal{Y}_{x^+}\\
= &\int_{\Sigma_+} \sigma^* \pi^*_{YZ}(\mathcal{F}^{\mu\nu} V_\mu d^3x_\nu)\\
= &\int_{\Sigma_+} (\mathcal{F}^{\mu\nu} V_\mu) \circ (\pi_{YZ} \circ \sigma)\sigma^* d^3x_\nu\\
= &\int_{\Sigma_+} (\mathcal{F}^{\mu 0} + \mathcal{F}^{\mu 1} V_\mu \circ (\pi_{YZ} \circ \sigma))d^3x_+
\end{aligned}$$

$$\begin{aligned}
\Rightarrow R_{x^+} : \mathcal{Z}_{x^+} &\to T^*\mathcal{Y}_{x^+}\\
\mathcal{F}^{\mu\nu} &\mapsto \mathfrak{E}^\mu = (\mathcal{F}^{\mu 0} + \mathcal{F}^{\mu 1}) \circ \sigma
\end{aligned}$$

We define $E_{x^+} : \mathcal{Z}_{x^+} \to \mathfrak{g}^*$ by

$$\begin{aligned}
\langle E_{x^+}(\sigma), \chi\rangle :&= \int_{\Sigma_+} \sigma^*\langle J, \chi\rangle\\
&= \int_{\Sigma_+} \sigma^*(\mathcal{F}^{\mu\nu}\chi_{,\nu} d^3x_\mu)\\
&= \int_{\Sigma_+} (\mathcal{F}^{\nu 0} + \mathcal{F}^{\nu 1})\chi_{,\nu} d^3x_+
\end{aligned}$$

Choosing a holonomic lift, σ, we obtain

$$\begin{aligned}
\langle \mathcal{E}_{x^+}(A, \mathfrak{E}), \chi\rangle &= \int_{\Sigma_+} \mathfrak{E}^\nu \chi_{,\nu} d^3x_+\\
&= \int_{\Sigma_+} (\mathfrak{E}^- \chi_{,-} + \mathfrak{E}^m \chi_{,m})d^3x_+
\end{aligned}$$

Having used $\mathfrak{E}^0 = \mathfrak{E}^+ + \mathfrak{E}^-$, $\mathfrak{E}^1 = \mathfrak{E}^+ - \mathfrak{E}^-$, $\partial_0 = \frac{1}{2}(\partial_+ + \partial_-)$.

Noting that $G = G_+$, we have $\mathcal{J}_{x^+} = \mathcal{E}_{x^+}$ and by Cor.7C.3

$$\begin{aligned}
\langle \mathcal{E}_{x^+}(A, \mathfrak{E}), \chi\rangle = \langle \mathfrak{E}, \chi_{\mathcal{Y}_{x^+}}\rangle &= \int_{\Sigma_+} \mathfrak{E}\left(\chi_{,\alpha} \frac{\delta}{\delta A_\alpha}\right)\\
&= \int_{\Sigma_+} \mathfrak{E}^- \chi_{,-} + \mathfrak{E}^m \chi_{,m} \text{ on } \mathcal{P}_{x^+}
\end{aligned}$$

where we used 74.C to obtain $\chi_{\mathcal{Y}_{x^+}}$. Alternatively, because G acts by $(f, (A, \mathfrak{E})) \mapsto (A + df|_{\Sigma_+}, \mathfrak{E})$, we have for $\chi \in \mathfrak{g}$

$$\chi_{T^*\mathcal{Y}_{x^+}} = \chi_+ \frac{\delta}{\delta A_+} + D_-\chi \frac{\delta}{\delta A_-} + D_-\chi \frac{\delta}{\delta A_i} \text{ , where } \chi_+ := \frac{\partial \chi}{\partial x^+}\Big|_{\Sigma_+}$$

We have $\mathcal{P}^2_{x^+} = C_{x^+} = \mathcal{E}^{-1}_{x^+}(0)$ as expected[50] as $\mathfrak{g}_{C_{x^+}}$ is full, i.e.

$$\mathfrak{g}_{C_{x^+}} = \mathfrak{X}(C_{x^+}) \cap \mathfrak{X}(C_{x^+})^\perp$$

where $\mathfrak{g}_{C_{x^+}} := \{\xi_{C_{x^+}} \mid \xi \in \mathfrak{g}\}$ with $\xi_{C_{x^+}}$ given by

$$\left(\xi_{C_{x^+}} \lrcorner \omega_{x^+} - d\langle \mathcal{E}_{x^+}, \xi\rangle\right)\Big|_{C_{x^+}} = 0$$

Explicitely, denoting $\chi_{C_{x^+}} = (F, \tilde{F})$, we require for all $(G, \tilde{G}) \in \mathfrak{X}(\mathcal{P}_{x^+})$:

$$\omega\Big((F, \tilde{F}), (G, \tilde{G})\Big) = (G, \tilde{G})\,[\langle \mathcal{E}_{x^+}, \chi\rangle] \text{ on } C_{x^+}$$

$$\begin{aligned}
\text{l.h.s. } &= \langle F_-, \tilde{G}_-\rangle - \langle G_-, \tilde{F}_-\rangle - \langle D_m F_m, G_-\rangle + 2\langle D_- F_m, G_m\rangle - \langle D_m F_-, G_m\rangle \\
\text{r.h.s. } &= (G, \tilde{G})\left[\int_{\Sigma_+} (\mathfrak{E}^- \chi_{,-} - (D_- A_m - D_m A_-) D_m \chi)\, d^3x_+\right] \\
&= \langle \tilde{G}_-, D_-\chi\rangle - \langle D_- G_m, D_m\chi\rangle + \langle D_m G_-, D_m \chi\rangle \\
&\Rightarrow F_- = D_-\chi \text{ , } F_m = D_m\chi \text{ , } \tilde{F}_- = 0 \\
&\Rightarrow \mathfrak{g}_{C_{x^+}} \text{ is full}
\end{aligned}$$

$$\begin{aligned}
P_{x^+} :&= \Big\{\xi \in \mathfrak{g}_{x^+} \,\Big|\, \xi_{Y_{x^+}} = 0 \text{ and } \underbrace{[\xi, \mathfrak{g}] \subset \mathfrak{g}_{x^+}}_{\text{trivial}}\Big\} \\
&= \Big\{\chi \in C^\infty(X) \,\Big|\, \chi_{Y_{x^+}} = 0\Big\} \\
&= \Big\{\chi \in C^\infty(X) \,\Big|\, \partial_\alpha \chi|_{\Sigma_+} = 0\Big\} \text{ , as } \chi_{Y_{x^+}} = \chi_{,\alpha} \frac{\partial}{\partial A_\alpha}
\end{aligned}$$

$$\begin{aligned}
\dot{\mathfrak{g}}_{x^+} :&= \mathrm{Span}_{C^\infty(Y_{x^+})}\Big\{[\zeta, \xi]_{Y_{x^+}} \,\Big|\, \xi \in P_{x^+}\Big\} \text{ , } \zeta\text{...generator of slicing} \\
&= \mathrm{Span}_{C^\infty(Y_{x^+})}\Big\{\Big[\chi_{Y_{x^+}}, \zeta_{Y_{x^+}}\Big] \,\Big|\, \chi \in P_{x^+}\Big\}
\end{aligned}$$

where in the last step we used a standard identity for lie algebras.

$$\zeta_{Y_{x^+}} = \zeta^\mu \frac{\partial}{\partial x^\mu} + (\varphi_{,\alpha} - A_\nu \zeta^\nu{}_{,\alpha}) \frac{\partial}{\partial A_\alpha} \text{ , for some } \varphi \in C^\infty(\Sigma_+)$$

[50] Check with requirement that constraints be first class. I.p. the char. constraint

$$\chi_{Y_{x^+}} = \chi_{,\alpha}\frac{\partial}{\partial A_\alpha}$$

Finally, we can calculate

$$\begin{aligned}
[\chi_{Y_{x^+}}, \zeta_{Y_{x^+}}] &= -\chi_{,\beta}\zeta^\beta{}_{,\alpha}\frac{\partial}{\partial A_\alpha} - \zeta^\mu\chi_{,\alpha\mu}\frac{\partial}{\partial A_\alpha} \\
&= -\left(\chi_{,\beta}\zeta^\beta{}_{,\alpha} + \zeta^\beta\chi_{,\alpha\beta}\right)\frac{\partial}{\partial A_\alpha} \\
&= -\zeta^\beta\chi_{,\alpha\beta}\frac{\partial}{\partial A_\alpha}\ , \text{ as } \partial_\alpha\chi|_{\Sigma_+} = 0 \\
&= \zeta^+\chi_{,++}\frac{\partial}{\partial A_+}\ , \text{ for the same reason}
\end{aligned}$$

whence

$$\begin{aligned}
\langle \dot{J}_{x^+}(A, \mathfrak{E}), \dot{\chi}_{T^*\mathcal{Y}_{x^+}} \rangle &= -\int_{\Sigma_+} \mathfrak{E}\left(\chi_{,++}\frac{\delta}{\delta A_+}\right) \\
&= \int_{\Sigma_+} \mathfrak{E}^+\chi_{,++} \\
&= 0\forall\chi \Leftrightarrow \mathfrak{E}^+ = 0 \\
&\Rightarrow \mathcal{P}_{x^+} = \dot{J}^{-1}_{x^+}(0) \cap \{\mathfrak{E}^m = -F_{-m}\}
\end{aligned}$$

Statement of Affirmation[51]

I hereby declare that the master thesis submitted was in all parts exclusively prepared on my own, and that other resources or other means (including electronic media and online sources), than those explicitly referred to, have not been utilized.

All implemented fragments of text, employed in a literal and/or analogous manner, have been marked as such.

Patrik Omland	Place, Date

[51] Taken from Universität Kassel MAHE Master program

References

[AMR88] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*. Springer, 1988.

[BSF06] Ernst Binz, Jedrzej Sniatycki, and Hans Fischer. *Geometry of Classical Fields*. Dover Publications, 2006.

[CM70] Paul R. Chernoff and Jerrold E. Marsden. *Properties of Infinite Dimensional Hamiltonian Systems*. Springer-Verlag, 1970.

[GLSW83] M. J. Gotay, Richard Lashof, J. Sniatycki, and Alan Weinstein. Closed forms on symplectic fibre bundles. *Comment. Math. Helvetici*, 58:617–621, 1983.

[GM06] Mark J. Gotay and Jerrold E. Marsden. *Momentum Maps and Classical Fields*. ArXiv: physics/9801019, 2006.

[GNH78] Mark J. Gotay, James M. Nester, and George Hinds. Presymplectic manifolds and the dirac-bergmann theory of constraints. *Journal of Mathematical Physics*, 19(11):2388–2399, 1978.

[Got83] Mark J. Gotay. On the validity of dirac's conjecture regarding first-class secondary constraints. *J.Phys.A: Math. Gen.*, 16:L141–L145, 1983.

[HE75] S.W. Hawking and G.F.R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics, 1975.

[Hoe83] Lars Hoermander. *The Analysis of Linear Partial Differential Operators 2*. Springer-Verlag, 1983.

[Lee12] John M. Lee. *Introduction to smooth manifolds*. Springer, 2012.

[Lun78] Lars-Erik Lundberg. The klein-gordon equation with light-cone data. *Communications in Mathematical Physics*, 62:107–118, 1978.

[MPS98] Jerrold E. Marsden, George W. Patrick, and Steve Shkoller. Multisymplectic geometry, variational integrators, and nonlinear pdes. *Communications in Mathematical Physics*, 199:351–395, 1998.

[Pal68] Richard S. Palais. *Foundations of Global Non-Linear Analysis*. W.A. Benjamin, 1968.

[Rau97] Jeffrey Rauch. *Partial Differential Equations*. Springer, 1997.

[Sei12] Werner M. Seiler. *Involution: The formal theory of partial differential equations*. Springer, 2012.

[ST95] Werner M. Seiler and Robin W. Tucker. Involution and constrained dynamics i: The dirac approach. *Journal of Physics A: Math. Gen.*, 28:4431–4451, 1995.

[Ste80] Paul Joseph Steinhardt. Problems of quantization in the infinite momentum frame. *Annals of Physics*, 128:425–447, 1980.

[Sun82] Kurt Sundermeyer. *Constrained Dynamics*. Springer-Verlag, 1982.

[Wik13] Wikipedia. As cited. *NA*, NA:NA, 2013.

[zH90] H. Müller zum Hagen. Characteristic initial value problem for hyperbolic systems of second order differential equations. *Annales De L I.H.P, Section A*, 53:2:159–216, 1990.

[zHS79] H. Mueller zum Hagen and H.-J. Seifert. The characteristic initial-value problem in general relativity. In J. Ehlers, editor, *Proceedings of the International School of Physics - Isolated Gravitating Systems in General Relativity*, 1979.