Radboud University Nijmegen

Faculty of Science

Symplectic derivation of the ADM mass in general relativity

Master Thesis in Mathematics

Supervisor:
dr. Annegret Burtscher

Author:
Floris Farro

Second reader:
dr. Peter Hochs

August 20, 2026

Abstract

This thesis studies the ADM energy–momentum using a symplectic framework based on Hamiltonian dynamical systems. The framework provides a systematic passage from a Lagrangian field theory to a Hamiltonian description. We develop the field-theoretic construction of Chruściel–Jezierski–Kijowski and supply arguments for statements left unproved in their work. We also reformulate part of their construction in more geometric terms using jet bundle theory.

We apply this framework to general relativity, with particular emphasis on asymptotically flat spacetimes. We show that the Hamiltonian energy–momentum obtained from the construction agrees with the ADM energy–momentum. We discuss invariance of the Hamiltonian and we compare this framework with the ADM formalism and Christodoulou's Lagrangian approach. Finally, we discuss the positive mass problem and survey recent developments for asymptotically flat, asymptotically hyperbolic, and asymptotically hyperboloidal settings.

Acknowledgements

I would like to express my sincere gratitude to Annegret Burtscher for her guidance, constructive feedback, and support throughout the process of writing this thesis. I am also grateful for her encouragement and advice regarding my future academic endeavours.
I would also like to thank Peter Hochs for serving as the second reader of this thesis, for attending my presentation, and for his support of my plans to pursue a career in mathematics.
Above all, I wish to thank my parents and my girlfriend for their constant support and encouragement. I am also grateful for the many practical ways in which they helped me during the writing process, whether by preparing dinner when I was working late or by providing me with a quiet place to work when the library was closed.

Contents

Chapter 1

Introduction

General relativity is a theory of gravity in which gravity is described in terms of the curvature of spacetime. The Einstein equations determine the behaviour of spacetime and hence of the gravitational field. Spacetime and therefore gravitational fields are described by a Lorentzian metric. Since general relativity is coordinate independent, the components $g_{\mu\nu}$ depend on the chosen coordinate system even though the underlying metric does not. This motivates separating the variables that contain the dynamical information from those that describe the choice of coordinates.

This separation can be achieved by casting the theory into canonical Hamiltonian form. Arnowitt, Deser and Misner [4, 5], commonly abbreviated as ADM, developed such a formulation using a $3+1$ splitting of spacetime. In this decomposition, the canonical variables are given by a spatial metric and its conjugate momentum, while the lapse and shift encode how the spatial hypersurfaces are propagated through spacetime. The ADM formulation was developed for asymptotically flat spacetimes, whose metric approaches the Minkowski metric at spatial infinity.

The ADM formulation also yields a Hamiltonian and a notion of total energy and momentum, known as the ADM energy-momentum. The corresponding invariant mass is called the ADM mass. These quantities are defined by integrals at spatial infinity and measure how the spacetime metric deviates from the Minkowski metric.

In this thesis, we study the geometric approach to Hamiltonian field theory developed by Chruściel, Jezierski and Kijowski [27]. We first examine the general construction for fields on an evolving hypersurface. We then apply this framework to general relativity, where the resulting Hamiltonian yields the ADM energy-momentum.

Geometric field theory construction

We begin with the geometric framework developed by Chruściel, Jezierski and Kijowski [27, Chapter 3]. Let M be an $(n+1)$-dimensional manifold and let $\pi\colon F \to M$ be a fibre bundle. Sections $f \in \Gamma(F)$ represent the fields of the theory. As a simple example, one may take $F = TM$, in which case the sections $X \in \Gamma(TM) = \mathfrak{X}(M)$ are vector fields.

To describe derivatives of sections in a coordinate-independent way, we use the first jet bundle J^1F, following Saunders [71]. A first-order Lagrangian is a bundle map

$$\mathbf{L}\colon J^1F \to \Lambda^{n+1}T^*M.$$

The dynamically admissible sections, meaning the physically relevant ones, are precisely those that satisfy the Euler–Lagrange equation.

Next, we introduce the momentum bundle

$$P := V^*F \otimes \tilde{\Lambda}^n T^*M,$$

where V^*F denotes the vertical cotangent bundle of F and $\tilde{\Lambda}^n T^*M$ the bundle of odd n-forms; see Chruściel–Jezierski–Kijowski [27, Appendix A]. Dynamically admissible fields can equivalently be described by appropriate sections of this momentum bundle P.

We study the evolution of the fields by describing them relative to a hypersurface moving through M. Let $\Sigma \subset M$ be a hypersurface and define its motion through M by a map

$$\psi \colon I \times \Sigma \to M,$$

such that ψ_t is an embedding for every $t \in I$. Furthermore, let

$$\Psi \colon I \times \Phi \to F$$

be a map covering ψ such that, for every $t \in I$, the map Ψ_t is an isomorphism from the fibres of Φ to those of F. Here we set $\Phi := \psi_0^* F$. We call the pair (ψ, Ψ) a reference system. The corresponding commutative diagram is shown in Figure 1.1.

$$\begin{tikzcd}
I \times \Phi \arrow[r, "\Psi"] \arrow[d, "\mathrm{id}_I \times \mathrm{pr}_{\Phi \to \Sigma}"'] & F \arrow[d, "\mathrm{pr}_{F \to M}"] \\
I \times \Sigma \arrow[r, "\psi"'] & M
\end{tikzcd}$$

Figure 1.1: The reference system

We also define the momentum bundle

$$\Pi := V^*\Phi \otimes \tilde{\Lambda}^n T^*\Sigma.$$

Using the Liouville one-form construction, we obtain an antisymmetric two-form on the space of sections $\Gamma(\Pi)$.

The reference system allows us to pull back dynamically admissible sections $p \in \Gamma(P)$ to sections $\pi \in \Gamma(\Pi)$. We define the phase space $\mathscr{P}_\Sigma$ to consist of these pulled-back sections. We then show that $\mathscr{P}_\Sigma$ has the structure of a dynamical system.

$$\begin{tikzcd}
J^1\Phi \arrow[r, "j^1\Psi_t"] \arrow[d] & J^1F|_{\psi_t(\Sigma)} \arrow[r, "\mathbf{L}"] \arrow[d] & \tilde{\Lambda}^{n+1}T^*M \\
\{t\} \times \Phi \arrow[r, "\Psi_t"] \arrow[d] & F|_{\psi_t(\Sigma)} \arrow[d] & \\
\{t\} \times \Sigma \arrow[r, "\psi_t"] & \psi_t(\Sigma) &
\end{tikzcd}$$

Figure 1.2: The pullback construction for the Lagrangian.

To construct a Hamiltonian dynamical system, we first pull back the Lagrangian $\mathbf{L}$, as illustrated in Figure 1.2. This formulation makes the role of the jet prolongation $j^1\Psi_t$ explicit and is developed in detail in Subsection 4.3.2.

After pulling back the Lagrangian $\mathbf{L} \circ j^1\Psi_t$ to the hypersurface Σ, we apply the Legendre transformation to obtain a Hamiltonian on the phase space $\mathscr{P}_\Sigma$. We show that its variation is related to the antisymmetric two-form Ω_Σ, so that $\mathscr{P}_\Sigma$ acquires the structure of a Hamiltonian dynamical system. The precise construction and the corresponding variation formula are developed in Chapter 4.

Application to general relativity

Our goal is to apply the field theory construction to asymptotically flat spacetimes and to show that the resulting Hamiltonian yields the ADM energy-momentum. Since the hypersurfaces under consideration are not compact, the relevant integrals are not automatically finite. Therefore, we must establish their convergence and show that the appropriate integrals vanish in order to obtain a Hamiltonian dynamical system.

In general relativity, the gravitational field is described by a Lorentzian metric. We consider a metric $g^{\mu\nu}$ of signature $(-,+,+,+)$ satisfying the vacuum Einstein equations

$$R_{\mu\nu} = 0.$$

These equations arise as the Euler–Lagrange equations of a variational principle based on the Hilbert Lagrangian with a background connection; see Chruściel [20].

To apply the field theory construction to general relativity, we modify the Hilbert Lagrangian while preserving its Euler–Lagrange equations. We replace the metric with its associated metric density and introduce a background connection. This allows us to define suitable momentum variables and construct a first-order Lagrangian. The dynamical variables of this Lagrangian are the metric density and its conjugate momentum. We show that this Lagrangian differs from the Hilbert Lagrangian by a total divergence and therefore yields the same Euler–Lagrange equations; see Anderson [1].

We construct a reference system by choosing a vector field X on M. Its flow ψ describes the motion of a hypersurface $\mathscr{S}$ through M. Having introduced the dynamical variables, the Lagrangian, and a reference system, we apply the framework of Chapter 4 to obtain a Hamiltonian

$$H(X, \mathscr{S}, p) = \int_{\mathscr{S}} \left(p^\mu_{\alpha\beta} \pounds_X \mathfrak{g}^{\alpha\beta} - X^\mu L \right) dS_\mu.$$

The corresponding antisymmetric two-form $\Omega_{\mathscr{S}}$ and the variation δH can likewise be expressed as integrals involving the dynamical variables.

When $\mathscr{S}$ is spacelike, these expressions simplify and can be analysed in the asymptotically flat setting. The asymptotic conditions imposed on the metric allow us to prove that the relevant integrals converge. They also imply that the boundary terms obstructing Hamiltonian evolution vanish. It follows that the resulting phase space has the structure of a Hamiltonian dynamical system.

We explicitly compute the energy and momentum induced by the Hamiltonian and show that they agree with the ADM energy-momentum. More precisely, we obtain

$$-H_{\text{boundary}}(\partial_0, \mathscr{S}) = \lim_{R\to\infty} \frac{1}{16\pi} \int_{S(R)} (g_{ij,j} - g_{jj,i})\, d\sigma_i$$

for the energy, and

$$-H_{\text{boundary}}(\partial_i, \mathscr{S}) = \lim_{R\to\infty} \frac{1}{8\pi} \int_{S(R)} P^{ij}\, d\sigma_j$$

for the momentum, where $i = 1, 2, 3$. Finally, using the results of Bartnik [7] and Chruściel [22], we explain why the resulting ADM energy-momentum is invariant under the appropriate asymptotic coordinate transformations. In Chapter 5, we develop this framework in detail.

To place this construction in a broader context, we compare the Chruściel–Jezierski–Kijowski framework [27] with the original ADM formalism [4, 5] and with the Lagrangian approach of Christodoulou [18]. We also discuss the positive mass problem and survey recent developments in asymptotically flat, asymptotically hyperbolic, and asymptotically hyperboloidal settings. These comparisons and results are treated in detail in Chapter 6. A more detailed overview of the chapters is given in the organisation of the thesis below.

Main results and contributions

This thesis is largely based on the framework of Chruściel, Jezierski and Kijowski [27]. A substantial part of the work therefore consists of checking details and providing arguments for statements left unproved in that source. We provide our own proofs of all results stated in Chapter 4, as these arguments are not available in the cited literature, except for Proposition 4.12, Lemma 4.35 and the sketch of the proof of Theorem 4.37. In addition, Chruściel, Jezierski and Kijowski gave an outline of the proof of Theorem 4.32, which we made rigorous.

The same applies to Chapter 5, all statements accompanied by a proof are proved by the author, except for Theorem 5.45, since this proof relies substantially on results that are not established here.

We also introduce a new jet bundle formulation for the pullback of the Lagrangian to the hypersurface. This geometric construction is discussed in detail in Subsection 4.3.2. It requires reformulating several arguments, namely those in Lemmas 4.24, 4.29 and 4.35 and proposition 4.31.

Moreover, we prove that the Hamiltonian energy-momentum agrees with the ADM energy-momentum; see Subsections 5.3.4 and 5.3.5 for details. To the best of our knowledge, this calculation has not previously appeared in the literature. Chruściel confirmed in personal communication that he did not write out the calculation elsewhere.

Goal of this thesis

The goal of this thesis is threefold. First, we study in depth the field theory construction developed by Chruściel, Jezierski and Kijowski [27]. We fill in missing details and formulate parts of the construction more geometrically.

Second, we review the application of this framework to general relativity, with particular emphasis on asymptotically flat spacetimes. We show that the energy-momentum induced by the Hamiltonian agrees with the ADM energy-momentum.

Third, we place this approach in a broader context by comparing it with other derivations of the ADM energy-momentum. We also discuss contemporary developments concerning positive mass theorems.

Organisation of the thesis

This thesis consists of two preliminary chapters, two chapters containing the main mathematical development, a comparison chapter, and a concluding chapter. Chapters 2 and 3 provide the background required for the thesis and may be skipped by readers already familiar with the material. The central construction and its application to general relativity are developed in Chapters 4 and 5. These results are placed in a broader context in Chapter 6, while Chapter 7 concludes the thesis and discusses directions for future research.

In Chapter 2, we introduce the geometric background needed throughout the thesis. Since we do not wish to restrict ourselves to orientable manifolds, we discuss integration using densities and odd forms and state a corresponding version of Stokes' theorem. We also introduce the bundle notation used later and develop the basic theory of jet bundles.

In Chapter 3, we adopt a historical approach in which symplectic geometry is developed from classical mechanics. We begin with Lagrangian mechanics and pass to Hamiltonian mechanics through the Legendre transformation. This leads naturally to a symplectic formulation of Hamiltonian mechanics. The symplectic form associates a Hamiltonian function with a Hamiltonian vector field, whose integral curves describe the solutions of the equations of motion. This structure motivates the definition of a Hamiltonian dynamical system used throughout the thesis. We also explain how the Liouville one-form induces the canonical symplectic form on the cotangent bundle.

In Chapter 4, the first main chapter of the thesis, we develop the geometric field theory framework of Chruściel–Jezierski–Kijowski [27, Chapter 3]. Using jet bundles, odd forms and a reference system, we construct a phase space associated with an evolving hypersurface and show that it has the structure of a Hamiltonian dynamical system. We also give a more geometric formulation of the pullback of the Lagrangian using jet bundle theory. At this stage, the construction requires neither a metric nor any asymptotic assumptions.

In Chapter 5, we apply the framework developed in Chapter 4 to general relativity. This is the second main chapter of the thesis and focuses in particular on asymptotically flat spacetimes. After adapting the construction to a spacelike hypersurface, we show that the resulting phase space carries the structure of a Hamiltonian dynamical system in the asymptotically flat setting. We show in full detail that its energy-momentum agrees with the ADM energy-momentum. We conclude by discussing the invariance of this construction using the results of Bartnik [7] and Chruściel [22].

In Chapter 6, we place the approach developed in this thesis in a broader context by comparing the Chruściel–Jezierski–Kijowski framework [27] with the original ADM formulation [4, 5] and the Lagrangian approach of Christodoulou [18]. We then give a historical account of positive mass theorems. This is especially relevant given the recent progress in the field. We begin with the classical results of Schoen–Yau [72, 73, 76] and Witten [86], and continue to recent developments across several asymptotic settings.

Finally, in Chapter 7, we summarize the main results of the thesis and discuss several directions for future research.

Data management

- ☒ This thesis is based entirely on existing data, literature, and theory.
- ☒ No new research data was collected or generated.
- ☒ No data sets were stored or processed as part of this thesis.
- ☒ No personal data was processed.

No data were generated or used for the preparation of this thesis.

Use of generative AI

In parts of the research, the LLMs ChatGPT and Google AI have been used as search engines to explore the literature. ChatGPT was used to generate Figure 5.2 based on a sketch by the author. ChatGPT was also used to improve the spelling and grammar of this thesis. Apart from spelling and grammar, the thesis is written without the use of any generative AI. All mathematical arguments presented in this thesis have been developed by the author and are the author's responsibility.

Chapter 2

Geometric preliminaries

This chapter discusses the prerequisites from differential geometry that are needed throughout this thesis but are not usually covered in introductory courses.

In general relativity, we do not assume that the underlying manifold is orientable. We therefore need a notion of integration and a version of Stokes' theorem that remain available on non-orientable manifolds. In Section 2.1, we develop these ideas using densities and odd forms. In Section 2.2, we discuss a generalization of vector bundles, namely fibre bundles. We give a brief overview of the relevant fibre bundle theory and introduce notation that is used throughout the thesis. We also introduce affine bundles and the vertical bundle, which are important special cases and constructions that will be used later in the thesis. Finally, in Section 2.3, we discuss jet bundles. These bundles encode information about a section and its derivatives. We introduce the relevant definitions and develop the basic theory of jet bundles needed for the remainder of this thesis. Jet bundles will be important for the formulation of field theories on manifolds in Chapter 4.

For Section 2.1, we use Lee [54, Chapter 16] and Chruściel, Jezierski, and Kijowski [27, Appendix A]. For Sections 2.2 and 2.3, we follow the exposition of Saunders [71].

2.1 Stokes' theorem on non-orientable manifolds

The goal of this section is to develop a notion of integration on non-orientable manifolds and to introduce a corresponding version of Stokes' theorem using odd forms.

In Subsection 2.1.1, we begin by recalling integration on orientable manifolds. We explain why differential forms of top degree are the natural objects to integrate and why orientability is required in this setting. The obstruction caused by non-orientability motivates the introduction of densities, which are discussed in Subsection 2.1.2. Densities allow one to define integration on non-orientable manifolds and can be viewed as the non-orientable analogue of differential forms of top degree. In Subsection 2.1.3, we introduce odd forms, which extend this construction to lower degrees. To describe them in local coordinates, we refine our earlier terminology by referring to the densities introduced above as scalar densities and placing them within the more general class of tensor densities. Finally, we state a corresponding version of Stokes' theorem.

For Subsections 2.1.1 and 2.1.2, we follow Lee's treatment of integration and densities [54, Chapter 16]. For Subsection 2.1.3, we follow Chruściel, Jezierski, and Kijowski [27, Appendix A].

2.1.1 Integration on orientable manifolds

This subsection reviews integration using differential forms. We first explain why differential forms of top degree naturally represent volume and are therefore the appropriate objects to integrate. We then define integration of differential forms on Euclidean space. The transformation law for integrals shows that the sign of the Jacobian determinant plays an essential role. When passing to manifolds, this requires compatible choices of orientation on overlapping coordinate charts, and hence leads to the notion of an orientable manifold.

To understand why the determinant appears in the transformation law for integration, we first recall its geometric interpretation as a measure of volume.

Lemma 2.1 *Let $v_1, \ldots, v_n$ be n linearly independent vectors in $\mathbb{R}^n$. They span the parallelepiped*

$$P = \left\{ \sum_{i=1}^{n} \lambda_i v_i \,\middle|\, 0 \leq \lambda_i \leq 1 \right\},$$

whose volume is given by

$$\mathrm{Vol}(P) = \left| \det(v_1, \ldots, v_n) \right|.$$

Proof. We apply the Gram–Schmidt process to the vectors $v_1, \ldots, v_n$ to obtain an orthogonal set of vectors $v_1, v_2^\perp, \ldots, v_n^\perp$. Each vector v_j is written as

$$\begin{aligned} v_1 &= v_1 \\ v_2 &= v_2^\perp + \mathrm{proj}_{v_1^\perp}(v_2) \\ &\;\vdots \\ v_n &= v_n^\perp + \sum_{i=1}^{n-1} \mathrm{proj}_{v_i^\perp}(v_n), \end{aligned}$$

where

$$\mathrm{proj}_{v_i^\perp}(v_j) = \frac{\langle v_j, v_i^\perp \rangle}{\langle v_i^\perp, v_i^\perp \rangle} v_i^\perp,$$

and $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product.

The determinant is multilinear in its columns and vanishes whenever two columns are linearly dependent. Consequently, when expanding the determinant using the above decompositions, all terms containing projection vectors vanish. Hence,

$$\begin{aligned} |\det(v_1, \ldots, v_n)| &= \left| \det \left(v_1, v_2^\perp + \mathrm{proj}_{v_1^\perp}(v_2), \ldots, v_n^\perp + \sum_{i=1}^{n-1} \mathrm{proj}_{v_i^\perp}(v_n) \right) \right| \\ &= |\det(v_1^\perp, v_2^\perp, \ldots, v_n^\perp)|. \end{aligned}$$

Since the vectors $v_1, v_2^\perp, \ldots, v_n^\perp$ are mutually orthogonal, equality holds in Hadamard's inequality; see Axler [6]. Therefore,

$$|\det(v_1^\perp, \ldots, v_n^\perp)| = \prod_{i=1}^{n} \|v_i^\perp\|.$$

The Gram–Schmidt procedure preserves the volume of the parallelepiped, since subtracting from each vector its projection onto the span of the preceding vectors does not change the corresponding perpendicular height. Hence, the original parallelepiped P has volume

$$\mathrm{Vol}(P) = \prod_{i=1}^{n} \|v_i^{\perp}\|.$$

We therefore conclude that

$$\mathrm{Vol}(P) = |\det(v_1, \dots, v_n)|.$$

□

The determinant provides a natural way to measure volume in $\mathbb{R}^n$. In particular, it is an alternating multilinear map

$$\det\colon \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ times}} \to \mathbb{R},$$

and its absolute value gives the volume of the parallelepiped spanned by its arguments.

This suggests how to generalize the notion of a volume element to a manifold. Let M be an n-dimensional manifold. At each point $p \in M$, an n-form

$$\omega_p \in \Lambda^n(T_p^*M)$$

is an alternating multilinear map

$$\omega_p\colon \underbrace{T_pM \times \cdots \times T_pM}_{n \text{ times}} \to \mathbb{R}.$$

Thus, an n-form has the same multilinear structure as the determinant and can be viewed as assigning a volume to n tangent vectors. This motivates the use of differential n-forms,

$$\omega \in \Omega^n(M),$$

as the natural objects to integrate on an n-dimensional manifold.

To use this construction to integrate functions, we relate smooth functions to differential forms of top degree. Recall that

$$\dim \Lambda^k T_p^* M = \binom{n}{k}. \tag{2.1}$$

In particular,

$$\dim \Lambda^n T_p^* M = 1.$$

If M is orientable, the bundle $\Lambda^n T^*M$ admits a nowhere vanishing section. Choosing such a section allows us to identify smooth functions with differential n-forms.

Lemma 2.2 *Let M be an n-dimensional orientable manifold and let $\mu \in \Omega^n(M)$ be a nowhere vanishing n-form. Then the map*

$$C^\infty(M) \to \Omega^n(M), \qquad f \mapsto f\mu,$$

is an isomorphism of $C^\infty(M)$-modules.

Proof. Notice that $\Lambda^n(T^*M)$ is a line bundle over M. Since M is orientable, there exists a nowhere vanishing n-form $\mu \in \Omega^n(M)$, and hence $\Lambda^n(T^*M)$ is a trivial line bundle.

The space of smooth functions

$$C^\infty(M) = \Gamma(M \times \mathbb{R})$$

can be viewed as the space of sections of the trivial line bundle $M \times \mathbb{R}$. Sending the generator $\mathbb{1}$ of $C^\infty(M)$ to the nowhere vanishing section μ defines an isomorphism of $C^\infty(M)$-modules,

$$C^\infty(M) \to \Omega^n(M), \quad f\mathbb{1} \mapsto f\mu.$$

Therefore,

$$C^\infty(M) \cong \Omega^n(M).$$

□

We now make this identification concrete by defining the integral of a top-degree form on Euclidean space.

Definition 2.3 (integral on $\mathbb{R}^n$) Let $D \subset \mathbb{R}^n$ be an open and bounded subset, and let $\omega \in \Omega^n(\overline{D})$, where $\overline{D}$ denotes the closure of D. There exists $f \in C^\infty(\overline{D})$ such that

$$\omega = f\, dx^1 \wedge \cdots \wedge dx^n.$$

We define the **integral of ω over D** by

$$\int_D \omega = \int_D f\, dV.$$

Equivalently, we write

$$\int_D f\, dx^1 \cdots dx^n := \int_D f\, dx^1 \wedge \cdots \wedge dx^n.$$

The following proposition shows how the integral behaves under pullback by a diffeomorphism. In particular, it makes explicit the sign change that occurs when the diffeomorphism reverses orientation. This sign change is the obstruction we must address when passing to non-orientable manifolds.

Proposition 2.4 *Suppose D and E are open and bounded in $\mathbb{R}^n$, and $G : \overline{D} \to \overline{E}$ is a smooth map that restricts to an orientation-preserving or orientation-reversing diffeomorphism from D to E. If $\omega \in \Omega^n(\overline{E})$, then*

$$\int_D G^*\omega = \begin{cases} \displaystyle\int_E \omega, & \text{if } G \text{ is orientation-preserving,} \\ \displaystyle -\int_E \omega, & \text{if } G \text{ is orientation-reversing.} \end{cases}$$

Proof. Let $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^n)$ be coordinates on D and E, respectively, and write

$$\omega = f\, dy^1 \wedge \cdots \wedge dy^n.$$

The transformation law for integrals gives

$$\int_E \omega = \int_E f\,dV = \int_D (f \circ G)|\det DG|\,dV,$$

where DG denotes the Jacobian of G with respect to the coordinates $(x^1, \dots, x^n)$.

Since G is either orientation-preserving or orientation-reversing, we have

$$|\det DG| = \pm \det DG,$$

where the plus sign corresponds to the orientation-preserving case and the minus sign to the orientation-reversing case. Hence,

$$\int_E \omega = \pm \int_D (f \circ G)(\det DG)\,dV.$$

The integrand on the right is exactly the coordinate expression for the pullback of the top-degree form ω. Therefore,

$$\int_E \omega = \pm \int_D G^*\omega. \qquad \square$$

We now extend the definition of integration from Euclidean space to an oriented manifold. The integral is defined locally in positively oriented charts, and these local contributions are combined using a partition of unity.

Definition 2.5 Let M be an oriented n-dimensional manifold and let $\omega \in \Omega^n(M)$ be compactly supported. Let $\{U_i\}$ be a finite cover of $\operatorname{supp}(\omega)$ by positively oriented coordinate charts (U_i, φ_i), and let $\{\psi_i\}$ be a partition of unity subordinate to $\{U_i\}$. We define the **integral of ω over M** by

$$\int_M \omega := \sum_i \int_{\varphi_i(U_i)} (\varphi_i^{-1})^*(\psi_i \omega).$$

Remark 2.6 The assumption that ω is compactly supported ensures that only finitely many local integrals are needed and avoids convergence issues. In applications to non-compact manifolds, we will also integrate forms that are not compactly supported. In that case, the integral must be shown to converge.

One can prove that this definition is independent of the choice of positively oriented charts and of the subordinate partition of unity. The integral also satisfies the usual properties, such as $\mathbb{R}$-linearity and invariance under orientation-preserving diffeomorphisms. We omit these results and their proofs. For details, see Lee [54, Chapter 16].

One of the fundamental results of integration theory on oriented manifolds is Stokes' theorem. We use the following formulation, adapted from Jost [47] and Lee [54].

Theorem 2.7 (Stokes' theorem) *Let M be an oriented m-dimensional differentiable manifold with smooth boundary, and let ω be a smooth $(m-1)$-form on M. If the integrals exist, then*

$$\int_M d\omega = \int_{\partial M} \omega.$$

2.1.2 Integration on non-orientable manifolds

The preceding discussion defined integration using top-degree differential forms, which requires a choice of orientation. Since such a choice is not available on a non-orientable manifold, we now introduce densities, whose transformation law involves the absolute value of the determinant. Densities therefore provide a notion of integration that does not depend on orientability. We first present the construction in linear algebra and then extend it to vector bundles. Most proofs are omitted; see [54, Chapter 16].

Definition 2.8 Let V be an n-dimensional vector space. An **n-density on V** is a function

$$\mu\colon V \times \cdots \times V \to \mathbb{R}, \tag{2.2}$$

such that for every linear map $T\colon V \to V$,

$$\mu(Tv_1, \ldots, Tv_n) = |\det T|\, \mu(v_1, \ldots, v_n). \tag{2.3}$$

Note that μ is not a tensor, since it is not $\mathbb{R}$-linear in any of its arguments. We denote the space of n-densities on V by $\tilde{\Lambda}^n V$. Although a density is not a multilinear map, the following proposition shows that every differential n-form determines a density.

Proposition 2.9 (Properties of densities) *Let V be a vector space of dimension $n \geq 1$.*

(a) *$\tilde{\Lambda}^n V$ is a vector space under the operations*

$$(c_1\mu_1 + c_2\mu_2)(v_1, \ldots, v_n) = c_1\mu_1(v_1, \ldots, v_n) + c_2\mu_2(v_1, \ldots, v_n).$$

(b) *If $\mu_1, \mu_2 \in \tilde{\Lambda}^n V$ and*

$$\mu_1(E_1, \ldots, E_n) = \mu_2(E_1, \ldots, E_n)$$

for some basis $(E_1, \ldots, E_n)$ of V, then $\mu_1 = \mu_2$.

(c) *If $\omega \in \Lambda^n(V^*)$, then the map*

$$|\omega|\colon V \times \cdots \times V \to \mathbb{R},$$

defined by

$$|\omega|(v_1, \ldots, v_n) = |\omega(v_1, \ldots, v_n)|,$$

is a density.

(d) *$\tilde{\Lambda}^n V$ is one-dimensional and is spanned by $|\omega|$ for any nonzero $\omega \in \Lambda^n(V^*)$.*

A density μ on V is said to be **positive** if

$$\mu(v_1, \ldots, v_n) > 0$$

whenever $\{v_1, \ldots, v_n\}$ is a linearly independent set. If $\omega \in \Lambda^n(V^*)$ is nonzero, then the density $|\omega|$ is positive.

We now extend the notion of a density to vector bundles by applying the preceding construction fibrewise.

Definition 2.10 Let M be a manifold. The set

$$\tilde{\Lambda}^n T^*M = \bigsqcup_{p \in M} \tilde{\Lambda}^n T_p^* M$$

is called the **density bundle** on M.

This set has a natural projection

$$\pi\colon \tilde{\Lambda}^n T^*M \to M, \qquad (\mu, p) \mapsto p.$$

The following proposition shows that the term density bundle is justified.

Proposition 2.11 *Let M be a manifold. Then its density bundle $\tilde{\Lambda}^n T^*M$ is a line bundle over M.*

Definition 2.12 Let M be a manifold. A section of $\tilde{\Lambda}^n T^*M$ is called a **density** on M.

The space of densities on M is denoted by

$$\tilde{\Omega}^n(M) := \Gamma(\tilde{\Lambda}^n T^*M).$$

A density μ on M is called **positive** if μ_p is positive for every $p \in M$.

Proposition 2.13 *Let M be a manifold. Then there exists a global smooth positive density μ on M. Choosing such a density defines an isomorphism*

$$C^\infty(M) \to \tilde{\Omega}^n(M), \qquad f \mapsto f\mu.$$

Recall that, on an oriented manifold, we defined integration by using the identification between smooth functions and top-degree forms. Similarly, using Proposition 2.13, we can identify smooth functions with densities and thereby define integration on Euclidean space.

Definition 2.14 Let $D \subset \mathbb{R}^n$ be open and bounded, and let $\mu \in \tilde{\Omega}^n(\overline{D})$. Then there exists a unique function $f\colon \overline{D} \to \mathbb{R}$ such that

$$\mu = f \left| dx^1 \wedge \cdots \wedge dx^n \right|.$$

We define the **integral of** μ over D by

$$\int_D \mu = \int_D f \, dV.$$

Equivalently, we write

$$\int_D f \, dx^1 \cdots dx^n := \int_D f \left| dx^1 \wedge \cdots \wedge dx^n \right|.$$

To extend the definition of integration to manifolds, we need to understand how densities behave under coordinate transformations.

Definition 2.15 Let $F\colon M \to N$ be a smooth map between n-dimensional manifolds and let $\mu \in \tilde{\Omega}^n(N)$. We define the **pullback density** $F^*\mu$ on M by

$$(F^*\mu)_p(v_1, \ldots, v_n) = \mu_{F(p)}\big(dF_p(v_1), \ldots, dF_p(v_n)\big).$$

We record some basic properties of the pullback density.

Lemma 2.16 *Let $G\colon P \to M$ and $F\colon M \to N$ be smooth maps between n-manifolds, with or without boundary, and let μ be a density on N.*

(a) For any $f \in C^\infty(N)$,

$$F^*(f\mu) = (f \circ F)\, F^*\mu.$$

(b) If ω is an n-form on N, then

$$F^*|\omega| = |F^*\omega|.$$

(c) If μ is smooth, then $F^*\mu$ is a smooth density on M.

(d)

$$(F \circ G)^*\mu = G^*(F^*\mu).$$

We show that densities have the same transformation law as integrals. In particular, the absolute value of the determinant appears in both cases.

Proposition 2.17 *Suppose $F\colon M \to N$ is a smooth map between n-manifolds. Let $\{x^i\}$ and $\{y^j\}$ be smooth coordinates on open subsets $U \subseteq M$ and $V \subseteq N$, respectively, and let u be a continuous real-valued function on V. Then, on $U \cap F^{-1}(V)$,*

$$F^*\big(u\,|dy^1 \wedge \cdots \wedge dy^n|\big) = (u \circ F)\,|\det DF|\,|dx^1 \wedge \cdots \wedge dx^n|,$$

where DF denotes the Jacobian of F in these coordinates.

This proposition follows directly from the properties in Lemma 2.16.

The transformation law in Proposition 2.17 shows that densities transform in the same way as the Euclidean integral. The following proposition makes this compatibility precise and is the key step that allows us to define integration without choosing an orientation.

Proposition 2.18 *Suppose U and O are open subsets of $\mathbb{R}^n$, and let $G\colon U \to O$ be a diffeomorphism. If μ is a compactly supported density on O, then*

$$\int_O \mu = \int_U G^*\mu.$$

This proposition shows that the value of the integral does not depend on the coordinate chart used to represent the density. Therefore, we define integration locally and then combine the local contributions using a partition of unity.

Definition 2.19 Let M be a manifold and let $\mu \in \tilde{\Omega}^n(M)$ have compact support contained in a chart (U, φ). We define the **integral of** μ over M by

$$\int_M \mu = \int_{\varphi(U)} (\varphi^{-1})^*\mu.$$

More generally, let $\{\psi_i\}$ be a partition of unity subordinate to an open cover of M by smooth coordinate charts. We define

$$\int_M \mu = \sum_i \int_M \psi_i\mu.$$

This definition is independent of the choice of coordinate charts and of the subordinate partition of unity. The resulting integral has the usual properties of integration, but no choice of orientation is required. We omit the proofs; they can be found in Lee [54, Chapter 16].

2.1.3 Odd forms

This subsection establishes the framework needed to formulate Stokes' theorem without assuming a global orientation. We introduce only the terminology and constructions needed in this thesis and do not develop the general theory. References for readers interested in a more detailed treatment are given below.

Densities can be viewed as the non-orientable analogue of differential forms of top degree. However, to formulate a version of Stokes' theorem without assuming orientability, we also need corresponding objects in lower degrees. Odd forms provide such a generalization. On an m-dimensional manifold, odd forms of top degree are naturally identified with the densities introduced in Subsection 2.1.2. In the terminology introduced below, these densities are called **scalar densities**.

To describe odd forms of lower degree, we require a more general class of objects that also includes scalar densities as a special case. We therefore refine our terminology and introduce **tensor densities**. This allows us to describe odd forms of arbitrary degree in local coordinates. We do not develop the full theory here, since this would go beyond the scope of this thesis.

Our presentation follows Chruściel, Jezierski, and Kijowski [27, Appendix A], from whom we also take the version of Stokes' theorem for odd forms. For a more detailed discussion, they refer to the Polish source [80]. We also refer to Bott [9] for a more detailed treatment of odd differential forms.

Definition 2.20 Let M be an m-dimensional manifold, let $p \in M$, and let U be an orientable neighbourhood of p. An **odd n-form** on U is an equivalence class

$$[(\alpha_n, \mathscr{O})],$$

where α_n is a differential n-form on U and $\mathscr{O}$ is an orientation of U. The equivalence relation is given by

$$(\alpha_n, \mathscr{O}) \sim (-\alpha_n, -\mathscr{O}),$$

where $-\mathscr{O}$ denotes the orientation opposite to $\mathscr{O}$.

Odd forms can be defined globally using a partition of unity, even for non-orientable manifolds.

For odd forms of top degree, the above equivalence relation leads to the same transformation behaviour as for the densities introduced in Subsection 2.1.2. Let M be an m-dimensional manifold and let

$$\widetilde{\alpha}_m = [(\alpha_m, \mathscr{O})]$$

be an odd m-form. In local coordinates $(x^1, \dots, x^m)$, an ordinary m-form is written as

$$\alpha_m = f\, dx^1 \wedge \cdots \wedge dx^m.$$

Under a change of coordinates, the ordinary top form transforms with the determinant of the Jacobian. If the coordinate transformation reverses orientation, this determinant changes sign. At the same time, the local orientation changes from $\mathscr{O}$ to $-\mathscr{O}$. By the equivalence relation

$$(\alpha_m, \mathscr{O}) \sim (-\alpha_m, -\mathscr{O}),$$

this change of orientation introduces an additional minus sign. Consequently, the two signs cancel, and the coefficient of an odd top-degree form transforms with the absolute value of the determinant.

Thus, in local coordinates, an odd m-form has a transformation law of the form

$$f_y = \left|\det\left(\frac{\partial x}{\partial y}\right)\right| f_x.$$

This is precisely the transformation law of a density. Hence, on an m-dimensional manifold, odd m-forms are naturally identified with the densities introduced in Subsection 2.1.2.

To describe odd forms of lower degree in local coordinates, we need a more general notion. The densities considered above are special cases of **tensor densities**, called **scalar densities**.

Definition 2.21 A **tensor density of weight** 1 is an object whose transformation law contains, in addition to the usual tensor transformation, a factor given by the determinant of the Jacobian of the coordinate transformation. More precisely, under a change of coordinates

$$F\colon (x^1, \dots, x^m) \longmapsto (y^1, \dots, y^m),$$

a tensor density of weight 1 transforms as the corresponding tensor, multiplied by the factor

$$\det DF,$$

where DF denotes the Jacobian of F.

More generally, the exponent of the determinant appearing in the transformation law is called the **weight** of the tensor density. In this thesis, whenever the weight is not specified, it is understood to be 1.

Locally, odd n-forms on an m-dimensional manifold are represented by antisymmetric contravariant tensor densities of rank

$$r = m - n$$

and weight 1. More precisely, let $(\alpha_n, \mathscr{O})$ be a local representative of an odd n-form, where $\mathscr{O}$ is the orientation determined by the frame $(x^1, \dots, x^m)$. Then α_n is written locally as

$$\alpha_n = T^{i_1 \dots i_r}\left(\frac{\partial}{\partial x^{i_1}} \wedge \cdots \wedge \frac{\partial}{\partial x^{i_r}}\right) \lrcorner \left(dx^1 \wedge \cdots \wedge dx^m\right),$$

where $T^{i_1 \dots i_r}$ is antisymmetric in its indices and transforms as a contravariant tensor density of weight 1.

Example 2.22 If $r = 0$, then $n = m$, so the tensor density has no tensor indices. In this case, one obtains a **scalar density**. Thus, odd forms of top degree correspond to scalar densities.

If $r = 1$, then $n = m - 1$. In this case, a contravariant vector density corresponds to an odd $(m-1)$-form.

For a more detailed treatment of tensor densities, see Plebański and Krasiński [67] or Schouten [77].

To define the integral of an odd form over a submanifold, we must keep track of two different notions of orientation. Besides the usual orientation of the tangent bundle of the submanifold, we also need an orientation of the directions normal to it. Together, these determine an orientation of the ambient manifold along the submanifold.

Definition 2.23 Let $\Sigma \subset M$ be a submanifold. An **exterior orientation**, denoted by $\mathscr{O}_{\text{ext}}$, is an orientation of the normal bundle $N\Sigma$. Here we denote

$$N\Sigma := TM|_{\Sigma} / T\Sigma.$$

An **interior orientation**, denoted by $\mathscr{O}_{\text{int}}$, is an orientation of the tangent bundle $T\Sigma$.

An **ambient orientation**, denoted by $\mathscr{O}$, is an orientation of $TM|_{\Sigma}$.

The relation

$$(\mathscr{O}_{\text{ext}}, \mathscr{O}_{\text{int}}) = \mathscr{O}$$

means that a positively oriented basis of the normal directions, followed by a positively oriented basis of $T\Sigma$, gives a positively oriented basis of the ambient tangent space.

Example 2.24 Let M be an m-dimensional manifold, and let $\Sigma \subset M$ be an n-dimensional submanifold. Let

$$(v_1, \ldots, v_n)$$

be a positively oriented local frame of $T\Sigma$ with respect to $\mathscr{O}_{\text{int}}$. Furthermore, let

$$(e_1, \ldots, e_{m-n})$$

be a positively oriented local frame of the normal bundle with respect to $\mathscr{O}_{\text{ext}}$. Then the concatenated basis

$$(e_1, \ldots, e_{m-n}, v_1, \ldots, v_n)$$

is positively oriented with respect to the ambient orientation $\mathscr{O}$.

One can integrate odd n-forms over externally oriented n-dimensional submanifolds.

Definition 2.25 Let $\widetilde{\alpha}_n = [(\alpha_n, \mathscr{O})]$ be an odd n-form, and let Σ be an n-dimensional submanifold with exterior orientation $\mathscr{O}_{\text{ext}}$. We define

$$\int_{(\Sigma, \mathscr{O}_{\text{ext}})} \widetilde{\alpha}_n := \int_{(\Sigma, \mathscr{O}_{\text{int}})} \alpha_n, \tag{2.4}$$

where $\mathscr{O}_{\text{int}}$ is the interior orientation of Σ determined by the condition

$$(\mathscr{O}_{\text{ext}}, \mathscr{O}_{\text{int}}) = \mathscr{O}.$$

In other words, $\mathscr{O}_{\text{int}}$ is chosen such that a positively oriented basis of the normal directions, followed by a positively oriented basis of $T\Sigma$, gives the ambient orientation $\mathscr{O}$.

The left-hand side of Equation (2.4) denotes the integral of an odd form over an externally oriented submanifold. The right-hand side is the integral of the differential form α_n over Σ, as defined in Subsection 2.1.1. The interior orientation used in this integral is determined by $\mathscr{O}_{\text{ext}}$ and $\mathscr{O}$.

We use the following result from Chruściel–Jezierski–Kijowski [27, Appendix A] without proof.

Theorem 2.26 (Stokes' theorem for odd forms) *Let M be a smooth m-dimensional manifold and let $\Sigma \subset M$ be a smooth n-dimensional submanifold with smooth boundary. Let*

$$\widetilde{\alpha}_{n-1} = [(\alpha_{n-1}, \mathscr{O})]$$

be an odd $(n-1)$-form. Then

$$\int_{(\Sigma,\mathscr{O}_{\mathrm{ext}})} d\widetilde{\alpha}_{n-1} = \int_{\partial(\Sigma,\mathscr{O}_{\mathrm{ext}})} \widetilde{\alpha}_{n-1}.$$

Here

$$d[(\alpha_{n-1},\mathscr{O})] := [(d\alpha_{n-1},\mathscr{O})],$$

and the boundary $\partial(\Sigma,\mathscr{O}_{\mathrm{ext}})$ is equipped with the exterior orientation induced by $(\Sigma,\mathscr{O}_{\mathrm{ext}})$.

This concludes our discussion of integration. The main point for this thesis is that odd forms can be integrated on non-orientable manifolds and admit a corresponding version of Stokes' theorem. For readers unfamiliar with odd forms, it may be helpful to think of them intuitively as differential forms adapted to settings in which no global orientation is available.

In the remainder of this thesis, odd forms will remain mostly in the background and will play only a minor role. This subsection therefore mainly records the framework that permits the use of integration and Stokes' theorem without imposing orientability. In the following chapters, we will use the corresponding version of Stokes' theorem without discussing the underlying orientation issues.

2.2 Bundles

This section introduces the standard definitions and notation used throughout this thesis. We also explain several mild abuses of notation that are adopted for the sake of readability. We assume familiarity with vector bundles and will therefore not review these concepts here.

In Subsection 2.2.1, we define a fibre bundle, discuss its basic properties and introduce the notation used throughout the thesis. We then consider two related constructions in more detail. In Subsection 2.2.2, we introduce affine bundles. This will be useful later because jet bundles naturally carry affine-bundle structures. In Subsection 2.2.3, we introduce the vertical bundle associated with a fibre bundle. Its fibres consist of vectors tangent to the fibres of the original bundle. The vertical bundle arises naturally in the study of fibre bundles and will appear in the field-theoretic constructions of Chapter 4.

This section is based on Saunders [70, Chapters 1–3].

2.2.1 Fibre bundles

Recall that a vector bundle is locally described by a product of the base manifold with a vector space. Since every finite-dimensional vector space is naturally a smooth manifold, this suggests replacing the vector space by an arbitrary manifold. This leads to the notion of a fibre bundle.

Definition 2.27 A **fibre bundle with typical fibre** F consists of smooth manifolds E, M, and F, together with a surjective smooth map

$$\pi\colon E \to M$$

that is locally trivial. More precisely, for every $p \in M$, there exists an open neighbourhood $U \subset M$ of p and a diffeomorphism

$$h_U\colon E|_U := \pi^{-1}(U) \longrightarrow U \times F$$

such that

$$\mathrm{pr}_1 \circ h_U = \pi|_{\pi^{-1}(U)},$$

where $\mathrm{pr}_1 \colon U \times F \to U$ denotes the projection onto the first factor. The map h_U is called a **local trivialization**.

The manifold M is called the **base space**, E the **total space**, and F the **typical fibre**. For each $p \in M$, the preimage

$$E_p := \pi^{-1}(p)$$

is called the **fibre over** p.

Remark 2.28 When the typical fibre is not relevant to the discussion or is clear from the context, we omit it from the notation and denote the fibre bundle by the triple

$$(E, \pi, M),$$

or simply by

$$\pi \colon E \to M.$$

Fibre bundles can locally be described as product manifolds. This local product structure allows us to reorganize the local coordinates such that points in the same fibre have the same first n coordinates and are distinguished by their last m coordinates, where $n = \dim(M)$ and $m = \dim(F)$.

Lemma 2.29 *Let (E, π, M) be a fibre bundle with typical fibre F. Around each $e \in E$ there exist local coordinates*

$$(x^1, \dots, x^n, u^1, \dots, u^m)$$

such that

$$\pi(x^1, \dots, x^n, u^1, \dots, u^m) = (x^1, \dots, x^n).$$

Proof. Let $e \in E$, let $W \subset F$ be a neighbourhood of $\mathrm{pr}_2(h(e))$ and let $U \subset M$ be a trivializing neighbourhood of $\pi(e)$ with a diffeomorphism

$$h \colon \pi^{-1}(U) \xrightarrow{\sim} U \times F.$$

Locally $U \subset M$ and $W \subset F$ have coordinates

$$\begin{aligned} x &\colon U \to \mathbb{R}^n \\ u &\colon W \to \mathbb{R}^m, \end{aligned}$$

where $\dim(M) = n$ and $\dim(F) = m$. We define $y \colon h^{-1}(U \times W) \to \mathbb{R}^{n+m}$ by

$$y := (x \circ \mathrm{pr}_1 \circ h, u \circ \mathrm{pr}_2 \circ h),$$

which defines local coordinates around $e \in E$. □

Remark 2.30 We write

$$y = (x^1, \dots, x^n, u^1, \dots, u^m) = (x^i, u^a).$$

We abuse notation and use the same symbol x^i both for the i-th coordinate function on $\pi(V)$ and for the composite $x^i \circ \pi$ on V, where $V \subset E$ is a trivializing neighbourhood.

These coordinates are called **adapted coordinates**. Since we may always choose such coordinates locally, we will use them whenever convenient without further mention.

Definition 2.31 Let (E, π, M) be a fibre bundle, and let $U \subset M$ be open. A map $\phi\colon U \to E$ is called a **local section** of E if $\pi \circ \phi = \mathrm{id}_U$. The set of local sections on U is denoted by $\Gamma(E|_U)$. Whenever $U = M$, the section is called a **global section**. The set of global sections is denoted by $\Gamma(E)$.

A section can be described in terms of coordinates. Suppose $\phi \in \Gamma(E)$, and (x^i, u^a) are coordinate functions around $\phi(p) \in E$, where $p \in M$. As mentioned before, we use x^i for the coordinate function on the base and for the composite $x^i \circ \pi$ on the total space. Hence,

$$\begin{aligned} x^i(\phi(p)) &= x^i \circ \pi(\phi(p)) \\ &= x^i(p). \end{aligned}$$

This shows the first n coordinates of $\phi(p)$ are determined by the coordinates of p. Hence, locally the section ϕ is determined by functions ϕ^a defined as

$$\phi^a = u^a \circ \phi.$$

Again, we abuse notation; we write ϕ for the restriction of ϕ to its coordinate neighbourhood.

A morphism between fibre bundles can be described as a pair of maps, one between the base spaces and one between the total spaces. The two maps are related by the bundle projections.

Definition 2.32 Let (E, π, M) and (H, ρ, N) be fibre bundles. A **bundle morphism** from $E \to H$ is a pair $(f, \overline{f})$ where $f\colon E \to H$ and $\overline{f}\colon M \to N$ with $\rho \circ f = \overline{f} \circ \pi$. The map $\overline{f}$ is called the **projection** of f.

$$\begin{array}{ccc} E & \xrightarrow{f} & H \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\rho} \\ M & \xrightarrow[\overline{f}]{} & N \end{array}$$

We characterize a bundle morphism in the following way.

Lemma 2.33 *Let $f\colon E \to H$ be a map between fibre bundles, possibly over different bases. There exists a bundle morphism $(f, \overline{f})$ if and only if for every $p \in M$ and all $a, b \in E|_p$ we have:*

$$\rho(f(a)) = \rho(f(b)).$$

Furthermore, the map $\overline{f}$ is unique.

This means that a bundle map must send fibres to fibres. Usually, we omit the projection map and write only the map between the bundles.

To describe bundle morphisms in local coordinates, let $(f, \overline{f})$ be a bundle morphism between (E, π, M) and (H, ρ, N). Suppose (y^j, v^A) is a coordinate system on (H, ρ, N). Then we have functions f^j, f^A defined by

$$\begin{aligned} f^j &= y^j \circ f \\ f^A &= v^A \circ f. \end{aligned}$$

Note that f^j is constant on the fibres since f maps fibres to fibres. In the same manner, we find a function $\overline{f}^j$ defined by

$$\overline{f}^j = y^j \circ \overline{f},$$

where y^j are local coordinates on N. Applying the definition of bundle morphisms, and using the fact that y^j may be regarded as coordinate functions on both N and H, we obtain

$$\begin{aligned} y^j \circ (\rho \circ f) &= y^j \circ (\overline{f} \circ \pi) \\ y^j \circ f &= (y^j \circ \overline{f}) \circ \pi \\ f^j &= \overline{f}^j \circ \pi. \end{aligned}$$

By a mild abuse of notation, we also denote the corresponding function by f^j on M and on E. More precisely, we write

$$f^j = y^j \circ \overline{f}.$$

To conclude this subsection, we introduce the notions of a fibred product bundle and a pullback bundle.

Definition 2.34 Let (E, π, M) and (H, ρ, M) be fibre bundles over the same base manifold M. Their **fibred product bundle** is the fibre bundle

$$(E \times_M H, \pi \times_M \rho, M),$$

where the total space is defined by

$$E \times_M H := \{(a, b) \in E \times H \mid \pi(a) = \rho(b)\},$$

and the projection is given by

$$\pi \times_M \rho \colon E \times_M H \to M, \qquad (a, b) \longmapsto \pi(a) = \rho(b).$$

The terminology agrees with the geometric intuition, since the fibre over each point $x \in M$ is

$$(E \times_M H)_x = E_x \times H_x.$$

The local trivializations of E and H combine to give local trivializations of $E \times_M H$, showing that it is indeed a fibre bundle.

Sections of the two original bundles naturally determine a section of their fibred product. Conversely, every section of the fibred product determines a section of each factor by composition with the canonical projections.

$$\begin{tikzcd} E \times_M H \arrow[r, "\mathrm{pr}_1"] \arrow[d, "\mathrm{pr}_2"'] & E \arrow[d, "\pi"] \\ H \arrow[r, "\rho"'] & M \end{tikzcd}$$

Figure 2.1: The commutative diagram defining the fibred product bundle.

The ordinary product projections restrict to maps

$$\mathrm{pr}_1\colon E \times_M H \to E, \qquad \mathrm{pr}_2\colon E \times_M H \to H.$$

These maps form the commutative diagram shown in Figure 2.1.

The pullback bundle is obtained from the same construction when one of the two maps to the common base is not required to be a bundle projection. More precisely, given a fibre bundle $\pi\colon E \to M$ and a smooth map $f\colon N \to M$, we form the fibred product of E and N over M.

Definition 2.35 Let (E, π, M) be a fibre bundle, and let $f\colon N \to M$ be a smooth map between manifolds. The **pullback bundle** of E along f is the fibre bundle

$$(f^*E, f^*\pi, N),$$

whose total space is

$$f^*E := \{(x, e) \in N \times E \mid f(x) = \pi(e)\},$$

and whose bundle projection is

$$f^*\pi\colon f^*E \to N, \qquad (x, e) \longmapsto x.$$

The restriction of the projection onto the second component gives a smooth map

$$\mathrm{pr}_2\colon f^*E \to E.$$

By the definition of f^*E, these maps satisfy

$$\pi \circ \mathrm{pr}_2 = f \circ f^*\pi,$$

as illustrated in Figure 2.2.

$$\begin{tikzcd} f^*E \arrow[r, "\mathrm{pr}_2"] \arrow[d, "f^*\pi = \mathrm{pr}_1"'] & E \arrow[d, "\pi"] \\ N \arrow[r, "f"'] & M \end{tikzcd}$$

Figure 2.2: The commutative diagram defining the pullback bundle.

If E has typical fibre F, then f^*E also has typical fibre F. Indeed, every local trivialization of E over an open set $U \subset M$ induces a local trivialization of f^*E over $f^{-1}(U) \subset N$.

Let $\phi \in \Gamma(E|_U)$ be a local section over an open set $U \subset M$. The **pullback section**

$$f^*\phi \in \Gamma\left(f^*E|_{f^{-1}(U)}\right)$$

is defined by

$$(f^*\phi)(a) := \big(a, \phi(f(a))\big), \qquad a \in f^{-1}(U).$$

2.2.2 Affine bundles

This subsection introduces affine bundles. The motivation for this material comes from jet bundles, which naturally carry affine-bundle structures. In particular, the affine structure underlies their coordinate transformation laws, which will be used later in Subsection 4.3.3. The connection with jet bundles will be discussed when they are introduced. The presentation follows Saunders [70, Chapter 2].

We begin by defining affine spaces and affine maps in the linear algebraic setting. We then extend these notions to bundles over manifolds and define affine bundles. Finally, we discuss some of their basic properties.

The idea of an affine space is to retain the notions of translation and of the difference between two points, without distinguishing a particular point as the origin. More precisely, we define an affine space as follows.

Definition 2.36 Let V be a vector space, A be a set and $\alpha\colon A \times V \to A$ be a function. The triple (A, V, α) is called an **affine space** if it satisfies the following three axioms:

1. for all $x \in A$, $\alpha(x, 0) = x$;
2. for all $x \in A$ and for all $v, w \in V$, $\alpha(\alpha(x, v), w) = \alpha(x, v + w)$;
3. if $x, y \in A$, then there exists a unique $v \in V$ satisfying $\alpha(x, v) = y$.

We define the dimension of an affine space to be equal to the dimension of the underlying vector space.

Often, we refer simply to A as the affine space rather than mentioning the triple (A, V, α) with underlying vector space V.

Example 2.37 The real number line is a one-dimensional affine space. More precisely, let A be the underlying set of $\mathbb{R}$, and let $V = \mathbb{R}$. Define

$$\alpha\colon A \times \mathbb{R} \to A, \qquad \alpha(x, v) = x + v.$$

Then $(A, \mathbb{R}, \alpha)$ is a one-dimensional affine space modelled on $\mathbb{R}$. Although the action is defined using addition, no point of A is distinguished as the origin.

The function α can be regarded as a displacement of $a \in A$ by a vector $v \in V$. We may suggestively write $+$ instead of α, in which case the axioms become

$$\begin{aligned} x + 0 &= x \\ (x + v) + w &= x + (v + w) \\ x + v = y &\iff v = y - x. \end{aligned}$$

Remark 2.38 This definition formalizes the intuition that an affine space is a vector space without an origin. Indeed, once we fix a point $p \in A$, one can assign to each vector $v \in V$ a unique element $x \in A$ by setting $x = \alpha(p, v)$. For $x, y \in A$ and $\lambda \in \mathbb{R}$, we define addition and scalar multiplication as

$$\begin{aligned} x + y &:= \alpha(p, (x - p) + (y - p)), \\ \lambda x &:= \alpha(p, \lambda(x - p)). \end{aligned}$$

When the model vector space V is finite-dimensional, an affine space admits coordinate systems after choosing a point that serves as an origin. The affine coordinates of a point $x \in A$ are the coordinates, with respect to a chosen basis of V, of the unique vector that translates the chosen origin to x.

Definition 2.39 Let A be an affine space modelled on an n-dimensional vector space V. Let $(e_1, \ldots, e_n)$ be a basis of V, and let $p \in A$. For each $x \in A$, the **affine coordinates** of x with respect to the affine frame $(p, e_1, \ldots, e_n)$ are the unique real numbers $(x^1, \ldots, x^n)$ such that

$$x = \alpha(p, \sum_{i=1}^{n} x^i e_i).$$

When the origin is changed while the basis of the model vector space is kept fixed, the affine coordinates transform by a translation. Let $q \in A$ be the new origin, with affine coordinates $(q^1, \ldots, q^n)$ with respect to the affine frame $(p, e_1, \ldots, e_n)$. Thus,

$$q = \alpha(p, \sum_{i=1}^{n} q^i e_i).$$

Let $x \in A$ have affine coordinates $(x^1, \ldots, x^n)$ with respect to $(p, e_1, \ldots, e_n)$. Starting from the original origin p, we rewrite the expression for x relative to the new origin q as follows:

$$\begin{aligned} x &= \alpha(p, \sum_{i=1}^{n} x^i e_i) \\ &= \alpha(p, \sum_{i=1}^{n} q^i e_i + \sum_{i=1}^{n} (x^i - q^i) e_i) \\ &= \alpha(\alpha(p, \sum_{i=1}^{n} q^i e_i), \sum_{i=1}^{n} (x^i - q^i) e_i) \\ &= \alpha(q, \sum_{i=1}^{n} (x^i - q^i) e_i). \end{aligned}$$

Therefore, the affine coordinates of x with respect to the affine frame $(q, e_1, \ldots, e_n)$ are

$$x^i - q^i, \qquad i = 1, \ldots, n.$$

Similarly, one can change to the basis $\{f_i\}$ obtained from $\{e_i\}$ by a linear transformation

$$e_i = T_i^j f_j.$$

It follows that x has affine coordinates $x^i T_i^j$ with respect to $\{f_j\}$.

We are interested in describing morphisms between affine spaces, since they will be essential for understanding coordinate transformations in affine bundles.

Definition 2.40 Let (A, V, α) and (B, W, β) be affine spaces. A map $T\colon A \to B$ is called an **affine morphism** if there exists a linear map

$$\overline{T}\colon V \to W$$

such that, for every $x \in A$ and $v \in V$,

$$T(\alpha(x, v)) = \beta\big(T(x), \overline{T}(v)\big).$$

Remark 2.41 The map $\overline{T}$ is called the **linear part** of T. It is uniquely determined by T. Furthermore, T is an affine isomorphism if and only if its linear part $\overline{T}$ is a linear isomorphism.

To obtain the coordinate representation of an affine morphism $T\colon A \to B$, choose affine frames $(p, e_1, \ldots, e_n)$ and $(q, f_1, \ldots, f_m)$ on A and B, respectively.

Let $(c^1, \ldots, c^m)$ be the affine coordinates of $T(p)$ with respect to the affine frame $(q, f_1, \ldots, f_m)$. Thus,

$$T(p) = \beta(q, c^j f_j).$$

Write the linear part of T in the chosen bases as

$$\overline{T}(e_i) = T_i{}^j f_j.$$

Now let $x \in A$ have affine coordinates $(x^1, \ldots, x^n)$ with respect to $(p, e_1, \ldots, e_n)$. Then

$$\begin{aligned}
T(x) &= T\big(\alpha(p, x^i e_i)\big) \\
&= \beta\big(T(p), \overline{T}(x^i e_i)\big) \\
&= \beta\big(\beta(q, c^j f_j), x^i \overline{T}(e_i)\big) \\
&= \beta\big(\beta(q, c^j f_j), x^i T_i{}^j f_j\big) \\
&= \beta\big(q, (c^j + T_i{}^j x^i) f_j\big).
\end{aligned}$$

Therefore, $(c^1 + T_i{}^1 x^i, \ldots, c^m + T_i{}^m x^i)$ are the affine coordinates of $T(x)$ with respect to $(q, f_1, \ldots, f_m)$. Thus, an affine morphism is represented by a linear transformation followed by a translation.

Having discussed affine spaces and their morphisms, we now introduce the corresponding notion of an affine bundle.

Definition 2.42 Let (E, π, M) be a vector bundle. An **affine bundle modelled on** π is a quadruple (A, ρ, M, α) such that

1. (A, ρ, M) is a bundle
2. (a) $\alpha\colon A \times_M E \longrightarrow A$ satisfies, for each $p \in M$, $\alpha(A_p \times E_p) \subset A_p$
   (b) for each $p \in M$, $(A_p, E_p, \alpha|_{A_p \times E_p})$ is an affine space
3. for each $p \in M$ there exists an open neighbourhood $W \subset M$ with an **affine local trivialization**

   $$h\colon A|_W := \rho^{-1}(W) \xrightarrow{\sim} W \times \mathbb{R}^n$$

   such that the restriction to each fibre

   $$h_x := h|_{A_x}\colon A_x \to \{x\} \times \mathbb{R}^n$$

   is an affine isomorphism onto $\{x\} \times \mathbb{R}^n$.

Example 2.43 A vector bundle (E, π, M) can be regarded as an affine bundle (E, π, M, α) by defining

$$\alpha\colon E \times_M E \to E$$

to be fibrewise addition. Thus, for $u, v \in E_p$,

$$\alpha(u, v) = u + v.$$

Conversely, a global section of an affine bundle determines a vector bundle structure on it by selecting an origin in each fibre. This is analogous to the fact that an affine space becomes a vector space once a base point has been chosen.

Let (A, ρ, M, α) be an affine bundle modelled on the vector bundle (E, π, M) with local coordinates (x^i, u^j). To describe this affine bundle locally, we introduce **local affine coordinates** (x^i, a^j). These coordinates (x^i, a^j) have the property that a^j is an affine map on each fibre, whose linear part is the corresponding vector bundle coordinate map u^j.

The last part of this subsection introduces affine bundle morphisms. These are bundle morphisms whose restrictions to the corresponding fibres are affine maps.

Definition 2.44 Let (A, ρ, M, α) and (B, σ, N, β) be affine bundles. A bundle morphism $(f, \overline{f})\colon A \to B$ is called an **affine bundle morphism** if, for each $p \in M$, $f|_{A_p}\colon A_p \to B_{\overline{f}(p)}$ is an affine morphism.

As in the case of vector bundles, one can define a local matrix representation of an affine bundle morphism. Let A and B be modelled on E and H, respectively. Analogously to the case of affine spaces, for a neighbourhood $U \subset M$ we choose an affine local frame (z, e_i) where $z \in \Gamma(A|_U)$ is a section that serves as a choice of base point in each fibre and $\{e_i\}$ is a local frame of $E|_U$. Moreover, let (s, h_j) be an affine local frame for $B|_{\overline{f}(U)}$. We express the matrix representation of the linear part of f, denoted by $\overline{f}$, for every $p \in U \subset M$ as

$$\overline{f}_p(e_j(p)) = M^i_j(p) h_i(\overline{f}(p)).$$

For $x \in A_p$ with affine coordinates x^j, the local matrix representation of the affine morphism f is then given by

$$\begin{aligned} f(x) &= f(\alpha(z(p), x^j e_j(p))) \\ &= \beta(f(z(p)), \overline{f}(x^j e_j(p))) \\ &= \beta(s, q^i h_i + x^j \overline{f}(e_j(p))) \\ &= \beta(s, (q^i + x^j M^i_j) h_i(\overline{f}(p))). \end{aligned}$$

The matrix with components $q^i + x^j M^i_j$ is the local representation of the affine bundle morphism f.

2.2.3 The vertical bundle

In the study of fibre bundles, questions concerning tangent vectors to the fibres arise naturally. This subsection introduces the vertical bundle and develops the basic tools needed to work with these fibrewise tangent spaces. The vertical bundle will later be used to reformulate the momentum bundle in Subsection 4.2.2.

We omit most proofs which can be found in Saunders [70, Chapter 3].

We begin by precisely defining tangent vectors along the fibres.

Definition 2.45 Let $\pi\colon E \to M$ be a fibre bundle over a base manifold M. For $e \in E$ we define the **vertical space** to be $V_eE = \ker(d\pi_e)$, where $d\pi_e\colon T_eE \to T_xM$ is the differential with $\pi(e) = x$. The collection of all vertical spaces is called the **vertical bundle**

$$VE = \bigsqcup_{e\in E} V_eE.$$

The following proposition shows that the vertical bundle is indeed a vector bundle.

Proposition 2.46 *Let $\pi\colon E \to M$ be a fibre bundle. Then VE is a vector subbundle of TE.*

Proof. We show that VE has a local frame which is a subset of the local frame of TE. Since π is a submersion, it follows by the rank-nullity theorem that the kernel has constant rank, i.e.,

$$\dim(T_eE) - \dim(T_xM) = \dim(\ker(d\pi_e)).$$

Let $U \subset M$ be a trivializing neighbourhood of the fibre bundle E so that $E|_U = \pi^{-1}(U) \cong U \times F$, where F is the typical fibre. By [54, Prop. 3.14], we find the tangent space in this neighbourhood to be

$$TE_{(x,f)}|_U \cong T_xU \oplus T_fF.$$

A calculation with curves shows that, on this trivializing neighbourhood $d\pi|_U = \mathrm{pr}_1$. We conclude that $\ker d\pi|_U = TF$. For a local frame $\{v_1, \dots, v_n, y_1, \dots, y_k\}$ of $TU \oplus TF \cong TE|_U$, we find that $V_eE = \ker(d\pi_e) = \mathrm{span}\{y_1(e), \dots, y_k(e)\}$ for all $e \in \pi^{-1}(U) \subset E$. This shows that $VE \subset TE$ is a subbundle. $\square$

As mentioned in the introduction, the vertical space at a point of the total space is precisely the tangent space to the fibre through that point. More precisely, we have the following lemma.

Lemma 2.47 *Let (E, π, M) be a bundle. If $a \in E$, then $T_a(E_{\pi(a)}) \cong V_aE$.*

The vertical bundle also provides a useful criterion for determining whether a smooth map between fibre bundles is a bundle morphism.

Lemma 2.48 *Let (E, π, M) and (F, ρ, N) be bundles, and let $f\colon E \to F$ be a smooth map. Then f is a bundle morphism if and only if $df(VE) \subset VF$.*

For proofs of these statements, see [70, Lemmas 3.1.2 and 3.1.3].

2.3 Jet Bundles

This section provides a brief introduction to the geometry of jet bundles. We define jet bundles, discuss their basic properties, and present several examples. Jet bundles encode derivatives of local sections in a coordinate-independent manner. Therefore, they provide a geometric framework for describing partial differential equations on manifolds. In this thesis, they play a central role in the geometric formulation of the Lagrangian as a bundle morphism in Subsection 4.2.1.

In Subsection 2.3.1, we introduce the equivalence relation on local sections that defines the first jet manifold. In Subsection 2.3.2, we make the affine structure of jet bundles explicit. We also explain how bundle morphisms extend to morphisms between jet bundles, called prolongations. These prolongations are used in Subsection 4.3.2 to formulate the construction more geometrically.

Our treatment follows Saunders [70, Chapter 4].

2.3.1 Equivalence of local sections

Recall that a tangent vector at a point can be defined as an equivalence class of curves through that point that have the same first derivative there. The notion of a first jet manifold generalizes this idea to sections of a bundle. Informally, two sections are equivalent at a point if they have the same value and the same derivative in every direction.

Definition 2.49 Let (E, π, M) be a bundle and let $p \in M$. Let $\phi, \psi \in \Gamma_p(E)$, where $\Gamma_p(E)$ denotes the set of local sections defined on a neighbourhood of $p \in M$. We say ϕ and ψ are **1-equivalent at** p if $\phi(p) = \psi(p)$ and if, in a coordinate system (x^i, u^a) around $\phi(p)$,

$$\left.\frac{\partial \phi^a}{\partial x^i}\right|_p = \left.\frac{\partial \psi^a}{\partial x^i}\right|_p \tag{2.5}$$

for $1 \leq i \leq m$ and $1 \leq a \leq n$. The equivalence class containing ϕ is called the **1-jet** of ϕ at p and is denoted by $j^1_p\phi$.

Although the definition is formulated in terms of local coordinates, the following lemma shows that the resulting notion of 1-equivalence is independent of the choice of coordinates.

Lemma 2.50 *Let (E, π, M) be a bundle, and let $p \in M$. Suppose that $\phi, \psi \in \Gamma_p(E)$ satisfy $\phi(p) = \psi(p)$. Let (x^i, u^a) and (y^j, v^b) be two coordinate systems around $\phi(p)$, and suppose also that*

$$\left.\frac{\partial (u^a \circ \phi)}{\partial x^i}\right|_p = \left.\frac{\partial (u^a \circ \psi)}{\partial x^i}\right|_p \qquad \text{for } 1 \leq i \leq m \text{ and } 1 \leq a \leq n.$$

Then

$$\left.\frac{\partial (v^b \circ \phi)}{\partial y^j}\right|_p = \left.\frac{\partial (v^b \circ \psi)}{\partial y^j}\right|_p \qquad \text{for } 1 \leq j \leq m \text{ and } 1 \leq b \leq n.$$

Proof. By the chain rule,

$$\left.\frac{\partial (v^b \circ \phi)}{\partial y^j}\right|_p = \left.\frac{\partial (v^b \circ \phi)}{\partial x^i}\right|_p \left.\frac{\partial x^i}{\partial y^j}\right|_p.$$

Since $v^b \circ \phi = v^b(x^i \circ \phi, u^a \circ \phi)$ and $(u^a \circ \phi)(x)$ depends on local coordinates x^i, we compute

$$\left.\frac{\partial (v^b \circ \phi)}{\partial x^i}\right|_p = \left.\frac{\partial v^b}{\partial x^i}\right|_{\phi(p)} \left.\frac{\partial (x^i \circ \phi)}{\partial x^i}\right|_p + \left.\frac{\partial v^b}{\partial u^a}\right|_{\phi(p)} \left.\frac{\partial (u^a \circ \phi)}{\partial x^i}\right|_p.$$

Note that $x^i \circ \phi = x^i$, hence

$$\left.\frac{\partial(v^b \circ \phi)}{\partial y^j}\right|_p = \left(\left.\frac{\partial v^b}{\partial x^i}\right|_{\phi(p)} + \left.\frac{\partial v^b}{\partial u^a}\right|_{\phi(p)} \left.\frac{\partial(u^a \circ \phi)}{\partial x^i}\right|_p \right) \left.\frac{\partial x^i}{\partial y^j}\right|_p .$$

The same computation holds for ψ. Since $\phi(p) = \psi(p)$, the partial derivatives of v^b agree, and by hypothesis

$$\left.\frac{\partial(u^a \circ \phi)}{\partial x^i}\right|_p = \left.\frac{\partial(u^a \circ \psi)}{\partial x^i}\right|_p .$$

It follows that

$$\left.\frac{\partial(v^b \circ \phi)}{\partial y^j}\right|_p = \left.\frac{\partial(v^b \circ \psi)}{\partial y^j}\right|_p .$$

This proves the claim. □

Having shown that 1-equivalence is independent of the choice of coordinates, we now consider the set of all equivalence classes and introduce natural coordinate charts on this set. Then we show in Proposition 2.53 that these charts define a smooth manifold structure.

Definition 2.51 The **first jet manifold** of a fibre bundle (E, π, M) is the set

$$J^1E := \{ j^1_p\phi \mid p \in M,\ \phi \in \Gamma_p(E) \}.$$

The functions π_1 and $\pi_{1,0}$, called the **source** and **target** projections, respectively, are defined by

$$\begin{aligned} \pi_1 \colon J^1E &\to M \\ j^1_p\phi &\mapsto p, \end{aligned}$$

and

$$\begin{aligned} \pi_{1,0} \colon J^1E &\to E \\ j^1_p\phi &\mapsto \phi(p). \end{aligned}$$

Definition 2.52 Let (E, π, M) be a bundle, and let (U, u) be a coordinate chart on E, where $u = (x^i, u^a)$ with $1 \le i \le m$ and $1 \le a \le n$. The **induced coordinate chart** (U^1, u^1) on J^1E is defined by

$$\begin{aligned} U^1 &= \{ j^1_p\phi \in J^1E \mid \phi(p) \in U \} \\ u^1 &= (x^i, u^a, u^a_i), \end{aligned}$$

where $x^i(j^1_p\phi) = x^i(p)$, $u^a(j^1_p\phi) = u^a(\phi(p))$, and the mn new functions

$$u^a_i \colon U^1 \to \mathbb{R}$$

are specified by

$$u^a_i(j^1_p\phi) = \left.\frac{\partial \phi^a}{\partial x^i}\right|_p$$

and are known as **derivative coordinates**.

The notation for the derivative coordinates is somewhat unfortunate. For example, upper and lower indices are also used for coefficients of covariant and contravariant tensors. In later chapters, there will be instances when upper and lower indices are also used in contexts unrelated to the jet bundle. To avoid confusion, we will write the derivative coordinates in full. Thus, instead of u_i^α, we write $\frac{\partial u^\alpha}{\partial x^i}$.

The following result justifies the terminology used in Definition 2.51.

Proposition 2.53 *Given a smooth structure on E consisting of the charts (U, u), there is an induced smooth structure on J^1E consisting of charts (U^1, u^1). Furthermore, J^1E is Hausdorff, second-countable and connected, making J^1E into a manifold.*

Sketch of proof. First note that every 1-jet $j_p^1\phi$ lies in the domain of some induced coordinate chart, since one can choose a coordinate neighbourhood U containing $\phi(p)$.

To obtain a smooth structure, it remains to verify that the transition functions between induced coordinate charts are smooth. Let $u^1 := (x^i, u^a, \frac{\partial u^a}{\partial x^i})$ be coordinates on U^1 and let $v^1 := (y^j, v^b, \frac{\partial v^b}{\partial y^j})$ be coordinates on V^1. On the overlap $u^1(U^1 \cap V^1)$ we have the transition function

$$v^1 \circ (u^1)^{-1}\big|_{u^1(U^1\cap V^1)}.$$

The first and second components of the transition function are given by

$$\begin{aligned} y^j \circ (u^1)^{-1} &= y^j \circ u^{-1} \circ \mathrm{pr}_1 \\ v^b \circ (u^1)^{-1} &= v^b \circ u^{-1} \circ \mathrm{pr}_1, \end{aligned}$$

and therefore are smooth.

For the third component, we compute

$$\begin{aligned} \frac{\partial v^b}{\partial y^j}(j_p^1\phi) &= \left.\frac{\partial(v^b\circ\phi)}{\partial y^j}\right|_p \\ &= \left(\left.\frac{\partial v^b}{\partial x^i}\right|_{\phi(p)} + \left.\frac{\partial v^b}{\partial u^a}\right|_{\phi(p)} \left.\frac{\partial u^a}{\partial x^i}\right|_p (j_p^1\phi)\right) \left.\frac{\partial x^i}{\partial y^j}\right|_p. \end{aligned}$$

Therefore each $\frac{\partial v^b}{\partial y^j} \circ (u^1)^{-1}$ is a smooth function. Hence, all components of the transition functions are smooth, and the induced coordinate charts define a smooth structure on J^1E.

The Hausdorff, second-countable, and connected properties of J^1E follow from topological arguments. We omit these details, since including them would not clarify the present discussion and they are not essential for the thesis. □

The following example, taken from [70, Chapter 4], illustrates concretely the information encoded by a first jet.

Example 2.54 Consider the trivial line bundle $(\mathbb{R}^2 \times \mathbb{R}, \mathrm{pr}_1, \mathbb{R}^2)$. Let (x^1, x^2, u^1) be global coordinates on $\mathbb{R}^2 \times \mathbb{R}$. Then the induced coordinates on $J^1(\mathbb{R}^2 \times \mathbb{R})$ are

$$\left(x^1, x^2, u^1, \frac{\partial u^1}{\partial x^1}, \frac{\partial u^1}{\partial x^2}\right),$$

where $\phi^1 = u^1 \circ \phi$. To each 1-jet $j^1_p\phi \in J^1(\mathbb{R}^2 \times \mathbb{R})$, where $p = (p^1, p^2) \in \mathbb{R}^2$, we associate a map $\overline{\psi}\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\overline{\psi}(q) = \phi^1(p) + \left.\frac{\partial \phi^1}{\partial x^1}\right|_p (q^1 - p^1) + \left.\frac{\partial \phi^1}{\partial x^2}\right|_p (q^2 - p^2),$$

where $q = (q^1, q^2) \in \mathbb{R}^2$. We then define a global section $\psi = (\mathrm{id}_{\mathbb{R}^2}, \overline{\psi})$, and it follows that $j^1_p\phi = j^1_p\psi$. But $\overline{\psi}$ is simply the first-order Taylor polynomial of ϕ at p. This example shows that the first jet captures, in a coordinate-free manner, the same derivative information as this polynomial.

2.3.2 The first jet bundle

Having introduced 1-jets and the first jet manifold, we now study its bundle structure. We show that the first jet manifold naturally carries the structure of an affine bundle. We then introduce first-order differential equations and illustrate the use of jet bundles with an example. Finally, we introduce prolongations, which extend bundle morphisms to morphisms between jet bundles.

We begin by describing the bundle structures carried by the first jet manifold. One can show that both the source and target projections, denoted by π_1 and $\pi_{1,0}$, respectively, are smooth surjective submersions. The affine-bundle structure associated with $\pi_{1,0}$ is described by the following theorem.

Theorem 2.55 *Let (E, π, M) be a smooth bundle, and let (x^i, u^α) be local coordinates on E around $a \in E$. Then $(J^1E, \pi_{1,0}, E, \sigma)$ is an affine bundle modelled on the vector bundle*

$$\pi^*(T^*M) \otimes VE \to E,$$

where the action

$$\sigma\colon J^1E \times_E \big(\pi^*(T^*M) \otimes VE\big) \to J^1E$$

is given by

$$(j^1_{\pi(a)}\phi, \xi) \longmapsto \sigma(j^1_{\pi(a)}\phi, \xi),$$

and transforms the derivative coordinates as follows:

$$\frac{\partial u^\alpha}{\partial x^i}\big(\sigma(j^1_{\pi(a)}\phi, \xi)\big) = \frac{\partial u^\alpha}{\partial x^i}(j^1_{\pi(a)}\phi) + \xi^\alpha_i = \left.\frac{\partial \phi^\alpha}{\partial x^i}\right|_{\pi(a)} + \xi^\alpha_i,$$

where

$$\xi = \xi^\alpha_i \left(dx^i \otimes \frac{\partial}{\partial u^\alpha}\right)_a.$$

Furthermore, for each chart (U, u) on E, we obtain an affine local trivialization

$$t_u\colon \pi^{-1}_{1,0}(U) \to U \times \mathbb{R}^{nm}, \qquad j^1_p\phi \mapsto \left(\phi(p), \left(\frac{\partial u^\alpha}{\partial x^i}(j^1_p\phi)\right)\right).$$

Proof. The coordinate independence of the action σ follows from a computation similar to the one in Lemma 2.50. It remains to show that t_u is an affine local trivialization. The map t_u is given by the composition

$$t_u = (u^{-1} \times \mathrm{id}_{\mathbb{R}^{mn}}) \circ u^1.$$

Since $\mathrm{pr}_1 \circ t_u = \pi_{1,0}|_{U^1}$, we conclude that t_u is a diffeomorphism.

To verify that the restriction of t_u to each fibre over $a \in U$ is an affine morphism, we define

$$t_{u;a} := \mathrm{pr}_2 \circ t_u\big|_{\pi_{1,0}^{-1}(a)} \colon \pi_{1,0}^{-1}(a) \to \mathbb{R}^{nm}$$

given by

$$t_{u;a}(j^1_{\pi(a)}\phi) = \left(\frac{\partial u^\alpha}{\partial x^i}(j^1_{\pi(a)}\phi)\right) = \left(\frac{\partial \phi^\alpha}{\partial x^i}\bigg|_{\pi(a)}\right).$$

To show that $t_{u;a}$ is an affine map, let

$$\beta \colon \mathbb{R}^{nm} \times \mathbb{R}^{nm} \to \mathbb{R}^{nm}, \qquad \beta(z, v) = z + v,$$

be the standard affine action on $\mathbb{R}^{nm}$. Let

$$\bar{t}_{u;a} \colon T^*_{\pi(a)}M \otimes V_aE \to \mathbb{R}^{nm}$$

be the linear map

$$\xi = \xi^\alpha_i \left(dx^i \otimes \frac{\partial}{\partial u^\alpha}\right)_a \mapsto (\xi^\alpha_i).$$

Then, for every $j^1_{\pi(a)}\phi \in \pi_{1,0}^{-1}(a)$ and every $\xi \in T^*_{\pi(a)}M \otimes V_aE$, we have

$$\begin{aligned}
t_{u;a}\big(\sigma(j^1_{\pi(a)}\phi, \xi)\big) &= \left(\frac{\partial u^\alpha}{\partial x^i}\sigma(j^1_{\pi(a)}\phi, \xi)\right) \\
&= \left(\frac{\partial u^\alpha}{\partial x^i}(j^1_{\pi(a)}\phi) + \xi^\alpha_i\right) \\
&= \beta\big(t_{u;a}(j^1_{\pi(a)}\phi), \bar{t}_{u;a}(\xi)\big).
\end{aligned}$$

Hence, $t_{u;a}$ preserves the affine structure and is therefore an affine morphism in the sense of Definition 2.40. Since this holds for every $a \in U$ and t_u is a diffeomorphism, we conclude that t_u is an affine local trivialization. □

The affine structure of the jet bundle has several further consequences. For example, the pullback bundle of TE along $\pi_{1,0}$ admits a canonical decomposition, and sections of this affine bundle correspond to connections on the bundle E. These topics lie beyond the scope of this thesis, but can be found in [70, Chapter 4].

We now return to the two projections of the first jet manifold. By Theorem 2.55, the target projection

$$\pi_{1,0} \colon J^1E \to E$$

makes J^1E into an affine bundle over E. The source projection

$$\pi_1 \colon J^1E \to M$$

also defines a fibre-bundle structure, as stated in the following proposition.

Proposition 2.56 *Let (E, π, M) be a fibre bundle. Then (J^1E, π_1, M) is a fibre bundle.*

The bundle structures established in Theorem 2.55 and proposition 2.56 fit into the following commutative diagram. The vertical maps are fibre-bundle projections, while $\pi_{1,0}$ is an affine-bundle projection.

$$\begin{tikzcd} J^1E \arrow[r, "\pi_{1,0}"] \arrow[d, "\pi_1"'] & E \arrow[d, "\pi"] \\ M \arrow[r, "\mathrm{id}_M"'] & M \end{tikzcd}$$

An important application of jet bundles is the formulation of differential equations in a coordinate-free manner.

Definition 2.57 Let (E, π, M) be a bundle. A **first-order differential equation** on E is a closed embedded submanifold S of the first jet manifold J^1E. A **solution** of the differential equation S is a local section $\phi \in \Gamma(E|_W)$, for some $W \subset M$, such that $j^1_p(\phi) \in S$ for every $p \in W$.

The following example, taken from [70, Chapter 4], illustrates this somewhat abstract definition.

Example 2.58 Consider the trivial line bundle $(\mathbb{R}^2 \times \mathbb{R}, \mathrm{pr}_1, \mathbb{R}^2)$ with global coordinates (x^1, x^2, u^1). The map $F\colon J^1E \to \mathbb{R}$ defined by

$$F = \frac{\partial u^1}{\partial x^1}\frac{\partial u^1}{\partial x^2} - 2x^2u^1,$$

gives rise to an embedded submanifold

$$S := \{j^1_p\phi \in J^1E \mid (\frac{\partial u^1}{\partial x^1}\frac{\partial u^1}{\partial x^2} - 2x^2u^1)(j^1_p\phi) = 0\}.$$

Hence, S is a differential equation. The section $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2 \times \mathbb{R}$ defined by

$$\phi(p^1, p^2) = (p^1, p^2, p^1(p^2)^2)$$

is a solution of the differential equation S.

We conclude this chapter by introducing the notion of prolongation. Intuitively, prolongation extends a morphism between fibre bundles to a morphism between their first jet bundles.

Definition 2.59 Let (E, π, M) and (F, ρ, N) be fibre bundles, and let $(f, \overline{f})$ be a bundle morphism, where

$$f\colon E \to F \qquad \text{and} \qquad \overline{f}\colon M \to N,$$

with $\overline{f}$ a diffeomorphism. The **first prolongation** of $(f, \overline{f})$ is the map

$$j^1f\colon J^1E \to J^1F$$

defined by

$$j^1f\big(j^1_p\phi\big) := j^1_{\overline{f}(p)}\big(\widetilde{f}(\phi)\big),$$

where, for a local section $\phi \in \Gamma(E|_W)$,

$$\widetilde{f}(\phi) := f \circ \phi \circ \big(\overline{f}|_W\big)^{-1}.$$

Using the chain rule, one can verify that this definition is independent of the choice of local section representing the jet $j^1_p\phi$.

Chapter 3

Hamiltonian mechanics and symplectic geometry

In this chapter, we discuss the concepts from mechanics and symplectic geometry that are used throughout this thesis. The chapter follows a broadly historical approach. More specifically, we show how Hamiltonian mechanics gives rise to symplectic geometry. We explain why symplectic geometry provides a natural framework for formulating Hamiltonian mechanics. Finally, we extend this discussion to field theories and introduce dynamical systems. These dynamical systems mimic the structure of Hamiltonian mechanics and therefore allow us to describe field theories within a Hamiltonian framework.

We begin in Section 3.1 by discussing Lagrangian mechanics. We introduce the Euler–Lagrange equations and explain how the Legendre transform is used to pass from Lagrangian to Hamiltonian mechanics. In Section 3.2, we discuss several important features of Hamiltonian mechanics, with particular emphasis on the Hamiltonian vector field. We then introduce symplectic geometry and formulate Hamiltonian mechanics in symplectic terms. Furthermore, we introduce the Liouville one-form construction, which will be critical for defining a symplectic form in Chapter 4. Finally, in Section 3.3, we introduce dynamical systems and explain how they mimic the Hamiltonian structure. These dynamical systems will later be used to formulate general relativity in Hamiltonian terms.

In the first part of this chapter, we mainly follow Arnold [3] and McDuff [63]. The final section is based on the Hamiltonian framework developed by Chruściel, Jezierski, and Kijowski in [27].

3.1 From Lagrangian to Hamiltonian mechanics

In this section, we introduce Hamiltonian mechanics. In Subsection 3.1.1, we introduce the calculus of variations and derive the equations of motion from the Lagrangian using variational methods. We also consider the example of a bead constrained to move along a circular path. This system would be difficult to analyse directly using Newtonian mechanics, but it can be treated relatively easily using Lagrangian mechanics. In Subsection 3.1.2, we show how to pass from Lagrangian mechanics to Hamiltonian mechanics using the Legendre transform. We then return to the bead example and show that the Hamiltonian formulation yields the same equations of motion. Finally, in Subsection 3.1.3, we define Hamiltonian vector fields and explain how they encode the equations of motion of a mechanical system. We conclude the section with an example illustrating this concept, since it will play an important role throughout the remainder of the chapter.

In this section, we use theorems and definitions from McDuff [63] and Arnold [3].

3.1.1 The Euler–Lagrange equation

In the calculus of variations, we study extremals of functions whose domains are infinite-dimensional spaces. Such functions are called **functionals**. An example of a functional is the length of a curve γ in the Euclidean plane.

Definition 3.1 A functional ϕ is called **differentiable** if

$$\phi(\gamma + h) - \phi(\gamma) = F + R,$$

where F depends linearly on h. The linear part $F(h)$ is called the **differential** or **variation**, and $R(h, \gamma) = O(h^2)$.

We restrict our attention to functionals of the following form. Let

$$\gamma = (t, x(t)) \in I \times \mathbb{R}^n$$

be a curve, and let $L(x(t), \dot{x}(t), t)$ be a function of three variables, where we write

$$\dot{x} := \frac{dx}{dt}.$$

We define the functional

$$\phi(\gamma) = \int_I L(x(t), \dot{x}(t), t)\, dt.$$

Definition 3.2 We call a curve γ an **extremal** of ϕ if $F(h) = 0$ for all h.

The extremals of the functional under consideration satisfy a particular differential equation.

Definition 3.3 The equation

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{x}}\Big) - \frac{\partial L}{\partial x} = 0 \tag{3.1}$$

is called the **Euler–Lagrange equation** associated with the functional

$$\phi = \int_I L(x, \dot{x}, t)\, dt.$$

The Euler–Lagrange equations characterize precisely when a curve is an extremal of the functional defined above.

Theorem 3.4 *A curve γ is an extremal of the functional*

$$\phi(\gamma) = \int_I L(x, \dot{x}, t)\, dt$$

if and only if the Euler–Lagrange equations are satisfied along γ.

In mechanics, the function $L(x, \dot{x}, t)$ is called the **Lagrangian** of the system whenever

$$L = T - U,$$

where T is the kinetic energy and U is the potential energy. In classical mechanics, one assumes the following principle.

Proposition 3.5 *A path γ through physical space is an extremal of the functional*

$$\phi(\gamma) = \int_I L(x, \dot{x}, t)\, dt.$$

This statement is known as **Hamilton's principle of least action**. It implies that the curves satisfying the Euler–Lagrange equations describe the equations of motion of the mechanical system.

To illustrate the theory, we consider the following example.

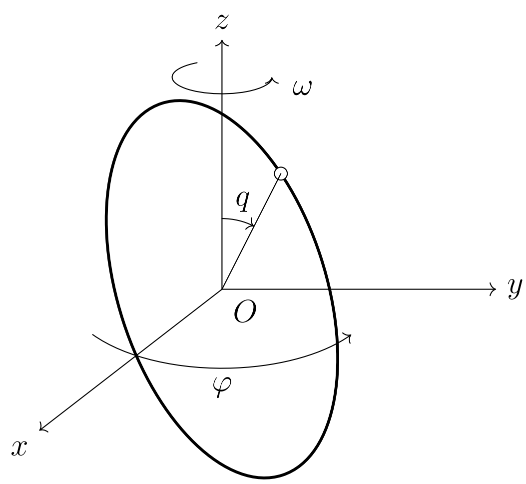

Figure 3.1: Bead on a rotating circle

Example 3.6 We illustrate the Lagrangian formalism with a bead constrained to move along a rotating vertical circle. This example is adapted from Arnold [3], but we present the calculations in greater detail to show explicitly how the Euler–Lagrange equations yield the equations of motion.

Consider a vertical circle of radius r centred at the origin in $\mathbb{R}^3$, as shown in Figure 3.1. We use Cartesian coordinates (x, y, z), with the z-axis vertical. The circle rotates about the z-axis with constant angular velocity ω, while a bead is free to move along the circle. We take the configuration manifold M to be the circle and denote by q the angular position of the bead, measured from its highest point. Let φ denote the angle between the plane of the circle and the vertical plane $y = 0$. Since the angular velocity is constant,

$$\varphi = \omega t.$$

The mapping

$$i\colon M \times \mathbb{R} \longrightarrow \mathbb{R}^3$$

is given by

$$i(q, t) = \big(r\sin(q)\cos(\omega t), r\sin(q)\sin(\omega t), r\cos(q)\big).$$

First, we calculate the kinetic energy

$$T = \frac{1}{2} m \left| \frac{d}{dt} i \right|^2,$$

where $\left|\frac{d}{dt}i\right|$ is the speed and m is the mass. Since $q = q(t)$, a straightforward calculation gives

$$T = \frac{m}{2} \left| \frac{d}{dt} i \right|^2 = \frac{m}{2} \left(\omega^2 r^2 \sin^2 q + r^2 \dot{q}^2 \right).$$

The height of the bead is $z = r\cos(q)$, so its gravitational potential energy is

$$U = mgh = mgr\cos(q).$$

The Lagrangian is therefore

$$\begin{aligned} L = T - U &= \frac{m}{2} \left(\omega^2 r^2 \sin^2 q + r^2 \dot{q}^2 \right) - mgr\cos(q) \\ &= \frac{m}{2} \dot{q}^2 r^2 - \left(mgr\cos(q) - \frac{m}{2} r^2 \sin^2(q) \omega^2 \right) \\ &= T_0 - V(q). \end{aligned}$$

Although the constraint depends explicitly on time, the resulting Lagrangian is independent of t. The final expression shows that the Lagrangian has the form of a one-dimensional mechanical system with kinetic energy

$$T_0 = \frac{M}{2} \dot{q}^2, \qquad M = mr^2,$$

and potential energy

$$V = A\cos q - B\sin^2 q, \qquad A = mgr, \qquad B = \frac{m}{2} \omega^2 r^2.$$

We now derive the equation of motion from the Euler–Lagrange equation.

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) = m\ddot{q} r^2, \qquad \frac{\partial L}{\partial q} = mr^2 \omega^2 \sin(q)\cos(q) + mgr\sin(q).$$

Substituting these formulas into the Euler–Lagrange equations gives

$$\ddot{q} - \omega^2 \sin(q)\cos(q) - \frac{g}{r}\sin(q) = 0.$$

Solving this differential equation determines the motion $q(t)$. We reconstruct the trajectory of the bead in $\mathbb{R}^3$ from

$$(x, y, z) = (r\sin(q(t))\cos(\omega t), r\sin(q(t))\sin(\omega t), r\cos(q(t))).$$

The preceding example illustrates how the Euler–Lagrange equation determines the motion of a mechanical system. To represent the possible states of such a system geometrically, one introduces its phase space. In the Lagrangian formulation, a state is described by its position and velocity. Rather than constructing the phase portrait of the rotating-bead system, we illustrate this idea using the simpler example of a harmonic oscillator.

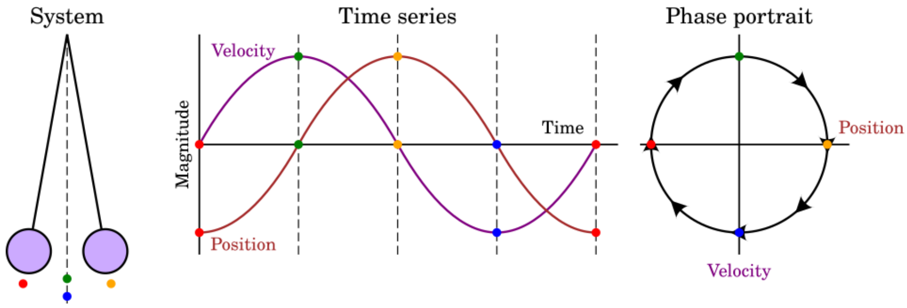

Figure 3.2: Phase portrait of a harmonic oscillator in phase space (position versus velocity). Adapted from [85].

Every mechanical system admits a phase space describing its possible states. In the Lagrangian formulation, a state is determined by its position and velocity, so the phase space is naturally even-dimensional. This observation has a natural geometric interpretation. If the configuration space of the mechanical system is a manifold M, then its velocity phase space is the tangent bundle TM. Although this geometric formulation does not yet offer significant advantages, it provides the natural setting for the transition to Hamiltonian mechanics, where the phase space will instead be the cotangent bundle T^*M.

3.1.2 The Legendre transform and the Hamiltonian

In the previous subsection, we introduced the Lagrangian formulation of mechanics. In this subsection, we explain how to pass from the Lagrangian formulation to the Hamiltonian formulation. In the Lagrangian setting, we derive the equations of motion from the Euler–Lagrange equations. The Legendre transform converts these equations into Hamilton's equations.

The **Legendre condition**

$$\det\left(\frac{\partial^2 L}{\partial \dot{x}_j \partial \dot{x}_k}\right) \neq 0$$

guarantees that the Euler–Lagrange equations form a regular system of second-order differential equations in the n variables $x_1, \ldots, x_n$. Using the **Legendre transform**, this system can locally be rewritten as a regular system of first-order differential equations in $2n$ variables.

Definition 3.7 Let f be a differentiable function and define

$$p := \frac{df}{dx}(x).$$

The **Legendre transform** of f is the function g defined by

$$g(p) = px - f(x),$$

where x is expressed as a function of p.

The Legendre condition ensures that the Legendre transform is locally well defined.

To conform with the standard notation in Hamiltonian mechanics, we now write the position coordinates as q instead of x. No mathematical change is intended.

Proposition 3.8 *Let $L(q, \dot{q}, t)$ be a Lagrangian satisfying the Legendre condition. Define the conjugate momenta by*

$$p_k := \frac{\partial L}{\partial \dot{q}_k}, \qquad k = 1, \dots, n.$$

The Legendre transform of L is given by

$$H(q, p, t) = \sum_{k=1}^{n} p_k \dot{q}_k - L(q, \dot{q}, t),$$

where $\dot{q}$ is expressed locally as a function of (q, p, t). The function H is called the ***Hamiltonian****.*

Proof. We regard $L(q, \dot{q}, t)$ as a function of $\dot{q}$, with q and t fixed. Its Legendre transform is therefore

$$\sum_{k=1}^{n} \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k - L(q, \dot{q}, t).$$

By the Legendre condition and the implicit function theorem, we can express $\dot{q}$ locally as a function of (q, p, t), denoted by

$$\dot{q}_k = G_k(q, p, t), \qquad k = 1, \dots, n.$$

Hence, the Hamiltonian depends only on (q, p, t) and is given by

$$H(q, p, t) = \sum_{k=1}^{n} p_k G_k(q, p, t) - L(q, G(q, p, t), t). \tag{3.2}$$

□

From (3.2) in the proof of Proposition 3.8, we obtain the relations

$$\frac{\partial H}{\partial q_k} = -\frac{\partial L}{\partial q_k}, \qquad \frac{\partial H}{\partial p_k} = G_k.$$

The Euler–Lagrange equations are rewritten as the **Hamiltonian equations of motion**

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}, \qquad \dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad k = 1, \dots, n.$$

Example 3.9 To verify that the Hamiltonian formulation yields the same equation of motion, we return to the mechanical system from Example 3.6, consisting of a bead constrained to move along a rotating circular path. Recall that the Lagrangian is given by

$$\begin{aligned} L &= \frac{m}{2}\dot{q}^2 r^2 - mgr\cos(q) + \frac{m}{2} r^2 \omega^2 \sin^2(q) \\ &= T_0 - V(q). \end{aligned}$$

The conjugate momentum is

$$p = \frac{\partial L}{\partial \dot{q}} = mr^2 \dot{q}.$$

Since $mr^2 > 0$, the Legendre transformation is invertible, with

$$\dot{q} = \frac{p}{mr^2}.$$

The Hamiltonian is defined by

$$\begin{aligned} H(q,p) &= p\dot{q} - L(q,\dot{q}) \\ &= \frac{p^2}{2mr^2} + mgr\cos(q) - \frac{m}{2}\omega^2 r^2 \sin^2(q). \end{aligned}$$

Equivalently,

$$H(q,p) = \frac{p^2}{2mr^2} + V(q).$$

Hamilton's equations are

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{mr^2}$$

and

$$\begin{aligned} \dot{p} &= -\frac{\partial H}{\partial q} \\ &= mgr\sin(q) + m\omega^2 r^2 \sin(q)\cos(q). \end{aligned}$$

Differentiating the first Hamilton equation and substituting the second gives

$$\ddot{q} = \frac{\dot{p}}{mr^2} = \frac{g}{r}\sin(q) + \omega^2 \sin(q)\cos(q).$$

Hence, we obtain the same equation of motion, namely

$$\ddot{q} - \omega^2 \sin(q)\cos(q) - \frac{g}{r}\sin(q) = 0.$$

In a Hamiltonian formulation of mechanics, the variables describing the system are position and momentum rather than position and velocity. Accordingly, the Hamiltonian phase space is formulated in terms of position and momentum. In Example 3.13, we illustrate this using the phase space of a harmonic oscillator. In geometric terms, this amounts to considering the cotangent bundle T^*M of the configuration manifold M instead of its tangent bundle TM.

3.1.3 Hamiltonian vector fields

In this subsection, we introduce a special vector field associated with a Hamiltonian, called the Hamiltonian vector field. Its integral curves describe the equations of motion of the mechanical system. We also introduce coordinate transformations that preserve the Hamiltonian structure. These transformations are called symplectomorphisms and foreshadow the introduction of symplectic geometry in the next section. Finally, we give an example involving a Hamiltonian vector field and a symplectomorphism, since both concepts will be important throughout the remainder of the chapter.

Let

$$H\colon \mathbb{R}^{2n} \to \mathbb{R}$$

be a Hamiltonian, and write

$$z = (q_1, \ldots, q_n, p_1, \ldots, p_n) \in \mathbb{R}^{2n}.$$

The Hamiltonian equations of motion are given by

$$-J_0 \nabla H = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix} = \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix},$$

where

$$J_0 := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

In the Hamiltonian formulation, the vector field associated with the Hamiltonian represents the equations of motion on phase space.

Definition 3.10 The **Hamiltonian vector field** associated with a Hamiltonian H is defined by

$$X_H := -J_0 \nabla H = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}. \tag{3.3}$$

We obtain the equations of motion from the flow of the Hamiltonian vector field. The flow of X_H is a map

$$\phi_H\colon \mathbb{R} \times \mathbb{R}^{2n} \to \mathbb{R}^{2n}, \quad (t, q, p) \mapsto \phi_H^t(q, p),$$

where $\phi_H^t(q, p) = z(t)$ is the trajectory through (q, p). The flow satisfies

$$\frac{\partial}{\partial t}\phi_H^t = X_H \circ \phi_H^t,$$
$$\phi_H^{t_1} \circ \phi_H^{t_0} = \phi_H^{t_1 + t_0}, \qquad \phi_H^0 = \mathrm{id}\,.$$

Such a flow is called a **Hamiltonian flow**.

The Hamiltonian vector field encodes the equations of motion. We now introduce transformations that preserve the Hamiltonian structure.

Definition 3.11 A **symplectomorphism**, historically called a **canonical transformation**, is a diffeomorphism

$$\psi\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$$

such that

$$d\psi(\zeta)^T J_0 d\psi(\zeta) = J_0,$$

for every $\zeta \in \mathbb{R}^{2n}$.

As a first example of a symplectomorphism, for every time for which it is defined, the map ϕ_H^t of a Hamiltonian flow is a symplectomorphism.

The following lemma, quoted from McDuff [63], describes how symplectomorphisms preserve the Hamiltonian structure.

Lemma 3.12 *Let $\psi\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a symplectomorphism and let $H\colon \mathbb{R}^{2n} \to \mathbb{R}$ be a smooth Hamiltonian function. Suppose that $\zeta(t)$ is a solution of the Hamiltonian differential equation*

$$\dot{\zeta} = X_{H\circ\psi}(\zeta).$$

Then $z(t) := \psi(\zeta(t))$ is a solution of $\dot{q} = \frac{\partial H}{\partial p}$, $\dot{p} = -\frac{\partial H}{\partial q}$. In other words,

$$X_{H\circ\psi} = \psi^* X_H.$$

The pullback of the vector field is well-defined and defines a vector field since ψ is a diffeomorphism.

Thus, a symplectomorphism transforms the trajectories of $H \circ \psi$ into trajectories of H. The following example verifies this relation explicitly for a harmonic oscillator and a linear symplectomorphism.

Example 3.13 Consider a harmonic oscillator without resistance, whose phase portrait is shown in Figure 3.2. This example has two purposes. First, we illustrate how the Hamiltonian vector field determines trajectories through phase space. Second, we show how a symplectomorphism transforms the Hamiltonian while preserving the relation between the vector field and its trajectories.

The Hamiltonian of the system is

$$H\colon \mathbb{R}^2 \to \mathbb{R}, \qquad H(q,p) = \frac{1}{2}(q^2 + p^2).$$

Hamilton's equations give

$$\dot{q} = p, \qquad \dot{p} = -q.$$

Consequently,

$$\frac{d}{dt}(q^2 + p^2) = 2q\dot{q} + 2p\dot{p} = 0.$$

Thus, the trajectories lie on the level sets of H, which are circles in phase space.

We now introduce the following symplectomorphism

$$\psi\colon \mathbb{R}^2 \to \mathbb{R}^2; \psi(q,p) = (\frac{1}{2}q, 2p).$$

The Jacobian is given by

$$d\psi = \begin{pmatrix} \frac{\partial \psi_1}{\partial q} & \frac{\partial \psi_1}{\partial p} \\ \frac{\partial \psi_2}{\partial q} & \frac{\partial \psi_2}{\partial p} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix}.$$

The transformed Hamiltonian is given by

$$(H \circ \psi)(q,p) = \frac{1}{2}(\frac{1}{4}q^2 + 4p^2).$$

Calculating the Hamiltonian vector field of $H \circ \psi$ gives

$$X_{H\circ\psi} = \begin{pmatrix} \frac{\partial (H\circ\psi)}{\partial p} \\ -\frac{\partial (H\circ\psi)}{\partial q} \end{pmatrix} = \begin{pmatrix} 4p \\ -\frac{1}{4}q \end{pmatrix}.$$

Next, we calculate the pullback of the Hamiltonian vector field X_H by ψ.

$$(\psi^* X_H)(q,p) = d\psi^{-1}(X_H \circ \psi(q,p)) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2p \\ -\frac{1}{2}q \end{pmatrix} = \begin{pmatrix} 4p \\ -\frac{1}{4}q \end{pmatrix}.$$

Here, we have used the inverse Jacobian

$$d\psi^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

together with the Hamiltonian vector field

$$X_H = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix} = \begin{pmatrix} p \\ -q \end{pmatrix}.$$

We therefore obtain $X_{H\circ\psi} = \psi^* X_H$. The two vector fields are shown in Figure 3.3. In both diagrams, the black curve represents a trajectory through phase space, and the arrows of the corresponding Hamiltonian vector field are tangent to this trajectory. The transformation ψ changes the circular trajectories into elliptical trajectories while preserving this relation.

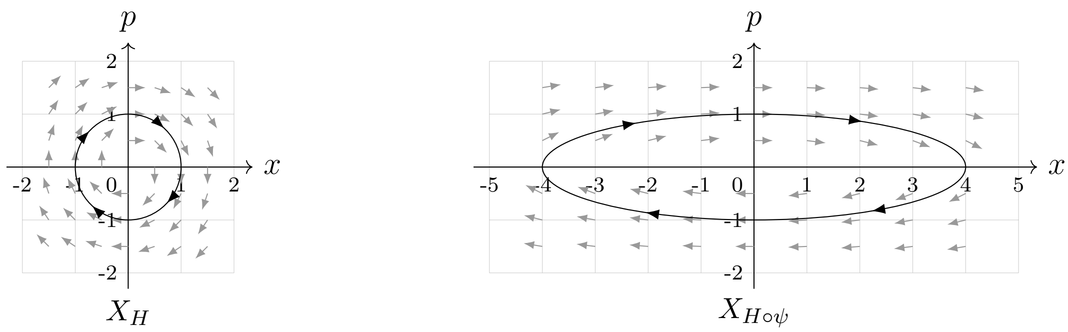

Figure 3.3: The Hamiltonian vector fields X_H and $X_{H\circ\psi}$.

3.2 From Hamiltonian mechanics to symplectic geometry

The previous section explained how to pass from a Lagrangian formulation of mechanics to a Hamiltonian formulation using the Legendre transform. We introduced the Hamiltonian vector field and explained how its flow determines the trajectories of the mechanical system in phase space. Finally, we discussed symplectomorphisms as transformations that preserve the Hamiltonian structure and illustrated their effect on Hamiltonian vector fields in an explicit example.

In this section, we reformulate Hamiltonian mechanics in the language of symplectic geometry. In Subsection 3.2.1, we introduce the canonical symplectic form on Euclidean space and use it to express Hamiltonian mechanics geometrically. We then introduce the historical definition of a symplectic manifold and show, using Darboux's theorem, that it is equivalent to the modern definition in terms of a closed non-degenerate two-form. Finally, in Subsection 3.2.2, we construct the Liouville one-form on a cotangent bundle. This construction will play a critical role in Chapter 4.

For a more detailed treatment of symplectic geometry, we refer to McDuff [63].

3.2.1 Symplectic formulation of Hamiltonian mechanics

We begin by introducing the canonical symplectic form on Euclidean space and use it to reformulate the Hamiltonian equations geometrically.

On $\mathbb{R}^{2n}$, with coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$, we define the 2-form

$$\omega_{\text{can}} := \sum_{j=1}^{n} dq_j \wedge dp_j.$$

This form is called the **canonical symplectic form**.

For vectors $X, Y \in \mathbb{R}^{2n}$, the canonical symplectic form is written as

$$\omega_{\text{can}}(X, Y) = -X^T J_0 Y,$$

where X^T denotes the transpose of X, and

$$J_0 := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

Using this identity, we express the Hamiltonian vector field in terms of the canonical symplectic form.

Lemma 3.14 *Let $H \colon \mathbb{R}^{2n} \to \mathbb{R}$ be a smooth Hamiltonian. Then the Hamiltonian vector field X_H is characterized by the following relation*

$$X_H \rfloor \omega_{\text{can}} = -dH, \tag{3.4}$$

where dH denotes the differential of H.

Proof. For $Y \in \mathbb{R}^{2n}$ we obtain

$$\begin{aligned}
\big(X_H \rfloor \omega_{\text{can}}\big)(Y) &= \omega_{\text{can}}(X_H, Y) \\
&= -X_H^T J_0 Y \\
&= -(J_0 \nabla H)^T J_0 Y \\
&= -(\nabla H)^T J_0^T J_0 Y \\
&= -(\nabla H)^T Y \\
&= -dH(Y).
\end{aligned}$$

Since $J_0^T J_0 = I_{2n}$ and the differential of H is given by

$$dH(Y) = \sum_{j=1}^{n} \frac{\partial H}{\partial q_j} Y_j + \sum_{j=1}^{n} \frac{\partial H}{\partial p_j} Y_{n+j} = (\nabla H)^T Y,$$

we obtain

$$-(\nabla H)^T Y = -dH(Y).$$

Since this holds for every $Y \in \mathbb{R}^{2n}$, we conclude that

$$X_H \rfloor \omega_{\text{can}} = -dH.$$

□

Hence, the symplectic form ω_{can} characterizes the Hamiltonian vector field and therefore determines the trajectories of the mechanical system. In Lemma 3.12, we showed that symplectomorphisms preserve the Hamiltonian structure. This suggests that a symplectomorphism should preserve the symplectic form itself, as the terminology indicates.

Lemma 3.15 *Let $\psi\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a diffeomorphism. Then ψ is a symplectomorphism if and only if*

$$\psi^*\omega_{\text{can}} = \omega_{\text{can}}.$$

Proof. Assume first that ψ is a symplectomorphism. For $X, Y \in \mathbb{R}^{2n}$, we have

$$\begin{aligned}
\psi^*\omega_{\text{can}}(X,Y) &= \omega_{\text{can}}(d\psi X, d\psi Y)\\
&= -(d\psi X)^T J_0 d\psi Y\\
&= -X^T d\psi^T J_0 d\psi Y\\
&= -X^T J_0 Y\\
&= \omega_{\text{can}}(X,Y).
\end{aligned}$$

Hence,

$$\psi^*\omega_{\text{can}} = \omega_{\text{can}}.$$

Conversely, assume that

$$\psi^*\omega_{\text{can}} = \omega_{\text{can}}.$$

Then, for all $X, Y \in \mathbb{R}^{2n}$,

$$-X^T d\psi^T J_0 d\psi Y = -X^T J_0 Y.$$

Therefore, it follows that

$$d\psi^T J_0 d\psi = J_0.$$

Hence, ψ is a symplectomorphism. □

We conclude that the symplectic form also characterizes symplectomorphisms.

Definition 3.16 Let M be a manifold equipped with an atlas whose transition maps are symplectomorphisms. We call such a manifold a **symplectic manifold**.

Thus, a symplectic manifold is a manifold whose coordinate changes preserve the symplectic, and hence the Hamiltonian, structure. This shows why symplectic geometry provides a natural framework for Hamiltonian mechanics.

We now relate the historical, atlas-based description of a symplectic manifold to the modern formulation in terms of differential forms.

Proposition 3.17 *Let M be a symplectic manifold. Then there exists a global closed non-degenerate two-form $\omega \in \Omega^2(M)$.*

Proof. Let (U_1, χ) and (U_2, ξ) be coordinate charts belonging to the symplectic atlas. On the overlap $U_1 \cap U_2$, the transition map is a symplectomorphism, so

$$(\chi \circ \xi^{-1})^*\omega_{\text{can}} = \omega_{\text{can}}.$$

Define

$$\omega_\chi = \chi^*\omega_{\text{can}} \qquad \text{and} \qquad \omega_\xi = \xi^*\omega_{\text{can}}$$

on U_1 and U_2, respectively. On the overlap $U_1 \cap U_2$, we have

$$\begin{aligned} \omega_\chi &= \chi^* \omega_{\mathrm{can}} \\ &= (\chi \circ \xi^{-1} \circ \xi)^* \omega_{\mathrm{can}} \\ &= \xi^* (\chi \circ \xi^{-1})^* \omega_{\mathrm{can}} \\ &= \xi^* \omega_{\mathrm{can}} \\ &= \omega_\xi. \end{aligned}$$

Hence, the locally defined forms agree on overlaps and therefore determine a global two-form ω on M. Since each local form is the pullback of ω_{can}, the form ω is closed and non-degenerate. □

We have shown that a symplectic atlas determines a global closed non-degenerate two-form. We now introduce the modern definition of a symplectic manifold and prove the converse implication.

Definition 3.18 A **symplectic form** on a manifold M is a two-form $\omega \in \Omega^2(M)$ that is closed and non-degenerate. More precisely,

$$d\omega = 0,$$

and, for every $p \in M$ and $X \in T_pM$,

$$\omega_p(X, Y) = 0 \quad \text{for all } Y \in T_pM \quad \Longrightarrow \quad X = 0.$$

A manifold equipped with a symplectic form is called a **symplectic manifold**.

Remark 3.19 The canonical form ω_{can} is a symplectic form on $\mathbb{R}^{2n}$.

To prove that this modern definition is equivalent to the historical atlas-based definition, we use Darboux's theorem.

Theorem 3.20 (Darboux's theorem) *Let (M, ω) be a $2n$-dimensional symplectic manifold and let $p \in M$. Then there exist neighbourhoods U of p and V of $0 \in \mathbb{R}^{2n}$, together with a diffeomorphism $\varphi \colon U \to V$, such that $\varphi(p) = 0$ and*

$$\varphi^* \omega_{\mathrm{can}} = \omega|_U.$$

A proof can be found in McDuff [63, Theorem 3.2.2].

Proposition 3.21 *Let M be a manifold equipped with a symplectic form $\omega \in \Omega^2(M)$. Then M admits a symplectic atlas.*

Proof. By Darboux's theorem, for every point of M there exists a chart (U_i, φ_i), where

$$\varphi_i \colon U_i \to V_i \subset \mathbb{R}^{2n},$$

such that

$$\varphi_i^* \omega_{\mathrm{can}} = \omega|_{U_i}.$$

On an overlap $U_i \cap U_j$, the transition map satisfies

$$\begin{aligned} (\varphi_i \circ \varphi_j^{-1})^* \omega_{\mathrm{can}} &= (\varphi_j^{-1})^* \varphi_i^* \omega_{\mathrm{can}} \\ &= (\varphi_j^{-1})^* \omega \\ &= \omega_{\mathrm{can}}. \end{aligned}$$

Hence, the transition maps are symplectomorphisms, and the charts (U_i, φ_i) form a symplectic atlas on M. □

Thus, the atlas-based and differential-form definitions of a symplectic manifold are equivalent. In this thesis, we use both formulations interchangeably.

We now define a Hamiltonian vector field in the modern setting.

Definition 3.22 Let (M, ω) be a symplectic manifold. A vector field $X \in \mathfrak{X}(M)$ is called a **Hamiltonian vector field** if the one-form $X \rfloor \omega$ is exact. More explicitly, there exists a smooth function $H\colon M \to \mathbb{R}$ such that

$$X \rfloor \omega = -dH. \tag{3.5}$$

Remark 3.23 Given a Hamiltonian H, the associated Hamiltonian vector field X_H is uniquely determined by the non-degeneracy of ω.

3.2.2 The Liouville one-form on a cotangent bundle

The previous subsection showed how Hamiltonian mechanics is expressed using the symplectic form. This subsection shows how to define a symplectic form on the cotangent bundle of any manifold. Since the phase space associated with a configuration manifold M is its cotangent bundle T^*M, this construction provides the natural symplectic structure used in Hamiltonian mechanics. We now introduce the Liouville one-form from which this symplectic form is obtained.

Proposition 3.24 *Let L be a manifold. Then $(T^*L, \omega_{\text{can}})$ is a symplectic manifold, where*

$$\omega_{\text{can}} = -d\lambda_{\text{can}}.$$

Proof. Let

$$\pi\colon T^*L \to L$$

denote the canonical projection. For $p = (x, \xi) \in T^*L$, where $\xi \in T^*_xL$, we define the Liouville one-form by

$$(\lambda_{\text{can}})_p\colon T_pT^*L \to \mathbb{R}, \quad (\lambda_{\text{can}})_p = (d\pi_p)^*\xi = \xi \circ d\pi_p.$$

We define

$$\omega_{\text{can}} := -d\lambda_{\text{can}}.$$

We verify in local coordinates that this is a symplectic form.

Let $U \subset L$ be an open set with coordinates $(x_1, \dots, x_n)$. Then $(dx_1, \dots, dx_n)$ is a basis for T^*_xL, and every $\xi \in T^*_xL$ can be written as

$$\xi = \sum_{i=1}^{n} \xi_i (dx_i)_x.$$

These coordinates induce coordinates on T^*U given by

$$\begin{aligned} &T^*U \to \mathbb{R}^{2n}, \\ &(x, \xi) \mapsto (x_1, \dots, x_n, \xi_1, \dots, \xi_n). \end{aligned}$$

Since $x_i \circ \pi = x_i$, Lee [54, Proposition 11.25] gives

$$\begin{aligned} (d\pi_p)^*(dx_i)_x &= d(x_i \circ \pi)_p \\ &= (dx_i)_p. \end{aligned}$$

Hence, the Liouville one-form is given in these coordinates by

$$\begin{aligned}(\lambda_{\text{can}})_p &= (d\pi_p)^*\xi \\ &= \sum_{i=1}^{n} \xi_i (dx_i)_p.\end{aligned}$$

Therefore,

$$\begin{aligned}\omega_{\text{can}} &= -d\lambda_{\text{can}} \\ &= -\sum_{i=1}^{n} d\xi_i \wedge dx_i \\ &= \sum_{i=1}^{n} dx_i \wedge d\xi_i.\end{aligned}$$

This is the canonical symplectic form on $\mathbb{R}^{2n}$. In particular, it is closed and non-degenerate, so ω_{can} is a symplectic form on T^*L. $\square$

3.3 From mechanics to field theory

The previous section formulated Hamiltonian mechanics in the language of symplectic geometry. In this section, we generalize the concepts introduced in Subsection 3.2.2 to describe field theories. So far, we have considered systems of particles. We now develop the tools to describe fields and their evolution. In this setting, the phase space is typically a function space and may therefore be infinite-dimensional.

A complete treatment of the infinite-dimensional case would require a careful discussion of the manifold structure of the spaces involved. In particular, one would need suitable notions of tangent vectors and differential forms, as well as notions of non-degeneracy and closedness for the symplectic form. Chruściel, Jezierski and Kijowski avoid these analytical and geometric complications by introducing dynamical systems that mimic the essential structure of Hamiltonian mechanics.

The following definitions and constructions are adapted from Chruściel, Jezierski and Kijowski's treatment of Hamiltonian field theory in [27]. We reproduce them here because they provide the framework used throughout the remainder of this thesis.

3.3.1 Dynamical system

We define the notion of a dynamical system and explain when such a system is Hamiltonian. The idea is to mimic the relation satisfied by a Hamiltonian vector field, so that the resulting dynamical system generates trajectories in phase space.

We first introduce the notion of a differentiable curve.

Definition 3.25 Let $\mathscr{P}$ be a vector space of differentiable functions on a set $\mathscr{O}$. A curve $f_\lambda \in \mathscr{P}$ is called a **differentiable curve** if the family of functions f_λ is differentiable both in $\lambda \in \mathbb{R}$ and in $x \in \mathscr{O}$.

We now introduce the notion of a dynamical system.

Definition 3.26 Let $\mathscr{P}$ be a vector space of differentiable functions on a set $\mathscr{O}$. A family of maps

$$T_\lambda \colon \mathscr{P} \to \mathscr{P}, \qquad \lambda \in \mathbb{R},$$

is called a **differentiable dynamical system** on $\mathscr{P}$ if the following three properties hold:

1. $T_0 = \mathrm{id}$, where id denotes the identity map on $\mathscr{P}$;
2. $T_a \circ T_b = T_{a+b}$ for all $a, b \in \mathbb{R}$;
3. for every $p \in \mathscr{P}$, the map
$$\lambda \mapsto T_\lambda(p)$$
is a differentiable curve in $\mathscr{P}$.

Furthermore, we define the **vector field generating the dynamics** as

$$\mathscr{X}(p) = \frac{d}{d\lambda} T_\lambda(p) \bigg|_{\lambda=0}.$$

It is important to note that the fundamental objects are the maps T_λ. No assumption is made that the family T_λ can be recovered from its generating vector field $\mathscr{X}$.

Before introducing Hamiltonian dynamical systems, we fix an antisymmetric two-form

$$\Omega \colon \mathscr{P} \times \mathscr{P} \to \mathbb{R}.$$

Remark 3.27 The precise non-degeneracy conditions that should be imposed on Ω require further study. In the framework adopted from Chruściel, Jezierski and Kijowski [27], these conditions are not stated explicitly.

Definition 3.28 Let $\mathscr{P}$ be a vector space of differentiable functions. A dynamical system T_λ on $\mathscr{P}$ is said to be **Hamiltonian** if there exists a function $H \colon \mathscr{P} \to \mathbb{R}$ such that, for every differentiable curve f_σ in $\mathscr{P}$, the map

$$\sigma \mapsto H(f_\sigma)$$

is differentiable at $\sigma = 0$ and satisfies

$$\frac{d}{d\sigma} H(f_\sigma) \bigg|_{\sigma=0} = -\Omega \left(\mathscr{X}, \frac{df_\sigma}{d\sigma} \bigg|_{\sigma=0} \right).$$

The function H is called a **Hamiltonian** for the dynamical system $(\mathscr{P}, T_\lambda)$.

Remark 3.29 The dynamical systems considered here are autonomous. Later, the Hamiltonian will be constructed on a spacelike hypersurface at a fixed value of the time parameter. For each fixed time, the resulting Hamiltonian is therefore treated as generating an autonomous dynamical system.

The family of maps T_λ plays the role of the flow of a vector field. However, since we have not introduced the required geometric structure on $\mathscr{P}$, we do not attempt to formulate this relation rigorously. Recall that, in finite-dimensional Hamiltonian mechanics, the flow of the Hamiltonian vector field determines the trajectories of the mechanical system. Here, the family T_λ similarly describes the evolution of the system, while the Hamiltonian condition

$$\frac{d}{d\sigma} H(f_\sigma)\bigg|_{\sigma=0} = -\Omega\left(\mathscr{X}(f_0), \frac{df_\sigma}{d\sigma}\bigg|_{\sigma=0}\right)$$

closely resembles the defining relation for a Hamiltonian vector field,

$$X_H \rfloor \omega = -dH.$$

The integral curves of X_H determine the trajectories of the corresponding mechanical system. In this sense, a Hamiltonian dynamical system mimics the essential structure of Hamiltonian mechanics.

A rigorous treatment of this construction in the infinite-dimensional setting would require additional functional analytic and differential geometric machinery and lies beyond the scope of this thesis.

Chapter 4

Geometry of Hamiltonian field theory

A field theory provides a framework for determining the evolution of a field in time. We begin with an $(n+1)$-dimensional manifold M, which may be regarded intuitively as an $(n+1)$-dimensional spacetime. However, no metric is assumed on M, so this interpretation serves only as motivation. To study fields on M, we consider their restrictions to hypersurfaces $\Sigma \subset M$ and examine how these restricted fields evolve in time. The goal of this chapter is to describe the evolution of fields on a hypersurface Σ by means of a Hamiltonian dynamical system, in the sense discussed in Section 3.3.

We begin in Section 4.1 by describing the geometric framework used throughout the chapter. We start with a manifold M equipped with a fibre bundle $F \to M$, called the **field bundle**. Sections $f \in \Gamma(F)$ represent the fields of the theory. Since we wish to study the field theory on hypersurfaces, we choose a **hypersurface** $\Sigma \subset M$ and pull back the bundle F to Σ along its embedding. To describe how Σ moves through M, we introduce a **reference system**. This construction tracks the motion of Σ in M and identifies the corresponding fibres of F with those of the fixed pullback bundle over Σ. The reference system may be viewed as analogous to a reference frame: just as particle dynamics is described relative to a chosen reference frame, the field dynamics will be described relative to the chosen reference system.

In Section 4.2, we explain how to obtain dynamics on the hypersurface. To define a field theory on M, we introduce a **Lagrangian**, which is a bundle map from the first jet bundle of F to the bundle of top-degree odd forms on M. The Lagrangian determines which sections satisfy the Euler–Lagrange equations and are therefore **dynamically admissible**. We then introduce a **momentum bundle** P associated with M. This momentum bundle allows us to express the Euler–Lagrange equations in terms of a section of P. We also introduce a momentum bundle Π associated with the hypersurface Σ. Using the reference system, dynamically admissible momentum data in P are transported to sections $\pi \in \Gamma(\Pi)$. We define the phase space over Σ in terms of the admissible hypersurface data obtained by this construction. Finally, we show that, under suitable assumptions on the evolution, this phase space carries the structure of a dynamical system.

In Section 4.3, we define a Hamiltonian $\mathscr{H}_\Sigma$ associated with the hypersurface Σ. We formulate geometrically the pullback of the Lagrangian to $I \times \Sigma$. Chruściel, Jezierski, and Kijowski [27] formulate the construction pointwise, we reformulate it geometrically using jet bundles and bundle maps. We then apply a Legendre transformation in the velocity variables to obtain a Hamiltonian on the phase space.

In Section 4.4, we show that the dynamical system on the phase space is Hamiltonian with respect to $\mathscr{H}_\Sigma$, in the sense of Section 3.3. We first construct an antisymmetric two-form on the space of sections of the momentum bundle. We then compute the variation $\delta\mathscr{H}_\Sigma$ and establish the relation between this variation, the antisymmetric two-form, and the vector field generating the dynamics. Finally, we consider an explicit reference system and derive the corresponding expressions for the Hamiltonian, the antisymmetric two-form, and the variation of the Hamiltonian. This reference system will be used in Chapter 5.

Some terminology, including *field bundle*, *phase bundle*, *Lagrangian*, and *Hamiltonian*, is borrowed from physics. At this stage, these terms are used only to name the relevant geometric objects; their physical interpretation will be discussed later.

This chapter relies heavily on fibre bundles, jet bundles, and dynamical systems. For a review of these concepts and the notation used throughout the chapter, we refer to Section 2.2, Section 2.3, and Section 3.3, respectively.

This chapter is based primarily on Chapter 3 of Chruściel, Jezierski, and Kijowski [27, Chapter 3]. The order of presentation has been modified for didactic purposes.

4.1 Geometric setup

In this section, we introduce the basic geometric objects used to describe a field theory on a hypersurface. The main goal is to define a reference system that allows us to describe the motion of a fixed hypersurface through the manifold M.

In Subsection 4.1.1, we begin with a fibre bundle $F \to M$ over an $(n+1)$-dimensional manifold M, which we call the **field bundle** over M. We then choose a hypersurface $\Sigma \subset M$ and construct, by pullback, a field bundle $\Phi \to \Sigma$.

In Subsection 4.1.2, we use the bundles introduced in Subsection 4.1.1 to define a reference system. We then show how the reference system transforms a section $f \in \Gamma(F)$ into a one-parameter family of sections $\phi_t \in \Gamma(\Phi)$.

4.1.1 Fields restricted to a hypersurface

Consider a fibre bundle F over an $(n+1)$-dimensional manifold M

$$\begin{tikzcd} F \arrow[d, "\mathrm{pr}_{F\to M}"] \\ M \end{tikzcd}$$

The map $\mathrm{pr}_{F\to M}$ denotes the surjective submersion associated with the fibre bundle F over M. We omit the subscript when it is clear from the context which bundle projection is meant.

Let Σ be an n-dimensional smooth manifold and let $i\colon \Sigma \to M$ be an embedding. Then $i(\Sigma)$ is an embedded smooth hypersurface of M. We define the pullback bundle of F along i by

$$\Phi := i^*F = \{(\xi, \chi) \in \Sigma \times F \mid i(\xi) = \mathrm{pr}_{F\to M}(\chi)\}. \tag{4.1}$$

This bundle is equipped with the projection map $\mathrm{pr}_{\Phi\to\Sigma}\colon \Phi \to \Sigma$ given by projection onto the first coordinate

$$\mathrm{pr}_{\Phi\to\Sigma}(\xi,\chi) = \xi.$$

We denote the fibre of Φ above $\xi \in \Sigma$ by Φ_ξ. By construction, this fibre is naturally identified with the fibre of F above $i(\xi)$

$$\Phi_\xi = (i^*F)_\xi \cong F_{i(\xi)}.$$

Lemma 4.1 *The map $h\colon \Phi \to F$, defined by projection onto the second coordinate,*

$$h(\xi,\chi) = \chi,$$

makes the following diagram commute:

$$\begin{tikzcd} \Phi \arrow[r, "h"] \arrow[d, "\mathrm{pr}_{\Phi\to\Sigma}"'] & F \arrow[d, "\mathrm{pr}_{F\to M}"] \\ \Sigma \arrow[r, "i"'] & M \end{tikzcd}$$

Hence h is a bundle map covering i. Furthermore, for every $\xi \in \Sigma$, the restriction

$$h_\xi := h|_{\Phi_\xi}\colon \Phi_\xi \to F_{i(\xi)}$$

is a fibre isomorphism. Consequently, sections of the restricted bundle $F|_{i(\Sigma)}$ are in one-to-one correspondence with sections of Φ.

Proof. Let $(\xi,\chi) \in \Phi$. Then

$$\begin{aligned} i \circ \mathrm{pr}_{\Phi\to\Sigma}(\xi,\chi) &= i(\xi), \\ \mathrm{pr}_{F\to M} \circ h(\xi,\chi) &= \mathrm{pr}_{F\to M}(\chi). \end{aligned}$$

By construction of the pullback bundle in Equation (4.1), we have

$$i(\xi) = \mathrm{pr}_{F\to M}(\chi).$$

Therefore

$$i \circ \mathrm{pr}_{\Phi\to\Sigma} = \mathrm{pr}_{F\to M} \circ h.$$

Thus the diagram commutes, and h is a bundle map from Φ to F covering i.

For fixed $\xi \in \Sigma$, the map

$$h_\xi\colon \Phi_\xi \to F_{i(\xi)}$$

is given by

$$h_\xi(\xi,\chi) = \chi.$$

Its inverse is

$$h_\xi^{-1}(\chi) = (\xi,\chi),$$

for $\chi \in F_{i(\xi)}$. Hence h_ξ is an isomorphism.

Let f be a smooth section of $F|_{i(\Sigma)}$. Then we define a section $\phi \in \Gamma(\Phi)$ by

$$\phi(\xi) := h_\xi^{-1}(f(i(\xi))) = (\xi, f(i(\xi))). \tag{4.2}$$

Conversely, if $\phi \in \Gamma(\Phi)$, then $h \circ \phi \circ i^{-1}|_{i(\Sigma)}$ defines a section of $F|_{i(\Sigma)}$. These two constructions are inverse to each other. This proves the claimed correspondence. □

4.1.2 Motions and reference systems

In the previous subsection, we chose a hypersurface $\Sigma \subset M$ and constructed the associated pullback bundle $\Phi \to \Sigma$. In this subsection, we describe how Σ may move through M and introduce a reference system that identifies the fibres of Φ with the corresponding fibres of F along this motion. This construction plays a role analogous to that of a reference frame and will allow us to represent fields on M as time-dependent sections of the fixed bundle Φ.

We begin by formalizing what it means for the hypersurface Σ to move through M.

Definition 4.2 Let $I \subset \mathbb{R}$ be an interval. A **motion of the hypersurface** Σ in M is a smooth map $\psi\colon I \times \Sigma \to M$ such that, for every $t \in I$, the map

$$\psi_t\colon \Sigma \to M, \quad \xi \mapsto \psi(t, \xi), \tag{4.3}$$

is an embedding. Equivalently, ψ_t is a diffeomorphism from Σ onto its image $\psi_t(\Sigma)$. Moreover, we require that

$$\psi_0 = i,$$

where $i\colon \Sigma \to M$ is the embedding introduced at the start of this chapter.

The map ψ describes a one-parameter family of hypersurfaces $\psi_t(\Sigma)$ in M. Notice that the full map $\psi\colon I \times \Sigma \to M$ does not need to be injective. For example, different hypersurfaces in the family may intersect. Thus, although each ψ_t parametrizes an embedded hypersurface, the map ψ need not define a diffeomorphism.

Example 4.3 Suppose that $X \in \mathfrak{X}(M)$ is a smooth vector field whose flow $\varphi\colon I \times M \to M$ is defined on all of $I \times M$. Then

$$\psi\colon I \times \Sigma \to M, \qquad \psi(t, \xi) := \varphi_t(i(\xi)),$$

defines a motion of Σ. Indeed, since each φ_t is a diffeomorphism of M, the map $\psi_t = \varphi_t \circ i$ is an embedding, and $\psi_0 = i$. This construction will be considered in more detail in Subsection 4.4.4.

Remark 4.4 For readers familiar with general relativity, the preceding definition is reminiscent of a spacetime splitting. If (M, g) is globally hyperbolic, the Geroch–Bernal–Sánchez characterization guarantees the existence of a Cauchy time function whose level sets are Cauchy surfaces; see Burtscher and García-Heveling [13, Theorem 1.3]. These level sets motivate a family of hypersurfaces analogous to $\{\psi_t(\Sigma)\}_{t\in I}$.

The definition of a motion used here is more general, it assumes neither a Lorentzian metric nor global hyperbolicity, and it does not require the hypersurfaces $\psi_t(\Sigma)$ to be disjoint or to cover M. Additional smoothness and identification data are therefore required to obtain a motion and a reference system from a spacetime splitting. These issues will be discussed in Chapter 5; see also Subsection 5.3.3.

To compare fields on different hypersurfaces, we need a way to identify the fibres of the field bundle $F \to M$ along these hypersurfaces with the fibres of the fixed bundle $\Phi \to \Sigma$. This is the role of a reference system.

Definition 4.5 A bundle map $\Psi\colon I \times \Phi \to F$ together with a motion ψ is called a **reference system** if the following diagram commutes

$$\begin{tikzcd} I \times \Phi \arrow[r, "\Psi"] \arrow[d, "\mathrm{id}_{\mathbb{R}} \times \mathrm{pr}_{\Phi\to\Sigma}"'] & F \arrow[d, "\mathrm{pr}_{F\to M}"] \\ I \times \Sigma \arrow[r, "\psi"'] & M \end{tikzcd}$$

Moreover, we assume that for every $(t,\xi) \in I \times \Sigma$, the induced map on fibres

$$\Psi_{t,\xi}\colon \Phi_\xi \to F_{\psi(t,\xi)} \tag{4.4}$$

is an isomorphism. We refer to the pair (Ψ, ψ) as a **reference system**.

We do not investigate the general existence of reference systems in this thesis. The construction in this chapter can be carried out for any fixed reference system and therefore does not require a particular choice of (Ψ, ψ).

The choice of a reference system is generally not unique, since both the motion ψ and the bundle map Ψ may be chosen in different ways. A given choice determines the field and momentum variables on the fixed hypersurface Σ, and hence the corresponding expression for the Hamiltonian. Different choices may therefore lead to different hypersurface variables and different expressions for the Hamiltonian. These should be regarded as different representations of the same field theory, relative to different choices of motion and fibrewise identification.

When applying the formalism to a specific field theory, one must choose an appropriate reference system and identify the abstract variables introduced here with the variables used in that theory. This is done for general relativity in Chapter 5. We postpone the construction of a particular reference system until Subsection 4.4.4, where the motion is induced by the flow of a vector field on M and the corresponding bundle map is obtained from a lift of this flow.

From this point onward, we fix a reference system (Ψ, ψ). Recall that fields on M are represented by sections $f \in \Gamma(F)$ of the field bundle $F \to M$. The reference system allows us to represent the restriction of such a field to each hypersurface $\psi_t(\Sigma)$ as a section of the fixed bundle $\Phi \to \Sigma$. More precisely, a section $f \in \Gamma(F)$ determines a one-parameter family of sections of Φ, as shown in the following proposition.

Proposition 4.6 *Let $f \in \Gamma(F)$. For every $t \in I$, the reference system defines a section $\phi_t \in \Gamma(\Phi)$ by*

$$\phi_t(\xi) := \Psi_{t,\xi}^{-1}(f(\psi(t,\xi))). \tag{4.5}$$

Hence, $\{\phi_t\}_{t\in I}$ is a one-parameter family of sections of Φ. We write $\phi(t,\xi) := \phi_t(\xi)$.

Proof. Since f is a section of F, we have

$$f(\psi(t,\xi)) \in F_{\psi(t,\xi)}.$$

By definition of a reference system, the map $\Psi_{t,\xi}\colon \Phi_\xi \to F_{\psi(t,\xi)}$ is an isomorphism. Therefore $\Psi_{t,\xi}^{-1}(f(\psi(t,\xi)))$ is a well-defined element of Φ_ξ. Hence ϕ_t is a section of Φ.

The construction can be summarized by the following diagram

$$\begin{tikzcd} \phi_t(\xi) = \Psi_{t,\xi}^{-1}(f(\psi(t,\xi))) & f(\psi(t,\xi)) \arrow[l, mapsto, "\Psi_{t,\xi}^{-1}"'] \\ \xi \in \Sigma \arrow[u, mapsto, "\phi_t"] \arrow[r, mapsto, "\psi_t"] & \psi(t,\xi) \in M \arrow[u, mapsto, "f"'] \end{tikzcd}$$

Since f, ψ, and Ψ are smooth, the resulting map $\phi\colon I \times \Sigma \to \Phi$ is smooth. Thus $\{\phi_t\}_{t\in I}$ is a smooth one-parameter family of sections of Φ. □

Remark 4.7 The relation between f and ϕ_t can be written without explicitly using $\Psi_{t,\xi}^{-1}$. For each $t \in I$, it is equivalent to

$$f \circ \psi_t = \Psi_t \circ \phi_t.$$

If a section $f \in \Gamma(F)$ and a section $\phi_t \in \Gamma(\Phi)$ satisfy this relation, we say that they are **related by a reference system**.

4.2 Lagrangian field theory and hypersurface dynamics

In the previous section, we introduced a reference system to describe the motion of a hypersurface through the ambient manifold M. Its essential role is that it allows geometric and field-theoretic data on the moving hypersurfaces in M to be represented as time-dependent data on a single fixed hypersurface Σ. In particular, a field on M can be transported by the reference system to a one-parameter family of fields on Σ. We therefore refer to Σ as the **model hypersurface**. Whenever a reference system and a hypersurface Σ are considered together, Σ will be understood to have this role, and we will usually omit the qualifier "model."

This fixed-hypersurface description is the basis of the constructions that follow. Because all transported field data are represented on the same hypersurface Σ, they can be organized into a single phase space, and their evolution can be described as a dynamical system on that phase space. In this section, we begin the construction of this hypersurface formulation. Starting from a Lagrangian field theory on M, we use the reference system to construct a phase space on Σ and to describe the induced field dynamics entirely in terms of time-dependent data on this fixed hypersurface.

In Subsection 4.2.1, we explain how to formulate a Lagrangian field theory on M. The Lagrangian determines which sections are dynamically admissible through the Euler–Lagrange equations. We then introduce the **momentum bundle** P associated with the field bundle $F \to M$ and show how sections of P can be used to describe dynamically admissible sections of F. Finally, we show an identity for the first variation of the Lagrangian, which will be used in Subsection 4.4.2 when passing to the Hamiltonian formulation.

In Subsection 4.2.2, we introduce the **momentum bundle** Π associated with Φ. We define Π as the disjoint union of the **momentum spaces** Π_ξ over all $\xi \in \Sigma$. This fibrewise description will later allow us to construct a canonical symplectic form on each momentum space. By integrating these forms over Σ, we obtain an antisymmetric two-form on the space of sections of Π. We postpone this construction until Subsection 4.4.1.

In Subsection 4.2.3, we define the phase space on Σ by transporting dynamically admissible sections of P to the hypersurface using the reference system. We then show that, under suitable assumptions on the evolution, this phase space carries the structure of a dynamical system.

4.2.1 Lagrangian field theory on M

To formulate a field theory on M, we define a Lagrangian on the first jet bundle J^1F. The Lagrangian determines which sections $f \in \Gamma(F)$ are dynamically admissible through the Euler–Lagrange equations. More precisely, we define the Lagrangian as follows.

Definition 4.8 Let $F \to M$ be a fibre bundle over an $(n+1)$-dimensional manifold M. A **Lagrangian** for the field theory on M is a smooth bundle map

$$\mathbf{L}\colon J^1F \to \tilde{\Lambda}^{n+1}T^*M.$$

After choosing local bundle coordinates and a local orientation, the Lagrangian can be written as

$$(x^\mu, f^A, \partial_\mu f^A) \mapsto L(x^\mu, f^A, \partial_\mu f^A)\, dx^1 \wedge \cdots \wedge dx^{n+1}, \tag{4.6}$$

where $(x^\mu, f^A, \partial_\mu f^A)$ are local coordinates on J^1F and $\partial_\mu f^A := \frac{\partial f^A}{\partial x^\mu}$.

For a review of first jet bundles and odd forms, see Section 2.3 and Subsection 2.1.3, respectively.

A section $f \in \Gamma(F)$ is called **dynamically admissible** if it satisfies the Euler–Lagrange equation

$$\partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu f^A)} \right) = \frac{\partial L}{\partial f^A}. \tag{4.7}$$

For a physical example of the Euler–Lagrange equations, see Example 3.6. We do not specify a particular Lagrangian in this chapter. For general relativity, however, an explicit Lagrangian is introduced in Subsection 5.1.2.

To construct the phase space, we need an alternative description of dynamically admissible fields. This is where the notion of a momentum bundle enters.

Definition 4.9 Let $F \to M$ be a fibre bundle. We define the **momentum bundle** P over M by

$$P := V^*F \otimes \tilde{\Lambda}^n T^*M. \tag{4.8}$$

The bundle $\tilde{\Lambda}^n T^*M$ of odd n-forms on M has rank

$$\binom{\dim M}{n} = \binom{n+1}{n} = n+1.$$

In local coordinates, a basis is given by

$$\left\{ \partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1} \right\}_{\mu=1}^{n+1},$$

where $\partial_\mu := \frac{\partial}{\partial x^\mu}$. Therefore, a local section p of P can be written as

$$p = p_A{}^\mu df^A \otimes \left(\partial_\mu \rfloor (dx^1 \wedge \cdots \wedge dx^{n+1}) \right).$$

Thus, a section p is locally described by the functions $p_A{}^\mu$ together with the component functions f^A.

Remark 4.10 Throughout this thesis, we use the notation of (4.8), which should be interpreted as defining a bundle over either F or M, according to the context. More precisely, V^*F is a vector bundle over F, while $\tilde{\Lambda}^n T^*M$ is a vector bundle over M. Since both interpretations will be used interchangeably, we briefly clarify this notational convention here.

When P is viewed as a vector bundle over F, the precise expression is

$$P = V^*F \otimes \mathrm{pr}^*_{F\to M}\left(\tilde{\Lambda}^n T^*M\right).$$

We express a local section $p\colon F \to P$ as

$$p(x,f) = p_A{}^\mu(x,f)\, df^A \otimes \left(\partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}\right).$$

If instead P is viewed as a bundle over M, then we work with the pullback bundle $f^*P := f^*(V^*F) \otimes \tilde{\Lambda}^n T^*M$ over M, along a section $f \in \Gamma(F)$. A local section p of f^*P can be written as[1]

$$\begin{aligned} p(x) &= \sum_{\mu=1}^{n+1}\left(\sum_A p_A{}^\mu(x)\, df^A|_{f(x)}\right) \otimes \left(\partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}\right)\Big|_x \\ &= p_A{}^\mu(x)\, df^A|_{f(x)} \otimes \left(\partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}\right)\Big|_x. \end{aligned} \tag{4.9}$$

Here we use the Einstein summation convention.

The momentum bundle P introduced above provides the geometric setting for momentum variables over M. We can rewrite the Euler–Lagrange equation in terms of this momentum bundle.

Definition 4.11 Define the **canonical momenta** associated with f by

$$p_A{}^\mu := \frac{\partial L}{\partial(\partial_\mu f^A)}. \tag{4.10}$$

In terms of the canonical momenta, the Euler–Lagrange equation becomes

$$\partial_\mu p_A{}^\mu = \frac{\partial L}{\partial f^A}. \tag{4.11}$$

For any section $f \in \Gamma(F)$, equation (4.10) determines an associated section $p \in \Gamma(P)$. It follows from Definition 4.11 that f is dynamically admissible if and only if the associated components $(f^A, p_A{}^\mu)$ satisfy equations (4.10)–(4.11). A dynamically admissible field may

[1] Strictly speaking, $df^A\big|_{f(x)}$ is a vertical covector at the point $f(x) \in F$, while a section of f^*P is defined over $x \in M$. The corresponding covector over x is obtained by pulling back along f:

$$f^*\left(df^A\big|_{f(x)}\right) = d(f^A \circ f)\big|_x.$$

From now on, we identify these two expressions by abuse of notation and write both simply as $df^A\big|_{f(x)}$, or just as df^A when the base point is clear. Furthermore, we use the same symbol f^A for the fibre coordinate on Φ and for the corresponding component function $f^A \circ f$ of the section f. In this notation,

$$f^A = f^A \circ f.$$

therefore be described in terms of the field variables f^A together with the momentum variables ${p_A}^\mu$.

Physically, ${p_A}^\mu$ is the canonical momentum associated with the derivative $\partial_\mu f^A$. Thus, the momentum bundle P provides the geometric setting in which the field variables and their associated momentum variables are combined. After choosing a hypersurface and a reference system, these variables give rise to the corresponding dynamical variables on Σ. This will be discussed in Subsection 4.2.3.

One can rewrite (4.10)–(4.11) in a compact form using variations. Let $p(\lambda)$ be a one-parameter family of sections of P, with underlying family of sections $f(\lambda) \in \Gamma(F)$. We write

$$\delta := \left.\frac{d}{d\lambda}\right|_{\lambda=0}, \qquad \delta f := \left.\frac{d}{d\lambda}\right|_{\lambda=0} f(\lambda).$$

Thus, δf is a vertical vector field along $f := f(0)$.

Proposition 4.12 *Let $p(\lambda)$ be a one-parameter family of sections of P, defined for λ in an interval containing 0, with local coordinate representation $\big({p_A}^\mu(\lambda), f^B(\lambda)\big)$. Suppose that these components satisfy the Euler–Lagrange equations (4.10)–(4.11). We write ${p_A}^\mu := {p_A}^\mu(0)$, $f^B := f^B(0)$. Then*

$$\delta L(x^\mu, f^A, \partial_\mu f^A) = \partial_\mu\big({p_A}^\mu \delta f^A\big). \tag{4.12}$$

Equivalently,

$$\delta \mathbf{L}(j^1 f) = \mathbf{d}\big(p(\delta f)\big), \tag{4.13}$$

where $\mathbf{d}$ is the exterior derivative on M and $p(\delta f) \in \Gamma\big(\tilde{\Lambda}^n T^*M\big)$ is the odd n-form obtained by pairing the vertical covector component of p with the vertical variation field δf. Pointwise, this pairing is given by

$$p(\delta f)(x) := p(x)\big(\delta f(x)\big).$$

In local coordinates,

$$p(\delta f) = {p_A}^\mu \delta f^A \left(\partial_\mu \rfloor \big(dx^1 \wedge \cdots \wedge dx^{n+1}\big)\right). \tag{4.14}$$

Proof. By the chain rule, we have

$$\delta L = \frac{\partial L}{\partial f^A}\delta f^A + \frac{\partial L}{\partial(\partial_\mu f^A)}\delta(\partial_\mu f^A).$$

Since $\delta(\partial_\mu f^A) = \partial_\mu(\delta f^A)$ and ${p_A}^\mu = \frac{\partial L}{\partial(\partial_\mu f^A)}$, this becomes

$$\delta L = \frac{\partial L}{\partial f^A}\delta f^A + {p_A}^\mu \partial_\mu(\delta f^A).$$

Using (4.11), $\partial_\mu {p_A}^\mu = \frac{\partial L}{\partial f^A}$, we obtain

$$\delta L = \partial_\mu {p_A}^\mu \delta f^A + {p_A}^\mu \partial_\mu(\delta f^A) = \partial_\mu({p_A}^\mu \delta f^A).$$

In terms of differential forms, this is precisely

$$\begin{aligned}
\delta \mathbf{L}(j^1 f) &= \partial_\mu({p_A}^\mu \delta f^A) dx^1 \wedge \cdots \wedge dx^{n+1} \\
&= \mathbf{d}({p_A}^\mu \delta f^A (\partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1})) \\
&= \mathbf{d}(p(\delta f)).
\end{aligned}$$

□

4.2.2 Momentum bundle on Σ

In the previous subsection, we showed that a dynamically admissible section $f \in \Gamma(F)$ can be represented by an associated section $p \in \Gamma(P)$ of the momentum bundle. We now construct the corresponding momentum bundle over Σ.

We formulate the momentum bundle on Σ as a disjoint union of **momentum spaces**. This fibrewise description will be used in Subsection 4.4.1 to construct an antisymmetric two-form on the space of sections of the momentum bundle. We then show that this disjoint union has the structure of the momentum bundle naturally associated with $\Phi \to \Sigma$.

Definition 4.13 For any point $\xi \in \Sigma$, we define the **momentum space** at ξ by

$$\Pi_\xi := T^*\Phi_\xi \otimes \tilde{\Lambda}^n T^*_\xi \Sigma, \tag{4.15}$$

where Φ_ξ is the fibre of Φ over ξ, and $\tilde{\Lambda}^n T^*_\xi \Sigma$ is the space of top-degree odd forms at ξ.

The collection of all momentum spaces Π_ξ over $\xi \in \Sigma$ will be denoted by Π and will be called the **momentum bundle**. We use the vertical bundle of $\Phi \to \Sigma$ to express this collection in a coordinate-free way.

Proposition 4.14 *Let $\Phi \to \Sigma$ be the field bundle of Σ. The momentum bundle Π associated to the field bundle is given by*

$$\Pi = \bigsqcup_{\xi \in \Sigma} \Pi_\xi \cong V^*\Phi \otimes \tilde{\Lambda}^n T^*\Sigma.$$

Proof. Recall that the tangent space of a fibre can be identified with the vertical space above that point. More precisely, by Lemma 2.47, for $x \in \Phi_\xi$ we have

$$V_x\Phi \cong T_x\Phi_\xi.$$

Dualizing both sides and taking the disjoint union yields

$$\bigsqcup_{x \in \Phi_\xi} V^*_x\Phi \cong \bigsqcup_{x \in \Phi_\xi} T^*_x\Phi_\xi = T^*\Phi_\xi.$$

Substituting this expression into the definition of Π_ξ, we obtain

$$\Pi_\xi = T^*\Phi_\xi \otimes \tilde{\Lambda}^n T^*_\xi \Sigma \cong (\bigsqcup_{x \in \Phi_\xi} V^*_x\Phi) \otimes \tilde{\Lambda}^n T^*_\xi \Sigma.$$

Taking the disjoint union over all $\xi \in \Sigma$, we find

$$\begin{aligned} \Pi &\cong \bigsqcup_{\xi \in \Sigma} \Big((\bigsqcup_{x \in \Phi_\xi} V^*_x\Phi) \otimes \tilde{\Lambda}^n T^*_\xi \Sigma \Big) \\ &= V^*\Phi \otimes \tilde{\Lambda}^n T^*\Sigma. \end{aligned} \tag{4.16}$$

This proves the claimed identification. □

Sections $\pi \in \Gamma(\Pi)$ can locally be expressed as

$$\pi = \pi_A d\phi^A \otimes d\xi^1 \wedge \cdots \wedge d\xi^n. \tag{4.17}$$

Hence, we can describe π by the collection of functions (π_A, ϕ^A).

Remark 4.15 As discussed in Remark 4.10, the expression

$$\Pi = V^*\Phi \otimes \tilde{\Lambda}^n T^*\Sigma$$

involves bundles that naturally live over different base spaces. Throughout this thesis, we interpret Π as a bundle over either Φ or Σ, according to the context. Since both interpretations will be used interchangeably, we briefly clarify this notational convention here.

If we view Π as a bundle over Φ, the correct expression is obtained by pulling back the bundle of odd forms along the projection $\mathrm{pr}_{\Phi\to\Sigma}\colon \Phi \to \Sigma$. Hence, we obtain

$$\Pi = V^*\Phi \otimes \mathrm{pr}^*_{\Phi\to\Sigma}\left(\tilde{\Lambda}^n T^*\Sigma\right).$$

A section $\pi\colon \Phi \to \Pi$ can then locally be expressed as

$$\begin{aligned}\pi(\xi,\phi) &= \pi_A(\xi,\phi) d\phi^A \otimes d(\xi^1 \circ \mathrm{pr}_{\Phi\to\Sigma}) \wedge \cdots \wedge d(\xi^n \circ \mathrm{pr}_{\Phi\to\Sigma}) \\ &= \pi_A(\xi,\phi) d\phi^A \otimes d\xi^1 \wedge \cdots \wedge d\xi^n,\end{aligned}$$

where we use the standard abuse of notation in which the coordinates ξ^i on Σ are also used for their pullbacks to Φ.

If we view Π as a bundle over Φ, along a section $\phi \in \Gamma(\Phi)$, we express the momentum bundle $\phi^*\Pi := \phi^*(V^*\Phi) \otimes \tilde{\Lambda}^n T^*\Sigma$ locally as

$$\pi(\xi) = \pi_A(\xi)\ d\phi^A\big|_{\phi(\xi)} \otimes d\xi^1 \wedge \cdots \wedge d\xi^n\big|_\xi .$$

Here (ξ^i, ϕ^A)[2] are local coordinates on Φ.

In this section, we constructed the geometric data associated with a fixed hypersurface $\Sigma \subset M$. Starting from the field bundle $F \to M$, we obtained the pulled-back bundle $\Phi \to \Sigma$, whose sections describe field configurations on the hypersurface. We then introduced the momentum bundle Π, which encodes the corresponding momentum variables.

The momentum bundle carries a natural symplectic structure, analogous to the canonical symplectic form on a cotangent bundle. We postpone its construction until it is needed for the Hamiltonian dynamics at the end of the chapter. In the next section, we allow the hypersurface to move through M and use a reference system to describe the dynamics on the fixed hypersurface Σ.

[2]Strictly speaking, $d\phi^A\big|_{\phi(\xi)}$ is a vertical covector at the point $\phi(\xi) \in \Phi$, while a section of $\phi^*\Pi$ is defined over $\xi \in \Sigma$. The corresponding covector over ξ is obtained by pulling back along ϕ:

$$\phi^*\left(d\phi^A\big|_{\phi(\xi)}\right) = d(\phi^A \circ \phi)\big|_\xi .$$

From now on, we identify these two expressions by abuse of notation and write both simply as $d\phi^A\big|_{\phi(\xi)}$, or just as $d\phi^A$ when the base point is clear. Furthermore, we use the same symbol ϕ^A for the fibre coordinate on Φ and for the corresponding component function $\phi^A \circ \phi$ of the section ϕ. In this notation,

$$\phi^A = \phi^A \circ \phi.$$

4.2.3 From fields on M to dynamics on Σ

In the previous subsections, we introduced the momentum bundles P and Π. In this subsection, we use the reference system (Ψ, ψ) introduced in Subsection 4.1.2 to define a phase space on Σ. We do this by transporting a dynamically admissible section $p \in \Gamma(P)$ to a one-parameter family of sections $\{\pi_t\}_{t \in I}$ of Π. After imposing the appropriate evolution condition on this family, we show that the resulting phase space carries the structure of a dynamical system.

We begin by establishing the construction of the family $\{\pi_t\}_{t\in I}$.

Proposition 4.16 *Let (Ψ, ψ) be a reference system, let $f \in \Gamma(F)$, and let $p \in \Gamma(P)$. Then (Ψ, ψ) determines a one-parameter family of sections $\pi_t \in \Gamma(\Pi)$, defined by*

$$\pi_t(\xi) := \left((\Psi_{t,\xi})^* \otimes \psi^*_{t,\xi}\right) p(\psi(t, \xi)). \tag{4.18}$$

Here $\Psi_{t,\xi} \colon \Phi_\xi \xrightarrow{\sim} F_{\psi(t,\xi)}$ denotes the fibrewise diffeomorphism from Φ_ξ to $F_{\psi(t,\xi)}$, and $\psi^*_{t,\xi}$ denotes the pullback of n-forms induced by $\psi_t := \psi(t, \cdot)$ at ξ.

In local coordinates, π_t is written as

$$\pi_t = \pi_A(t) d\phi^A(t) \otimes d\xi^1 \wedge \cdots \wedge d\xi^n,$$

where $\phi_t = (\phi^B(t))$ is related to f by the reference system. As before, when no ambiguity can occur, we write $\pi(t, \xi) := \pi_t(\xi)$.

Proof. The definition of $\pi_t(\xi)$ is fibrewise. We first check that the right-hand side of equation (4.18) lies in the correct fibre. The map $\Psi_{t,\xi}$ induces a diffeomorphism between tangent bundles

$$d\Psi_{t,\xi} \colon T\Phi_\xi \xrightarrow{\sim} TF_{\psi(t,\xi)}.$$

We use the standard identification of the vertical space with the tangent space of the fibre, and after dualizing we obtain the diffeomorphism

$$d\Psi^\vee_{t,\xi} \colon \; V^*F|_{F_{\psi(t,\xi)}} \xrightarrow{\sim} T^*\Phi_\xi,$$

where we denote $VF|_{F_{\psi(t,\xi)}} := \bigsqcup_{a \in F_{\psi(t,\xi)}} V_aF$ and use $\vee$ to denote the dual, to avoid overloading the notation. Hence elements in the pullback bundle $(\Psi_{t,\xi})^* \left(V^*F|_{F_{\psi(t,\xi)}}\right)$ can be identified with elements in $T^*\Phi_\xi$.

For the second tensor factor of P, note that the map $\psi_t \colon \Sigma \to M$ is an embedding. Therefore, for each $\xi \in \Sigma$, its differential is injective

$$d_\xi\psi_t \colon T_\xi\Sigma \hookrightarrow \sim T_{\psi(t,\xi)}M.$$

It induces the pullback map

$$\psi^*_{t,\xi} = \psi^*_{t,\xi} \colon \tilde{\Lambda}^n T^*_{\psi(t,\xi)}M \xrightarrow{\sim} \tilde{\Lambda}^n T^*_\xi \Sigma.$$

Hence, the second factor of $p(\psi(t, \xi))$ is pulled back to an odd n-form at ξ.

Combining the two pullbacks, we obtain

$$\pi_t(\xi) \in T^*\Phi_\xi \otimes \tilde{\Lambda}^n T^*_\xi \Sigma = \Pi_\xi.$$

The maps involved in this construction depend smoothly on ξ. Since p, Ψ, and ψ are smooth, the map

$$\xi \mapsto \pi_t(\xi)$$

is smooth for every fixed $t \in I$. Hence, π_t is a section of Π.

It remains to describe the local form. Let (x^μ, f^A) be local coordinates on F, and write

$$p(x) = p_A{}^\mu(x)\, df^A|_{f(x)} \otimes \left(\partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}\right)\big|_x.$$

The reference system transports the fibre coordinates f^A on F to fibre coordinates on Φ by

$$\phi_t^A := f^A \circ \Psi_t.$$

Therefore

$$(\Psi_{t,\xi})^* \left(df^A|_{f(\psi(t,\xi))}\right) = d\phi_t^A|_{\phi_t(\xi)}.$$

After choosing compatible local orientations, the pullback of each odd n-form can be written as

$$\psi_{t,\xi}^* \left(\left(\partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}\right)\big|_{\psi(t,\xi)}\right) = J_\mu(t,\xi)\, d\xi^1 \wedge \cdots \wedge d\xi^n,$$

for suitable smooth functions $J_\mu(t,\xi)$[3]. Consequently,

$$\begin{aligned}\pi_t(\xi) &= p_A{}^\mu(\psi(t,\xi)) J_\mu(t,\xi)\, d\phi_t^A|_{\phi_t(\xi)} \otimes d\xi^1 \wedge \cdots \wedge d\xi^n \\ &= \pi_A(t,\xi)\, d\phi_t^A|_{\phi_t(\xi)} \otimes d\xi^1 \wedge \cdots \wedge d\xi^n,\end{aligned}$$

where

$$\pi_A(t,\xi) := p_A{}^\mu(\psi(t,\xi)) J_\mu(t,\xi).$$

Using the convention of Remark 4.10, we write this locally as

$$\pi_t = \pi_A(t)\, d\phi^A(t) \otimes d\xi^1 \wedge \cdots \wedge d\xi^n. \qquad \square$$

The construction in Proposition 4.16 associates with every section p of P a one-parameter family of sections $\{\pi_t\}_{t\in I}$ of Π. To define the phase space, we restrict this construction to sections p satisfying the field equations (4.10)–(4.11) and consider the resulting data at $t = 0$.

Definition 4.17 We define the **phase space on** Σ, denoted by $\mathscr{P}_\Sigma$, as the set of all sections of Π obtained at $t = 0$ by the construction of Proposition 4.16 from sections of P satisfying the field equations (4.10)–(4.11).

[3] An analogous coordinate transformation is carried out in the proof of Lemma 4.22, where the corresponding coefficient functions are written explicitly in terms of the Jacobian determinant and the derivatives of the coordinate transformation.

Remark 4.18 This definition implicitly requires suitable regularity conditions on the sections of P. These conditions are needed to ensure that the Euler–Lagrange equations (4.10)–(4.11) are well defined and admit solutions with the required regularity. We do not address these analytical issues here, but return to them in the setting of general relativity.

There, the canonical variables on a hypersurface are the induced Riemannian metric γ_{ij} and its canonically conjugate momentum π^{ij}. Rather than proving a general existence result, we work with Cauchy data induced by a smooth Lorentzian metric satisfying the Einstein equations. Under this assumption, γ_{ij} and π^{ij} have the required regularity. See Section 5.2 for further details.

We interpret the phase space $\mathscr{P}_\Sigma$ as the space of admissible initial data on the reference hypersurface Σ. Given a section of P, Proposition 4.16 produces a one-parameter family of sections of Π,

$$t \longmapsto \pi_t \in \Gamma(\Pi).$$

The reference system identifies each hypersurface $\psi_t(\Sigma)$ with the fixed hypersurface Σ. Consequently, every π_t is represented as hypersurface data on the same fixed manifold Σ. By shifting the origin of the time parameter from 0 to t, the section π_t may therefore be regarded as new initial data, provided that $\pi_t \in \mathscr{P}_\Sigma$.

Definition 4.19 Suppose that for every $\pi(0) \in \mathscr{P}_\Sigma$ there is a unique transported family $\pi(t) \in \mathscr{P}_\Sigma$, defined for all $t \in I$, arising from the construction of Proposition 4.16. Then we define the **time-evolution map**

$$T_t \colon \mathscr{P}_\Sigma \to \mathscr{P}_\Sigma$$

by

$$T_t(\pi_A(s), \phi^B(s)) := (\pi_A(s+t), \phi^B(s+t)),$$

whenever $s, t, s+t \in I$.

Remark 4.20 The assumption in Definition 4.19 is essential for the time-evolution map to be well defined. For a discussion of the difficulties that may arise, see [27, p. 32]. In our treatment of general relativity, we do not assume in advance that the construction yields a well-defined time-evolution map. Instead, we address this issue when applying the construction to general relativity in Subsection 5.3.3.

Theorem 4.21 *Suppose that the time-evolution maps $\{T_t\}_{t\in I}$ satisfy the conditions of Definition 4.19. Then $\{T_t\}_{t\in I}$ is a dynamical system on $\mathscr{P}_\Sigma$ as defined in Subsection 3.3.1. Furthermore, the vector field generating the dynamics is locally given by*

$$\mathscr{X}(\pi) = (\dot{\pi}_A, \dot{\phi}^B).$$

Proof. By definition, $T_0 = \mathrm{id}_{\mathscr{P}_\Sigma}$. Let $(\pi_A(s), \phi^B(s)) \in \mathscr{P}_\Sigma$. Then, whenever $t_1, t_2, t_1 + t_2 \in I$, we have

$$\begin{aligned} T_{t_1+t_2}(\pi_A(s), \phi^B(s)) &= (\pi_A(s+t_1+t_2), \phi^B(s+t_1+t_2)), \\ T_{t_1} \circ T_{t_2}(\pi_A(s), \phi^B(s)) &= T_{t_1}(\pi_A(s+t_2), \phi^B(s+t_2)) \\ &= (\pi_A(s+t_1+t_2), \phi^B(s+t_1+t_2)). \end{aligned}$$

Hence $T_{t_1+t_2} = T_{t_1} \circ T_{t_2}$.

Moreover, we check that for $\pi \in \mathscr{P}_\Sigma$ the map $t \mapsto T_t(\pi)$ is a differentiable curve in $\mathscr{P}_\Sigma$. By construction, the curve $t \mapsto \pi(t)$ is obtained from the smooth section p and the smooth reference system (Ψ, ψ). Hence $t \mapsto \pi(t)$ depends smoothly on t, and therefore $t \mapsto T_t(\pi)$ is differentiable. Thus $\{T_t\}_{t\in I}$ is a dynamical system on $\mathscr{P}_\Sigma$.

Lastly, the generating vector field is defined by differentiating the time-evolution map at $t = 0$:

$$\mathscr{X}(\pi) := \frac{d}{dt}\bigg|_{t=0} T_t(\pi).$$

In local coordinates, if $\pi = (\pi_A, \phi^B)$, then

$$\begin{aligned}\mathscr{X}(\pi) &= \left(\frac{d}{dt}\bigg|_{t=0} \pi_A(t),\, \frac{d}{dt}\bigg|_{t=0} \phi^B(t) \right) \\ &= (\dot{\pi}_A, \dot{\phi}^B).\end{aligned}$$

□

4.3 From Lagrangian to Hamiltonian

In Subsection 4.2.3, we showed how a reference system allows us to transport dynamically admissible field data on M to time-dependent data on the fixed hypersurface Σ. This led to the definition of the phase space $\mathscr{P}_\Sigma$ as the space of admissible initial data on Σ. Under the assumption that the transported data remain in $\mathscr{P}_\Sigma$ and are uniquely determined by their initial values, we obtained a family of time-evolution maps $\{T_t\}_{t\in I}$ and hence a dynamical system on $\mathscr{P}_\Sigma$.

The goal of this chapter is to show that the phase space $\mathscr{P}_\Sigma$ carries the structure of a Hamiltonian dynamical system. As an intermediate step toward this goal, this section constructs the Hamiltonian associated with the dynamics on the hypersurface Σ.

In Subsection 4.3.1, we introduce an auxiliary geometric object that records the momentum data obtained from a section p of P after transporting it to $I \times \Sigma$ using the reference system. This construction will be used later to show that the dynamical system defined above is Hamiltonian.

In Subsection 4.3.2, we explain how the Lagrangian $\mathbf{L}$ on M gives rise to a pulled-back Lagrangian on $I \times \Sigma$, and hence to a time-dependent Lagrangian on the fixed hypersurface Σ. This construction uses the first prolongation of the reference system to the corresponding jet bundles, as discussed in Subsection 2.3.2. Chruściel, Jezierski, and Kijowski [27] formulate the construction pointwise, we reformulate it geometrically using jet bundles and bundle maps.

Finally, in Subsection 4.3.3, we construct the Hamiltonian on $\mathscr{P}_\Sigma$ by applying the Legendre transformation to the hypersurface Lagrangian with respect to the velocity variables. More precisely, for a one-parameter family of field configurations $\{\phi_t\}_{t\in I}$, we define the velocity of the component ϕ^A by the variation

$$\dot{\phi}^A := \frac{\partial \phi^A}{\partial t}.$$

We then perform the Legendre transformation with respect to the variables $\dot{\phi}^A$. For a review of the Legendre transformation, see Subsection 3.1.2.

4.3.1 Total pullback of the momentum section

In Proposition 4.16, we transported a section $p \in \Gamma(P)$ to the model hypersurface Σ at each fixed time $t \in I$. The resulting fixed-time pullback retains only the data obtained after restricting to the time slice $\{t\} \times \Sigma$. To record simultaneously the time dependence and all components of p transported by the reference system, we instead pullback before restricting to a fixed time. The resulting object, denoted by $\boldsymbol{\pi}$, encodes all the information carried by p. We call $\boldsymbol{\pi}$ the **total pullback** of p.

Lemma 4.22 *Let Ψ and ψ be a reference system as in Definition 4.5. For a section $p\colon F \to P$, we define*

$$\boldsymbol{\pi} := (\Psi^* \otimes \psi^*)p. \tag{4.19}$$

In local coordinates, after choosing a local orientation, $\boldsymbol{\pi}$ can be written as

$$\boldsymbol{\pi} = \pi_A d\phi^A \otimes d\xi^1 \wedge \cdots \wedge d\xi^n + \pi_A{}^a d\phi^A \otimes \partial_a \rfloor dt \wedge d\xi^1 \wedge \cdots \wedge d\xi^n, \tag{4.20}$$

Proof. Recall that a section $p\colon F \to P$ can be written locally as

$$p = p_A{}^\mu df^A \otimes \partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1},$$

where (x^μ, f^A) are local coordinates on F. Applying the pullback $(\Psi^* \otimes \psi^*)$ gives

$$\begin{aligned}\boldsymbol{\pi} &= (\Psi^* \otimes \psi^*)\left(p_A{}^\mu df^A \otimes \partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}\right)\\ &= \sum_{\mu=1}^{n+1} (p_A{}^\mu \circ \Psi) d(f^A \circ \Psi) \otimes d(x^1 \circ \psi) \wedge \cdots \wedge d\widehat{(x^\mu \circ \psi)} \wedge \cdots \wedge d(x^{n+1} \circ \psi), \end{aligned} \tag{4.21}$$

where the hat indicates the omitted term.

Since p takes values in the vertical cotangent bundle, we only keep track of the vertical part of the first factor. Writing $\Psi^A := f^A \circ \Psi$, we obtain

$$d(f^A \circ \Psi) = \frac{\partial \Psi^A}{\partial \phi^B} d\phi^B. \tag{4.22}$$

For the second tensor factor, write $y := (t, \xi^i)$ and $\psi^\mu := x^\mu \circ \psi$. The coordinate transformation gives

$$d\psi^1 \wedge \cdots \wedge \widehat{d\psi^\mu} \wedge \cdots \wedge d\psi^{n+1} = \det(D\psi) \sum_\nu \frac{\partial y^\nu}{\partial \psi^\mu} dy^1 \wedge \cdots \wedge \widehat{dy^\nu} \wedge \cdots \wedge dy^{n+1}.$$

Hence the coefficient of $d\xi^1 \wedge \cdots \wedge d\xi^n$ in $\psi^*(\partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1})$ is

$$\det(D\psi)\frac{\partial t}{\partial \psi^\mu}, \tag{4.23}$$

while the coefficient of $\partial_a \rfloor dt \wedge d\xi^1 \wedge \cdots \wedge d\xi^n$ is

$$\det(D\psi)\frac{\partial \xi^a}{\partial \psi^\mu}. \tag{4.24}$$

Substituting (4.22), (4.23) and (4.24) into (4.21) gives

$$\boldsymbol{\pi} = \det(D\psi) \sum_{\mu,A} (p_A{}^\mu \circ \Psi) \frac{\partial \Psi^A}{\partial \phi^B} \frac{\partial t}{\partial \psi^\mu} d\phi^B \otimes d\xi^1 \wedge \cdots \wedge d\xi^n$$

$$+ \det(D\psi) \sum_{\mu,A} (p_A{}^\mu \circ \Psi) \frac{\partial \Psi^A}{\partial \phi^B} \frac{\partial \xi^a}{\partial \psi^\mu} d\phi^B \otimes \partial_a \rfloor dt \wedge d\xi^1 \wedge \cdots \wedge d\xi^n. \tag{4.25}$$

We define π_B and $\pi_B{}^a$ to be the coefficients that absorb the coordinate change. Therefore, we obtain

$$\boldsymbol{\pi} = \pi_B d\phi^B \otimes d\xi^1 \wedge \cdots \wedge d\xi^n + \pi_B{}^a d\phi^B \otimes \partial_a \rfloor dt \wedge d\xi^1 \wedge \cdots \wedge d\xi^n. \qquad \square$$

4.3.2 Lagrangian on the hypersurface Σ

In this subsection, we construct a Lagrangian on the model hypersurface Σ from the Lagrangian on the total manifold M. Our construction follows the approach of Chruściel, Jezierski, and Kijowski [27, pp. 32–33], but we formulate it geometrically in terms of bundle maps rather than pointwise expressions.

Since the Lagrangian is defined on the first jet bundle J^1F, we first discuss the prolongation of the reference system. This prolongation defines the bundle map between the relevant first jet bundles that is needed to transport the Lagrangian from M to $I \times \Sigma$, and hence to the hypersurface Σ at each fixed time.

Let (Ψ, ψ) be a reference system as in Definition 4.5, and fix $t \in I$. Restricting the reference system to this time gives a bundle morphism

$$(\Psi_t, \psi_t) \colon \Phi \longrightarrow F,$$

where $\Psi_t := \Psi(t, \cdot)$ and $\psi_t := \psi(t, \cdot)$. By definition of a motion, ψ_t is a diffeomorphism from Σ onto its image $\psi_t(\Sigma)$.

We now introduce the first prolongation of this bundle morphism, as defined in Definition 2.59. It defines a bundle morphism

$$j^1\Psi_t \colon J^1\Phi \longrightarrow J^1\big(F|_{\psi_t(\Sigma)}\big)\,.$$

For $\phi \in \Gamma(\Phi)$ and $\xi \in \Sigma$, it is given by

$$j^1\Psi_t\big(j^1_\xi \phi\big) = j^1_{\psi_t(\xi)}\big(\widetilde{\psi}_t(\phi)\big),$$

where

$$\widetilde{\psi}_t(\phi) := \Psi_t \circ \phi \circ \psi_t^{-1} \in \Gamma\big(F|_{\psi_t(\Sigma)}\big)\,. \tag{4.26}$$

For a section ϕ defined only on an open subset $U \subset \Sigma$, the corresponding transported section $\widetilde{\psi}_t$ is defined on $\psi_t(U)$. We suppress these domains when they are clear from the context.

The prolongation $j^1\Psi_t$ is a bundle morphism. Composing $j^1\Psi_t$ with the Lagrangian $\mathbf{L}$ yields a bundle map from $J^1\Phi$ to $\widetilde{\Lambda}^{n+1}T^*M$, as illustrated in the diagram below.

$$\begin{array}{ccccc}
J^1\Phi & \xrightarrow{j^1\Psi_t} & J^1F|_{\psi_t(\Sigma)} & \xrightarrow{\mathbf{L}} & \tilde{\Lambda}^{n+1}T^*M \\
\downarrow & & \downarrow & & \\
\{t\}\times\Phi & \xrightarrow{\Psi_t} & F|_{\psi_t(\Sigma)} & & \\
\downarrow & & \downarrow & & \\
\{t\}\times\Sigma & \xrightarrow{\psi_t} & \psi_t(\Sigma) & &
\end{array}$$

We assume that the sections $\phi_t \in \Gamma(\Phi)$ and $f \in \Gamma(F)$ are related by the reference system, as discussed in Remark 4.7. Recall that this means that

$$f \circ \psi_t = \Psi_t \circ \phi_t.$$

This relation simplifies the composition with $j^1\Psi_t$. Let $j^1_\xi(\phi_t) \in J^1\Phi$ be a point in our jet bundle. We compute

$$\begin{aligned}
\mathbf{L} \circ j^1\Psi_t(j^1_\xi\phi_t) &= \mathbf{L}(j^1_{\psi_t(\xi)}(\widetilde{\psi}_t(\phi_t))) \\
&= \mathbf{L}(j^1_{\psi(t,\xi)}(\Psi_t \circ \phi_t \circ \psi_t^{-1})) \\
&= \mathbf{L}(j^1_{\psi(t,\xi)}(f \circ \psi_t \circ \psi_t^{-1})) \\
&= \mathbf{L}(j^1_{\psi(t,\xi)}f),
\end{aligned} \tag{4.27}$$

where $\widetilde{\psi}_t(\phi_t)$ is defined as in (4.26). Since Ψ, ψ, ϕ, and $\mathbf{L}$ are smooth, the composition $\mathbf{L} \circ j^1\Psi_t$ depends smoothly on t. This construction transports the Lagrangian on the total manifold M to the hypersurface Σ.

Definition 4.23 Let (Ψ, ψ) be a reference system, and let $\mathbf{L}$ be the Lagrangian on M. We define a one-parameter family of **pulled-back Lagrangians** on Σ by

$$\mathscr{L}_t := \psi_t^*(\dot{\psi}_t \rfloor \mathbf{L} \circ j^1\Psi_t), \tag{4.28}$$

where

$$\dot{\psi}_t(\xi) := d\psi_{(t,\xi)}\left(\frac{\partial}{\partial t}\right) = \frac{\partial \psi}{\partial t}(t,\xi) \in T_{\psi_t(\xi)}M$$

Since the construction is smooth in t, this gives a smooth one-parameter family of Lagrangians, denoted by $\{\mathscr{L}_t\}_{t\in I}$. More explicitly, let (x^μ, f^A) be local coordinates on F and $f \in \Gamma(F)$ and $\phi_t \in \Gamma(\Phi)$ be sections related by the reference system. The pullback (4.28) is pointwise given by

$$\begin{aligned}
\mathscr{L}_t(j^1_\xi(\phi_t)) &= \psi_t^*(\dot{\psi}_t \rfloor \mathbf{L}(j^1_{\psi(t,\xi)}f)) \\
&= L\left(x^\mu(j^1_{\psi(t,\xi)}f), f^A(j^1_{\psi(t,\xi)}f), \frac{\partial f^A}{\partial(x^\mu \circ \psi)}(j^1_{\psi(t,\xi)}f)\right)\psi_t^*(\dot{\psi}_t \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}) \\
&= L\left(\psi^\mu(t,\xi), (f^A \circ f)(\psi(t,\xi)), \left.\frac{\partial(f^A \circ f)}{\partial \psi^\mu}\right|_{\psi(t,\xi)}\right)\det(D\psi)d\xi^1 \wedge \cdots \wedge d\xi^n.
\end{aligned} \tag{4.29}$$

Here $\psi^\mu = x^\mu \circ \psi$, and $D\psi$ denotes the Jacobian of ψ. We denote the coefficient of the Lagrangian $\mathscr{L}_t$ by

$$\widetilde{\mathscr{L}_t} := L\left(\psi^\mu(t,\xi), (f^A \circ f)(\psi(t,\xi)), \frac{\partial(f^A \circ f)}{\partial \psi^\mu}\bigg|_{\psi(t,\xi)}\right) \det(D\psi). \tag{4.30}$$

Thus we write

$$\mathscr{L}_t = \widetilde{\mathscr{L}_t} d\xi^1 \wedge \cdots \wedge d\xi^n. \tag{4.31}$$

A section $f \in \Gamma(F)$ gives rise, by means of the reference system, to a one-parameter family of sections $\{\phi_t\}_{t\in I}$ of Φ. However, multiple sections of F may induce the same one-parameter family $\{\phi_t\}_{t\in I}$. This does not cause a problem, because $j^1_{\psi(t,\xi)} f$ is uniquely determined by $j^1_{(t,\xi)}\phi$. We make this statement precise in the following lemma.

Lemma 4.24 *Let $f \in \Gamma(F)$ be a section and let $\{\phi_t\}_{t\in I}$ be a one-parameter family in $\Gamma(\Phi)$, defined pointwise by $\phi(t,\xi) = \Psi^{-1}_{(t,\xi)} f(\psi(t,\xi))$. The pulled-back Lagrangian*

$$\mathscr{L}_t(j^1_\xi \phi_t) = \psi^*(\dot{\psi}_t \rfloor \mathbf{L} \circ j^1\Psi_t)(j^1_\xi(\phi_t)) = \psi^*\left(\dot{\psi}_t \rfloor \mathbf{L}(j^1_{\psi(t,\xi)} f)\right),$$

does not depend on the choice of f. Hence, $\mathscr{L}_t$ is uniquely determined.

Proof. We argue pointwise. Assume that $d\psi_{(t,\xi)}$ is invertible. By the inverse function theorem, ψ_t is a local diffeomorphism near ξ. Hence, in a neighbourhood of $\psi(t,\xi)$, the section f is determined by ϕ_t through

$$f = \Psi_t \circ \phi_t \circ \psi_t^{-1}.$$

Therefore $j^1_{\psi(t,\xi)} f$ is uniquely determined by $j^1_{(t,\xi)}\phi$. Since $\mathbf{L}$ depends only on $j^1 f$, it follows that the Lagrangian is uniquely determined by $j^1_{(t,\xi)}\phi$.

Now assume that $d\psi_{(t,\xi)}$ is not invertible. Then ψ is not a local diffeomorphism at (t,ξ), and the pullback vanishes

$$\psi^*_{(t,\xi)}(\dot{\psi}_t \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}) = \det\big(d\psi_{(t,\xi)}\big) d\xi^1 \wedge \cdots \wedge d\xi^n = 0.$$

The last computation implies that $\mathscr{L}_t(j^1_\xi \phi_t)$ is zero. Hence, it does not depend on the choice of f. Thus in both cases the pulled-back Lagrangian density is uniquely determined. $\square$

4.3.3 Hamiltonian on the phase space $\mathscr{P}_\Sigma$

In this subsection, we obtain the Hamiltonian function on the hypersurface. We derive it by applying the Legendre transform, as discussed in Subsection 3.1.2, to the local coefficient $\widetilde{\mathscr{L}_t}$ of the pulled-back Lagrangian. This coefficient is defined in (4.31) and expressed locally in (4.30) and (4.29).

The following lemma identifies the momentum conjugate to ϕ^A with the derivative of $\widetilde{\mathscr{L}_t}$ with respect to the velocity $\dot{\phi}^A$, thereby providing the relation needed to perform the Legendre transform.

Lemma 4.25 *Let $\{\phi_t\}_{t\in I}$ be a one-parameter family of sections of Φ related to $f \in \Gamma(F)$ by the reference system. The coefficient $\widetilde{\mathscr{L}_t}$ of the Lagrangian $\mathscr{L}$ on the hypersurface satisfies*

$$\frac{\partial \widetilde{\mathscr{L}_t}}{\partial \dot{\phi}^A} = \pi_A,$$

where we denote $\dot{\phi}^A := \frac{\partial \phi^A}{\partial t}$ and where π_A is obtained from the pullback discussed in Lemma 4.22.

Proof. Let $(y^i) = (t, \xi^i)$ be local coordinates on $I \times \Sigma$, and assume that p is an admissible section of P. We write p locally as $(p_A{}^\mu, f^A)$, where

$$p_A{}^\mu = \frac{\partial L}{\partial(\partial_\mu f^A)}.$$

Using equation (4.30), the derivative of $\widetilde{\mathscr{L}_t}$ is given by

$$\frac{\partial \widetilde{\mathscr{L}_t}}{\partial \dot{\phi}^A} = \det(D\psi) \frac{\partial L}{\partial \dot{\phi}^A}.$$

The function L depends on the local jet coordinates of $J^1 F$. More precisely, in local coordinates the Lagrangian is given by

$$(x^\mu, f^A, \partial_\mu f^A) \mapsto L(x^\mu, f^A, \partial_\mu f^A)\, dx^1 \wedge \cdots \wedge dx^{n+1}.$$

For details, see Definition 4.8. Since the jet coordinates depend on the derivatives of the section, the chain rule implies

$$\frac{\partial \widetilde{\mathscr{L}_t}}{\partial \dot{\phi}^A} = \det(D\psi) \sum_{\mu, B} \frac{\partial L}{\partial(\partial_\mu \Psi^B)} \frac{\partial(\partial_\mu \Psi^B)}{\partial \dot{\phi}^A}, \tag{4.32}$$

where we denote $\partial_\mu := \frac{\partial}{\partial \psi^\mu}$ and $\Psi^B := f^B \circ \Psi$.

The functions Ψ^B appear because the reference system induces an affine coordinate change on the jet coordinates:

$$\begin{aligned}
\partial_\mu f^B(j^1_{\psi(t,\xi)} f) &= \left.\frac{\partial(f^B \circ f)}{\partial \psi^\mu}\right|_{\psi(t,\xi)} \\
&= \left.\frac{\partial \Psi^B}{\partial \psi^\mu}\right|_{\phi(t,\xi)} && (4.33)\\
&= \left(\left.\frac{\partial \Psi^B}{\partial y^i}\right|_{\phi(t,\xi)} + \left.\frac{\partial \Psi^B}{\partial \phi^A}\right|_{\phi(t,\xi)} \frac{\partial \phi^A}{\partial y^i}(j^1_{\psi(t,\xi)} f) \right) \left.\frac{\partial y^i}{\partial \psi^\mu}\right|_{\psi(t,\xi)}. && (4.34)
\end{aligned}$$

We now evaluate both the derivatives $\frac{\partial L}{\partial(\partial_\mu \Psi^B)}$ and $\frac{\partial(\partial_\mu \Psi^B)}{\partial \dot{\phi}^A}$ of equation (4.32).

Since only the last argument of the function L depends on Ψ, taking the derivative with respect to $f^A \circ \Psi$ is the same as composing the derivative of f^A with Ψ. Expressed in formulas, this yields

$$\frac{\partial L}{\partial(\partial_\mu \Psi^B)} = \frac{\partial L}{\partial(\partial_\mu f^B)} \circ \Psi. \tag{4.35}$$

Moreover, we use equations (4.33) and (4.34) to compute the derivative of $\partial_\mu \Psi^B$ with respect to $\dot{\phi}^A$. We obtain

$$\frac{\partial(\partial_\mu \Psi^B)}{\partial \dot{\phi}^A} = \frac{\partial \Psi^B}{\partial \phi^A} \frac{\partial t}{\partial \psi^\mu}. \tag{4.36}$$

Combining equations (4.32), (4.35) and (4.36), we obtain

$$\begin{aligned} \frac{\partial \widetilde{\mathscr{L}_t}}{\partial \dot{\phi}^A} &= \det(D\psi) \sum_{\mu,B} \left(\frac{\partial L}{\partial(\partial_\mu f^B)} \circ \Psi \right) \frac{\partial \Psi^B}{\partial \phi^A} \frac{\partial t}{\partial \psi^\mu} \\ &= \det(D\psi) \sum_{\mu,B} ({p_B}^\mu \circ \Psi) \frac{\partial \Psi^B}{\partial \phi^A} \frac{\partial t}{\partial \psi^\mu} \end{aligned}$$

By the computation in Lemma 4.22, this is precisely the coefficient π_A. Therefore,

$$\frac{\partial \widetilde{\mathscr{L}_t}}{\partial \dot{\phi}^A} = \pi_A$$

as desired. □

Lemma 4.25 identifies the variables needed to apply the Legendre transformation described in Definition 3.7. The time derivative $\dot{\phi}^A$ plays the role of the original variable with respect to which the transformation is performed, while the transformed variable is the momentum

$$\pi_A = \frac{\partial \widetilde{\mathscr{L}_t}}{\partial \dot{\phi}^A}.$$

Thus, the pair $(\dot{\phi}^A, \pi_A)$ is the field-theoretic analogue of the pair of variables appearing in the Legendre transformation discussed in Proposition 3.8. It follows that the Hamiltonian density on the hypersurface is given locally by

$$\left(\pi_A(t)\dot{\phi}^A(t) - \widetilde{\mathscr{L}_t} \right) d\xi^1 \wedge \cdots \wedge d\xi^n.$$

Since the dynamical variables are fields on Σ, the corresponding Hamiltonian function is obtained by integrating this density over Σ.

Definition 4.26 Let $\pi(t)$ be a one-parameter family of sections of Π obtained from an admissible section of P. For fixed $t \in I$, we define the Hamiltonian on the phase space $\mathscr{P}_\Sigma$ by

$$\mathscr{H}_\Sigma(\pi, t) := \mathscr{H}_\Sigma(\pi_A, \phi^B, t) := \int_\Sigma (\pi_A(t)\dot{\phi}^A(t) - \widetilde{\mathscr{L}_t}) d^n\xi, \tag{4.37}$$

where $\widetilde{\mathscr{L}_t}$ denotes the coefficient of the pulled-back Lagrangian at fixed time $t \in I$.

Remark 4.27 We have considered the possibility of reformulating the Legendre transformation using the language of jet bundles. Such a reformulation could provide a more intrinsic explanation of the passage from the Lagrangian theory on spacetime to the Hamiltonian theory on the model hypersurface Σ.

In studying this question, we considered the geometric treatment of Legendre transformations developed by Saunders [71] and the related field-theoretic constructions discussed by Forger and Romero [35]. Recall that, in classical mechanics, the cotangent bundle is the natural target of the Legendre transformation. The idea would be to extend this construction to jet bundles. Since jet bundles have the structure of affine bundles, one must consider an appropriate affine dual in order to define a Legendre transformation. The precise application of these ideas to the hypersurface-based framework of Chruściel, Jezierski, and Kijowski remains to be worked out.

In the final section of this chapter, we will relate this Hamiltonian $\mathscr{H}_\Sigma$ to the dynamical system on the phase space $\mathscr{P}_\Sigma$.

4.4 Symplectic structure and Hamiltonian dynamics

The goal of this final section is to show that the Hamiltonian $\mathscr{H}_\Sigma$ constructed above equips the dynamical system on the phase space, introduced in Theorem 4.21, with the structure of a Hamiltonian dynamical system in the sense of Definition 3.28. To this end, we relate the Hamiltonian function $\mathscr{H}_\Sigma$, the symplectic form on the phase space, and the vector field generating the dynamics.

In Subsection 4.4.1, we use the Liouville one-form construction to obtain a symplectic form on the momentum space defined in Definition 4.13. By integrating the resulting fibrewise symplectic form over Σ, we construct an antisymmetric two-form on the space of sections of the momentum bundle Π associated with Φ.

In Subsection 4.4.2, we establish a technical identity for the first variation of the pulled-back Lagrangian. Readers who are primarily interested in the final result may skip the details of the calculation and proceed directly to Proposition 4.31.

In Subsection 4.4.3, we prove Theorem 4.32, which shows that $\mathscr{H}_\Sigma$ equips the phase space with the structure of a Hamiltonian dynamical system.

Finally, in Subsection 4.4.4, we construct a reference system from the flow of a global vector field on M. This construction yields a corresponding set of dynamical variables and allows us to determine the resulting form of the Hamiltonian and its variation. The Hamiltonian obtained in this setting will provide the foundation for the Hamiltonian formulation of general relativity developed in Chapter 5.

4.4.1 Antisymmetric two-form on the momentum bundle

In this subsection, we associate an antisymmetric two-form with the momentum bundle. We first construct the form fibrewise. For each $\xi \in \Sigma$, we define a canonical symplectic form Ω_ξ on Π_ξ using the Liouville one-form construction. Readers who are unfamiliar with this construction may consult Subsection 3.2.2. We then integrate these fibrewise forms over Σ to obtain a two-form Ω_Σ on the space of sections of the momentum bundle Π.

It is not a priori clear that Ω_Σ is itself a symplectic form, as this requires additional properties, in particular nondegeneracy, to be verified on the space of sections. Such properties generally have to be checked on a case-by-case basis. Since the arguments below require only that Ω_Σ be an antisymmetric two-form, we do not address this question in full generality.

We begin by constructing a symplectic form on the momentum space, as defined in Definition 4.13. Let (ξ^i) be local coordinates on Σ and let (ϕ^A, ξ^i) be local coordinates on Φ. For each $\xi \in \Sigma$, recall that

$$\Pi_\xi = T^*\Phi_\xi \otimes \tilde{\Lambda}^n T^*_\xi \Sigma. \tag{4.38}$$

For fixed ξ, the density factor $\tilde{\Lambda}^n T^*_\xi \Sigma$ is independent of the point of the fibre Φ_ξ. Thus Π_ξ is a density-valued cotangent bundle over Φ_ξ. Since $\tilde{\Lambda}^n T^*_\xi \Sigma$ is one-dimensional, a choice of local orientation gives a local generator $d\xi^1 \wedge \cdots \wedge d\xi^n$. Hence every element of Π_ξ can be written locally as

$$\pi_A \, d\phi^A \otimes (d\xi^1 \wedge \cdots \wedge d\xi^n).$$

The coefficients π_A, together with the coordinates (ξ^i, ϕ^A), form the local coordinates (ξ^i, ϕ^A, π_A) on Π.

Following Subsection 3.2.2, we now construct the Liouville 1-form on a cotangent bundle to each fibre. Let $\mathrm{pr}_{\Pi \to \Phi} \colon \Pi \to \Phi$ denote the canonical projection. For $\pi \in \Pi_\xi$, the Liouville 1-form Θ_ξ on Π_ξ is defined by the pullback

$$\Theta_\xi(\pi) := \mathrm{pr}^*_{\Pi \to \Phi} \, \pi. \tag{4.39}$$

The canonical symplectic form on Π_ξ is then defined by

$$\Omega_\xi := d\Theta_\xi, \tag{4.40}$$

where d denotes the exterior derivative on the manifold Π_ξ, for fixed ξ.

Let us compute Θ_ξ locally. A tangent vector $Y \in T_\pi \Pi_\xi$ is of the form

$$Y = u^A \frac{\partial}{\partial \pi_A} + v^A \frac{\partial}{\partial \phi^A}.$$

The projection forgets the coordinates π_A, so

$$d \, \mathrm{pr}_{\Pi \to \Phi}(Y) = v^A \frac{\partial}{\partial \phi^A}.$$

Therefore

$$\begin{aligned} \Theta_\xi(\pi)(Y) &= \pi(d \, \mathrm{pr}_{\Pi \to \Phi}(Y)) \\ &= \pi(v^A \frac{\partial}{\partial \phi^A}) \\ &= \pi_A v^A (d\xi^1 \wedge \cdots \wedge d\xi^n). \end{aligned}$$

Hence, locally the Liouville 1-form reads

$$\Theta_\xi = \pi_A \, d\phi^A \otimes (d\xi^1 \wedge \cdots \wedge d\xi^n),$$

and the canonical symplectic form is

$$\Omega_\xi = (d\pi_A \wedge d\phi^A) \otimes (d\xi^1 \wedge \cdots \wedge d\xi^n). \tag{4.41}$$

We now integrate these fibrewise symplectic forms over Σ. Let $\sigma \in \Gamma(\Pi)$, and let X_1 and X_2 be tangent vectors to the space of sections at σ. Thus, for each $\xi \in \Sigma$, the vectors $X_1(\xi)$ and $X_2(\xi)$ are based at the same point $\sigma(\xi) \in \Pi_\xi$. We define

$$\Omega_\Sigma(X_1, X_2) := \int_\Sigma \Omega_\xi(X_1(\xi), X_2(\xi)). \tag{4.42}$$

In local coordinates, write

$$X_a(\xi) = \delta_a \pi_A(\xi) \left.\frac{\partial}{\partial \pi_A}\right|_{\sigma(\xi)} + \delta_a \phi^A(\xi) \left.\frac{\partial}{\partial \phi^A}\right|_{\sigma(\xi)}, \qquad a = 1, 2,$$

where $\delta_a \pi_A$ and $\delta_a \phi^A$ denote the coefficient functions of the corresponding components of X_a. We use this convention throughout the thesis. It follows that

$$\begin{aligned}\Omega_\Sigma(X_1, X_2) &= \int_\Sigma (d\pi_A \wedge d\phi^A)\big(X_1(\xi), X_2(\xi)\big)\, d\xi^1 \wedge \cdots \wedge d\xi^n \\ &= \int_\Sigma \big(\delta_1 \pi_A(\xi)\delta_2\phi^A(\xi) - \delta_2\pi_A(\xi)\delta_1\phi^A(\xi)\big)\, d\xi^1 \wedge \cdots \wedge d\xi^n, \end{aligned} \tag{4.43}$$

provided that the integrand belongs to $L^1(\Sigma)$, so that the integral is well defined. Since the fibrewise canonical symplectic form (4.41) is defined independently of the choice of coordinates, the local expression (4.43) is also coordinate independent.

Remark 4.28 In the construction considered here, this form is obtained by integrating a family of symplectic forms over Σ. To place this argument on a firmer geometric foundation, one would first need to determine conditions under which integration preserves closedness and nondegeneracy.

In Chapter 5, it will be necessary to express the antisymmetric two-form Ω_Σ in terms of the variables of the momentum bundle P restricted to the hypersurface $i(\Sigma) \subset M$. Recall that (π_A, ϕ^A) is obtained by pulling back the corresponding data $({p_A}^\mu, f^A)$ along ψ_t, and that $\psi_0 = i$. We may therefore rewrite (4.43) as an antisymmetric two-form on variations of the momentum data along $i(\Sigma)$.

Let X_a, for $a = 1, 2$, be vector fields as above, and let $\widetilde{X}_a$ denote the corresponding vector fields along $i(\Sigma)$. Locally, these are written as

$$\widetilde{X}_a = (\delta_a p)_A{}^\mu \frac{\partial}{\partial {p_A}^\mu} + (\delta_a f)^A \frac{\partial}{\partial f^A}.$$

We then obtain

$$\begin{aligned}\Omega_\Sigma(X_1, X_2) &= \int_\Sigma \big(\delta_1 \pi_A\, \delta_2 \phi^A - \delta_2 \pi_A\, \delta_1 \phi^A\big)\, d\xi^1 \wedge \cdots \wedge d\xi^n \\ &= \int_{i(\Sigma)} \big[(\delta_1 p)_A{}^\mu (\delta_2 f)^A - (\delta_2 p)_A{}^\mu (\delta_1 f)^A\big]\, dS_\mu \\ &=: \Omega_{i(\Sigma)}(\widetilde{X}_1, \widetilde{X}_2), \end{aligned} \tag{4.44}$$

where dS_μ denotes the odd n-form

$$dS_\mu := \frac{\partial}{\partial x^\mu} \lrcorner \left(dx^1 \wedge \cdots \wedge dx^{n+1}\right).$$

4.4.2 First variation of the pulled-back Lagrangian

This subsection proves the identity

$$\delta\mathscr{L}(j^1\phi(t)) = d\boldsymbol{\pi}(\delta\phi(t)), \tag{4.45}$$

together with its local coordinate expression. These identities are stated and proved in Proposition 4.31.

Let (Ψ, ψ) be a reference system and let $p_\lambda \in \Gamma(P)$ be a one-parameter family of sections satisfying the field equations (4.10)–(4.11). By Proposition 4.16, this induces a two-parameter family of sections $\pi_{t,\lambda} \in \Gamma(\Pi)$. The family p_λ also induces sections of F by

$$f_\lambda := \mathrm{pr}_{P\to F}(p_\lambda).$$

Throughout this subsection, δ denotes differentiation with respect to λ at $\lambda = 0$. To keep the notation readable, we suppress the λ-dependence when no ambiguity can arise. Thus we write f, p, $\phi(t)$, and $\boldsymbol{\pi}$ for the corresponding λ-dependent families, with the understanding that after applying δ all quantities are evaluated at $\lambda = 0$. We keep the dependence on t visible, because the pulled-back Lagrangian is defined on $I \times \Sigma$.

The proof of Proposition 4.31 is split into two lemmas: the first shows that variation commutes with the pullback by ψ, and the second identifies the pulled-back term $p(\delta f)$ with $\boldsymbol{\pi}(\delta\phi)$.

Lemma 4.29 *Let δ denote differentiation with respect to λ at $\lambda = 0$, and let $\mathbf{L}$ be the Lagrangian for the field theory. Suppose that the reference system (Ψ, ψ) is independent of λ. Then, for each $t \in I$,*

$$\delta\left(\psi_t^*\left(\dot{\psi}_t \rfloor \mathbf{L}(j^1 f)\right)\right) = \psi_t^*\left(\dot{\psi}_t \rfloor \delta\big(\mathbf{L}(j^1 f)\big)\right).$$

Moreover, let $\{\phi(t)\}_{t\in I}$ be the one-parameter family of sections of Φ related to f through the reference system, as in Remark 4.7. Then, for every $t \in I$ and $\xi \in \Sigma$,

$$\delta\mathscr{L}_t\big(j^1_\xi\phi(t)\big) = \psi_t^*\left(\dot{\psi}_t \rfloor \delta\big(\mathbf{L}(j^1_{\psi(t,\xi)} f)\big)\right).$$

Proof. We verify the identity pointwise. Let $v_1, \ldots, v_n \in T_\xi\Sigma$. By the definitions of contraction and pullback,

$$\begin{aligned}
&\psi_t^*\left(\dot{\psi}_t \rfloor \mathbf{L}(j^1_{\psi(t,\xi)} f)\right)(v_1, \ldots, v_n) \\
&\qquad = \left(\mathbf{L}(j^1_{\psi(t,\xi)} f)\right)\left(\dot{\psi}_t(\xi), d\psi_t|_\xi(v_1), \ldots, d\psi_t|_\xi(v_n)\right).
\end{aligned}$$

Since the reference system does not depend on λ, both ψ_t and its first variation $\dot{\psi}_t$ are independent of λ. Differentiating at $\lambda = 0$ therefore gives

$$\delta\left(\psi_t^*\left(\dot{\psi}_t \rfloor \mathbf{L}(j^1 f)\right)\right) = \psi_t^*\left(\dot{\psi}_t \rfloor \delta\mathbf{L}(j^1 f)\right).$$

Since f and $\phi(t)$ are related by the reference system, the definition of the pulled-back Lagrangian gives

$$\mathscr{L}_t\big(j^1_\xi\phi(t)\big) = \psi_t^*\left(\dot{\psi}_t \rfloor \mathbf{L}(j^1_{\psi(t,\xi)} f)\right).$$

Differentiating this identity with respect to λ at $\lambda = 0$ proves the second statement. □

Lemma 4.30 *With the setup of Lemma 4.29, we have*

$$\psi^*(p(\delta f)) = \boldsymbol{\pi}(\delta\phi(t)). \tag{4.46}$$

Proof. We write the variations locally as

$$\begin{aligned}\delta f &= \delta f^A \frac{\partial}{\partial f^A}\\ \delta\phi &= \delta\phi^A(t)\frac{\partial}{\partial \phi^A}.\end{aligned}$$

From an earlier equation (4.14), we have

$$p(\delta f) = p_A{}^\mu \delta f^A \partial_\mu \rfloor (dx^1 \wedge \cdots \wedge dx^{n+1}),$$

where we regard p as a section from $M \to P$. Hence, $p_A{}^\mu = p_A{}^\mu \circ f$ and $f^A = f^A \circ f$. Pulling this expression back along ψ gives

$$\psi^*(p(\delta f)) = (p_A{}^\mu \circ f \circ \psi)(\psi^* \delta f^A)\psi^*\left(\partial_\mu \rfloor (dx^1 \wedge \cdots \wedge dx^{n+1})\right). \tag{4.47}$$

Since ψ does not depend on λ, it can be moved through the variation

$$\begin{aligned}\psi^*\delta f^A &= \left.\frac{d(f^A\circ f)}{d\lambda}\right|_{\lambda=0}\circ\psi\\ &= \left.\frac{d(f^A\circ f\circ\psi)}{d\lambda}\right|_{\lambda=0}.\end{aligned}$$

Using the reference system and the chain rule, with $\Psi^A := f^A \circ \Psi$, we find

$$\begin{aligned}\psi^*\delta f^A &= \left.\frac{d(f^A\circ\Psi\circ\phi(t))}{d\lambda}\right|_{\lambda=0}\\ &= \frac{\partial\Psi^A}{\partial\phi^B}\left.\frac{d\phi^B(t)}{d\lambda}\right|_{\lambda=0}\\ &= \frac{\partial\Psi^A}{\partial\phi^B}\delta\phi^B(t).\end{aligned}$$

Substituting this into the previous expression (4.47) and using the definitions of the coefficients π_B and $\pi_B{}^a$ from the total pullback, we obtain

$$\begin{aligned}\psi^*(p(\delta f)) = {}& \pi_B\delta\phi^B(t)\, d\xi^1\wedge\cdots\wedge d\xi^n\\ &+ \pi_B{}^a\delta\phi^B(t)\,\partial_a\rfloor(dt\wedge d\xi^1\wedge\cdots\wedge d\xi^n).\end{aligned}$$

This is precisely the total pullback $\boldsymbol{\pi}$ evaluated on $\delta\phi$, so

$$\psi^*(p(\delta f)) = \boldsymbol{\pi}(\delta\phi(t)).$$ □

We now combine the Lemmas 4.29 and 4.30 to obtain the desired first-variation identity.

Proposition 4.31 *With the setup of Lemma 4.29, let*

$$\iota_t \colon \Sigma \longrightarrow I \times \Sigma, \qquad \iota_t(\xi) := (t, \xi).$$

Then

$$\delta \mathscr{L}_t\big(j^1\phi(t)\big) = \psi_t^*\left(\dot{\psi}_t \rfloor \delta\big(\mathbf{L}(j^1 f)\big)\right) = \psi_t^*\left(\dot{\psi}_t \rfloor \mathbf{d}\big(p(\delta f)\big)\right) = \iota_t^*\left(\frac{\partial}{\partial t} \rfloor \mathbf{d}\big(\boldsymbol{\pi}(\delta\phi(t))\big)\right). \tag{4.48}$$

In local coordinates (t, ξ^a) on $I \times \Sigma$, this becomes

$$\delta \mathscr{L}_t\big(j^1_\xi \phi(t)\big) = \left(\frac{\partial\big(\pi_A(t)\delta\phi^A(t)\big)}{\partial t} + \frac{\partial\big(\pi_A{}^a(t)\delta\phi^A(t)\big)}{\partial \xi^a}\right) d\xi^1 \wedge \cdots \wedge d\xi^n. \tag{4.49}$$

Proof of Proposition 4.31. By the definition of the pulled-back Lagrangian and Lemma 4.29, we have

$$\delta \mathscr{L}_t\big(j^1\phi(t)\big) = \psi_t^*\left(\dot{\psi}_t \rfloor \delta\big(\mathbf{L}(j^1 f)\big)\right).$$

By Proposition 4.12,

$$\delta\big(\mathbf{L}(j^1 f)\big) = \mathbf{d}\big(p(\delta f)\big).$$

It follows that

$$\delta \mathscr{L}_t\big(j^1\phi(t)\big) = \psi_t^*\left(\dot{\psi}_t \rfloor \mathbf{d}\big(p(\delta f)\big)\right).$$

Since $\psi_t = \psi \circ \iota_t$ and

$$\dot{\psi}_t = d\psi\left(\frac{\partial}{\partial t}\right),$$

contraction and pullback satisfy

$$\psi_t^*\left(\dot{\psi}_t \rfloor \alpha\right) = \iota_t^*\left(\frac{\partial}{\partial t} \rfloor \psi^* \alpha\right)$$

for every differential form α on M. Taking $\alpha = \mathbf{d}(p(\delta f))$ and using the commutativity of the exterior derivative with pullback gives

$$\begin{aligned} &\psi_t^*\left(\dot{\psi}_t \rfloor \mathbf{d}\big(p(\delta f)\big)\right) \\ &\qquad = \iota_t^*\left(\frac{\partial}{\partial t} \rfloor \mathbf{d}\left(\psi^*\big(p(\delta f)\big)\right)\right). \end{aligned}$$

By Lemma 4.30,

$$\psi^*\big(p(\delta f)\big) = \boldsymbol{\pi}\big(\delta\phi(t)\big).$$

Therefore,

$$\delta \mathscr{L}_t\big(j^1\phi(t)\big) = \iota_t^*\left(\frac{\partial}{\partial t} \rfloor \mathbf{d}\big(\boldsymbol{\pi}(\delta\phi(t))\big)\right).$$

It remains to compute the local expression. We have

$$\boldsymbol{\pi}\big(\delta\phi(t)\big) = \pi_A(t)\delta\phi^A(t)\, d\xi^1 \wedge \cdots \wedge d\xi^n + \pi_A{}^a(t)\delta\phi^A(t)\, \partial_a \lrcorner \left(dt \wedge d\xi^1 \wedge \cdots \wedge d\xi^n\right).$$

Taking the exterior derivative yields

$$\mathbf{d}\big(\boldsymbol{\pi}(\delta\phi(t))\big) = \left(\frac{\partial\big(\pi_A(t)\delta\phi^A(t)\big)}{\partial t} + \frac{\partial\big(\pi_A{}^a(t)\delta\phi^A(t)\big)}{\partial \xi^a}\right) dt \wedge d\xi^1 \wedge \cdots \wedge d\xi^n.$$

Contracting with $\partial/\partial t$ and pulling back along ι_t removes the factor dt. Consequently,

$$\delta\mathscr{L}_t\big(j^1_\xi\phi(t)\big) = \left(\frac{\partial\big(\pi_A(t)\delta\phi^A(t)\big)}{\partial t} + \frac{\partial\big(\pi_A{}^a(t)\delta\phi^A(t)\big)}{\partial \xi^a}\right) d\xi^1 \wedge \cdots \wedge d\xi^n.$$

This proves the local expression and completes the proof. □

4.4.3 Hamiltonian dynamics

In the previous subsection, we showed that the first variation of the pulled-back Lagrangian can be expressed in terms of the exterior derivative of $\boldsymbol{\pi}(\delta\phi(t))$. In this subsection, we use this identity to relate the variation of the Hamiltonian $\mathscr{H}_\Sigma$ to the antisymmetric two-form Ω_Σ. This relation shows that the dynamical system on the phase space is Hamiltonian.

We restate the definition of the dynamical system using the notation introduced in this chapter and summarize the results established so far.

The phase space $\mathscr{P}_\Sigma$ associated with Σ, defined in Definition 4.17, together with the time-evolution maps

$$T_t \colon \mathscr{P}_\Sigma \to \mathscr{P}_\Sigma, \qquad t \in I,$$

defined by

$$T_t(\pi_A(s), \phi^B(s)) = (\pi_A(s+t), \phi^B(s+t)),$$

was shown in Theorem 4.21 to form a **dynamical system**. In particular, the following three properties are satisfied:

1. $T_0 = \mathrm{id}$, where id denotes the identity map on $\mathscr{P}_\Sigma$;
2. $T_a \circ T_b = T_{a+b}$ for all $a, b \in I$;
3. for every $\pi \in \mathscr{P}_\Sigma$, the map
$$t \mapsto T_t(\pi)$$
is a differentiable curve in $\mathscr{P}_\Sigma$.

Moreover, we showed that the vector field generating the dynamics was given by

$$\mathscr{X}(\pi) = (\dot{\pi}_A, \dot{\phi}^B).$$

Recall that the dynamical system $(T_t, \mathscr{P}_\Sigma)$ is said to be **Hamiltonian** if for every **curve**

$$\lambda \to \pi_\lambda$$

in $\mathscr{P}_\Sigma$, the function $\mathscr{H}_\Sigma(\pi_\lambda, t)$ is differentiable at $\lambda = 0$ and satisfies

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} \mathscr{H}_\Sigma(\pi_\lambda, t) = -\Omega_\Sigma\left(\mathscr{X}, \left.\frac{d\pi_\lambda}{d\lambda}\right|_{\lambda=0}\right) = -\left((\dot\pi_A, \dot\phi^B), (\delta\pi_A, \delta\phi^B)\right).$$

To show that the phase space has the structure of a Hamiltonian dynamical system, we must assume that the boundary integral over $\partial\Sigma$ vanishes and that $\mathscr{H}_\Sigma$ is differentiable along the admissible curves in phase space. Neither issue is treated in full generality here; both will be addressed in Subsection 5.3.3.

Theorem 4.32 *Let $T_t\colon \mathscr{P}_\Sigma \to \mathscr{P}_\Sigma$ be the dynamical system, defined by*

$$T_t(\pi_A(s), \phi^B(s)) := (\pi_A(s+t), \phi^B(s+t)).$$

Suppose the hypersurface Σ is compact, and that the boundary term

$$\int_{\partial\Sigma} \pi_A{}^a \delta\phi^A \, d\sigma_a$$

vanishes. Then the Hamiltonian function $\mathscr{H}_\Sigma$ makes this dynamical system into a Hamiltonian dynamical system. More precisely,

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} \mathscr{H}_\Sigma(\pi_\lambda, t) = -\Omega_\Sigma\left(\mathscr{X}, \left.\frac{d\pi_\lambda}{d\lambda}\right|_{\lambda=0}\right), \tag{4.50}$$

where $\mathscr{X}$ is the vector field generating the dynamics.

Proof. To improve readability, we suppress the index λ in π_λ. We first evaluate the right-hand side of (4.50). By Theorem 4.21, the vector field generating the dynamics T_t is locally given by

$$\mathscr{X}(\pi) = (\dot\pi_A, \dot\phi^B).$$

The variation of the curve π is written locally as

$$\left.\frac{d\pi}{d\lambda}\right|_{\lambda=0} = \delta\pi = (\delta\pi_A, \delta\phi^B).$$

Using the local expression (4.43) for Ω_Σ, we obtain

$$\Omega_\Sigma((\dot\pi_A, \dot\phi^B), (\delta\pi_A, \delta\phi^B)) = \int_\Sigma \left(\dot\pi_A \delta\phi^A - \dot\phi^A \delta\pi_A\right) d\xi^1 \wedge \cdots \wedge d\xi^n.$$

We now evaluate the left-hand side of (4.50). Differentiation with respect to the parameter λ means taking the first variation. Hence

$$-\left.\frac{d\mathscr{H}_\Sigma(\pi, t)}{d\lambda}\right|_{\lambda=0} = -\delta\mathscr{H}_\Sigma(\pi_A, \phi^B, t) = -\delta\int_\Sigma \left(\pi_A\dot\phi^A - \widetilde{\mathscr{L}_t}\right) d\xi^1 \wedge \cdots \wedge d\xi^n.$$

Since the domain Σ is compact, we may differentiate under the integral sign. This yields

$$-\left.\frac{d\mathscr{H}_\Sigma(\pi, t)}{d\lambda}\right|_{\lambda=0} = -\int_\Sigma \delta\left(\pi_A\dot\phi^A - \widetilde{\mathscr{L}_t}\right) d\xi^1 \wedge \cdots \wedge d\xi^n. \tag{4.51}$$

Proposition 4.31 gave

$$\delta\mathscr{L}_t\big(j^1_\xi\phi(t)\big) = \widetilde{\mathscr{L}_t}d\xi^1\wedge\cdots\wedge d\xi^n\left(\frac{\partial\big(\pi_A(t)\delta\phi^A(t)\big)}{\partial t}+\frac{\partial\big(\pi_A{}^a(t)\delta\phi^A(t)\big)}{\partial\xi^a}\right)d\xi^1\wedge\cdots\wedge d\xi^n.$$

Hence, the variation of the integrand

$$\begin{aligned}\delta(\pi_A\dot\phi^A-\widetilde{\mathscr{L}_t}) &= \dot\phi^A\delta\pi_A+\pi_A\delta\dot\phi^A-\delta\widetilde{\mathscr{L}_t}\\ &= \dot\phi^A\delta\pi_A+\pi_A\delta\dot\phi^A-\frac{\partial(\pi_A(t)\delta\phi^A(t))}{\partial t}-\frac{\partial(\pi^a{}_A(t)\delta\phi^A(t))}{\partial\xi^a}\\ &= \dot\phi^A\delta\pi_A+\pi_A\delta\dot\phi^A-(\dot\pi_A\delta\phi^A+\pi_A\delta\dot\phi^A)-\partial_a(\pi_A{}^a\delta\phi^A)\\ &= \dot\phi^A\delta\pi_A-\dot\pi_A\delta\phi^A-\partial_a(\pi_A{}^a\delta\phi^A).\end{aligned}$$

Substituting this expression into (4.51) and applying Stokes' theorem Theorem 2.26 yield

$$\begin{aligned}-\left.\frac{d\mathscr{H}_\Sigma(\pi,t)}{d\lambda}\right|_{\lambda=0} &= -\int_\Sigma\delta(\pi_A\dot\phi^A-\widetilde{\mathscr{L}_t})d\xi^1\wedge\cdots\wedge d\xi^n\\ &= \int_\Sigma\Big(\dot\pi_A\delta\phi^A-\dot\phi^A\delta\pi_A\Big)d^n\xi+\int_{\partial\Sigma}\pi_A{}^a\delta\phi^A d\sigma_a,\end{aligned}\tag{4.52}$$

where

$$d\sigma_a := \partial_a\rfloor(d\xi^1\wedge\cdots\wedge d\xi^n)$$

is the surface element induced on $\partial\Sigma$. By assumption, the boundary integral vanishes. Therefore,

$$-\left.\frac{d\mathscr{H}_\Sigma(\pi,t)}{d\lambda}\right|_{\lambda=0} = \int_{\partial\Sigma}\pi_A{}^a\delta\phi^A d\sigma_a = \Omega_\Sigma((\dot\pi_A,\dot\phi^B),(\delta\pi_A,\delta\phi^B)).$$

Hence, (4.50) holds with $\mathscr{X} = (\dot\pi_A,\dot\phi^B)$, and we have shown that the dynamical system is Hamiltonian. □

Remark 4.33 The theorem is stated here in a simplified setting that is sufficient for the present discussion. In the noncompact case, both differentiation under the integral sign and the application of Stokes' theorem require additional care. To interchange the variation and the integral, one must impose suitable regularity and integrability conditions on the one-parameter families under consideration. In the asymptotically flat setting relevant to this thesis, additional issues concerning compactification and regularity arise. In particular, requiring the compactification to be smooth is generally too restrictive. We do not discuss these matters here; see Van Voorthuizen [81, Section 3.3] for further details.

On a noncompact hypersurface, Stokes' theorem is typically applied to an exhaustion of Σ by compact submanifolds. The resulting boundary integral must then be understood as a limit, and the convergence of this limit must be verified. The corresponding analysis in the asymptotically flat case will be discussed in Section 5.3.

4.4.4 Hamiltonian for reference system for general relativity

We conclude this chapter by specifying the Hamiltonian for a particular choice of reference system. The Hamiltonian has already been formulated on the model hypersurface Σ. In this section, we show that this specific reference system allows both the Hamiltonian and its variation to be rewritten as integrals over the corresponding embedded hypersurface $\mathscr{S} = \psi_t(\Sigma) \subset M$. This reformulation is needed in Section 5.1, where both the model hypersurface Σ and the embedded hypersurface $\mathscr{S}$ are used.

Let $X \in \mathscr{X}(M)$ be a vector field on M and let $\psi\colon I \times \Sigma \to M$ be its flow. Then ψ_t is an embedding. Therefore, ψ defines a motion of the hypersurface $\Sigma \subset M$. We use the following result from Chruściel, Jezierski, and Kijowski [27, p. 27].

Proposition 4.34 *Let F be a tensor bundle over M, and let ψ be the flow of a vector field $X \in \mathscr{X}(M)$. Then the diffeomorphisms ψ_t induce a canonical lifted flow*

$$\Psi : I \times F \to F.$$

Thus, after restricting the lifted flow to the pullback bundle $\Phi = \psi_0^* F$, we obtain a reference system (Ψ, ψ), as illustrated below.

$$\begin{tikzcd}
I \times \Phi \arrow[r, "\Psi"] \arrow[d] & F \arrow[d] \\
I \times \Sigma \arrow[r, "\psi"'] & M
\end{tikzcd}$$

The following lemma is useful for expressing the Hamiltonian as an integral over the embedded hypersurface $\psi_t(\Sigma) \subset M$.

Lemma 4.35 *Let $\psi\colon I \times \Sigma \to M$ be induced by the flow of a vector field $X \in \mathscr{X}(M)$. Let $f \in \Gamma(F)$ and $\{\phi_t\}_{t\in I}$ be related by the reference system. Then, for every $t \in I$, the Lagrangian defined in Definition 4.23 is given by*

$$\mathscr{L}_t\big(j^1_\xi \phi_t\big) = \psi_t^* \left(X_{\psi(t,\xi)} \rfloor \mathbf{L}\big(j^1_{\psi(t,\xi)} f\big)\right). \tag{4.53}$$

Equivalently, using the coefficient $\widetilde{\mathscr{L}_t}$ defined in (4.31), we obtain

$$\widetilde{\mathscr{L}_t}\, d\xi^1 \wedge \cdots \wedge d\xi^n = \psi_t^* \left(X \rfloor \mathbf{L}(j^1 f)\right). \tag{4.54}$$

Proof. By definition of the pulled-back Lagrangian,

$$\mathscr{L}_t\big(j^1_\xi \phi_t\big) = \psi_t^* \left(\dot{\psi}_t \rfloor \mathbf{L}\big(j^1_{\psi(t,\xi)} f\big)\right),$$

where

$$\dot{\psi}_t(\xi) = \frac{\partial \psi}{\partial t}(t, \xi).$$

Since ψ is the flow of X, we have

$$\dot{\psi}_t(\xi) = X_{\psi(t,\xi)}.$$

Substituting this identity into the definition of $\mathscr{L}_t$ gives

$$\mathscr{L}_t(j^1_\xi \phi_t) = \psi_t^* \left(X_{\psi(t,\xi)} \rfloor \mathbf{L}(j^1_{\psi(t,\xi)} f) \right).$$

Finally, by the definition of the local coefficient $\widetilde{\mathscr{L}_t}$,

$$\mathscr{L}_t = \widetilde{\mathscr{L}_t}\, d\xi^1 \wedge \cdots \wedge d\xi^n.$$

Hence

$$\widetilde{\mathscr{L}_t}\, d\xi^1 \wedge \cdots \wedge d\xi^n = \psi_t^* \left(X \rfloor \mathbf{L}(j^1 f) \right). \qquad \square$$

For fixed $t \in I$, we rewrite the Hamiltonian as an integral over the embedded hypersurface $\psi_t(\Sigma) \subset M$ in a form that will be used in Section 5.1.

Theorem 4.36 *Let M be an oriented $(n+1)$-dimensional manifold, and let $F \to M$ be a tensor bundle representing the fields of the theory. Let $f \in \Gamma(F)$ be a field, and let $p \in \Gamma(P)$ be a corresponding momentum section, locally written as*

$$p = p_A{}^\mu df^A \otimes (\partial_\mu \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}).$$

Let $\mathbf{L}$ be the Lagrangian, locally given by

$$\mathbf{L}(j^1_x f) = L(f^A, \frac{\partial f^A}{\partial x^\mu}, x^\mu) dx^1 \wedge \cdots \wedge dx^{n+1}.$$

Suppose that ψ is a motion, as in Definition 4.2, obtained from the flow of $X \in \mathscr{X}(M)$, and that Ψ is the canonical lift of this flow to the tensor bundle F. Then on $\mathscr{S} := \psi_t(\Sigma)$ the Hamiltonian can be written as

$$\mathscr{H}_\Sigma(\pi, t) = \int_{\mathscr{S}} p_A{}^\mu \pounds_X f^A - X^\mu L \, dS_\mu,$$

where $\pounds_X f^A$ denotes the Lie derivative of f with respect to X, and

$$dS_\mu := \frac{\partial}{\partial x^\mu} \rfloor dx^1 \wedge \cdots \wedge dx^{n+1}.$$

Proof. We transport the Hamiltonian to spacetime using the reference system (Ψ, ψ). By Proposition 4.6,

$$\phi^A(t, \xi) = \Psi_t^{-1}\big(f^A(\psi(t,\xi))\big).$$

Since Ψ is the canonical lift to a tensor bundle, we have

$$d\Psi_t \frac{\partial}{\partial t} \left(\Psi_t^{-1} f^A(\psi(t,\xi)) \right) = \pounds_X f^A.$$

Using Lemma 4.35 for the Lagrangian term, we obtain

$$\begin{aligned}
\mathscr{H}_\Sigma(\pi, t) &= \int_\Sigma \pi_A(t) \frac{\partial \phi^A(t)}{\partial t} - \widetilde{\mathscr{L}_t} d^n \xi \\
&= \int_\Sigma \pi_A(t) \frac{\partial \Psi^{-1}_{(t,\xi)} f(\psi(t,\xi))}{\partial t} - \psi_t^*(X \rfloor \mathbf{L}(j^1 f)) \\
&= \int_{\psi_t(\Sigma)} p_A{}^\mu \Psi_* \frac{\partial \Psi^{-1}_{(t,\xi)} f(\psi(t,\xi))}{\partial t} - X \rfloor (L(f^A, \frac{\partial f^A}{\partial x^\mu}, x^\mu) dx^1 \wedge \cdots \wedge dx^{n+1}) \\
&= \int_{\mathscr{S}} p_A{}^\mu \pounds_X f(\psi(t,\xi)) - X^\mu L \frac{\partial}{\partial x^\mu} \rfloor (dx^1 \wedge \cdots \wedge dx^{n+1}) \\
&= \int_{\mathscr{S}} p_A{}^\mu \pounds_X f - X^\mu L dS_\mu.
\end{aligned} \qquad \square$$

To indicate the dependence on X, we denote the Hamiltonian as follows:

$$H(X, \mathscr{S}, p) := \int_{\mathscr{S}} p_A{}^\mu \pounds_X f - X^\mu L dS_\mu. \tag{4.55}$$

To establish a Hamiltonian dynamical system in general relativity, as developed in Chapter 5, we compute the variation of the Hamiltonian function H. We first clarify the meaning of $\dot{\pi}$, which denotes the time derivative of the canonical momentum after it has been pulled back along the reference system (Ψ, ψ). The reference system allows us to identify and compare the momentum data on hypersurfaces corresponding to different times. As shown in Proposition 4.16, the pulled-back momentum therefore defines a curve

$$t \longmapsto \pi(t)$$

in $\mathscr{P}_\Sigma$, and hence

$$\dot{\pi} := \frac{d\pi}{dt}$$

is well defined.

This should not be confused with a canonical spacetime derivative of the full momentum $p_A{}^\mu$. In general, differentiating $p_A{}^\mu$ along the curve

$$\gamma_\xi(t) := \psi(t, \xi)$$

requires the choice of a connection on the momentum bundle P, and hence a covariant derivative along γ_ξ. There is no canonical choice of such a derivative. Although this appears to introduce an arbitrariness, we will show that it does not affect the final variation formula, because only the terms corresponding to $\dot{\pi}$ contribute after integration.

In the present setting, the motion is induced by the flow of a global spacetime vector field X, and the relevant bundles are tensor bundles. The flow induces a canonical Lie derivative $\pounds_X$ on the fields. We denote the corresponding derivative on sections of the momentum bundle P by $\widetilde{\pounds}_X$. Following Chruściel–Jezierski–Kijowski [27, p. 35], we require $\widetilde{\pounds}_X$ to be compatible with $\pounds_X$ in the sense that

$$\psi_t^* \left((\widetilde{\pounds}_X p)(f) + p(\pounds_X f) \right) = \frac{\partial}{\partial t} \left(\psi_t^*(p(f)) \right). \tag{4.56}$$

This is the only property of $\widetilde{\pounds}_X$ needed for rewriting the variation of the Hamiltonian in Theorem 4.37.

Theorem 4.37 *Let M be an oriented $(n+1)$-dimensional manifold, and let $F \to M$ be a tensor bundle representing the fields of the theory. Let $f \in \Gamma(F)$ be a field, and let $p \in \Gamma(P)$ be a corresponding momentum section, locally written as*

$$p = p_A{}^\mu df^A \otimes (\partial_\mu \lrcorner dx^1 \wedge \cdots \wedge dx^{n+1}).$$

Let $\mathbf{L}$ be the Lagrangian, locally given by

$$\mathbf{L}(j^1_x f) = L(f^A, \frac{\partial f^A}{\partial x^\mu}, x^\mu) dx^1 \wedge \cdots \wedge dx^{n+1}.$$

Suppose that ψ is a motion, as in Definition 4.2, obtained from the flow of $X \in \mathscr{X}(M)$, and that Ψ is the canonical lift of this flow to the tensor bundle F. Then the variation of the Hamiltonian H, defined in (4.55), is given by

$$\begin{aligned} -\delta H &= \int_{\mathscr{S}} \left\{ (\pounds_X p_A{}^\mu)\delta f^A - (\pounds_X f^A)\delta p_A{}^\mu \right\} dS_\mu + \int_{\partial\mathscr{S}} X^{[\mu} p_A{}^{\nu]} \delta f^A dS_{\mu\nu} \\ &= \Omega_{\mathscr{S}}((\pounds_X p_A{}^\mu, \pounds_X f^A), (\delta p_A{}^\mu, \delta f^A)) + \int_{\partial\mathscr{S}} X^{[\mu} p_A{}^{\nu]} \delta f^A dS_{\mu\nu}. \end{aligned} \tag{4.57}$$

Here the odd form $dS_{\mu\nu}$ is defined by

$$\begin{aligned} dS_{\mu\nu} &= (\partial_\mu \wedge \partial_\nu) \rfloor dx^1 \wedge \cdots \wedge dx^{n+1} \\ &= dx^1 \wedge \cdots \wedge \widehat{dx^\mu} \wedge \cdots \wedge \widehat{dx^\nu} \wedge \cdots \wedge dx^{n+1}, \end{aligned}$$

where the hat indicates that we omit that term.

Sketch of proof. We prove the result for a scalar field f. The general case, $f \in \Gamma(F)$, follows analogously.

Taking the first variation of (4.55), while keeping X and $\mathscr{S}$ fixed, gives

$$\delta H = \int_{\mathscr{S}} [(\pounds_X f)\, \delta p^\mu + p^\mu\, \delta(\pounds_X f) - X^\mu \delta L]\, dS_\mu.$$

Since X is independent of the variation parameter, the variation commutes with the Lie derivative. Moreover, because f is a scalar field,

$$\pounds_X f = X^\alpha \partial_\alpha f, \qquad \pounds_X(\delta f) = X^\alpha \partial_\alpha(\delta f).$$

Using Proposition 4.12, namely $\delta L = \partial_\alpha\big(p^\alpha \delta f\big)$, we obtain

$$\delta H = \int_{\mathscr{S}} \left[(\pounds_X f)\, \delta p^\mu + p^\mu X^\alpha \partial_\alpha(\delta f) - X^\mu \partial_\alpha\big(p^\alpha \delta f\big) \right] dS_\mu. \tag{4.58}$$

The momentum p^μ is a vector; see Subsection 2.1.3. Its Lie derivative is given by

$$\pounds_X p^\mu = \partial_\alpha\big(p^\mu X^\alpha\big) - p^\alpha \partial_\alpha X^\mu. \tag{4.59}$$

A direct calculation using (4.59) yields

$$p^\mu X^\alpha \partial_\alpha(\delta f) - X^\mu \partial_\alpha\big(p^\alpha \delta f\big) = \partial_\alpha \left[\big(p^\mu X^\alpha - X^\mu p^\alpha\big) \delta f \right] - (\pounds_X p^\mu)\delta f.$$

Substituting this identity into (4.58), we find

$$\delta H = \int_{\mathscr{S}} [(\pounds_X f)\, \delta p^\mu - (\pounds_X p^\mu)\, \delta f]\, dS_\mu + \int_{\mathscr{S}} \partial_\alpha \left[\big(p^\mu X^\alpha - X^\mu p^\alpha\big) \delta f \right] dS_\mu. \tag{4.60}$$

Applying Stokes' theorem to the final term gives

$$-\delta H = \Omega_{\mathscr{S}}\left(\big(\pounds_X p^\mu, \pounds_X f\big), \big(\delta p^\mu, \delta f\big) \right) + \int_{\partial\mathscr{S}} X^{[\mu} p^{\nu]} \delta f\, dS_{\mu\nu}. \qquad \square$$

Chapter 5

Hamiltonian formulation of asymptotically flat spacetimes

The goal of this chapter is to apply the Hamiltonian framework developed in Chapter 4 to general relativity, with particular emphasis on the asymptotically flat case. In this setting, the Hamiltonian is related to the ADM energy-momentum. We do not give a systematic account of general relativity here. For a more detailed treatment, we refer to the standard textbooks [14, 19, 50, 79, 82].

The general construction from Chapter 4 already provides the abstract Hamiltonian framework. To apply it to general relativity, however, we must specify the geometric data to which the framework is applied. This means identifying the field bundle F, choosing the gravitational field variables, constructing the corresponding momenta, and specifying the Lagrangian. We must also describe the reference systems appropriate for these variables. Once this is done, the general Hamiltonian formalism is applied to general relativity on arbitrary spacetimes.

The asymptotically flat case requires additional work. In the general framework, the phase space and the reference system were described abstractly. For asymptotically flat spacetimes, we must specify the asymptotic conditions, define a suitable phase space, and choose appropriate reference systems. We must also check the convergence of the relevant integrals, and the boundary terms that should vanish must indeed vanish. These checks are what allow the Hamiltonian construction to define a Hamiltonian dynamical system in the asymptotically flat setting.

We begin in Section 5.1 by formulating vacuum general relativity as a Lagrangian field theory. Following Chruściel, Jezierski, and Kijowski [27], we use the contravariant metric density as the field variable. We introduce a first-order Lagrangian obtained from the Hilbert Lagrangian by subtracting a total divergence. This formulation naturally determines the field bundle used in the Hamiltonian construction. After introducing this Lagrangian field theory, we construct the appropriate reference system. This allows us to apply the construction from Chapter 4 to general relativity.

In Section 5.2, we reduce the resulting spacetime formulas to a spacelike hypersurface and compare them with the usual ADM variables. Section 5.3 then specializes this construction to asymptotically flat spacetimes. We introduce the relevant asymptotic conditions and use the reduced formulas from Section 5.2 to prove convergence and vanishing of the relevant integrals. We also discuss the phase space on which the Hamiltonian formulation is defined, and study the Hamiltonian terms that arise in this setting. These Hamiltonian terms are then related to the ADM energy-momentum.

Finally, in Section 5.4, we discuss the invariance of the ADM mass on a spacelike hypersurface in an asymptotically flat manifold. We also examine its invariance under changes of spacetime foliation. For the invariance on a fixed hypersurface, we use the work of Bartnik [7]. For the spacetime invariance, we follow Chruściel's work on the invariant mass conjecture [22].

The exposition follows Chruściel, Jezierski, and Kijowski [27, Chapter 5], supplemented by the work of Bartnik [7] and Chruściel [22] on the invariance of the ADM mass.

5.1 General relativity as a Hamiltonian field theory

In this section we apply the construction from Chapter 4 to general relativity. We consider vacuum general relativity on a four-dimensional spacetime manifold M. The field configurations are Lorentzian metrics g on M with signature $(-,+,+,+)$. In the vacuum case, the corresponding field equations are the vacuum Einstein equations

$$R_{\mu\nu}(g) = 0. \tag{5.1}$$

For details on variational principles that generate the Einstein equations, see Wald or Landsman [50, 82]. The aim of this section is to describe these equations within the Hamiltonian framework developed earlier.

Subsection 5.1.1 introduces the field bundle used to describe gravitational fields. Instead of working directly with the metric tensor, we use the contravariant metric density as the field variable. Subsection 5.1.2 introduces the momentum tensor and constructs the first-order Lagrangian used in the field theory. This subsection is rather technical, and most of the proofs consist of direct coordinate calculations with Christoffel symbols and metric densities. For this reason, the main text only states the required identities; the details are given in Appendix A. The purpose of this subsection is to show that the first-order Lagrangian is equivalent to the Hilbert Lagrangian up to a total divergence. Moreover, we show that the Lagrangian identifies the momentum tensor as the canonical momentum conjugate to the metric density.

Finally, Subsection 5.1.3 introduces the appropriate reference system and outlines the transition from the Lagrangian formulation to the Hamiltonian formulation. Using the field bundle, the first-order Lagrangian, and the reference system, we construct the Hamiltonian associated with general relativity. The resulting expression is written in terms of the Freud superpotential [36].

Throughout this section we follow the background-connection approach of Chruściel, Jezierski, and Kijowski, as developed in [27, Chapter 5].

5.1.1 Metric density and the gravitational field bundle

Throughout this section, M denotes a four-dimensional spacetime.

We begin by identifying the field variable and the corresponding configuration bundle. Instead of working directly with a Lorentzian metric g on M, we use the contravariant metric density

$$\mathfrak{g}^{\mu\nu} := \frac{1}{16\pi}\sqrt{-\det g}\, g^{\mu\nu}. \tag{5.2}$$

For Lorentzian metrics, describing g is equivalent to describing $\mathfrak{g}^{\mu\nu}$. We will therefore regard $\mathfrak{g}^{\mu\nu}$ as the field variable describing the gravitational field on M.

In Chapter 4, fields were modelled as sections of a fibre bundle over the base manifold. We now identify the corresponding bundle for the gravitational field.

Lemma 5.1 *The object $\mathfrak{g}^{\mu\nu}$ is a symmetric contravariant tensor density. Hence, it defines a section of*

$$S_0T^2M \otimes \tilde{\Lambda}^4 T^*M,$$

where S_0T^2M denotes the bundle of non-degenerate symmetric contravariant tensors over M.

Proof. We show that $\mathfrak{g}^{\mu\nu}$ is a symmetric tensor density by showing it satisfies the correct transformation laws.

Let (x^μ) and $(\tilde{x}^\sigma)$ be two coordinate systems on an overlapping domain. Since

$$dx^\mu = \frac{\partial x^\mu}{\partial \tilde{x}^\sigma} d\tilde{x}^\sigma,$$

the metric components transform as

$$\tilde{g}_{\sigma\alpha} = \frac{\partial x^\mu}{\partial \tilde{x}^\sigma} \frac{\partial x^\nu}{\partial \tilde{x}^\alpha} g_{\mu\nu}.$$

Hence

$$\det(\tilde{g}) = \det\left(\frac{\partial x}{\partial \tilde{x}}\right)^2 \det(g),$$

where $\frac{\partial x}{\partial \tilde{x}}$ denotes the Jacobian matrix of the coordinate transform. Therefore, we obtain

$$\sqrt{-\det(\tilde{g})} = \left|\det\left(\frac{\partial x}{\partial \tilde{x}}\right)\right| \sqrt{-\det(g)}.$$

Thus $\sqrt{-\det(g)}$ transforms as a density.

The inverse metric transforms contravariantly

$$\tilde{g}^{\sigma\alpha} = \frac{\partial \tilde{x}^\sigma}{\partial x^\mu} \frac{\partial \tilde{x}^\alpha}{\partial x^\nu} g^{\mu\nu}.$$

Combining this with the transformation law for $\sqrt{-\det(g)}$, we obtain

$$\tilde{\mathfrak{g}}^{\sigma\alpha} = \left|\det\left(\frac{\partial x}{\partial \tilde{x}}\right)\right| \frac{\partial \tilde{x}^\sigma}{\partial x^\mu} \frac{\partial \tilde{x}^\alpha}{\partial x^\nu} \mathfrak{g}^{\mu\nu}.$$

This is precisely the transformation law for a symmetric tensor density. Since g is non-degenerate and Lorentzian, $\mathfrak{g}^{\mu\nu}$ defines a section of

$$S_0T^2M \otimes \tilde{\Lambda}^4 T^*M. \qquad \square$$

We conclude that gravitational fields can be viewed as sections of the tensor bundle

$$\begin{array}{c} F := S_0T^2M \otimes \tilde{\Lambda}^4 T^*M \\ \Big\downarrow \mathrm{pr}_{F\to M} \\ M \end{array}$$

Hence F is the natural choice for the configuration bundle over M in the gravitational field theory.

5.1.2 The gravitational Lagrangian

The goal of this subsection is to construct the first-order Lagrangian used for the Hamiltonian formulation of general relativity. We identified the gravitational field with the metric density $\mathfrak{g}^{\mu\nu}$, viewed as a section of

$$F = S_0 T^2 M \otimes \tilde{\Lambda}^4 T^* M.$$

We now describe the corresponding Lagrangian density. Following the background-connection formalism from Chruściel, Jezierski, and Kijowski [27], we replace the Hilbert Lagrangian density by a first-order Lagrangian that differs by a total divergence from the Hilbert Lagrangian. Both Lagrangians therefore give the same Euler–Lagrange equations.

The construction proceeds as follows. We introduce a symmetric connection B, called the background connection, and compare it with the Levi–Civita connection Γ of the metric g. From the difference between these two connections we define a tensor $p^{\alpha}_{\mu\nu}$. Once the background connection B is fixed, this tensor contains the same information as the Levi–Civita connection Γ. It is then used to define the first-order Lagrangian density L. Finally, we show that the canonical momentum associated with the field variable $\mathfrak{g}^{\mu\nu}$ is precisely $p^{\lambda}_{\mu\nu}$.

This subsection is technical, most of the intermediate identities are direct coordinate calculations involving Christoffel symbols, metric densities, and the background connection. To keep the main text readable, we state the required identities here and give the detailed calculations in Appendix A. When an identity is used only as a standard computational tool, we briefly explain it. For identities that are central to the construction, we include a proof sketch in the main text. Readers interested only in the final outcome of the construction may skip to Theorem 5.12.

We first recall the coordinate expression for the Ricci tensor. It is obtained by contracting the Riemann curvature tensor, and in local coordinates it is given by

$$R_{\mu\nu} = \partial_\alpha \Gamma^{\alpha}_{\mu\nu} - \partial_\mu \Gamma^{\alpha}_{\alpha\nu} - \left[\Gamma^{\sigma}_{\alpha\nu}\Gamma^{\alpha}_{\mu\sigma} - \Gamma^{\sigma}_{\mu\nu}\Gamma^{\alpha}_{\alpha\sigma}\right], \tag{5.3}$$

where $\Gamma^{\alpha}_{\mu\nu}$ denotes the Christoffel symbols of the Levi–Civita connection of g.

Using the Ricci tensor, the Hilbert Lagrangian density is locally written as

$$\tilde{L} := \frac{1}{16\pi}\sqrt{-\det g}R = \frac{1}{16\pi}\sqrt{-\det g}g^{\mu\nu}R_{\mu\nu} = \mathfrak{g}^{\mu\nu}R_{\mu\nu}. \tag{5.4}$$

To manipulate this density, we now rewrite it in terms of Christoffel symbols.

Lemma 5.2 *For a metric density $\mathfrak{g}^{\mu\nu}$ one can express the Hilbert Lagrangian as follows*

$$\mathfrak{g}^{\mu\nu}R_{\mu\nu} = \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^{\alpha}_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^{\alpha}_{\alpha\nu}) + \mathfrak{g}^{\mu\nu}\left[\Gamma^{\alpha}_{\sigma\mu}\Gamma^{\sigma}_{\alpha\nu} - \Gamma^{\sigma}_{\alpha\sigma}\Gamma^{\alpha}_{\mu\nu}\right]. \tag{5.5}$$

The proof of Lemma 5.2 uses the following coordinate identity for the derivative of the metric density

$$\partial_\alpha \mathfrak{g}^{\mu\nu} = \mathfrak{g}^{\mu\nu}\Gamma^{\sigma}_{\alpha\sigma} - \mathfrak{g}^{\sigma\nu}\Gamma^{\mu}_{\alpha\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^{\nu}_{\sigma\alpha}. \tag{5.6}$$

This identity follows from metric compatibility of the Levi–Civita connection, together with the standard formula for the derivative of the determinant of the metric. The full coordinate calculation is given in Lemma A.1.

We now introduce the background connection. Let B be a symmetric connection on M, and denote its Ricci tensor by $r_{\mu\nu}$. In local coordinates,

$$r_{\mu\nu} = \partial_\alpha B^{\alpha}_{\mu\nu} - \partial_\mu B^{\alpha}_{\alpha\nu} - \left[B^{\sigma}_{\alpha\nu}B^{\alpha}_{\mu\sigma} - B^{\sigma}_{\mu\nu}B^{\alpha}_{\alpha\sigma}\right].$$

Remark 5.3 The construction used here resembles the construction of Chruściel [20], where one assumes the existence of a background metric $b_{\mu\nu}$ and takes the background connection B to be the Levi–Civita connection of b. We follow Chruściel, Jezierski, and Kijowski [27], who generalize this approach somewhat by requiring only that the background connection be symmetric. In particular, the background connection need not arise from a metric.

Using equation (5.6), we obtain the contraction $\mathfrak{g}^{\mu\nu} r_{\mu\nu}$.

Lemma 5.4 *Let B be a symmetric background connection on M, and let $r_{\mu\nu}$ be its Ricci tensor. The contraction is locally written as*

$$\begin{aligned}\mathfrak{g}^{\mu\nu} r_{\mu\nu} &= \partial_\alpha(\mathfrak{g}^{\mu\nu} B^\alpha_{\mu\nu}) - \partial_\mu(\mathfrak{g}^{\mu\nu} B^\alpha_{\alpha\nu}) - \mathfrak{g}^{\mu\nu}\left[B^\sigma_{\alpha\nu} B^\alpha_{\mu\sigma} - B^\sigma_{\mu\nu} B^\alpha_{\alpha\sigma}\right] \\ &\quad + \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\nu\alpha} B^\alpha_{\mu\sigma} + \Gamma^\sigma_{\alpha\mu} B^\alpha_{\sigma\nu} - \Gamma^\sigma_{\nu\mu} B^\alpha_{\alpha\sigma} - \Gamma^\sigma_{\alpha\sigma} B^\alpha_{\mu\nu}\right]\end{aligned} \tag{5.7}$$

Proof sketch. Multiply the coordinate expression for $r_{\mu\nu}$ by $\mathfrak{g}^{\mu\nu}$ and apply the product rule to the terms $\mathfrak{g}^{\mu\nu}\partial_\alpha B^\alpha_{\mu\nu}$ and $\mathfrak{g}^{\mu\nu}\partial_\mu B^\alpha_{\alpha\nu}$. The derivatives of $\mathfrak{g}^{\mu\nu}$ are then eliminated using (5.6). After relabelling dummy indices, one obtains (5.7). The full coordinate calculation is given in Lemma A.2. □

Before introducing the tensor $p^\lambda_{\mu\nu}$, we fix some notation. For a tensor with two indices, we define symmetrization by

$$T_{(\mu\nu)} := \frac{1}{2}\left(T_{\mu\nu} + T_{\nu\mu}\right).$$

More generally, symmetrization over n indices is defined by

$$T_{(t_1 \dots t_n)} := \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} T_{t_{\sigma(1)} \dots t_{\sigma(n)}},$$

where $\mathfrak{S}_n$ is the symmetric group of degree n.

Lemma 5.5 *The Ricci tensor is expressed using symmetrization, as follows*

$$R_{\mu\nu} = \partial_\alpha\left[\Gamma^\alpha_{\mu\nu} - \delta^\alpha_{(\mu}\Gamma^\kappa_{\nu)\kappa}\right] - \left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right].$$

This is only a rewriting of the coordinate expression for the Ricci tensor using the symmetrization convention introduced above. The full calculation is given in Lemma A.3.

We now introduce the tensor $p^\lambda_{\mu\nu}$. Let B be the fixed symmetric background connection introduced above, and let Γ be the Levi–Civita connection of g. We define

$$p^\alpha_{\mu\nu} := \left(B^\alpha_{\mu\nu} - \delta^\alpha_{(\mu} B^\kappa_{\nu)\kappa}\right) - \left(\Gamma^\alpha_{\mu\nu} - \delta^\alpha_{(\mu}\Gamma^\kappa_{\nu)\kappa}\right). \tag{5.8}$$

Although this expression is written in terms of connection coefficients, $p^\alpha_{\mu\nu}$ is a tensor, since it is obtained from the difference between the two symmetric connections B and Γ. This tensor is central to the construction, it will be used to define the first-order Lagrangian and will later be identified with the momentum canonically conjugate to $\mathfrak{g}^{\mu\nu}$.

Once the background connection B is fixed, the tensor $p^\alpha_{\mu\nu}$ contains the same information as the Levi–Civita connection Γ. This is made explicit by the following lemma.

Lemma 5.6 *The Christoffel symbols of g are recovered from the background connection $B^\alpha_{\mu\nu}$ and the tensor $p^\alpha_{\mu\nu}$ by*

$$\Gamma^\alpha_{\mu\nu} = B^\alpha_{\mu\nu} - p^\alpha_{\mu\nu} + \frac{2}{3}\delta^\alpha_{(\mu}p^\kappa_{\nu)\kappa}. \tag{5.9}$$

Proof sketch. From the definition of $p^\alpha_{\mu\nu}$ we solve for $\Gamma^\alpha_{\mu\nu}$ as

$$\Gamma^\alpha_{\mu\nu} = B^\alpha_{\mu\nu} - p^\alpha_{\mu\nu} - \delta^\alpha_{(\mu}B^\kappa_{\nu)\kappa} + \delta^\alpha_{(\mu}\Gamma^\kappa_{\ \nu)\kappa}.$$

It remains to rewrite the last two terms. Contracting the definition of $p^\alpha_{\mu\nu}$ over α and ν gives, using $\dim M = 4$,

$$p^\kappa_{\mu\kappa} = -\frac{3}{2}B^\kappa_{\mu\kappa} + \frac{3}{2}\Gamma^\kappa_{\ \mu\kappa}.$$

Hence

$$\frac{2}{3}\delta^\alpha_{(\mu}p^\kappa_{\nu)\kappa} = -\delta^\alpha_{(\mu}B^\kappa_{\nu)\kappa} + \delta^\alpha_{(\mu}\Gamma^\kappa_{\ \nu)\kappa}.$$

Substituting this into the expression for $\Gamma^\alpha_{\mu\nu}$ gives

$$\Gamma^\alpha_{\mu\nu} = B^\alpha_{\mu\nu} - p^\alpha_{\mu\nu} + \frac{2}{3}\delta^\alpha_{(\mu}p^\kappa_{\nu)\kappa}.$$

The full calculation is given in Lemma A.4. □

We now introduce the density that will be the Lagrangian of the field theory

$$L := \mathfrak{g}^{\mu\nu}\big[(\Gamma^\alpha_{\sigma\mu} - B^\alpha_{\mu\sigma})(\Gamma^\sigma_{\nu\alpha} - B^\sigma_{\alpha\nu}) - (\Gamma^\alpha_{\mu\nu} - B^\alpha_{\mu\nu})(\Gamma^\sigma_{\alpha\sigma} - B^\sigma_{\alpha\sigma}) + r_{\mu\nu}\big]. \tag{5.10}$$

The following proposition shows that this density differs from the Hilbert Lagrangian density only by a divergence.

Proposition 5.7 *The Hilbert Lagrangian density is expressed in terms of p and L as*

$$\mathfrak{g}^{\mu\nu}R_{\mu\nu} = -\partial_\alpha(\mathfrak{g}^{\mu\nu}p^\alpha_{\mu\nu}) + L, \tag{5.11}$$

where L is defined by (5.10).

Proof sketch. The identity is obtained by subtracting the contraction (5.7) from the expression (5.5) for the Hilbert Lagrangian density. The derivative terms combine as

$$\mathfrak{g}^{\mu\nu}R_{\mu\nu} - \mathfrak{g}^{\mu\nu}r_{\mu\nu} = -\partial_\alpha(\mathfrak{g}^{\mu\nu}p^\alpha_{\mu\nu}) + K,$$

where we collect all the remaining terms in K. To complete the proof we need to show $\mathfrak{g}^{\mu\nu}r_{\mu\nu} + K = L$. This calculation is given in Proposition A.5. □

The Lagrangian L and the Hilbert Lagrangian $\tilde{L}$ differ by a total divergence. By Anderson [1], total divergences are variationally trivial for the Euler–Lagrange equations. Hence, L and $\tilde{L}$ define the same equations of motion.

Remark 5.8 This is enough for the present purpose, because the goal is to replace the Hilbert Lagrangian by a first-order Lagrangian suitable for the Hamiltonian construction.

The corresponding variational problems are more delicate. To obtain the same variational problem, one has to impose boundary conditions that make the resulting boundary variation vanish. Indeed, using Proposition 5.7, the Einstein–Hilbert action is written as

$$\begin{aligned}S_{\tilde{L}}(g) &= \int_M \mathfrak{g}^{\mu\nu} R_{\mu\nu}\, d^4x \\ &= -\int_M \partial_\alpha \left(\mathfrak{g}^{\mu\nu} p^\alpha_{\mu\nu}\right) d^4x + \int_M L\, d^4x \\ &= -\int_{\partial M} \mathfrak{g}^{\mu\nu} p^\alpha_{\mu\nu} (\partial_\alpha \rfloor d^4x) + S_L(g).\end{aligned}$$

Thus the two actions differ by the boundary integral

$$-\int_{\partial M} \mathfrak{g}^{\mu\nu} p^\alpha_{\mu\nu} (\partial_\alpha \rfloor d^4x).$$

Consequently, their first variations differ by the variation of this boundary integral

$$\delta S_{\tilde{L}}(g) - \delta S_L(g) = -\int_{\partial M} \delta\left(\mathfrak{g}^{\mu\nu} p^\alpha_{\mu\nu}\right)(\partial_\alpha \rfloor d^4x).$$

Therefore, although the total divergence does not change the Euler–Lagrange equations, the two variational problems are identical only after imposing boundary conditions that make this boundary variation vanish. We do not analyze these boundary conditions here. For compactly supported variations, see Landsman [50, Section 7.2]. For boundary conditions of a non-compact case, see Chruściel–Jezierski–Kijowski [27, Section 2.2].

This replacement is useful because L can be expressed in terms of the first derivatives of $\mathfrak{g}^{\mu\nu}$. We state this explicitly in Proposition 5.11. Before proving this proposition, we introduce two auxiliary lemmas.

Lemma 5.9 *Let ∇^B denote the covariant derivative associated to the background connection B. Then*

$$\mathfrak{g}^{\mu\nu}{}_{;\alpha} = \mathfrak{g}^{\mu\nu}(\Gamma^\sigma_{\alpha\sigma} - B^\sigma_{\alpha\sigma}) - \mathfrak{g}^{\mu\sigma}(\Gamma^\nu_{\sigma\alpha} - B^\nu_{\sigma\alpha}) - \mathfrak{g}^{\nu\sigma}(\Gamma^\mu_{\sigma\alpha} - B^\mu_{\sigma\alpha}), \tag{5.12}$$

where $\mathfrak{g}^{\mu\nu}{}_{;\alpha} := \nabla^B_\alpha \mathfrak{g}^{\mu\nu}$.

This follows by applying the covariant derivative to a contravariant tensor density and then using (5.6). The full coordinate calculation is given in Lemma A.6.

To write L in terms of $\mathfrak{g}^{\mu\nu}$, we first express $p^\alpha_{\mu\nu}$ in terms of the background-covariant derivative of $\mathfrak{g}^{\mu\nu}$.

Lemma 5.10 *The tensor $p^\alpha_{\mu\nu}$ can be written as*

$$p^\lambda_{\mu\nu} = \frac{1}{2}\mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\nu} + \frac{1}{2}\mathfrak{g}_{\nu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\mu} - \frac{1}{2}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\sigma\mu}\mathfrak{g}_{\rho\nu}\mathfrak{g}^{\sigma\rho}{}_{;\alpha} + \frac{1}{4}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\mu\nu}\mathfrak{g}_{\sigma\rho}\mathfrak{g}^{\sigma\rho}{}_{;\alpha}. \tag{5.13}$$

Here $\mathfrak{g}_{\mu\nu}$ denotes the inverse of $\mathfrak{g}^{\mu\nu}$, so that $\mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\alpha\nu} = \delta^\nu_\mu$.

Proof sketch. Lemma 5.9 expresses the background covariant derivative $\mathfrak{g}^{\mu\nu}{}_{;\alpha}$ in terms of the difference between the Levi–Civita connection Γ and the background connection B. Computing the contractions and using the fact that $\dim M = 4$ gives the formula above. The full calculation is given in Lemma A.7. □

Using the expression for $p^{\lambda}_{\mu\nu}$ obtained in Lemma 5.10, we rewrite the Lagrangian L entirely in terms of the background covariant derivatives of the metric density $\mathfrak{g}^{\mu\nu}$.

Proposition 5.11 *The Lagrangian L is written in terms of the background covariant derivatives of $\mathfrak{g}^{\mu\nu}$ as*

$$L = \frac{1}{2}\mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\mu\nu}{}_{;\lambda}\mathfrak{g}^{\lambda\alpha}{}_{;\nu} - \frac{1}{4}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\sigma\mu}\mathfrak{g}_{\rho\nu}\mathfrak{g}^{\mu\nu}{}_{;\lambda}\mathfrak{g}^{\sigma\rho}{}_{;\alpha} \tag{5.14}$$

$$+ \frac{1}{8}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\mu\nu}\mathfrak{g}^{\mu\nu}{}_{;\lambda}\mathfrak{g}_{\sigma\rho}\mathfrak{g}^{\sigma\rho}{}_{;\alpha} + \mathfrak{g}^{\mu\nu}r_{\mu\nu}. \tag{5.15}$$

Proof. Starting from the definition of the Lagrangian

$$L = \mathfrak{g}^{\mu\nu}\big[(\Gamma^{\alpha}_{\sigma\mu} - B^{\alpha}_{\mu\sigma})(\Gamma^{\sigma}_{\nu\alpha} - B^{\sigma}_{\alpha\nu}) - (\Gamma^{\alpha}_{\mu\nu} - B^{\alpha}_{\mu\nu})(\Gamma^{\sigma}_{\alpha\sigma} - B^{\sigma}_{\alpha\sigma})\big] + \mathfrak{g}^{\mu\nu}r_{\mu\nu}, \tag{5.16}$$

we rewrite the differences $\Gamma - B$ in terms of $p^{\lambda}_{\mu\nu}$ using Lemma 5.6

$$\Gamma^{\alpha}_{\mu\nu} - B^{\alpha}_{\mu\nu} = -p^{\alpha}_{\mu\nu} + \frac{2}{3}\delta^{\alpha}_{(\mu}p^{\kappa}_{\nu)\kappa}. \tag{5.17}$$

We then substitute the expression for $p^{\lambda}_{\mu\nu}$, from Lemma 5.10, into (5.17). We expand the products in (5.16) and collect the resulting quadratic factors. Rewriting these factors in terms of $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$ gives the stated expression. The calculation is a direct substitution and involves no additional geometric idea. □

The preceding lemmas and proposition allow us to identify the canonical momenta associated with the Lagrangian L.

Theorem 5.12 *Let (M, g) be a $(3+1)$-dimensional spacetime equipped with a connection Γ, and let B be a symmetric background connection on M. Define the metric density by*

$$\mathfrak{g}^{\mu\nu} = \frac{1}{16\pi}\sqrt{-\det g}\, g^{\mu\nu},$$

and define the tensor $p^{\alpha}_{\mu\nu}$ by

$$p^{\alpha}_{\mu\nu} := \Big(B^{\alpha}_{\mu\nu} - \delta^{\alpha}_{(\mu}B^{\kappa}_{\nu)\kappa}\Big) - \Big(\Gamma^{\alpha}_{\mu\nu} - \delta^{\alpha}_{(\mu}\Gamma^{\kappa}{}_{\nu)\kappa}\Big).$$

Let the gravitational Lagrangian be

$$L := \mathfrak{g}^{\mu\nu}\big[(\Gamma^{\alpha}_{\sigma\mu} - B^{\alpha}_{\mu\sigma})(\Gamma^{\sigma}_{\nu\alpha} - B^{\sigma}_{\alpha\nu}) - (\Gamma^{\alpha}_{\mu\nu} - B^{\alpha}_{\mu\nu})(\Gamma^{\sigma}_{\alpha\sigma} - B^{\sigma}_{\alpha\sigma}) + r_{\mu\nu}\big].$$

Then $p^{\lambda}_{\mu\nu}$ is the canonical momentum associated with the metric density $\mathfrak{g}^{\mu\nu}$, that is,

$$\frac{\partial L}{\partial \mathfrak{g}^{\mu\nu}{}_{,\lambda}} = \frac{\partial L}{\partial \mathfrak{g}^{\mu\nu}{}_{;\lambda}} = p^{\lambda}_{\mu\nu}. \tag{5.18}$$

Here, $\mathfrak{g}^{\mu\nu}{}_{,\lambda}$ denotes the partial derivative of the metric density, whereas $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$ denotes its covariant derivative with respect to the background connection B.

Proof sketch. By definition of the background covariant derivative ∇^B, we have

$$\mathfrak{g}^{\mu\nu}{}_{;\lambda} = \mathfrak{g}^{\mu\nu}{}_{,\lambda} + B^{\mu}_{\lambda\sigma}\mathfrak{g}^{\sigma\nu} + B^{\nu}_{\lambda\sigma}\mathfrak{g}^{\mu\sigma} - B^{\sigma}_{\lambda\sigma}\mathfrak{g}^{\mu\nu}.$$

Thus $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$ differs from $\mathfrak{g}^{\mu\nu}{}_{,\lambda}$ only by terms depending on $\mathfrak{g}^{\mu\nu}$ and the background connection B. Therefore, differentiating L with respect to $\mathfrak{g}^{\mu\nu}{}_{,\lambda}$ is equivalent to differentiating it with respect to $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$.

We now differentiate the expression for L obtained in Proposition 5.11. The term $\mathfrak{g}^{\mu\nu}r_{\mu\nu}$ does not depend on $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$, so it does not contribute. The remaining terms give precisely $p^{\lambda}_{\mu\nu}$. For a full calculation, see Theorem A.8. □

5.1.3 Hamiltonian dynamical system for general relativity

Subsection 5.1.2 identified the gravitational Lagrangian L with the abstract Lagrangian framework developed in Chapter 4. In this subsection, we construct a reference system and apply the Hamiltonian framework from Chapter 4 to general relativity.

In Subsection 4.4.4, we described the Hamiltonian associated with a reference system on a tensor bundle. Since the gravitational field is represented by the metric density $\mathfrak{g}^{\mu\nu}$, viewed as a section of the tensor bundle

$$F = S_0T^2M \otimes \tilde{\Lambda}^4T^*M,$$

we apply that construction to the gravitational variables $\mathfrak{g}^{\mu\nu}$ and $p^{\lambda}_{\mu\nu}$.

The goal of this subsection is to specialize the general Hamiltonian formula from Chapter 4 to the gravitational Lagrangian L. The resulting Hamiltonian is written compactly using the Freud superpotential [36]. We will use this superpotential only as a computational tool. We state the formula needed below, but do not derive it.

Let $\Sigma \subset M$ be a hypersurface, let $X \in \mathscr{X}(M)$ be a global vector field on M, and let $\psi : I \times \Sigma \to M$ be its flow. Since

$$F = S_0T^2M \otimes \tilde{\Lambda}^4T^*M$$

is a tensor bundle, the flow ψ_t induces a flow Ψ_t on F. Thus, as in Subsection 4.4.4, we obtain a reference system

$$\begin{tikzcd} I \times \Phi \arrow[r, "\Psi"] \arrow[d] & F \arrow[d] \\ I \times \Sigma \arrow[r, "\psi"'] & M \end{tikzcd}$$

where $\Phi = (\psi_0)^*F$.

We now specialize the general Hamiltonian construction to the gravitational variables. The field variable is

$$f^A = \mathfrak{g}^{\mu\nu},$$

and, by Theorem 5.12, the corresponding canonical momentum is

$$p_A{}^{\lambda} = p^{\lambda}_{\mu\nu}.$$

The Lagrangian density is the gravitational Lagrangian L obtained in Proposition 5.11.

Recall that the spacetime momentum bundle is

$$P := V^*F \otimes \tilde{\Lambda}^4T^*M.$$

In the present case, a section of P is locally described by the variables $(\mathfrak{g}^{\mu\nu}, p^\lambda_{\mu\nu})$, and is written as

$$p = p^\lambda_{\mu\nu} d\mathfrak{g}^{\mu\nu} \otimes \left(\partial_\lambda \rfloor dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3\right).$$

As discussed in Definition 4.17, the phase space on the hypersurface Σ is obtained by pulling back dynamically admissible sections to that hypersurface. A dynamically admissible gravitational field is described by a metric density $\mathfrak{g}^{\mu\nu}$ together with its canonical momentum $p^\lambda_{\mu\nu}$ that satisfy the Euler–Lagrange equation associated to L.

Remark 5.13 (Convention one-parameter families and first variations.) The use of one-parameter families is implicit throughout this chapter. Recall from Subsection 4.4.2 that, in order to establish the Hamiltonian structure of the phase space $\mathscr{P}_\Sigma$, we introduced a one-parameter family of dynamical variables

$$\left(f^A(s), p^\mu_A(s)\right).$$

The dependence on the parameter s was subsequently suppressed to improve readability.

The same convention is used in the present chapter. Thus, the gravitational variables

$$\left(\mathfrak{g}^{\alpha\beta}(s), p^\mu_{\alpha\beta}(s)\right)$$

and the metric g are always understood to belong to a differentiable one-parameter family, even when the parameter s is not displayed explicitly. Every quantity constructed from these variables is likewise understood to depend on s.

For any such one-parameter family $q(s)$, its first variation is denoted by

$$\delta q := \left.\frac{d}{ds}\right|_{s=0} q(s).$$

Accordingly, every occurrence of δ in this chapter refers to differentiation along the underlying one-parameter family at $s = 0$.

Using the construction outlined at the end of Subsection 4.4.1, we transport the antisymmetric two-form on the phase space of Σ to the hypersurface $\mathscr{S} \subset M$. Substituting the dynamical variables $(p^\lambda_{\mu\nu}, \mathfrak{g}^{\mu\nu})$ into (4.44), we obtain

$$\Omega_{\mathscr{S}}((\delta_1 p^\lambda_{\mu\nu}, \delta_1 \mathfrak{g}^{\alpha\beta}), (\delta_2 p^\lambda_{\mu\nu}, \delta_2 \mathfrak{g}^{\alpha\beta})) = \int_{\mathscr{S}} (\delta_1 p^\mu_{\alpha\beta} \delta_2 \mathfrak{g}^{\alpha\beta} - \delta_2 p^\mu_{\alpha\beta} \delta_1 \mathfrak{g}^{\alpha\beta}) dS_\mu. \tag{5.19}$$

Here the odd form dS_μ is defined by

$$\frac{\partial}{\partial x^\mu} \rfloor dx^0 \wedge \cdots \wedge dx^n,$$

and we have abbreviated[1] the vector fields

$$X_a = \delta_a p^\lambda_{\mu\nu} \frac{\partial}{\partial p^\lambda_{\mu\nu}} + \delta_a \mathfrak{g}^{\alpha\beta} \frac{\partial}{\partial \mathfrak{g}^{\alpha\beta}}$$

by their components $(\delta_a p^\lambda_{\mu\nu}, \delta_a \mathfrak{g}^{\alpha\beta})$.

[1]We abbreviate vector fields by their components using δ_a. We use this convention throughout the chapter.

Substituting the gravitational field variable $f^A = \mathfrak{g}^{\alpha\beta}$ and its canonical momentum $p_A{}^\mu = p^\mu_{\alpha\beta}$ into the Hamiltonian formula (4.55), we obtain

$$H(X, \mathscr{S}, p) = \int_{\mathscr{S}} (p^\mu_{\alpha\beta} \pounds_X \mathfrak{g}^{\alpha\beta} - X^\mu L) dS_\mu, \tag{5.20}$$

where $p = (\mathfrak{g}^{\mu\nu}, p^\lambda_{\alpha\beta})$. To shorten notation, we will suppress the dependence on p and write $H(X, \mathscr{S})$.

The corresponding variation formula is obtained from (4.57) by substituting the gravitational field variables. This gives

$$\begin{aligned} -\delta H = &\int_{\mathscr{S}} \left(\pounds_X p^\lambda_{\mu\nu} \delta \mathfrak{g}^{\mu\nu} - \pounds_X \mathfrak{g}^{\mu\nu} \delta p^\lambda_{\mu\nu} \right) dS_\lambda \\ &+ \int_{\partial\mathscr{S}} X^{[\mu} p^{\nu]}_{\alpha\beta} \delta \mathfrak{g}^{\alpha\beta} dS_{\mu\nu}, \end{aligned} \tag{5.21}$$

where $dS_{\mu\nu}$ is defined as

$$dS_{\mu\nu} = (\partial_\mu \wedge \partial_\nu) \rfloor dx^0 \wedge \cdots \wedge dx^3.$$

Following Chruściel [20], the integrand of the Hamiltonian is decomposed into a divergence term involving a Freud-type superpotential[2] plus an additional term. In the notation used here, this identity is

$$p^\mu_{\alpha\beta} \pounds_X \mathfrak{g}^{\alpha\beta} - X^\mu L = \partial_\alpha \mathbb{W}^{\mu\alpha} - 2\mathfrak{g}^{\beta[\gamma} \delta^{\mu]}_\sigma (X^\sigma{}_{;\beta\gamma} - B^\sigma{}_{\beta\gamma\kappa} X^\kappa), \tag{5.22}$$

$$\mathbb{W}^{\nu\lambda} = \left(2\mathfrak{g}^{\mu[\nu} p^{\lambda]}_{\mu\beta} - 2\delta^{[\nu}_\beta p^{\lambda]\mu}_{\mu\sigma} \mathfrak{g}^{\mu\sigma} - \frac{2}{3} \mathfrak{g}^{\mu[\nu} \delta^{\lambda]}_\beta p^\sigma_{\sigma\mu} \right) X^\beta - 2\mathfrak{g}^{\alpha[\nu} \delta^{\lambda]}_\beta X^\beta{}_{,\alpha}. \tag{5.23}$$

Here a semicolon denotes the covariant derivative with respect to the background connection B. Furthermore, $B^\sigma{}_{\beta\gamma\kappa}$ is the curvature tensor of the connection $B^\sigma_{\beta\gamma}$ and the square brackets denote the antisymmetrization with a factor $\frac{1}{2}$ when two indices are involved.

We rewrite the Hamiltonian in a form that does not contain the background connection. For this purpose, define

$$A^\alpha{}_{\mu\nu} := \Gamma^\alpha{}_{\mu\nu} - \delta^\alpha_{(\mu} \Gamma^\kappa{}_{\nu)\kappa}, \tag{5.24}$$

and, similarly, for the background connection,

$$\mathring{A}^\alpha{}_{\mu\nu} := B^\alpha{}_{\mu\nu} - \delta^\alpha_{(\mu} B^\kappa{}_{\nu)\kappa}. \tag{5.25}$$

By the definition of the canonical momentum, we have

$$p^\alpha{}_{\mu\nu} = \mathring{A}^\alpha{}_{\mu\nu} - A^\alpha{}_{\mu\nu}. \tag{5.26}$$

We use the following formula for the Lie derivative of a connection

$$\pounds_X B^\lambda{}_{\mu\nu} = X^\lambda{}_{;\mu\nu} - B^\lambda{}_{\mu\nu\sigma} X^\sigma.$$

---

[2] We use (5.22) as a quoted identity. For more details on the Freud superpotential (5.23), see Freud [36].

Here the semicolon denotes covariant differentiation with respect to the background connection B. A direct calculation shows that

$$\mathfrak{g}^{\mu\nu}\pounds_X \mathring{A}^{\lambda}{}_{\mu\nu} = 2\mathfrak{g}^{\beta[\nu}\delta^{\lambda]}_{\sigma}\left(X^{\sigma}{}_{;\beta\gamma} - B^{\sigma}{}_{\beta\gamma\kappa}X^{\kappa}\right). \tag{5.27}$$

Assume that $\mathscr{S}$ is a compact hypersurface. Using this identity and Stokes' theorem, the Hamiltonian is decomposed as

$$H(X,\mathscr{S}) = H_{\text{boundary}}(X,\mathscr{S}) + H_{\text{volume}}(X,\mathscr{S}), \tag{5.28}$$

$$H_{\text{boundary}}(X,\mathscr{S}) = \frac{1}{2}\int_{\partial\mathscr{S}} \mathbb{W}^{\alpha\mu}\, dS_{\alpha\mu}, \tag{5.29}$$

$$\begin{aligned} H_{\text{volume}}(X,\mathscr{S}) &= -\int_{\mathscr{S}} \mathfrak{g}^{\mu\nu}\,\pounds_X \mathring{A}^{\lambda}{}_{\mu\nu}\, dS_{\lambda} \\ &= -\int_{\mathscr{S}} 2\mathfrak{g}^{\beta[\nu}\delta^{\lambda]}_{\sigma}\left(X^{\sigma}{}_{;\beta\gamma} - B^{\sigma}{}_{\beta\gamma\kappa}X^{\kappa}\right)\, dS_{\lambda}. \end{aligned} \tag{5.30}$$

For the variation formula, we use (4.57) and rewrite the volume integrand in terms of (5.24).

$$\begin{aligned} -\delta H_{\text{boundary}}(X,\mathscr{S}) = &\int_{\mathscr{S}} \left(\pounds_X \mathfrak{g}^{\mu\nu}\,\delta A^{\lambda}{}_{\mu\nu} - \pounds_X A^{\lambda}{}_{\mu\nu}\,\delta\mathfrak{g}^{\mu\nu}\right)\, dS_{\lambda} \\ &+ \int_{\partial\mathscr{S}} X^{[\mu}p^{\nu]}_{\alpha\beta}\,\delta\mathfrak{g}^{\alpha\beta}\, dS_{\mu\nu}. \end{aligned} \tag{5.31}$$

In the asymptotically flat setting, H_{boundary} will be used to recover the ADM energy-momentum as defined by Arnowitt, Deser, and Misner in [4, 5].

Remark 5.14 The assumption that $\mathscr{S}$ is compact is sufficient for the present discussion. In applications to asymptotically flat spacetimes, one has to deal with additional issues concerning compactification and regularity. In particular, smoothness of the compactification is generally too strong an assumption. We do not discuss these details here; see Van Voorthuizen [81, Section 3.3] for further discussion.

5.2 Reduction to a spacelike hypersurface

In Section 5.1, we identified the correct tensor bundle F, the Lagrangian L, and the appropriate reference system. This allowed us to apply the construction from Chapter 4 to general relativity. In this section we reduce the formulas for the Hamiltonian and antisymmetric two-form, using the assumption that the hypersurface $\mathscr{S}$ is spacelike.

In Subsection 5.2.1, we rewrite the antisymmetric two-form $\Omega_{\mathscr{S}}$. We also introduce the Gauss–Codazzi constraint equations. In Subsection 5.2.2, we decompose the spacetime metric into the induced metric on the hypersurface and the lapse and shift variables governing its evolution through spacetime. We use this decomposition to rewrite the variation of the Hamiltonian in a form suited for the asymptotically flat setting.

5.2.1 Reduction of the antisymmetric two-form and Gauss–Codazzi equations

From now on, we assume that the embedded hypersurface $\mathscr{S} \subset M$ is spacelike. We choose local coordinates (x^0, x^1, x^2, x^3) on $\mathscr{S}$, so that x^0 on $\mathscr{S}$ is constant. In these coordinates, the spacetime expression for the antisymmetric two-form simplifies.

Lemma 5.15 *Let (x^0, x^1, x^2, x^3) be coordinates on $\mathscr{S}$ such that x^0 is constant on $\mathscr{S}$. Then the antisymmetric two-form Ω_Σ, given in (5.19), is written as*

$$\begin{aligned}\Omega_{\mathscr{S}}((\delta_1 p^\lambda_{\mu\nu}, \delta_1 \mathfrak{g}^{\alpha\beta}), (\delta_2 p^\lambda_{\mu\nu}, \delta_2 \mathfrak{g}^{\alpha\beta})) &= \int_{\mathscr{S}} (\delta_1 p^0_{\alpha\beta} \delta_2 \mathfrak{g}^{\alpha\beta} - \delta_2 p^0_{\alpha\beta} \delta_1 \mathfrak{g}^{\alpha\beta}) d^3x \\ &= -\int_{\mathscr{S}} (\delta_1 A^0_{\alpha\beta} \delta_2 \mathfrak{g}^{\alpha\beta} - \delta_2 A^0_{\alpha\beta} \delta_1 \mathfrak{g}^{\alpha\beta}) d^3x. \end{aligned} \tag{5.32}$$

Proof. The first coordinate x^0 is constant on the hypersurface $\mathscr{S}$, consequently the form $dx^0|_{\mathscr{S}} = 0$. This simplifies the odd form dS_μ, because only terms without dx^0 remain. Hence we find

$$\begin{aligned}\sum_{\mu=0}^{3} dS_\mu &= \sum_{\mu=0}^{3} \frac{\partial}{\partial x^\mu} \rfloor dx^0 \wedge \cdots \wedge dx^3 \\ &= dx^1 \wedge dx^2 \wedge dx^3 \\ &= d^3x. \end{aligned}$$

Substituting this into equation (5.19), we obtain

$$\Omega_{\mathscr{S}}((\delta_1 p^\lambda_{\mu\nu}, \delta_1 \mathfrak{g}^{\alpha\beta}), (\delta_2 p^\lambda_{\mu\nu}, \delta_2 \mathfrak{g}^{\alpha\beta})) = \int_{\mathscr{S}} (\delta_1 p^0_{\alpha\beta} \delta_2 \mathfrak{g}^{\alpha\beta} - \delta_2 p^0_{\alpha\beta} \delta_1 \mathfrak{g}^{\alpha\beta}) d^3x.$$

Recall that the background metric is fixed, so its variation vanishes. From $\delta p^0_{\alpha\beta} = \delta \mathring{A}^0_{\alpha\beta} - \delta A^0_{\alpha\beta} = -\delta A^0_{\alpha\beta}$, it follows that

$$\Omega_{\mathscr{S}}((\delta_1 p^\lambda_{\mu\nu}, \delta_1 \mathfrak{g}^{\alpha\beta}), (\delta_2 p^\lambda_{\mu\nu}, \delta_2 \mathfrak{g}^{\alpha\beta})) = -\int_{\mathscr{S}} (\delta_1 A^0_{\alpha\beta} \delta_2 \mathfrak{g}^{\alpha\beta} - \delta_2 A^0_{\alpha\beta} \delta_1 \mathfrak{g}^{\alpha\beta}) d^3x. \tag{5.33}$$

This proves the claim. □

Note that the information about the momentum $p^0_{\alpha\beta}$ canonically conjugate to $\mathfrak{g}^{\alpha\beta}$ is encoded in the ten components $A^0_{\alpha\beta}$ of the connection Γ. However, these components are not independent. They satisfy constraints coming from the metricity condition (5.6).

Lemma 5.16 *For $k > 0$ and $l > 0$, we obtain the following two constraints*

$$A^0_{00} = \frac{1}{\mathfrak{g}^{00}} (\partial_k \mathfrak{g}^{0k} + A^0_{kl} \mathfrak{g}^{kl}), \tag{5.34}$$

$$A^0_{0k} = -\frac{1}{2\mathfrak{g}^{00}} (\partial_k \mathfrak{g}^{00} + 2A^0_{kl} \mathfrak{g}^{0l}). \tag{5.35}$$

Proof. Proving these constraint equations is a straightforward calculation, therefore we only give the calculation of equation (5.34).

Recall that $k > 0$ and $l > 0$, while other dummy variables may be zero. Note the following identities

$$A^0_{kl} = \Gamma^0_{kl} - \delta^0_{(k} \Gamma^\sigma_{l)\sigma} = \Gamma^0_{kl}$$

and

$$A^0_{00} = \Gamma^0_{00} - \delta^0_{(0} \Gamma^\sigma_{0)\sigma} = -\Gamma^k_{0k}$$

We prove (5.34) by rewriting Christoffel symbols and introducing $l > 0$ as a dummy variable. Using (5.6), we find

$$\begin{aligned}\partial_k \mathfrak{g}^{0k} &= \mathfrak{g}^{0k}\Gamma^\sigma_{k\sigma} - \mathfrak{g}^{0\sigma}\Gamma^k_{\sigma k} - \mathfrak{g}^{k\sigma}\Gamma^0_{\sigma k} \\ &= \mathfrak{g}^{0k}\Gamma^\sigma_{k\sigma} - \mathfrak{g}^{00}\Gamma^k_{0k} - \mathfrak{g}^{0l}\Gamma^k_{lk} - \mathfrak{g}^{0k}\Gamma^0_{0k} - \mathfrak{g}^{kl}\Gamma^0_{lk}\end{aligned}$$

We collect the terms $\mathfrak{g}^{0k}$ and interchange the dummy indices k and l in the third term. This gives

$$\begin{aligned}\partial_k \mathfrak{g}^{0k} &= \mathfrak{g}^{0k}\Gamma^l_{lk} - \mathfrak{g}^{00}\Gamma^k_{0k} - \mathfrak{g}^{0k}\Gamma^l_{kl} - \mathfrak{g}^{kl}\Gamma^0_{lk} \\ &= -\mathfrak{g}^{00}\Gamma^k_{0k} - \mathfrak{g}^{kl}\Gamma^0_{kl} \\ &= \mathfrak{g}^{00}A^0_{00} - \mathfrak{g}^{kl}A^0_{kl}.\end{aligned}$$

Rearranging the terms proves (5.34). □

We now use a result from Chruściel, Jezierski and Kijowski [27] which rewrites $\Omega_{\mathscr{S}}$. For a more detailed treatment, see Kijowski [49, Section 5]. The theorem below shows that, up to a boundary term, the spacetime antisymmetric two-form can be written in terms of the induced metric g_{kl} on the hypersurface $\mathscr{S}$ and its conjugate momentum P^{kl}.

Theorem 5.17 *Suppose that $\mathscr{S}$ is a hypersurface of M which has compact closure and smooth compact boundary $\partial\mathscr{S}$. On an open subset $\mathscr{O} \subset M$ we take coordinates (x^0, x^1, x^2, x^3) that satisfy $\mathscr{S} \cap \mathscr{O} = \{x^0 = 0\}$ and $\partial\mathscr{S} \cap \mathscr{O} = \{x^0 = 0, x^3 = 1\}$. The antisymmetric two-form $\Omega_{\mathscr{S}}$ from equation (5.19) reduces to*

$$\begin{aligned}\Omega_{\mathscr{S}}((\delta_1 p^\lambda_{\mu\nu}, \delta_1\mathfrak{g}^{\alpha\beta}), (\delta_2 p^\lambda_{\mu\nu}, \delta_2\mathfrak{g}^{\alpha\beta})) &= \frac{1}{16\pi}\int_{\mathscr{S}} (\delta_1 g_{kl}\delta_2 P^{kl} - \delta_2 g_{kl}\delta_1 P^{kl}) d^3x \\ &+ \frac{1}{16\pi}\int_{\partial\mathscr{S}} \left(\delta_1 N^3 \delta_2 \frac{\sqrt{\det g_{kl}}}{N} - \delta_2 N^3 \delta_1 \frac{\sqrt{\det g_{kl}}}{N}\right) d^2x,\end{aligned} \tag{5.36}$$

where

$$N = \frac{1}{\sqrt{-g^{00}}}, \;\; N_k = g_{0k}, \;\; N^3 = (g^{3k} - \frac{g^{03}g^{0k}}{g^{00}})N_k$$

and $P^{kl} := \sqrt{\det g_{mn}}({}^3g^{ij}K_{ij}{}^3g^{kl} - K^{kl})$ with $K_{kl} = -1/\sqrt{|g^{00}|}A^0_{kl}$, and with ${}^3g^{kl}$ being the three-dimensional inverse of the induced metric g_{kl} on $\mathscr{S}$.

Remark 5.18 The theorem is stated here in a simplified setting sufficient for the present discussion. In applications to asymptotically flat spacetimes, one has to deal with additional issues concerning compactification and regularity. In particular, smoothness of the compactification is generally too strong an assumption. We do not discuss these details here; see Van Voorthuizen [81, Section 3.3] for further discussion.

In Subsection 5.3.2, in the asymptotically flat setting, we will see that the boundary integral vanishes due to the asymptotic conditions on the metric.

The induced data on the hypersurface $\mathscr{S}$ satisfy the Gauss–Codazzi constraint equations. We follow the convention of Chruściel–Jezierski–Kijowski [27].

Proposition 5.19 *The tensor fields (g_{kl}, P^{kl}) satisfy the Gauss-Codazzi constraint equations*

$$P_i{}^l{}_{|}l = 8\pi\sqrt{\det g_{mn}}T_{i\mu}n^\mu, \tag{5.37}$$

$$(\det g_{mn})\mathscr{R} - P^{kl}P_{kl} + \frac{1}{2}(P^{kl}g_{kl})^2 = 16\pi(\det g_{mn})T_{\mu\nu}n^\mu n^\nu, \tag{5.38}$$

where $T_{\mu\nu}$ denotes the energy-momentum tensor of matter fields, $\mathscr{R}$ denotes the three-dimensional scalar curvature of g_{kl}, $n = n^\mu\partial_\mu$ is a future timelike four-vector normal to the hypersurface $\mathscr{S}$ and "$|$" denotes the three-dimensional covariant derivative induced by g_{kl}.

Remark 5.20 The Gauss–Codazzi equations arise from the geometry of an embedded hypersurface. When the ambient manifold satisfies the Einstein equations, they give rise to the constraint equations (5.37)–(5.38).

5.2.2 Lapse function and shift vector field

In this subsection, we describe how a spacelike hypersurface moves through spacetime. We pull back the metric g to $I \times \Sigma$. Using the assumption that each embedded hypersurface is spacelike, we obtain a Riemannian metric γ on Σ. We then introduce the lapse function ν and the shift vector field β, which allow us to write $\psi^* g$ in terms of γ, ν, and β. We also introduce a momentum π^{kl} canonically conjugate to γ_{kl}. These variables satisfy certain relations, which will be used to reduce the variation of the Hamiltonian H_{boundary}, given by (5.31).

Let ψ be a motion, as defined in Definition 4.2, and let Σ be a model hypersurface in M. We assume that for every $t \in I$, the image $\psi_t(\Sigma)$ is a spacelike hypersurface in M. This ensures that for every time $t \in I$, pulling back the metric g along the embedding ψ_t gives a Riemannian metric γ_t on Σ. We denote this Riemannian metric by

$$\gamma_t = \gamma_{t,ij}dx^i dx^j = (\psi_t)^* g.$$

To describe how the hypersurface evolves in time, we define a one-parameter[3] family of lapse functions on Σ.

Definition 5.21 Let $\psi\colon I \times \Sigma \to M$ be a map such that for all $t \in I$, the map ψ_t is an embedding with spacelike image. The **lapse functions** are the one-parameter family of functions $\nu_t\colon \Sigma \to M$ defined by

$$\nu_t = g(d\psi\frac{\partial}{\partial t}, n_t) \circ \psi_t,$$

where n_t denotes the unit future timelike four-vector normal to the hypersurface $\psi_t(\Sigma)$.

[3]Recall from Remark 5.13 that the metric $g = g(s)$ and the corresponding dynamical variables $\mathfrak{g}^{\alpha\beta}(s)$ and $p^\mu_{\alpha\beta}(s)$ are understood to belong to differentiable one-parameter families, although their dependence on the variation parameter s is suppressed. Consequently, the quantities introduced through the time evolution are, strictly speaking, two-parameter families. For example,

$$\gamma(s,t) = \psi_t^* g(s),$$

and similarly for the lapse, shift, and canonical momentum. Here s is the variation parameter, whereas t is the evolution parameter. We continue to suppress the dependence on s and write these quantities as γ_t, ν_t, β_t, and π_t. Accordingly, δ denotes differentiation with respect to s, while a dot denotes differentiation with respect to t.

The lapse function measures the normal component of the time-evolution vector field. Thus the lapse vanishes when the hypersurface is moved tangentially to itself.

Lemma 5.22 *For each $t \in I$, the vector field given by*

$$d\psi \frac{\partial}{\partial t} + \nu_t \cdot (n_t \circ \psi_t)$$

is orthogonal to n_t.

Proof. Let $t \in I$. Using the bilinearity of the metric we find

$$\begin{aligned} g(d\psi \frac{\partial}{\partial t} + \nu_t \cdot (n_t \circ \psi_t), n_t \circ \psi_t) &= g(d\psi \frac{\partial}{\partial t}, n_t \circ \psi_t) + g(\nu_t \cdot (n_t \circ \psi_t), n_t \circ \psi_t) \\ &= g(d\psi \frac{\partial}{\partial t}, n_t) \circ \psi_t + \nu_t g(n_t \circ \psi_t, n_t \circ \psi_t) \\ &= \nu_t - \nu_t = 0 \end{aligned}$$

Since n_t is a unit timelike normal to the hypersurface $\psi_t(\Sigma)$, $g(n_t, n_t) = -1$. This completes the proof. □

We conclude that the vector field

$$d\psi \frac{\partial}{\partial t} + \nu_t \cdot (n_t \circ \psi_t)$$

is tangent to $\psi_t(\Sigma)$. Hence there exists a one-parameter family of vector fields $\beta(t, \cdot)$ on Σ such that

$$d\psi \frac{\partial}{\partial t} + \nu_t \cdot (n_t \circ \psi_t) = d\psi_t \beta(t, \cdot).$$

The vector field $d(\psi_t)\beta(t, \cdot)$ is called the **shift vector field**.

The lapse function and shift vector field now allow us to decompose the pulled back metric $\psi^* g$ into the metric induced on the hypersurfaces and the variables describing the choice of time evolution. The following formula gives this decomposition.

Proposition 5.23 *In local coordinates on $I \times \Sigma$, the pulled back metric is given by*

$$\psi^* g = -\nu^2 dt^2 + \gamma_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt). \tag{5.39}$$

Equivalently,

$$\psi^* g = -\nu^2 dt^2 + \gamma_{ij} dx^i dx^j + \gamma_{ij}\beta^j dx^i dt + \gamma_{ij}\beta^i dx^j dt + \gamma_{ij}\beta^i\beta^j dt^2, \tag{5.40}$$

where $i = 1, 2, 3$ in both formulas.

For a derivation of this decomposition, see Gourgoulhon [37, Section 4.2.3].

Although γ_t is Riemannian on each hypersurface Σ, the metric $\psi^* g$ on $I \times \Sigma$ is only Lorentzian whenever the lapse is non-zero.

We show a few elementary identities that will be used in Section 5.3. They follow directly from the decomposition of the metric $\psi^* g$ and from the compatibility of the Levi–Civita connection of γ_{ij}. Recording them here keeps the proofs in Section 5.3 from being interrupted by coordinate calculations.

Lemma 5.24 *In local coordinates, the decomposition of $\psi^* g$ is given by*

$$(\psi^* g)_{00} = -\nu^2 + \gamma_{ij}\beta^i\beta^j, \tag{5.41}$$
$$(\psi^* g)_{ij} = \gamma_{ij}, \tag{5.42}$$
$$(\psi^* g)_{0i} = \gamma_{ij}\beta^j. \tag{5.43}$$

In particular,

$$\beta_i = (\psi^* g)_{0i}, \qquad \beta^i = \gamma^{ij}(\psi^* g)_{0j}. \tag{5.44}$$

Proof. The second identity of equation (5.44) is obtained by using the inverse metric. The first identity is obtained by lowering indices

$$\beta_i = \gamma_{ij}\beta^j = (\psi^* g)_{0i}.$$

The identities (5.42) and (5.43) are obtained directly from the decomposition of the metric $\psi^* g$. □

The induced metric γ_{ij} and its canonically conjugate momentum π^{ij} satisfy the standard ADM evolution equations. We will only need the following relation

$$\dot\gamma_{kl} = \frac{2\nu}{\sqrt{\det \gamma_{ij}}}\left(\pi_{kl} - \frac{1}{2}\gamma_{kl}\,\mathrm{tr}(\pi)\right) + \beta_{k|l} + \beta_{l|k}, \tag{5.45}$$

where $|$ denotes the covariant derivative with respect to γ_{ij}, and

$$\mathrm{tr}(\pi) = \gamma^{kl}\pi_{kl}.$$

For a derivation of the full set of constraints, see Misner–Thorne–Wheeler [79, p. 525].

Lemma 5.25 *The covariant derivative of the lowered shift vector is given by*

$$\beta_{k|l} = \gamma_{ki}\beta^i{}_{|l}. \tag{5.46}$$

Proof. Since $\beta_k = \gamma_{ki}\beta^i$, the components β_k define a one-form. We therefore expand $\beta_{k|l}$ as the covariant derivative of a one-form, and then use the metric compatibility of the Levi–Civita connection of γ. This gives

$$\begin{aligned}\beta_{k|l} &= \nabla_l(\gamma_{ij}\beta^j)\\ &= \partial_l(\gamma_{ik}\beta^i) - \Gamma^m_{lk}\gamma_{mi}\beta^i\\ &= (\partial_l\gamma_{ik})\beta^i + \gamma_{ik}(\partial_l\beta^i) - \Gamma^m_{lk}\gamma_{mi}\beta^i\end{aligned} \tag{5.47}$$

By metric compatibility

$$\partial_l\gamma_{ki} = \Gamma^n_{lk}\gamma_{ni} + \Gamma^n_{li}\gamma_{kn}.$$

Substituting this into (5.47) gives

$$\begin{aligned}\beta_{k|l} &= \Gamma^n_{li}\gamma_{nk}\beta^i + \Gamma^n_{lk}\gamma_{in}\beta^i + \gamma_{ik}(\partial_l\beta^i) - \Gamma^m_{lk}\gamma_{mi}\beta^i\\ &= \gamma_{ik}(\partial_l\beta^i + \Gamma^k_{ln}\beta^i)\\ &= \gamma_{ki}\beta^i{}_{|l}.\end{aligned}$$

This proves the desired identity. □

We now quote the corresponding reduction of the variation of the Hamiltonian, analogous to the reduction of the antisymmetric two-form above. We do not reproduce the coordinate calculation here, since it will not be needed later. For a proof, see Kijowski [49, Section 5.6 and Appendix A]. We follow the formulation of Chruściel, Jezierski, and Kijowski [27].

Theorem 5.26 (Reduced variation formula for the gravitational boundary Hamiltonian) *Let (M, g) be a four-dimensional vacuum spacetime, let B be a fixed symmetric connection on M, and let Σ be a compact*[4] *three-dimensional manifold with smooth boundary.*

Let

$$\psi\colon I \times \Sigma \longrightarrow M$$

be a smooth motion generated by a vector field $X \in \mathfrak{X}(M)$. Set $\mathscr{S}_t := \psi_t(\Sigma)$, $\mathscr{S} := \mathscr{S}_0$, and suppose that $\mathscr{S}_t$ is spacelike for every $t \in I$.

Define the metric density and the corresponding canonical momentum by

$$\mathfrak{g}^{\alpha\beta} := \frac{1}{16\pi}\sqrt{-\det g}\, g^{\alpha\beta},$$
$$p^{\lambda}{}_{\mu\nu} := \left(B^{\lambda}{}_{\mu\nu} - \delta^{\lambda}{}_{(\mu}B^{\kappa}{}_{\nu)\kappa}\right) - \left(\Gamma^{\lambda}{}_{\mu\nu} - \delta^{\lambda}{}_{(\mu}\Gamma^{\kappa}{}_{\nu)\kappa}\right),$$

where Γ is the Levi–Civita connection of g.

Write the pullback of the spacetime metric to $I \times \Sigma$ in the form

$$\psi^* g = -\nu^2\, dt^2 + \gamma_{ij}\left(dx^i + \beta^i\, dt\right)\left(dx^j + \beta^j\, dt\right),$$

where ν, β^i, and γ_{ij}[5] *are, respectively, the lapse, the shift, and the induced Riemannian metric. Let K_{ij} denote the second fundamental form of $\mathscr{S}_t$, and define the ADM momentum density by*

$$\pi^{ij} := \sqrt{\det \gamma_{mn}}\left(K\gamma^{ij} - K^{ij}\right), \qquad K := \gamma^{ij}K_{ij}.$$

Set

$$\sigma := \beta^3, \qquad \mu := \frac{\sqrt{\det \gamma_{kl}}}{\nu}.$$

Recall that we suppress the notation for the variation of the dynamical variables $(\mathfrak{g}, p)$. The same is true for the variables obtained from this pair. For every quantity $q(\lambda)$ induced by this family, define

$$\delta q := \left.\frac{d}{d\lambda}\right|_{\lambda=0} q(\lambda).$$

The background connection B, the vector field X, and the motion ψ are held fixed under the variation. A dot denotes differentiation with respect to the evolution parameter t:

$$\dot{q} := \left.\frac{\partial q}{\partial t}\right|_{t=0}.$$

Let $H_{\text{boundary}}(X, \mathscr{S})$ be the gravitational boundary Hamiltonian defined in (5.29). Then

$$-\,\delta H_{\text{boundary}}(X, \mathscr{S}) = \frac{1}{16\pi}\int_{\Sigma}\left(\dot{\pi}^{kl}\,\delta\gamma_{kl} - \dot{\gamma}_{kl}\,\delta\pi^{kl}\right)d^3x$$
$$+\,\frac{1}{16\pi}\int_{\partial\Sigma}\left(\dot{\mu}\,\delta\sigma - \dot{\sigma}\,\delta\mu\right)d^2x + \int_{\partial\mathscr{S}}\left(X^0 p^3{}_{\alpha\beta}\delta\mathfrak{g}^{\alpha\beta} - X^3 p^0{}_{\alpha\beta}\delta\mathfrak{g}^{\alpha\beta}\right)d^2x. \tag{5.48}$$

[4] For a discussion on this assumption, see Remark 5.18

[5] We suppress notation of the evolution parameter t.

5.3 Asymptotically flat spacetime

In Section 5.2, we assumed the hypersurface $\mathscr{S}$ was spacelike. This reduced the formulas for the antisymmetric two-form $\Omega_{\mathscr{S}}$ and the variation of the Hamiltonian $\delta H_{\text{boundary}}$ substantially. In this section, we use Theorem 5.26 and Theorem 5.17 to construct a Hamiltonian dynamical system for asymptotically flat spacetimes.

We now apply the Hamiltonian formulation of Chruściel, Jezierski, and Kijowski [27] to asymptotically flat spacetimes, schematically illustrated in Figure 5.1. Roughly speaking, asymptotic flatness allows the geometry to be highly nontrivial in a bounded region while requiring it to approach that of Minkowski spacetime at spatial infinity. More precisely, for a spacelike hypersurface $\mathscr{S}$, there exists a compact set $K \subset \mathscr{S}$ such that, outside K, the metric approaches the Minkowski metric under the prescribed asymptotic conditions.

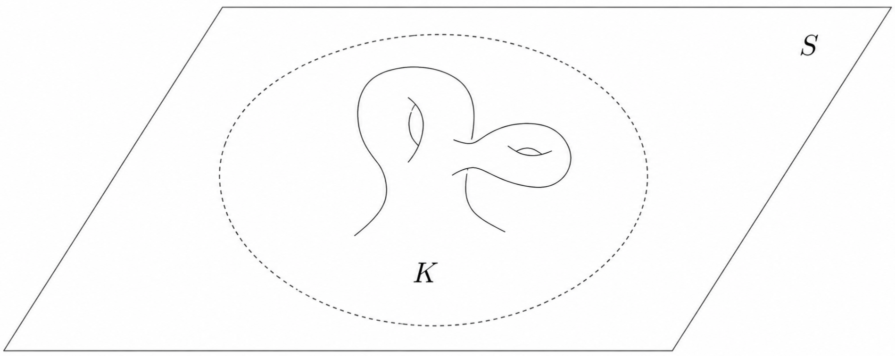

Figure 5.1: Schematic picture of an asymptotically flat hypersurface $\mathscr{S}$. The compact region K represents the part where the geometry may be strongly curved or otherwise nontrivial. Outside this region, the metric approaches the flat Minkowski metric.

In Subsection 5.3.1, we introduce the precise asymptotic conditions and begin the construction of the corresponding phase space. Subsection 5.3.2 continues this construction by imposing further restrictions on the initial data. We then use an idea by Lichnerowicz [56] to rewrite part of the antisymmetric two-form and prove the convergence and vanishing of the relevant integrals.

In Subsection 5.3.3, we show that, under the asymptotic conditions chosen earlier, the variation of the Hamiltonian agrees with the antisymmetric two-form. This gives a Hamiltonian dynamical system on the asymptotically flat phase space. Finally, in Subsection 5.3.4, we show that the ADM four-momentum is obtained from the Hamiltonian. We first give a sketch of the argument. Since we did not find this calculation in the literature, we include the full computation in Subsection 5.3.5.

5.3.1 Asymptotic conditions

We begin this subsection by specifying what we mean by an asymptotically flat spacetime. We then introduce the corresponding initial data, which will be used in Subsection 5.3.3 to construct the phase space.

Definition 5.27 A 4-dimensional spacetime (M, g) is **asymptotically flat** if there exists a compact set $K \subset M$ such that $M \setminus K$ is a finite union of **ends** $M_1, ..., M_l$, and for each end M_i, there exists a diffeomorphism

$$\Phi_i \colon M_i \to \mathbb{R} \times \mathbb{R}^3 \setminus B_R(0),$$

where $B_R(0)$ is the closed ball with radius R centred at the origin. We regard Φ_i as an asymptotically flat coordinate chart, with coordinates t, x, y, z. In this coordinate chart, the metric g satisfies

$$|g_{\mu\nu} - \eta_{\mu\nu}| = O(r^{-1}), \ |\partial g_{\mu\nu}| = O(r^{-2}), \ |\partial^2 g_{\mu\nu}| = O(r^{-3}), \tag{5.49}$$

where $\eta_{\mu\nu}$ denotes the Minkowski metric and r is the Euclidean radial coordinate on the asymptotic end.

Remark 5.28 The asymptotic conditions used here are the classical fall-off conditions. They can be optimized in more refined treatments. For further discussion, see Chruściel–Jezierski–Kijowski [27].

Example 5.29 The Schwarzschild spacetime provides a standard example of an asymptotically flat spacetime. In suitable asymptotic coordinates, its metric satisfies the prescribed decay conditions.

In this thesis, we restrict attention to asymptotically flat spacetimes with a single end. As Chruściel, Jezierski, and Kijowski note in [27, p. 76], the extension to finitely many ends is a straightforward generalization.

We assume that Σ, the **model space for our hypersurfaces**, is an asymptotically flat Riemannian manifold. Thus Σ can be decomposed as

$$\Sigma = \Sigma_{\text{int}} \cup \Sigma_{\text{ext}},$$

where Σ_{int} is compact and

$$\Sigma_{\text{ext}} \cong \mathbb{R}^3 \setminus B_R(0).$$

We now choose a fixed background metric on the model spacetime $\mathbb{R} \times \Sigma$. On the asymptotic region $\mathbb{R} \times \Sigma_{\text{ext}}$, we choose this background metric to be the Minkowski metric,

$$b = -dt^2 + dx^2 + dy^2 + dz^2.$$

The restriction of b to Σ_{ext} is the Euclidean metric

$$dx^2 + dy^2 + dz^2.$$

Remark 5.30 The physical metric g is defined on the spacetime M, whereas the background metric b is defined on the model spacetime $\mathbb{R} \times \Sigma$. To compare these metrics, one has to transport g to $\mathbb{R} \times \Sigma$. We return to this point in Subsection 5.3.3.

We will define the phase space in several steps. We begin by choosing **initial reference data**, which is a pair

$$(\mathring{\gamma}_{ij}, \mathring{\pi}^{kl}), \tag{5.50}$$

where $\mathring{\gamma}_{ij}$ is a Riemannian metric on Σ and $\mathring{\pi}^{kl}$ is its conjugate momentum on Σ. We require this pair to satisfy the vacuum Gauss–Codazzi constraint equations (5.37)–(5.38), with the induced metric and momentum variables replaced by $\mathring{\gamma}_{ij}$ and $\mathring{\pi}^{ij}$, and with the energy-momentum tensor of the matter fields $T_{\mu\nu} = 0$.

In addition, we assume that the pair $(\mathring{\gamma}_{ij}, \mathring{\pi}^{ij})$ arises from an asymptotically flat spacetime. Thus it is induced by a spacetime metric satisfying the asymptotic conditions in (5.49). We denote

$$\operatorname{tr}(\mathring{\pi}) := \mathring{\gamma}_{ij}\mathring{\pi}^{ij}. \tag{5.51}$$

We approximate the trace by $\operatorname{tr}(\mathring{\pi}) = O(r^{-2})$.

Lemma 5.31 *The trace of the initial reference data satisfies the fall-off*

$$\operatorname{tr}(\mathring{\pi}) = O(r^{-2}).$$

Proof. To avoid overloading the notation, we omit the circles in this proof. Recall equation (5.45)

$$\dot{\gamma}_{kl} = \frac{2\nu}{\sqrt{\det \gamma_{ij}}}\left(\pi_{kl} - \frac{1}{2}\gamma_{kl}\operatorname{tr}(\pi)\right) + \beta_{k|l} + \beta_{l|k}. \tag{5.52}$$

By (5.46), we have

$$\beta_{k|l} = \gamma_{ik}\beta^i_{|l}.$$

Contracting (5.52) with γ^{kl} gives

$$\begin{aligned}
\gamma^{kl}\dot{\gamma}_{kl} &= \gamma^{kl}\left(\frac{2\nu}{\sqrt{\det \gamma_{ij}}}\left(\pi_{kl} - \frac{1}{2}\gamma_{kl}\pi\right) + 2\gamma_{ik}\beta^i_{|l}\right) \\
&= \frac{2\nu}{\sqrt{\det \gamma_{ij}}}\left(\gamma^{kl}\pi_{kl} - \frac{1}{2}\gamma^{kl}\gamma_{kl}\operatorname{tr}(\pi)\right) + 2\gamma^{kl}\gamma_{ik}\beta^i_{|l} \\
&= \frac{2\nu}{\sqrt{\det \gamma_{ij}}}\left(\operatorname{tr}(\pi) - \frac{3}{2}\operatorname{tr}(\pi)\right) + 2\beta^i_{|i}
\end{aligned}$$

Rearranging gives

$$\operatorname{tr}(\pi) = \frac{\sqrt{\det \gamma_{ij}}}{\nu}\left(2\beta^i_{|i} - \gamma^{kl}\dot{\gamma}_{kl}\right). \tag{5.53}$$

We now estimate the right-hand side using the asymptotic conditions (5.49). Since

$$\gamma^{ij} - \eta^{ij} = O(r^{-1}),$$

we also have

$$\gamma^{ij} = \eta^{ij} + O(r^{-1}).$$

Since the Minkowski metric contributes the leading term, we conclude that $\gamma^{ij} = O(1)$. Moreover, the asymptotic conditions imply $\dot{\gamma}_{ij} = O(r^{-2})$. Hence

$$\gamma^{kl}\dot{\gamma}_{kl} = O(1)O(r^{-2}). \tag{5.54}$$

It remains to estimate ${\beta^i}_{|i}$. By (5.44), we have

$$\beta_i = (\psi^* g)_{0i}.$$

Thus, for the approximation of $\beta^k_{|k}$ we find

$$\begin{aligned} \partial_k \beta^k &= \partial_k(\gamma^{ki})\beta_i + \gamma^{ik}\partial_k(\beta_i) \\ &= O(r^{-2})O(1) + O(1)O(r^{-2}) \\ &= O(r^{-2}). \end{aligned} \tag{5.55}$$

Furthermore, we have

$$\frac{\sqrt{\det \gamma_{ij}}}{\nu} = O(1). \tag{5.56}$$

Substituting (5.56), (5.55) and (5.54) into (5.53) gives

$$\operatorname{tr}(\pi) = O(1)\left(2O(r^{-2}) - O(r^{-2})\right) = O(r^{-2}). \qquad \square$$

5.3.2 Convergence of antisymmetric two-form

In Subsection 5.3.1 we fixed the asymptotic setting and chose the initial reference data around which the phase space will be constructed. The goal of this subsection is to determine the conditions under which the reduced antisymmetric two-form $\Omega_{\mathscr{S}}$, given in (5.57), is well-defined on a non-compact hypersurface $\mathscr{S}$. Moreover, we show that the boundary term vanishes under these conditions. This vanishing of the boundary integral is important because it allows the antisymmetric two-form $\Omega_{\mathscr{S}}$ to be identified with the variation of the Hamiltonian $\delta H_{\text{boundary}}$.

We do not give a full proof of the convergence of the antisymmetric two-form. Instead, following Chruściel, Jezierski, and Kijowski [27], we state the fall-off conditions that ensure convergence of the volume term and vanishing of the boundary term. The main idea comes from a method due to Lichnerowicz [56]. One uses a conformal transformation of the spatial metric to rewrite the integrand of the volume term. This separates the different contributions to the integrand and makes clear which fall-off conditions are needed to control them. These conditions will then be incorporated into the definition of the asymptotically flat phase space.

Recall from (5.36) that, after pulling back to the model hypersurface Σ, the antisymmetric two-form takes the form

$$\begin{aligned} \Omega_{\mathscr{S}}((\delta_1 p^\lambda_{\mu\nu}, \delta_1 \mathfrak{g}^{\alpha\beta}), (\delta_2 p^\lambda_{\mu\nu}, \delta_2 \mathfrak{g}^{\alpha\beta})) &= \frac{1}{16\pi}\int_\Sigma \left(\delta_1\gamma_{kl}\,\delta_2\pi^{kl} - \delta_2\gamma_{kl}\,\delta_1\pi^{kl}\right)\, d^3x \\ &\quad + \frac{1}{16\pi}\int_{\partial\mathscr{S}} \left(\delta_1 N^3\,\delta_2\frac{\sqrt{\det g_{kl}}}{N} - \delta_2 N^3\,\delta_1\frac{\sqrt{\det g_{kl}}}{N}\right)\, d^2x. \end{aligned} \tag{5.57}$$

Here, we replace the induced metric g_{kl} and momentum P^{kl} by the Riemannian metric γ_{kl} and the momentum π^{kl} canonically conjugate to it.

Definition 5.32 Let Σ be a manifold. Two Riemannian metrics γ_{ij} and $\hat{\gamma}_{ij}$ on Σ are called **conformal** if there exists a smooth positive function $f\colon \Sigma \to (0,\infty)$ such that

$$\gamma_{ij} = f\hat{\gamma}_{ij}.$$

The Lichnerowicz conformal method rewrites the Gauss–Codazzi scalar constraint equation (5.38) as an elliptic equation for a conformal factor. This has the following consequence. For more details, see Lichnerowicz [56].

Theorem 5.33 *Let γ_{ij} be an Riemannian metric. Suppose that the metric γ_{ij} is written in the conformal form*

$$\gamma_{ij} = e^{\phi}\hat{\gamma}_{ij}.$$

Then the integrand of

$$\int_{\Sigma} \left(\delta_1\gamma_{kl}\,\delta_2\pi^{kl} - \delta_2\gamma_{kl}\,\delta_1\pi^{kl}\right)\, d^3x$$

is written as

$$\begin{aligned}&(\delta_1\gamma_{kl}\delta_2\pi^{kl} - \delta_1 \leftrightarrow \delta_2) = \\ &\delta_1\phi\delta_2(\gamma_{kl}\pi^{kl}) - e^{\phi}[\pi^{kl}\delta_1\phi\delta_2\hat{\gamma}_{kl} - \delta_1\hat{\gamma}_{kl}\delta_2\pi^{kl}] - \delta_1 \leftrightarrow \delta_2.\end{aligned} \tag{5.58}$$

Here $\delta_1 \leftrightarrow \delta_2$ denotes the preceding expression with δ_1 and δ_2 interchanged.

In the asymptotically flat setting, the conformal factor can in general be chosen so that

$$\phi = O(r^{-1}).$$

The fall-off of the conformal metric $\hat{\gamma}_{ij}$ can be freely chosen. Since $\phi = O(r^{-1})$, the Taylor expansion of e^{ϕ} gives

$$e^{\phi} = 1 + O(r^{-1}),$$

and in particular $e^{\phi} = O(1)$. Thus the factor e^{ϕ} does not affect the integrability of the corresponding terms in (5.58).

Under suitable fall-off assumptions on $\hat{\gamma}_{ij}$ and π^{kl}, the terms involving

$$e^{\phi}\left(\pi^{kl}\delta_1\phi\,\delta_2\hat{\gamma}_{kl} - \delta_1\hat{\gamma}_{kl}\delta_2\pi^{kl}\right)$$

are integrable. Thus, the only problematic term in equation (5.58) is

$$\delta_1\phi\,\delta_2(\gamma_{kl}\pi^{kl}) - (\delta_1 \leftrightarrow \delta_2).$$

We therefore impose an extra condition on the fall-off of $\gamma_{kl}\pi^{kl}$, or more precisely on its deviation from the trace of the chosen initial reference data $(\mathring{\gamma}_{ij}, \mathring{\pi}_{ij})$.

We now continue the construction of the phase space from the initial reference data introduced in Subsection 5.3.1. Fix the reference data

$$(\mathring{\gamma}_{ij}, \mathring{\pi}^{ij}).$$

The phase space $\mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$ consists of initial data

$$(\gamma_{ij}, \pi^{ij}) \in \mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$$

that satisfy the following asymptotic conditions on Σ_{ext}:

$$\phi = O(r^{-1}),\ \frac{\gamma_{kl}\pi^{kl} - \mathrm{tr}(\mathring{\pi})}{r} \in L^1(\Sigma_{\mathrm{ext}}) \tag{5.59}$$

where ϕ is the conformal factor determined by $\gamma_{ij} = e^{\phi}\hat{\gamma}_{ij}$. Furthermore, we assume that

$$\hat{\gamma}_{ij} - \delta_{ij} \in L^2(\Sigma_{\mathrm{ext}}),\ \pi^{ij} \in L^2(\Sigma_{\mathrm{ext}}). \tag{5.60}$$

Remark 5.34 The definition of the phase space is not yet complete. We will still need to impose conditions on the evolution of the initial data, which will be done in Subsection 5.3.3. The assumptions introduced here are sufficient for the purpose of this subsection, namely to prove convergence of the antisymmetric two-form.

Example 5.35 To make the assumptions more concrete, we give fall-off conditions that imply (5.59) and (5.60). The conditions in (5.59) are satisfied if

$$\phi = O(r^{-1}), \qquad \gamma_{kl}\pi^{kl} - \operatorname{tr}(\mathring{\pi}) = O(r^{-\alpha})$$

for some $\alpha > 2$. Indeed, in this case

$$\frac{\gamma_{kl}\pi^{kl} - \operatorname{tr}(\mathring{\pi})}{r} = O(r^{-\alpha-1}),$$

which is integrable on Σ_{ext} when $\alpha > 2$.

Similarly, the conditions in (5.60) are satisfied if

$$\hat{\gamma}_{ij} - \delta_{ij} = O(r^{-\delta}), \qquad \pi^{ij} = O(r^{-\beta})$$

for some $\delta > 3/2$ and $\beta > 3/2$.

Following Chruściel, Jezierski, and Kijowski, we use the following consequence of the fall-off conditions. The statement will be used without proof.

Theorem 5.36 (Well-definedness of the antisymmetric form) *Let (M, g) be a vacuum asymptotically flat spacetime, where g satisfies the asymptotic conditions (5.49). Let Σ be the model hypersurface, and let*

$$i \colon \Sigma \longrightarrow M$$

be an embedding whose image $\mathscr{S} := i(\Sigma)$ is spacelike.

Fix a reference initial data set $\left(\mathring{\gamma}_{ij}, \mathring{\pi}^{ij}\right)$ on Σ satisfying the vacuum Gauss–Codazzi equations (5.37) and (5.38), and denote the trace as

$$\mathring{\pi} := \mathring{\gamma}_{kl}\mathring{\pi}^{kl}.$$

Let (γ_{ij}, π^{ij}) be an initial data set on Σ satisfying the vacuum constraint equations. On the asymptotic region Σ_{ext}, write

$$\gamma_{ij} = e^{\phi}\widehat{\gamma}_{ij}.$$

Suppose that

$$\phi = O(r^{-1}), \quad \frac{\gamma_{kl}\pi^{kl} - \operatorname{tr}(\mathring{\pi})}{r} \in L^1(\Sigma_{\text{ext}}), \tag{5.61}$$

and

$$\hat{\gamma}_{ij} - \delta_{ij} \in L^2(\Sigma_{\text{ext}}), \quad \pi^{ij} \in L^2(\Sigma_{\text{ext}}). \tag{5.62}$$

Then the volume integral defining the reduced antisymmetric form $\Omega_{\mathscr{S}}$ is convergent. Moreover, the boundary term in (5.57) vanishes. Consequently,

$$\Omega_{\mathscr{S}}\left((\delta_1 p^{\lambda}_{\mu\nu}, \delta_1 \mathfrak{g}^{\alpha\beta}), (\delta_2 p^{\lambda}_{\mu\nu}, \delta_2 \mathfrak{g}^{\alpha\beta})\right) = \frac{1}{16\pi}\int_{\Sigma} \left(\delta_1\gamma_{kl}\,\delta_2\pi^{kl} - \delta_2\gamma_{kl}\,\delta_1\pi^{kl}\right) d^3x. \tag{5.63}$$

5.3.3 Hamiltonian dynamical system

In this subsection we use the Hamiltonian formalism developed in Section 5.1 to construct a Hamiltonian dynamical system for asymptotically flat spacetimes.

The construction is done in several steps. We first complete the definition of the local phase space $\mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$ associated with fixed reference initial data $(\mathring{\gamma}_{ij}, \mathring{\pi}^{kl})$, as defined in (5.50). In Subsection 5.3.2, we imposed the required conditions on initial data sets (γ_{ij}, π^{ij}). It remains to specify the evolution of these data and to verify that the prescribed conditions are preserved by this evolution.

We then pass from these local phase spaces to the ADM phase space, by taking the union of the local phase spaces. In this sense, the local phase spaces may be regarded as local patches of the full ADM phase space. This interpretation is intended only as a heuristic, since we do not develop a manifold structure on the ADM phase space.

Finally, we state a convergence result for the variation of the Hamiltonian. Together with Theorem 5.36, this shows that the ADM phase space carries the structure of a Hamiltonian dynamical system.

We begin by addressing the problem of identifying the physical spacetime with the model spacetime $I \times \Sigma$. Let (γ_{ij}, π^{ij}) be a candidate initial data pair satisfying the conditions (5.59) and (5.60). By Choquet-Bruhat and Geroch [17, Theorem 3], we associate with this data a maximal globally hyperbolic development

$$(M_{(\gamma,\pi)}, g),$$

which is unique up to isometry. We also fix an embedding

$$i \colon \Sigma \to M_{(\gamma,\pi)}$$

such that

$$i^* g = \gamma, \qquad i^* P = \pi,$$

where the momentum P is defined as in Theorem 5.17. Let $\mathscr{O}_{(\gamma,\pi)} \subset M$ be a neighbourhood of $i(\Sigma)$ such that $i(\Sigma)$ is a Cauchy surface for $\mathscr{O}_{(\gamma,\pi)}$. To construct the desired identification with $I \times \Sigma$, we first introduce the notion of a time function.

Definition 5.37 A differentiable function $f \colon M \to \mathbb{R}$ is called a **time function** if its gradient ∇f is timelike.

Let

$$\tau \colon \mathscr{O}_{(\gamma,\pi)} \to \mathbb{R}$$

be a time function. The level sets of τ give a foliation of $\mathscr{O}_{\gamma,\pi}$. The integral curves of $\nabla\tau$ then define a diffeomorphism

$$\xi_{(\gamma,\pi)} \colon \mathscr{O}_{(\gamma,\pi)} \to (t_-, t_+) \times \Sigma, \qquad p \mapsto (t, x),$$

where $t = \tau(p)$, and the point $x \in \Sigma$ is obtained as follows: Let θ^p denote the integral curve of $\nabla\tau$ through p. This curve intersects $i(\Sigma)$ in a unique point, say

$$\theta^p(I) \cap i(\Sigma) = \{i(x)\}.$$

We then set

$$\xi_{(\gamma,\pi)}(p) = (\tau(p), x).$$

The following Figure 5.2 illustrates this construction.

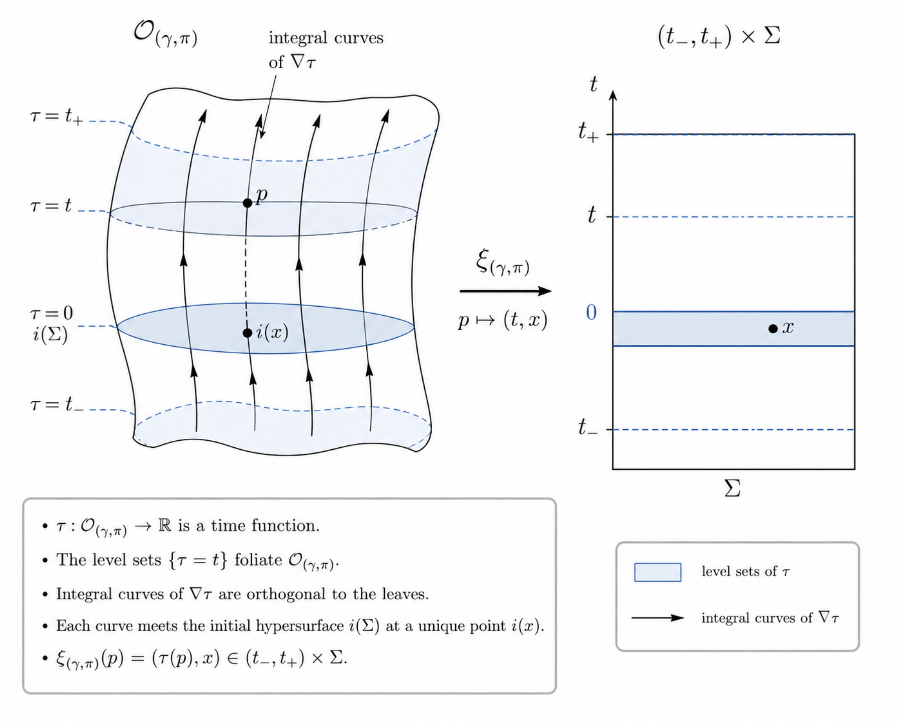

Figure 5.2: Schematic construction of the diffeomorphism $\xi_{(\gamma,\pi)} \colon \mathscr{O}_{(\gamma,\pi)} \to (t_-, t_+) \times \Sigma$. The integral curves of the timelike vector field $\nabla\tau$ intersect the initial hypersurface $i(\Sigma)$ and are used to assign to each point $p \in \mathscr{O}_{(\gamma,\pi)}$ a time coordinate $t = \tau(p)$ and a spatial point $x \in \Sigma$.

Remark 5.38 The construction is not canonical, as it depends in particular on the choice of time function and hence on the induced foliation. These uniqueness issues are related to the invariance of the ADM mass and will be discussed in Section 5.4.

The map $\xi_{(\gamma,\pi)}$ identifies a neighbourhood of the initial hypersurface with the model spacetime $(t_-, t_+) \times \Sigma$. Equivalently, its inverse

$$\psi_{(\gamma,\pi)} \colon (t_-, t_+) \times \Sigma \to \mathscr{O}_{(\gamma,\pi)} \tag{5.64}$$

defines a motion of hypersurfaces in the sense of Definition 4.2. For each $t \in (t_-, t_+)$, the map

$$\psi_t \colon \Sigma \to M_{(\gamma,\pi)}, \qquad \psi_t(x) := \psi_{(\gamma,\pi)}(t, x),$$

embeds $\Sigma \subset M$ as the hypersurface at $\tau = t$.

Thus, the time function determines a foliation of a neighbourhood of $i(\Sigma)$. Using the associated identification of each level set with Σ, the spacetime metric g induces a one-parameter family of Cauchy data on the fixed model hypersurface Σ.

In this way the chosen time function determines a curve in the space of initial data, as follows:

$$t \mapsto (\gamma_{ij}(t), \pi^{ij}(t)).$$

Here we denote $\gamma(t) := \psi_t^* g$, and $\pi(t) := \psi_t^* P$, where P is defined as in Theorem 5.17.

We now require this curve to remain in the phase space. In other words, the conditions imposed on the initial data at $t = 0$ should also be satisfied by the evolved data

$$(\gamma_{ij}(t), \pi^{ij}(t)).$$

Following Chruściel–Jezierski–Kijowski [27, p. 79], we enforce this requirement by choosing the time function so that the trace $\gamma_{ij}\pi^{ij}$ is preserved along the evolution:

$$\gamma_{ij}(t,x)\pi^{ij}(t,x) = \gamma_{ij}(0,x)\pi^{ij}(0,x). \tag{5.65}$$

We now complete the definition of the phase space $\mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$.

Definition 5.39 The **local phase space** $\mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$ consists of the initial data such that the asymptotic conditions (5.49) together with the assumptions (5.59)–(5.60) are preserved by the implicit evolution equation (5.65).

Remark 5.40 The choice of time function allows us to represent the spacetime evolution as a curve of initial data on the fixed hypersurface Σ. Since the definition of $\mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$ requires this curve to remain in the phase space, the evolution defines a dynamical system on $\mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$.

The phase space $\mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$ was constructed for a fixed reference initial data set $(\mathring{\gamma}_{ij}, \mathring{\pi}^{ij})$. Since there are many possible choices of reference data, we obtain the full asymptotically flat phase space by taking the union over all admissible reference traces.

Definition 5.41 We define the **ADM phase space** as

$$\mathscr{P}_{\mathrm{ADM}} := \bigcup_{\mathrm{tr}(\mathring{\pi})} \mathscr{P}_{\mathrm{tr}(\mathring{\pi})}, \tag{5.66}$$

where the union is taken over all initial data $(\mathring{\gamma}_{ij}, \mathring{\pi}^{ij})$ such that

$$\mathrm{tr}(\mathring{\pi}) = O(r^{-2}).$$

Remark 5.42 Chruściel, Jezierski, and Kijowski note that $\mathscr{P}_{\mathrm{ADM}}$ may be viewed as a manifold in a loose sense, with the spaces $\mathscr{P}_{\mathrm{tr}(\mathring{\pi})}$ playing the role of local coordinate patches. The idea is that, near a given reference trace $\mathrm{tr}(\mathring{\pi})$, the vacuum Einstein evolution allows one to pass from one admissible trace to nearby admissible traces. Thus the local phase spaces are not isolated objects, but overlap through the dynamics.

We will not develop this analytic manifold structure in detail. For our purposes, it is enough to use these local phase spaces as local descriptions of the full ADM phase space. The Hamiltonian will be described locally on each such phase space, and the compatibility of these local descriptions is understood in the sense explained by Chruściel, Jezierski, and Kijowski [27, p. 79].

We showed that, with each initial data set (γ, π), we can associate a reference system $\psi_{(\gamma,\pi)}$, given in (5.64). The field bundle F, the dynamical variables $(\mathfrak{g}^{\mu\nu}, p^{\lambda}_{\mu\nu})$, and the gravitational Lagrangian L have already been specified in Section 5.1. The choice of reference system therefore provides the remaining ingredient needed to apply the Hamiltonian construction of Subsection 5.1.3 to asymptotically flat manifolds. Hence, we obtain the Hamiltonian

$$H(X, \mathscr{S}) = H_{\mathrm{boundary}}(X, \mathscr{S}) + H_{\mathrm{volume}}(X, \mathscr{S}).$$

Here, the boundary and volume Hamiltonian are given by

$$H_{\mathrm{boundary}}(X, \mathscr{S}) = \frac{1}{2}\int_{\partial\mathscr{S}} \mathbb{W}^{\alpha\mu}\, dS_{\alpha\mu},$$

and

$$H_{\text{volume}}(X, \mathscr{S}) = -\int_{\mathscr{S}} \mathfrak{g}^{\mu\nu} \, \pounds_X \mathring{A}^{\lambda}{}_{\mu\nu} \, dS_\lambda,$$

where

$$\mathring{A}^{\alpha}{}_{\mu\nu} := B^{\alpha}_{\mu\nu} - \delta^{\alpha}_{(\mu} B^{\kappa}_{\nu)\kappa}.$$

Remark 5.43 Strictly speaking, the preceding construction yields a Hamiltonian on the maximal globally hyperbolic development of the initial data set (γ, π). Using Remark 5.42, these local Hamiltonian descriptions may be regarded as fitting together to define a Hamiltonian on the full ADM phase space.

In Subsection 5.3.1, we chose the background metric b to be the Minkowski metric. Hence, $B^{\alpha}_{\mu\nu}$ is the Levi–Civita connection of b. We now choose X to be a translation vector field of the form

$$X := a^\mu \partial_\mu,$$

where the coefficients a^μ are constant. Translation vector fields are Killing vector fields of the Minkowski metric b. Consequently,

$$\pounds_X B^{\alpha}_{\mu\nu} = 0.$$

Moreover,

$$\pounds_X \mathring{A}^{\lambda}{}_{\mu\nu} = \pounds_X \left(B^{\lambda}_{\mu\nu} - \delta^{\lambda}_{(\mu} B^{\kappa}_{\nu)\kappa} \right) = 0.$$

Hence, for every translation vector field X, we have

$$H_{\text{volume}}(X, \mathscr{S}) = 0.$$

Therefore, only H_{boundary} remains, when working with translation vector fields X. The conditions imposed on the initial data in $\mathscr{P}_{\text{ADM}}$ guarantee that the integral defining the variation $\delta H_{\text{boundary}}$ converges. This is made precise by the following result from Chruściel–Jezierski–Kijowski [27, p. 79].

Proposition 5.44 *Let the data belong to $\mathscr{P}_{\text{ADM}}$, and let X be a translation vector field. Then the volume integral in (5.48) is convergent, and the two boundary integrals vanish. Consequently, the variation of the boundary Hamiltonian is well-defined and is given by*

$$-\delta H_{\text{boundary}}(X, \mathscr{S}) = \frac{1}{16\pi} \int_{\Sigma} \left(\dot{\pi}^{kl} \delta\gamma_{kl} - \dot{\gamma}_{kl} \delta\pi^{kl} \right) d^3x. \tag{5.67}$$

Furthermore, one can show that the boundary Hamiltonian

$$H_{\text{boundary}}(X, \mathscr{S}) = \frac{1}{2} \int_{\partial \mathscr{S}} \mathbb{W}^{\alpha\mu} \, dS_{\alpha\mu}$$

converges; for details, see Wald [83].

Using Proposition 5.44, we establish the Hamiltonian character of $\mathscr{P}_{\text{ADM}}$.

Theorem 5.45 *Let (M, g) be an asymptotically flat $(3+1)$-dimensional spacetime, and let $\mathscr{P}_{\mathrm{ADM}}$ be the ADM phase space defined in Definition 5.41. Let $(\gamma_{ij}, \pi^{ij}) \in \mathscr{P}_{\mathrm{ADM}}$ be initial data on Σ, and let $\mathscr{S} := i(\Sigma)$ be the corresponding spacelike hypersurface in M. Suppose that the evolution is determined by a motion*

$$\psi_{(\gamma,\pi)} \colon (t_-, t_+) \times \Sigma \longrightarrow \mathscr{O}_{(\gamma,\pi)},$$

where $\mathscr{O}_{(\gamma,\pi)} \subset M$ is a spacetime neighbourhood of $\mathscr{S}$ and $\psi_{(\gamma,\pi),0} = i$.

Let X be a translation vector field, that is, $X = a^\mu \partial_\mu$, where the coefficients a^μ are constant.

Recall that we suppress the dependence on the variation parameter of the dynamical variables $(\mathfrak{g}^{\alpha\beta}, p^\lambda_{\mu\nu})$. Thus, for a differentiable one-parameter family $s \mapsto q(s)$, we write

$$\delta q := \left.\frac{d}{ds}\right|_{s=0} q(s).$$

The vector field X and the motion $\psi_{(\gamma,\pi)}$ are held fixed under the variation. Likewise, we suppress the dependence on the evolution parameter t and write

$$\dot{q} := \left.\frac{\partial q}{\partial t}\right|_{t=0}.$$

Then for data in $\mathscr{P}_{\mathrm{ADM}}$, we obtain

$$-\delta H_{\mathrm{boundary}}(X, \mathscr{S}) = \Omega_{\mathscr{S}}(\mathscr{X}, (\delta\gamma_{ij}, \delta\pi^{ij})),$$

where $\mathscr{X} = (\dot{\gamma}_{ij}, \dot{\pi}^{ij})$ is the vector field generated by the dynamics. Hence, $\mathscr{P}_{\mathrm{ADM}}$ carries the structure of a Hamiltonian dynamical system, with Hamiltonian $H_{\mathrm{boundary}}(X, \mathscr{S})$.

Proof. By the construction of the time evolution and Theorem 4.21, the dynamics determines the vector field

$$\mathscr{X} = (\dot{\gamma}_{ij}, \dot{\pi}^{ij})$$

on $\mathscr{P}_{\mathrm{ADM}}$. From Proposition 5.44, we have

$$-\delta H_{\mathrm{boundary}}(X, \mathscr{S}) = \frac{1}{16\pi} \int_\Sigma \left(\dot{\pi}^{kl} \delta\gamma_{kl} - \dot{\gamma}_{kl} \delta\pi^{kl}\right) d^3x.$$

On the other hand, Theorem 5.36 gives

$$\Omega_{\mathscr{S}}(\mathscr{X}, \delta\pi) = \frac{1}{16\pi} \int_\Sigma \left(\dot{\pi}^{kl} \delta\gamma_{kl} - \dot{\gamma}_{kl} \delta\pi^{kl}\right) d^3x.$$

These two expressions agree. Hence

$$-\delta H_{\mathrm{boundary}}(X, \mathscr{S}) = \Omega_{\mathscr{S}}(\mathscr{X}, \delta\pi),$$

which is the Hamiltonian relation. Therefore $\mathscr{P}_{\mathrm{ADM}}$ carries the structure of a dynamical system. $\square$

5.3.4 Calculation of ADM four-momentum

The previous subsection showed that the boundary Hamiltonian gives rise to a Hamiltonian dynamical system on $\mathscr{P}_{\text{ADM}}$. We now use this Hamiltonian to define the energy-momentum of the system.

Definition 5.46 We define the four-momentum p^μ as

$$p_\mu a^\mu := -H_{\text{boundary}}(a^\mu \partial_\mu, \mathscr{S}),$$

for every constant vector field $a^\mu \partial_\mu$.

We verify that this four-momentum coincides with the Arnowitt–Deser–Misner four-momentum defined in [4, 5]. We do this by evaluating the boundary Hamiltonian on the vector fields ∂_μ separately. Because the computation is rather long, this subsection gives a sketch of the argument and records the main steps. Subsection 5.3.5 completes the proofs by calculating the expressions in full.

In the proofs below, Greek indices range from 0 to 3, while Latin indices range from 1 to 3. Furthermore, we denote

$$dS_{\alpha\mu} := (\partial_\alpha \wedge \partial_\mu) \rfloor dx^0 \wedge \cdots \wedge dx^3. \tag{5.68}$$

Theorem 5.47 *Let $\mathscr{S}$ be an asymptotically flat hypersurface. We choose coordinates for which x^0 is constant on $\mathscr{S}$. Then the ADM energy, or first component of the ADM energy-momentum vector, agrees with $-H_{\text{boundary}}(\partial_0, \mathscr{S})$, explicitly*

$$-H_{\text{boundary}}(\partial_0, \mathscr{S}) = E_{\text{ADM}} := \lim_{R\to\infty} \frac{1}{16\pi} \int_{S(R)} (g_{ij,j} - g_{jj,i}) d\sigma_i, \tag{5.69}$$

where $d\sigma_i := dS_{0i}$, as defined in (5.68), and $S(R)$ is the 2-sphere with radius R.

Sketch of the proof. We prove the theorem by calculating

$$-H_{\text{boundary}}(\partial_0, \mathscr{S}) = -\frac{1}{2} \int_{\partial \mathscr{S}} \mathbb{W}^{\alpha\mu} dS_{\alpha\mu}, \tag{5.70}$$

where $dS_{\alpha\mu}$ is defined as

$$dS_{\alpha\mu} := (\partial_\alpha \wedge \partial_\mu) \rfloor dx^0 \wedge \cdots \wedge dx^3.$$

The details of the calculation are outlined in the proof of Theorem 5.52. The calculation involves rewriting Christoffel symbols and using the asymptotic conditions to approximate the metric to get the ADM formula.

Step 1. *Reduction of the superpotential.* The integral (5.70) consists of 16 components. Since x^0 is constant on $\mathscr{S}$, the only nonzero terms are $\mathbb{W}^{0i} dS_{0i}$ and $\mathbb{W}^{i0} dS_{i0}$, for $i = 1, 2, 3$. By antisymmetry of both $\mathbb{W}^{0i}$ and dS_{0i}, we only need to calculate $\mathbb{W}^{0i} dS_{0i}$. Furthermore, since the background metric b is the Minkowski metric in the asymptotic coordinates, the corresponding connection coefficients $B^\mu_{\alpha\beta}$ vanish. Hence the superpotential reduces to

$$\mathbb{W}^{0i} = \left(2\mathfrak{g}^{\mu[0} p^{i]}_{\mu 0} - 2\delta^{[0}_0 p^{i]}_{\mu\sigma} \mathfrak{g}^{\mu\sigma} - \frac{2}{3} \mathfrak{g}^{\mu[0} \delta^{i]}_0 p^\sigma_{\mu\sigma} \right). \tag{5.71}$$

Step 2. *Christoffel symbol manipulations.* We now rewrite $\mathbb{W}^{i0}$. First, we expand the antisymmetrization in (5.71), we find

$$\mathbb{W}^{0i} = \mathfrak{g}^{\mu 0} p^i_{\mu 0} - \mathfrak{g}^{\mu i} p^0_{\mu 0} - \mathfrak{g}^{\mu\sigma} p^i_{\mu\sigma} + \frac{1}{3}\mathfrak{g}^{\mu i} p^\sigma_{\mu\sigma}.$$

After calculating this termwise we obtain the formula

$$\mathbb{W}^{0i} = \mathfrak{g}^{\mu j}\Gamma^i_{\mu j} - \mathfrak{g}^{\mu i}\Gamma^j_{\mu j},$$

where $j = 1, 2, 3$.

Step 3. *Using the asymptotic conditions to approximate* $\mathbb{W}^{i0}$. Recall that in the asymptotic end the metric $g = \eta + h$, and it satisfies the asymptotic conditions (5.49). Using Lemmas 5.50 and 5.51, we approximate

$$\mathfrak{g}^{\mu\nu}\Gamma^\lambda_{\rho\sigma} = \frac{1}{16\pi}\eta^{\mu\nu}\Gamma^\lambda_{\rho\sigma} + O(h\partial h).$$

Since the Minkowski metric acts as a Kronecker δ when the indices are greater than zero, we approximate the potential by

$$\mathbb{W}^{0i} = \frac{1}{16\pi}(\Gamma^i_{jj} - \Gamma^j_{ij}) + O(h\partial h).$$

Rewriting the Christoffel symbols in terms of the metric and using Lemma 5.49 gives

$$\mathbb{W}^{0i} = \frac{1}{16\pi}\left(\partial_j g_{ij} - \partial_i g_{jj}\right) + O(h\partial h).$$

The antisymmetry of both the potential $\mathbb{W}^{0i}$ and the odd form dS_{0i} give

$$\mathbb{W}^{i0} dS_{i0} = -\mathbb{W}^{0i}(-dS_{0i}) = \mathbb{W}^{0i} dS_{0i}.$$

Since $O(h\partial h) = O(r^{-3})$, the integral over $O(h\partial h)$ vanishes at infinity. Subsection 5.3.3 already addressed the issue of convergence of these integrals. Hence, we find

$$\begin{aligned} E_{\mathrm{ADM}} &= \frac{1}{2}\int_{\partial\mathscr{S}} \mathbb{W}^{\alpha\mu} dS_{\alpha\mu} \\ &= \frac{1}{16\pi}\int_{\partial\mathscr{S}} \partial_j g_{ij} - \partial_i g_{jj} + O(h\partial h) dS_{0i} \\ &= \lim_{R\to\infty}\frac{1}{16\pi}\int_{S(R)} \left(g_{ij,j} - g_{jj,i}\right) dS_{0i}. \end{aligned}$$

□

The remaining components of the four-momentum are obtained by applying the same boundary Hamiltonian formula to the spatial vector fields ∂_l. The calculation is similar in spirit to the one in the proof of Theorem 5.47.

Theorem 5.48 *Let $\mathscr{S}$ be an asymptotically flat hypersurface. We choose coordinates for which x^0 is constant on $\mathscr{S}$. Then the ADM momentum P^l_{ADM} agrees with $-H_{\mathrm{boundary}}(\partial_l, \mathscr{S})$, explicitly*

$$-H_{\mathrm{boundary}}(\partial_l, \mathscr{S}) = p^l_{\mathrm{ADM}} := \lim_{R\to\infty}\frac{1}{8\pi}\int_{S(R)} P^{lj} d\tilde{\sigma}_j, \tag{5.72}$$

where $d\tilde{\sigma}_j := -dS_{0j}$, as defined in (5.68), and $S(R)$ is the 2-sphere with radius R.

Sketch of proof. The momentum component is obtained by evaluating the boundary Hamiltonian on the vector field ∂_l. Thus we calculate

$$-H_{\text{boundary}}(\partial_l, \mathscr{S}) = -\frac{1}{2}\int_{\partial\mathscr{S}} \mathbb{W}^{\alpha\mu} dS_{\alpha\mu}. \tag{5.73}$$

The details of the calculation are outlined in the proof of Theorem 5.54. The computation is similar to the one in Theorem 5.47, but now the vector field is ∂_l, with $l = 1, 2, 3$, rather than ∂_0.

Step 1. *Reduction of the superpotential.* As in the proof of Theorem 5.47, the antisymmetry of $\mathbb{W}^{\alpha\mu}$ and $dS_{\alpha\mu}$ reduces the calculation to the components $\mathbb{W}^{0i} dS_{0i}$. For $X = \partial_l$, the reduced expression for the superpotential is

$$\mathbb{W}^{0i} = \mathfrak{g}^{\mu 0} p^i_{\mu l} - \mathfrak{g}^{\mu i} p^0_{\mu l} + \delta^i_l \left(\mathfrak{g}^{\mu\sigma} p^0_{\mu\sigma} - \frac{1}{3}\mathfrak{g}^{\mu 0} p^\sigma_{\mu\sigma} \right).$$

Here the factor δ^i_l cannot be simplified further, since i is summed over the forms dS_{0i} while l denotes the fixed spatial translation direction.

Step 2. *Christoffel symbol manipulations.* Rewriting the momenta $p^\lambda_{\mu\nu}$ in terms of Christoffel symbols gives

$$\mathbb{W}^{0i} = -\mathfrak{g}^{\mu 0}\Gamma^i_{\mu l} + \mathfrak{g}^{\mu i}\Gamma^0_{\mu l} + \delta^i_l(-\mathfrak{g}^{\mu\sigma}\Gamma^0_{\mu\sigma} + \mathfrak{g}^{\mu 0}\Gamma^\kappa_{\mu\kappa}).$$

Step 3. *Using the asymptotic conditions to approximate* $\mathbb{W}^{i0}$. Using Lemma 5.51 to approximate the metric density together with the fact that the Minkowski metric acts as a Kronecker delta on spatial indices, we obtain

$$\mathbb{W}^{0i} = \frac{1}{16\pi}(\Gamma^i_{0l} + \Gamma^0_{il} + \delta^i_l(\Gamma^0_{00} - \Gamma^0_{jj} - \Gamma^\mu_{0\mu})) + O(h\partial h)$$

where $j = 1, 2, 3$. We use Lemma 5.50 to approximate the Christoffel symbols. This gives

$$\mathbb{W}^{0i} = \frac{1}{16\pi}\left[\partial_0 g_{il} - \partial_i g_{0l} + \delta^i_l(\partial_j g_{0j} - \partial_0 g_{jj})\right] + O(h\partial h).$$

On the other hand, using Lemma 5.53 we approximate the ADM momentum P^{il} by

$$P^{il} = \delta^i_l(\partial_j g_{j0} - \frac{1}{2}\partial_0 g_{jj}) + \frac{1}{2}(-\partial_i g_{l0} - \partial_l g_{i0} + \partial_0 g_{il}) + O(h\partial h)$$

Hence, we express the potential $\mathbb{W}^{0i}$ in terms of the ADM momentum P^{il} as

$$16\pi\mathbb{W}^{0i} = 2P^{il} + \partial_j(\delta^j_l g_{i0} - \delta^i_l g_{j0}) + O(h\partial h)$$

Using both the antisymmetry of $dS_{\alpha\mu}$ and $\mathbb{W}^{\alpha\mu}$, we obtain

$$\begin{aligned} p^l &= -\int_{\partial\mathscr{S}} \mathbb{W}^{0i} dS_{0i} \\ &= -\frac{1}{8\pi}\int_{\partial\mathscr{S}} P^{il} + O(h\partial h) dS_{0i} + \frac{1}{16\pi}\int_{\partial\mathscr{S}} \partial_j(\delta^j_l g_{i0} - \delta^i_l g_{j0}) dS_{0i} \\ &= \lim_{R\to\infty}\frac{1}{8\pi}\int_{\partial\mathscr{S}} P^{il} d\tilde{\sigma}_i. \end{aligned}$$

The second integral vanishes in the last equality due to the boundary of the boundary being empty, and the $O(h\partial h)$ term has vanishing surface integral under the ADM fall-off assumptions. This completes the proof. □

5.3.5 Details of the ADM four-momentum calculation

This subsection fills in the gaps left in the proof sketches of the two theorems in Subsection 5.3.4. We also prove the auxiliary lemmas that were used there. These lemmas mainly provide the asymptotic approximations for the quantities appearing in the boundary Hamiltonian.

We prove the lemmas first, before giving the complete proofs of the theorems. This order allows the final calculations to proceed without interrupting them with repeated approximation arguments.

Lemma 5.49 *In asymptotic coordinates, suppose that*

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu},$$

where $h_{\mu\nu} = O(r^{-1})$. *Then*

$$g^{\lambda\rho} = \eta^{\lambda\rho} + O(h).$$

More precisely,

$$g^{\lambda\rho} = \eta^{\lambda\rho} - \eta^{\lambda\mu}\eta^{\rho\nu}h_{\mu\nu} + O(h^2).$$

Proof. In matrix notation, we write

$$g = \eta + h = \eta(I + \eta^{-1}h).$$

Set

$$A := \eta^{-1}h.$$

Since $h = O(r^{-1})$, we have $A = O(r^{-1})$. Hence, all eigenvalues satisfy $|\lambda_i| < 1$. We approximate $(I + A)^{-1}$ with the Neumann series; see [66, Section 3.4]. This gives

$$(I + A)^{-1} = I - A + A^2 - \cdots .$$

Thus

$$g^{-1} = (I + A)^{-1}\eta^{-1} = (I - A + O(A^2))\eta^{-1}.$$

Since $A = \eta^{-1}h$, we obtain

$$g^{-1} = \eta^{-1} - \eta^{-1}h\eta^{-1} + O(h^2).$$

In components, this is

$$g^{\lambda\rho} = \eta^{\lambda\rho} - \eta^{\lambda\mu}\eta^{\rho\nu}h_{\mu\nu} + O(h^2).$$

In particular,

$$g^{\lambda\rho} = \eta^{\lambda\rho} + O(h).$$

□

Lemma 5.50 *In asymptotic coordinates, the Christoffel symbols satisfy*

$$\Gamma^{\lambda}_{\mu\nu} = O(\partial h)$$

Proof. Recall that the Christoffel symbols are given by

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}(\partial_\mu g_{\rho\nu} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}).$$

Using Lemma 5.49 to approximate the inverse metric, we obtain

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}(\eta^{\lambda\rho} + O(h))(\partial_\mu g_{\rho\nu} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu})$$

Since $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ and $\eta_{\mu\nu}$ is constant in the asymptotic coordinates, we have

$$\eta^{\lambda\rho}\partial_\mu g_{\rho\nu} = O(\partial h), \qquad O(h)\partial_\mu g_{\rho\nu} = O(h\partial h).$$

Hence, we find

$$\Gamma^{\lambda}_{\mu\nu} = O(\partial h) + O(h\partial h) = O(\partial h).$$

□

Lemma 5.51 *In asymptotic coordinates, the metric density satisfies*

$$\mathfrak{g}^{\mu\nu} = \frac{1}{16\pi}\eta^{\mu\nu} + O(h)$$

Proof. We write the metric in matrix form

$$g = \eta + h = \eta(I + \eta^{-1}h).$$

Set $A := \eta^{-1}h$. As in the proof of Lemma 5.49, we have

$$A = O(h) = O(r^{-1}).$$

Taking determinants gives

$$\begin{aligned} -\det(g) &= -\det(\eta)\det(I + A) \\ &= 1 + \operatorname{tr}(A) + O(A^2) \\ &= 1 + \operatorname{tr}(A) + O(h^2). \end{aligned}$$

Since $\operatorname{tr}(A) = O(h)$, the Taylor expansion

$$\sqrt{1+x} = 1 + \frac{1}{2}x + O(x^2)$$

gives

$$\sqrt{-\det(g)} = 1 + \frac{1}{2}\operatorname{tr}(A) + O(h^2) = 1 + O(h).$$

By Lemma 5.49, we also have

$$g^{\mu\nu} = \eta^{\mu\nu} + O(h).$$

Substituting these estimates into the definition of the metric density gives

$$\mathfrak{g}^{\mu\nu} = \frac{1}{16\pi}\sqrt{-\det(g)}\, g^{\mu\nu} = \frac{1}{16\pi}(1 + O(h))(\eta^{\mu\nu} + O(h)).$$

Therefore

$$\mathfrak{g}^{\mu\nu} = \frac{1}{16\pi}\eta^{\mu\nu} + O(h).$$

□

We use Lemmas 5.49 to 5.51 to prove the following theorem.

Theorem 5.52 *Let $\mathscr{S}$ be an asymptotically flat hypersurface. We choose coordinates for which x^0 is constant on $\mathscr{S}$. Then the ADM energy, or first component of the ADM energy-momentum vector, agrees with $-H_{\text{boundary}}(\partial_0, \mathscr{S})$, explicitly*

$$-H_{\text{boundary}}(\partial_0, \mathscr{S}) = E_{\text{ADM}} := \lim_{R\to\infty} \frac{1}{16\pi} \int_{S(R)} (g_{ij,j} - g_{jj,i}) d\sigma_i, \tag{5.74}$$

where $d\sigma_i := dS_{0i}$, as defined in (5.68), and $S(R)$ is the 2-sphere with radius R.

Proof. We follow the three steps outlined in the proof sketch.

Step 1. *Reduction of the superpotential.* We compute

$$H_{\text{boundary}}(\partial_0, \mathscr{S}) = \frac{1}{2} \int_{\partial\mathscr{S}} \mathbb{W}^{\alpha\mu} dS_{\alpha\mu}.$$

This integral is a sum over the components $\mathbb{W}^{\alpha\mu} dS_{\alpha\mu}$. Recall

$$dS_{\alpha\mu} := (\partial_\alpha \wedge \partial_\mu) \rfloor \left(dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3\right).$$

Since x^0 is constant on $\mathscr{S}$, the only components contributing on $\partial\mathscr{S}$ are the components dS_{0i} and dS_{i0}, with $i = 1, 2, 3$. By antisymmetry of both $\mathbb{W}^{\alpha\mu}$ and $dS_{\alpha\mu}$, it suffices to compute only the terms $\mathbb{W}^{0i} dS_{0i}$.

Since the background metric b is the Minkowski metric in the asymptotic coordinates, the corresponding background connection has vanishing connection coefficients. Thus

$$X^\beta{}_{;\alpha} = \lambda_\alpha(\partial_0) = 0,$$

and (5.23), the formula for $\mathbb{W}^{0i}$, reduces to

$$\begin{aligned}
\mathbb{W}^{0i} &= \left(2\mathfrak{g}^{\mu[0} p^{i]}_{\mu 0} - 2\delta^{[0}_0 p^{i]}_{\mu\sigma} \mathfrak{g}^{\mu\sigma} - \frac{2}{3} \mathfrak{g}^{\mu[0} \delta^{i]}_0 p^\sigma_{\mu\sigma}\right) \cdot 1 - 0 \\
&= \mathfrak{g}^{\mu 0} p^i_{\mu 0} - \mathfrak{g}^{\mu i} p^0_{\mu 0} - \left(\delta^0_0 p^i_{\mu\sigma} \mathfrak{g}^{\mu\sigma} - \underbrace{\delta^i_0 p^0_{\mu\sigma} \mathfrak{g}^{\mu\sigma}}_{=0}\right) \\
&\quad - \frac{1}{3}\left(\underbrace{\mathfrak{g}^{\mu 0} \delta^i_0 p^\sigma_{\mu\sigma}}_{=0} - \mathfrak{g}^{\mu i} \delta^0_0 p^\sigma_{\mu\sigma}\right) \\
&= \mathfrak{g}^{\mu 0} p^i_{\mu 0} - \mathfrak{g}^{\mu i} p^0_{\mu 0} - \mathfrak{g}^{\mu\sigma} p^i_{\mu\sigma} + \frac{1}{3} \mathfrak{g}^{\mu i} p^\sigma_{\mu\sigma}.
\end{aligned} \tag{5.75}$$

Step 2. *Christoffel symbol manipulations.* To keep the notation readable, set

$$T_\mu := \Gamma^\kappa_{\mu\kappa}.$$

We compute (5.75) term by term.

For the first term, we have

$$\begin{aligned}
\mathfrak{g}^{\mu 0} p^i_{\mu 0} &= \mathfrak{g}^{\mu 0}\big[-\Gamma^i_{\mu 0} + \frac{1}{2}\delta^i_\mu T_0 + \underbrace{\frac{1}{2}\delta^i_0 T_\mu}_{=0}\big] \\
&= -\mathfrak{g}^{\mu 0} \Gamma^i_{\mu\ 0} + \frac{1}{2} \mathfrak{g}^{\mu 0} \delta^i_\mu T_0 \\
&= -\mathfrak{g}^{\mu 0} \Gamma^i_{\mu\ 0} + \frac{1}{2} \mathfrak{g}^{i0} T_0.
\end{aligned}$$

For the second term, we have

$$\begin{aligned}-\mathfrak{g}^{\mu i}p^0_{\mu 0} &= \mathfrak{g}^{\mu i}\left[\Gamma^0_{\mu 0} - \frac{1}{2}\delta^0_\mu T_0 - \frac{1}{2}\delta^0_0 T_\mu\right]\\ &= \mathfrak{g}^{\mu i}\Gamma^0_{\mu 0} - \frac{1}{2}\mathfrak{g}^{0i}T_0 - \frac{1}{2}\mathfrak{g}^{\mu i}T_\mu.\end{aligned}$$

For the third term, we have

$$\begin{aligned}-\mathfrak{g}^{\mu\sigma}p^i_{\mu\sigma} &= \mathfrak{g}^{\mu\sigma}\left[\Gamma^i_{\mu\sigma} - \frac{1}{2}\delta^i_\mu T_\sigma - \frac{1}{2}\delta^i_\sigma T_\mu\right]\\ &= \mathfrak{g}^{\mu\sigma}\Gamma^i_{\mu\sigma} - \frac{1}{2}\mathfrak{g}^{i\sigma}T_\sigma - \frac{1}{2}\mathfrak{g}^{\mu i}T_\mu\\ &= \mathfrak{g}^{\mu\sigma}\Gamma^i_{\mu\sigma} - \mathfrak{g}^{\mu i}T_\mu.\end{aligned}$$

For the fourth term, we have

$$\begin{aligned}\frac{1}{3}\mathfrak{g}^{\mu i}p^\sigma_{\mu\sigma} &= \frac{1}{3}\mathfrak{g}^{\mu i}\left[-T_\mu + \frac{1}{2}\delta^\sigma_\mu T_\sigma + \frac{1}{2}\delta^\sigma_\sigma T_\mu\right]\\ &= \frac{1}{3}\mathfrak{g}^{\mu i}\left[-T_\mu + \frac{1}{2}T_\mu + \frac{1}{2}4T_\mu\right]\\ &= \frac{1}{2}\mathfrak{g}^{\mu i}T_\mu.\end{aligned}$$

Substituting all the terms back into (5.75) gives

$$\begin{aligned}\mathbb{W}^{0i} &= \mathfrak{g}^{\mu 0}p^i_{\mu 0} - \mathfrak{g}^{\mu i}p^0_{\mu 0} - \mathfrak{g}^{\mu\sigma}p^i_{\mu\sigma} + \frac{1}{3}\mathfrak{g}^{\mu i}p^\sigma_{\mu\sigma}\\ &= (-\mathfrak{g}^{\mu 0}\Gamma^i_{\mu\ 0} + \frac{1}{2}\mathfrak{g}^{i0}T_0) + (\mathfrak{g}^{\mu i}\Gamma^0_{\mu 0} - \frac{1}{2}\mathfrak{g}^{0i}T_0 - \frac{1}{2}\mathfrak{g}^{\mu i}T_\mu)\\ &\quad + (\mathfrak{g}^{\mu\sigma}\Gamma^i_{\mu\sigma} - \mathfrak{g}^{\mu i}T_\mu) + (\frac{1}{2}\mathfrak{g}^{\mu i}T_\mu)\\ &= -\mathfrak{g}^{\mu 0}\Gamma^i_{\mu\ 0} + \mathfrak{g}^{\mu i}\Gamma^0_{\mu 0} + \mathfrak{g}^{\mu\sigma}\Gamma^i_{\mu\sigma} - \mathfrak{g}^{\mu i}T_\mu.\end{aligned}$$

Splitting the time and space indices, we obtain

$$\begin{aligned}\mathbb{W}^{0i} &= -\mathfrak{g}^{\mu 0}\Gamma^i_{\mu\ 0} + \mathfrak{g}^{\mu i}\Gamma^0_{\mu 0} + \mathfrak{g}^{\mu\sigma}\Gamma^i_{\mu\sigma} - \mathfrak{g}^{\mu i}\Gamma^\kappa_{\mu\kappa}\\ &= \mathfrak{g}^{\mu j}\Gamma^i_{\mu j} - \mathfrak{g}^{\mu i}\Gamma^j_{\mu j},\end{aligned}$$

where $j = 1, 2, 3$. **Step 3.** *Using the asymptotic conditions to approximate* $\mathbb{W}^{0i}$. We approximate the metric density and the Christoffel symbols using Lemmas 5.50 and 5.51, respectively. We obtain

$$\begin{aligned}\mathfrak{g}^{\mu\nu}\Gamma^\lambda_{\rho\sigma} &= (\frac{1}{16\pi}\eta^{\mu\nu} + O(h))\Gamma^\lambda_{\rho\sigma}\\ &= \frac{1}{16\pi}\eta^{\mu\nu}\Gamma^\lambda_{\rho\sigma} + O(h)\Gamma^\lambda_{\rho\sigma}\\ &= \frac{1}{16\pi}\eta^{\mu\nu}\Gamma^\lambda_{\rho\sigma} + O(h\partial h).\end{aligned}$$

Hence, we approximate $\mathbb{W}^{0i}$ by

$$\mathbb{W}^{0i} = \frac{1}{16\pi}(\eta^{\mu j}\Gamma^i_{\mu j} - \eta^{\mu i}\Gamma^j_{\mu j}) + O(h\partial h).$$

Since i and j are spatial indices and η is the Minkowski metric, we obtain

$$\begin{aligned}\mathbb{W}^{0i} &= \frac{1}{16\pi}(\delta^\mu_j\Gamma^i_{\mu j} - \delta^\mu_i\Gamma^j_{\mu j}) + O(h\partial h) \\ &= \frac{1}{16\pi}(\Gamma^i_{jj} - \Gamma^j_{ij}) + O(h\partial h). \end{aligned} \tag{5.76}$$

We now compute the two Christoffel symbols appearing in (5.76). Using Lemma 5.49 to approximate the inverse metric and the fact that $i > 0$ yields

$$\begin{aligned}\Gamma^i_{jj} &= \frac{1}{2}g^{i\sigma}\left(\partial_j g_{\sigma j} + \partial_j g_{\sigma j} - \partial_\sigma g_{jj}\right) \\ &= \frac{1}{2}g^{i\sigma}\left(2\partial_j g_{\sigma j} - \partial_\sigma g_{jj}\right) \\ &= \delta^i_\sigma\partial_j g_{\sigma j} - \frac{1}{2}\delta^i_\sigma\partial_\sigma g_{jj} + O(h\partial h) \\ &= \partial_j g_{ij} - \frac{1}{2}\partial_i g_{jj} + O(h\partial h), \end{aligned} \tag{5.77}$$

and

$$\begin{aligned}\Gamma^j_{ij} &= \frac{1}{2}g^{j\sigma}\left(\partial_i g_{\sigma j} + \partial_j g_{\sigma i} - \partial_\sigma g_{ij}\right) \\ &= \frac{1}{2}\left(\delta^j_\sigma\partial_i g_{\sigma j} + \delta^j_\sigma\partial_j g_{\sigma i} - \delta^j_\sigma\partial_\sigma g_{ij}\right) + O(h\partial h) \\ &= \frac{1}{2}\left(\partial_i g_{jj} + \partial_j g_{ji} - \partial_j g_{ij}\right) + O(h\partial h) \\ &= \frac{1}{2}\partial_i g_{jj} + O(h\partial h). \end{aligned} \tag{5.78}$$

Substituting (5.77) and (5.78) into (5.76) gives

$$\begin{aligned}\mathbb{W}^{0i} &= \frac{1}{16\pi}\left(\partial_j g_{ij} - \frac{1}{2}\partial_i g_{jj} - \frac{1}{2}\partial_i g_{jj} + O(h\partial h)\right) + O(h\partial h) \\ &= \frac{1}{16\pi}\left(\partial_j g_{ij} - \partial_i g_{jj}\right) + O(h\partial h).\end{aligned}$$

Using the antisymmetry of both $\mathbb{W}^{\alpha\mu}$ and $dS_{\alpha\mu}$, we have

$$\mathbb{W}^{i0}dS_{i0} = -\mathbb{W}^{0i}(-dS_{0i}) = \mathbb{W}^{0i}dS_{0i}.$$

Hence we find the formula

$$\begin{aligned}p^0 = -p_0 &= H_{\text{boundary}}(\partial_0, \mathscr{S}) \\ &= \frac{1}{2}\int_{\partial\mathscr{S}}\mathbb{W}^{\alpha\mu}dS_{\alpha\mu} \\ &= \frac{1}{16\pi}\int_{\partial\mathscr{S}}\left(\partial_j g_{ij} - \partial_i g_{jj} + O(h\partial h)\right)dS_{0i} \\ &= \lim_{R\to\infty}\frac{1}{16\pi}\int_{S(R)}\left(g_{ij,j} - g_{jj,i}\right)dS_{0i}. \end{aligned} \tag{5.79}$$

The integral of the error term vanishes in the limit, since $O(h\partial h) = O(r^{-3})$. This completes the proof. $\square$

Before giving the Hamiltonian momentum, we prove an auxiliary lemma that approximates the ADM momentum using the asymptotic conditions.

Lemma 5.53 *In asymptotic coordinates, the ADM momentum satisfies*

$$P^{il} = \delta^{li}(\partial_j g_{j0} - \frac{1}{2}\partial_0 g_{jj}) - \frac{1}{2}(\partial_i g_{l0} + \partial_l g_{i0} - \partial_0 g_{il}) + O(h\partial h)$$

Proof. Recall from Theorem 5.17 that

$$P^{kl} := \sqrt{\det g_{mn}}({}^3g^{ij}K_{ij}{}^3g^{kl} - K^{kl}),$$

where ${}^3g^{ij}$ denotes the induced inverse metric on $\mathscr{S}$ and K_{ij} is given by

$$K_{ij} = -\frac{1}{\sqrt{|g^{00}|}}\Gamma^0_{ij}. \tag{5.80}$$

We first estimate the Christoffel symbol

$$\begin{aligned}
\Gamma^0_{ij} &= \frac{1}{2}(g^{0\rho})(\partial_i g_{j\rho} + \partial_j g_{i\rho} - \partial_\rho g_{ij}) \\
&= \frac{1}{2}(\eta^{0\rho} + O(h))(\partial_i g_{j\rho} + \partial_j g_{i\rho} - \partial_\rho g_{ij}) \\
&= -\frac{1}{2}\delta^0_\rho(\partial_i g_{j\rho} + \partial_j g_{i\rho} - \partial_\rho g_{ij}) + O(h\partial h) \\
&= -\frac{1}{2}(\partial_i g_{j0} + \partial_j g_{i0} - \partial_0 g_{ij}) + O(h\partial h).
\end{aligned}$$

Since $g^{00} = -1 + O(h)$, the Taylor expansion

$$(1+x)^{-1/2} = 1 - \frac{1}{2}x + O(x^2)$$

gives

$$\frac{1}{\sqrt{|g^{00}|}} = 1 + O(h).$$

Substituting this into (5.80), we obtain

$$\begin{aligned}
K_{ij} &= -(1 + O(h))\left[-\frac{1}{2}(\partial_i g_{j0} + \partial_j g_{i0} - \partial_0 g_{ij}) + O(h\partial h)\right] \\
&= \frac{1}{2}(\partial_i g_{j0} + \partial_j g_{i0} - \partial_0 g_{ij}) + O(h\partial h).
\end{aligned}$$

We now approximate the trace term $K = {}^3g^{ij}K_{ij}$. The induced metric gives

$${}^3g^{ij} = \delta^{ij} + O(h).$$

The leading term of K_{ij} is $O(\partial h)$. Thus

$$\begin{aligned}
K &= (\delta^{ij} + O(h))K_{ij} \\
&= \delta^{ij}K_{ij} + O(h\partial h),
\end{aligned}$$

Thus, we find

$$K = \frac{1}{2}(2\partial_j g_{j0} - \partial_0 g_{jj}) + O(h\partial h),$$

where $j = 1, 2, 3$.

We also approximate K^{ij}

$$\begin{aligned} K^{ji} &= {}^3g^{jm3}g^{in}K_{mn} \\ &= (\delta^{jm} + O(h))(\delta^{in})K_{mn} \\ &= K_{ij} + O(h\partial h). \end{aligned}$$

From the proof of Lemma 5.51, we have

$$\sqrt{\det g_{mn}} = 1 + O(h).$$

Thus

$$\begin{aligned} P^{li} &= \sqrt{\det g}(K^3g^{li} - K^{li}) \\ &= (1 + O(h))(K(\delta^{li} + O(h)) - K_{li} + O(h\partial h)) \\ &= \delta^{li}K - K_{li} + O(h\partial h). \end{aligned}$$

Substituting the expressions for K and K_{li} gives

$$P^{il} = \delta^{li}(\partial_j g_{j0} - \frac{1}{2}\partial_0 g_{jj}) - \frac{1}{2}(\partial_i g_{l0} + \partial_l g_{i0} - \partial_0 g_{il}) + O(h\partial h).$$

□

We now prove that the ADM momentum coincides with the momentum defined by the boundary Hamiltonian.

Theorem 5.54 *Let $\mathscr{S}$ be an asymptotically flat hypersurface. We choose coordinates for which x^0 is constant on $\mathscr{S}$. Then the ADM momentum P^l_{ADM} agrees with $-H_{\text{boundary}}(\partial_l, \mathscr{S})$, explicitly*

$$-H_{\text{boundary}}(\partial_l, \mathscr{S}) = p^l_{\text{ADM}} := \lim_{R\to\infty} \frac{1}{8\pi}\int_{S(R)} P^{lj} d\sigma_j, \tag{5.81}$$

where $d\tilde{\sigma}_j := -dS_{0j}$, as defined in (5.68), and $S(R)$ is the 2-sphere with radius R.

Proof. Step 1. *Reduction of the superpotential.* We use the same assumptions as in Theorem 5.52. Hence, it is enough to compute $\mathbb{W}^{0i}$, defined in (5.23), for the vector field $X = \partial_l$, where $l = 1, 2, 3$ and $i = 1, 2, 3$. We find

$$\begin{aligned} \mathbb{W}^{0i} &= \left(2\mathfrak{g}^{\mu[0}p^{i]}_{\mu l} - 2\delta^{[0}_l p^{i]}_{\mu\sigma}\mathfrak{g}^{\mu\sigma} - \frac{2}{3}\mathfrak{g}^{\mu[0}\delta^{i]}_l p^\sigma_{\mu\sigma}\right) \\ &= \mathfrak{g}^{\mu 0}p^i_{\mu l} - \mathfrak{g}^{\mu i}p^0_{\mu l} - \left(\underbrace{\delta^0_l p^i_{\mu\sigma}\mathfrak{g}^{\mu\sigma}}_{=0} - \delta^i_l p^0_{\mu\sigma}\mathfrak{g}^{\mu\sigma}\right) \\ &\quad - \frac{1}{3}\left(\mathfrak{g}^{\mu 0}\delta^i_l p^\sigma_{\mu\sigma} - \underbrace{\delta^0_l \mathfrak{g}^{\mu i}p^\sigma_{\mu\sigma}}_{=0}\right) \\ &= \mathfrak{g}^{\mu 0}p^i_{\mu l} - \mathfrak{g}^{\mu i}p^0_{\mu l} + \delta^i_l\left(\mathfrak{g}^{\mu\sigma}p^0_{\mu\sigma} - \frac{1}{3}\mathfrak{g}^{\mu 0}p^\sigma_{\mu\sigma}\right). \end{aligned} \tag{5.82}$$

Step 2. *Christoffel symbol manipulations.* We compute (5.82) term by term. The first term

$$\begin{aligned}\mathfrak{g}^{\mu 0}p^i_{\mu l} &= \mathfrak{g}^{\mu 0}\left[-\Gamma^i_{\mu l} + \frac{1}{2}\delta^i_\mu T_l + \frac{1}{2}\delta^i_l T_\mu\right] \\ &= -\mathfrak{g}^{\mu 0}\Gamma^i_{\mu l} + \frac{1}{2}\mathfrak{g}^{i0}T_l + \frac{1}{2}\delta^i_l \mathfrak{g}^{\mu 0}T_\mu.\end{aligned}$$

The second term

$$\begin{aligned}-\mathfrak{g}^{\mu i}p^0_{\mu l} &= \mathfrak{g}^{\mu i}\left[\Gamma^0_{\mu l} - \frac{1}{2}\delta^0_\mu T_l - \frac{1}{2}\delta^0_l T_\mu\right] \\ &= \mathfrak{g}^{\mu i}\Gamma^0_{\mu l} - \frac{1}{2}\mathfrak{g}^{0i}T_l - \underbrace{\frac{1}{2}\delta^0_l \mathfrak{g}^{\mu i}T_\mu}_{=0} \\ &= \mathfrak{g}^{\mu i}\Gamma^0_{\mu l} - \frac{1}{2}\mathfrak{g}^{0i}T_l.\end{aligned}$$

The third term

$$\begin{aligned}\delta^i_l \mathfrak{g}^{\mu\sigma}p^0_{\mu\sigma} &= \delta^i_l \mathfrak{g}^{\mu\sigma}\left[-\Gamma^0_{\mu\sigma} + \frac{1}{2}\delta^0_\mu T_\sigma + \frac{1}{2}\delta^0_\sigma T_\mu\right] \\ &= -\delta^i_l \mathfrak{g}^{\mu\sigma}\Gamma^0_{\mu\sigma} + \frac{1}{2}\delta^i_l \mathfrak{g}^{0\sigma}T_\sigma + \frac{1}{2}\delta^i_l \mathfrak{g}^{\mu 0}T_\mu \\ &= -\delta^i_l \mathfrak{g}^{\mu\sigma}\Gamma^0_{\mu\sigma} + \delta^i_l \mathfrak{g}^{\mu 0}T_\mu.\end{aligned}$$

The fourth term

$$\begin{aligned}-\delta^i_l \frac{1}{3}\mathfrak{g}^{\mu 0}p^\sigma_{\mu\sigma} &= \delta^i_l \frac{1}{3}\mathfrak{g}^{\mu 0}\left[\Gamma^\sigma_{\mu\sigma} - \frac{1}{2}\delta^\sigma_\mu T_\sigma - \frac{1}{2}\delta^\sigma_\sigma T_\mu\right] \\ &= \delta^i_l \frac{1}{3}\mathfrak{g}^{\mu 0}\left[T_\mu - \frac{1}{2}T_\mu - 2T_\mu\right] \\ &= -\frac{1}{2}\delta^i_l \mathfrak{g}^{\mu 0}T_\mu.\end{aligned}$$

We substitute all the terms into (5.82) and find

$$\begin{aligned}\mathbb{W}^{0i} &= (-\mathfrak{g}^{\mu 0}\Gamma^i_{\mu l} + \frac{1}{2}\mathfrak{g}^{i0}T_l + \frac{1}{2}\delta^i_l \mathfrak{g}^{\mu 0}T_\mu) + (\mathfrak{g}^{\mu i}\Gamma^0_{\mu l} - \frac{1}{2}\mathfrak{g}^{0i}T_l) \\ &\quad + (-\delta^i_l \mathfrak{g}^{\mu\sigma}\Gamma^0_{\mu\sigma} + \delta^i_l \mathfrak{g}^{\mu 0}T_\mu) - (\frac{1}{2}\delta^i_l \mathfrak{g}^{\mu 0}T_\mu) \\ &= -\mathfrak{g}^{\mu 0}\Gamma^i_{\mu l} + \mathfrak{g}^{\mu i}\Gamma^0_{\mu l} + \delta^i_l(-\mathfrak{g}^{\mu\sigma}\Gamma^0_{\mu\sigma} + \mathfrak{g}^{\mu 0}T_\mu).\end{aligned}$$

Step 3. *Using the asymptotic conditions to approximate* $\mathbb{W}^{0i}$. Using the approximation of the metric density in Lemma 5.51, we obtain

$$\begin{aligned}\mathbb{W}^{0i} &= \frac{1}{16\pi}(-\eta^{\mu 0}\Gamma^i_{\mu l} + \eta^{\mu i}\Gamma^0_{\mu l} + \delta^i_l(-\eta^{\mu\sigma}\Gamma^0_{\mu\sigma} + \eta^{\mu 0}T_\mu)) + O(h\partial h) \\ &= \frac{1}{16\pi}(\delta^0_\mu\Gamma^i_{\mu l} + \delta^\mu_i\Gamma^0_{\mu l} + \delta^i_l(\Gamma^0_{00} - \Gamma^0_{jj} - \delta^\mu_0 T_\mu)) + O(h\partial h) \\ &= \frac{1}{16\pi}(\Gamma^i_{0l} + \Gamma^0_{il} + \delta^i_l(\Gamma^0_{00} - \Gamma^0_{jj} - \Gamma^\mu_{0\mu})) + O(h\partial h),\end{aligned} \tag{5.83}$$

where $j = 1, 2, 3$.

We approximate the Christoffels in (5.83) using Lemma 5.50. The first term gives

$$\begin{aligned}\Gamma^i_{0l} &= \frac{1}{2}\eta^{i\rho}(\partial_0 g_{\rho l} + \partial_l g_{\rho 0} - \partial_\rho g_{0l}) + O(h\partial h) \\ &= \frac{1}{2}\delta^i_\rho(\partial_0 g_{\rho l} + \partial_l g_{\rho 0} - \partial_\rho g_{0l}) + O(h\partial h) \\ &= \frac{1}{2}(\partial_0 g_{il} + \partial_l g_{i0} - \partial_i g_{0l}) + O(h\partial h).\end{aligned}$$

For the second term, we have

$$\begin{aligned}\Gamma^0_{il} &= \frac{1}{2}\eta^{0\rho}(\partial_i g_{\rho l} + \partial_l g_{\rho i} - \partial_\rho g_{il}) + O(h\partial h) \\ &= -\frac{1}{2}\delta^0_\rho(\partial_i g_{\rho l} + \partial_l g_{\rho i} - \partial_\rho g_{il}) + O(h\partial h) \\ &= \frac{1}{2}(-\partial_i g_{0l} - \partial_l g_{0i} + \partial_0 g_{il}) + O(h\partial h).\end{aligned}$$

For the third term, we have

$$\begin{aligned}\Gamma^0_{00} &= \frac{1}{2}\eta^{0\rho}(\partial_0 g_{\rho 0} + \partial_0 g_{\rho 0} - \partial_\rho g_{00}) + O(h\partial h) \\ &= -\frac{1}{2}\delta^0_\rho(\partial_0 g_{\rho 0} + \partial_0 g_{\rho 0} - \partial_\rho g_{00}) + O(h\partial h) \\ &= \frac{1}{2}(-\partial_0 g_{00} - \partial_0 g_{00} + \partial_0 g_{00}) + O(h\partial h) \\ &= -\frac{1}{2}\partial_0 g_{00} + O(h\partial h).\end{aligned}$$

For fourth term, we have

$$\begin{aligned}-\Gamma^0_{jj} &= -\frac{1}{2}\eta^{0\rho}(\partial_j g_{\rho j} + \partial_j g_{\rho j} - \partial_\rho g_{jj}) + O(h\partial h) \\ &= \frac{1}{2}\delta^0_\rho(\partial_j g_{\rho j} + \partial_j g_{\rho j} - \partial_\rho g_{jj}) + O(h\partial h) \\ &= \partial_j g_{0j} - \frac{1}{2}\partial_0 g_{jj} + O(h\partial h).\end{aligned}$$

For fifth term, we have

$$\begin{aligned}-\Gamma^\mu_{0\mu} &= -\frac{1}{2}\eta^{\mu\rho}(\partial_0 g_{\mu\rho} + \partial_\mu g_{0\rho} - \partial_\rho g_{\mu 0}) + O(h\partial h) \\ &= \frac{1}{2}(-\delta^0_\rho + \delta^j_\rho)(-\partial_0 g_{\mu\rho} - \partial_\mu g_{0\rho} + \partial_\rho g_{\mu 0}) + O(h\partial h) \\ &= -\frac{1}{2}\delta^0_\rho(-\partial_0 g_{0\rho} - \partial_0 g_{0\rho} + \partial_\rho g_{00}) \\ &\quad + \frac{1}{2}\delta^j_\rho(-\partial_0 g_{j\rho} - \partial_j g_{0\rho} + \partial_\rho g_{j0}) + O(h\partial h) \\ &= \frac{1}{2}(\partial_0 g_{00} + \partial_0 g_{00} - \partial_0 g_{00}) \\ &\quad + \frac{1}{2}(-\partial_0 g_{jj} - \partial_j g_{0j} + \partial_j g_{j0}) + O(h\partial h) \\ &= \frac{1}{2}\partial_0 g_{00} - \frac{1}{2}\partial_0 g_{jj} + O(h\partial h).\end{aligned}$$

We substitute all the terms back into (5.83) and obtain

$$\begin{aligned}\mathbb{W}^{0i} &= \frac{1}{16\pi}\Big[\frac{1}{2}(\partial_0 g_{il} + \partial_l g_{i0} - \partial_i g_{0l}) + \frac{1}{2}(-\partial_i g_{0l} - \partial_l g_{0i} + \partial_0 g_{il}) \\ &\quad + \delta^i_l(-\frac{1}{2}\partial_0 g_{00} + \partial_j g_{0j} - \frac{1}{2}\partial_0 g_{jj} + \frac{1}{2}\partial_0 g_{00} - \frac{1}{2}\partial_0 g_{jj})\Big] + O(h\partial h) \\ &= \frac{1}{16\pi}\big[\partial_0 g_{il} - \partial_i g_{0l} + \delta^i_l(\partial_j g_{0j} - \partial_0 g_{jj})\big] + O(h\partial h).\end{aligned}$$

Using Lemma 5.53, we approximate the ADM momentum as

$$P^{il} = \delta^{li}(\partial_j g_{j0} - \frac{1}{2}\partial_0 g_{jj}) + \frac{1}{2}(-\partial_i g_{l0} - \partial_l g_{i0} + \partial_0 g_{il}) + O(h\partial h).$$

Expressing $\mathbb{W}^{0i}$ in terms of the ADM momentum gives

$$\begin{aligned}16\pi\mathbb{W}^{0i} &= 2P^{il} - \delta^i_l\partial_j g_{j0} + \partial_l g_{i0} + O(h\partial h) \\ &= 2P^{il} + \partial_j(\delta^j_l g_{i0} - \delta^i_l g_{j0}) + O(h\partial h).\end{aligned}$$

Finally, we compute the momentum component

$$\begin{aligned}p^l = p_l &= -H_{\text{boundary}}(\partial_l, \mathscr{S}) \\ &= -\frac{1}{2}\int_{\partial\mathscr{S}} \mathbb{W}^{\alpha\mu} dS_{\alpha\mu}.\end{aligned}$$

Using both the antisymmetry of $dS_{\alpha\mu}$ and $\mathbb{W}^{\alpha\mu}$, we obtain

$$\begin{aligned}p^l &= -\frac{1}{2}\int_{\partial\mathscr{S}} \mathbb{W}^{0i} dS_{0i} \\ &= -\frac{1}{8\pi}\int_{\partial\mathscr{S}} P^{il} + O(h\partial h) dS_{0i} + \frac{1}{16\pi}\int_{\partial\mathscr{S}} \partial_j(\delta^j_l g_{i0} - \delta^i_l g_{j0}) dS_{0i} \\ &= \lim_{R\to\infty}\frac{1}{8\pi}\int_{S(R)} P^{il} d\tilde{\sigma}_i.\end{aligned}$$

The second integral vanishes in the last equality due to the boundary of the boundary being empty, and the $O(h\partial h)$ term has vanishing surface integral under the ADM fall-off assumptions. This completes the proof. □

5.4 Invariance of the ADM mass

One can use the ADM energy-momentum vector to define the ADM mass. The ADM mass is an invariant of asymptotically flat manifolds. This section outlines the definitions and theorems needed to prove these invariance results.

Before discussing the invariance results in detail, we distinguish between two settings: asymptotically flat Riemannian manifolds and asymptotically flat spacetimes. In an asymptotically flat spacetime, the ADM energy-momentum is described by the four-vector

$$p^\mu_{\text{ADM}} = \left(E_{\text{ADM}}, p^1_{\text{ADM}}, p^2_{\text{ADM}}, p^3_{\text{ADM}}\right),$$

where E_{ADM} and p^l_{ADM} are given in Theorem 5.47 and Theorem 5.48, respectively, with $l = 1, 2, 3$.

The ADM mass is defined as the Lorentz-invariant norm of this four-vector. We obtain

$$M_{\mathrm{ADM}} := \sqrt{E_{\mathrm{ADM}}^2 - |P_{\mathrm{ADM}}|^2}, \tag{5.84}$$

where $P_{\mathrm{ADM}} = (p^1_{\mathrm{ADM}}, p^2_{\mathrm{ADM}}, p^3_{\mathrm{ADM}})$. For a more detailed treatment of the ADM mass in the spacetime setting, see Jaramillo and Gourgoulhon [45].

For an asymptotically flat Riemannian 3-manifold, the ADM mass is defined as

$$E_{\mathrm{ADM}} := \lim_{R\to\infty} \frac{1}{16\pi} \int_{S(R)} (g_{ij,j} - g_{jj,i}) d\sigma_i, \tag{5.85}$$

where $d\sigma_i := dS_{0i}$, as defined in (5.68), and $S(R)$ is the 2-sphere with radius R. This quantity corresponds to the energy component of the ADM energy-momentum in the spacetime setting. For more details about the ADM mass in the Riemannian setting, see Lee [51, Chapter 3]. It is important to keep the distinction between these two settings in mind, even though both quantities are often referred to as the ADM mass or ADM energy.

In Subsection 5.4.1, we address the coordinate invariance of the ADM mass on an asymptotically flat Riemannian manifold. Since the definition of the ADM mass uses coordinates, one must show that the resulting value is independent of the choice of asymptotic coordinate system. This is achieved by defining asymptotic coordinates on an end of an asymptotically flat manifold and then proving that the masses computed in any two such coordinate systems are well defined and coincide. This exposition follows Bartnik [7].

In Subsection 5.4.2, we show that the ADM mass is an invariant of an asymptotically flat spacetime. The key ingredient is the asymptotic symmetries theorem, which establishes the invariance of the ADM mass under changes of spacetime coordinates. Our exposition follows the treatment of Chruściel [22].

5.4.1 Invariance of ADM mass on a Riemannian manifold

It is natural to begin by studying the invariance of the ADM mass on Riemannian manifolds, since, in the spacetime setting, one restricts attention to spacelike hypersurfaces. Equivalently, the metric induced on each hypersurface is required to be Riemannian.

Bartnik [7] proves the coordinate invariance of the ADM mass, defined in (5.85), by formulating the asymptotic flatness conditions in terms of weighted Sobolev spaces. The invariance of the ADM mass is then proved for asymptotically flat manifolds satisfying the appropriate analytical conditions.

For completeness, we define weighted Sobolev spaces.

Definition 5.55 The **weighted Lebesgue spaces** L^p_δ, L'^p_δ, $1 \le p \le \infty$, with weight $\delta \in \mathbb{R}$, are the spaces of measurable functions in $L^p_{\mathrm{loc}}(\mathbb{R}^n)$, $L^p_{\mathrm{loc}}(\mathbb{R}^n \setminus \{0\})$, respectively, such that the norms $\|\cdot\|_{p,\delta}$, $\|\cdot\|'_{p,\delta}$ defined by

$$L^p_\delta: \qquad \|u\|_{p,\delta} = \begin{cases} \left(\int_{\mathbb{R}^n} |u|^p \sigma^{-\delta p - n}\, dx \right)^{1/p}, & p < \infty, \\ \operatorname{ess\,sup}_{\mathbb{R}^n} \left(\sigma^{-\delta} |u| \right), & p = \infty, \end{cases} \tag{5.86}$$

and

$$L'^p_\delta: \qquad \|u\|'_{p,\delta} = \begin{cases} \left(\int_{\mathbb{R}^n\setminus\{0\}} |u|^p r^{-\delta p - n}\, dx\right)^{1/p}, & p < \infty, \\ \operatorname{ess\,sup}_{\mathbb{R}^n\setminus\{0\}} \left(r^{-\delta}|u|\right), & p = \infty, \end{cases} \tag{5.87}$$

are finite. Here, the essential supremum is defined by

$$\operatorname{ess\,sup}_{x\in\mathbb{R}^n} \left(\sigma(x)^{-\delta}|u(x)|\right) := \inf\{C \geq 0 : \sigma(x)^{-\delta}|u(x)| \leq C \text{ for almost every } x \in \mathbb{R}^n\}.$$

The **weighted Sobolev spaces** are defined as

$$W^{k,p}_\delta: \qquad \|u\|_{k,p,\delta} = \sum_{j=0}^{k} \|D^j u\|_{p,\delta-j}, \tag{5.88}$$

and

$$W'^{k,p}_\delta: \qquad \|u\|'_{k,p,\delta} = \sum_{j=0}^{k} \|D^j u\|'_{p,\delta-j}. \tag{5.89}$$

Bartnik uses a more analytic definition of asymptotic flatness. In particular, he reformulates the asymptotic conditions in terms of weighted Sobolev spaces.

Definition 5.56 A smooth n-dimensional manifold (M, g) with complete Riemannian metric $g \in W^{1,q}_{\text{loc}}(M)$ for some $n < q < \infty$ is said to be **asymptotically flat** if there is a compact subset $K \subset M$ such that $M \setminus K$ has a **structure of infinity**: There is an $R \geq 1$ and a diffeomorphism $\Phi\colon M \setminus K \to \mathbb{R}^n \setminus B_R$, where B_R is the closed ball with radius R and centre 0. We denote $\mathbb{R}^n \setminus B_R := E_R$. The diffeomorphism Φ satisfies the following axioms

1. The differential $(\Phi g)_{ij}$ is uniformly equivalent to the flat metric η_{ij} on $\mathbb{R}^n \setminus B_R$. More precisely, there exists a $\lambda \geq 1$ such that

$$\frac{1}{\lambda}|\xi|^2 \leq ((d\Phi)g)_{ij}(x)\xi^i\xi^j \leq \lambda|\xi|^2 \text{for all } x \in E_R,\ \xi^i \in \mathbb{R}^n$$

2. $(d\Phi g)_{ij} - \eta_{ij} \in W^{1,q}_{-t}(E_R)$ for some decay rate $t > 0$.

The convention for the weights differs from some commonly used conventions, but has the advantage that it directly describes the growth at infinity. The condition

$$g_{ij} - \eta_{ij} \in W^{1,q}_{-t}(E_R)$$

ensures that the Riemannian metric approaches the flat metric sufficiently fast at infinity. Bartnik's exposition assumes that the asymptotically flat manifold has one end and that the metric is complete. The definition above gives a structure at infinity, denoted by (Φ, x), where $x^i = \Phi^i(m)$ for $m \in M$.

One can show that two structures at infinity differ only by a rigid motion and by terms which decay as $o(r^{1-t})$, for a decay rate $t > 0$. In this sense, the structure at infinity is almost unique. To show that the mass is well-defined, we impose one additional assumption.

Definition 5.57 A Riemannian manifold (M, g) satisfies the **mass decay conditions** if there is an asymptotic structure Φ such that

1. $((d\Phi)g - \eta) \in W^{2,q}_{-t}(E_{R_0})$ for some $R_0 > 1,\ q > n$ and $t \geq \frac{1}{2}(n-2)$,
2. The scalar curvature must be integrable. Hence, we have $R(g) \in L^1(M)$.

Using this mass decay condition, one defines the mass of an asymptotic structure.

Proposition 5.58 *Suppose that (M, g, Φ) satisfies the mass decay conditions. Let $\{D_k\}_{k=1}^{\infty}$ be an exhaustion of M by closed sets, and set*

$$S_k := \Phi(\partial D_k).$$

Assume that each S_k is a connected $(n-1)$-dimensional C^1 submanifold without boundary in $\mathbb{R}^n$, and that

$$\begin{aligned} R_k := \inf\big\{|x| \colon x \in S_k\big\} \longrightarrow \infty \qquad &\text{as } k \to \infty, \\ R_k^{-(n-1)}\operatorname{area}(S_k) \quad &\text{is bounded as } k \to \infty. \end{aligned}$$

Suppose also that $R_1 \geq R_0$. Then the limit

$$\operatorname{mass}(g, \Phi) := \lim_{k\to\infty} \oint_{S_k} ((d\Phi g)_{ij,j} - (d\Phi g)_{jj,i})\, dS^i$$

exists and is independent of the sequence $\{S_k\}$.

Using Proposition 5.58 and the almost uniqueness of the structure at infinity gives the uniqueness of mass.

Theorem 5.59 *Let (Φ, x) and (Ψ, z) be two structures of infinity for (M, g) in the sense of Definition 5.56. Suppose that both (Φ, x) and (Ψ, z) satisfy the mass decay conditions, as in Definition 5.57, with decay rates t_1, t_2 respectively such that $t = \min(t_1, t_2) \geq \frac{1}{2}(n-2)$. Then $\operatorname{mass}(g, \Phi)$ and $\operatorname{mass}(g, \Psi)$ are well defined and equal.*

This last theorem shows that the mass is indeed an invariant of a Riemannian asymptotically flat manifold (M, g).

5.4.2 Invariance under changes of foliation

In this subsection, we return to the spacetime setting. We outline how to show that the ADM mass, as defined in (5.84), is an invariant of an asymptotically flat spacetime. Our exposition follows Chruściel [22].

First, we show that a Lorentz transformation preserves the ADM mass. Then, we explain why restrictions on the allowed coordinate transformations are necessary, by exhibiting coordinate transformations that generate non-zero mass in Minkowski spacetime. Finally, we introduce the asymptotic symmetry theorem and the regularity assumptions under which such pathological transformations are excluded.

We begin by rigorously defining a Lorentz transformation.

Definition 5.60 A **Lorentz transformation** is a linear transformation Λ such that

$$\Lambda^T \eta \Lambda = \eta,$$

where η is the Minkowski metric.

A **boost** is a Lorentz transformation from the reference frame of an observer A to the reference frame of an observer B moving with constant velocity relative to A.

Lemma 5.61 *If two ADM energy-momenta are related by a Lorentz transformation, then they determine the same ADM mass.*

Proof. Let $p^{\mu\prime}_{\mathrm{ADM}}$ and p^{ν}_{ADM} be two ADM energy-momenta related by a Lorentz transformation:

$$p^{\mu\prime}_{\mathrm{ADM}} = \Lambda_{\mu}{}^{\nu} p^{\nu}_{\mathrm{ADM}}. \tag{5.90}$$

For notational simplicity, we suppress the ADM subscript. Since Λ is a Lorentz matrix, it satisfies

$$\eta_{\alpha\beta}\Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu} = \eta_{\mu\nu}.$$

It follows that

$$p'^{\mu}p'_{\mu} = \eta_{\alpha\beta}p'^{\alpha}p'^{\beta} = \eta_{\alpha\beta}\Lambda^{\alpha}{}_{\mu}p^{\mu}\Lambda^{\beta}{}_{\nu}p^{\nu} = \eta_{\alpha\beta}\Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu}p^{\mu}p^{\nu} = \eta_{\mu\nu}p^{\mu}p^{\nu} = p^{\mu}p_{\nu}. \tag{5.91}$$

Recall that we use the $(-,+,+,+)$ convention for the Lorentzian metric. Thus, the contraction is

$$p^{\mu}p_{\nu} = \eta_{\mu\nu}p^{\mu}p^{\nu} = -(p^0)^2 + \sum_{i=1}^{3}(p^i)^2.$$

Consequently, we rewrite the ADM energy equation (5.84) as

$$M_{\mathrm{ADM}} = \sqrt{-p^{\mu}p_{\mu}}.$$

Applying (5.91), we obtain

$$M'_{\mathrm{ADM}} = \sqrt{-p'^{\mu}p'_{\mu}} = \sqrt{-p^{\mu}p_{\mu}} = M_{\mathrm{ADM}}.$$

Therefore, the two ADM energy-momenta determine the same ADM mass. □

Let $\Sigma_1 \subset M$ and $\Sigma_2 \subset M$ be two spacelike hypersurfaces with initial data sets

$$(\gamma^{(1)}_{ij}, \pi^{ij}_{(1)}) \qquad \text{and} \qquad (\gamma^{(2)}_{ij}, \pi^{ij}_{(2)}),$$

respectively. The two hypersurfaces are associated with different foliations of the spacetime. Let $H_{\mathrm{boundary},1}$ and $H_{\mathrm{boundary},2}$ denote the corresponding boundary Hamiltonians, expressed in the asymptotic coordinate systems x^{μ} and y^{μ}, respectively. Suppose that the coordinate systems x^{μ} and y^{μ} are related by an **asymptotically Lorentzian transformation**, that is,

$$y^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu} + \xi^{\mu}, \qquad \xi^{\mu} = O_2(r^{1-\alpha})^{6}, \tag{5.92}$$

where Λ is a Lorentz matrix and $\alpha > 1/2$. Then the two ADM four-momenta, expressed using the boundary Hamiltonians, are related by the Lorentz transformation Λ. Hence, their ADM masses agree. For more details, see Chruściel [21].

[6] We use a version of the Big-O notation that also records the decay of derivatives. We write $f = O_k(r^{\alpha})$ if

$$\partial^l f = O(r^{\alpha-l})$$

for all derivatives of order $l \leq k$.

When no restrictions are imposed on the coordinate transformations, one can construct mass generating coordinate transformation. For example, in Minkowski spacetime the ADM energy-momentum four-vector is zero, and hence the ADM mass is zero. However, if one performs the coordinate transformation

$$y^0 = x^0 + a\sqrt{r}, y^i = \left(1 + \frac{b}{\sqrt{r}}\right) x^i, \quad a, b \in \mathbb{R},$$

where r denotes the radius from the origin, then the resulting energy-momentum four-vector p^μ_{ADM} is nonzero whenever $a \neq \pm b$. Thus, without suitable restrictions on the allowed coordinate transformations, even Minkowski spacetime can be assigned nonzero ADM mass.

The following two conjectures describe the geometric mechanism that is expected to underlie the invariance of the ADM mass. The first concerns the asymptotic form of the coordinate transformations. The second expresses the resulting independence of the mass from the choice of asymptotically flat hypersurface.

We include these conjectures to motivate the more precise results below. They will not be used as established statements. Chruściel proves a refined asymptotic symmetry theorem under additional hypotheses and derives an invariant mass result in the situations covered by those hypotheses.

Conjecture 5.62 (Asymptotic symmetries) *Let x^μ and y^μ be two coordinate systems in which the metric satisfies*

$$g_{\mu\nu} - \eta_{\mu\nu} = O_1(r^{-\alpha}), \qquad \frac{1}{2} < \alpha \leq 1.$$

Then every C^2 coordinate transformation between these coordinate systems that preserves this asymptotic form is of the form

$$y^\mu = \Lambda^\mu{}_\nu x^\nu + \chi^\mu,$$

where $\Lambda^\mu{}_\nu$ is a Lorentz matrix and

$$\chi^\mu = O_2(r^{1-\alpha}).$$

The preceding conjecture shows that all twice differentiable asymptotic coordinate transformations are in fact asymptotically Lorentz. Hence, these coordinate changes transform the ADM energy–momentum as a Lorentz transformation. Therefore, both coordinate systems determine the same ADM mass. We conclude that the asymptotic symmetries conjecture implies the invariance of the ADM mass. More precisely, we obtain the following conjecture.

Conjecture 5.63 (Invariant mass) *Let M be an asymptotically flat end of a Lorentzian manifold $(\mathscr{M}, g)$ satisfying the vacuum Einstein equations. Let $N_1, N_2 \subset M$ be two asymptotically flat three-ends, and let $m(N_1)$ and $m(N_2)$ denote their associated ADM masses. Then*

$$m(N_1) = m(N_2).$$

To obtain a rigorous result, we introduce the additional hypotheses. The resulting statement is a refined version of the preceding conjectures. We formulate this result as a theorem.

The situation is restricted to four-dimensional asymptotic regions whose time-width grows linearly with the spatial radius. More precisely, we define a **boost-type domain** as

$$\Omega_{\theta,R,T} = \{x^\mu \mid r \geq R, |t| \leq \theta r + T\}, \text{ for some } \theta, R \geq 0,\ T \in \mathbb{R}.$$

It is boost-type because the region $\Omega_{\theta,R,T}$ is large enough to contain hypersurfaces that are related by a **finite** Lorentz boost. Here finite means that the time grows at most linearly with spatial radius.

Boost-type domains provide enough overlap to compare hypersurfaces related by finite boosts. However, they do not control the asymptotic behaviour of the corresponding coordinate systems. We therefore impose conditions on each asymptotic coordinate system individually. Their role in establishing coordinate invariance becomes apparent when two admissible coordinate systems are compared. The conditions on the individual systems constrain the transition map between them, thereby excluding pathological coordinate changes.

Definition 5.64 A coordinate system $\Phi_x = \{x^\mu\}$ defined on a subset M_x of a Lorentzian manifold $\mathscr{M}$ will be called **α-admissible** or, shortly, **admissible**, if there exist $\theta_x, R_x \in \mathbb{R}^+$, $T_x \in \mathbb{R}$ such that:

1. $\Phi_x(M_x) = \Omega_{\theta_x,R_x,T_x}$, where $\Omega_{\theta_x,R_x,T_x} = \{x^\mu : r \geq R_x,\ |t| \leq \theta_x r + T_x\}$
2. In $\mathscr{O}_x = \Phi_x(M_x)$ we have, with $g_{\mu\nu} = g\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right) \in C^2(\mathscr{O}_x)$,
$$0 < -g_{00}, \quad \forall X^i \in \mathbb{R}^{n-1}, \quad 0 < g_{ij}X^iX^j$$
3. For $0 < \Psi < \theta_x$, for $x \in \Omega_{\Psi,R_x,T_x}$, with $\sigma(x) = \left(1 + \sum_i (x^i)^2\right)^{1/2}$, we have
$$|g_{\mu\nu}(x) - \eta_{\mu\nu}| \leq C_x(\Psi)\sigma(x)^{-\alpha}, \qquad \left|\frac{\partial g_{\mu\nu}(x)}{\partial x^\sigma}\right| \leq C_x(\Psi)\sigma(x)^{-\alpha-1},$$
with some function $C_x(\Psi) < \infty$.

The first condition ensures that the coordinate system covers a large, boost-type region. The second requires t to be a regular time coordinate and that the slices $t = \text{const}$ are spacelike hypersurfaces, so that their induced metrics are Riemannian. The final condition ensures that the metric tends uniformly to the flat one in spacelike directions.

To compare the ADM quantities computed in two admissible coordinate systems, it is not enough that each system separately satisfies the required asymptotic conditions. We must also ensure that the asymptotic hypersurface associated with one coordinate system lies within the region covered by the other, and that it remains strictly inside the latter's boost-type domain. This motivates the following notion of a three-end lying within an asymptotic four-end.

Definition 5.65 Let x^μ and y^μ be admissible coordinate systems, with appropriate constants $\theta_x, R_x, \theta_y, R_y, T_x, T_y \in \mathbb{R}$. The three-end $N^x_R = \{x^\mu : x^0 = 0,\ r(x) \geq R\}, R \geq R_x$, will be said to **lie within the asymptotic four-end** defined by the coordinates y^μ and the three-end $N^y_{R_y} = \{y^\mu : y^0 = 0,\ r(y) \geq R_y\}$ if the following two conditions are satisfied:

1. $\Phi_x^{-1}(N^x_R) \subset M_y$, where $M_y \subset M$ is the subset on which the coordinates y^μ are defined.

2. $\Phi_y \circ \Phi_x^{-1}(N_R^x) \subset \Omega^y_{\theta,R_y}$ for some $\theta < \theta_y$.

If $\theta_y > 1$, we shall say that the metric g is a no-radiation metric.

Having specified both the admissibility conditions for the coordinate systems and the required inclusion of one asymptotic end within the other, we are now in a position to formulate the corresponding asymptotic symmetries theorem.

Theorem 5.66 (Asymptotic Symmetries at Spatial Infinity) *Let the coordinates $\{x^\mu\}$ and $\{y^\mu\}$ be α-admissible, $0 < \alpha \leq 1$. Let $y(x)$ be twice differentiable and suppose that there exist constants $R \equiv R_x$ and $\theta < \theta_y$ such that*

$$N_R^x = \left\{x^\mu : x^0 = 0,\ r(x) \geq R\right\} \subseteq \Omega^y_{\theta,R_y,T_y} = \left\{y^\mu : r(y) \geq R_y,\ |y^0| \leq \theta r(y) + T_y\right\}.$$

There exists a Lorentz matrix Λ and a constant C such that, for every $x \in N_R^x$,

$$\begin{aligned} &y^\mu = \Lambda^\mu{}_\nu x^\nu + \zeta^\mu, \quad |\zeta^\mu{}_{,\nu}| \leq C(1+r)^{-\alpha}, \quad |\zeta^\mu{}_{,\sigma\nu}| \leq C(1+r)^{-\alpha-1}, \\ &|\zeta^\mu| \leq C(1+r)^{1-\alpha} \qquad \textit{for } 0<\alpha<1, \qquad |\zeta^\mu| \leq C\big(1+\ln(1+r)\big) \quad \textit{if } \alpha = 1. \end{aligned} \tag{3.21}$$

Thus, under the assumptions of Theorem 5.66, the transition map between the two admissible coordinate systems, and hence between the foliations defined by the level sets of their respective time coordinates, is asymptotically Lorentz in the sense of (5.92). Its leading linear part is a Lorentz transformation, while the remainder is of lower order and its derivatives decay at the stated rates. The inclusion condition ensures that the corresponding spatial ends can be compared within a common boost-type region. Therefore, the ADM four-momenta, and hence the boundary Hamiltonians, associated with the two asymptotic coordinate systems are related by a Lorentz transformation. Consequently, they determine the same ADM mass.

Chapter 6

Comparison with other approaches and positive mass results

This chapter places the results of this thesis in a broader context. In Section 6.1, we discuss several derivations of and motivations for the ADM mass. We first review the original approach developed by Arnowitt, Deser, and Misner [4, 5]. We then consider an alternative approach due to Christodoulou [18], which uses the gravitational Lagrangian and Noether's theorem to motivate the ADM mass. In the final subsection, we compare these approaches with the Hamiltonian framework developed by Chruściel, Jezierksi, and Kijowski [27]. We also briefly discuss work by Michel [64], which identifies an algebraic mechanism underlying, among other asymptotic invariants, the ADM mass.

In Section 6.2, we discuss positive mass theorems in several Riemannian and Lorentzian asymptotic settings. We begin with a historical overview of the positive mass theorem for asymptotically flat Riemannian manifolds and asymptotically flat spacetimes. We first review the work of Schoen and Yau [72, 73, 76] on the positive mass theorem, including their generalization up to dimension 7. We then discuss Witten's alternative proof for spin manifolds [86] and briefly consider Reula's work [68] on the analytic foundations of Witten's argument. We subsequently turn to recent developments, including preprints appearing in 2026 by Bi, Hao, He, Shi, Zhu and Jintian [8] and by Brendle and Wang [10, 11]. We also highlight several results concerning metrics of lower regularity. Next, we discuss notions of mass for asymptotically hyperbolic Riemannian manifolds and asymptotically anti-de Sitter spacetimes, together with an overview of positive mass results in these settings. Finally, we mention two additional notions of mass or energy arising in these geometries and provide references for further reading.

6.1 Hamiltonian and Lagrangian perspectives on the ADM mass

In 1916, Einstein [34] published his paper on the foundations of general relativity. In this theory, gravity is described by the Lorentzian geometry of a four-dimensional spacetime. The curvature of spacetime gives rise to gravitational effects. The gravitational field is represented by a Lorentzian metric, and the Einstein equations determine how this metric is related to the distribution of matter and energy.

General relativity is invariant under changes of coordinates. More precisely, metrics related by a diffeomorphism represent the same physical geometry. This creates a difficulty when one attempts to describe the evolution of the gravitational field through time. Suitable initial data satisfying the Einstein constraint equations determine a spacetime development only up to diffeomorphism. Consequently, a canonical description must distinguish genuine dynamical evolution from changes arising solely from the choice of coordinates.

This section discusses several perspectives on this problem and on the resulting definition of gravitational energy. In Subsection 6.1.1, we review the Hamiltonian formulation introduced by Arnowitt, Deser, and Misner [5]. Their construction rewrites the Einstein equations as a constrained Hamiltonian system and leads, for asymptotically flat initial data, to the asymptotic energy now known as the ADM mass [4].

In Subsection 6.1.2, we discuss Christodoulou's Lagrangian approach to motivating the ADM mass [18]. This approach uses Noether's theorem to obtain conserved quantities and identifies the ADM mass with the quantity associated with asymptotic time translations. Finally, in Subsection 6.1.3, we compare the original ADM construction and Christodoulou's approach with the Hamiltonian framework of Chruściel, Jezierksi, and Kijowski used in this thesis. The aim is to clarify where the three approaches agree, how they differ, and what each perspective contributes.

Throughout this section, Greek indices range from 0 to 3, whereas Latin indices range from 1 to 3. To facilitate comparison, we express the constructions of Arnowitt, Deser, and Misner and of Christodoulou in the notation used throughout this thesis.

6.1.1 The original ADM Hamiltonian formulation

Arnowitt, Deser, and Misner introduced a Hamiltonian formulation of general relativity in which spacetime is decomposed into a family of three-dimensional hypersurfaces. Their construction separates the variables describing the gravitational field on each hypersurface from the quantities describing how the hypersurfaces are related within spacetime. In this way, the ADM construction rewrites the Einstein equations as a constrained Hamiltonian system with first-order evolution equations.

The ADM formulation therefore has two main features. First, it singles out a time coordinate and recasts the four-dimensional theory in a $3+1$-dimensional form. Second, it provides a Hamiltonian description from which, under suitable asymptotic conditions, the total energy of the gravitational system can be obtained.

The following exposition is based on the treatment of Arnowitt, Deser, and Misner [5], which republishes their original 1962 work.

To obtain a first-order formulation of the field equations, ADM begin with the Einstein–Hilbert action

$$I = \int \sqrt{-\det g}\, R\, d^4x. \tag{6.1}$$

Here, we omit the normalization factor $\frac{1}{16\pi}$ and adopt the units of ADM in this subsection. The integrand in (6.1) is the Hilbert Lagrangian density introduced in Subsection 5.1.2. ADM then rewrite this action in the first-order Palatini form

$$I = \int \widetilde{L}\, d^4x = \int \mathfrak{g}^{\mu\nu} R_{\mu\nu}\, d^4x,$$

where

$$\mathfrak{g}^{\mu\nu} := \sqrt{-\det g}\, g^{\mu\nu}.$$

Here $\widetilde{L}$ denotes the Palatini Lagrangian density, as in Subsection 5.1.2. In this formulation, $\mathfrak{g}^{\mu\nu}$ and $\Gamma^{\alpha}_{\mu\nu}$ are treated as independent variables. Variation with respect to $\mathfrak{g}^{\mu\nu}$ gives the vacuum Einstein equations, while variation with respect to $\Gamma^{\alpha}_{\mu\nu}$ recovers the usual relation between the connection and the metric. In this way, the field equations are expressed as a first-order system.

The next step is to single out the time coordinate and perform a $3+1$ decomposition. To remain consistent with the notation used throughout this thesis, we denote the induced spatial metric by ${}^3g_{ij}$ and its conjugate momentum P^{ij}_{ADM}, together with the lapse N and shift N^i. ADM denote the canonical momentum introduced here by π^{ij}. We use P^{ij}_{ADM} to remain consistent with the notation of this thesis. The sign convention for the canonical momentum differs from the convention used by Chruściel–Jezierksi–Kijowski. In this subsection, we follow the ADM convention, since our exposition is based on their paper. These quantities are given by

$$\begin{aligned} {}^3g_{ij} := g_{ij}, \quad N := \frac{1}{\sqrt{-g^{00}}}, \quad N_j := g_{0j}, \quad N^i := {}^3g^{ij}N_j. \\ P^{ij}_{\mathrm{ADM}} := \sqrt{-\det g}(\Gamma^0_{pq} - {}^3g_{pq}\Gamma^0_{rs}{}^3g^{rs}){}^3g^{ip3}g^{jq}. \end{aligned}$$

Here $g_{\mu\nu}$ and $\Gamma^{\alpha}_{\mu\nu}$ denote four-dimensional quantities, whereas ${}^3g_{ij}$ denotes the induced three-dimensional metric. The function N is called the lapse, and N^i is called the shift vector. Chruściel–Jezierksi–Kijowski use the symbols ν and β^i for the corresponding quantities in the model-space formulation developed earlier in this thesis. To distinguish that construction from the historical ADM presentation, we retain the notation N and N^i here.

Remark 6.1 In the asymptotically flat setting considered by ADM, the induced metric ${}^3g_{ij}$ approaches the Euclidean metric δ_{ij}, while the conjugate momentum P^{ij}_{ADM} satisfies an appropriate fall-off condition. Under suitable asymptotic assumptions, the resulting Hamiltonian gives rise to the ADM mass.

Several important observations now follow. The spatial metric ${}^3g_{ij}$ is geometrically defined on a fixed $t = \mathrm{const}$ hypersurface and does not depend on how the coordinates are continued away from this hypersurface. By contrast, the lapse N and shift N_i describe this continuation of the coordinate system.

The canonical momentum can also be expressed in terms of the second fundamental form K_{ij}. Loosely speaking, K_{ij} measures how the normal to the hypersurface changes from point to point and therefore describes how the hypersurface is embedded in the four-dimensional spacetime. Like the spatial metric, it does not depend on how the coordinates are continued away from the hypersurface. Using the sign convention of ADM, one obtains

$$P^{ij}_{\mathrm{ADM}} = -\sqrt{\det {}^3g}\left(K^{ij} - {}^3g^{ij}K\right),$$

where

$$K := {}^3g^{ij}K_{ij}.$$

Subject to four constraint equations, the variables $({}^3g_{ij}, P^{ij}_{\mathrm{ADM}})$ form a complete set of Cauchy data. After discarding a spatial divergence and a total time derivative, ADM rewrite the Lagrangian density as

$$\widetilde{L} = P^{ij}_{\mathrm{ADM}}\partial_t{}^3g_{ij} - NR^0 - N_iR^i,$$

where R^0 and R^i are the ADM constraint functions. We will not require the explicit form of these constraints here.

ADM subsequently solve the four constraint equations and impose four coordinate conditions. We state the result of this reduction without proof. Using the transverse-traceless decomposition of the spatial metric and its conjugate momentum, the reduced Lagrangian density becomes

$$\widetilde{L}_{\text{red}} = P^{ij\,TT}_{\text{ADM}} \partial_t {}^3g^{TT}_{ij} + \mathfrak{T}^0{}_0.$$

Here TT denotes the transverse-traceless components obtained from the linear orthogonal decomposition of symmetric tensors. The quantity $\mathfrak{T}^0{}_0$ is determined implicitly by solving the constraint equations and satisfies

$$\mathfrak{T}^0{}_0 = \nabla^2 g^T,$$

where g^T denotes the trace of the transverse part of the spatial metric and ∇^2 is the flat-space Laplacian. For details of the decomposition and the subsequent reduction, see Sections 4.2 and 4.5 of [5].

Comparing the reduced Lagrangian with the standard Legendre relation

$$L = p\dot{q} - H,$$

we find that the reduced Hamiltonian density is

$$\mathscr{H}_{\text{red}} = -\mathfrak{T}^0{}_0 = -\nabla^2 g^T.$$

Consequently, the reduced Hamiltonian is

$$H = -\int_{\mathscr{S}_t} \mathfrak{T}^0{}_0 \, d^3x = -\int_{\mathscr{S}_t} \nabla^2 g^T \, d^3x,$$

where $\mathscr{S}_t$ denotes the spacelike hypersurface corresponding to the time coordinate t.

For a particular solution, the numerical value of this Hamiltonian equals the gravitational energy. Applying the divergence theorem, ADM obtain

$$E = -\int_{\mathscr{S}_t} \nabla^2 g^T \, d^3x = -\lim_{r\to\infty} \oint_{S_r} \partial_i g^T \, dS_i = \lim_{r\to\infty} \oint_{S_r} \left(\partial_j {}^3g_{ij} - \partial_i {}^3g_{jj}\right) dS_i.$$

Here dS_i denotes the oriented two-dimensional surface element on the coordinate sphere S_r, and

$$g^T = {}^3g_{ii} - \frac{1}{\nabla^2} \partial_i \partial_j {}^3g_{ij},$$

with $\frac{1}{\nabla^2}$ denoting the inverse flat-space Laplacian.

In a similar manner, ADM obtain the total spatial momentum. It is given by

$$P^i = -2 \lim_{r\to\infty} \oint_{S_r} P^{ij}_{\text{ADM}} \, dS_j.$$

In [4], ADM compare several expressions for gravitational energy and examine their behaviour under coordinate transformations at spatial infinity. Their analysis shows that coordinate invariance requires suitable asymptotic assumptions. The coordinate invariance of the ADM mass was later established under precise decay and regularity assumptions by Bartnik [7] and Chruściel [22]. For a more detailed discussion, we refer to Section 5.4.

In summary, the ADM formulation rewrites general relativity as a constrained Hamiltonian system on a family of spacelike hypersurfaces. After solving the constraint equations and imposing suitable coordinate conditions, ADM express the gravitational energy and momentum as surface integrals at spatial infinity. The resulting energy integral is the quantity now known as the ADM mass.

6.1.2 Christodoulou's motivation for the ADM energy

Christodoulou gives an alternative, Lagrangian motivation for the ADM mass. Unlike the original ADM construction, his approach does not begin by rewriting general relativity as a constrained Hamiltonian system. Instead, he applies Noether's theorem to the gravitational Lagrangian and associates conserved currents with symmetries of a reference Minkowski metric. Under suitable asymptotic conditions, these currents yield well-defined surface integrals, and the integrals associated with translations agree with the ADM mass.

The starting point of Christodoulou's construction is Noether's theorem. We recall the following standard form of the theorem without proof.

Theorem 6.2 (Noether's theorem) *In the framework of a Lagrangian theory, to each continuous group of transformations leaving the Lagrangian invariant there corresponds a quantity which is conserved. In particular:*

- *Energy is conserved when the Lagrangian is left invariant under time translations.*
- *Linear momentum is conserved when the Lagrangian is left invariant under space translations.*

Again, we use the notation adopted throughout this thesis, which follows Chruściel–Jezierksi–Kijowski, but retain the normalization used by Christodoulou. In particular, $\widetilde{L}$ denotes the Hilbert Lagrangian density, whereas L will denote the corresponding first-order Lagrangian density.

Christodoulou begins with the action

$$I = \int_M \widetilde{L}\, d^4x = -\frac{1}{4}\int_M R\, d\mu_g,$$

where

$$\widetilde{L} = -\frac{1}{4}\sqrt{-\det g}\, R.$$

Here R is the scalar curvature of the spacetime metric g, and

$$d\mu_g = \sqrt{-\det g}\, d^4x$$

is its volume form.

The Hilbert Lagrangian density contains second derivatives of the metric. The version of Noether's theorem used in Christodoulou's construction is most conveniently applied to a Lagrangian density depending only on the fields and their first derivatives. Following Einstein and Weyl, Christodoulou therefore removes a total divergence and introduces the first-order Lagrangian density

$$L := \widetilde{L} - \partial_\alpha I^\alpha,$$

where

$$I^\alpha := -\frac{1}{4}\sqrt{-\det g}\left(g^{\mu\nu}\Gamma^\alpha_{\mu\nu} - g^{\mu\alpha}\Gamma^\nu_{\mu\nu}\right).$$

The second-derivative terms contained in $\widetilde{L}$ are absorbed into the divergence $\partial_\alpha I^\alpha$, so that L depends only on the metric and its first derivatives.

Since L and $\widetilde{L}$ differ by a total divergence, they give rise to the same Euler–Lagrange equations. One can use an argument similar to that in Subsection 5.1.2; for details, see [1].

Using the first-order Lagrangian density L, we introduce the field momentum and the associated canonical stress.

Definition 6.3 The **momentum conjugate** to the derivative $\partial_\mu g_{\alpha\beta}$ is defined by

$$p^{\mu\alpha\beta} := \frac{\partial L}{\partial(\partial_\mu g_{\alpha\beta})}.$$

The corresponding **canonical stress** is

$$T^\mu{}_\nu := p^{\mu\alpha\beta}\partial_\nu g_{\alpha\beta} - \delta^\mu{}_\nu L.$$

Let X be a vector field whose flow leaves the Lagrangian form $L\,d^n x$ invariant. The Noether current associated with X is defined by

$$J^\mu_X := T^\mu_\nu X^\nu.$$

Furthermore, this current satisfies

$$\partial_\mu J^\mu_X = 0.$$

Equivalently, the $(n-1)$-form associated with J^μ_X is closed.

Example 6.4 As an example, suppose that the Lagrangian L is invariant under translations

$$x^\mu \longmapsto x^\mu + c^\mu,$$

where the c^μ are constants. The corresponding translation vector field has constant components. Hence, its Noether current satisfies

$$\partial_\mu J^\mu_X = 0.$$

Since

$$J^\mu_X = T^\mu{}_\nu c^\nu$$

and the constants c^ν are arbitrary, it follows that

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

This expresses the conservation of the canonical stress.

The Noether current associated with translations yields conserved energy and momentum. More precisely, the current satisfies

$$\partial_\mu J^\mu = 0.$$

In differential-form notation, the components J^μ determine the $(n-1)$-form

$$J := J^\mu \iota_{\partial_\mu}(d^n x),$$

and the conservation law is equivalent to

$$dJ = 0.$$

Thus, J is a closed form.

Integrating J over a spacelike hypersurface yields the corresponding conserved quantity. To express this quantity as an integral over the boundary of the hypersurface, as required for comparison with the ADM surface integral, one seeks an $(n-2)$-form G such that

$$J = dG.$$

Stokes' theorem then gives

$$\int_\Omega J = \int_{\partial\Omega} G.$$

However, a closed form need not be exact. Additional assumptions or a more specific construction are therefore required to obtain an exact Noether current.

Christodoulou then introduces a Minkowski background metric. Its isometry group is the ten-parameter Poincaré group, whose infinitesimal generators consist of four translations, three spatial rotations, and three boosts. For each Killing vector field X of the background metric, Christodoulou shows that, along solutions of the vacuum Einstein equations, the associated Noether current is exact:

$$J_X = dG_X,$$

where J_X is a 3-form and G_X is a 2-form.

Let Ω be a compact region in a spacelike hypersurface and let $S = \partial\Omega$. Christodoulou defines the surface integral associated with X by

$$Q_X(S) := \int_S G_X.$$

By Stokes' theorem,

$$Q_X(S) = \int_{\partial\Omega} G_X = \int_\Omega J_X.$$

For the Killing vector field generating time translations, the corresponding asymptotic surface integral is the total energy. If S_r denotes a coordinate sphere of radius r, then

$$E = \lim_{r\to\infty} \int_{S_r} e^i \, dS_i,$$

where

$$e^i = \frac{\sqrt{\det {}^3 g}}{4} \left({}^3 g^{jm\,3} g^{in} - {}^3 g^{ij\,3} g^{mn} \right) \partial_j {}^3 g_{mn}.$$

Here ${}^3 g_{ij}$ denotes the induced three-dimensional metric and dS_i is the Euclidean oriented surface element on S_r. Christodoulou mentions that this expression agrees with the ADM energy.

Remark 6.5 Under the prescribed asymptotic decay assumptions, Christodoulou defines the surface integral at spatial infinity by

$$Q_X(S_\infty) := \lim_{r\to\infty} Q_X(S_r).$$

Here S_∞ is notation for the limiting process. Christodoulou proves that, for translation vector fields, this limiting surface integral is a geometric invariant. His argument follows the approach of Bartnik [7]. For details, see Sections 3.1 and 3.2 of [18].

Christodoulou's construction uses Noether's theorem to associate conserved quantities with the symmetries of a reference Minkowski metric. For asymptotic translations, the corresponding surface integrals yield the total energy and momentum, and the energy integral agrees with the ADM mass. This approach therefore provides a clear geometric motivation for interpreting the ADM mass as an invariant of asymptotically flat initial data.

6.1.3 Comparison with the ADM formalism and Christodoulou's construction

In 1959, Arnowitt, Deser, and Misner formulated general relativity as a canonical Hamiltonian system. Their approach begins directly with a $3+1$ decomposition of spacetime and proceeds by identifying the constraint equations, imposing coordinate conditions, and reducing the system to its independent dynamical variables.

The construction of Chruściel, Jezierksi, and Kijowski [27], which is followed in this thesis, starts from a more general and geometric Hamiltonian framework for Lagrangian field theories and then specializes it to general relativity. The construction follows the same principle of studying field theories by examining their behaviour on a hypersurface. This seems to be a natural generalization of the ADM approach, which studies general relativity by splitting spacetime into $3+1$ dimensions. This requires a more abstract setup at the outset, but it has the advantage that the Hamiltonian and its boundary terms arise naturally from the underlying structure. The same framework can also be applied in several different asymptotic settings.

When specialized to general relativity, the two approaches begin with closely related Lagrangian formulations, as discussed in Section 5.1. Their emphases are different. ADM carry out an explicit canonical reduction by solving the constraint equations and imposing coordinate conditions, whereas Chruściel–Jezierksi–Kijowski derive a general Hamiltonian identity and then study the resulting boundary terms and the conditions under which the corresponding integrals converge. Under the asymptotically flat assumptions considered here, both constructions lead to the same expression for the ADM mass, up to the normalization and sign conventions adopted in each treatment. The ADM approach is more direct and historically foundational, while the formulation by Chruściel, Jezierksi, and Kijowski makes the underlying geometric structure and the role of the Hamiltonian more explicit.

Christodoulou's construction differs substantially from the Hamiltonian approach of Chruściel, Jezierksi, and Kijowski. Christodoulou works directly with the Lagrangian and applies Noether's theorem, rather than first passing to a Hamiltonian formulation. This elegantly motivates interpreting the ADM mass as the conserved surface integral associated with asymptotic translations.

The main advantage of the construction by Chruściel, Jezierksi, and Kijowski is its versatility. Their Hamiltonian framework is formulated for general Lagrangian field theories and can be applied in a variety of asymptotic spacetime settings. By contrast, in the discussion presented here, Christodoulou's construction is applied to the Poincaré symmetries of asymptotically flat spacetimes, with particular emphasis on the energy associated with asymptotic time translations.

It is also worth mentioning Michel's work [64], which identifies a common algebraic mechanism underlying the geometric invariance of several asymptotic mass quantities. Although we do not discuss this framework, it provides a broader explanation for why many mass and energy constructions in different asymptotic settings share similar structural properties.

6.2 Positive mass theorems

From a physical perspective, one expects the total energy of a gravitational system to be nonnegative. However, rigorously proving that the total energy is nonnegative is a difficult mathematical problem. Results of this type are collectively known as positive mass theorems. The first proofs were obtained by Schoen and Yau in the late 1970s, initially in dimension three, and their arguments were later extended to dimensions up to seven. Recent preprints from 2026 present new arguments for the Riemannian and spacetime positive mass theorems in arbitrary dimensions [10, 11].

In this section, we review the historical development of positive mass theorems. We state several of the principal results but omit their proofs. In Subsection 6.2.1, we focus on the asymptotically flat setting and highlight several important developments, including results concerning metrics of lower regularity. In Subsection 6.2.2, we discuss notions of mass for asymptotically hyperbolic Riemannian manifolds and asymptotically anti-de Sitter spacetimes, together with an overview of corresponding positive mass results.

For several results of Schoen and Yau, we use the modern formulations presented by Lee [52]. These formulations allow the original results to be stated consistently with the notation and terminology used throughout this thesis.

6.2.1 The positive mass theorem for asymptotically flat spacetimes

In the asymptotically flat setting, the positive mass theorem concerns the ADM mass of a Riemannian manifold or the ADM energy-momentum of a Lorentzian manifold. Under suitable geometric conditions, the theorem asserts that this mass or energy–momentum is nonnegative. Furthermore, the statement characterizes the case of zero mass with the absence of curvature.

The curvature and asymptotic hypotheses are essential. In the Riemannian setting, one assumes in particular that the scalar curvature is nonnegative. Standard definitions of asymptotic flatness frequently also require the scalar curvature to be integrable,

$$R \in L^1(M),$$

or impose stronger decay assumptions from which this integrability follows. Consequently, the integrability condition may be contained implicitly in the phrase asymptotically flat rather than stated separately in each positive mass theorem.

The first proof of the positive mass theorem was obtained by Schoen and Yau in 1979 [72]. Their paper contains two principal Riemannian results, which we state in the modern form presented by Lee [51].

Theorem 6.6 (Riemannian positive mass theorem) *Let (M, g) be a complete asymptotically flat three-dimensional Riemannian manifold with nonnegative scalar curvature. Then the ADM mass of each end of M is nonnegative.*

The same paper also established a rigidity result for the zero-mass case under an additional asymptotic regularity and decay assumption. Writing

$$g_{ij} - \delta_{ij} = h_{ij},$$

and assuming that the metric satisfies the asymptotic conditions specified in Subsection 5.3.1, Schoen and Yau impose the additional condition

$$|\partial\partial\partial h_{ij}| + |\partial\partial\partial\partial h_{ij}| + |\partial\partial\partial\partial\partial h_{ij}| \leq \frac{k}{1+r^5}, \tag{6.2}$$

where k is a positive constant.

Theorem 6.7 (Positive mass rigidity) *Let (M, g) be a complete asymptotically flat three-dimensional Riemannian manifold with nonnegative scalar curvature. Suppose, in addition, that (6.2) holds on some end M_k and that the ADM mass of M_k is zero. Then g is flat. In fact, (M, g) is isometric to $\mathbb{R}^3$ equipped with the standard Euclidean metric.*

Later in 1979, Schoen and Yau [76] announced a higher-dimensional extension of their argument.

Theorem 6.8 *Let (M, g) be a complete asymptotically flat Riemannian manifold of dimension $3 \leq n \leq 7$ with nonnegative scalar curvature. Then the ADM mass of each end of M is nonnegative. Furthermore, if the ADM mass of one end is zero, then g is flat.*

In 1981, Schoen and Yau gave a proof of the positive mass theorem for general $3+1$-dimensional asymptotically flat initial data sets [73]. An initial data set consists of a three-dimensional manifold M, a Riemannian metric g, and a symmetric two-tensor K representing the second fundamental form. The associated local energy density μ and momentum density J are related to (g, K) through the Einstein constraint equations. Under the relevant asymptotic conditions, the total mass of each end is nonnegative. If one end has zero total mass, then the initial data arise as the metric and second fundamental form induced on a spacelike hypersurface in Minkowski spacetime.

In 1981, Witten [86] proved the positive energy theorem by an alternative spinorial method. In the Riemannian setting, his method yields the positive mass theorem for asymptotically flat spin manifolds in arbitrary dimension. We state this result in its modern Riemannian form.

Theorem 6.9 (Positive mass theorem for spin manifolds) *Let (M^n, g) be a complete asymptotically flat spin manifold of dimension $n \geq 3$ with nonnegative scalar curvature. Then the ADM mass of each end of M is nonnegative. Moreover, if the ADM mass of any end is zero, then (M, g) is isometric to Euclidean space $(\mathbb{R}^n, \delta)$.*

In 1982, Parker and Taubes [65] gave a complete and mathematically rigorous treatment of Witten's argument. In a related contribution published in the same year, Reula [68] proved an existence theorem for solutions of Witten's equation, thereby addressing the solvability required in the spinorial proof. Both Witten and Yau received the Fields Medal in part for their work on this topic.

An important obstruction to positive scalar curvature provides one route to the Riemannian positive mass theorem. We state the result without proof and summarize its historical development below.

Theorem 6.10 (Torus connected-sum obstruction) *Let T^n be the n-dimensional torus, and let M be a closed n-dimensional manifold. Then $T^n \# M$ cannot carry a metric of positive scalar curvature.*

Remark 6.11 The operation # denotes the connected sum. Informally, it produces a new manifold from two manifolds by removing the interior of a small ball from each and gluing the resulting boundary spheres together. In the present setting, this construction can be written schematically as

$$T^n \# M = (T^n \setminus \mathrm{int}(B^n)) \cup_{S^{n-1}} (M \setminus \mathrm{int}(B^n)),$$

where B^n is an n-dimensional ball and $S^{n-1} = \partial B^n$ is its $(n-1)$-dimensional boundary sphere.

For $n \leq 7$, Theorem 6.10 was proved by Schoen and Yau in their 1979 work on manifolds of positive scalar curvature [74]. In 1980, Gromov and Lawson proved the corresponding result in arbitrary dimensions under a spin assumption [38]. In 1993, Smale [78] proved a result that implied the $n = 8$ case.

Lohkamp later presented an independent all-dimensional approach in a series of preprints beginning in 2006 and developed further through 2016 [57–60]. Independently, Schoen and Yau [75] proved Theorem 6.10 in all dimensions. For an explanation of how this scalar-curvature obstruction can be used to prove the Riemannian positive mass theorem, see Section 3.3 of Lee [52].

The year 2026 saw several major developments concerning the positive mass theorem in higher dimensions. In March, Bi, Hao, He, Shi, and Zhu proved the Riemannian positive mass theorem in dimensions up to 19 [8], building on the work of He, Shi, and Yu [40], as well as Chodosh, Mantoulidis and Schulze [15, 16].

In April 2026, Brendle and Wang presented a proof of the Riemannian positive mass theorem in arbitrary dimensions [10]. Their proof develops a new dimension-descent scheme and uses an argument inspired by the work of Bi, Hao, He, Shi, and Zhu [8]. Later that month, Brendle and Wang extended these ideas to the spacetime positive energy theorem in arbitrary dimensions [11]. At the time of writing, these results are available as arXiv preprints.

In June 2026, Khuri, Jian Wang, and Jinmin Wang proved an all-dimensional Riemannian positive mass theorem for asymptotically flat L^∞-metrics with subcritical singular sets [48]. Their proof uses arguments from both Bi, Hao, He, Shi, and Zhu [8] and Brendle and Wang [10].

Before turning to asymptotically hyperbolic geometry, it is worth mentioning the active study of low-regularity metrics in mathematical general relativity. In the Lorentzian setting, weak notions of curvature allow one to study spacetimes containing impulsive gravitational waves. For an overview, see LeFloch [55]. In the Riemannian setting, related techniques make it possible to formulate notions of scalar curvature and mass for metrics that are not smooth.

Lundgren and Meco [61] generalized the ADM mass to asymptotically Euclidean manifolds with metrics of local Sobolev regularity

$$W^{1,2}_{\mathrm{loc}} \cap L^\infty.$$

Under suitable asymptotic assumptions, they prove that this mass is finite and invariant under admissible changes of coordinates at infinity, and that it agrees with the classical ADM mass in the smooth setting. They also derive an expression involving the Ricci tensor which, for sufficiently regular metrics, agrees with the Ricci-tensor formulation of the ADM mass studied by Herzlich [43].

In February 2026, Hafemann [39] established a low-regularity Riemannian positive mass theorem for non-spin manifolds. More precisely, he considers asymptotically flat metrics in

$$C^0 \cap W^{1,n}_{\mathrm{loc}}$$

that are smooth outside a compact set and have nonnegative scalar curvature. His work builds on the formulation of scalar curvature and mass developed by Lee and LeFloch [53], and the work of Jiang, Sheng, and Zhang [46].

In June 2026, Mazurowski and Yao [62] introduced a harmonic mass for continuous asymptotically flat metrics on $\mathbb{R}^3$. This quantity agrees with the usual ADM mass whenever the metric is sufficiently regular and satisfies the appropriate fall-off conditions. It can also be recovered as a limit of the C^0 local mass introduced by Burkhardt-Guim [12]. Under their approximation and scalar-curvature hypotheses, Mazurowski and Yao prove that the harmonic mass is nonnegative and that vanishing harmonic mass implies flatness.

6.2.2 Positive mass results in asymptotically hyperbolic geometries

We now turn to positive mass results in asymptotically hyperbolic and asymptotically anti-de Sitter geometries. Although these settings differ substantially from the asymptotically flat case, they form an active area of research in mathematical general relativity and geometric analysis. In the Riemannian setting, one considers asymptotically hyperbolic manifolds, whereas in the Lorentzian setting the corresponding geometries are asymptotically anti-de Sitter spacetimes.

We do not rigorously define mass in these settings. Roughly speaking, in the Riemannian case one compares the metric with a fixed hyperbolic background metric and evaluates the asymptotic deviation against hyperbolic static potentials V. This leads to a boundary integral at infinity which defines the corresponding mass functional or energy–momentum vector.

For a more detailed treatment of mass in asymptotically hyperbolic Riemannian geometry, several sources are available. Wang's 2001 paper [84] is one of the foundational works in this area. Wang defines an invariant called the total mass and proves a positive mass theorem using spinorial methods. Chruściel and Herzlich [26] subsequently introduced mass integrals for a broader class of asymptotically hyperbolic Riemannian manifolds. For a more expository account of these mass formulas, see also Herzlich [42].

In the Lorentzian setting, the work of Henneaux and Teitelboim [41] is an important early source on asymptotically anti-de Sitter spacetimes. Chruściel and Nagy [28, 29] later gave a Hamiltonian definition of mass for spacelike hypersurfaces in spacetimes that are asymptotic to anti-de Sitter spacetime. They also proved that the resulting mass defines a geometric invariant of the spacelike hypersurface. For a discussion of how the Hamiltonian framework reviewed earlier in this thesis applies in this setting, see Section 5.5 of [27]. From this point onward, we abbreviate anti-de Sitter as AdS.

Andersson, Cai, and Galloway [2] proved an important non-spin positive mass result in the asymptotically hyperbolic Riemannian setting. They prove a scalar-curvature rigidity theorem and derive a positive mass theorem without assuming that the manifold is spin.

In 2019, Chruściel and Delay [23] proved a positive mass theorem for asymptotically hyperbolic Riemannian manifolds with spherical conformal infinity. They related the asymptotically hyperbolic energy–momentum vector to the corresponding positive mass theorem for asymptotically flat initial data sets satisfying the appropriate energy conditions.

In 2021, Sakovich [69] solved the Jang equation for three-dimensional asymptotically hyperbolic initial data sets and used this construction to prove the positive mass theorem without spinorial methods. For the well-definedness and coordinate invariance of the mass vector, Sakovich refers to the general framework developed by Michel [64].

Later that year, Chruściel and Galloway [25] proved positive mass theorems for asymptotically hyperbolic and asymptotically locally hyperbolic Riemannian manifolds with black-hole-type boundaries.

Finally, Chruściel, Delay, and Wutte [24] derived a formula for the energy of an asymptotically locally hyperbolic manifold obtained by gluing two such manifolds at conformal infinity.

In 2023, Dahl, Kröncke, and McCormick [32] introduced a new geometric quantity for asymptotically hyperbolic manifolds, called the volume-renormalized mass. This quantity is well defined under weaker fall-off assumptions than those generally required for the usual mass surface integrals. The authors also establish several positivity results for the volume-renormalized mass. In a subsequent paper [33], they show that this mass can be recovered from a reduced Hamiltonian framework.

In 2026, Hirsch, Khuri, Lesourd, and Zhang [44] used the recent Riemannian positive mass theorem of Brendle and Wang [10] to prove the spacetime positive mass theorem for asymptotically flat and asymptotically hyperboloidal initial data sets in arbitrary dimensions.

Recent preprints by Chruściel and Wutte study holographic energy in spacetimes with negative cosmological constant [30, 31]. Under specific hypotheses on the conformal boundary, they prove positivity of a weighted holographic energy. They also show that a relative holographic energy agrees with the corresponding relative Hamiltonian energy. We do not discuss these constructions in further detail here.

Chapter 7

Conclusion and outlook

Summary and main results

In this thesis, we have studied the Hamiltonian formulation of general relativity developed by Chruściel, Jezierski, and Kijowski [27], with particular emphasis on its geometric and symplectic structure and on the resulting notion of energy. The thesis's principal contributions are that it derives this construction in detail, reformulates some of its geometric ingredients using jet bundles, and explicitly verifies that the resulting Hamiltonian energy-momentum agrees with the ADM energy-momentum. To the best of our knowledge, this calculation has not previously appeared in the literature. Chruściel has confirmed in personal communication that he did not write out the calculation elsewhere.

The first four chapters develop the geometric and analytical framework needed for the main construction. In Chapter 2, we introduced the relevant tools from differential geometry, including integration on orientable and non-orientable manifolds, fibre bundles, and jet bundles. In particular, differential odd forms provide a coordinate-invariant theory of integration on non-orientable manifolds, while jet bundles encode derivatives of sections and therefore provide a natural setting for first-order Lagrangian field theories.

In Chapter 3, we reviewed the passage from Lagrangian mechanics to Hamiltonian mechanics and introduced the required concepts from symplectic geometry, including Hamiltonian vector fields and the Liouville one-form. We then formulated the notion of a dynamical system used throughout the thesis. Rather than treating the space of fields as an infinite-dimensional symplectic manifold, we work directly with a phase space of sections, a Hamiltonian function, and an antisymmetric two-form that encodes the corresponding Hamiltonian evolution.

In Chapter 4, we studied in detail the Hamiltonian field-theory framework developed by Chruściel, Jezierski, and Kijowski [27]. We used jet bundle theory to formulate the Lagrangian field theory, while a reference system identified the evolving fields with fields on a fixed model hypersurface Σ. Pulling the Lagrangian back to Σ and applying a Legendre transformation then produced a Hamiltonian formulation on the associated phase space. As part of this discussion, we reformulated the pointwise pullback construction in a coordinate-invariant jet bundle language and worked out several details of the resulting Hamiltonian dynamical system. We concluded by specialising the construction to the reference system later used for asymptotically flat spacetimes.

In Chapter 5, we applied this framework to general relativity. We represented gravitational fields as sections of an appropriate tensor bundle and constructed a reference system from the flow of a global spacetime vector field. After restricting to spacelike hypersurfaces, we reduced the symplectic form and the variation of the Hamiltonian to expressions involving the induced geometric data. In the asymptotically flat setting, the imposed decay conditions ensured the convergence of the relevant integrals and the vanishing of the required boundary terms. We then explicitly verified that the Hamiltonian energy-momentum arising from this construction agrees with the ADM energy-momentum. Finally, drawing on the work of Bartnik [7] and Chruściel [22], we examined the invariance of the resulting asymptotic quantities.

Broader context

In Chapter 6, we placed the Hamiltonian construction studied in the preceding chapters within the broader theory of gravitational energy. We first compared the original ADM Hamiltonian formulation [4, 5] with Christodoulou's [18] Lagrangian motivation and with the Hamiltonian framework of Chruściel, Jezierski, and Kijowski [27]. Although these approaches give different geometric perspectives, they lead to the same notion of energy in the asymptotically flat setting.

We then reviewed the historical development of positive mass theorems in asymptotically flat, asymptotically hyperbolic, and asymptotically AdS geometries. We particularly emphasized recent progress in arbitrary dimensions. These developments illustrate both the robustness of mass as a geometric invariant and the wide range of geometric and analytical settings in which positive mass results have been established.

Outlook

Several questions arising from the constructions studied in this thesis suggest directions for further investigation. First, the Legendre transformation used in Chapter 4 could be reformulated more systematically in the language of jet bundles. In studying this question, we considered the geometric treatment of Legendre transformations developed by Saunders [71] and the related field-theoretic constructions discussed by Forger and Romero [35]. Their precise application to the hypersurface-based framework of Chruściel, Jezierski, and Kijowski remains to be worked out. Such a reformulation could provide a more intrinsic explanation of the passage from the Lagrangian theory on spacetime to the Hamiltonian theory on the model hypersurface Σ.

A second question concerns the antisymmetric two-form on the phase space. In the construction considered here, this form is obtained by integrating a family of symplectic forms over Σ. It would be interesting to determine conditions under which integration preserves closedness and nondegeneracy. Furthermore, the phase space appearing in the construction is not the full space of sections of the momentum bundle associated to Σ, but a distinguished subset consisting of pulled back sections satisfying the conditions imposed by the theory. Consequently, a symplectic form on the full section space would not by itself provide a symplectic structure on the phase space. These issues are not addressed explicitly in the framework as presented by Chruściel, Jezierski, and Kijowski and provide a natural direction for further investigation.

A final direction is to extend the Hamiltonian framework to hypersurfaces or spacetimes with boundary. This would require identifying the additional boundary contributions in the variation of the Hamiltonian, determining suitable boundary conditions, and studying how these choices affect the resulting Hamiltonian and notion of energy.

Appendix A

Technical details of Subsection 5.1.2

In this appendix we restate the lemmas of Subsection 5.1.2 and give the corresponding proofs.

Lemma A.1 *For a metric density $\mathfrak{g}^{\mu\nu}$ one can express the Hilbert–Lagrangian as follows*

$$\mathfrak{g}^{\mu\nu} R_{\mu\nu} = \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\alpha\nu}) + \mathfrak{g}^{\mu\nu}\left[\Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\alpha\nu} - \Gamma^\sigma_{\alpha\sigma}\Gamma^\alpha_{\mu\nu}\right]. \tag{A.1}$$

Proof. Using the coordinate expression for the Ricci tensor, we obtain

$$\begin{aligned}\mathfrak{g}^{\mu\nu} R_{\mu\nu} &= \mathfrak{g}^{\mu\nu}\Big(\partial_\alpha\Gamma^\alpha_{\mu\nu} - \partial_\mu\Gamma^\alpha_{\alpha\nu} - \left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right]\Big)\\ &= \mathfrak{g}^{\mu\nu}\partial_\alpha\Gamma^\alpha_{\mu\nu} - \mathfrak{g}^{\mu\nu}\partial_\mu\Gamma^\alpha_{\alpha\nu} - \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right]\end{aligned}$$

We can move the metric density inside the derivative by using the product rule

$$\mathfrak{g}^{\mu\nu}\partial_\alpha\Gamma^\alpha_{\mu\nu} = \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - (\partial_\alpha\mathfrak{g}^{\mu\nu})\Gamma^\alpha_{\mu\nu} \tag{A.2}$$

To evaluate the partial derivative of our density we again use the product rule

$$\begin{aligned}\partial_\alpha\mathfrak{g}^{\mu\nu} &= \frac{1}{16\pi}\Big((\partial_\alpha\sqrt{-\det g})g^{\mu\nu} + \partial_\alpha(g^{\mu\nu})\sqrt{-\det g}\Big)\\ &= \frac{1}{16\pi}\Big(\frac{1}{2\sqrt{-\det g}}\partial_\alpha(-\det g)g^{\mu\nu} + (-g^{\sigma\nu}\Gamma^\mu_{\alpha\sigma} - g^{\mu\sigma}\Gamma^\nu_{\alpha\sigma})\sqrt{-\det g}\Big)\end{aligned} \tag{A.3}$$

The last equality is because we are using the Levi-Civita connection, which in particular is compatible with our metric g and g^{-1}. This is show below in local coordinates:

$$0 = g^{\mu\nu}{}_{;\alpha} = \partial_\alpha(g^{\mu\nu}) + g^{\sigma\nu}\Gamma^\mu_{\alpha\sigma} + g^{\mu\sigma}\Gamma^\nu_{\alpha\sigma} \tag{A.4}$$

$$0 = g_{\mu\nu;\alpha} = \partial_\alpha(g_{\mu\nu}) - g_{\sigma\nu}\Gamma^\sigma_{\alpha\mu} - g_{\mu\sigma}\Gamma^\sigma_{\alpha\nu}. \tag{A.5}$$

We continue the rewriting of equation A.3, notice the trace we obtained from the derivative of the determinant expressed as a contraction in local coordinates

$$\begin{aligned}\partial_\alpha\mathfrak{g}^{\mu\nu} &= \frac{1}{16\pi}\Big(\frac{1}{2}\sqrt{-\det g}g^{\mu\nu}\partial_\alpha g_{\mu\nu}g^{\mu\nu} + (-g^{\sigma\nu}\Gamma^\mu_{\alpha\sigma} - g^{\mu\sigma}\Gamma^\nu_{\alpha\sigma})\sqrt{-\det g}\Big)\\ &= \mathfrak{g}^{\mu\nu}\frac{1}{2}(\Gamma^\sigma_{\alpha\mu}g_{\sigma\nu} + \Gamma^\sigma_{\alpha\nu}g_{\mu\sigma})g^{\mu\nu} - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\alpha\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\sigma\alpha}\\ &= \mathfrak{g}^{\mu\nu}\frac{1}{2}(\Gamma^\sigma_{\alpha\mu}\delta^\mu_\sigma + \Gamma^\sigma_{\alpha\nu}\delta^\nu_\sigma) - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\alpha\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\sigma\alpha}\\ &= \mathfrak{g}^{\mu\nu}\Gamma^\sigma_{\alpha\sigma} - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\alpha\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\sigma\alpha}\end{aligned}$$

With this expression we can go back and rewrite the Hilbert-Lagrange density A.2

$$\begin{aligned}
\tilde{L} &= \mathfrak{g}^{\mu\nu}\partial_\alpha\Gamma^\alpha_{\mu\nu} - \mathfrak{g}^{\mu\nu}\partial_\mu\Gamma^\alpha_{\alpha\nu} - \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right] \\
&= \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\alpha\nu}) \\
&- (\partial_\alpha\mathfrak{g}^{\mu\nu})\Gamma^\alpha_{\mu\nu} + (\partial_\mu\mathfrak{g}^{\mu\nu})\Gamma^\alpha_{\alpha\nu} - \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right] \\
&= \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\alpha\nu}) \\
&- (\mathfrak{g}^{\mu\nu}\Gamma^\sigma_{\alpha\sigma} - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\alpha\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\sigma\alpha})\Gamma^\alpha_{\mu\nu} + (\mathfrak{g}^{\mu\nu}\Gamma^\sigma_{\mu\sigma} - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\mu\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\sigma\mu})\Gamma^\alpha_{\alpha\nu} \\
&- \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right] \\
&= \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\alpha\nu}) \\
&- \mathfrak{g}^{\mu\nu}\Gamma^\sigma_{\alpha\sigma}\Gamma^\alpha_{\mu\nu} + \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\alpha\sigma}\Gamma^\alpha_{\mu\nu} + \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\sigma\alpha}\Gamma^\alpha_{\mu\nu} \\
&+ \mathfrak{g}^{\mu\nu}\Gamma^\sigma_{\mu\sigma}\Gamma^\alpha_{\alpha\nu} - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\mu\sigma}\Gamma^\alpha_{\alpha\nu} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\sigma\mu}\Gamma^\alpha_{\alpha\nu} \\
&- \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right] \\
&= \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\alpha\nu}) \\
&+ \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\mu\sigma}\Gamma^\alpha_{\alpha\nu} - \Gamma^\sigma_{\alpha\sigma}\Gamma^\alpha_{\mu\nu}\right] + \mathfrak{g}^{\sigma\nu}\left[\Gamma^\mu_{\alpha\sigma}\Gamma^\alpha_{\mu\nu} - \Gamma^\mu_{\mu\sigma}\Gamma^\alpha_{\alpha\nu}\right] \\
&+ \mathfrak{g}^{\mu\sigma}\left[\Gamma^\nu_{\sigma\alpha}\Gamma^\alpha_{\mu\nu} - \Gamma^\nu_{\sigma\mu}\Gamma^\alpha_{\alpha\nu}\right] - \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right] \\
&= \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\alpha\nu}) \\
&+ \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\mu\sigma}\Gamma^\alpha_{\alpha\nu} - \Gamma^\sigma_{\alpha\sigma}\Gamma^\alpha_{\mu\nu}\right] + \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\alpha\mu}\Gamma^\alpha_{\sigma\nu} - \Gamma^\sigma_{\sigma\mu}\Gamma^\alpha_{\alpha\nu}\right] \\
&+ \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\nu\alpha}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\nu\mu}\Gamma^\alpha_{\alpha\sigma}\right] - \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\sigma} - \Gamma^\sigma_{\mu\nu}\Gamma^\alpha_{\alpha\sigma}\right] \\
&= \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\alpha\nu}) + \mathfrak{g}^{\mu\nu}\left[\Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\alpha\nu} - \Gamma^\alpha_{\mu\nu}\Gamma^\sigma_{\alpha\sigma}\right].
\end{aligned}$$

Notice that we switched around the Γ's which can be done since both are smooth functions. In conclusion we find $\tilde{L} = \partial_\alpha\big(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\mu\nu}\big) - \partial_\mu(\mathfrak{g}^{\mu\nu}\Gamma^\alpha_{\alpha\nu}) + \mathfrak{g}^{\mu\nu}\left[\Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\alpha\nu} - \Gamma^\sigma_{\alpha\sigma}\Gamma^\alpha_{\mu\nu}\right]$. □

Lemma A.2 *Let B be a symmetric background connection on M, and let $r_{\mu\nu}$ be its Ricci tensor. The contraction is locally written as*

$$\begin{aligned}
\mathfrak{g}^{\mu\nu}r_{\mu\nu} &= \partial_\alpha(\mathfrak{g}^{\mu\nu}B^\alpha_{\mu\nu}) - \partial_\mu(\mathfrak{g}^{\mu\nu}B^\alpha_{\alpha\nu}) - \mathfrak{g}^{\mu\nu}\left[B^\sigma_{\alpha\nu}B^\alpha_{\mu\sigma} - B^\sigma_{\mu\nu}B^\alpha_{\alpha\sigma}\right] && \text{(A.6)} \\
&+ \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\nu\alpha}B^\alpha_{\mu\sigma} + \Gamma^\sigma_{\alpha\mu}B^\alpha_{\sigma\nu} - \Gamma^\sigma_{\nu\mu}B^\alpha_{\alpha\sigma} - \Gamma^\sigma_{\alpha\sigma}B^\alpha_{\mu\nu}\right] && \text{(A.7)}
\end{aligned}$$

Proof. Starting from the coordinate expression for the Ricci tensor of B, we have

$$\begin{aligned}
\mathfrak{g}^{\mu\nu}r_{\mu\nu} &= \mathfrak{g}^{\mu\nu}\left(\partial_\alpha B^\alpha_{\mu\nu} - \partial_\mu B^\alpha_{\alpha\nu} - \left[B^\sigma_{\alpha\nu}B^\alpha_{\mu\sigma} - B^\sigma_{\mu\nu}B^\alpha_{\alpha\sigma}\right]\right) \\
&= \mathfrak{g}^{\mu\nu}\partial_\alpha B^\alpha_{\mu\nu} - \mathfrak{g}^{\mu\nu}\partial_\mu B^\alpha_{\alpha\nu} - \mathfrak{g}^{\mu\nu}\left[B^\sigma_{\alpha\nu}B^\alpha_{\mu\sigma} - B^\sigma_{\mu\nu}B^\alpha_{\alpha\sigma}\right].
\end{aligned}$$

Applying the product rule to the first two terms gives

$$\mathfrak{g}^{\mu\nu}r_{\mu\nu} = \partial_\alpha(\mathfrak{g}^{\mu\nu}B^\alpha_{\mu\nu}) - \partial_\mu(\mathfrak{g}^{\mu\nu}B^\alpha_{\alpha\nu}) - B^\alpha_{\mu\nu}\partial_\alpha\mathfrak{g}^{\mu\nu} + B^\alpha_{\alpha\nu}\partial_\mu\mathfrak{g}^{\mu\nu} - \mathfrak{g}^{\mu\nu}\left[B^\sigma_{\alpha\nu}B^\alpha_{\mu\sigma} - B^\sigma_{\mu\nu}B^\alpha_{\alpha\sigma}\right].$$

Recall the equivalent of the metricity condition of Γ, given in equation (5.6)

$$\partial_\beta\mathfrak{g}^{\mu\nu} = \mathfrak{g}^{\mu\nu}\Gamma^\sigma_{\beta\sigma} - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\beta\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\beta\sigma}.$$

Substituting this for $\beta = \alpha$ and $\beta = \mu$, we obtain

$$\begin{aligned}
\mathfrak{g}^{\mu\nu}r_{\mu\nu} &= \partial_\alpha(\mathfrak{g}^{\mu\nu}B^\alpha_{\mu\nu}) - \partial_\mu(\mathfrak{g}^{\mu\nu}B^\alpha_{\alpha\nu}) - \mathfrak{g}^{\mu\nu}\left[B^\sigma_{\alpha\nu}B^\alpha_{\mu\sigma} - B^\sigma_{\mu\nu}B^\alpha_{\alpha\sigma}\right] \\
&- \left(\mathfrak{g}^{\mu\nu}\Gamma^\sigma_{\alpha\sigma} - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\alpha\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\alpha\sigma}\right)B^\alpha_{\mu\nu} \\
&+ \left(\mathfrak{g}^{\mu\nu}\Gamma^\sigma_{\mu\sigma} - \mathfrak{g}^{\sigma\nu}\Gamma^\mu_{\mu\sigma} - \mathfrak{g}^{\mu\sigma}\Gamma^\nu_{\mu\sigma}\right)B^\alpha_{\alpha\nu}.
\end{aligned}$$

After relabelling dummy indices and using the symmetry of $\mathfrak{g}^{\mu\nu}$, the terms involving Γ and B become

$$\mathfrak{g}^{\mu\nu}\left[\Gamma^{\sigma}_{\nu\alpha}B^{\alpha}_{\mu\sigma}+\Gamma^{\sigma}_{\alpha\mu}B^{\alpha}_{\sigma\nu}-\Gamma^{\sigma}_{\nu\mu}B^{\alpha}_{\alpha\sigma}-\Gamma^{\sigma}_{\alpha\sigma}B^{\alpha}_{\mu\nu}\right].$$

Therefore

$$\begin{aligned}\mathfrak{g}^{\mu\nu}r_{\mu\nu}=\partial_\alpha(\mathfrak{g}^{\mu\nu}B^{\alpha}_{\mu\nu})-\partial_\mu(\mathfrak{g}^{\mu\nu}B^{\alpha}_{\alpha\nu})-\mathfrak{g}^{\mu\nu}\left[B^{\sigma}_{\alpha\nu}B^{\alpha}_{\mu\sigma}-B^{\sigma}_{\mu\nu}B^{\alpha}_{\alpha\sigma}\right]\\+\mathfrak{g}^{\mu\nu}\left[\Gamma^{\sigma}_{\nu\alpha}B^{\alpha}_{\mu\sigma}+\Gamma^{\sigma}_{\alpha\mu}B^{\alpha}_{\sigma\nu}-\Gamma^{\sigma}_{\nu\mu}B^{\alpha}_{\alpha\sigma}-\Gamma^{\sigma}_{\alpha\sigma}B^{\alpha}_{\mu\nu}\right],\end{aligned}$$

which proves the desired identity. □

Lemma A.3 *The Ricci tensor can be expressed using symmetrization, as follows*

$$R_{\mu\nu}=\partial_\alpha\left[\Gamma^{\alpha}_{\mu\nu}-\delta^{\alpha}_{(\mu}\Gamma^{\kappa}_{\nu)\kappa}\right]-\left[\Gamma^{\sigma}_{\alpha\nu}\Gamma^{\alpha}_{\mu\sigma}-\Gamma^{\sigma}_{\mu\nu}\Gamma^{\alpha}_{\alpha\sigma}\right].$$

Proof. We only need to prove that

$$\partial_\alpha\left(\delta^{\alpha}_{(\mu}\Gamma^{\kappa}_{\nu)\kappa}\right)=\partial_\mu\Gamma^{\alpha}_{\alpha\nu}.$$

By the definition of symmetrization, we have

$$\begin{aligned}\partial_\alpha\left(\delta^{\alpha}_{(\mu}\Gamma^{\kappa}_{\nu)\kappa}\right)&=\frac{1}{2}\partial_\alpha\left(\delta^{\alpha}_{\mu}\Gamma^{\kappa}_{\nu\kappa}+\delta^{\alpha}_{\nu}\Gamma^{\kappa}_{\mu\kappa}\right)\\&=\frac{1}{2}\left(\partial_\mu\Gamma^{\kappa}_{\nu\kappa}+\partial_\nu\Gamma^{\kappa}_{\mu\kappa}\right).\end{aligned}$$

Since the Levi-Civita connection is torsion-free, we have $\Gamma^{\kappa}_{\nu\kappa}=\Gamma^{\kappa}_{\kappa\nu}$. Moreover, for the Levi-Civita connection,

$$\Gamma^{\kappa}_{\kappa\nu}=\frac{1}{2\det g}\partial_\nu\det g=\partial_\nu\ln\sqrt{-\det g}.$$

Hence

$$\begin{aligned}\partial_\nu\Gamma^{\kappa}_{\mu\kappa}&=\partial_\nu\Gamma^{\kappa}_{\kappa\mu}\\&=\partial_\nu\partial_\mu\ln\sqrt{-\det g}\\&=\partial_\mu\partial_\nu\ln\sqrt{-\det g}\\&=\partial_\mu\Gamma^{\kappa}_{\kappa\nu}.\end{aligned}$$

Therefore

$$\begin{aligned}\partial_\alpha\left(\delta^{\alpha}_{(\mu}\Gamma^{\kappa}_{\nu)\kappa}\right)&=\frac{1}{2}\left(\partial_\mu\Gamma^{\kappa}_{\kappa\nu}+\partial_\mu\Gamma^{\kappa}_{\kappa\nu}\right)\\&=\partial_\mu\Gamma^{\kappa}_{\kappa\nu}.\end{aligned}$$

Relabelling the dummy index κ as α gives

$$\partial_\alpha\left(\delta^{\alpha}_{(\mu}\Gamma^{\kappa}_{\nu)\kappa}\right)=\partial_\mu\Gamma^{\alpha}_{\alpha\nu}.$$

This proves the claimed rewriting of the Ricci tensor. □

Lemma A.4 *The Christoffel symbols of g can be recovered from the background connection $B^{\alpha}_{\mu\nu}$ and the tensor $p^{\alpha}_{\mu\nu}$ by*

$$\Gamma^{\alpha}_{\mu\nu} = B^{\alpha}_{\mu\nu} - p^{\alpha}_{\mu\nu} + \frac{2}{3}\delta^{\alpha}_{(\mu} p^{\kappa}_{\nu)\kappa}. \tag{A.8}$$

Proof. We first begin by rewriting our tensor p to get the Christoffel symbols Γ

$$\begin{aligned} p^{\alpha}_{\mu\nu} &= \left(B^{\alpha}_{\mu\nu} - \delta^{\alpha}_{(\mu} B^{\kappa}_{\nu)\kappa}\right) - \left(\Gamma^{\alpha}_{\mu\nu} - \delta^{\alpha}_{(\mu}\Gamma^{\kappa}_{\nu)\kappa}\right) \\ \Rightarrow \Gamma^{\alpha}_{\mu\nu} &= B^{\alpha}_{\mu\nu} - p^{\alpha}_{\mu\nu} - \delta^{\alpha}_{(\mu} B^{\kappa}_{\nu)\kappa} + \delta^{\alpha}_{(\mu}\Gamma^{\kappa}_{\nu)\kappa}. \end{aligned}$$

We now look at the only tensor that differs in both expressions

$$\begin{aligned} \frac{2}{3}\delta^{\alpha}_{(\mu} p^{\kappa}_{\nu)\kappa} &= \frac{2}{3}\big[\frac{1}{2}\delta^{\alpha}_{\mu} p^{\kappa}_{\nu\kappa} + \frac{1}{2}\delta^{\alpha}_{\nu} p^{\kappa}_{\mu\kappa}\big] \\ &= \frac{1}{3}\delta^{\alpha}_{\mu}\big[\left(B^{\kappa}_{\nu\kappa} - \delta^{\kappa}_{(\nu} B^{\sigma}_{\kappa)\sigma}\right) - \left(\Gamma^{\kappa}_{\nu\kappa} - \delta^{\kappa}_{(\nu}\Gamma^{\sigma}_{\kappa)\sigma}\right)\big] \\ &+ \frac{1}{3}\delta^{\alpha}_{\nu}\big[\left(B^{\kappa}_{\mu\kappa} - \delta^{\kappa}_{(\mu} B^{\sigma}_{\kappa)\sigma}\right) - \left(\Gamma^{\kappa}_{\mu\kappa} - \delta^{\kappa}_{(\mu}\Gamma^{\sigma}_{\kappa)\sigma}\right)\big]. \end{aligned}$$

We now apply that we work in a 4 dimensional spacetime and work out the symmetric tensors involving B

$$\begin{aligned} \delta^{\kappa}_{(\nu} B^{\sigma}_{\kappa)\sigma} &= \frac{1}{2}\big[\delta^{\kappa}_{\nu} B^{\sigma}_{\kappa\sigma} + \delta^{\kappa}_{\kappa} B^{\sigma}_{\nu\sigma}\big] = \frac{1}{2}\big[B^{\sigma}_{\nu\sigma} + 4B^{\sigma}_{\nu\sigma}\big] = \frac{5}{2}B^{\sigma}_{\nu\sigma} \\ \delta^{\kappa}_{(\mu} B^{\sigma}_{\kappa)\sigma} &= \frac{1}{2}\big[\delta^{\kappa}_{\mu} B^{\sigma}_{\kappa\sigma} + \delta^{\kappa}_{\kappa} B^{\sigma}_{\mu\sigma}\big] = \frac{5}{2}B^{\sigma}_{\mu\sigma}. \end{aligned}$$

We do the same for the symmetric tensors involving Γ

$$\begin{aligned} \delta^{\kappa}_{(\nu}\Gamma^{\sigma}_{\kappa)\sigma} &= \frac{1}{2}\big[\delta^{\kappa}_{\nu}\Gamma^{\sigma}_{\kappa\sigma} + \delta^{\kappa}_{\kappa}\Gamma^{\sigma}_{\nu\sigma}\big] = \frac{5}{2}\Gamma^{\sigma}_{\nu\sigma} \\ \delta^{\kappa}_{(\mu}\Gamma^{\sigma}_{\kappa)\sigma} &= \frac{1}{2}\big[\delta^{\kappa}_{\mu}\Gamma^{\sigma}_{\kappa\sigma} + \delta^{\kappa}_{\kappa}\Gamma^{\sigma}_{\mu\sigma}\big] = \frac{5}{2}\Gamma^{\sigma}_{\mu\sigma}. \end{aligned}$$

Notice that the σ and κ are dummy variables so we can switch them. We then recover our original symmetric tensors

$$\begin{aligned} \frac{2}{3}\delta^{\alpha}_{(\mu} p^{\kappa}_{\nu)\kappa} &= \frac{1}{3}\delta^{\alpha}_{\mu}\big[\left(B^{\kappa}_{\nu\kappa} - \frac{5}{2}B^{\kappa}_{\nu\kappa}\right) - \left(\Gamma^{\kappa}_{\nu\kappa} - \frac{5}{2}\Gamma^{\kappa}_{\nu\kappa}\right)\big] \\ &+ \frac{1}{3}\delta^{\alpha}_{\nu}\big[\left(B^{\kappa}_{\mu\kappa} - \frac{5}{2}B^{\kappa}_{\mu\kappa}\right) - \left(\Gamma^{\kappa}_{\mu\kappa} - \frac{5}{2}\Gamma^{\kappa}_{\mu\kappa}\right)\big] \\ &= \frac{1}{3}\left(\delta^{\alpha}_{\mu}\big[-\frac{3}{2}B^{k}_{\nu\kappa} + \frac{3}{2}\Gamma^{\kappa}_{\nu\kappa}\big] + \delta^{\alpha}_{\nu}\big[-\frac{3}{2}B^{\kappa}_{\mu\kappa} + \frac{3}{2}\Gamma^{\kappa}_{\nu\kappa}\big]\right) \\ &= -\frac{1}{2}\big[\delta^{\alpha}_{\mu} B^{\kappa}_{\nu\kappa} + \delta^{\alpha}_{\nu} B^{\kappa}_{\mu\kappa}\big] + \frac{1}{2}\big[\delta^{\alpha}_{\mu}\Gamma^{\kappa}_{\nu\kappa} + \delta^{\alpha}_{\mu}\Gamma^{\kappa}_{\nu\kappa}\big] \\ &= -\delta^{\alpha}_{(\mu} B^{\kappa}_{\nu)\kappa} + \delta^{\alpha}_{(\mu}\Gamma^{\kappa}_{\nu)\kappa}. \end{aligned}$$

□

Proposition A.5 *We can express the Hilbert Lagrangian, using the momentum p and the Lagrangian L as follows*

$$\mathfrak{g}^{\mu\nu} R_{\mu\nu} = -\partial_{\alpha}(\mathfrak{g}^{\mu\nu} p^{\alpha}_{\mu\nu}) + L, \tag{A.9}$$

where L is defined as above.

Proof. We can obtain the identity by substracting the contraction 5.7 on our connection B from the Hilbert Lagrangian density 5.5. More precisely, we find

$$\begin{aligned}
\mathfrak{g}^{\mu\nu}R_{\mu\nu} - \mathfrak{g}^{\mu\nu}r_{\mu\nu} &= \partial_\alpha\big(\mathfrak{g}^{\mu\nu}(\Gamma^\alpha_{\mu\nu} - \delta^\alpha_{(\mu}\Gamma^\kappa_{\nu)\kappa})\big) + \mathfrak{g}^{\mu\nu}\left[\Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\alpha\nu} - \Gamma^\sigma_{\alpha\sigma}\Gamma^\alpha_{\mu\nu}\right] \\
&\quad - \Big[\partial_\alpha\big[\mathfrak{g}^{\mu\nu}(B^\alpha_{\mu\nu} - \delta^\alpha_{(\mu}B^\kappa_{\nu)})\big] - \mathfrak{g}^{\mu\nu}\left[B^\sigma_{\alpha\nu}B^\alpha_{\mu\sigma} - B^\sigma_{\mu\nu}B^\alpha_{\alpha\sigma}\right] \\
&\quad + \mathfrak{g}^{\mu\nu}\left[\Gamma^\sigma_{\nu\alpha}B^\alpha_{\mu\sigma} + \Gamma^\sigma_{\alpha\mu}B^\alpha_{\sigma\nu} - \Gamma^\sigma_{\nu\mu}B^\alpha_{\alpha\sigma} - \Gamma^\sigma_{\alpha\sigma}B^\alpha_{\mu\nu}\right]\Big] \\
&\quad = -\partial_\alpha(\mathfrak{g}^{\mu\nu}p^\alpha_{\mu\nu}) + K,
\end{aligned}$$

where we catch all the rest terms with the variable K. To complete the proof we need to show $\mathfrak{g}^{\mu\nu}r_{\mu\nu} + K = L$. This is a straight forward calculation, shown below

$$\begin{aligned}
\mathfrak{g}^{\mu\nu}r_{\mu\nu} + K = \mathfrak{g}^{\mu\nu}\Big[&(\Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\alpha\nu} - \Gamma^\alpha_{\sigma\mu}B^\sigma_{\alpha\nu} - B^\alpha_{\mu\sigma}\Gamma^\sigma_{\nu\alpha} + B^\alpha_{\mu\sigma}B^\sigma_{\alpha\nu}) \\
&- (\Gamma^\sigma_{\alpha\sigma}\Gamma^\alpha_{\mu\nu} - \Gamma^\sigma_{\alpha\sigma}B^\alpha_{\mu\nu} - \Gamma^\alpha_{\mu\nu}B^\sigma_{\sigma\alpha} + B^\alpha_{\mu\nu}B^\sigma_{\sigma\alpha}) + r_{\mu\nu}\Big] \\
= \mathfrak{g}^{\mu\nu}\Big[&(\Gamma^\alpha_{\sigma\mu} - B^\alpha_{\mu\sigma})(\Gamma^\sigma_{\nu\alpha} - B^\sigma_{\alpha\nu}) - (\Gamma^\alpha_{\mu\nu} - B^\alpha_{\mu\nu})(\Gamma^\sigma_{\alpha\sigma} - B^\sigma_{\alpha\sigma}) + r_{\mu\nu}\Big]
\end{aligned}$$

□

Lemma A.6 *Let ∇^B denote the covariant derivative associated to the background connection B. Then*

$$\mathfrak{g}^{\mu\nu}{}_{;\alpha} = \mathfrak{g}^{\mu\nu}(\Gamma^\sigma_{\alpha\sigma} - B^\sigma_{\alpha\sigma}) - \mathfrak{g}^{\mu\sigma}(\Gamma^\nu_{\sigma\alpha} - B^\nu_{\sigma\alpha}) - \mathfrak{g}^{\nu\sigma}(\Gamma^\mu_{\sigma\alpha} - B^\mu_{\sigma\alpha}), \tag{A.10}$$

where $\mathfrak{g}^{\mu\nu}{}_{;\alpha} := \nabla^B_\alpha\mathfrak{g}^{\mu\nu}$.

Proof. We compute the covariant derivative of $\mathfrak{g}^{\mu\nu}$ with respect to the background connection B. Since

$$\mathfrak{g}^{\mu\nu} = \frac{1}{16\pi}\sqrt{-\det g}\, g^{\mu\nu},$$

the product rule gives

$$\mathfrak{g}^{\mu\nu}{}_{;\alpha} = \frac{1}{16\pi}\left[\left(\nabla^B_\alpha\sqrt{-\det g}\right)g^{\mu\nu} + \sqrt{-\det g}\,\nabla^B_\alpha g^{\mu\nu}\right]. \tag{A.11}$$

We first compute the covariant derivative of the density factor. Since $\nabla^B_\alpha\partial_\rho = B^\sigma_{\alpha\rho}\partial_\sigma$, the induced covariant derivative on densities gives

$$\nabla^B_\alpha\sqrt{-\det g} = \partial_\alpha\sqrt{-\det g} - B^\sigma_{\alpha\sigma}\sqrt{-\det g}.$$

Moreover,

$$\partial_\alpha\sqrt{-\det g} = \frac{1}{2}\sqrt{-\det g}\, g^{\rho\sigma}\partial_\alpha g_{\rho\sigma}.$$

Because Γ is the Levi-Civita connection of g, metric compatibility gives

$$\partial_\alpha g_{\rho\sigma} = g_{\lambda\sigma}\Gamma^\lambda_{\alpha\rho} + g_{\rho\lambda}\Gamma^\lambda_{\alpha\sigma}.$$

Hence

$$\begin{aligned}\frac{1}{2}g^{\rho\sigma}\partial_\alpha g_{\rho\sigma} &= \frac{1}{2}g^{\rho\sigma}\left(g_{\lambda\sigma}\Gamma^\lambda_{\alpha\rho} + g_{\rho\lambda}\Gamma^\lambda_{\alpha\sigma}\right)\\ &= \Gamma^\lambda_{\alpha\lambda}.\end{aligned}$$

Therefore

$$\nabla^B_\alpha\sqrt{-\det g} = \left(\Gamma^\sigma_{\alpha\sigma} - B^\sigma_{\alpha\sigma}\right)\sqrt{-\det g}. \tag{A.12}$$

Next, since $g^{\mu\nu}$ is a contravariant tensor, its covariant derivative with respect to B is

$$\nabla^B_\alpha g^{\mu\nu} = \partial_\alpha g^{\mu\nu} + B^\mu_{\alpha\sigma}g^{\sigma\nu} + B^\nu_{\alpha\sigma}g^{\mu\sigma}.$$

Using metric compatibility of the Levi-Civita connection again,

$$\partial_\alpha g^{\mu\nu} = -\Gamma^\mu_{\alpha\sigma}g^{\sigma\nu} - \Gamma^\nu_{\alpha\sigma}g^{\mu\sigma}.$$

Thus

$$\nabla^B_\alpha g^{\mu\nu} = -\left(\Gamma^\mu_{\alpha\sigma} - B^\mu_{\alpha\sigma}\right)g^{\sigma\nu} - \left(\Gamma^\nu_{\alpha\sigma} - B^\nu_{\alpha\sigma}\right)g^{\mu\sigma}. \tag{A.13}$$

Substituting both expressions (A.12)–(A.13) into the equation (A.11) gives

$$\mathfrak{g}^{\mu\nu}{}_{;\alpha} = \mathfrak{g}^{\mu\nu}\left(\Gamma^\sigma_{\alpha\sigma} - B^\sigma_{\alpha\sigma}\right) - \mathfrak{g}^{\sigma\nu}\left(\Gamma^\mu_{\alpha\sigma} - B^\mu_{\alpha\sigma}\right) - \mathfrak{g}^{\mu\sigma}\left(\Gamma^\nu_{\alpha\sigma} - B^\nu_{\alpha\sigma}\right).$$

This proves the identity. □

Lemma A.7 *The tensor $p^\alpha_{\mu\nu}$ can be written as*

$$p^\lambda_{\mu\nu} = \frac{1}{2}\mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\nu} + \frac{1}{2}\mathfrak{g}_{\nu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\mu} - \frac{1}{2}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\sigma\mu}\mathfrak{g}_{\rho\nu}\mathfrak{g}^{\sigma\rho}{}_{;\alpha} + \frac{1}{4}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\mu\nu}\mathfrak{g}_{\sigma\rho}\mathfrak{g}^{\sigma\rho}{}_{;\alpha}. \tag{A.14}$$

Here $\mathfrak{g}_{\mu\nu}$ denotes the inverse matrix of $\mathfrak{g}^{\mu\nu}$, so that $\mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\alpha\nu} = \delta^\nu_\mu$.

Proof. Define

$$C^\alpha_{\mu\nu} := \Gamma^\alpha_{\mu\nu} - B^\alpha_{\mu\nu}.$$

Since both Γ and B are symmetric connections, $C^\alpha_{\mu\nu}$ is symmetric in the lower indices. In terms of C, the definition of $p^\lambda_{\mu\nu}$ becomes

$$p^\lambda_{\mu\nu} = -C^\lambda_{\mu\nu} + \delta^\lambda_{(\mu}C^\kappa_{\nu)\kappa}.$$

By the previous lemma, we have

$$\mathfrak{g}^{\mu\nu}{}_{;\alpha} = \mathfrak{g}^{\mu\nu}C^\sigma_{\alpha\sigma} - \mathfrak{g}^{\mu\sigma}C^\nu_{\sigma\alpha} - \mathfrak{g}^{\nu\sigma}C^\mu_{\sigma\alpha}.$$

We now compute the four terms appearing in the claimed expression. First,

$$\begin{aligned}\frac{1}{2}\mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\nu} &= \frac{1}{2}\mathfrak{g}_{\mu\alpha}\left(\mathfrak{g}^{\lambda\alpha}C^\sigma_{\nu\sigma} - \mathfrak{g}^{\lambda\sigma}C^\alpha_{\sigma\nu} - \mathfrak{g}^{\alpha\sigma}C^\lambda_{\sigma\nu}\right)\\ &= \frac{1}{2}\left(\delta^\lambda_\mu C^\sigma_{\nu\sigma} - \mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\lambda\sigma}C^\alpha_{\sigma\nu} - C^\lambda_{\mu\nu}\right).\end{aligned}$$

Similarly,

$$\frac{1}{2}\mathfrak{g}_{\nu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\mu} = \frac{1}{2}\left(\delta^\lambda_\nu C^\sigma_{\mu\sigma} - \mathfrak{g}_{\nu\alpha}\mathfrak{g}^{\lambda\sigma}C^\alpha_{\sigma\mu} - C^\lambda_{\mu\nu}\right).$$

Next,

$$\mathfrak{g}_{\sigma\mu}\mathfrak{g}_{\rho\nu}\mathfrak{g}^{\sigma\rho}{}_{;\alpha} = \mathfrak{g}_{\mu\nu}C^\tau_{\alpha\tau} - \mathfrak{g}_{\rho\nu}C^\rho_{\mu\alpha} - \mathfrak{g}_{\sigma\mu}C^\sigma_{\nu\alpha}.$$

Hence

$$\begin{aligned}-\frac{1}{2}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\sigma\mu}\mathfrak{g}_{\rho\nu}\mathfrak{g}^{\sigma\rho}{}_{;\alpha} = &-\frac{1}{2}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\mu\nu}C^\tau_{\alpha\tau} + \frac{1}{2}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\rho\nu}C^\rho_{\mu\alpha}\\ &+ \frac{1}{2}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\sigma\mu}C^\sigma_{\nu\alpha}.\end{aligned}$$

Finally, since $\dim M = 4$,

$$\begin{aligned}\mathfrak{g}_{\sigma\rho}\mathfrak{g}^{\sigma\rho}{}_{;\alpha} &= \mathfrak{g}_{\sigma\rho}\left(\mathfrak{g}^{\sigma\rho}C^\tau_{\alpha\tau} - \mathfrak{g}^{\sigma\tau}C^\rho_{\tau\alpha} - \mathfrak{g}^{\rho\tau}C^\sigma_{\tau\alpha}\right)\\ &= 4C^\tau_{\alpha\tau} - C^\rho_{\rho\alpha} - C^\sigma_{\sigma\alpha}\\ &= 2C^\tau_{\alpha\tau}.\end{aligned}$$

Therefore

$$\frac{1}{4}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\mu\nu}\mathfrak{g}_{\sigma\rho}\mathfrak{g}^{\sigma\rho}{}_{;\alpha} = \frac{1}{2}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\mu\nu}C^\tau_{\alpha\tau}.$$

Adding the four expressions, the terms involving $\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\mu\nu}C^\tau_{\alpha\tau}$ cancel. The remaining mixed terms also cancel after relabelling dummy indices and using the symmetry of $C^\alpha_{\mu\nu}$ in its lower indices. We are left with

$$-C^\lambda_{\mu\nu} + \frac{1}{2}\delta^\lambda_\mu C^\sigma_{\nu\sigma} + \frac{1}{2}\delta^\lambda_\nu C^\sigma_{\mu\sigma}.$$

This is precisely

$$-C^\lambda_{\mu\nu} + \delta^\lambda_{(\mu}C^\sigma_{\nu)\sigma} = p^\lambda_{\mu\nu}.$$

Hence the claimed identity follows. □

Theorem A.8 *Let (M, g) be a $(3+1)$-dimensional spacetime equipped with a connection Γ, and let B be a symmetric background connection on M. Define the metric density by*

$$\mathfrak{g}^{\mu\nu} = \frac{1}{16\pi}\sqrt{-\det g}\, g^{\mu\nu}.$$

and define the tensor $p^\alpha_{\mu\nu}$ by

$$p^\alpha_{\mu\nu} := \left(B^\alpha_{\mu\nu} - \delta^\alpha_{(\mu}B^\kappa_{\nu)\kappa}\right) - \left(\Gamma^\alpha_{\mu\nu} - \delta^\alpha_{(\mu}\Gamma^\kappa_{\nu)\kappa}\right).$$

Let the gravitational Lagrangian be

$$L := \mathfrak{g}^{\mu\nu}\left[(\Gamma^\alpha_{\sigma\mu} - B^\alpha_{\mu\sigma})(\Gamma^\sigma_{\nu\alpha} - B^\sigma_{\alpha\nu}) - (\Gamma^\alpha_{\mu\nu} - B^\alpha_{\mu\nu})(\Gamma^\sigma_{\alpha\sigma} - B^\sigma_{\alpha\sigma}) + r_{\mu\nu}\right].$$

Then $p^\lambda_{\mu\nu}$ is the canonical momentum associated with the metric density $\mathfrak{g}^{\mu\nu}$, that is,

$$\frac{\partial L}{\partial \mathfrak{g}^{\mu\nu}{}_{,\lambda}} = \frac{\partial L}{\partial \mathfrak{g}^{\mu\nu}{}_{;\lambda}} = p^\lambda_{\mu\nu}, \tag{A.15}$$

Here, $\mathfrak{g}^{\mu\nu}{}_{,\lambda}$ denotes the partial derivative of the metric density, whereas $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$ denotes its covariant derivative with respect to the background connection B.

Proof. Calculating the covariant derivative of the background connection gives

$$\mathfrak{g}^{\mu\nu}{}_{;\lambda} = \mathfrak{g}^{\mu\nu}{}_{,\lambda} + B^{\mu}_{\lambda\sigma}\mathfrak{g}^{\sigma\nu} + B^{\nu}_{\lambda\sigma}\mathfrak{g}^{\mu\sigma} - B^{\sigma}_{\lambda\sigma}\mathfrak{g}^{\mu\nu}.$$

Since $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$ differs from $\mathfrak{g}^{\mu\nu}_{,\lambda}$ by terms depending only on $\mathfrak{g}^{\mu\nu}$ and B, differentiating with respect to $\mathfrak{g}^{\mu\nu}{}_{,\lambda}$ is equivalent to differentiating with respect to $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$. We now differentiate the expression for L obtained in lemma 5.11. The term $\mathfrak{g}^{\mu\nu}r_{\mu\nu}$ does not depend on $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$, so it does not contribute.

First consider the term

$$\frac{1}{2}\mathfrak{g}_{\rho\alpha}\mathfrak{g}^{\rho\beta}{}_{;\eta}\mathfrak{g}^{\eta\alpha}{}_{;\beta}.$$

When differentiating with respect to $\mathfrak{g}^{\mu\nu}{}_{;\lambda}$, this term gives two contributions, because the derivative variable occurs in both factors. Differentiating the first factor gives

$$\frac{1}{2}\mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\nu},$$

while differentiating the second factor gives

$$\frac{1}{2}\mathfrak{g}_{\nu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\mu}.$$

Hence

$$\frac{\partial}{\partial\mathfrak{g}^{\mu\nu}{}_{;\lambda}}\left(\frac{1}{2}\mathfrak{g}_{\rho\alpha}\mathfrak{g}^{\rho\beta}{}_{;\eta}\mathfrak{g}^{\eta\alpha}{}_{;\beta}\right) = \frac{1}{2}\mathfrak{g}_{\mu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\nu} + \frac{1}{2}\mathfrak{g}_{\nu\alpha}\mathfrak{g}^{\lambda\alpha}{}_{;\mu}.$$

Next consider the term

$$-\frac{1}{4}\mathfrak{g}^{\eta\alpha}\mathfrak{g}_{\sigma\rho}\mathfrak{g}_{\tau\beta}\mathfrak{g}^{\rho\beta}{}_{;\eta}\mathfrak{g}^{\sigma\tau}{}_{;\alpha}.$$

Again the derivative variable occurs in both factors. Differentiating the first factor gives

$$-\frac{1}{4}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\sigma\mu}\mathfrak{g}_{\tau\nu}\mathfrak{g}^{\sigma\tau}{}_{;\alpha},$$

and differentiating the second factor gives

$$-\frac{1}{4}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\sigma\mu}\mathfrak{g}_{\rho\nu}\mathfrak{g}^{\sigma\rho}{}_{;\alpha}.$$

After relabelling dummy indices, these two terms are equal. Therefore

$$\frac{\partial}{\partial\mathfrak{g}^{\mu\nu}{}_{;\lambda}}\left(-\frac{1}{4}\mathfrak{g}^{\eta\alpha}\mathfrak{g}_{\sigma\rho}\mathfrak{g}_{\tau\beta}\mathfrak{g}^{\rho\beta}{}_{;\eta}\mathfrak{g}^{\sigma\tau}{}_{;\alpha}\right) = -\frac{1}{2}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\sigma\mu}\mathfrak{g}_{\rho\nu}\mathfrak{g}^{\sigma\rho}{}_{;\alpha}.$$

Finally, consider

$$\frac{1}{8}\mathfrak{g}^{\eta\alpha}\mathfrak{g}_{\rho\beta}\mathfrak{g}^{\rho\beta}{}_{;\eta}\mathfrak{g}_{\sigma\tau}\mathfrak{g}^{\sigma\tau}{}_{;\alpha}.$$

Differentiating the first factor gives

$$\frac{1}{8}\mathfrak{g}^{\lambda\alpha}\mathfrak{g}_{\mu\nu}\mathfrak{g}_{\sigma\tau}\mathfrak{g}^{\sigma\tau}{}_{;\alpha},$$

and differentiating the second factor gives the same expression after relabelling dummy indices. Hence

$$\frac{\partial}{\partial \mathfrak{g}^{\mu\nu}{}_{;\lambda}} \left(\frac{1}{8} \mathfrak{g}^{\eta\alpha} \mathfrak{g}_{\rho\beta} \mathfrak{g}^{\rho\beta}{}_{;\eta} \mathfrak{g}_{\sigma\tau} \mathfrak{g}^{\sigma\tau}{}_{;\alpha} \right) = \frac{1}{4} \mathfrak{g}^{\lambda\alpha} \mathfrak{g}_{\mu\nu} \mathfrak{g}_{\sigma\rho} \mathfrak{g}^{\sigma\rho}{}_{;\alpha}.$$

Combining these three computations, we obtain

$$\begin{aligned} \frac{\partial L}{\partial \mathfrak{g}^{\mu\nu}{}_{;\lambda}} &= \frac{1}{2} \mathfrak{g}_{\mu\alpha} \mathfrak{g}^{\lambda\alpha}{}_{;\nu} + \frac{1}{2} \mathfrak{g}_{\nu\alpha} \mathfrak{g}^{\lambda\alpha}{}_{;\mu} \\ &\qquad - \frac{1}{2} \mathfrak{g}^{\lambda\alpha} \mathfrak{g}_{\sigma\mu} \mathfrak{g}_{\rho\nu} \mathfrak{g}^{\sigma\rho}{}_{;\alpha} + \frac{1}{4} \mathfrak{g}^{\lambda\alpha} \mathfrak{g}_{\mu\nu} \mathfrak{g}_{\sigma\rho} \mathfrak{g}^{\sigma\rho}{}_{;\alpha}. \end{aligned}$$

By Lemma A.7, the right-hand side is precisely $p^{\lambda}_{\mu\nu}$. Therefore

$$\frac{\partial L}{\partial \mathfrak{g}^{\mu\nu}{}_{,\lambda}} := \frac{\partial L}{\partial \mathfrak{g}^{\mu\nu}{}_{;\lambda}} = p^{\lambda}_{\mu\nu}.$$

This proves the theorem. $\square$

Bibliography

[1] Ian M. Anderson, *The variational bicomplex*, 1989. Preliminary draft.

[2] Lars Andersson, Mingliang Cai, and Gregory J Galloway, *Rigidity and positivity of mass for asymptotically hyperbolic manifolds*, Annales henri poincaré, 2008, pp. 1–33.

[3] Vladimir Igorevich Arnol'd, *Mathematical methods of classical mechanics*, Vol. 60, Springer Science & Business Media, 2013.

[4] Richard Arnowitt, Stanley Deser, and Charles W Misner, *Coordinate invariance and energy expressions in general relativity*, Physical Review **122** (1961), no. 3, 997.

[5] ______, *Republication of: The dynamics of general relativity*, General Relativity and Gravitation **40** (2008), no. 9, 1997–2027.

[6] Sheldon Axler, *Linear algebra done right*, 3rd ed., Springer, 2015.

[7] Robert Bartnik, *The mass of an asymptotically flat manifold*, Communications on pure and applied mathematics **39** (1986), no. 5, 661–693.

[8] Yuchen Bi, Tianze Hao, Shihang He, Yuguang Shi, and Jintian Zhu, *A proof for the riemannian positive mass theorem up to dimension 19*, arXiv preprint arXiv:2603.02769 (2026).

[9] Raoul Bott and Loring W Tu, *Differential forms in algebraic topology*, Vol. 82, Springer Science & Business Media, 2013.

[10] Simon Brendle and Yipeng Wang, *A dimension descent scheme for the positive mass theorem in arbitrary dimension*, arXiv preprint arXiv:2604.08473 (2026).

[11] ______, *On the spacetime positive energy theorem in arbitrary dimension*, arXiv preprint arXiv:2604.18561 (2026).

[12] Paula Burkhardt-Guim, *Pointwise lower scalar curvature bounds for C^0 metrics via regularizing Ricci flow*, arXiv preprint arXiv:1907.13116 (2019).

[13] Annegret Burtscher and Leonardo García-Heveling, *Global hyperbolicity through the eyes of the null distance*, Communications in Mathematical Physics **405** (2024), no. 4, 90.

[14] Sean M Carroll, *Spacetime and geometry*, Cambridge University Press, 2019.

[15] Otis Chodosh, Christos Mantoulidis, and Felix Schulze, *Generic regularity for minimizing hypersurfaces in dimensions 9 and 10* (2023), available at 2302.02253. Final version, to appear in Publications Mathématiques de l'IHÉS.

[16] ______, *Improved generic regularity of codimension-1 minimizing integral currents*, arXiv preprint arXiv:2306.13191 (2023).

[17] Yvonne Choquet-Bruhat and Robert Geroch, *Global aspects of the cauchy problem in general relativity*, Communications in Mathematical Physics **14** (1969), no. 4, 329–335.

[18] Demetrios Christodoulou, *Mathematical problems of general relativity I*, Vol. 1, European Mathematical Society, 2008.

[19] Piotr T Chruściel, *Elements of general relativity*, Springer Nature, 2020.

[20] Piotr Tadeusz Chruściel, *On the relation between the einstein and the komar expressions for the energy of the gravitational field*, Annales de l'ihp physique théorique, 1985, pp. 267–282.

[21] ———, *A remark on the positive-energy theorem*, Classical and Quantum Gravity **3** (1986), no. 6, L115–L121.

[22] ———, *On the invariant mass conjecture in general relativity*, Communications in mathematical physics **120** (1988), no. 2, 233–248.

[23] Piotr Tadeusz Chruściel and Erwann Delay, *The hyperbolic positive energy theorem*, arXiv preprint arXiv:1901.05263 (2019).

[24] Piotr Tadeusz Chruściel, Erwann Delay, and Raphaela Wutte, *Hyperbolic energy and maskit gluings*, arXiv preprint arXiv:2112.00095 (2021).

[25] Piotr Tadeusz Chruściel and Gregory J Galloway, *Positive mass theorems for asymptotically hyperbolic riemannian manifolds with boundary*, arXiv preprint arXiv:2107.05603 (2021).

[26] Piotr Tadeusz Chruściel and Marc Herzlich, *The mass of asymptotically hyperbolic Riemannian manifolds*, Pacific journal of mathematics **212** (2003), no. 2, 231–264.

[27] Piotr Tadeusz Chruściel, Jacek Jezierski, and Jerzy Kijowski, *Hamiltonian field theory in the radiating regime*, Springer, 2002.

[28] Piotr Tadeusz Chruściel and Gabriel Nagy, *The hamiltonian mass of asymptotically anti-de sitter space-times*, Classical and Quantum Gravity **18** (2001), no. 9, L61–L68.

[29] ———, *The mass of spacelike hypersurfaces in asymptotically anti-de sitter space-times*, arXiv preprint gr-qc (2001).

[30] Piotr Tadeusz Chruściel and Raphaela Wutte, *Holographic is hamiltonian, relatively*, arXiv preprint arXiv:2604.10877 (2026).

[31] ———, *Positivity of holographic energy*, Physical Review Letters **136** (2026), no. 18, 181401.

[32] Mattias Dahl, Klaus Kroencke, and Stephen McCormick, *A volume-renormalized mass for asymptotically hyperbolic manifolds*, arXiv preprint arXiv:2307.06196 (2023).

[33] ———, *The volume-renormalized mass from a hamiltonian perspective*, arXiv preprint arXiv:2506.11841 (2025).

[34] Albert Einstein, *Die Grundlagen der allgemeinen*, Relat. Teorie Ann. Phys **49** (1916), 769–822.

[35] Michael Forger and Sandro Vieira Romero, *Covariant poisson brackets in geometric field theory*, Communications in Mathematical Physics **256** (2005), no. 2, 375–410.

[36] Ph. Freud, *Über die ausdrücke der gesamtenergie und des gesamtimpulses eines materiellen systems in der allgemeinen relativitätstheorie*, Annals of Mathematics (1939), 417–419.

[37] Eric Gourgoulhon, *3+1 formalism and bases of numerical relativity*, 2007. Lecture notes, version 1.

[38] Mikhael Gromov and H Blaine Lawson Jr, *Spin and scalar curvature in the presence of a fundamental group. i*, Annals of Mathematics (1980), 209–230.

[39] Eduardo Hafemann, *A low-regularity riemannian positive mass theorem for non-spin manifolds with distributional curvature*, arXiv preprint arXiv:2602.03451 (2026).

[40] Shihang He, Yuguang Shi, and Haobin Yu, *Singularity removal rigidity theorems for minimal hypersurfaces in manifolds with nonnegative scalar curvature*, arXiv preprint arXiv:2602.23705 (2026).

[41] Marc Henneaux and Claudio Teitelboim, *Asymptotically anti-de sitter spaces*, Communications in Mathematical Physics **98** (1985), no. 3, 391–424.

[42] Marc Herzlich, *Mass formulae for asymptotically hyperbolic manifolds*, Ads/cft correspondence: Einstein metrics and their conformal boundaries, 2005, pp. 103–121.

[43] ———, *Computing asymptotic invariants with the ricci tensor on asymptotically flat and asymptotically hyperbolic manifolds*, Annales henri poincaré, 2016, pp. 3605–3617.

[44] Sven Hirsch, Marcus Khuri, Martin Lesourd, and Yiyue Zhang, *The hyperboloidal and spacetime positive mass theorem in all dimensions*, arXiv preprint arXiv:2604.24746 (2026).

[45] José Luis Jaramillo and Eric Gourgoulhon, *Mass and angular momentum in general relativity*, Mass and motion in general relativity, 2011, pp. 87–124.

[46] Wenshuai Jiang, Weimin Sheng, and Huaiyu Zhang, *Weak scalar curvature lower bounds along ricci flow*, Science China Mathematics **66** (2023), no. 6, 1141–1160.

[47] Jürgen Jost, *Riemannian geometry and geometric analysis*, Springer, 2005.

[48] Marcus Khuri, Jian Wang, and Jinmin Wang, *Riemannian positive mass theorem in all dimensions in the presence of low-codimension singularities*, arXiv preprint arXiv:2606.23529 (2026).

[49] Jerzy Kijowski, *A simple derivation of canonical structure and quasi-local hamiltonians in general relativity*, General Relativity and Gravitation **29** (1997), no. 3, 307–343.

[50] Klaas Landsman, *Foundations of general relativity: From einstein to black holes*, Second corrected and expanded printing, Radboud University Press, Nijmegen, 2022.

[51] Dan A Lee, *Geometric relativity*, Vol. 201, American Mathematical Society, 2021.

[52] ———, *Geometric relativity*, Vol. 201, American Mathematical Society, 2021.

[53] Dan A Lee and Philippe G LeFloch, *The positive mass theorem for manifolds with distributional curvature*, Communications in Mathematical Physics **339** (2015), no. 1, 99–120.

[54] John M Lee, *Smooth manifolds*, Introduction to smooth manifolds, 2003, pp. 1–29.

[55] Philippe G. LeFloch, *Einstein spacetimes with weak regularity*, Advances in lorentzian geometry, 2011, pp. 81–96.

[56] André Lichnerowicz, *L'intégration des équations de la gravitation relativiste et le problème des n corps*, Journal de Mathématiques Pures et Appliquées **23** (1944), 37–63.

[57] Joachim Lohkamp, *The higher dimensional positive mass theorem I*, arXiv preprint (2006), available at math/0608795.

[58] ———, *Skin structures in scalar curvature geometry*, arXiv preprint arXiv:1512.08252 (2015).

[59] ———, *Skin structures on minimal hypersurfaces*, arXiv preprint arXiv:1512.08249 (2015).

[60] ———, *The higher dimensional positive mass theorem II*, arXiv preprint arXiv:1612.07505 (2016).

[61] Stig Lundgren and Benjamin Meco, *A generalization of the adm mass for asymptotically euclidean manifolds of weak regularity: S. lundgren et al.*, Geometriae Dedicata **220** (2026), no. 2, 18.

[62] Liam Mazurowski and Xuan Yao, *A positive mass theorem for continuous metrics*, arXiv preprint arXiv:2606.19123 (2026).

[63] Dusa McDuff and Dietmar Salamon, *Introduction to symplectic topology*, Oxford University Press, 2017.

[64] Benoıt Michel, *Geometric invariance of mass-like asymptotic invariants*, Journal of mathematical physics **52** (2011), no. 5.

[65] Thomas Parker and Clifford Henry Taubes, *On Witten's proof of the positive energy theorem*, Communications in Mathematical Physics **84** (1982), no. 2, 223–238.

[66] Kaare Brandt Petersen and Michael Syskind Pedersen, *The matrix cookbook*, Technical University of Denmark, 2012. Version 20121115.

[67] Jerzy Plebański and Andrzej Krasiński, *An introduction to general relativity and cosmology*, Cambridge University Press, 2006.

[68] Oscar Alejandro Reula, *Existence theorem for solutions of witten's equation and non-negativity of total mass.*, Ph.D. Thesis, 1983.

[69] Anna Sakovich, *The jang equation and the positive mass theorem in the asymptotically hyperbolic setting*, Communications in Mathematical Physics **386** (2021), no. 2, 903–973.

[70] David J Saunders, *The geometry of jet bundles*, Vol. 142, Cambridge University Press, 1989.

[71] DJ Saunders, *A note on legendre transformations*, Differential Geometry and its Applications **1** (1991), no. 2, 109–122.

[72] Richard Schoen and Shing-Tung Yau, *On the proof of the positive mass conjecture in general relativity*, Communications in Mathematical Physics **65** (1979), no. 1, 45–76.

[73] ———, *Proof of the positive mass theorem. ii*, Communications in Mathematical Physics **79** (1981), no. 2, 231–260.

[74] ———, *The structure of manifolds with positive scalar curvature*, Directions in partial differential equations, 1987, pp. 235–242.

[75] ———, *Positive scalar curvature and minimal hypersurface singularities*, arXiv preprint arXiv:1704.05490 (2017).

[76] Richard M Schoen and Shing-Tung Yau, *Complete manifolds with nonnegative scalar curvature and the positive action conjecture in general relativity*, Proceedings of the National Academy of Sciences **76** (1979), no. 3, 1024–1025.

[77] Jan Arnoldus Schouten, *Tensor analysis for physicists*, Courier Corporation, 1989.

[78] Nathan Smale, *Generic regularity of homologically area minimizing hypersurfaces in eight dimensional manifolds*, Communications in Analysis and Geometry **1** (1993), no. 2, 217–228.

[79] Kip S Thorne, Charles W Misner, and John Archibald Wheeler, *Gravitation*, Freeman San Francisco, 2000.

[80] Piotr Urbański, *Analiza dla studentów fizyki*, Uniwersytet Warszawski, Warsaw, 1998 (Polish). In Polish.

[81] Jorn van Voorthuizen, *Multipole moments in stationary spacetimes*, Journal of Mathematical Physics **66** (2025), no. 7.

[82] Robert M Wald, *General relativity*, University of Chicago press, 2024.

[83] Robert M Wald and Andreas Zoupas, *General definition of "conserved quantities" in general relativity and other theories of gravity*, Physical Review D **61** (2000), no. 8, 084027.

[84] Xiaodong Wang, *The mass of asymptotically hyperbolic manifolds*, Journal of Differential Geometry **57** (2001), no. 2, 273–299.

[85] Wikipedia contributors, *Phase space — wikipedia, the free encyclopedia*, 2026. Accessed: 2026-07-27.

[86] Edward Witten, *A new proof of the positive energy theorem*, Communications in Mathematical Physics **80** (1981), no. 3, 381–402.