Contents

1 Overview

I'm not sure now, at the beginning, how this tale ends. With luck I'm going to learn something new. Maybe you will, too.

These notes develop an intrinsically coordinate-free approach to classical mechanics and field theory. We are motivated by the Einstein equations of general relativity, where there is no preferred notion of time and the subject is permeated with issues arising from diffeomorphism group gauge freedom. In order to understand these difficulties better, and perhaps come to some resolution of old problems, we return to the foundations and rebuild mechanics from the ground up with a perspective that does not have in-built preferences for a limited choice of time variables. Of course, coordinate-based results, including those involving time, are critical and we show at every step how to recover standard results once a time function and attendant other gauge choices are made. But the aim is to create a theory that focuses on coordinate-independent constructions and generates gauge-dependent formulations after the fact.

As a by-product of this work, we construct a formulation of Lagrangian mechanics, and its Hamiltonian counterpart, that smoothly transitions from single particle dynamics to classical fields (e.g, scalar fields, Yang-Mills connections), dust and other isentropic fluids (both filling spacetime and configurations with boundaries), as well as Vlasov matter distributions. Gauge theories are emphasized along the way, even in settings that would not normally fall under the umbrella of gauge theory. The hope is that by approaching easy things like particle mechanics the hard way, we gain insight into how to attack the truly harder problems stemming from general relativity.

These notes are not for everyone. There's a good chance that they might not even be for you. In particular, they are not intended to be a first exposure to abstract mechanics. There are many excellent texts already for this, including [GPS08] and [Ar97]. For readers already well-versed in modern differential geometry at the level of [Le13] we can recommend [**?**] and especially the delightful [Sp10]. For readers already familiar with the mechanics, I hope that the route taken through what is already well-traveled territory will feel unusual.

2 Actions on Curves

The standard notion of a Lagrangian on a manifold M^d is simply a real-valued function on TM. Let $I = [t_0, t_1]$ be a closed interval with coordinate t. Then the action of a curve $\gamma : I \to M$ is

$$S_L[\gamma] = \int_I L(\dot{\gamma})\, dt.$$

Lagrangian mechanics considers curves such that the action is stationary under perturbation, associated Hamiltonian theory, and so forth; we assume that this is already a familiar story.

The role of the time variable t is somewhat special in this formalism: there is a specific 1-form dt that appears in the integrand, whereas no particular coordinate on M is essential. If we were to reparameterize I with a different coordinate, $t = t(\tilde{t})$ say, then generically $\tilde{t}$ would appear explicitly in the integrand. In particular, the integrand would no longer be a Lagrangian in the sense above. One can handle this situation by introducing time-dependent Lagrangians, maps on $TM \times I$ with associated actions $\int_I L(\dot{\gamma}, t)\, dt$. The class of time-dependent Lagrangians is closed under time reparameterization, but now the difference between space and time has become even more pronounced.

Our first goal is to give a "coordinate-free" description of Lagrangians on curves where all coordinates are treated on an equal footing. With luck, this approach clarifies the theory in a manner similar to, e.g., the coordinate-free description of the tangent bundle of a manifold. Moreover, it allows us to build up an apparatus used to smoothly transition from Lagrangians associated with particles to Lagrangians associated with fields and fluids. All this comes, however, at the expense of some abstraction. To keep things simple, in this section we treat the case of Lagrangians over curves, essentially particle mechanics. Section 4 addresses the case of Lagrangians on higher-dimensional surfaces needed for field theory, and the machinery there is a modest generalization of the material from this section.

Before moving on, we summarize some basic definitions.

awkward

Definition 2.1. Let M^d be a manifold and let $I \subset \mathbb{R}$ be an interval. A smooth map from $TM \times I$ to $\mathbb{R}$ is a **classical (time-dependent) Lagrangian**. We use this same term for smooth maps $U \times I \to \mathbb{R}$ where $U \subset TM$ is open. If L is a classical Lagrangian and $L(v, t)$ is independent of $t \in I$, we say that L is **time independent**. If J is a compact subinterval of of I and $\gamma : J \to M$ is a curve, the **action** of γ with respect to L is

$$S_L[\gamma] = \int_J L(\dot{\gamma}, t)\, dt. \tag{2.1}$$

2.1 Densities on Lines

Let V be a one dimensional vector space, which we will call a line. A density on V is a map $\mu : V \to \mathbb{R}$ satisfying

$$\mu(cv) = |c|\, \mu(v) \tag{2.2}$$

for any $v \in V$. The space $\mathcal{D}V$ of densities on V is closed under taking linear combinations, and it is straightforward to see that the space is one-dimensional. Indeed, there is a close relationship between densites on V and elements of the dual space V^*. If $\eta \in V^*$ then η determines a density $|\eta|$ via

$$|\eta|(v) = |\eta(v)|.$$

Using equation (2.2), it's easy to see that a density is determined by its action on a single $v \in V$ and as a consequence, if $\eta \in V^* \setminus \{0\}$ then any density on V is a multiple of $|\eta|$.

Let Γ be a one-dimensional manifold. The density bundle over Γ is

$$\mathcal{D}(T\Gamma) = \coprod_{q \in \Gamma} \mathcal{D}(T_q M)$$

which can be given the structure of a line bundle over M in the usual way. If μ is a section of $\mathcal{D}(T\Gamma)$ we can integrate μ over compact sets of Γ. As a concrete example, suppose $I = [t_0, t_1] \subset \mathbb{R}$ and $\gamma : I \to \Gamma$ is a smooth curve with $\dot{\gamma} \neq 0$ everywhere. Then

$$\int_{\gamma(I)} \mu = \int_{t_0}^{t_1} \mu(\dot{\gamma})\, dt.$$

For full details on integration of densities, including the construction in higher dimensions, see [Le13] Chapter 16.

2.2 Densities over Projective Tangent Bundles

Let V be a d-dimensional vector space. The projective space over V is the set of lines through the origin in V, and we denote it by $\mathrm{G}_1(V)$. More generally, the set of k-dimensional subspaces of V is called the Grassmannian of order k over V and is written $\mathrm{G}_k(V)$, but for now we will only be interested in the case $k = 1$.

Let M^d be a manifold. The projective tanget bundle of M is

$$\mathrm{G}_1(TM) = \coprod_{q \in M} \mathrm{G}_1(T_q M)$$

which can be given the structure of a fiber bundle over M with fiber $\mathbb{R}P^d$.

Each line $\ell \in \mathrm{G}_1(TM)$ has an associated space $\mathcal{D}(\ell)$ of densities over it. The set of all such densities form the **line densities over** M,

$$\mathcal{D}\mathrm{G}_1(TM) = \coprod_{\ell \in \mathrm{G}_1(TM)} \mathcal{D}(\ell)$$

and a standard construction gives $\mathcal{D}\mathrm{G}_1(TM)$ the structure of a line bundle over $\mathrm{G}_1(TM)$. The notation here is more formidable than the concept it describes. Each element of $\mathcal{D}\mathrm{G}_1(TM)$ can be thought of as a line in some $T_q M$ along with a particular density on that line.

2.3 Fundamental Lagrangians

The machinery reviewed up to this point allows for a simple coordinate-free description of Lagrangians defined on curves.

Definition 2.2. Let M be a manifold. A **fundamental Lagrangian** on M is a section $\mathcal{L}$ of $\mathcal{D}\mathrm{G}_1(TM)$ over an open subset of $\mathrm{G}_1(TM)$.

In Section 4 we generalize this definition to Lagrangians over higher-dimensional objects involving $\mathrm{G}_k(TM)$ with $k > 1$. But for the remainder of this section, a fundamental Lagrangian refers to the object from Definition 2.2. The word Lagrangian by itself ought to refer to a fundamental Lagrangian, in the sense that classical Lagrangians are coordinate-dependent representations of a fundamental Lagrangian (Section 2.6). Nevertheless, we bow to historical precedent and refrain from using the word Lagrangian without a modifier. Indeed, the following section introduces a third related object, a *pragmatic Lagrangian*, which can be used to represent a fundamental Lagrangian. Unlike classical Lagrangians, these are coordinate independent objects. Moreover, they are easier to work with than the rather abstract notion of a fundamental Lagrangian, but they suffer from the drawback that a fundamental Lagrangian is represented by many pragmatic Lagrangians. The point here is that other varieties of Lagrangian each describe a fundamental Lagrangian but do so with some kind of deficiency. Nevertheless, all three objects are important. To distinguish them notationally, we use roman L to for classical Lagrangians, script $\mathcal{L}$ for fundamental Lagrangians, and double struck $\mathbb{L}$ for the yet-to-be defined pragmatic Lagrangians.

Mechanically, a fundamental Lagrangian is a map $\mathcal{L}$ that consumes a line in some T_qM and yields a density on the very same line. We allow the section to not be defined globally on all of $\mathrm{G}_1(TM)$ because we might need to restrict the allowable lines to, say, timelike lines in a Lorenzian manifold. The main job of a Lagrangian is to assign a number, the action, to allowable curves. To describe this process, let $\mathcal{L}$ be a Lagrangian on M and let Γ be a 1-dimensional submanifold of M diffeomorphic to a compact interval in $\mathbb{R}$; we call Γ an interval in M. An interval Γ in M such that each $T_q\Gamma$ lies in the domain of $\mathcal{L}$ is **admissible** with respect to $\mathcal{L}$.

Suppose Γ is an admissible interval in M with respect to a fundamental Lagrangian $\mathcal{L}$. Let $\iota : \Gamma \to M$ be the natural embedding and fix $q \in \Gamma$. Observe that $\iota_*(T_q\Gamma)$ is a line in T_qM. Hence $\mathcal{L}[\iota_*(T_p\Gamma)]$ is a density on $\iota_*(T_p\Gamma)$; we use square brackets whenever an extended body such as a subspace or a manifold is the argument of a function. If $v \in T_q\Gamma$ then $\iota_* v$ lies in $\iota_*(T_q\Gamma)$ and is therefore a suitable argument for $\mathcal{L}[\iota_*(T_q\Gamma)]$. We define

$$\iota^*\mathcal{L}(v) = (\iota_* v) \lrcorner \mathcal{L}[\iota_*(T_q\Gamma)]$$

and it is easy to see that $\iota^*\mathcal{L}$ is indeed a density on Γ. The **action** of Γ with respect to $\mathcal{L}$ is simply

$$S_{\mathcal{L}}[\Gamma] = \int_\Gamma \iota^*\mathcal{L}. \tag{2.3}$$

In practice, to compute $S[\Gamma]$ we parameterize Γ with a map $\gamma : [\lambda_0, \lambda_1] \to \mathcal{M}$ with $\dot\gamma \neq 0$. Then, upon unwinding definitions, we find

$$S_{\mathcal{L}}[\Gamma] = \int_{\lambda_0}^{\lambda_1} \dot\gamma \lrcorner \mathcal{L}[\operatorname{span}\dot\gamma]\, d\lambda. \tag{2.4}$$

Because the action from (2.3) depends only on Γ, the quantity on the right-hand side of equation (2.4) is inherently independent of the choice of parameterization.

We are interested in intervals in M for which $S_{\mathcal{L}}$ is stationary. It will be some time (Section 4.10) before we have the "right" tools for finding stationary intervals. But at this point we can at least define what we are looking for.

Definition 2.3. Let $\mathcal{L}$ be a fundamental Lagrangian on M^d. An admissible interval Γ in M is **stationary** for $S_{\mathcal{L}}$ if whenever Ψ_s is a smooth family of maps $\Gamma \to M$ such that $\Psi_0 = \iota$ (where ι is the natural embedding) and such that $\Psi_s|_{\partial\Gamma} = \iota|_{\partial\Gamma}$ for all s,

$$\left.\frac{d}{ds}\right|_{s=0} S_{\mathcal{L}}[\Psi_s(\Gamma)] = 0.$$

Note that a compactness argument shows that $\Psi_s(\Gamma)$ in Definition (2.3) is an admissible interval for s sufficiently small and hence $S_{\mathcal{L}}[\Psi_s(\Gamma)]$ is well-defined.

The integrand appearing in equation (2.4) appears sufficiently frequently that we introduce notation for it. The **action density** associated with a fundamental Lagrangian is the real-valued map

$$\mathbb{s}_{\mathcal{L}}(v) = v \lrcorner \mathcal{L}[\operatorname{span} v] \tag{2.5}$$

defined for $v \in TM$ such that $\operatorname{span} v$ is in the domain of $\mathcal{L}$. In effect, $\mathbb{s}_{\mathcal{L}}$ is the infinitesimal contribution to the action of a curve as it is traversed with tangent v. The following result is an immediate consequence of the definitions and is related to parameterization invariance of the action.

Lemma 2.4. *Let $\mathbb{s}_{\mathcal{L}}$ be the action density of a fundamental Lagrangian L. For any v such that* $\operatorname{span} v$ *is in the domain of $\mathcal{L}$,*

$$\mathbb{s}_{\mathcal{L}}(\alpha v) = |\alpha|\, \mathbb{s}_{\mathcal{L}}(v) \tag{2.6}$$

for all nonzero $\alpha \in \mathbb{R}$.

In fact, one can show that a map satisfying equation (4.7) uniquely determines a fundamental Lagrangian, but this perspective generalizes less naturally to fundamental Lagrangians over higher-dimensional surfaces.

2.4 Pragmatic Lagrangians

As tidy as Definition 2.2 is, it is cumbersome to use in practice, and the following construction is easier to work with.

Definition 2.5. Let M be a manifold, and let U be an open subset of TM consisting of nonzero vectors such that whenever $v \in U$, $\alpha v \in U$ for all $\alpha \neq 0$. A **pragmatic Lagrangian** on U is a smooth map $\mathbb{L} : U \to T^*M$ satisfying the following:

1. If $v \in T_qM$ then $\mathbb{L}(v) \in T_q^*(M)$; i.e., $\mathbb{L}$ is a bundle map.
2. For any $v \in T_qM$ and $\alpha \in \mathbb{R}_{>0}$, $\mathbb{L}(\alpha v) = \mathbb{L}(v)$
3. $\mathbb{L}(-v) = -\mathbb{L}(v)$.

More succinctly, a pragmatic Lagrangian is a bundle map $\mathbb{L} : U \to T^*M$ satisfying

$$\mathbb{L}(\alpha V) = \frac{\alpha}{|\alpha|}\mathbb{L}(V) \tag{2.7}$$

for all $\alpha \neq 0$.

Every pragmatic Lagrangian $\mathbb{L}$ determines a fundamental Lagrangian $\mathcal{L}_{\mathbb{L}}$ as follows. Let $\ell \in \mathrm{G}_1(TM)$, so ℓ is a line in some T_qM. For any vector $w \in \ell$, $\mathbb{L}(w)$ is an element of $T_q^*(M)$ and we define

$$\mu_\ell(w) = w \lrcorner \mathbb{L}(w)$$

with the convention that $\mathbb{L}(0) = 0$. Note that if $\alpha \neq 0$ then

$$\mu_\ell(\alpha w) = (\alpha w) \lrcorner \mathbb{L}(\alpha w) = \alpha\frac{\alpha}{|\alpha|}(w \lrcorner \mathbb{L}(w)) = |\alpha|\mu_\ell(w).$$

Since we have declared $\mathbb{L}(0) = 0$ by fiat it follows that $\mu_\ell(0) = 0$ and we conclude μ_ℓ is indeed a density on ℓ. We define $\mathcal{L}_{\mathbb{L}}[\ell] = \mu_\ell$ and call $\mathcal{L}_{\mathbb{L}}$ the **fundamental Lagrangian induced** by the pragmatic Lagrangian $\mathbb{L}$.

What Definition 2.5 gains in concreteness it loses in redundancy. More than one pragmatic Lagrangian can induce the same fundamental Lagrangian, and indeed much of the information in a pragmatic Lagrangian is lost when passing to the fundamental Lagrangian it induces. To see this, note that by definition, if $\ell \in \mathrm{G}_1(TM)$ and $w \in \ell$ then

$$w \lrcorner \mathcal{L}_{\mathbb{L}}[\ell] = w \lrcorner \mathbb{L}(w). \tag{2.8}$$

Consequently the covector $\mathbb{L}(w)$ impacts $\mathcal{L}_{\mathbb{L}}$ only in its application to the line $\boldsymbol{\ell}$ and the addition of a covector that annihilates $\boldsymbol{\ell}$ would leave the induced fundamental Lagrangian unchanged. The following result shows, however, that every fundamental Lagrangian is induced by a pragmatic Lagrangian; in effect the space of fundamental Lagrangians is a quotient of the space of pragmatic Lagrangians.

Lemma 2.6. *Let $\mathcal{L}$ be a fundamental Lagrangian on M. There exists a pragmatic Lagrangian $\mathbb{L}$ on M that induces $\mathcal{L}$:*

$$\mathcal{L} = \mathcal{L}_{\mathbb{L}}.$$

Proof. Let g be a metric on M. For any $w \in TM$ and $\operatorname{span} w$ in the domain of $\mathcal{L}$ define

$$\mathbb{L}(w) = (w \lrcorner \mathcal{L}[\operatorname{span} w])\frac{1}{|w|_g^2} w^\flat$$

where $w^\flat$ is the covector dual (via g) to w. Note that $\operatorname{span} 0$ is never in the domain of $\mathcal{L}$ and hence there is no division by zero in this definition. Since $\mathbb{L}$ is evidently a bundle map of the right type, to show that it is a pragmatic Lagrangian we need only show that it satisfies the scaling property (2.7). But if $\alpha \neq 0$ then

$$\mathbb{L}(\alpha w) = ((\alpha w) \lrcorner \mathcal{L}[\operatorname{span} \alpha w])\frac{1}{|\alpha w|_g^2}(\alpha w)^\flat = |\alpha|(w \lrcorner \mathcal{L}[\operatorname{span} w])\ \frac{1}{|\alpha|^2 |w|_g^2}\alpha w^\flat = \frac{\alpha}{|\alpha|}\mathbb{L}(w)$$

as required.

To see that $\mathcal{L} = \mathcal{L}_{\mathbb{L}}$, consider some line $\boldsymbol{\ell}$ in the domain of $\mathcal{L}$ and pick $w \in \boldsymbol{\ell}$ with $w \neq 0$. By definition

$$\begin{aligned} w \lrcorner \mathcal{L}_{\mathbb{L}}[\boldsymbol{\ell}] = w \lrcorner \mathbb{L}(w) &= (w \lrcorner \mathcal{L}[\operatorname{span} w])\frac{1}{|w|_g^2} w^\flat(w) \\ &= w \lrcorner \mathcal{L}[\operatorname{span} w] \\ &= w \lrcorner \mathcal{L}[\boldsymbol{\ell}]. \end{aligned}$$

Two line densities are the same if they agree for a single nonzero vector and hence $\mathcal{L}[\boldsymbol{\ell}] = \mathcal{L}_{\mathbb{L}}[\boldsymbol{\ell}]$. □

Note that the proof of Lemma 2.6 relies on a choice of metric, which reflects the fact that fundamental Lagrangians are not uniquely represented by pragmatic Lagrangians. This phenomenon generally does not cause difficulty, but see Section **??** where nonuniqueness needs to be taken into account. ref

The action of in interval with respect to a fundamental Lagrangian can be easily computed in terms of a pragmatic Lagrangian that induces it. Indeed, let $\mathbb{L}$ be a pragmatic Lagrangian on M with induced fundamental Lagrangian $\mathcal{L}_{\mathbb{L}}$ and let Γ an $\mathcal{L}$-admissible interval in M. Unwinding definitions, if $\gamma : [\lambda_0, \lambda_1] \to \Gamma$ is a parameterization of Γ with $\dot\gamma \neq 0$ then

$$S_{\mathcal{L}_{\mathbb{L}}}[\Gamma] = \int_{\lambda_0}^{\lambda_1} \dot\gamma \lrcorner \mathbb{L}(\dot\gamma)\ dt. \tag{2.9}$$

2.5 Examples

2.5.1 Arclength

Let (M, g) be a Riemannian manifold. For any $v \in TM$ define

$$\mathbb{L}_{\mathrm{arc}}(v) = \frac{1}{|v|_g} v^\flat = \frac{1}{|v|_g} g_{ab} v^b.$$

One readily verifies that $\mathbb{L}_{\mathrm{arg}}$ is indeed a pragmatic Lagrangian and we denote its induced Lagrangian by $\mathcal{L}_{\mathrm{arc}}$. If Γ is an interval in M parameterized by a curve $\gamma : [t_0, t_1] \to M$ then equation (2.9) implies

$$S_{\mathcal{L}_{\mathrm{arc}}}[\Gamma] = \int_{t_0}^{t_1} \dot\gamma \lrcorner \mathbb{L}_{\mathrm{arc}}(\dot\gamma)\ dt = \int_{t_0}^{t_1} \frac{1}{|\dot\gamma|_g} |\dot\gamma|_g^2\ dt = \int_{t_0}^{t_1} |\dot\gamma|_g\ dt,$$

which is precisely the arclength of Γ.

Alternatively, consider a line ℓ in some T_qM and let ι_ℓ be the natural embedding of ℓ into T_qM. Then $g_\ell := \iota_\ell^* g$ is a Riemannian metric on ℓ with Riemannian density dV_{g_ℓ} and it is easy to see that $\mathcal{L}_{\mathrm{arc}}[\ell] = dV_{g_\ell}$ check Jack's notation

2.5.2 Proper time

This example is a minor modification of the previous example, but shows how the domain of a Lagrangian may need to be a restricted set of lines.

Let (M, g) be a Lorentzian manifold with signature $(-, +, \cdots, +)$. For timelike vectors $v \in TM$ with $v \neq 0$ define

$$\mathbb{L}_{\text{time}}(v) = -\frac{1}{(-g(v,v))^{-1/2}} v^\flat = -\frac{1}{(-g(v,v))^{-1/2}} g_{ab} v^b.$$

A calculation analogous to the one above for arclength shows that $\mathbb{L}_{\text{time}}$ is a pragmatic Lagrangian inducing a fundamental Lagrangian $\mathcal{L}_{\text{time}}$. If Γ is a timelike interval parameterized by a curve $\gamma : [t_0, t_1] \to M$, then

$$S_{\mathcal{L}_{\text{time}}}[\Gamma] = \int_{t_0}^{t_1} \left(-g(\dot\gamma, \dot\gamma)\right)^{1/2} \, dt$$

is the proper time elapsed along Γ. Moreover, if ℓ is a timelike line in some tangent space T_qM then $-g$ induces a Riemannian metric on ℓ and $\mathcal{L}_{\text{time}}$ is its Riemannian density on ℓ.

2.5.3 Classical Mechanics

Consider $M^{d+1} = \Sigma^d \times I$ where $I = [t_0, t_1]$ is an interval with coordinate t. Let π_Σ be projection onto the first factor and let ∂_t be the unique vector field with $(\pi_\Sigma)_* \partial_t = 0$ such that $dt(\partial_t) = 1$.

Let $L : T\Sigma \times I \to \mathbb{R}$ be a classical Lagrangian. We define a pragmatic Lagrangian as follows. Suppose $v \in T_qM$ such that $dt(v) \neq 0$ and decompose $v = \beta(w + \partial t)$ where $dt(w) = 0$. Setting $q = (x, t)$ we can identify w as an element of $T_x\Sigma$ and we define

$$\mathbb{L}(v) := \frac{\beta}{|\beta|} L(w, t)\, dt. \tag{2.10}$$

Note that β and w are implicit functions of v: $\beta = dt(v) \neq 0$ and $w = \beta^{-1} v - \partial_t$. If $\alpha \neq 0$ then

$$\mathbb{L}(\alpha v) = \frac{\alpha\beta}{|\alpha\beta|} L(w,t)\, dt = \frac{\alpha}{|\alpha|} \frac{\beta}{|\beta|} L(w,t)\, dt = \frac{\alpha}{|\alpha|} \mathbb{L}(v).$$

Hence $\mathbb{L}$ defines a pragmatic Lagrangian.

Let Γ be the diffeomorphic image of a curve $\gamma : [t_0, t_1] \to M$ parameterized by t. So $\gamma(t) = (x(t), t)$ for some curve $x(t)$ in Σ. From equations (2.9) and (2.10) along with the identity $dt(\dot\gamma) = 1$ we compute

$$S_{\mathcal{L}_{\mathbb{L}}}[\Gamma] = \int_{t_0}^{t_1} (\dot x + \partial_t) \lrcorner \mathbb{L}(\dot x + \partial_t) \, dt = \int_{t_0}^{t_1} L(\dot x, t)\, dt,$$

precisely the classical action of the curve $x(t)$ with respect to the classical time-dependent Lagrangian L.

2.5.4 Actions from Covectors

Let M be a manifold and suppose τ is a non-vanishing 1-form on M that, as described below, induces a "time orientation". If M is equipped with a time-orientable Lorenzian metric we can let $\tau = g(T, \cdot)$ for any non-vanishing timelike vector field. Alternatively, for the classical Lagrangians of Section 2.5.3 we can take $\tau = dt$. In these settings, every covector on M determines an action that is, in effect, integration of the covector.

To see this, consider a 1-form η on M and define

$$\mathbb{L}(v) = \frac{\tau(v)}{|\tau(v)|}\eta.$$

for vectors v with $\tau(v) \neq 0$. It is easy to see that $\mathbb{L}$ is a pragmatic Lagrangian. Suppose Γ is an interval in M with tangent spaces transverse to the kernel of τ. Without loss of generality we can parameterize Γ with a curve $\gamma : [t_0, t_1] \to M$ satisfying $\tau(\dot{\gamma}) > 0$ and we find

$$S_{\mathcal{L}_{\mathbb{L}}}[\Gamma] = \int_{t_0}^{t_1} \eta(\dot{\gamma})\, dt. \tag{2.11}$$

Alternatively, we induce an orientation on the tangent spaces of Γ by declaring v is positively oriented if $\tau(v) > 0$. Giving Γ this orientation, the computation of equation (2.11) shows

$$S_{\mathcal{L}_{\mathbb{L}}}[\Gamma] = \int_{\Gamma} \eta.$$

In particular, Stokes's Theorem can be used to show that closed 1-forms lead to actions that are constant under perturbations of Γ leaving its endpoints fixed. Hence we arrive at a rendition of the standard observation that the stationary curves of a classical action are unchanged if an exact differential is added to the integrand.

2.6 Time Evolution

The coordinate-free description of a curve Lagrangian is conceptually simple but fails to

This section isn't ready for prime time.

capture the spirit of classical mechanics where the state of a system evolves in time. In this section we show how to recover this perspective from a fundamental Lagrangian, and that essentially any coordinate can play the role of the time coordinate. This freedom to allow any variable "be time" shows up in orbital mechanics where the evolution variable is an angle.

We begin by showing that if a manifold has already been decomposed into space and time, then the action associated with a fundamental Lagrangian can be computed in terms of a uniquely determined classical Lagrangian, and vice-versa, so long as we restrict attention to curves that are transverse to the time function.

Proposition 2.7. *Consider a product manifold $M^{d+1} = \Sigma^d \times I$ where $I \subset \mathbb{R}$ is an interval and let t be the coordinate on I.*

If $\mathcal{L}$ is a fundamental Lagrangian on M there exists a unique classical Lagrangian $L : T\Sigma \times I \to \mathbb{R}$ such that for any interval Γ in M transverse to the level sets of t,

$$S_{\mathcal{L}}[\Gamma] = \int_{t_0}^{t_1} L(\dot{x}, t)\, dt \tag{2.12}$$

where $x : [t_0, t_1] \to \Sigma$ is the unique path in Σ such that $\gamma(t) = (x(t), t)$ is a parameterization of Γ.

Conversely, suppose $L : T\Sigma \times I \to \mathbb{R}$ is a classical Lagrangian on M. There exists a unique fundamental Lagrangian on M defined on tangent lines transverse to the level sets of t such that equation (2.12) holds for any interval Γ in M transverse to the level sets of t, where again $x : [t_0, t_1] \to \Sigma$ is the unique curve determined by a parameterization of Γ by t.

Proof. Let $\mathcal{L}$ be a fundamental Lagrangian on M and let $\mathbb{L}$ be a pragmatic Lagrangian that induces it. For $(w, t) \in T\Sigma \times I$ define

$$L(w, t) = (\partial_t + w) \lrcorner \mathbb{L}(\partial_t + w).$$

If Γ is an interval in M transverse to the level sets of t it admits a unique parameterization of the form $\gamma(t) = (x(t), t)$ and equation (2.9) implies

$$S_{\mathcal{L}}[\Gamma] = \int_{t_0}^{t_1} (\partial_t + \dot{x}) \lrcorner \mathbb{L}(\partial_t + \dot{x})\, dt = \int_{t_0}^{t_1} L(\dot{x}, t)\, dt.$$

Conversely, suppose $L : T\Sigma \times I \to \mathbb{R}$ is a classical Lagrangian. For $v \in TM$ with $dt(v) \neq 0$ define

$$\mathbb{L}(v) = \frac{\beta_v}{|\beta_v|} L(\beta_v^{-1} w_v, t)\, dt$$

where $\beta_v = dt(v)$ and where $w_v = v - dt(v)\partial_t \in T\Sigma$. As shown in Section 2.5.3, $\mathbb{L}$ is a pragmatic Lagrangian. Suppose Γ is an interval in M transverse to the level sets of t and let $\gamma : [t_0, t_1] \to M$ be its unique parameterization of the form $\gamma(t) = (x(t), t)$. Equation (2.9) and the observation $dt(\dot\gamma) = 1$ imply

$$S_{\mathcal{L}_{\mathbb{L}}}[\Gamma] = \int_{t_0}^{t_1} (\partial_t + \dot x) \lrcorner \mathbb{L}(\partial_t + \dot x)\, dt = \int_{t_0}^{t_1} L(\dot x, t)\, dt.$$

Uniqueness in both cases follows from a localization argument. Suppose L_1 and L_2 are classical Lagrangians with $L_1(w^*, t^*) \neq L_2(w^*, t^*)$ at some point. Then we can find a curve $(x(t), t)$ for t near t^* such that $\dot x(t^*) = w^*$ and such that the actions of the two curves differ from each other. But then L_1 and L_2 cannot both satisfy (2.12) for the same fundamental Lagrangian $\mathcal{L}$.

The argument showing that two fundamental Lagrangians $\mathcal{L}_1$ and $\mathcal{L}_2$ cannot both satisfy (2.12) for the same classical Lagrangian L is essentially similar: if they disagree on some line ℓ transverse to the level sets of t we can find a small interval Γ for which ℓ is a tangent line such that the actions differ. But then (2.12) can't hold for both $\mathcal{L}_1$ and $\mathcal{L}_2$. □

As would be expected, uniqueness fails in Proposition 2.7 if we reconstruct pragmatic Lagrangians rather than fundamental Lagrangians. Nevertheless, it's insightful to see this failure explicitly. Starting with a pragmatic Lagrangian $\mathbb{L}$, we can pass to the fundamental Lagrangian determined by it and then apply the constructions of Proposition 2.7 to generate an associated classical Lagrangian and subsequently recover from it an associated pragmatic Lagrangian $\hat{\mathbb{L}}$. Working through these details one finds that if v is a vector with $dt(v) \neq 0$ then

$$\hat{\mathbb{L}}(v) = (v \lrcorner \mathbb{L}(v))\, \frac{1}{dt(v)}\, dt.$$

As pragmatic Lagrangians $\mathbb{L}$ and $\hat{\mathbb{L}}$ are generically different from each other: there is no reason that $\mathbb{L}$ would only take on values that are multiples of dt. But the abstract Lagrangians they induce are identical: if $\boldsymbol{\ell} \in \mathrm{G}_1(TM)$ is transverse to the level sets of t and if $v \in \boldsymbol{\ell}$ then

$$v \lrcorner \mathcal{L}_{\hat{\mathbb{L}}}[\boldsymbol{\ell}] = v \lrcorner \hat{\mathbb{L}}(v) = (v \lrcorner \mathbb{L}(v)) \frac{1}{dt(v)}\, dt(v) = v \lrcorner \mathbb{L}(v) = v \lrcorner \mathcal{L}_{\mathbb{L}}[\boldsymbol{\ell}].$$

A **local spacetime decomposition** of M^{d+1} is a diffeomorphism from an open set W in M to a product $\Sigma^d \times I$. Given such a gauge we use the following notation:

- t is projection onto the second coordinate,
- ∂_t is the unique vector field on W with $dt(\partial_t) = 1$ that is annihilated by projection onto Σ.

Using the implicit function theorem, one can show that there is vast flexibility in finding a local spacetime gauges in a neighborhood of any point.

Lemma 2.8. *Let $q \in M$ and suppose $T \in T_qM$ and $\eta \in T_q^*M$ satisfy $\eta(T) = 1$. Then there exists a local spacetime decomposition defined on a neighborhood of q such that at q, $\partial_t = T$ and $dt = \eta$.*

Lemma 2.8 can be used in tandem with Proposition 2.7 to compute the action of an interval Γ by splitting Γ into finitely many pieces, each contained in a local spacetime decompositions such that Γ is transverse to the time function on it. The action of Γ can then be computed by summing actions of the pieces with respect to classical Lagrangians as in Proposition 2.7. Alternatively, given an interval Γ, one can find a thin tubular neighborhood of Γ that is the domain of a local spacetime decomposition on it where Γ is transverse to the time coordinate. A single application of Proposition 2.7 then allows the action of Γ to be computed in terms of a single classical Lagrangian.

2.7 Momentum

Consider a classical time-dependent Lagrangian

$$L : M^d = T\Sigma^{d-1} \times I \to \mathbb{R}.$$

Choose local coordinates q^a for Σ with induced coordinates v^a on each $T_q\Sigma$, so (q^a, v^a) are local coordinates for $T\Sigma$. A familiar computation shows that a curve $(x(t), t)$ is stationary for the action determined by L if

$$\frac{d}{dt}\left[\frac{\partial L}{\partial v^a}\right] = \frac{\partial L}{\partial q^a}; \tag{2.13}$$

these are the Euler-Lagrange equations and the quantity $\partial L/\partial v^a$ appearing in it is known as the momentum conjugate to q^a. The key insight of the Hamiltonian approach to mechanics is to treat momentum as a dynamical variable on an equal footing with position, and that doing so leads to structural simplifications of the equations of motion.

The momenta appearing in equation (2.13) are gauge dependent quantities. First, coordinates have been imposed on Σ so that the derivatives $\partial L/\partial v^a$ are well-defined. This is not an essential drawback and there is a standard geometric description of these momenta. More seriously, the classical Lagrangian itself is rooted in a particular spacetime decomposition of M. Our first task therefore is to describe a notion of momentum associated with a fundamental Lagrangian that is independent of any gauge but which, when a particular gauge is imposed, reduces to the classical notion.

Let $\mathcal{L}$ be a fundamental Lagrangian and let $U \subset TM$ be the set of tangent vectors v such that span v is in the domain of $\mathcal{L}$. Recall the action density of $\mathcal{L}$ from equation (2.5)

$$\mathbb{s}_{\mathcal{L}}(v) = v \lrcorner \mathcal{L}[\operatorname{span} v].$$

If $\mathcal{L}$ is described in terms of a pragmatic Lagrangian $\mathbb{L}$ then $\mathbb{s}_{\mathcal{L}}$ can be computed from $\mathbb{L}$ directly via equation (2.9), namely $\mathbb{s}_{\mathcal{L}}(v) = v \lrcorner \mathbb{L}(v)$.

Now fix some $q \in M$ and consider the restriction $\mathbb{s}_{\mathcal{L},q}$ of $\mathbb{s}_{\mathcal{L}}$ to the tangent space T_qM. The derivative of $\mathbb{s}_{\mathcal{L},q}$ at some $v \in T_qM$ is a linear map $T_vT_qM \to \mathbb{R}$. But we can identify T_vT_qM with T_qM itself and hence it is a linear map from T_qM to $\mathbb{R}$, i.e. an element of T^*M. The **momentum** determined by v is the derivative of $\mathbb{s}_{\mathcal{L},q}$ at v:

$$\mathbb{P}_{\mathcal{L}}(v) = \mathrm{D}\mathbb{s}_{\mathcal{L},q}\,|_v \in T_q^*M.$$

Computationally,

$$w \lrcorner \mathbb{P}_{\mathcal{L}}(v) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(v + \epsilon w).$$

In some contexts, analogs of the map $\mathbb{P}_{\mathcal{L}}$ that convert velocity to momentum are called Legendre transformations (e.g. [Si01] Chapter 20). But usually a Legendre transformation denotes a subtly different operation related to the inverse of this map. We (unimaginatively) call it the **velocity-to-momentum map**.

Before examining properties of $\mathbb{P}_{\mathcal{L}}$, we start with two examples to illustrate the general case.

2.7.1 Momentum of a Relativistic Free Particle

The action for a free particle in a Lorentzian manifold (M^d, g) is, up to a factor of $-m$, the proper time Lagrangian of Section 2.5.2. We describe it in terms of a pragmatic Lagrangian

$$\mathbb{L}(v) = \frac{m}{\sqrt{-\mathrm{g}(v,v)}}\mathrm{g}(v,\cdot) \tag{2.14}$$

with induced fundamental Lagrangian $\mathcal{L} = \mathcal{L}_{\mathbb{L}}$.

In order to compute $\mathbb{P}_{\mathcal{L}}$ we first compute the associated action density

$$\mathbb{s}_{\mathcal{L}}(v) = -m\sqrt{-\mathrm{g}(v,v)}.$$

If $w \in T_qM$ then

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(v+\epsilon w) = \frac{m}{\sqrt{-\mathrm{g}(v,v)}}\mathrm{g}(v,w)$$

and consequently

$$\mathbb{P}_{\mathcal{L}}(v) = \frac{m}{\sqrt{-\mathrm{g}(v,v)}}\mathrm{g}(v,\cdot). \tag{2.15}$$

The fact that the right-hand sides of equations (2.14) and (2.15) are the same is striking, but keep in mind that $\mathbb{P}_{\mathcal{L}}$ is uniquely determined from $\mathcal{L}$ whereas $\mathbb{L}$ is merely one of a vast family of pragmatic Lagrangians that induce $\mathcal{L}$. Nevertheless, the connection is not an accident and we return to this point shortly.

In order to connect $\mathbb{P}_{\mathcal{L}}$ with familiar relativistic notions of momentum, let ∂_t be a unit timelike vector in some T_qM; we think of ∂_t as an observer at q. The timelike spacetime velocity v can be decomposed as $v = \beta(\partial_t + w)$ for some $\beta \neq 0$ and some vector w orthogonal to ∂_t satisfying $|w|_g < 1$. The spacelike vector w is the velocity of the particle witnessed by the observer ∂_t. On the other hand it turns out that the scale factor β is effectively unimportant and impacts $\mathbb{P}_{\mathcal{L}}(v)$ only via its sign; we assume for simplicity that $\beta > 0$ noting the other case follows from the easy consequence of equation (2.15) that $\mathbb{P}_{\mathcal{L}}(-v) = -v$.

Applying the covector $\mathbb{P}(v)$ to $-\partial_t$ we find

$$-\partial_t \lrcorner \mathbb{P}_{\mathcal{L}}(v) = \frac{m}{\sqrt{1-|w|_{\mathrm{g}}^2}},$$

which is the relativistic energy of a particle with velocity w measured by the observer. On the other hand, if u is a vector orthogonal to ∂_t then

$$u \lrcorner \mathbb{P}_{\mathcal{L}}(v) = \frac{m}{\sqrt{1-|w|_g^2}} g(w,u).$$

Hence the restriction of $\mathbb{P}_{\mathcal{L}}(v)$ to the subspace orthogonal to ∂_t recovers the usual relativistic momentum seen by ∂_t.

Returning to similarities between equations (2.14) and (2.15) we make the following related observations:

1. If $\alpha \in \mathbb{R}$ and $\alpha \neq 0$ then
$$\mathbb{P}_{\mathcal{L}}(\alpha v) = \frac{\alpha}{|\alpha|}\mathbb{P}_{\mathcal{L}}(v)$$
for any timelike vector v. Up to a sign, the momentum determined by a vector depends only on the line spanned by v.

2. In fact, the previous observation shows that $\mathbb{P}_{\mathcal{L}}$ is a pragmatic Lagrangian. More than that, the Lagrangian it determines is exactly $\mathcal{L}$ since
$$v \lrcorner \mathbb{P}_{\mathcal{L}}(v) = -m\sqrt{-g(v,v)} = v \lrcorner \mathbb{L}(v).$$
Of the many pragmatic Lagrangians that induce $\mathcal{L}$, we find that $\mathbb{P}_{\mathcal{L}}$ is a natural representative in the sense that it can be computed from $\mathcal{L}$ without making any choices.

3. The collection $\mathbb{P}_{\mathcal{L}}(TM)$ of all possible momenta consists of the covectors p in T^*M with $g(p,p) = -m^2$. In particular, it is a hypersurface of T^*M. The momentum of a particle cannot be freely prescribed, but must instead satisfy a constraint.

We show in Section **??** that the first two of these observations are universal. They stem from the fact that the action of an interval Γ in M depends on its geometry but is unrelated to any particular parameterization. Although the last observation does not always hold, it is closely related to the first, which implies that $\mathbb{P}$ cannot be full rank.

2.7.2 Momenta Associated with Classical Lagrangians

Consider a classical Lagrangian on $M^d = \Sigma^{d-1} \times I$,

$$L : T\Sigma \times I \to \mathbb{R}.$$

As usual, let t be the coordinate on I, let ∂_t be the vector satisfying $\partial_t t = 1$ that is annihilated by projection onto Σ. Working locally near some $q \in M$ let $q_1, \dots, q_{d-1}$ be coordinates on Σ with associated vector fields ∂_k and let $w_1, \dots, w_{d-1}$ be the induced coordinates on the fibers of TM.

Suppose $v = \beta(\partial_t + w) \in T_q M$ where $\beta \neq 0$ and where $dt(w) = 0$. Note that w is the velocity in Σ determined by v when using t and ∂_t to measure time. Following the construction of Section 2.5.3 we define the pragmatic Lagrangian

$$\mathbb{L}(v) = \frac{\beta}{|\beta|} L(w,t) dt$$

with its associated action density

$$\mathbb{s}_{\mathcal{L}}(v) = v \lrcorner \mathbb{L}(v) = |\beta| L(w,t).$$

We now compute $\mathbb{P}_{\mathcal{L}}(v)$ assuming $\beta > 0$; the case $\beta < 0$ follows from an analogous computation that shows $\mathbb{P}_{\mathcal{L}}(-v) = -\mathbb{P}_{\mathcal{L}}(v)$. For the spatial directions we find

$$\partial_k \lrcorner \mathbb{P}_{\mathcal{L}}(v) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(v + \epsilon \partial_k) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \beta L(w + (\epsilon/\beta)\partial_k, t) = \frac{\partial}{\partial w_k} L(w,t). \quad (2.16)$$

Hence $\partial_k \lrcorner \mathbb{P}_{\mathcal{L}}(v)$ is the classical momentum conjugate to q^k, and we denote it below by p_k.

It remains to compute $\partial_t \lrcorner \mathbb{P}_{\mathcal{L}}(v)$ and we start by noting that if $v = \beta(\partial_t + w)$ then

$$v + \epsilon \partial_t = (\beta + \epsilon)\left(\partial_t + \frac{1}{1 + \epsilon/\beta} w\right).$$

Hence

$$\begin{aligned}
-\partial_t \lrcorner \mathbb{P}_{\mathcal{L}}(v) &= -\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(v + \epsilon \partial_t) \\
&= -\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (\kappa + \epsilon)\, L((1 + \epsilon/\kappa)^{-1} w, t) \\
&= \left(\frac{\partial}{\partial w^k} L(w,t)\right) w^k - L(w,t)
\end{aligned}$$

Recalling the shorthand $p_k = \frac{\partial}{\partial w^k} L(w,t)$ for the 'spatial' momenta computed above we find

$$-\partial_t \lrcorner \mathbb{P}_{\mathcal{L}}(v) = p_k w^k - L(w,t). \tag{2.17}$$

The quantity appearing on the right-hand side of equation (2.17) is exactly the classical Hamiltonian associated with the classical Lagrangian and we denote it by H. It is related[1] to the Legendre transformation of L, an important operation that is nevertheless difficult to motivate, as is the definition of the Hamiltonian which appears *deus ex machina* in many treatments. By contrast, the Hamiltonian emerges here as a consequence of an easy principle: it is $-\partial_t \lrcorner \mathbb{P}_{\mathcal{L}}$ for the fundamental Lagrangian associated with a classical Lagrangian.

At the end of section 2.7.1 we listed three properties of the relativistic free-particle momentum associated with parameterization invariance. We revisit these properties in the current context.

1. If $v(dt) > 0$ (and hence $\varkappa > 0$ in the notation above), the expressions for $p_k = \partial_k \lrcorner \mathbb{P}_{\mathcal{L}}(v)$ and $H = -\partial_t \lrcorner \mathbb{P}_{\mathcal{L}})v)$ computed in equations (2.16) and (2.17) depend on v only via the spatial velocity w. Since w remains fixed if v is replaced by αv for $\alpha > 0$, we find that $\mathbb{P}(\alpha v) = \mathbb{P}(v)$ if $\alpha > 0$. Combining this with our earlier assertion that $\mathbb{P}(-v) = -\mathbb{P}(v)$ we have
$$\mathbb{P}(\alpha v) = \frac{\alpha}{|\alpha|} P(v)$$
for all $\alpha \neq 0$. Hence $\mathbb{P}$ is a pragmatic Lagrangian.

2. To compute the fundamental Lagrangian it determines it is enough to consider the case $\beta = v(dt) > 0$, in which case we find
$$\begin{aligned} v \lrcorner \mathbb{P}(v) &= \beta\left(\partial_t \lrcorner \mathbb{P}(v) + w \lrcorner \mathbb{P}(v)\right) \\ &= \beta\left(-H + w^k p_k\right) \\ &= \beta\left(L(w,t) - p_k w^k + w^k p_k\right) \\ &= \beta\, L(w,t) \\ &= v \lrcorner \mathbb{L}(v). \end{aligned}$$
That is, $\mathbb{P}_{\mathcal{L}}$ induces the fundamental Lagrangian it was derived from.

[1] Strictly speaking, the Hamiltonian is the Legendre transformation of L once the variables w^k have been rewritten in terms of the variables p_k, assuming this inverse procedure is possible. Nevertheless, the quantity on the right-hand side of equation (2.17) is well defined regardless of whether such an inversion is possible.

3. Without knowing more about L we cannot recover the fact observed for free relativistic particles that $\mathbb{P}_{\mathcal{L}}(TM)$ is a hypersurface of T^*M. Nevertheless, because the of the scaling property $\mathbb{P}(cv) = \mathbb{P}(v)$ the derivative of $\mathbb{P}$ cannot be full rank, and $\mathbb{P}_{\mathcal{L}}(TM)$ is a Lebesgue null set.

2.8 Consequences of Parameterization Invariance

In Section 2.3 we showed that the action density $\mathbb{s}_{\mathcal{L}}$ of a fundamental Lagrangian $\mathcal{L}$ satisfies

$$\mathbb{s}_{\mathcal{L}}(\alpha v) = |\alpha|\, \mathbb{s}_{\mathcal{L}}(\alpha v)$$

and remarked that it is the source of the parameterization invariance of the action $S_{\mathcal{L}}$. This same scaling property is the root of the following result.

Proposition 2.9. *Let $\mathcal{L}$ be a fundamental Lagrangian on a manifold M^d. Its associated velocity-to-momentum map $\mathbb{P}_{\mathcal{L}}$ is a pragmatic Lagrangian that induces $\mathcal{L}$. That is, for all $v \in TM$ such that* span v *is in the domain of $\mathcal{L}$,*

$$v \lrcorner \mathbb{P}(v) = v \lrcorner \mathcal{L}[\operatorname{span} v]. \tag{2.18}$$

Proof. If $\alpha > 0$ and if v and δv are vectors in TM with span v in the domain of $\mathcal{L}$ then

$$\begin{aligned}
\mathbb{P}_{\mathcal{L}}(\alpha v)(\delta v) &= \frac{d}{d\epsilon}|_{\epsilon=0}\mathbb{s}(\alpha v + \epsilon\delta v) \\
&= |\alpha|\,\frac{d}{d\epsilon}|_{\epsilon=0}\mathbb{s}(v + (\epsilon/\alpha)\delta v) \\
&= |\alpha|\,(\delta v/\alpha) \lrcorner \mathbb{P}_{\mathcal{L}}(v) \\
&= \frac{\alpha}{|\alpha|}\,\delta v \lrcorner \mathbb{P}_{\mathcal{L}}(v).
\end{aligned}$$

Hence $\mathbb{P}_{\mathcal{L}}$ is a pragmatic Lagrangian. Moreover,

$$v \lrcorner \mathbb{P}_{\mathcal{L}}(v) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(v + \epsilon v) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} (1+\epsilon)\mathbb{s}_{\mathcal{L}}(v) = \mathbb{s}_{\mathcal{L}}(v) = v \lrcorner \mathcal{L}[\operatorname{span} v].$$

Hence $\mathbb{P}_{\mathcal{L}}$ induces $\mathcal{L}$. $\square$

Equation (2.18) can alternatively be written

$$v \lrcorner \mathbb{P}_{\mathcal{L}}(v) - v \lrcorner \mathcal{L}[\operatorname{span} v] = 0 \tag{2.19}$$

and the left-hand side can be interpreted as a version of the classical Hamiltonian appearing on the right-hand side of equation (2.17). Equation (2.19) is the source of statements to the effect of "the Hamiltonian of a parameterization invariant Lagrangian vanishes identically". We prefer equation (2.18), however, which will be the key to transitioning to Hamiltonian mechanics.

The **feasible momenta** associated with $\mathcal{L}$ is the image $\mathcal{P}_{\mathcal{L}} = \mathbb{P}_{\mathcal{L}}(U) \subset T^*M$, where U is the set of tangent vectors v with span v in the domain of $\mathcal{L}$. It is the set of all possible momenta associated with any admissible direction of travel in M, and Proposition 2.9 implies that it is a thin subset of T^*M. Indeed, since $\mathbb{P}_{\mathcal{L}}$ is a pragmatic Lagrangian, $\mathbb{P}_{\mathcal{L}}(\alpha v) = \mathbb{P}_{\mathcal{L}}(v)$ for any $\alpha > 0$ and hence the derivative of $\mathbb{P}_{\mathcal{L}}$ has less than full rank everywhere. Consequently, $\mathcal{P}_{\mathcal{L}}$ is a Lebesgue null set in T^*M. The maximum possible rank for $\mathbb{P}_{\mathcal{L}}$ is $2d - 1$, where d is the dimension of M; this is the generic case. If $\mathcal{P}_{\mathcal{L}}$ has constant rank $2d - 1$ then $\mathcal{P}_{\mathcal{L}}$ is an immersed hypersurface of T^*M, and for local questions we can assume $\mathcal{P}_{\mathcal{L}}$ is an embedded hypersurface. Although the case where $\mathcal{P}_{\mathcal{L}}$ has rank less than $2d-1$ is important (such as when additional "non-dynamical" variables akin to the parameterization variable are present), we will frequently assume for now that $\mathcal{P}_{\mathcal{L}}$ is an embedded submanifold.

2.9 Action from the Hamiltonian Perspective

Let $\mathcal{L}$ be a Lagrangian on M^d and let $\pi : T^*M \to M$ be the projection. Suppose Γ is an $\mathcal{L}$-admissible interval in M. If γ is a parameterization of Γ then $\dot{\gamma}$ is a curve in TM and equation (2.18) implies

$$S_{\mathcal{L}}[\Gamma] = \int_{\lambda_0}^{\lambda_1} \dot{\gamma} \lrcorner \mathcal{L}[\operatorname{span} \dot{\gamma}]\, d\lambda = \int_{\lambda_0}^{\lambda_1} \dot{\gamma} \lrcorner \mathbb{P}_{\mathcal{L}}(\dot{\gamma})\, d\lambda. \tag{2.20}$$

This last integrand can be reinterpreted in terms of a natural 1-form on T^*M known (among other names) as the tautological 1-form, which we we recall now.

Suppose $V \in T_p T^*M$ for some $p \in T^*M$. Projecting T^*M onto M it determines $v = \pi_* V \in TM$. Since p and v are both based at the same $q \in M$, we can then form $v \lrcorner p$.

Using these operations we define $\Theta_p \in T_p^* T^*M$ by

$$\Theta_p(V) = (\pi_* V) \lrcorner\, p.$$

This is the **tautological 1-form** at p. If q^k are local coordinates on M with induced coordinates p_k on the fibers of T^*M, one readily shows $\Theta = p_k dq^k$.

Returning to equation (2.20) we find

$$S_{\mathcal{L}}[\Gamma] = \int_{\mathbb{P}_{\mathcal{L}} \circ \dot{\gamma}} \Theta. \tag{2.21}$$

At first glance, this is a promising equation. We are interested in computing stationary intervals Γ, and we can computed action in terms of integrating a universal one-form over curves in T^*M. Hence one can bring standard tools such as the exterior derivative to bear on examining variations of the action. Motivated by equation (2.21) we define the **action** of an arbitrary curve $\tilde{\gamma}$ in T^*M to be

$$S[\tilde{\gamma}] = \int_{\tilde{\gamma}} \Theta,$$

a quantity that depends that depends on the curve but is unrelated to any Lagrangian. Now suppose $\tilde{\gamma}_s(\lambda)$ is a variation of curves in T^*M for λ in some interval $I = [\lambda_0, \lambda_1]$ and $s \in (-\epsilon, \epsilon)$ for some $\epsilon > 0$. Letting $X = \frac{d}{ds}|_{s=0}\gamma_s$ we have

check sign!

$$\left.\frac{d}{ds}\right|_{s=0} S[\tilde{\gamma}_s] = \Theta(X(\lambda_1)) - \Theta(X(\lambda_0)) + \int_{\tilde{\gamma}_0} X \lrcorner\, d\Theta$$

In particular, for variations that fix the endpoints,

$$\left.\frac{d}{ds}\right|_{s=0} S[\tilde{\gamma}_s] = -\int_{\tilde{\gamma}_0} X \lrcorner\, \Omega \tag{2.22}$$

where $\Omega = -d\Theta$ is known as the symplectic 2-form.

Without any constraints, there are no curves in T^*M for which action is stationary. Indeed, equation (2.22) implies that such a curve would have to satisfy $\Omega(\dot{\tilde{\gamma}}, \cdot) = 0$ at each point along the curve. Otherwise, if there exists a vector X such that $\Omega(\dot{\theta}, X) \neq 0$ at some point we could exploit it to construct a variation of curves about $\tilde{\gamma}$ for which the action is not stationary. However, using the coordinate representation

$$\Omega = dq_k \wedge dp^k$$

it is easy to see that for any non-zero vector T, there exists a vector X with $\Omega(T, X) \neq 0$.

When working with a specific Lagrangian, however, we do not vary over arbitrary curves in T^*M. Rather, a curve γ in M lifts to a curve $\mathbb{P}_{\mathcal{L}} \circ \dot{\gamma}$ in the set $\mathcal{P}_{\mathcal{L}}$ of feasible momenta, which we assume for now is an embedded hypersurface in T^*M. This suggests the following task:

Problem 2.10. *Find curves $\tilde{\gamma}$ in $\mathcal{P}_{\mathcal{L}}$ such that the action $S[\tilde{\gamma}]$ is stationary among curves with the same endpoints in $\mathcal{P}_{\mathcal{L}}$.*

With a little thought, however, this seems like an line of attack with almost no chance of success. First, the category of curves we are examining has been massively increased. If we find a curve $\tilde{\gamma}$ in $\mathcal{P}_{\mathcal{L}}$ for which the action is stationary, there is no reason to expect that $\tilde{\gamma}$ is the lift $\mathbb{P}_{\mathcal{L}} \circ \dot{\gamma}$ of some curve in M. At the same time, although the action of a curve in M can be computed using Θ via equation (2.20), variations among curves in $\mathcal{P}$ is broader than variations among curves that are lifts of curves in M. There is no reason to expect that stationarity within the smaller class implies stationarity with respect to full variations. In summary, we are potentially admitting solutions we do not want, and simultaneously potentially ruling out solutions we would like to keep.

In fact, stationary action in the Hamiltonian sense of Problem 2.10 is, in the generic case, equivalent to stationary action in the Lagrangian sense and relies on a structural feature of $\mathcal{P}_{\mathcal{L}}$. To illustrate it, consider a relativistic particle with mass m on a Lorenzian manifold (M, g) as in Section 2.7.1. Equation (2.15) implies that in this case $\mathbb{P}_{\mathcal{L}}(v) = m/\sqrt{-g(v,v)}\, v^\flat$ and therefore $\mathcal{P}_{\mathcal{L}}$ consists of the covectors p with $g(p,p) = -m^2$. A familiar computation shows that the tangent space $T_p\mathcal{P}_{\mathcal{L}}$ consists of the covectors orthogonal to p. Moreover, if δp is one of these orthogonal covectors, then

$$\delta p(v) = g(\delta p, v^\flat) = \frac{\sqrt{-g(v,v,)}}{m} g(\delta p, p) = 0.$$

That is, $T_{\mathbb{P}_{\mathcal{L}}(v)}\mathcal{P}_{\mathcal{L}}$ consists of $v^\perp$, the set of covectors that annihilate v. This potentially coincidental feature of $\mathcal{P}_{\mathcal{L}}$ is a consequence of the following result and is again related to parameterization invariance via Proposition 2.9.

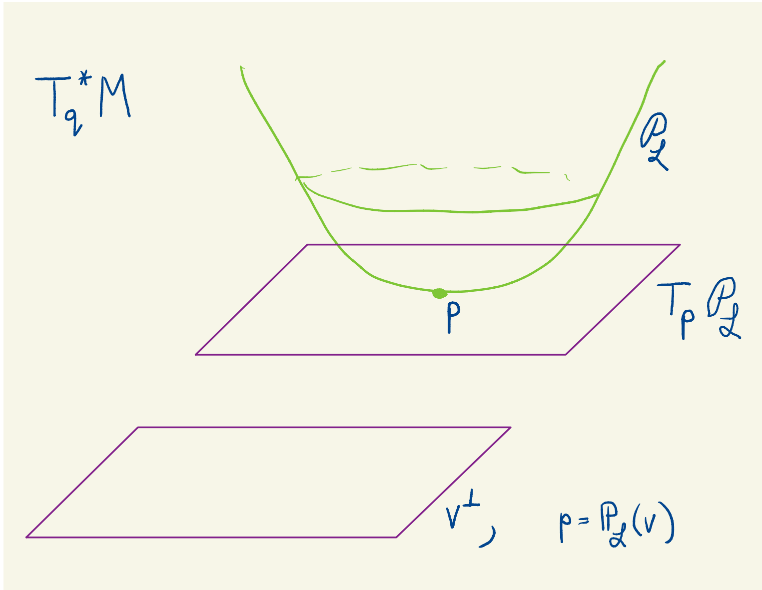

Figure 1: Feasible Momenta for a Free Particle

Proposition 2.11. *Let $\mathcal{L}$ be a fundamental Lagrangian on M^d with velocity-to-momentum map $\mathbb{P}_{\mathcal{L}}$. Fix some $q \in M$ and some $v \in T_qM$ with* span v *in the domain of $\mathcal{L}$. The derivative* $\mathrm{D}\,\mathbb{P}_{\mathcal{L}}$ *satisfies*

$$(\mathrm{D}\,\mathbb{P}_{\mathcal{L}})(T_vT_qM) \subset v^{\perp}.$$

Proof. Suppose $w \in T_vT_qM = T_qM$. From the definition of $\mathbb{P}_{\mathcal{L}}(v)$,

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(v+\epsilon w) = w \lrcorner \mathbb{P}_{\mathcal{L}}(v). \tag{2.23}$$

On the other hand, Proposition 2.9 implies

$$\mathbb{s}_{\mathcal{L}}(v+\epsilon w) = (v+\epsilon w) \lrcorner \mathbb{P}_{\mathcal{L}}(v+\epsilon w)$$

and hence

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(v+\epsilon w) = w \lrcorner \mathbb{P}_{\mathcal{L}}(v) + v \lrcorner (D\mathbb{P}_q)_v(w) \tag{2.24}$$

Comparing equations (2.23) and (2.24) we conclude $v \lrcorner (D\mathbb{P}_{\mathcal{L}})_v(w) = 0$ as claimed. ☐

The statement of Proposition 2.11 is a little fussy, mainly because we have assumed essentially nothing about $\mathbb{P}_{\mathcal{L}}$. If we assume that $\mathbb{P}_{\mathcal{L}}$ has the largest rank possible subject to the restriction from Proposition 2.9, we obtain a more straightforward result. In particular, we say $\mathcal{L}$ on M^d satisfies the **maximum rank condition** if $\mathbb{P}_{\mathcal{L}}$ has rank $2d-1$ at every point *and* if $\mathcal{P}_{\mathcal{L}}$ is an embedded submanifold of T^*M. Of course, if $\mathbb{P}_{\mathcal{L}}$ is constant rank, its image is an immersed submanifold and in local arguments we can assume $\mathcal{P}_{\mathcal{L}}$ is embedded by restricting our attention[2] to a small enough neighborhood in TM. The following is an easy consequence Proposition 2.11 and a dimension count argument when the maximum rank condition holds.

Corollary 2.12. *Suppose $\mathcal{L}$ satisfies the maximum rank condition. Let $p \in \mathcal{P}_{\mathcal{L}}$, so $p = \mathbb{P}_{\mathcal{L}}(v)$ for some $v \in TM$. Then*

$$T_p\mathcal{P}_{\mathcal{L}} = v^{\perp}$$

Corollary 2.12 shows that $\mathbb{P}_{\mathcal{L}}$ is effectively invertible in the maximum rank setting. Indeed, if $v, w \in T_qM$ for some q, and if $v^{\perp} = w^{\perp}$, then v and w are colinear. But we already knew that $\mathbb{P}_{\mathcal{L}}(\lambda v) = \pm\mathbb{P}_{\mathcal{L}}(v)$ depending on the sign of λ. The sign ambiguity is essential, since $\mathcal{P}_{\mathcal{L}} = \mathcal{P}_{-\mathcal{L}}$ and since momentum changes sign when $\mathcal{L}$ does. But the sign ambiguity is usually easy to manage and indeed can be eliminated in settings where there is a time orientation; see Section **??**.

The following two results validate the Hamiltonian agenda of examining stationary action by working in phase space, so long as $\mathcal{L}$ satisfies the maximum rank condition.

Proposition 2.13. *Suppose $\mathcal{L}$ satisfies the maximum rank condition. If Γ is a stationary interval in M and if γ is a parameterization of Γ then the lift $\tilde{\gamma} = \mathbb{P}_{\mathcal{L}} \circ \dot{\gamma}$ is stationary in $\mathcal{P}_{\mathcal{L}}$.*

Proof. Sketch. Consider an arbitrary variation of $\tilde{\gamma}_s$ of $\tilde{\gamma} = \mathbb{P}_{\mathcal{L}} \circ \dot{\gamma}$ in $\mathcal{P}_{\mathcal{L}}$. In local coordinates, Sloppy here about $\dot{\gamma}$ needing to not vanish downstairs.

$$\Theta(\dot{\tilde{\gamma}}_s) = p_k(\lambda, s)\dot{q}^k(\lambda, s).$$

For fixed λ, $p_k(\lambda, s)\, dq^k$ is a curve in $\mathcal{P}$ and hence

$$\left.\frac{d}{ds}\right|_{s=0} p_k(\lambda, s) dq^k \in T_{\mathbb{P}_{\mathcal{L}}\circ\dot{\gamma}(\lambda)}\mathcal{P}.$$

[2] Wouldn't it be nice to show that the embedded hypothesis is superfluous? This is less unreasonable than it sounds. Regardless, not critical for local work.

Since $\dot{q}^k \partial_k$ is the coordinate representation of $\dot{\gamma}$, Corollary 2.12 then implies

$$\left[\frac{d}{ds}\Big|_{s=0} p_k(\lambda, s) \right] \dot{q}^k(\lambda, s) = 0$$

and we conclude

$$\frac{d}{ds}\Big|_{s=0} \left[p_k(\lambda, s)\dot{q}^k(\lambda, s) \right] = p_k(\lambda, 0) \frac{d}{ds}\Big|_{s=0} \dot{q}^k(\lambda, s).$$

As a consequence,

$$\frac{d}{ds}\Big|_{s=0} S[\tilde{\gamma}_s] = \int_{\lambda_0}^{\lambda_1} X \lrcorner \tilde{\gamma}(\lambda) \, d\lambda \tag{2.25}$$

where $X(\lambda) = d/ds|_{s=0} \pi \circ \tilde{\gamma}_s$ is an "infintesimal" variation of γ.

Now consider the variation $\gamma_s := \pi \circ \tilde{\gamma}_s$ of γ. By stationarity of γ,

$$\frac{d}{ds}\Big|_{s=0} S[\mathbb{P}_{\mathcal{L}} \circ \dot{\gamma}_s] = 0.$$

But $\mathbb{P}_{\mathcal{L}} \circ \dot{\gamma}_s$ is another variation of $\tilde{\gamma}$, so equation (2.25) applies equally to it with the possibility that the vector field X may be different. But $\pi \circ \mathbb{P}_{\mathcal{L}} \circ \dot{\gamma}_s = \gamma_s = \pi \circ \tilde{\gamma}_s$, so X is unchanged as well. We conclude that the integral on the right-hand side of equation (2.25) vanishes and consequently $\tilde{\gamma}$ is stationary in $\mathcal{P}_{\mathcal{L}}$. □

Fix the sign ambiguity problem here!

Proposition 2.14. *Suppose $\mathcal{L}$ satisfies the maximum rank condition. If $\tilde{\gamma}$ is stationary in $\mathcal{P}_{\mathcal{L}}$ then it is the lift of a curve in M.*

Proof. Sketch. Suppose $\tilde{\gamma}$ is not the lift of a curve in M and let $\gamma = \pi \circ \tilde{\gamma}$. Then there is parameter λ such that $\tilde{\gamma}(\lambda) \neq \mathbb{P}_{\mathcal{L}}(\dot{\gamma}(\lambda))$. Let $p = \tilde{\gamma}(\lambda^*)$ and $v = \dot{\gamma}(\lambda^*)$. Now $p = \mathbb{P}_{\mathcal{L}}(w)$ for some $w \in TM$, and w is not colinear with v, for otherwise $\mathbb{P}_{\mathcal{L}}(w) = \pm\mathbb{P}_{\mathcal{L}}(v)$. Hence there exists $\delta p \in w^{\perp}$ such that $v \lrcorner \delta p \neq 0$. Because $\mathcal{P}_{\mathcal{L}}$ is a hypersurface, δp considered as an element of TT^*M belongs to $T_p\mathcal{P}_{\mathcal{L}}$. Moreover

$$-\Omega(\delta p, \dot{\tilde{\gamma}}) = \dot{\tilde{\gamma}} \lrcorner \delta p \neq 0.$$

Hence $\tilde{\gamma}$ is not stationary. □

2.10 Equations of Motion

At long last we develop the equations of motion for stationary action, assuming that the set $\mathcal{P}_{\mathcal{L}}$ of feasible momenta is an embedded hypersurface. In the previous section we showed that under this hypothesis, the stationary action problem reduces to finding curves γ in $\mathcal{P}_{\mathcal{L}}$ such that the generic action

$$S[\gamma] = \int_\gamma \Theta$$

is stationary among variations of curves constrained to $\mathcal{P}_{\mathcal{L}}$. We emphasize that $S[\cdot]$ depends on the tautological one-form Θ but is unrelated to any Lagrangian; the role that the Lagrangian plays is to dictate, via the map $\mathbb{P}_{\mathcal{L}}$, the set $\mathcal{P}_{\mathcal{L}}$ of feasible momenta. Having done its job, we forget about $\mathcal{L}$ at this point, and the remainder of this section concerns embedded hypersurfaces of T^*M.

Proposition 2.15. *Let $\mathcal{P}$ be an embedded hypersurface in T^*M. A curve $\gamma : I \to \mathcal{P}$ defined on a compact interval $I = [\lambda_0, \lambda_1]$ is stationary in $\mathcal{P}$ if and only if*

$$d\Theta(\dot{\gamma}(\lambda), V) = 0 \tag{2.26}$$

for all $\lambda \in I$ and all $V \in T_{\gamma(\lambda)}\mathcal{P}$.

Proof. Let γ_s be a variation of γ. Then

$$\frac{d}{ds}\bigg|_{s=0} S[\gamma_s] = \Theta(X(\lambda_1)) - \Theta(X(\lambda_0)) + \int_I d\Theta(\dot{\gamma}, X)\, d\lambda$$

where $X = d/ds|_{s=0}\gamma_s$ is the usual infinitesimal variation along γ. Because variations are fixed at the endpoints, the boundary terms vanish. So if equation (2.26) holds, γ is stationary. Conversely, suppose at some point p along the curve there exists a vector Z with $d\Theta(\dot{\gamma}, Z) \neq 0$; without loss of generality we assume the value is positive. We can then construct a variation of γ such that $X = Z$ at p, and such that $X = 0$ away from an arbitrarily small region near p so that $d\Theta(\dot{\gamma}, X) \geq 0$ at all points. For such a variation, $\int_\gamma d\Theta(\cdot, X) > 0$ and hence γ is not stationary. □

Proposition 2.15 indicates the key role that the symplectic 2-form $\Omega = -d\Theta$ plays in detecting stationary curves. As discussed in the following section, it is a more fundamental object on T^*M than Θ itself. Indeed, if we modify Θ by a closed one-form, although the action of individual curves change, the symplectic 2-form does not change, nor does the condition (2.26) needed for stationarity.

Following [Ar97] Sec ??, we now show that embedded hypersurfaces of T^*M are uniquely foliated by curves with tangent lines ℓ such that for $V \in \ell$, $\Omega|_{\mathcal{P}}(V, \cdot) = 0$.

Lemma 2.16. *Let B be a nondegenerate alternating bilinear form on an even-dimensional vector space V, and let W be a hyperspace of V. Then the annihilator of $B|_W$ is one-dimensional.*

Proof. Let g be an arbitrary metric on V and let $A : V \to V$ be the linear map defined by $g(v, Aw) = B(v, w)$. Since B is non-degenerate, A has trivial kernel and is hence invertible. Observe that $x \in \operatorname{Ann} B|_W$ if and only if $x \in W$ and $Ax \in W^\perp$, where $W^\perp$ is the orthogonal complement of W with respect to g. That is, $\operatorname{Ann} B|_W = A^{-1}(W^\perp) \cap W$. Since W is a hypersurface, $W^\perp$ is one dimensional, as is $A^{-1}(W^\perp)$. So $\operatorname{Ann} B|_W$ is at most one dimensional. Since W is has odd dimension, $\operatorname{Ann} B|_W$ is non-trivial, and hence is exactly one dimensional. □

flesh this last bit out.

Around here, talk about the fact that the distribution is integrable, foliations, easy case of Frobenious, etc.

Following [Ar97], we call an embedded submanifold of $\mathcal{P}$ with tangent space equal to the annihilator of $\Omega|_{\mathcal{P}}$ a **vortex curve**. A defining function for $\mathcal{P}$ yields a parameterization for vortex curves by selecting a preferred tangent vector at each point.

Lemma 2.17. *Let $\mathcal{P}$ be an embedded hypersurface of T^*M, let $p \in \mathcal{P}$, and let $\mathcal{H}$ be a local defining function for $\mathcal{P}$ near p. Then $X_{\mathcal{H}}$ defined by*

$$X_{\mathcal{H}} \lrcorner \Omega = d\mathcal{H}$$

spans $\operatorname{Ann} \Omega|_{T_p\mathcal{P}}$ *at p.*

Proof. First, observe that $X_{\mathcal{H}}$ is defined on an open neighborhood of p in T^*M, not just in $\mathcal{P}$. Moreover,

$$X_{\mathcal{H}}\mathcal{H} = d\mathcal{H}(X_{\mathcal{H}}) = \Omega(X_{\mathcal{H}}, X_{\mathcal{H}}) = 0$$

and hence $X_{\mathcal{H}}$ is tangent to the level sets of $\mathcal{H}$. In particular, on $\mathcal{P}$, $X_{\mathcal{H}}$ is tangent to $\mathcal{P}$.

Now suppose $W \in T_p\mathcal{P}$. Then $\Omega(X_{\mathcal{H}}, W) = -d\mathcal{H}(W) = 0$ since $\mathcal{H}$ is constant on $\mathcal{P}$. Thus $X_{\mathcal{H}}$ at p is in the annihilator of Ω restricted to $T_p\mathcal{P}$. Since $d\mathcal{H} \neq 0$ at p, $X_{\mathcal{H}} \neq 0$ there as well and hence Lemma 2.16 implies it spans $\operatorname{Ann} \Omega|_{T_p\mathcal{P}}$. □

Big picture. Through every point $p \in \mathcal{P}$ there is a unique maximal vortex curve. A curve in $\mathcal{P}$ is stationary if and only if it parameterizes a piece of a vortex curve. Moreover, we can generate such curves by selecting a defining function $\mathcal{H}$ for $\mathcal{P}$, in which case the vortex curves are exactly the integral curves of $X_{\mathcal{H}}$. The choice of defining function is essentially a choice of how the vortex curves are parameterized.

Theorem 2.18 (Hamilton's Equations)**.** *Let q^k be local coordinates on M and let p_k be the induced coordinates on the fibers of T^*M. Suppose $\mathcal{H}$ is a defining function for an embedded hypersurface $\mathcal{P}$. Then $X_{\mathcal{H}} = \dot{p}_k \partial_{p_k} + \dot{q}^k \partial_{q^k}$ is given by*

$$\begin{aligned} \dot{q}^k &= \frac{\partial \mathcal{H}}{\partial p_k} \\ \dot{p}^k &= -\frac{\partial \mathcal{H}}{\partial q^k} \end{aligned} \tag{2.27}$$

Proof. Short proof here. □

We call a defining function $\mathcal{H}$ for $\mathcal{P}$ a **Hamiltonian** for it, but some caution is needed to distinguish it from the Hamiltonian $H = -\partial_t \lrcorner \mathbb{P}_{\mathcal{L}}$ discussed in Section 2.7.2. The function $\mathcal{H}$ is defined on a neighborhood of $\mathcal{P}$ in T^*M and assumes a constant value on $\mathcal{P}$. By contrast, as described in detail in Section **??**, after a choice of time function t and time flow ∂_t have been made on M, H is a non-trivial function on $\mathcal{P}$. Both $\mathcal{H}$ and H induce motion on a symplectic manifold via the mechanism to be described in Section **??**, but in the case of H the manifold has two fewer dimensions than T^*M. Notably, both notions of Hamiltonian are gauge dependent quantities, with $\mathcal{H}$ reflecting a choice of parameterization along vortex lines, and H arising only after a time gauge has been imposed. The gauge-independent object from the Hamiltonian perspective is $\mathcal{P}_{\mathcal{L}}$, the hyperspace of feasible momenta.

Summary: a hyperspace $\mathcal{P} \subset T^*M$ is equipped with a preferred foliation by vortex curves. Tangent vectors T to vortex curves are characterized by $T \lrcorner \Omega|_{\mathcal{P}} = 0$. The curves γ into $\mathcal{P}$ that parameterize the vortex curves are exactly the curves for which action is stationary subject to the variations constrained to $\mathcal{P}$. A defining function $\mathcal{H}$ for $\mathcal{P}$ determines a vector field $X_{\mathcal{H}}$ along $\mathcal{P}$ tangent to the vortex curves and integrating $X_{\mathcal{H}}$ yields (parameterized) stationary curves. Different choices of defining functions determine different parameterizations of the vortex curves.

2.11 Hamilton's Equations, Concretely

The Hamiltonian approach described up to this point for finding stationary action curves is tortuous. First, starting with a fundamental Lagrangian $\mathcal{L}$, compute the velocity-to-momentum transformation $\mathbb{P}_{\mathcal{L}}$. With luck the full rank condition holds, in which case the image $\mathcal{P}_{\mathcal{L}}$ of this map is a hypersurface of T^*. Now find, somehow, a defining function $\mathcal{H}$ for the hypersurface. If all this can be done, then Theorem 2.18 can be applied to find stationary curves in $\mathcal{P}_{\mathcal{L}}$, which project down to stationary curves for the original action.

In this section we carry out this program with two different methods. In the first example, a defining function for $\mathcal{P}_{\mathcal{L}}$ is immediately apparent. This kind of serendipity can't always be expected, and the second example shows how to use an essentially arbitrary choice of time function on M to generate a local defining function $\mathcal{H}$.

2.11.1 Relativistic Free Particle

We revisit a particle of mass m on a Lorenzian manifold (M^d, g) and recall that in Section 2.7.1 we computed the velocity-to-momentum map

$$\mathbb{P}_{\mathcal{L}}(v) = \frac{m}{\sqrt{-g(v, v,)}} g(v, \cdot)$$

for time-like vectors v. But now it is apparent that

$$\mathcal{P}_{\mathcal{L}} = \{p \in T^*M \,:\, g(p, p) = -m^2\}.$$

Hence $\mathcal{H}(p) = \frac{1}{2} g(p, p)$ is a defining function for $\mathcal{P}_{\mathcal{L}}$. In fact, it is a defining function for the level sets associated with every mass m. Hamilton's equations (2.27) become

$$\begin{aligned}
\dot{q}^a &= g^{ab} p_b \\
\dot{p}_a &= -\frac{1}{2} \frac{\partial g^{cd}}{\partial q^a} p_c p_d
\end{aligned}$$

which are the geodesic equations. For these curves,

$$\begin{aligned}\frac{d}{d\lambda}|\dot{q}|_g^2 &= \frac{d}{d\lambda}g_{ab}\dot{q}^a\dot{q}^b \\ &= \frac{d}{d\lambda}g^{ab}p^a p_b \\ &= \frac{\partial g^{ab}}{\partial q^c}\dot{q}^c p^a p_b + 2g^{ab}p_a\dot{p}_b \\ &= \frac{\partial g^{ab}}{\partial q^c}g^{cd}p_d p_a p_b - g^{ab}p_a\frac{\partial g^{cd}}{\partial q^b}p_c p_d \\ &= 0.\end{aligned}$$

That is, this choice of $\mathcal{H}$ dictates a constant "speed" parameterization along the curves, and more than that links the speed to the particle mass:

$$g(\dot{q}, \dot{q}) = g(p, p) = -m^2.$$

Somewhere I should show how to get back to a classical action on M with classical Lagrangian $L(\dot{q}) = \frac{1}{2}g(\dot{q}, \dot{q})$ from here. The point is that once you've committed yourself to $\mathcal{H}$, then you've married yourself to *how* the curves are parameterized and nothing is lost in re-extracting a classical Lagrangian compatible with the agreed upon choice of time parameterization.

2.11.2 Selection of a Time Gauge

In the previous example, we were lucky: we spotted a defining function for $\mathcal{P}_{\mathcal{L}}$, and a fortuitous one at that. We now exhibit a technique for constructing a local defining function on $\mathcal{P}_{\mathcal{L}}$ based on the choice of an essentially arbitrary spacetime gauge (t, ∂_t) on M. The local defining function $\mathcal{H}$ generated by this method parameterizes vortex curves by t.

Let $\mathcal{L}$ be a fundamental Lagrangian on $\mathcal{P}_{\mathcal{L}}$ with velocity-to-momentum map $\mathbb{P}_{\mathcal{L}}$. Consider some $v_0 \in TM$ based at some point q_0 and let $p_0 = \mathbb{P}_{\mathcal{L}}(v_0)$. We assume that $\mathcal{L}$ satisfies the maximum rank condition, so $\mathcal{P}_{\mathcal{L}}$ is an embedded hypersurface, and we work locally to construct vortex curves in $\mathcal{P}_{\mathcal{L}}$ near p_0.

Pick a function t on M near q such that $dt(v_0) > 0$ at q_0 and such that $dt \neq 0$ on its domain. At the same time, pick a vector field ∂_t on M near q such that $dt(\partial_t) = 1$. With these choices made, we may as well assume that $M^d = \Sigma^{d-1} \times I$ for some interval I, with t being the coordinate along I and ∂_t pointing along I. Since $p_0 = \mathbb{P}_{\mathcal{L}}(v_0) = \mathbb{P}_{\mathcal{L}}(\lambda v_0)$ for any $\lambda > 0$, we can assume without loss of generality that $dt(v_0) = 1$.

The spacetime decomposition allows us to write each $p \in T^*M$ as a sum $p_\Sigma + sdt$ where p_Σ is the pullback of an element of $T^*\Sigma$ and $s \in \mathbb{R}$. Using the pullback, we simply identify elements of $T^*\Sigma$ as elements of T^*M. Our aim is to write $\mathcal{P}_{\mathcal{L}} \cap T_q M$ at $q = (q_\Sigma, t)$ as a graph over $T^*_{q_\Sigma}\Sigma$. That is, we want to find a function $H(q_\Sigma, p_\Sigma, t)$ such that each $p \in T_{(q_\Sigma,t)}M \cap \mathcal{P}_{\mathcal{L}}$ has the form

$$p = p_\Sigma - H(q_\Sigma, p_\Sigma, t)dt.$$

Assuming such a function H can be found, since elements of $T^*\Sigma$ annihilate ∂_t, it follows that p belongs to $\mathcal{P}_{\mathcal{L}} \cap T^*_q M$ if and only if

$$\partial_t \lrcorner\, p = -H(q_\Sigma, p_\Sigma(p), t)$$

where $p_\Sigma(p)$ is the projection of p onto $T_{q_\Sigma}\Sigma$ determined by the spacetime decomposition. Define

$$s(p) = \partial_t \lrcorner\, p = \partial_t \lrcorner\, \Theta_p \tag{2.28}$$

and let

$$\mathcal{H} = s + H,$$

which is a function on T^*M defined near p_0. Then $\mathcal{P}_{\mathcal{L}}$ is the level set $\mathcal{H} = 0$. Moreover, in terms of the coordinates (p_Σ, s) on each fiber $T_q M$, H is independent of s and hence $d\mathcal{H} = ds + dH \neq 0$. So $\mathcal{H}$ is a local defining function for $\mathcal{P}_{\mathcal{L}}$. If q^k_Σ are coordinates on Σ inducing coordinates $p_{\Sigma,k}$ on the fibers of $T^*\Sigma$, then the coordinates (q^k_Σ, t) induce coordinates $(p_{\Sigma,k}, s)$ on T^*M, where s is the function defined in equation (2.28). In terms of these coordinates, Hamilton's equations (2.27) become

$$\begin{aligned}
\dot{q}^k_\Sigma &= \frac{\partial \mathcal{H}}{\partial p_{\Sigma,k}} = \frac{\partial H}{\partial p_{\Sigma,k}} \\
\dot{p}^k_\Sigma &= -\frac{\partial \mathcal{H}}{\partial q^k_\Sigma} = -\frac{\partial H}{\partial q^k_\Sigma} \\
\dot{t} &= 1 \\
\dot{s} &= -\frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial H}{\partial t}
\end{aligned} \tag{2.29}$$

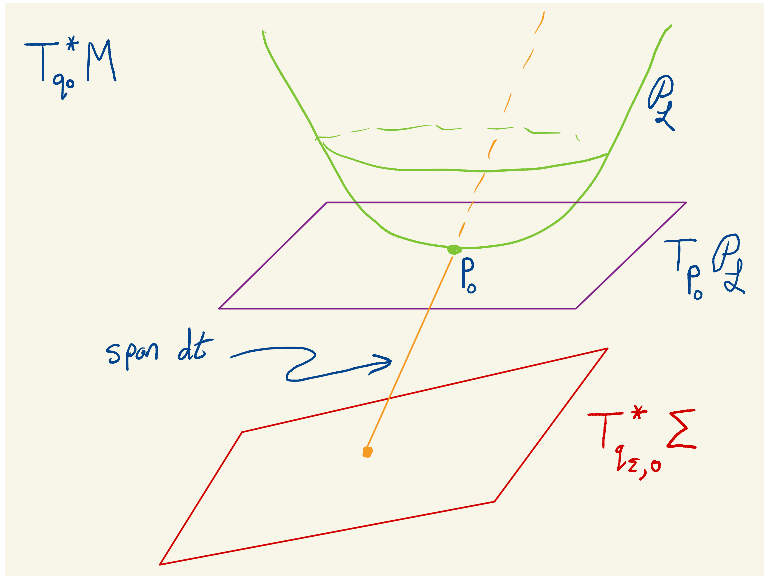

Figure 2: Local geometry of $\mathcal{P}_\mathcal{L}$.

Hence, as claimed earlier, with this choice of defining function $\mathcal{H}$, vortex curves are parameterized by t. The function H is independent of s and hence the first two equations decouple from the last equation. Moreover, the projection of a vortex curve back down into M depends only on q_Σ and t along the curve, and hence the final equation plays no role in these projections and can be ignored. Hence the first two equations are the primary evolution equations, and these are the familiar equations of motion stemming from a time-dependent Hamiltonian.

It remains to show that $\mathcal{P}_\mathcal{L}$ has the claimed structure as a graph near p_0 and to compute the function H from $\mathcal{L}$. Corollary 2.12 implies that

$$T_{p_0}\mathcal{P} = v_0^\perp.$$

Since $dt(v_0) \neq 0$, dt is transverse to $T_{p_0}\mathcal{P}_\mathcal{L}$. Since dt spans the kernel of the projection $T^*M \to T^*\Sigma$, it follows that the restriction of this projection to $\mathcal{P}_\mathcal{L}$ near p_0 is full rank and locally a diffeomorphism. Its inverse function takes p_Σ to $p_\Sigma - H(q_\sigma, p_\Sigma, t)dt$ for some function H, which establishes the local graph structure.

As far as the value of H goes, if p on $\mathcal{P}_\mathcal{L}$ then

$$\partial_t \lrcorner\, p = \partial_t \lrcorner\, (p_\Sigma - H(q_\Sigma, p_\Sigma, t)dt) = -H(q_\Sigma, p_\Sigma, t)$$

But each p on $\mathcal{P}_{\mathcal{L}}$ has the form $\mathbb{P}_{\mathcal{L}}(v)$ for some $v \in TM$ and

$$H = -\partial_t \lrcorner\, p = -\partial_t \lrcorner\, \mathbb{P}_{\mathcal{L}}(v).$$

That is, up to a sign, H is the momentum conjugate to ∂_t.

We would like to write H more explicitly as a function of p_Σ. To do this, we introduce a transformation $T\Sigma \to T^*\Sigma$ as follows. For $v_\Sigma \in T_{q_\Sigma}\Sigma$ first consider the map

$$v_\Sigma \to \mathbb{P}_{\mathcal{L}}(\partial_t + v_\Sigma)$$

into $\mathcal{P}_{\mathcal{L}}$. For $\partial_t + v_\Sigma$ near v_0, this map has full rank since the kernel of $D\mathbb{P}_{\mathcal{L}}$ at v_0 is spanned by v_0, which is transverse to $T\Sigma$. We have already established that near $p_0 = \mathbb{P}_{\mathcal{L}}(v_0)$ that projection from $\mathcal{P}_{\mathcal{L}}$ to $T^*\Sigma$ is a local diffeomorphism and hence, near v_0 we obtain diffeomorphism $\mathbb{p}(v_\Sigma) = \mathbb{P}_{\mathcal{L}}(\partial_t + v_\Sigma)|_{T\Sigma}$. This is, in effect, a velocity to momentum transformation relative to the spacetime gauge, and we write its inverse function as $\mathbb{v}$. In terms of $\mathbb{v}$,

$$H(q_\Sigma, p_\Sigma, t) = -\partial_t \lrcorner\, \mathbb{P}_{\mathcal{L}}(\partial_t + \mathbb{v}(p_\Sigma)).$$

We can make this expression appear more familiar by introducing the classical Lagrangian L induced by $\mathcal{L}$ and the spacetime gauge via the construction of Section **??**. Using the notation of this section,

$$\begin{aligned} L(v_\Sigma, t) &= (\partial_t + v_\Sigma) \lrcorner\, \mathcal{L}[\operatorname{span}(\partial_t + v_\Sigma)] \\ &= \mathbb{s}_{\mathcal{L}}(\partial_t + v_\Sigma). \end{aligned}$$

The map $\mathbb{p}$ is, in fact, the classical construction of spatial momenta. Indeed, if $\delta v_\Sigma \in T\Sigma$,

$$\begin{aligned} \delta v_\Sigma \lrcorner\, \mathbb{p}(v_\Sigma) &= \delta v_\Sigma \lrcorner\, \mathbb{P}_{\mathcal{L}}(\partial_t + v_\Sigma) \\ &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(\partial_t + v_\Sigma + \epsilon\delta v_\Sigma) \\ &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} L(\partial_t + v_\Sigma + \epsilon\delta v_\Sigma, t) \\ &= \delta v_\Sigma \lrcorner\, \frac{\partial L}{\partial v_\Sigma}(v_\Sigma, t). \end{aligned}$$

That is, using the local coordinates q^k_Σ introduced previously

$$\mathbb{p}(v_\Sigma) = \frac{\partial L}{\partial v^k_\Sigma} dq^k_\Sigma,$$

which is the classical momentum determined by v_Σ and L. The map $\mathbb{v}$ is the usual local inverse, which we have shown exists under the hypothesis that $\mathcal{L}$ satisfies the maximum rank condition and $dt(v_0) > 0$. The computation of $\partial_t \lrcorner \mathbb{P}_\mathcal{L}$ in Section **??** then shows

$$\begin{aligned} H &= -\partial_t \mathbb{P}_\mathcal{L}(\partial_t + \mathbb{v}(p_\Sigma)) \\ &= \mathbb{v}(p_\Sigma) \lrcorner p_\Sigma - L(\mathbb{v}(p_\Sigma), t) \\ &= v_\Sigma^k(p_\Sigma) p_{\Sigma,k} - L(v_\Sigma(p\Sigma), t). \end{aligned}$$

This is the classical time-dependent Hamiltonian determined from a time-dependent Lagrangian, and is the quantity appearing in equations (2.29).

2.12 Summary

1. Start with a fundamental Lagrangian $\mathcal{L}$. Perhaps it is expressed using a pragmatic Lagrangian or a classical Lagrangian.

2. Compute the velocity-to-momentum map $\mathbb{P}_\mathcal{L}$ by taking derivatives of the associated action density $\mathbb{s}_\mathcal{L}$. Both $\mathcal{L}$ and $\mathbb{s}_\mathcal{L}$ are gauge-independent, as is $\mathbb{P}_\mathcal{L}$.

3. The image of $\mathbb{P}_\mathcal{L}$, i.e. the set $\mathcal{P}_\mathcal{L}$ of feasible momenta, is a thin subset of T^*M, and we assume $\mathcal{L}$ satisfies the maximum rank condition so that locally it is an embedded hypersurface of T^*M.

4. T^*M posesses a universal 1-form Θ; the value of $\int_\gamma \Theta$ of a curve in T^*M is the action of the curve, a quantity independent of any Lagrangian.

5. Intervals Γ in M having satationary action correspond (by lifting parameterizations via $\mathbb{P}_\mathcal{L}$) to curves in $\mathcal{P}_\mathcal{L}$ for which the universal action S on T^*M is stationary, but only with respect to variations of curves in $\mathcal{P}_\mathcal{L}$.

6. Because action in the Hamiltoian setting reduces to integrating a universal one-form, variations of action are easy to compute and stationarity reduces to an algebraic condition on tangent vectors T along the cuve. Setting $\Omega = -d\Theta$, we require $T \lrcorner \Omega = 0$ when restricted to $\mathcal{P}_\mathcal{L}$.

7. In fact, Ω has a 1-dimensional annihilator on hypersurfaces of T^*M, which are then foliated by stationary curves. These are the vortex curves.

8. Until this point, no gauge choices have been made.

9. If a local defining function $\mathcal{H}$ for $\mathcal{P}_{\mathcal{L}}$ is known, then the vector field $X_{\mathcal{H}}$ defined by $X_{\mathcal{H}} \lrcorner \Omega = d\mathcal{H}$ is tangent to $\mathcal{P}_{\mathcal{H}}$ and is parallel to the vortex curves. We can integrate it to obtain stationary curves in $\mathcal{P}_{\mathcal{L}}$.

10. The choice of $\mathcal{H}$ selects, via $X_{\mathcal{H}}$, a parameterization of the vortex curves.

11. If in addition to $\mathcal{H}$ we make a choice of local coordinates, then the parameterized vortex curves satisfy Hamilton's equations, (2.27).

12. If a local defining function is not apparent, one can always be constructed by introducing a spacetime gauge (t, ∂_t) on M, which decomposes $M = \Sigma \times I$.

13. The surface $\mathcal{P}_{\mathcal{L}}$ is then expressed as a graph over $T^*\Sigma \times I$ via a function $H(q_\Sigma, p_\Sigma, t)$ that is, up to a sign, the momentum conjugate to ∂_t along $\mathcal{P}_{\mathcal{L}}$. That is, $H = -\partial_t \lrcorner \mathbb{P}_{\mathcal{L}}$.

14. In terms of the classical Lagrangian L determined by $\mathcal{L}$ and the spacetime gauge, we computed that $H = p_{\Sigma,k} q^k_\Sigma - L$. This expression was derived from computing $\partial_t \lrcorner \mathbb{P}_{\mathcal{L}}$ directly and was not given *a priori*.

15. The local defining function is $\mathcal{H} = s + H$ where $s = \partial_t \lrcorner \Theta$ and Hamilton's equations become equations (2.29), the classical equations of motion associated with a time-dependent Hamiltonian on $T^*\Sigma$. With this choice of $\mathcal{H}$, the vortex curves are parameterized by t.

3 Elements of Symplectic Geometry

4 Actions on Higher-Dimensional Surfaces

In Section **??** we introduced fundamental Lagrangians which, in a coordinate independent fashion, assign a number (the action) to compact one-dimensional submanifolds with boundary of an ambient space. Field theory requires an analogous device which assigns an action to higher-dimensional submanfolds.

Before developing the abstract machinery, it is helpful to consider at a concrete example from classical field theory, a massive scalar field with mass $m > 0$. Let (M^{d+1}, g) be a Lorenzian manfiold and suppose $u : M \to \mathbb{R}$. If K is a compact subset of M we can assign an action to the pair (K, u):

$$S[K,u] = \int_K \left[-\frac{1}{2}g(du,du) - mu^2\right] \mu_g. \tag{4.1}$$

We can think of action as a quantity, depending on u, that is assigned to arbitrary compact regions of spacetime. In a similar way, a classical Lagrangian can be thought of as assigning an action, depending on a curve γ, to intervals I

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

In this comparison, $(\mathbb{R}, I, \gamma)$ directly corresponds to (M, K, u).

Although coordinates in the traditional sense do not show up in the action (4.1), they are there implicitly. There is a surface $\Gamma \subset M \times \mathbb{R}$ consisting of the points $(p, u(p))$ and we have used $p \in M$ and $u \in \mathbb{R}$ as coordinates for this surface. To be clear, in this context, these are completely reasonable coordinates to use. Moreover, this perspective can be generalized to fiber bundles $E \to M$ and one can develop field theory based on jet bundles of sections of E [**?**]. As geometrically natural as this approach is for suitable models, it is nevertheless a coordinate-dependent description of field theory very much analogous to the coordinate-dependent description of classical Lagrangians on curves. The point is that action should be a number assgined to compact subset of Γ independent of how it is parameterized. If a parameterization of Γ as a graph over M is expedient for computation, one is welcome use it. But there are contexts, such as fluids, where a graph-over-spacetime description is not always natural and there is no single distinguished alternative parameterization of the surface associated with the field (see Section **??**).

Returning to the massive scalar field example, consider its associated graph Γ. We can assign a density at each point of Γ as follows. We have a projection $\pi_M : M \times \mathbb{R} \to M$. Using this projection, Γ acquires a Lorenzian metric $\tilde{g} = \pi^* g$. Similarly, we have a projection $\pi_{\mathbb{R}}$ and a function $\tilde{u} = \pi_{\mathbb{R}}^* \,\mathrm{id}$. At each point $w \in \Gamma$ we have a density on $T_w\Gamma$, namely

$$-\left[\frac{1}{2}\tilde{g}(d\tilde{u}, d\tilde{u}) + m\tilde{u}^2\right] \mu_{\tilde{g}} \tag{4.2}$$

At w, this density depends on Γ only via its tangent space $T_w\Gamma$: two surfaces Γ and Γ' passing through w such that $T_w\Gamma = T_w\Gamma'$ also share the same density on this common tangent space.

The benefit of this transformation is perhaps not evident: we've taken a familiar and servicable action (4.2) and decorated it with a number of tildes while making it more abstract. But this machine that consumes a tangent space and yields a density on the same tangent space, generalizes the fundamental Lagrangians of section **??** and yields a theory that applies equally to sections of fiber bundles over spacetime as well as fluids and related fields where the fiber bundle perspective is less natural. This section develops this generalization of fundamental Lagragians from Section **??**.

4.1 Densities

Densities on one-dimensional vector spaces played a central role in Section **??**, and we now require their higher-dimensional analogs. This section lightly reviews the facts we require, and we refer the reader to [Le13] Chapter ?? and for further details.

need another reference

A **density** on an n-dimensional vector space V is a map $\mu : V^n \to \mathbb{R}$ satisfying for any linear map $T : V \to V$

$$\mu(Tv_1, \dots, Tv_n) = |\det T|\, \mu(v_1, \dots, v_n). \tag{4.3}$$

It assigns a notion of "area" to the vector space: the area of the solid spanned by vectors v_k is $\mu(v_1, \dots, v_k)$. Unlike an n-form, the value of μ always has the same sign whenever its arguments are linearly independent. Nevertheless, just as for n-forms, we can assign a linear structure to the set $\mathcal{D}(V)$ of densities on V, and indeed it is one dimensional. The linear structure is the natural one: the sum of two densities is evidently again a density, and a scalar multiple of a density is also a density. Moreover, equation (4.3) implies that the value of a density is zero whenever its arguments are linearly dependent, and that if its value is known for some basis, then its value is known for every set of linearly independent arguments. It follows that the space of densities is one-dimensional so long as there is some non-zero density. In fact, any nonzero n-form ω determines a nonzero density by the following construction. We define the density $|\omega|$ by

$$|\omega|(v_1, \dots, v_n) := |\omega(v_1, \dots, v_n|$$

for arbitrary vectors $v_i \in V$. The analogous formula to (4.3) for n-forms shows that $|\omega|$ is a density on V, and it equals zero if and only if $\omega = 0$.

If $v \in V$ and if W is an $n-1$ dimensional subspace of V, we obtain an $n-1$ density $v \lrcorner \mu$ on W by

$$(v \lrcorner \mu)(w_1, \ldots, w_{n-1}) = \mu(w_1, \ldots, w_{n-1}, v).$$

Repeated interior products yield densities on lower dimensional subspaces.

The tangent bundle of a manifold M aquires a bundle of densities over it by a standard construction. Following [Le13], we denote it by

$$\mathcal{D}M = \coprod_{p \in M} \mathcal{D}(T_pM).$$

It is a line bundle over M and indeed is trivial. For example, the volume density of any metric on M determines a global section.

A section μ of $D(M)$ can be integrated over compact subsets of M^d. For example, suppose $D \subset \mathbb{R}^n$ is a domain of integration (a bounded subset with boundary having measure zero) and $\phi : D \to M$ is a diffeomorphism onto its image. Then

$$\int_K \mu = \int_K \mu(\phi_* \partial_1, \ldots, \phi_* \partial_n)\, dV_{\mathbb{R}^d}$$

where the latter integral is with respect to standard Lebesgue measure on $\mathbb{R}^d$. Equation (4.3) and the usual change-of-variables formula ensures that that this definition is independent of the choice of parameterization. A partition of unity argument allows for integration over compact sets in M that are not diffeomorphic to compact subsets of $\mathbb{R}^n$.

If $f : X^n \to Y^n$ is smooth, we can pull a density μ on Y back to $f^*\mu$ on X in the usual way: $f^*\mu(v_1, \ldots, v_n) = \mu(f_* v_1, \ldots, f_* v_n)$. Since densities form a linear space, and since they admit a suitable notion of pullback, the standard formula for the Lie derivative extends to densities.

Stokes's theorem for differential forms is intrinsically linked to orientation and does not generalize fully to densities. Nevertheless, we can define the divergence of a vector field X with respect to a nonvanishing density μ by

$$(\operatorname{div}_\mu X)\mu = \operatorname{Lie}_X \mu.$$

Moreover, a form of the divergence theorem holds. For a compact manifold with boundary Ω,

$$\int_\Omega \mathrm{Lie}_X\, \mu = \int_{\partial\Omega} \mathrm{sgn}(X)\, X \lrcorner\, \mu.$$

where $\mathrm{sgn}(X) = -1, 0, 1$ at each point of $\partial\Omega$ depending on whether X points to the interior, tangentially, or to the exterior of $\partial\Omega$. In particular, if μ_g is the volume density of a Riemannian metric and n is the outward pointing unit normal to $\partial\Omega$,

$$\int_\Omega \mathrm{div} X \mu) g = \int_{\partial\Omega} g(X, n)\, n \lrcorner\, \mu.$$

4.2 Blades and Multivectors

A tangent vector represents an infinitesimal parameterized curve. We require a kind of analogous construction for higher-dimensional parameterized surfaces; these are the k-blades which are dual versions of decomposable k-exterior forms. This section summarizes their construction.

Let V^d be a finite dimensional vector space. A rank k contravariant tensor $\alpha \in T^k V$ is canonically identified with a multilinear map $V^* \times \cdots \times V^* \to \mathbb{R}$ and it is **alternating** if its value changes sign when two of its arguments is interchanged. An alternating rank k contranvariant tensor is a $(k-)$**multivector** and the set of all k-multivectors is $\Lambda^k V$. Similarly, ΛV is the set of all multivectors of any rank.

This construction is likely more familiar in the covariant case: ΛV^* is the set of exterior forms on V and properties of multivectors can be deduced from those of exterior forms by exchanging V and V^*. In particular, ΛV is a graded algebra with the exterior product defined on vectors by

$$v \wedge w = v \otimes w - w \otimes v$$

and generalized to arbitrary multivectors in the same fashion as for exterior forms.

An element of $\Lambda^k V$ of the form

$$v_1 \wedge \cdots \wedge v_k \tag{4.4}$$

where each $v_j \in V$ is a k-**blade** (and is also more commonly known as a **decomposable** element of ΛV). We write $B^k V$ for the set of k-blades. These linearly span $\Lambda^k V$, which is a $\binom{d}{k}$ dimensional subspace of ΛV for $k \leq d$ and otherwise is a zero-dimensional subspace.

If v_j, $j = 1..k$ is a collection of vectors and if $w_i = T_i^j v_j$ for some matrix T of coefficients, then

$$w_1 \wedge \cdots \wedge w_k = \det T\, v_1 \wedge \cdots \wedge v_k. \tag{4.5}$$

This formula is the basis of half of the proof of the following.

Lemma 4.1. *Suppose* $\mathbf{v} = v_1 \wedge \cdots \wedge v_k$ *and* $\mathbf{w} = w_1 \wedge \cdots \wedge w_k$ *are nonzero elements of* $B^k V$ *and that* $\mathbf{v} = \alpha \mathbf{w}$ *for some* $\alpha \in \mathbb{R}$. *Then* $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(w_1, \dots, w_k)$, *and conversely.* maybe prove this

Hence a nonzero k-blade $v_1 \wedge \cdots \wedge v_k$ determines a k-dimensional subspace of V:

$$\langle v_1 \wedge \cdots \wedge v_k \rangle = \operatorname{span}(V_1, \dots, V_k). \tag{4.6}$$

If $\alpha \in \Lambda^k V$ and $\omega \in \Lambda^k V^*$ we define the interior product

$$\alpha \lrcorner \omega = \frac{1}{k!} \alpha \cdot \omega$$

where $\cdot$ denotes tensor contraction. For example, if $v_i \in V$ and $\omega^j \in V^*$, $j = 1, 2$, then

$$\begin{aligned} v_1 \wedge v_2 \lrcorner \omega^1 \wedge \omega^2 &= \frac{1}{2}\left[(v_1 \otimes v_2 - v_2 \otimes v_1) \cdot (\omega^1 \otimes \omega^2 - \omega^2 \otimes \omega^1)\right] \\ &= \omega^1(v_1)\omega^2(v_2) - \omega^2(v_1)\omega^1(v_2) \\ &= \det(\omega^j(v_i)) \end{aligned}$$

More generally,

$$v_1 \wedge \cdots \wedge v_k \lrcorner \omega^1 \wedge \cdots \omega^k = \det(\omega^j(v_i))$$

which motivates the coefficient $1/k!$ appearing in the definition of the interior product. We can also form the interior product of a d-blade α and a density on V itself: if μ is such a density, then $\mu = c|\omega|$ for some k-form ω and some $c \in \mathbb{R}$ and we define

$$\alpha \lrcorner \mu = c|\alpha \lrcorner \omega|.$$

One readily shows that this definition is indpendent of the choice of ω.

The linear space construction extends to the tangent space of a manifold M and have the bundle of k-multivectors

$$\Lambda^k TM = \coprod_{p \in M} \Lambda^k T_p M$$

analogous to the bundle of k-forms $\Lambda^k T^*M$. The set of sections of this bundle is $\mathfrak{X}^k(M)$, generalizing the set $\mathfrak{X}(M)$ of vector fields. Similarly,

$$B^k TM = \coprod_{p \in M} B^k T_p M$$

is the bundle of k-blades over M and is an embedded submanifold of $\Lambda^k TM$.

If $f : M \to N$ is smooth it induces maps $f_* \Lambda^k TM \to \Lambda^k TN$ and $f_* B^k TM \to B^k TN$ as follows. For a blade $\mathbf{v} = v_1 \wedge \cdots \wedge v_k$,

$$f_* \mathbf{v} = (f_* v_1) \wedge \cdots \wedge (f_* v_k).$$

Lemma 4.1 ensures that f_* is well defined. Every k-multivector is a linear combination of blades, and we extend f_* to $\Lambda^k TM$ by linearity. One needs to show that this is well defined.

do this for self.

A tangent vector V represents an infinitesimal parameterized curve and an ordered collection of tangent vectors $(V_1, \dots, V_k)$ can be thought of as an infinitesimal parameterized k-surface. A k-blade extracts just the part needed from the parameterization needed to compute integrals. Indeed, suppose $\omega \in \Omega^k(M)$ is a k-form and that $\Phi : D \to M$ is a smooth function that is a diffeomorphism onto it image. Then

$$\int_{\Phi(D)} \omega = \int_D \partial_1 \Phi \wedge \cdots \wedge \partial_k \Phi \lrcorner \omega \, dV_{\mathbb{R}^d}.$$

Because of this we introduce the notation

$$\mathcal{J}\Phi = \partial_1 \Phi \wedge \cdots \wedge \partial_k \Phi.$$

Similarly, suppose $\mu \in \Gamma(\mathcal{D}TM, M)$ and suppose $\Phi : D \subset \mathbb{R}^d \to M$ is a diffeomorphism onto its image. Then

$$\int_{\Phi(D)} \mu = \int_D \mathcal{J} \lrcorner \mu \, dV_{\mathbb{R}^d}.$$

4.3 Densities over Grassman Bundles

In Section 2.2 we described the bundle of densities over projective tangent bundles, which we now generalize to higher-dimensional subspaces.

use bold math for blades, not greek

Let V^d be a vector space. For each $0 \leq k \leq d$ we have the Grassman manifold $\mathrm{G}_k(V)$ of k-dimensional subspaces of V. As a set, the Grassman bundle of order k over some M^d with $d \geq k$ is

$$\coprod_{p \in M} \mathrm{G}_k(T_p M)$$

and it acquires the structure of a smooth bundle over M denoted by $\mathrm{G}_k TM$.

Each element $\mathfrak{p} \in \mathrm{G}_k(T_p M)$ is a k-dimensional vector space and hence has a bundle of densities $\mathcal{D}(\mathfrak{p})$ over it. Hence we obtain the bundle

$$\mathcal{D}\mathrm{G}_k(TM) = \coprod_{\mathfrak{p} \in \mathrm{G}_k(TM)} \mathcal{D}(\mathfrak{p})$$

which, in a standard construction, becomes a line bundle (consisting of densities) over $\mathrm{G}_k(TM)$. As in the lower-dimensional setting, the object being described here is more approachable than the machinery and notation needed to make it rigorous. Each element of $\mathcal{D}\mathrm{G}_k(TM)$ is simply a density on a k-dimensional subspace of a tangent space of M.

4.4 Fundamental Lagrangians

At least initially, the theory of Lagrangians on higher-dimensional surfaces closely parallels what we have already laid out for Lagrangians on curves in Section 2. So we proceed briskly for now.

Definition 4.2. Let M^d be a manifold and suppose $1 \leq k \leq d$. A **fundamental Lagrangian** of order k on M is a section $\mathcal{L}$ of $\mathcal{D}\mathrm{G}_k(TM)$ over an open subset of $\mathrm{G}_k(TM)$.

Suppose $\mathcal{L}$ is a fundamental Lagrangian of order k on some manifold M let F be a compact k-dimensional submanifold with boundary of M such that $T_q F$ is in the domain of $\mathcal{L}$ for all $q \in F$. Such a submanfiold is called **admissible** with respect to $\mathcal{L}$. Let ι be the natural embedding of F into M. At each $q \in F$, $\iota_*(T_q F)$ is a k-dimensional subspace of M and $\mathcal{L}[\iota_*(T_q\Gamma)]$ is a density on $\iota_*(T_q(\Gamma))$. Moreover, it determines a density $\iota^*\mathcal{L}$ on $T_q(F)$ via

$$\iota^*\mathcal{L}(v_1, \ldots, v_k) = \mathcal{L}(T_q\Gamma)(\iota_* v_1, \ldots, \iota_* v_k).$$

The **action** of F with respect to $\mathcal{L}$ is

$$S_{\mathcal{L}}[\Gamma] = \int_\Gamma \iota^*\mathcal{L}.$$

If $\Phi : D \to F$ is a diffeomorphism from a domain of integration $D \subset \mathbb{R}^k$ onto its image then
$$S_{\mathcal{L}}[\Phi(D)] = \int_D \mathcal{J}\Phi \lrcorner \mathcal{L}[\langle \mathcal{J}\Phi \rangle] dV_{\mathbb{R}^d}$$
where $\langle J\Phi \rangle$ is the subspace determined by $J\Phi$ as defined in equation (4.6).

If F is a compact k-dimensional submanifold with boundary in M, a **variation** of F is a smooth family of maps $\Phi_s : F \to M$ for $s \in (-\epsilon, \epsilon)$ for some $\epsilon > 0$ satisfying

mismatch: domains of integration are open. Their images are never compact.

1. $\Phi_0 = \mathrm{id}$,
2. $\Phi_s|_{\partial F} = \mathrm{id}$ for all $s \in (-\epsilon, \epsilon)$.

We say F is stationary for the action if
$$\left.\frac{d}{ds}\right|_{s=0} S_{\mathcal{L}}[\Psi_s(\Gamma)] = 0.$$
for every variation of F.

The action density associated with $\mathcal{L}$ is the map $\mathbb{s}_{\mathcal{L}}$ defined by
$$\mathbb{s}_{\mathcal{L}}(\mathbf{v}) = \alpha \lrcorner \mathcal{L}[\langle \mathbf{v} \rangle]$$
for all k-blades $\mathbf{v}$ such that $\langle \mathbf{v} \rangle$ lies in the domain of $\mathcal{L}$. Hence it is a smooth map from an open subset of $B^k TM$ to $\mathbb{R}$. The following analog of Lemma 4.3 follows from the definitions.

Lemma 4.3. *Let $\mathbb{s}_{\mathcal{L}}$ be the action density of a fundamental Lagrangian L. For any k-blade* $\mathbf{v}$ *such that $\langle \mathbf{v} \rangle$ is in the domain of $\mathcal{L}$,*
$$\mathbb{s}_{\mathcal{L}}(\alpha\, \mathbf{v}) = |\alpha|\, \mathbb{s}_{\mathcal{L}}(\mathbf{v}) \tag{4.7}$$
for all nonzero $\alpha \in \mathbb{R}$.

4.5 Pragmatic Lagrangians

The following definition parallels the definition of pragmatic Lagrangian from section 2.4.

Definition 4.4. Let M^d be a manifold, and let U be an open subset of B^kTM consisting of nonzero k-blades such that whenever $\mathbf{v} \in U$, $\alpha\mathbf{v} \in U$ for all $\alpha \neq 0$. A **pragmatic Lagrangian** on U is a smooth bundle map map $\mathbb{L} : U \to \Lambda^k T^*M$ satisfying

$$\mathbb{L}(\alpha\mathbf{v}) = \frac{\alpha}{|\alpha|}\mathbb{L}(\mathbf{v})$$

for all $\mathbf{v} \in U$ and all nonzero $\alpha \in \mathbb{R}$.

This definition reduces to that of Definition **??** when $k = 1$ but it is important to observe that the cases $k = 1$ and $k = d - 1$ are special. In general, the k-blades are a strictly embedded submanifold of the bundle of k-multivectors. However, $B^1TM = \Lambda^1TM = TM$ and $B^{d-1}TM = \Lambda^{d-1}TM$ and in these cases the fibers over M are vector spaces. This leads to some significant simplifications in these cases and I don't have the full theory worked out for the intermediate values of k yet. Yet.

Let $\mathbb{L}$ be a practical Lagrangian. It induces a fundamental Lagrangian with domain consisting of the elements of G_kTM of the form $\langle\mathbf{v}\rangle$ for some $\mathbf{v}$ in the domain of $\mathbb{L}$ as follows. Consider some $\langle\mathbf{v}\rangle$ and let $w_1, \dots, w_k$ be vectors in $\langle\mathbf{v}\rangle$. We define

$$\mu_{\langle\mathbf{v}\rangle}(w_1, \dots, w_k) = (w_1 \wedge \cdots \wedge w_k) \lrcorner \mathbb{L}(w_1 \wedge \cdots \wedge w_k)$$

with the understanding that $\mathbb{L}(0) = 0$. To see that this defines a density on $\langle\mathbf{v}\rangle$, consider some linear map $T : \langle\mathbf{v}\rangle \to \langle\mathbf{v}\rangle$. If $\det T = 0$ or if $w_1 \wedge \cdots \wedge w_k = 0$ then $\mu_{\langle\mathbf{v}\rangle}(Tw_1, \dots, Tw_k) = |\det T|\mu_{\langle\mathbf{v}\rangle}(w_1, \dots, w_k)$ trivially. Otherwise,

$$\begin{aligned}
\mu_{\langle\mathbf{v}\rangle}(Tw_1, \dots, Tw_k) &= (Tw_1 \wedge \cdots \wedge Tw_k) \lrcorner \mathbb{L}(Tw_1 \wedge \cdots \wedge Tw_k) \\
&= (\det Tw_1 \wedge \cdots \wedge w_k) \lrcorner \mathbb{L}(\det Tw_1 \wedge \cdots \wedge w_k) \\
&= \det T(w_1 \wedge \cdots \wedge w_k) \lrcorner \frac{\det T}{|\det T|}\mathbb{L}(w_1 \wedge \cdots \wedge w_k) \\
&= |\det T|\mu_{\langle\mathbf{v}\rangle}(w_1, \dots, w_k)
\end{aligned}$$

as required.

A metric on a manifold M induces, fiberwise, a metric on B^kTM. Using this fact one readily adapts the argument of Lemma 2.6 and we obtain the following.

Lemma 4.5. *Let $\mathcal{L}$ be a fundamental Lagrangian on M. There exists a pragmatic Lagrangian $\mathbb{L}$ on M that induces $\mathcal{L}$:*

$$\mathcal{L} = \mathcal{L}_{\mathbb{L}}.$$

Concretely, if g is a metric on M then we can take

$$\mathbb{L}(\mathbf{v}) = (\mathbf{v} \lrcorner \mathcal{L}[\langle \mathbf{v} \rangle]) \frac{\mathbf{v}^\flat}{|\mathbf{v}|_g^2}$$

where $\mathbf{v}^\flat$ is the k-form dual, via g, to the k-blade $\mathbf{v}$.

If $\mathbb{L}$ induces $\mathcal{L}$ then

$$\mathbb{s}(\mathbf{v}) = \mathbf{v} \lrcorner \mathbb{L}(\mathbf{v}).$$

If $\Phi : D \to M$ is a parameterization of $\Phi(D)$ for some domain of integragion $D \subset \mathbb{R}^k$ then

$$S_{\mathcal{L}}[\Phi(D)] = \int_D \mathcal{J}\Phi \lrcorner \mathbb{L}(J\Phi)\, dV_{\mathbb{R}^k}.$$

4.6 Examples

4.6.1 Surface Area

Let (M^d, g) be a Riemannian metric and suppose $1 \le k \le d$. Each k-dimensional subspace of a tangent space T_qM inherits a Riemannian density $\mu_{g,k}$, and this is a fundamental Lagrangian. We can represent it in terms of a pragmatic Lagrangian as follows:

$$\mathbb{L}(\mathbf{v}) = \frac{\mathbf{v}^\flat}{|\mathbf{v}|_g}$$

where $\mathbf{v}^\flat$ is the k-form dual (via g) to $\mathbf{v}$.

4.6.2 Scalar Fields

The ambient manifold is $M \times \mathbb{R}$ where (M^{d+1}, g) is a Lorenzian metric. Let $\mathscr{k}$ be an element of $G_{d+1}T(M \times \mathbb{R})$ with a full rank projection onto M (i.e., $(\pi_M)_*|_{\mathscr{k}}$ is full rank). Then $\mathscr{k}$ acquires two constructs: a Lorenzian metric $\tilde{g} = \pi_M^* g$ and a covector $\widetilde{du} = \pi_{\mathbb{R}}^* du$, where u is the coordinate on $\mathbb{R}$. Then

$$\mathcal{L}[\mathscr{k}] = -\tilde{g}(\widetilde{du}, \widetilde{du})\mu_{\tilde{g}}.$$

We can describe this Lagrangian in terms of a pragmatic Lagranian presented in local coordinates $(x^0, \dots, x^d, u)$ on $M \times \mathbb{R}$. Let $\partial_0, \dots, \partial_d, \partial_u$ and $dx^0, \dots, dx^d, du$ be the associated coordinate vectors and 1-forms. A generic $(d+1)$-blade that projects fully onto M has the form

$$\mathbf{v} = \alpha((dx_0 + u_0\partial_u) \wedge \cdots \wedge (dx_d + u_d\partial_u)) = \alpha\left(\mathbf{w} + \sum_{j=0}^{d} u_j \partial_u \wedge (dx^j \lrcorner \mathbf{w}_x)\right)$$

where α and u_j are coefficients and where $\mathbf{w}_x = \partial_0 \wedge \cdots \wedge \partial_d$. Then

$$\mathbb{L}(\mathbf{v}) = -\frac{\alpha}{2|\alpha|} g^{ij} u_i u_j \sqrt{-\det g}\, dx^0 \wedge \cdots \wedge dx^d.$$

This same pragmatic Lagrangian can be described without coordinates as follows. Let G be the metric $g + du^2$ on $M \times \mathbb{R}$. Let dV_g be one of the two volume forms on M and pull it back, with the same name, to $M \times \mathbb{R}$. The metric on $M \times \mathbb{R}$ induces a metric on k-blades such that $G(e_1 \wedge \cdots \wedge e_k, e_1 \wedge \cdots \wedge e_k) = g(e_1, e_1) \cdots g(e_k, e_k)$ whenever the e_j's are mutually orthogonal. Let $\mathbf{w}$ be a $d+1$ blade in $M \times \mathbb{R}$ that is parallel to M such that $dV_g(\mathbf{w}) = 1$; this specifies $\mathbf{w}$ completely. Then

$$\mathbb{L}(\mathbf{v}) = \frac{1}{2}\frac{G(\mathbf{v}, \mathbf{w})}{|G(\mathbf{v}, \mathbf{w})|} G\left(\frac{\mathbf{v} + G(\mathbf{v}, \mathbf{w})\mathbf{w}}{G(\mathbf{v}, \mathbf{w})}, \frac{\mathbf{v} + G(\mathbf{v}, \mathbf{w})\mathbf{w}}{G(\mathbf{v}, \mathbf{w})}\right) dV_g.$$

4.6.3 Dust

4.6.4 Charged Scalar Fields

4.7 The Tangent Space and Cotangent Space of $B^k V$

In this section explain

- The k-multivectors on V^d are a $\binom{d}{k}$ dimensional vector space.
- The k-blades are, by contrast, a $k(d-k)+1$ dimensional submanifold of the k multivectors. At $\mathbf{v} = e_1 \wedge \cdots \wedge e_k$ the tangent space is spanned by $\mathbf{v}$ itself along with $e_j \wedge (e^i \lrcorner \mathbf{v}$ with $1 \le i \le k$ and $k+1 \le j \le d$.

- If $k = 0$, $k = 1$, $k = d - 1$ or $k = d$ then the dimension counts coencide and the k-blades are exactly the k-multivectors and acquire a vector space structure.
- When $k = 0$ the k-blades are just the scalars, which are the 0-multivectors and the space is 1-dimensional.
- If $k = 1$ then the 1-blades comprise V itself, and the space is d dimensional. This is the setting of particle mechanics.
- If $k = d - 1$ then the space of blades is again d-dimensional and is the entirety of $\Lambda^{k-1}V$.
- When $k = d$ then the space of blades is simply the one-dimensional space of d-multivectors.
- For these four cases, $T_{\mathbf{v}}B^kV$ can be identified with Λ^kV and its dual space is naturally identitified with Λ^kV^*.
- Otherwise, $T_{\mathbf{v}}B^kV$ is a subspace of Λ^kV.
- Every element of Λ^kV^* determines an element of $T_{\mathbf{v}}B^k$.
- But some nonzero elements of Λ^kV^* vanish on $T_{\mathbf{v}}B^k$. Let's call that space $K_{\mathbf{v}}$. It is a linear subspace and indeed

$$\dim K_{\mathbf{v}} = \binom{d}{k} - k(d-k) - 1.$$

- We can identify $T_{\mathbf{v}}B^kV$ with $\Lambda^kV^*/K_{\mathbf{v}}$.
- This identification depends on $\mathbf{v}$, which is inconvenient and does not allow us to easily parallel the construction of momenta as elements of a space that is independent of $\mathbf{v}$. For particle mechanics, a momentum corresponding to $\mathbf{v} \in T_qM$ is an element of T_q^*M with no reference to the $\mathbf{v}$ that generated it.
- We go down the least worst path: momenta will admit more than one representation. Momenta will be elements of Λ^kT^*M and it will simply be the case that there is a whole subspace of equivalent momenta once a corresponding 'conjugate velocity' has been determined.

4.8 Momentum

Let $\mathcal{L}$ be a fundamental Lagrangian of order k and, for simplicity of discussion assume that its domain is all of $B^k TM$. The action density assocated with $\mathcal{L}$ is

$$\mathbb{s}_{\mathcal{L}}(\mathbf{v}) = \mathbf{v} \lrcorner \mathcal{L}[\langle \mathbf{v} \rangle]$$

which is a smooth map $B^k TM \to \mathbb{R}$. Fix $q \in M$ and let $\mathbb{s}_{\mathcal{L},q}$ be the restriction of $\mathbb{s}_{\mathcal{L}}$ to $B^k T_q(M)$. In Section 2.7 we considered the case $k = 1$ and defined

$$\mathbb{P}_{\mathcal{L}}(\mathbf{v}) = D\mathbb{s}_{\mathcal{L},q}|_{\mathbf{v}} \in T^*_{\mathbf{v}} T_q M = T^*_q M = (\Lambda^1 T^*)_q M.$$

In this way we obtained a bundle map

$$B^1 TM = TM \to T^* M = \Lambda^1 T^* M.$$

We would like to replicate this procedure for other values of k, but in general we cannot identify

$$T^* B^k T_q M \sim (\Lambda^k T^*)_q M$$

Instead, as discussed in the previous section, every element of $T^* B^k T_q M$ can be represented by an elements of $\Lambda^k T^*_q M$, but for $k \neq 0, 1, d-1, d$ there an entire subspace of $\Lambda^k T^*_q M$ that represents the same element of $T^* B^k T_q M$.

With this important distinction in mind, we define define

$$\mathbb{P}_{\mathcal{L}}(\mathbf{v}) = \left\{ \omega \in \Lambda^k T^*_q M \,:\, \omega|_{T_{\mathbf{v}} B^k T_q M} = D\mathbb{s}_{\mathcal{L},q}\big|_{\mathbf{v}} \right\}.$$

The momentum determined by the k-blade $\mathbf{v}$ at $q \in M$ is an entire affine subspace of $\Lambda^k T^*_q M$, all of which represent the same element of $T^*_{\mathbf{v}} B^k T_q M$.

If $\mathbf{w} \in T_{\mathbf{v}} B^k TM$ we can unambiguously define

$$\mathbf{w} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathbf{v}) = \mathbf{w} \lrcorner \omega$$

where ω is any element of $\mathbb{P}_{\mathcal{L}}(\mathbf{v})$ since any two such elements differ by an element of $\Lambda^k T^*_q M$ that vanishes on $T_{\mathbf{v}} B^k TM$.

Let γ be a curve in $B_k T_q M$ with $\gamma(0) = \mathbf{v}$ and $\dot{\gamma}(0) = \mathbf{w} \in T_{\mathbf{v}} B_k T_q M$. Then

$$\mathbf{w} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathbf{v}) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \mathbb{s}_{\mathcal{L}}(\gamma(\epsilon)). \tag{4.8}$$

4.9 Consequences of Parameterization Invariance

Lemma 4.6. *Suppose $\mathcal{L}$ is a fundamental Lagrangian of order k on M. Its associated velocity-to-momentum map $\mathbb{P}_{\mathcal{L}}$ satisfies*

$$\mathbb{P}_{\mathcal{L}}(\alpha\mathbf{v}) = \frac{\alpha}{|\alpha|}\mathbb{P}_{\mathcal{L}}(\mathbf{v})$$

for all $\alpha \neq 0$ and all $\mathbf{v} \in B^k TM$ with $\langle\mathbf{v}\rangle$ in the domain of $\mathcal{L}$. Moreover,

$$\mathbf{v} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathbf{v}) = \mathbf{v} \lrcorner \mathcal{L}[\langle\mathbf{v}\rangle] = \mathbb{s}_{\mathcal{L}}(\mathbf{v}).$$

Proof. Let γ be a curve in $B^k TM$ with $\gamma(0) = \mathbf{v}$ and $\dot{\gamma}(0) = \delta\mathbf{v} \in T_{\mathbf{v}} B^k TM$. Then

$$\begin{aligned}
(\delta\mathbb{v}) \lrcorner (\mathbb{P}_{\mathcal{L}}(\alpha\mathbf{v})) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \mathbb{s}(\alpha\gamma(\epsilon/\alpha)) \\
&= |\alpha| \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \\
&= |\alpha| \frac{\delta\mathbf{v}}{\alpha} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathbf{v}) \\
&= \frac{\alpha}{|\alpha|}\delta\mathbf{v} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathbf{v}).
\end{aligned}$$

Moreover,

$$\mathbf{v} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathbf{v}) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \mathbb{s}(\mathbf{v} + \epsilon\mathbf{v}) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} |1+\epsilon|\mathbb{s}(\mathbf{v}) = \mathbb{s}(\mathbf{v}) = \mathbf{v} \lrcorner \mathcal{L}[\langle\mathbf{v}\rangle].$$

□

Feasible Momenta The feasible momenta are

$$\mathcal{P}_{\mathcal{L}} = \cup_{\mathrm{v}\in U}\{\mathbb{P}_{\mathcal{L}}(\mathbf{v})\}$$

which is, on the face of things, a subset of $\Lambda^k T^* M$. Somewhere around here talk about the "rank" of $\mathbb{P}_{\mathcal{L}}$. Maybe via an intermediate map into $T^* B_k T_q M$. And establish that in the generic case, $\mathcal{P}_{\mathcal{L}}$ is locally a hypersurface in $\Lambda^k M$.

4.10 Action From the Hamiltonian Perspective

If $\Psi : D \to F$ is a parameterization,

$$S_{\mathcal{L}}[F] = \int_D \mathcal{J}\Phi \lrcorner L[\langle \mathcal{J}\Phi \rangle].$$

Using Lemma 4.6 we can rewrite this as

$$S_{\mathcal{L}}[F] = \int_D \mathcal{J}\Phi \lrcorner \mathbb{P}_{\mathcal{L}}(\mathcal{J}\Phi) dV_{\mathbb{R}^k}.$$

Tautological k-form: At $p \in \Lambda^k T_q M$

$$\Theta^k(v_1, \dots, v_k) = p(\pi_* v_1, \dots, \pi_* v_k).$$

In terms of coordinates,

$$\Theta = p_I dq^I$$

where I is an increasing sequence of integers in $1, \dots, d$ and where $dq^I = dq^{I_1} \wedge \cdots \wedge dq^{I_k}$.

We say $\tilde{\Psi} : M \to \Lambda^k T^* M$ is a lift of Ψ if

$$\tilde{\Psi}(q) \in \mathbb{P}_{\mathcal{L}}(\mathcal{J}\Psi(q))$$

at each point q in the domain of Ψ. For such a lift,

$$S_{\mathcal{L}}[F] = \int_D \mathcal{J}\Phi \lrcorner \tilde{\Psi}(q) dV_{\mathbb{R}^k} = \int_{\tilde{\Psi}} \Theta^k.$$

Motivated by this, for any map $\tilde{\Psi}$ from a domain of integration in $\mathbb{R}^k$ into $\Lambda^k T^* M$ we define the action

$$S[\tilde{\Psi}] = \int_{\tilde{\Psi}} \Theta^k$$

which depends on the map $\tilde{\Psi}$ but is unrelated to any Lagrangian.

Now look at a variation of maps $\tilde{\Psi}_s$. Let $X = d/ds|_{s=0} \tilde{\Psi}_s$. Then

$$\frac{d}{ds}|_{s=0} S[\tilde{\Psi}_s] = \int_{\partial \tilde{\Psi}} X \lrcorner \Theta + \int_{\tilde{\Psi}} X \lrcorner d\Theta^k.$$

For variations that fix the boundary,

$$\frac{d}{ds}|_{s=0} S[\tilde{\Psi}_s] = \int_{\tilde{\Psi}} X \lrcorner d\Theta^k.$$

The form

$$\Omega = -d\Theta$$

is the so-called multisymplectic $k+1$-form on $\Lambda^k T^* M$.

This form is non-degenerate: if $X = p^I \partial_{p^I} + q^k \partial_{q^k}$ then

$$X \lrcorner \Omega = p_I dq^I + dp^I \wedge q^k (\partial_k \lrcorner dq^I)$$

So for this to vanish it is clear that all the p_I's must vanish. But then some q^k does not vanish and without loss of generality we can assume $k = 1$. We can apply $X \lrcorner \Omega$ to $\partial_{p_{I_0}} \wedge \partial_2 \cdots \partial_k$ where $I_0 = (1, 2, \dots, k)$ and obtain a value of $q^k \neq 0$.

Proposition 4.7. *Suppose $v \in B^k T_q M$ has a span $\langle \mathbf{v} \rangle$ in the domain of $\mathcal{L}$. Let γ be a curve in $B^k T_q M$ with $\gamma(0) = \mathbf{v}$ and suppose $\tilde{\gamma}$ is a curve in $\Lambda^k T_q^* M$ with $\tilde{\gamma}(\lambda) \in \mathbb{P}_{\mathcal{L}}(\gamma(\lambda))$ for all λ. Then*

$$\dot{\tilde{\gamma}}(0) \in \mathbf{v}^{\perp}.$$

Proof. For notational convenience, let $\mathbf{w} = \dot{\gamma}(0)$. From equation (4.8) and the assumption $\tilde{\gamma}(\lambda) \in \mathbb{P}_{\mathcal{L}} \circ \gamma(\lambda)$ for all parameters λ we have

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} \mathbb{s}_{\mathcal{L}}(\gamma(\lambda)) = \mathbf{w} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathbf{v}) = \mathbf{w} \lrcorner \tilde{\gamma}(0). \tag{4.9}$$

On the other hand, Lemma 4.6 along with the previously mentioned assumption on $\tilde{\gamma}$ imply

$$\mathbb{s}_{\mathcal{L}}(\gamma(\lambda)) = \gamma(\lambda) \lrcorner \mathbb{P}_{\mathcal{L}}(\gamma(\lambda)) = \gamma(\lambda) \lrcorner \tilde{\gamma}(\lambda).$$

Taking a derivative and using equation (4.9) we find

$$\mathbf{w} \lrcorner \tilde{\gamma}(0) = \dot{\gamma}(0) \lrcorner \tilde{\gamma}(0) + \gamma(0) \lrcorner \dot{\tilde{\gamma}}(0) = \mathbf{w} \lrcorner \tilde{\gamma}(0) + \mathbf{v} \lrcorner \dot{\tilde{\gamma}}(0).$$

Hence

$$\mathbf{v} \lrcorner \dot{\tilde{\gamma}}(0) = 0.$$

□

Maximum rank condition: About each $\mathbf{v} \in B^k TM$ with a span in the domain of $\mathcal{L}$ there is an open set such that $\mathbb{P}_{\mathcal{L}}$ as a map into the affine subspaces of dimension $\binom{d}{k} - k(d-k) - 1$ of $\Lambda^k T^*M$ has rank $k(d-k)$ (i.e.,)and moreover $\mathcal{P}_{\mathcal{L}}$ is an embedded hypersurface of $\Lambda^k T^*M$.

Corollary 4.8. *Suppose $\mathcal{L}$ satisfies the maximum rank condition. Let $p \in \mathcal{P}_{\mathcal{L}}$, so $p \in \mathbb{P}_{\mathcal{L}}(\mathbf{w})$ for some $\mathbf{w} \in B^k TM$. Then*

$$T_p \mathcal{P}_{\mathcal{L}} = v^{\perp}$$

Lemma: If $\tilde{\Psi}$ is a lift of Ψ then $\pi_* \mathcal{J} \tilde{\Psi} = \mathcal{J} \Psi$.

Proposition 4.9. *Suppose $\mathcal{L}$ satisfies the maximum rank condition. If E is a stationary [needs tacking down] in M and if Ψ is a parameterization of E then any lift $\tilde{\Psi}$ of Ψ is stationary in $\mathcal{P}_{\mathcal{L}}$.*

Proof. Sketch. Consider an arbitrary variation $\tilde{\Psi}_s$ of $\tilde{\Psi}$ in $\mathcal{P}_{\mathcal{L}}$. Now

Sloppy here about $\dot{\gamma}$ needing to not vanish downstairs.

$$\Theta(\mathcal{J} \tilde{\Psi}_s) = \pi_*(\mathcal{J} \tilde{\Psi}_s) \lrcorner \tilde{\Psi}_s$$

and hence

$$\begin{aligned} \left.\frac{d}{ds}\right|_{s=0} \Theta(\mathcal{J} \tilde{\Psi}_s) &= \left(\left.\frac{d}{ds}\right|_{s=0} \pi_*(\mathcal{J} \tilde{\Psi}_s) \right) \lrcorner \tilde{\Psi}_0 + \pi_*(\mathcal{J} \tilde{\Psi}_0) \lrcorner \left.\frac{d}{ds}\right|_{s=0} \tilde{\Psi}_s \\ &= \left(\left.\frac{d}{ds}\right|_{s=0} \pi_*(\mathcal{J} \tilde{\Psi}_s) \right) \lrcorner \mathbb{P}_{\mathcal{L}}(\mathcal{J}\Psi) + \mathcal{J}\Psi \lrcorner \left.\frac{d}{ds}\right|_{s=0} \tilde{\Psi}_s. \end{aligned}$$

For fixed q, $\tilde{\Psi}_s(q)$ is a curve in $\mathcal{P}_{\mathcal{L}}$ starting at $\Psi(q) \in \mathbb{P}_{\mathcal{L}}(\mathcal{J}\Psi(q))$. So Proposition 4.7 implies

$$\mathcal{J}\Psi \lrcorner \left.\frac{d}{ds}\right|_{s=0} \tilde{\Psi}_s = 0$$

and therefore

$$\left.\frac{d}{ds}\right|_{s=0} \Theta(\mathcal{J} \tilde{\Psi}_s) = \left(\left.\frac{d}{ds}\right|_{s=0} \pi_*(\mathcal{J} \tilde{\Psi}_s) \right) \lrcorner \mathbb{P}_{\mathcal{L}}(\mathcal{J}\Psi).$$

As a consequence,

$$\left.\frac{d}{ds}\right|_{s=0} S[\tilde{\Psi}_s] = \int_D \mathbf{w} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathcal{J}\Psi) \tag{4.10}$$

where

$$\mathbf{w} = \left.\frac{d}{ds}\right|_{s=0} \pi_* \mathcal{J} \tilde{\Psi}_s = \left.\frac{d}{ds}\right|_{s=0} \mathcal{J}(\pi \circ \tilde{\Psi}_s).$$

maybe flesh out some of this in a lemma?

Let $\Phi_s = \pi \circ \tilde{\Psi}_s$, so Φ_s is a variation of Ψ. Let $\tilde{\Phi}_s$ be a family of lifts of Φ_s. Then, by stationarity,

$$\frac{d}{dt} S[\tilde{\Psi}_s] = 0$$

Really need a lemma that lifts exist and that lifts of families exist.

Equation (4.10) applies equally to the variation $\tilde{\Phi}_s$ and hence

$$0 = \int_D \hat{\mathbf{w}} \lrcorner \mathbb{P}_{\mathcal{L}}(\mathcal{J}\Phi)$$

where

$$\hat{\mathbf{w}} = \frac{d}{ds}\bigg|_{s=0} \pi_* \mathcal{J}\tilde{\Phi}_s = \frac{d}{ds}\bigg|_{s=0} \mathcal{J}\pi \circ \tilde{\Phi}_s = \frac{d}{ds}\bigg|_{s=0} \mathcal{J}\pi \circ \tilde{\Psi}_s$$

We conclude that the integral on the right-hand side of equation (4.10) vanishes and consequently $\tilde{\Psi}$ is stationary in $\mathcal{P}_{\mathcal{L}}$. $\square$

Fix the sign ambiguity problem here!

Proposition 4.10. *Suppose $\mathcal{L}$ satisfies the maximum rank condition. If $\tilde{\Psi}$ is stationary in $\mathcal{P}_{\mathcal{L}}$ then it is the lift of a parameterisation of a k-surface in M.*

Proof. Sketch. Suppose $\tilde{\Psi}$ is not such a lift and let $\Psi = \pi \circ \tilde{\Psi}$. Then there is some $x^* \in D$ such that $\tilde{\Psi}(x^*) \notin \mathbb{P}_{\mathcal{L}}(\mathcal{J}\Psi(x^*))$. Set $p^* = \tilde{\Psi}(x^*)$ and $\mathbf{v}^* = \mathcal{J}\Psi(x^*)$. Now $p^* \in \mathcal{P}_{\mathcal{L}}$ so there exists a k-blade $\mathbf{w}$ at $\Psi(x^*)$ such that $p^* \in \mathbb{P}_{\mathcal{L}}(\mathbf{w})$. Now $\mathbf{w}$ is not a multiple of $\mathbf{v}$ for otherwise $\mathbb{P}_{\mathcal{L}}(\mathbf{v}) = \pm\mathbb{P}_{\mathcal{L}}(\mathbf{w})$ (oops. Same sign ambiguity as before). Hence there exists $p \in \mathbf{w}^\perp$ with $\mathbf{v}^* \lrcorner p \neq 0$. Corollary 4.8 implies $\mathbf{w}^\perp = T_{p^*}\mathcal{P}_{\mathcal{L}}$ and hence we can consider p as an element of $T_{p^*}\mathcal{P}_{\mathcal{L}}$. But then

$$d\Theta(p \wedge \mathcal{J}\tilde{\Psi}) = p(\pi_* \mathcal{J}\tilde{\Psi}) = p(\mathcal{J}\Phi) = \mathbf{v}^* \lrcorner p \neq 0.$$

Need a lemma: $\Omega(p \wedge w) = -p(\pi_* w)$ if p is vertical.

Since p is tangent to $\mathcal{P}_{\mathcal{L}}$ at $\tilde{\Psi}(x^*)$ it follows that $\tilde{\Psi}$ is not stationary. $\square$

4.11 Equations of Motion

We forget about $\mathcal{L}$ and work with embedded hypersurfaces of $\Lambda^k T^* M$.

The following is a straightforward adaptation of Proposition 2.15.

Proposition 4.11. *Let $\mathcal{P}$ be an embedded hypersurface in $\Lambda^k T^*M$. A parameterization $\Psi :$ $D \to \mathcal{P}$ is stationary if and only if*

$$d\Theta(V \wedge \mathcal{J}\Psi(x)) = 0$$

for all $x \in D$ and all $V \in T_{\Psi(x)}\mathcal{P}$.

References

[Ar97] V. I. Arnol'd, *Mathematical methods of classical mechanics*, 2nd ed ed., Graduate texts in mathematics, no. 60, Springer, New York, 1997 (eng).

[GPS08] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical mechanics*, 3. ed., [nachdr.] ed., Addison Wesley, San Francisco Munich, 2008 (eng).

[Le13] J. M. Lee, *Introduction to smooth manifolds*, 2nd ed ed., Graduate texts in mathematics, no. 218, Springer, New York ; London, 2013, OCLC: ocn800646950.

[Si01] A. C. d. Silva, *Lectures on symplectic geometry*, Lecture notes in mathematics, no. 1764, Springer, Berlin ; New York, 2001.

[Sp10] M. Spivak, *Physics for mathematicians: mechanics I*, Publish or Perish, Houston, Tex., 2010 (eng).