Gradient flows and a generalized Wasserstein distance in the space of Cartesian currents

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation vorgelegt von

Malte Kampschulte, M.Sc.

aus Duisburg

Berichter: Univ.-Prof. Dr. Christof Melcher
Univ.-Prof. Dr. Michael Westdickenberg
Prof. Robert L. Jerrard, PhD (University of Toronto)

Tag der mündlichen Prüfung: 2. November 2017

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.
Contents

1 Introduction 1
  1.1 Structure of this work ....................... 5
  1.2 Notation .................................. 6
  1.3 Acknowledgments ........................... 7

I A generalized Wasserstein distance 8

2 The homogeneous flat metric 9
  2.1 The homogeneous flat metric .................. 9
  2.2 Theorems adapted to the boundary-less case .... 13

3 Curves of currents and the Lie-derivative 16
  3.1 The Lie-derivative of currents ................ 16
  3.2 Lie derivatives and the normal part ............ 18
  3.3 Curves of currents ........................... 22

4 Introduction to Cartesian currents 26
  4.1 Bubbling and discontinuities .................. 26
  4.2 Graphs and vertical parts ..................... 29
  4.3 Tangents and derivatives ..................... 34
  4.4 Energies and polyconvexity ................... 37
  4.5 Degree and topology .......................... 45

5 Length and Wasserstein-distance 48
  5.1 Length and generalized Wasserstein distance .... 48
  5.2 A vertical Wasserstein distance ................ 52

6 The trace of a curve 58
  6.1 The spatial and the space-time trace ............ 58
  6.2 Length and area of the trace ................... 64
  6.3 Rectifiability ............................... 67
CONTENTS

7 An equivalence theorem 72
  7.1 Flat metric equals 1-distance 72
  7.2 Generalizations, corollaries and remarks 78

8 Comparing vertical Wasserstein with $L^p$ 79
  8.1 The vector space structure in relation to $L^p$ 79
  8.2 Upper bounds 81
  8.3 Lower bounds 83

9 Limits of curves, lower semicontinuity 90
  9.1 Nonexistence of minimizers 90
  9.2 Limits of curves 92
  9.3 Lower semicontinuity 97
  9.4 Relaxed Energy 98

10 Including the boundary 102
  10.1 Wasserstein distance with boundary 102
  10.2 Correspondence with the Flat metric 107

II Gradient Flows in spaces of currents 109

11 Minimizing movements 110
  11.1 Gradient flows 110
  11.2 Abstract gradient flows 111
  11.3 Minimizing movements 113

12 The case of currents 115
  12.1 The minimizing movement iteration for currents 115
  12.2 Space-time solution 117
  12.3 The convergence theorem for currents 120
  12.4 The convergence theorem in the Cartesian case 123

13 Gradient flows in $W_2$ 126
  13.1 Minimizing movements for $W_2$ 126
  13.2 Possible applications 129

Appendices 133

A Differential geometry 134

B Geometric measure theory 140

Main index 145
Chapter 1

Introduction

The main aim of this thesis is to explore a possible approach to gradient flows for Cartesian (and other) currents. We will do so using the method of minimizing movements. The key ingredient for this turns out to be the right choice of metric. For this we propose a generalization of the Wasserstein $p$-distance from measures to currents. We show that in the case of Cartesian currents this distance turns out to be the correct geometric replacement for $L^p$. We also can prove that in the case $p = 1$, this distance coincides with the Flat metric of currents.

Let us start by stating our initial motivation. When considering nonlinear evolution equations, there is often a problem with singularities. While in most cases it is not hard to prove short time existence, showing existence of a solution for arbitrary long time intervals is harder. As a matter of fact, there are quite reasonable equations for which there is no long time existence in the classical sense.

A well known example of this is found in the harmonic map heat flow. For this, we deal with maps $u : \mathcal{M} \rightarrow \mathcal{N}$ between differentiable manifolds and consider the corresponding gradient flow for the Dirichlet energy $E(u) := \int_\mathcal{M} |\nabla u|^2 \, dx$, that is the problem

$$\begin{cases}
\partial_t u(t) = -\nabla_{L^2} E(u(t)) & \text{for } t \in (t_0, T) \\
u(t_0) = u_0.
\end{cases}$$

In this representation of the problem, its non-linearity still is a bit hidden, since it is due to the geometry of $\mathcal{N}$. For functions taking their values in $\mathbb{R}^N$, the above equation is linear. However since it is a manifold, $\mathcal{N}$ has no linear structure. This influences the gradient, which will turn out to be equal to the Laplace-Beltrami operator.\(^1\) For example in the case of $\mathcal{M} \subset \mathbb{R}^n$, $\mathcal{N} = \mathbb{S}^N$.

\(^1\)This equation was first studied by Eells and Sampson [ES64]. A complete theory was later developed by Chen and Struwe [Str85, Str88, CS89]. We also refer to the treatment in [Str96, III.6].
CHAPTER 1. INTRODUCTION

treated as a submanifold of \( \mathbb{R}^{N+1} \) we will get
\[
-\nabla E(u(t)) = \Delta u + |\nabla u|^2 u.
\]

There is no long time existence for this equation. In fact it is possible to find solutions of the harmonic map heat flow, which diverge in a finite amount of time (See [CG89, CD90, CDY92]). Since this can happen even if \( \mathcal{N} \) is closed and bounded, what happens is not the classical divergence toward infinity. Instead what occurs is a so called topological singularity. To understand this, we need to think of \( v : \mathcal{M} \to \mathcal{N} \) in terms of differential topology. Let us assume that \( \mathcal{M} \) and \( \mathcal{N} \) are of the same dimension. For any \( U \subset \mathcal{M} \), the image \( v(U) \) covers part of \( \mathcal{N} \). A single point \( y_0 \) in \( \mathcal{N} \) may be the image of several points in \( \mathcal{M} \), so we can count the number of times, \( y_0 \) is covered. In fact around any regular point (that is, a point \( x_0 \in \mathcal{M} \) such that \( Du(x_0) \) is invertible), the map \( v \) forms a local diffeomorphism and thus has an orientation, which we can use as a sign when counting covers. The resulting number of course is the mapping degree \( \text{deg}(v, U, y_0) \).

To see how this relates to the harmonic map heat flow, consider the simplest interesting case, where \( \mathcal{M} \subset \mathbb{R}^2 \) and \( \mathcal{N} = S^2 \). It is possible to find solutions \( u(t) \) where, when \( t \) converges to \( T \), for increasingly smaller discs \( B_{\varepsilon_t}(x_t) \) (with \( \varepsilon_t \to 0 \) and \( x_t \to x_0 \)), the image \( u(t)(B_{\varepsilon_t}(x_t)) \) already covers the whole sphere. But then in any reasonable limit \( u(T) \), this whole cover of the sphere will need to vanish as \( u(T)(x_0) \) can only take a single value. As a result, the mapping degree will change, which already tells us that \( u(t) \) cannot converge in any strong sense. In fact, consider the called topological charge, given by
\[
j_u := u \cdot \partial_1 u \times \partial_2 u,
\]
which can be seen as the density of the topological degree. In the above limit, this charge will concentrate in \( x_0 \) and form a Dirac-delta.

This phenomenon is called “bubbling” and is perhaps the prototypical example of a topological singularity. This is a situation, in which a topological quantity, which is something that should be constant, changes in the limit. Topological singularities are also often inherent to the problem, in the sense that they are not an artifact of the solution process, but an actual feature, not only in the sense that it is possible to construct specific cases, in which they occur, but also in the physical reality from which many problems are derived, as for example evidenced by singularities created due to vortex/anti-vortex collisions in thin magnetic films (See for example [KMMS09]). In fact, as the physical interest in topological quantities is steadily on the rise, from the above mentioned vortices, over Skyrmions [Sky61, Sky62], Hopfions [War99] and other related topological solitons (see also [MS04]) up to the 2016 Nobel price in physics\(^2\), we can only expect the number of examples of topological singularities to increase in the future.

The question now is, how to work with topological singularities instead of working around them. As we have already seen, the notion of a function is not

---

\(^2\)Awarded to David J. Thouless F. Duncan M. Haldane and J. Michael Kosterlitz “for theoretical discoveries of topological phase transitions and topological phases of matter”.

ideally equipped to describe the occurring singularities. A reasonable way to combat this problem then is to relax the notion of a function somewhat.

Ultimately, our difficulties are rooted in the non-linearity of the problem. A popular way to relax non-linearities is to elevate the image to something similar to the domain. The classical way to do so is the concept of Young measures, where we treat $u(x)$ as a probability measure $\nu_x$ on $\mathcal{N}$ instead of a single point. As a result, any nonlinear function $f(u(x))$ then is turned into a linear functional $\int_{\mathcal{N}} f(y) d\nu_x(y)$. Yet young measures are a pointwise way to look at things, something which does not work that well in concert with the topological structure.\(^3\)

However there is a more topological way to relax the concept of functions known under the name of Cartesian currents.\(^4\) The idea behind this is to identify each function $u$ with its graph $\mathcal{G}(u)$ as a submanifold of $\mathcal{M} \times \mathcal{N}$ and to treat this manifold as an integer rectifiable current. In a similar way to the Young measures, this deals with any problems arising due to nonlinearities, but as an added bonus, convergence in the sense of currents will always preserve topological invariants. This will be the setting which we want to have in mind.

Having thus set the stage, we can describe the contents of this thesis. The problem with Cartesian currents until now is that, while there are a lot of results most of them are in a way static. That is, since they are motivated by the calculus of variations, they all relate to convergence of sequences and to minimizers. The main aim of this thesis is to try to extend the theory of Cartesian currents towards evolution problems, or more specifically gradient flows, building on the fact that energies in general are well understood for Cartesian currents.

To achieve this aim, we would like to employ the technique known as “minimizing movements”\(^5\), where a time-discrete solution to the gradient flow is constructed via iterated minimization problems of the form

$$u_{k+1} := \arg \min_u \frac{1}{2h} d(u_k, u)^2 + E(u).$$

This is something which works quite well, not only since minimization problems are well understood for Cartesian currents, but also since they have good compactness properties which will be helpful in finding a limit of the time-discrete solutions. Yet before we can get there, we need to settle on the correct choice of a metric.

---

\(^3\)It bears noting however, that the original ideas of Young [You37, You42a, You42b] are much more geometric in nature and that his work is much closer to the concept of varifolds. The modern concept of a Young measure seems to be more rooted in the works of Ball, Murat and Tartar, see [Tar79, Bal89].

\(^4\)See [GMS89a, GMS89c, GMS89b] as well as the two volumes [GMS98].

\(^5\)In a way this is just an interpretation of the backwards Euler-scheme for metric spaces, which has been formalized in [dG93, Amb95]. For a more throughout discussion we refer to the book [AGS08].
The most common choice of metric in gradient flows of course is $L^2$. So while the flat metric is an obvious choice of metric for integer rectifiable currents, it is not hard to see that as an “area between curves”, it will not behave like $L^2$. What is needed is a metric, which is defined on Cartesian currents, yet which will behave familiar in $L^2$ distance in situations where it is defined, such as for graphs of functions.\(^6\)

The key idea to finding such a metric interestingly enough comes from a different direction, namely the notion of Wasserstein-distance in optimal transport, or to be more specific a certain way to formulate the Wasserstein-distance, which goes back to Benamou and Brenier [BB00]. Take two probability measures $\rho_0$ and $\rho_1$ defined on $\mathbb{R}^n$. We can think of $\rho_0$ being transported into $\rho_1$ along some path of probability measures $t \mapsto \rho(t)$ with $\rho(0) = \rho_0$ and $\rho(1) = \rho_1$. If this path is reasonably regular, then its derivative can also be represented as the measure being transported along a vector field $v$, that is via the conservation of mass formula
\[
\partial_t \rho(t) + \nabla \cdot (v(t)\rho(t)) = 0.
\]

It is then possible to assign a length to the curve $\rho$ via the integral $L_2(\rho) := \int_0^1 \int |v|^2 \, d\rho(t) \, dt$. The 2-Wasserstein-distance then can be defined as the square root of the infimum of this cost, taken over all possible paths $\rho$.\(^7\)

This can be generalized to currents. However in order to do so, we need to identify what it means for a current to be transported along a vector field. The fundamental insight here is that while transport of measures needs to preserve mass, transport of a current instead should preserve its multiplicity. This leads to defining a Lie-derivative for currents via
\[
\mathcal{L}_v S\langle \omega \rangle = S\langle \mathcal{L}_{-v} \omega \rangle
\]
where $\mathcal{L}_v \omega$ is the Lie-derivative of forms. This is a quite natural object, which extends the previous concept of conservation of mass to a conservation of multiplicity\(^8\), in the form of
\[
\partial_t S + \mathcal{L}_v S = 0.
\]

---

\(^6\)We note however that this new distance does not have to completely coincide with $L^2$ whenever both are defined, as only local behaviour of the metrics influences the gradient flow. Furthermore there are some situations where, as illustrated by the phenomenon of bubbling, two functions are close in $L^2$ eventhough they should be separate due to topology. In fact for our distance this will no longer be the case.

\(^7\)The best way to see, how this relates to $L^2$ is to consider the case of point masses $\rho_0 = \delta_a$ and $\rho_1 = \delta_b$. In this case the optimal way of transport $\rho$ is a constant movement of the mass from $a$ to $b$ along the straight line in between. Then $v$ will be equal to $b - a$ and thus $L_v = |b - a|^2$.

\(^8\)Note that, as $\mathcal{L}_v \omega = d_i v \wedge + i_v d\omega$, where $i_v$ is the contraction with $v$, for 0-currents we have $\mathcal{L}_v T\langle \phi \rangle = -T\langle v \cdot \nabla \phi \rangle$ which is nothing more than a weak formulation of $\nabla \cdot (vT)$. 

---
1.1. Structure of this work

Using this, it is then possible to define a generalized Wasserstein-distance for currents in general and by a slight modification in a variant more suitable for Cartesian currents.

The main idea is that a $p$-length for a curve of currents $S : I \rightarrow I_k(\Omega)$ then can similar to before be defined by

$$L_p(S) := \int_I \int_I |v|^p d\|S\| dt$$

where for each time $v$ satisfies (or rather is the smallest vector field satisfying) $\dot{S} + \mathcal{L}_v S = 0$. One can then try to minimize amongst all such curves to define a distance. We then show in Theorem 7.1 that for $p = 1$ this distance corresponds to the flat metric and that a vertical version of this distance mirrors the behaviour of the $L^p$-distance for Cartesian currents (see Theorem 8.10).

The details are of course a bit more involved and one encounters some issues with degeneracy if $p > 1$ (see Example 5.4 and Example 5.11). These will also continue to plague us when discussing lower semicontinuity in Chapter 9.

Finally having identified a suitable metric, the second part of this thesis will focus on applying minimizing movements to currents. Here it turns out that under reasonable compatibility conditions with the flat metric, a solution of a gradient flow can always be understood as a current in space-time itself. In fact, using the natural compactness properties of currents, we can derive a rather general convergence result in the form of Theorem 12.12.

1.1 Structure of this work

This work consists of two separate parts. Both are related, yet essentially self-contained. In the first part, we generalize the Wasserstein distance to currents in general and Cartesian currents as an important special case. We try to show that this establishes a useful metric. The second part then is an discussion on how to apply the method of minimizing movements to currents, in order to show existence results for gradient flows.

Going more into detail, in Part I we start off with Chapter 2, where we offer several remarks on the special case of currents without any boundary, on which we will focus our studies. In Chapter 3 we will explore the notion of curves of currents and most importantly interpret their motion as a variant of the Lie derivative, which will allow us to associate a corresponding vector field. Chapter 4 will then form a short interlude, in which we try to give a rather brief overview of the theory of Cartesian currents, which acts as an important special case for the following chapters. In Chapter 5, we will use curves to define an analogue of the Wasserstein distance for currents. There, we will also discuss how to adapt this idea to Cartesian currents in the form of a vertical Wasserstein distance.
In Chapter 6, we will then define the trace of a curve and explore some of its more useful properties, for example rectifiability. This also gives us a tool that we will use in several of the following chapters. In Chapter 7 we compare the general Wasserstein distance to the more traditional flat metric and specifically show that they are equivalent for \( p = 1 \). The theme of comparison then continues into Chapter 8, where we similarly compare the vertical Wasserstein distance with the \( L^p \) distance between functions and contrast the situations in which they differ or coincide. As the penultimate chapter of the first part, Chapter 9 then discusses the existence (or lack thereof) of minimizing curves and addresses the open question of lower-semicontinuity, which is needed to satisfy the prerequisites of the theorems in the second part. To finally Chapter 10 will be giving a short overview on how the results of the previous chapters can be extended to currents with boundaries.

In Part II we will then focus our attention on the concept of minimizing movements. After giving a short overview of the basic ideas and the existing theory in Chapter 11, we then consider what happens in the special case of currents and how the general theory can be improved in this special case in Chapter 12. We will then try to tie this part into the first in Chapter 13, where we will discuss how to improve the minimizing movements iteration in order to make the best use out of the distance we defined in the part one.

1.2 Notation

It is said that “Differential geometry is the study of properties that are invariant under change of notation”\(^9\). As such there are several competing notations for nearly everything, both in the concepts from differential geometry we apply, as well as in the study of currents. The author has tried to choose wisely from the options available. To list them all here would however exceed the size of this section. For a full list of notations, definitions and the specific versions of some theorems used, the reader is referred to the appendices.

For now let us just remark some general conventions. If not specified otherwise, we use \( \Omega \) to denote either a domain in \( \mathbb{R}^n \) or a oriented and bounded Riemannian manifold, as this often does not make a difference. In general we do assume that manifolds are isometrically embedded, although we try to work intrinsically whenever possible.

When dealing with Cartesian currents, we are especially interested in Cartesian products of the form \( \Omega \times \mathcal{M} \) where \( \Omega \subset \mathbb{R}^n \) is a bounded domain and \( \mathcal{M} \subset \mathbb{R}^N \) is a manifold (or sometimes \( \Omega \times \mathbb{R}^N \)). Of course then \( \Omega \times \mathcal{M} \) is a manifold as well and so this case is also covered. In such Cartesian products we denote the coordinates by \( (x, y) = (x_1, ..., x_n, y_1, ..., y_N) \) in order to distinguish between

\(^9\)An unsourced quote in the introduction of [Lee12].
the coordinates of both factors. We will also use the shorthand notation $dx := dx^1 \wedge ... \wedge dx^n$ and $dy := dy^1 \wedge ... \wedge dy^N$.

In a similar vein, we will sometimes consider currents in space time as well. In this case we always consider time as the zeroth component and denote the coordinates by $(t, x_1, ..., x_n)$ or $(t, x_1, ..., x_n, y_1, ..., y_N)$ respectively.

For the sake of clarity we adopt the convention of enclosing linear arguments in angle brackets. For example the differential $k$-form $\omega$ at point $x$ evaluated against the vectors $v_i$ would be denoted by

$$\omega(x)(v_1, ..., v_k),$$

while something like evaluating the $i$-th entry in a sequence of curves of currents at time $t$ with the form $\omega$ may look like

$$S_i(t)(\omega).$$

1.3 Acknowledgments

The author would like to thank all of those who supported him. First and foremost this includes all the family and friends who either showed their support directly, or indirectly, even if only by putting up with my general state in the end of writing this. An equal amount of thanks is due to all the colleagues which provided ideas and insights in long discussions and offered help whenever I was stuck. Last but not least I have to thank my advisor for allowing me to pick such an interesting speculative topic and for all the help, the mathematical education and of course him and the RWTH Aachen for providing funding for all of those years, especially in the form of the Dean’s seed fund.
Part I

A generalized Wasserstein distance
Chapter 2

Boundary-less currents and the homogeneous flat metric

When looking at later applications, especially at Cartesian currents, one of the more notable features is that we will only deal with currents which have no boundary. While most of the theory developed in the next few chapters can be adapted to include the case of currents with a boundary, it seems prudent to only consider the boundary-less currents for now. This will in many cases turn out to be much more tractable, as there is often the need for additional terms in order to accommodate the effect of the boundary.

As the boundary however plays a role in many classic definitions and results, it is reasonable to adapt some of them to the boundary-less currents, as this will remove the need for additional considerations in many cases.

2.1 The homogeneous flat metric

Instead of the full flat metric

\[ \mathcal{F}(T) := \sup\{ M(S) + M(R) \mid T = \partial S + R, S \in I_{k+1}(\Omega), R \in I_k(\Omega) \} \]

(see Definition B.14), we will deal with a simpler variant, where we will remove the boundary term. Central to that idea will be the following modification of the flat metric.

**Definition 2.1.** Let \( T \in I_k(\Omega) \) be an integral rectifiable current. We will then call

\[ \tilde{\mathcal{F}}(T) := \inf\{ M(S) \mid S \in I_{k+1}(\Omega), \partial S = T \}. \]

the **homogeneous flat metric**.\(^1\)

\(^1\)A note about naming. The word homogeneous here is used in the sense of dropping
It is not hard to see that $\dot{\mathcal{F}}(T)$ can only be finite if $\partial T = 0$. Thus it seems prudent to introduce a notation for the space of boundaryless currents:

**Definition 2.2.** The space of boundaryless integer rectifiable currents is given by
\[
\dot{\mathcal{I}}_k(\Omega) := \{ T \in \mathcal{I}_k(\Omega) \mid \partial T = 0 \}.
\]

Before we discuss some of the properties of this metric, it seems prudent to remind the reader about some peculiarities of the boundary of currents in spaces such as $\mathcal{D}_k(\Omega)$.

**Remark 2.3.** The definition of a boundary can be seen as a generalization of Stokes’ theorem. If $T$ corresponds to integration along a manifold, then $\partial T$ will as a direct result correspond to integration along its boundary.

In general, the boundary of a current should always be understood relatively to the domain $\Omega$. As we are only allowed to test against forms $\omega$ with compact support in $\Omega$, neither $\omega$ nor $d\omega$ can reach to the topological boundary $\partial \Omega$. To give an example, for an arbitrary $\Omega \subset \mathbb{R}^n$, we have $\partial \llbracket \Omega \rrbracket = 0$ when we treat $\llbracket \Omega \rrbracket$ as a current in $\mathcal{D}_n(\Omega)$. On the other hand, if we treat $\llbracket \Omega \rrbracket$ as a current in the whole space, that is in $\mathcal{D}_n(\mathbb{R}^n)$, we end up with
\[
\partial \llbracket \Omega \rrbracket = \llbracket \partial \Omega \rrbracket,
\]
assuming that $\partial \Omega$ is sufficiently regular.

At this point, one should also note, that while in general it is not clear, if it is possible to treat a current $T \in \mathcal{D}_k(\Omega)$ as a current in $\mathcal{D}_k(\mathbb{R}^n)$, there is no such problem if $T$ is representable by integration and of finite mass. In this case we can still use the integral
\[
T(\omega) = \int_{\Omega} \langle \omega(x), \xi(x) \rangle \, d\mu \quad \forall \omega \in C^\infty(\Omega; \wedge^k)
\]
for $\omega \in C^\infty(\mathbb{R}^n; \wedge^k)$. This case of course also includes the integer rectifiable currents in which we are interested.

In general when nothing else is mentioned, we will always refer to the boundary with respect to $\Omega$.

The definition of $\dot{\mathcal{F}}$ also works if there is no such current $S$, for example if $\partial T \neq 0$. In this case $\dot{\mathcal{F}}(T) = \infty$. As a consequence, we immediately get a variant of a well known result for $\mathcal{F}$.

---

The term of lower order, so in this case of lower dimension. Its meaning may become more apparent when looking at the flat norm and dualities (Remark 2.12), as well as in conjunction with the similarities to the $W^{-1,1}$-norm (Remark 2.8), which is also the source of the dot.
Lemma 2.4. Let $\tilde{T} \in I_k(\Omega)$ be a fixed current and define
\[ I := \left\{ T \in I_k(\Omega) \mid \tilde{\mathcal{F}}(T - \tilde{T}) < \infty \right\}. \]
Then $d(T_0, T_1) := \tilde{\mathcal{F}}(T_1 - T_0)$ is a metric on $I$.

Proof. The proof is essentially the same as the proof for the flat metric. \qed

![Illustrations of (a) flat metric and (b),(c),(d) homogeneous flat metric. Note that in (c) $T_0$ and $T_1$ have the same non-vanishing boundary and in (d) both currents have a boundary in the whole space but not in $\Omega$. Both situations are admissible for the homogeneous flat metric.](image.png)

Remark 2.5. A few things are worth noticing when considering the metric spaces associated with the homogeneous flat metric:

1. It can be shown that the metric space $\mathcal{I}$ in Lemma 2.4 is connected. This will be a consequence of Theorem 7.1. Furthermore, if some bounds on the mass are imposed it is also compact.

2. The class $\mathcal{I}$ can also be defined as $\{ T \in I_k(\Omega) \mid \partial T = R \}$ for some fixed $R \in I_{k-1}(\Omega)$. However then there are problems if the corresponding homology-class of $\Omega$ is not trivial. Consider $0$ and $[\Omega]$ as $n$-currents who are both without boundary in $\Omega$, but $\tilde{\mathcal{F}}(0 - [\Omega]) = \infty$.

3. Another approach here would be to stretch the definition of metric to also allow $\infty$ as a value. Then $d(\ldots)$ can be defined on all of $I_k(\Omega)$ and we get $d(T_0, T_1) < \infty$ if $T_0$ and $T_1$ are in the same connected component. \qed

While using the homogeneous flat metric simplifies most situations by removing all the additional terms related to the boundary, it does not reduce everything to a toy problem. In fact in many cases, when one is not really interested in dealing with boundary terms, it is easier to work with this definition than...
to take the full flat metric and then show nonexistence of boundary terms. In
general both are closely related, as the next lemma shows.

**Lemma 2.6.** Let \( T \in I_k(\mathbb{R}^n) \). Then the following relations hold

1. \( \mathcal{F}(T) \leq \hat{\mathcal{F}}(T) \)
2. For \( k < n \) we have \( \hat{\mathcal{F}}(T) < \infty \) if and only if \( \partial T = 0 \).
3. If \( \hat{\mathcal{F}}(T) < \infty \) then the infimum in the definition is obtained.
4. There exists \( C = C(n,k) \) such that \( \hat{\mathcal{F}}(T) = \mathcal{F}(T) \) for all \( T \in \hat{I}_k(\mathbb{R}^n) \) such that \( \mathcal{F}(T) < C \)

**Proof.**

1. This is a direct consequence of the definition, as the flat metric is
given by the infimum over a strictly larger class than the homogeneous flat
metric.

2. If \( \hat{\mathcal{F}}(T) < \infty \) then there exists \( S \in I_{k+1}(\mathbb{R}^n) \) such that \( T = \partial S \). However
then \( \partial T = \partial \partial S = 0 \). If \( \partial T = 0 \) then by the isoperimetric inequality B.18
there exists \( S \in I_{k+1}(\mathbb{R}^n) \) such that \( \partial S = T \) and \( M(S) \leq \gamma M(T)^{(k+1)/k} < \infty \).

3. Existence of a minimizer is a simple application of the direct method. Let
\( (S_k)_{k \in \mathbb{N}} \) be a minimizing sequence, that is \( M(S_k) \to \hat{\mathcal{F}}(T) \) and \( \partial S_k = T \)
for all \( k \in \mathbb{N} \). Then by the compactness closure theorem B.17 there is a
sub-sequence (not relabeled) and a current \( S \in I_{k+1}(\mathbb{R}^n) \) such that \( S_k \to S \).
Now since \( M \) is lower-semicontinuous, we have \( M(S) \leq \lim \inf_{k \to \infty} M(S_k) = \hat{\mathcal{F}}(T) \) and by continuity of the boundary-operator \( \partial S = T \). Therefore \( S \) is
the desired minimizer.

4. We know that for \( \mathcal{F}(T) \) the minimum is also obtained. So let \( S \in I_{k+1}(\mathbb{R}^n) \),
\( R \in \hat{I}_k(\mathbb{R}^n) \) with \( T = \partial S + R \) with \( M(S) + M(R) = \mathcal{F}(T) \). We assume that
\( R \neq 0 \). By the isoperimetric inequality there is a current \( S' \in I_{k+1}(\mathbb{R}^n) \)
such that \( \partial S' = R \) and \( M(S') \leq \gamma M(R)^{(k+1)/k} \) where \( \gamma \) only depends on
\( k \) and \( n \). But then

\[
\mathcal{F}(T) \leq \hat{\mathcal{F}}(T) \leq M(S) + M(S') \leq M(S) + \gamma M(R)^{(k+1)/k} \\
< M(S) + M(R) = \mathcal{F}(T)
\]

if \( M(R) < C(\gamma,n) = C(n,k) \), which certainly is the case for \( \mathcal{F}(T) < C(n,k) \).
Therefore \( R = 0 \) and thus \( \mathcal{F}(T) = \hat{\mathcal{F}}(T) \). \( \square \)

**Remark 2.7.** Note that hidden in this lemma, there are three different kinds
of topological aspects. Firstly, the equivalence in the second condition is directly
related to the fact that on \( \mathbb{R}^n \) all the relevant homology-groups are trivial. If
we would deal with currents defined on some manifold, then we would still
2.2. **THEOREMS ADAPTED TO THE BOUNDARY-LESS CASE**

have $\partial^2 = 0$ so the only if direction will still hold. However for the other direction, $\partial T = 0$ does not imply existence of any $S$ such that $\partial S = T$. The simplest counterexample here is $T$ being a circle winding around a torus.

The second topological aspect is seen when considering the space $\dot{I}_k(\Omega)$, or rather the subspace

$$\left\{ T \in \dot{I}_k(\Omega) \bigg| \dot{F}(T) < \infty \right\}.$$ 

Not only do $\dot{F}$ and $F$ both define metrics on this space, also, as seen in the last lemma, both metrics do not only define the same topology but are in fact equal for small distances. However for larger distances note that by the trivial choice $S = 0, R = T$ in $T = \partial S + R$, we have the upper bound

$$F(T) \leq M(T).$$

while $\dot{F}(T)$ can be arbitrarily large in comparison to $M(T)$, as can be seen by the scaling of the isoperimetric inequality.

The third aspect is found in the final point. The proof will work not only in $\mathbb{R}^n$, but on any $\Omega$, where we have a version of the isoperimetric inequality. The usual reason not to have such an inequality are of course boundaryless currents which are in itself not the boundary of another currents. Or in other words a choice of $\Omega$ with non-trivial homology. However, since we only consider currents $R$ with $M(R) \leq C$, we only need an isoperimetric inequality for currents with small mass, which still holds for many nontrivial $\Omega$.

**Remark 2.8.** In the case of $k = 0$, there is another way of understanding the homogeneous flat metric. An integer rectifiable 0-current is nothing more than a finite sum of integer-weighted Dirac-distributions. Now for two points $a$ and $b \in \Omega$ it is easy to derive that

$$\dot{F}(\delta_a - \delta_b) = \min \{|a - b|, \text{dist}(a, \partial \Omega) + \text{dist}(b, \partial \Omega)\} = \|\delta_a - \delta_b\|_{W^{1,1}}.$$  

Iterating on the arguments behind this, one can similarly prove that in fact $\dot{F}(T) = \|T\|_{W^{1,1}}$ for all $T \in I_0(\Omega)$.  

2.2 **Theorems adapted to the boundary-less case**

We will now conclude this section by stating the boundary-less versions of some of the central theorems of geometric measure theory. While they are direct corollaries and do not require an independent proof, it is still helpful to consider them separately.

Maybe the most fundamental technical tool in the study of currents is the deformation theorem. We state this theorem in its full form in the appendix as Theorem B.19. If the current that is to be deformed, has no boundary, many of the conclusions of the theorem however simplify quite a bit resulting in the following:
Corollary 2.9 (Deformation theorem, boundary-less case). Let $T \in \mathcal{D}_k(\mathbb{R}^n)$ be a normal current with $\partial T = 0$ and $\varepsilon > 0$. Then there exist $P \in \mathcal{D}_k(\mathbb{R}^n)$, $S \in \mathcal{D}_{k+1}(\mathbb{R}^n)$ such that for $C = 2n^{2k+2}$ the following holds:

1. $T = P + \partial S$ and $\partial P = 0$
2. If $k \neq 0$ then $\mathcal{M}(P)/\varepsilon^k \leq C\mathcal{M}(T)/\varepsilon^k$, $\mathcal{M}(S)/\varepsilon^{k+1} \leq C\mathcal{M}(T)/\varepsilon^k$
3. If $k = 0$ then $\mathcal{M}(P) \leq \mathcal{M}(T)$, $\mathcal{M}(S)/\varepsilon \leq \mathcal{M}(T)$
4. $\text{supp } P \cup \text{supp } S \subset B_{2n\varepsilon}(\text{supp } T)$
5. $\text{supp } P \subset W_k$, where $W_k$ is a regular $k$-skeleton of size $\varepsilon$. In other words, $P$ is a polyhedral chain.
6. If $T \in \dot{I}_k(\mathbb{R}^n)$, then $P \in \dot{I}_k(\mathbb{R}^n)$, $S \in I_{k+1}(\mathbb{R}^n)$.
7. If $T$ is a polyhedral chain, then so is $S$.

Remark 2.10. Since we are interested in compact manifolds $\mathcal{M}$ as well as the Cartesian product $\Omega \times \mathcal{M}$ for $\Omega \subset \mathbb{R}^n$, it should be noted, that the deformation theorem also holds in this setting, if $\varepsilon < \varepsilon_0(\mathcal{M})$ is small enough. The only additional change is in the constant $C$ which will in general depend on the dimensions as well as $\mathcal{M}$.

An easy proof of this assertion for $\mathcal{M}$ can be done by embedding $\mathcal{M}$ in $\mathbb{R}^l$. As $\mathcal{M}$ is assumed to be compact, this can be done such a way that there is a retraction map $\pi : U \rightarrow \mathcal{M}$ for some $\varepsilon_0$-neighbourhood $U$ of $\mathcal{M}$ in $\mathbb{R}^l$ with $\pi|_{\mathcal{M}} = id$. Now let $T \in \mathcal{D}_k(\mathcal{M})$ a current on $\mathcal{M}$. Then if $\varepsilon < \varepsilon_0$, we can deform $T$ in $\mathbb{R}^l$ using the deformation theorem. The resulting currents $P$ and $S$ will then still be in $U$ and can thus be projected back onto $\mathcal{M}$ and their projection will fulfill similar conditions, with the masses only changing by a factor depending on the Jacobian subdeterminants of $\pi$, which can be bounded.

Proving the theorem in this way, the $k$-skeleton $W_k$ in the proof will turn into the projection of a $k$-skeleton of $\mathbb{R}^l$ under $\pi$. However it is maybe more insightful to instead start with a triangulation of $\mathcal{M}$ by simplices. While the proof of the deformation theorem is commonly done using cubes, it easily adapts to simplices and so it is also possible to prove the assertion by repeating the original proof mutatis mutandis on each simplex of the triangulation. The constant $C$ in the resulting theorem will then depend on the quality of the triangulation, that is how far the simplices deviate from being isometric embeddings of a regular simplex. It is known that for $\varepsilon < \varepsilon_0(\mathcal{M})$, we can always find a triangulation of cell size smaller than $\varepsilon$ such that this deviation and thus $C$ is bounded uniformly.\(^2\)

The second, equally important theorem is the Federer-Fleming compactness closure theorem (see Theorem B.17). This too adapts into a similar form.\(^2\)

\(^2\)For details see [Whi40, Cai36, Cai34] as well as the book [Whi57].
Corollary 2.11 (Compactness-closure, boundary-less case). Let the sequence $(T_i)_{i \in \mathbb{N}} \subset \dot{I}_k(\Omega)$ be such that $M(T_i)$ is uniformly bounded. Then there exists a current $T \in \dot{I}_k(\Omega)$ and a sub-sequence $(T_i)_{i \in \mathbb{N}}$ (not relabeled) such that

$$T_i \to T$$

in $\mathcal{F}$.

In other words, for all $C > 0$ the set $\{T \in I_k(\Omega) \mid M(T) \leq C, \partial T = 0\}$ is sequentially compact.

The important observation here is that by Lemma 2.6 we can, in sufficiently nice cases, assume the convergence in $\dot{\mathcal{F}}$ instead of $\mathcal{F}$.

Finally let us remark a bit on dualities

Remark 2.12. If one drops the rectifiability in the definition of the flat metric, one ends up with

$$\mathbf{F}(T) := \min \{ M(S) + M(Q) \mid T = \partial S + Q, S \in D_{k+1}(\Omega), Q \in D_k(\Omega) \}$$

which is also often called flat norm. This similar definition works on a wider class of currents but retains a lot of similar properties.

This flat norm has a dual formulation [Fed69, 4.1.12]:

$$\mathbf{F}(T) = \sup \left\{ T(\omega) \biggm| \omega \in C^\infty(\Omega; \wedge^k), \|\omega(x)\| \leq 1, \|d\omega(x)\| \leq 1 \text{ for all } x \in \Omega \right\}$$

In a similar way, we can define a homogeneous version, which also has a dual formulation:

$$\hat{\mathbf{F}}(T) := \min \{ M(S) \mid T = \partial S, S \in D_{k+1}(\Omega) \}$$

$$= \sup \left\{ T(\omega) \biggm| \omega \in C^\infty(\Omega; \wedge^k) \text{ with } \|d\omega(x)\| \leq 1 \text{ for all } x \in \Omega \right\}$$

In all this note that while for $k = n$ (and $k = 0$) it is a direct consequence of the constancy theorem (or a simple calculation respectively) that $\hat{\mathbf{F}}(T) = \hat{\mathcal{F}}(T)$ for all $T \in I_k(\Omega)$, for co-dimension and dimension of two or higher\(^3\) there are counterexamples (see [You63]) which show that in general the two notions are not equal. In for any $C > 0$ there exists a current $T \in I_k(\Omega)$ with $\hat{\mathcal{F}}(T) > C\mathbf{F}(T)$ (see [Mor84, Whi84]).

\(^3\)The cases $k = 1$ and $k = n - 1$ are still open.
Chapter 3

Curves of currents and the Lie-derivative

The main goal of our approach is to define the distance between two currents as a transportation problem. Since the conserved quantity in our situation will not be the mass but the multiplicity, we cannot take the conventional approach of tallying which part of $T_0$ is moved to where in $T_1$. Instead we need to consider some sort of path along which $T_0$ is transported, or to be more precise a curve in the space of currents linking $T_0$ to $T_1$. Giving the right definitions to do so is the main aim of this chapter. However, before we go into what constitutes sufficiently smooth curves of currents, we need to elaborate on what constitutes an infinitesimal transport of a current.

3.1 The Lie-derivative of currents

Probably the most basic yet natural way to represent the movement of a geometric quantity is the Lie-derivative, which tells us what happens if we move something in the direction of a vector field. This idea is easily extended to currents by making use of the duality. Moving the current into the direction of $v$ should in first order correspond to moving any test-form into the opposite direction $-v$. This leads to the following:

Definition 3.1. Let $T \in I_k(\Omega)$ be a current. Then we formally denote the Lie derivative of a current by

$$(L_v T)(\omega) := T(L_{-v} \omega) = -T(L_v \omega) = -T(i_v d\omega + d i_v \omega)$$

where we assume $\omega$ and $v$ to be sufficiently regular. The fact that $L_v \omega = i_v d\omega + d i_v \omega$ is known as Cartan’s formula (see Lemma A.20).

Remark 3.2. Before we continue, we need to clarify what is meant by sufficient regularity of $v$. Since we are working with integer rectifiable currents,
which are per definition representable by integration, we are not restricted to testing with smooth forms. On the contrary, to ensure well definedness we only need to guarantee that $\mathcal{L}_v \omega \in L^1(\mathcal{H}^k \text{ supp } T; \wedge^k)$, which can be enforced by rather weak conditions on $v$, for example $v \in W^{1,1}(\mathcal{H}^k \text{ supp } T)$, where some care needs to be taken since $d$ also involves derivatives tangential to $\text{ supp } T$, which need to be worked around.

Even better however is the following observation: We can further reformulate $T \langle i_v d\omega + di_v \omega \rangle = T \langle i_v d\omega \rangle + \partial T \langle i_v \omega \rangle$.

Then since we are only interested in currents which are either boundary-less (so $\partial T = 0$) or have a stationary boundary (which will result in $v|_{\text{ supp } \partial T} = 0$), the second term vanishes and we can make sense of $\mathcal{L}_v T$ even if we only have $v \in L^1(\mathcal{H}^k \text{ supp } \Omega)$. ■

However, for the purpose of simplifying some of the proofs we will restrict ourselves to a stricter space.

**Definition 3.3.** Let $S \in I_k(\Omega)$. We will write $v \in C^1_{pw}(\text{ supp } S; \mathbb{R}^n)$ if there is a family of continuously differentiable manifolds $(\mathcal{M}_i)_{i \in \mathbb{N}}$ with regular boundaries such that $S = \sum_{i=1}^{\infty} \alpha_i \|\mathcal{M}_i\|$ for some integer weights $(\alpha_i)_{i \in \mathbb{N}}$ and if $v|_{\mathcal{M}_i} \in C^1(\mathcal{M}_i; \mathbb{R}^n)$ for all $i \in \mathbb{N}$ and $v|_{\mathcal{M}_i}$ can be extended continuously onto the boundary of $\mathcal{M}_i$.

The main idea here is that we can do most of our calculations separately on each manifold, where $v$ is continuously differentiable and thus can be treated classically. Now let us show that this definition indeed corresponds to moving our current along a vector field in the way of a Lie-derivative.

**Lemma 3.4.** Let $T \in I_k(\Omega)$, $v \in C^1(\Omega; \mathbb{R}^n)$ a continuously differentiable vector field. Let $\phi_t(x)$ be the flow corresponding to $v$, that is for each $x \in \Omega$, $t \mapsto \phi_t(x)$ solves $\frac{\partial}{\partial t} \phi_t(x) = v(\phi_t(x)), \phi_0(x) = x$. Then

$$\mathcal{L}_v T \langle \omega \rangle = \lim_{t \to 0} \left( \frac{\phi_t \# T - T}{t} \right) \langle \omega \rangle$$

for all $\omega \in C^\infty(\Omega; \wedge^k)$.

**Proof.** By the definition of the pushforward (See Definition B.7) and linearity, we can write

$$\left( \frac{\phi_t \# T - T}{t} \right) \langle \omega \rangle = \phi_t \# T \langle \omega \rangle - T \langle \omega \rangle = T \left( \phi_t \# \omega - \omega \right) = T \left( \frac{\phi_t \# \omega - \omega}{t} \right)$$
Then since $T$ is a normal current, it is continuous under uniform convergence, so we can take the limit in the argument and find that
\[
\lim_{t \to 0} \left( \frac{\phi_t \# T - T}{t} \right) \langle \omega \rangle = T \left( \lim_{t \to 0} \frac{\phi_t \# \omega - \omega}{t} \right) = T(\mathcal{L}_v \omega) = \mathcal{L}_v T(\omega). \quad \square
\]

Finally, the following formula will be useful in several calculations:

**Proposition 3.5.** Let $f : \Omega \to \mathbb{R}^N$ be a Lipschitz map, $T \in I_k(\Omega)$ and $v \in C^1_{\text{pw}}(\text{supp} T; \mathbb{R}^k)$. Then
\[
f_\# \left( \mathcal{L}_v T \right) = \mathcal{L}_{f_\# v} f_\# T.
\]

**Proof.** Let $\omega \in C^\infty(\mathbb{R}^N; \wedge^k)$. First assume that $f$ is of class $C^2$. Then
\[
f_\# \left( \mathcal{L}_v T \right) \langle \omega \rangle = -T \inner{ \mathcal{L}_v f_\# \omega } = -T \inner{ f_\# \mathcal{L}_v \omega }.
\]

Thus it is enough to show
\[
\mathcal{L}_v f_\# \omega = f_\# \mathcal{L}_{f_\# v} \omega
\]
$\mathcal{H}^k$-almost everywhere on $\text{supp} T$. Splitting $T$ into parts, we can then restrict ourselves to the case where $T$ is a $C^1$-manifold on which $v$ is continuously differentiable. In this situation the identity is the result of a simple calculation.

Finally, let us consider the case where $f$ is only Lipschitz. Then according to Definition B.7 we approximate $f$ uniformly by a sequence of smooth functions with bounded derivatives $(f_i)_{i \in \mathbb{N}}$. Then
\[
f_\# \left( \mathcal{L}_v T \right) = \lim_{i \to \infty} (f_i)_\# \left( \mathcal{L}_v T \right) = \lim_{i \to \infty} \mathcal{L}_{(f_i)_\# v} (f_i)_\# T. \quad \square
\]

### 3.2 Lie-derivatives and the normal part of a vector field

It is a straightforward observation that pushing a current in a tangential direction does not produce a change. Consider as a simple example the current $T = \left[ S^2 \right]$ and a family of diffeomorphisms $\phi_t : x \mapsto Q_t x$ where $Q_t$ is a smooth enough family of orthogonal matrices with $Q_0 = I$. Then $(\phi_t)_\# T = T$ for all $t$ and thus $\mathcal{L}_{\frac{\partial}{\partial t}} \phi_t T = 0$. It is also easy to see that $\frac{\partial}{\partial t} \phi_t(x)$ is a tangential vector to $S^2$ in the point $x$. 

3.2. LIE DERIVATIVES AND THE NORMAL PART

Now since the Lie derivative corresponds to the transport in direction of \( v \), it should be an easy consequence that then \( \mathcal{L}_v T = 0 \) if \( v \) is tangential to \( T \). In fact apart from some difficulties involving the boundary of \( T \) this is indeed true. It is then possible to use the linearity of the Lie-derivative to conclude that in fact we only ever need to consider the part of \( v \) which points out of \( T \) in a normal direction.

**Lemma 3.6.** Let \( T \in I_k(\Omega) \) and \( v \in C^1_{pw}(\text{supp} T; \mathbb{R}^n) \). We define \( v^\perp(x) \) as the part of \( v(x) \) that is orthogonal to the tangential space of \( \text{supp} T \) at \( x \). Then \( v^\perp: \text{supp} T \to \mathbb{R}^n \) is \( H^k \) \( L \)-supp \( T \)-a.e. well defined and \( v^\perp \in C^1_{pw}(\text{supp} T; \mathbb{R}^n) \). Furthermore \( \mathcal{L}_{v^\perp} T \) is well defined and

\[
\mathcal{L}_v T(\omega) = \mathcal{L}_{v^\perp} T(\omega) - \partial T(\langle i_{v^\perp} \omega \rangle)
\]

holds for all \( \omega \in C^\infty(\Omega; \wedge^k) \).

\[\begin{array}{c}
\approx h \cdot v(t) \\
\approx h \cdot (\pi h)_\# T
\end{array}\]

Figure 3.1: For small deformations moving by \( v \) and \( v^\perp \) is equivalent in the interior.

**Proof.** First of all let us note, that \( T \) can be decomposed into a family of smooth manifolds \( (\mathcal{M}_i)_{i \in \mathbb{N}} \) such that \( T = \sum_{i \in \mathbb{N}} \alpha_i [\mathcal{M}_i] \) (see Definition B.10) and we can choose those manifolds in such a way that they do not intersect. Additionally, due to rectifiability \( \text{supp} T \setminus \bigcup_{i \in \mathbb{N}} \mathcal{M}_i \) is a \( H^k \) zero-set. Therefore \( v^\perp \) is defined \( H^k \) \( L \) \( \text{supp} T \)-almost everywhere.

Furthermore the projection maps \( \pi_i: \mathcal{M}_i \times \mathbb{R}^n \to \mathbb{R}^n; (x, w) \mapsto w^\perp \), that project \( w \) onto its part orthogonal to the tangential space at \( x \), are smooth with \( |w^\perp| \leq |w| \). We can also assume that \( v \) is continuously differentiable when restricted to each of the \( \mathcal{M}_i \). But since \( v^\perp(x) = \pi_i(x, v(x)) \) for \( H^k \)-a.e., \( v^\perp \) is continuously differentiable when restricted to \( \mathcal{M}_i \) and thus \( v^\perp \in C^1_{pw}(\text{supp} T; \mathbb{R}^n) \). Furthermore for any \( \omega \in C^\infty(\Omega; \wedge^k) \), the form \( \mathcal{L}_{v^\perp} \omega = \text{di}_{v^\perp} \omega + i_{v^\perp} d\omega \in C^0(\Omega; \wedge^k) \) is well defined and bounded \( H^k \)-almost everywhere.

So \( T(\mathcal{L}_{v^\perp} \omega) < \infty \) and thus \( \mathcal{L}_{v^\perp} T \) is a well defined current.

For the final statement we need to show that \( T(\mathcal{L}_{v} \omega) = T(\mathcal{L}_{v^\perp} \omega) - \partial T(\langle i_{v^\perp} \omega \rangle) \) for an arbitrary \( \omega \in C^\infty(\Omega; \wedge^k) \). Fix \( x \in \text{supp} T \) such that the tangential \( k \)-vector \( \xi(x) \) exists. Now we use the linearity of the objects involved:

\[
\langle \xi, \mathcal{L}_{v} \omega \rangle = \langle \xi, i_v d\omega + \text{di}_v \omega \rangle = \langle \xi \wedge v, d\omega \rangle + \langle \xi, \text{di}_v \omega \rangle + \langle \xi, i_{v^\perp} \omega \rangle
\]
where by anti-symmetry we have $\xi \wedge v = \xi \wedge v^\perp$ and thus
\[
\xi \wedge v = \langle \xi, v^\perp \rangle + \langle \xi, \partial_t v^\perp \omega \rangle = \langle \xi, L_{v^\perp} \omega + \partial_t v^\perp \omega \rangle.
\]
Finally integrating over the measure associated with $T$ we get
\[
L_v T \langle \omega \rangle = -T \langle L_v \omega \rangle = -T \langle L_{v} \omega + \partial_t v^\perp \omega \rangle = L_{v^\perp} T \langle \omega \rangle - \partial T \langle i_{v^\perp} \omega \rangle.
\]
As a direct consequence of this we note that in our case $\partial T = 0$, so the second term vanishes. But then since $|v^\perp| \leq |v|$ and we will in general be interested in finding
\[
\inf \left\{ \int |v|^p d\|S(t)\| \left| \dot{S}(t) + L_v S(t) = 0 \right| \right\},
\]
for some curve of currents $S$, we can always choose $v = v^\perp$. In fact, we can state the following minimization result:

**Lemma 3.7.** Let $T \in \dot{I}_k(\Omega)$ and $D \in D_k(\Omega)$. Define
\[
\mathcal{A} = \left\{ v : \text{supp} \, T \to \mathbb{R}^n \mid v \in C^1_{pw}(\text{supp} \, T; \mathbb{R}^n) \text{ and } D + L_v T = 0 \right\}
\]
Then if $\mathcal{A} \neq \emptyset$ we have
\[
\inf_{v \in \mathcal{A}} \int |v|^p d\|T\| = \int |v^\perp|^p d\|T\|
\]
where $v^\perp$ is the unique minimizer in $\mathcal{A}$, constructed from any $v \in \mathcal{A}$ according to Lemma 3.6. Furthermore, $\mathcal{A} - v^\perp$ is the space of tangential $C^1_{pw}$-vector fields.

**Proof.** We will show that $\mathcal{A}$ is an affine vector space. The statement then is a direct consequence of Lemma 3.6 and the fact that $|v^\perp| \leq |v|$.

For this let $V$ be the space of all tangential vector fields to supp $T$. We want to show that $\mathcal{A} = v + V$ for some $v \in \mathcal{A}$. For this let $v, w \in \mathcal{A}$. We know that $L_v T = L_w T$, so by linearity $L_{v-w} T = 0$ and by Lemma 3.6 $L_{(v-w)^\perp} T = 0$. We have $v - w \in V$ if and only if $(v - w)^\perp = 0$. Thus we need only to show that
\[
L_v T = 0 \iff v^\perp = 0 \text{ a.e.}
\]
or in other words if and only if $v \in V$.

The reverse direction is a direct consequence of Lemma 3.6. For the other implication assume that there is $v \notin V$ such that $L_v T = 0$. Then we have $v^\perp \neq 0$ and $L_{v^\perp} T = 0$. Pick $x_0 \in \text{supp} \, T$ such that such that supp $T$ has a tangential space and $v^\perp \neq 0$ on a small set of positive measure around $x_0$. 

3.2. LIE DERIVATIVES AND THE NORMAL PART

Let us assume for now that \( x_0 = 0 \) and \( T = \|[U \times \{0\}]\|, U \times B_\varepsilon(0) \subset \mathbb{R}^n \) where \( U \subset \mathbb{R}^k \) open, \( 0 \in U \). Let us write as a convention \((x, y) = (x_1, \ldots, x_k, y_1, \ldots, y_{n-k}) \in \mathbb{R}^k \times \mathbb{R}^{n-k}\). Then we can treat \( v = v^+(x, y) \) as a function

\[
v : U \equiv U \times \{0\} \to \mathbb{R}^{n-k} \equiv \{0\} \times \mathbb{R}^{n-k} \subset \mathbb{R}^n, x \mapsto v(x).
\]

Let \( \phi \in C_\infty^\infty(U \times B_\varepsilon(0); [0, 1]) \) such that \( \phi|_{B_\varepsilon(0)} = 1 \) for some small \( \delta > 0 \). Now define

\[
\omega(x, y) := \phi(x, y) \langle v(x), y \rangle \, dx^1 \wedge \ldots \wedge dx^k.
\]

Then we can calculate (using summation convention):

\[
(L_\omega v)(x, 0) = (d i_v \omega + i_v d \omega)(x, 0)
\]

Here the first term vanishes since \( v(x) \perp x_i \) for all \( i \in \{1, \ldots, k\} \) implies \( i_v dx^1 \wedge \ldots \wedge dx^k = 0 \).

\[
\begin{align*}
= & \left. i_v(x) d(\phi(x, y) \langle v(x), y \rangle) \right|_{y = 0} \wedge dx^1 \wedge \ldots \wedge dx^k \\
= & i_v(x) (\langle v(x), 0 \rangle d\phi(x) + \phi(x, 0) dv) \wedge dx^1 \wedge \ldots \wedge dx^k \\
= & i_v(x) (\phi(x, 0) v_j(x) dy^j) \wedge dx^1 \wedge \ldots \wedge dx^k \\
= & \phi(x, 0) v_j(x) (i_v(x) dy^j) \wedge dx^1 \wedge \ldots \wedge dx^k = \phi(x) v(x)^2 dx^1 \wedge \ldots \wedge dx^k.
\end{align*}
\]

But then \( T(L_\omega v) = \int \phi(x, 0) v^2(x) d\|T\| > 0 \) which is a contradiction.

Now let \( f : U \times B_\varepsilon(0) \to \bar{\Omega} \subset \Omega \) be a \( C^1 \)-diffeomorphism such that \( T \llcorner \bar{U} = \alpha f_\# \|[U \times \{0\}]\| \) for some \( \alpha \in \mathbb{Z} \setminus \{0\} \). Define \( \omega \) as above but using \( (f^{-1})_\# v \) instead of \( v \). Then \( f^\# \omega \) can be extended by 0 onto the whole of \( \Omega \) and we have, using Proposition 3.5

\[
T \left( L_{\omega} f^\# \omega \right) = (T \llcorner \bar{\Omega}) \left( f^\# L_{(f^{-1})_\# v} \omega \right)
= (f_\# T \llcorner \bar{\Omega}) \left( L_{(f^{-1})_\# v} \omega \right) = \alpha \|[U \times \{0\}]\| \left( L_{(f^{-1})_\# v} \omega \right)
\]

which by the argument above is only finite if \( (f^{-1})_\# v \) is zero \( \mathcal{H}^k \)-almost everywhere, which since \( f \) is a diffeomorphism is only the case if \( v \) vanishes around \( x_0 \), which is a contradiction.

As it will turn out, this norm of a vector field will form the crucial ingredient of our generalization of the Wasserstein distance. Yet before we develop the details, we will need to consider curves of currents a bit more.
3.3 Curves of currents

When considering the distance between two points \( a, b \), it is often useful to have some means of interpolating between them efficiently, that is having a curve \( s(t) \) such that

\[
d(a, s(t)) = td(a, b) \quad \text{and} \quad d(s(t), b) = (1 - t)d(a, b).
\]

In a vector space, where the distance is derived from a norm, this is easily done by a linear interpolation

\[
s(t) := (1 - t)a + tb.
\]

In a manifold on the other hand, linear interpolation generally not possible. Moreover, if our distance is inherited from an ambient norm, there will in most cases be no valid way to interpolate between two points as desired.

In a way this is our situation. The space of currents is a linear space, where linear interpolation is possible. We are however specifically interested in integer rectifiable currents where this does not work. Consider

\[
S : t \mapsto (1 - t) \cdot \{\{0\} \times [0, 1]\} + t \cdot \{\{1\} \times [0, 1]\}
\]

which is the linear interpolation between \( \{\{0\} \times [0, 1]\} \) and \( \{\{1\} \times [0, 1]\} \) in \( D_k(\mathbb{R}^2) \). Then \( S(t) \notin I_k(\mathbb{R}^2) \) for all \( t \in (0, 1) \). The kind of interpolation we need would look more like

\[
S : t \mapsto \{\{t\} \times [0, 1]\}
\]

which is in \( I_k(\mathbb{R}^2) \) for all \( t \). Note that the same idea plays an important role in optimal transport, where the linear interpolation also needs to be avoided in favor of “displacement interpolation”, which is somewhat similar to what we are doing (See for example [Vil09, Chapter 7]).

As a general idea, the integer rectifiable currents can be seen as a “submanifold” of the linear space of currents.\(^1\) The general approach when dealing with distances on a manifold is to consider the length of minimal geodesics instead. We will define a notion of what length of a curve means in the next chapter, but in order to do so, as a preparation, we need to clarify the notion of curves first.

**Definition 3.8.** Let \( I \subset \mathbb{R} \) be an interval, \( S : I \to I_k(\Omega) \). We say that \( S \) is a continuous curve, or \( S \in C^0(I; I_k(\Omega)) \), if \( M(S) \) is locally bounded and \( S \) is continuous in the topology of currents. We define \( C^0(I; A) \) similarly for subsets \( A \subset I_k(\Omega) \).

\(^1\)One needs to be a bit cautious not to take this idea to far, as things get strange rather quickly. For example if \( S(t) \) is a curve with \( S(0) = T_0 \), then \( \tilde{S}(t) := S(t) - T_0 + T_1 \) is a curve with \( \tilde{S}(t) = T_1 \) but the same formal derivative. Thus the “tangential space” of \( I_k(\Omega) \) is in fact independent of the base point. Yet instead of \( I_k(\Omega) \) being equal to this tangential space, their only common element is 0.
3.3. CURVES OF CURRENTS

Remark 3.9. In the above definition we could also have used continuity in the \( \mathcal{F} \) metric instead, since due to the mass bound, this is the same notion of convergence by Lemma B.15. Furthermore if \( \partial S(t) = 0 \) for all \( t \), it is possible to use continuity in the \( \mathcal{F} \) metric instead, by applying Lemma 2.6.

In the end, using the previous definition however emphasizes that \( S \) is a continuous curve if and only if \( \mathcal{M}(S) \) is bounded and

\[
 t \mapsto S(t)(\omega)
\]

is continuous for all \( \omega \in C_0^\infty(\Omega; \wedge^k) \).

As a next step, let us discuss the concept of time derivatives a bit more.

Remark 3.10. In the subsequent discussion, we will consider differentiable curves in the space of integer rectifiable currents. It is rather easy to define a continuous curve as we did above. We are however interested in differentiable curves. For this we will need some considerations.

First of all note that in order to write down a difference quotient, we actually need a vector space. This of course is no obstacle, as the metric space \( I_k(\Omega) \) is for example trivially embedded in the space of normal currents. However this is not sufficient. In fact for \( S \) to be differentiable in the normal currents, the derivative \( \dot{S} \) needs to be a normal current and thus representable by integration, that is of the form

\[
 \dot{S}(t)(\omega) = \int_{\Omega} \langle \xi(x), \omega(x) \rangle \, d\mu(x)
\]
for some $k$-vector field $\xi$ and a measure $\mu$. However, as we have seen, even something as simple as moving a current in the direction of a vector field results in a Lie-derivative, which instead is of the form

$$ \dot{S}(t)(\omega) = \int_{\Omega} \langle \xi(x), d\omega(x) \rangle \, d\mu(x) $$

which cannot be represented by integration anymore, due to the derivative.

The solution here is to treat $I_k(\Omega)$ as a subset of $D_k(\Omega)$. While $D_k(\Omega)$ does not have a metric, it is still a topological space. So it is possible to define a derivative just by convergence of the difference quotient. For more details see for example [Bou04]. Thus we can define the space of continuously differentiable curves of integer rectifiable currents as $C^1([0, 1]; I_k(\Omega)) \subset C^1([0, 1]; D_k(\Omega))$.

Furthermore let us remark that the weak topology allows us to make sense of the differentiability in a “point-wise” fashion. Let $S \in C^1([0, 1]; I_k(\Omega))$ and $\omega \in C^\infty(\omega; \Lambda^k)$. Then for any $t \in [0, 1]$, the difference quotient $1/h(S(t + h) - S(t))$ converges to $\dot{S}(t)$. But then per definition of convergence of currents we also have

$$ \lim_{h \to 0} \frac{S(t + h) - S(t)}{h}(\omega) = \dot{S}(t)(\omega). $$

In other words, $S \in C^1([0, 1]; I_k(\Omega))$ implies $S(\cdot)(\omega) \in C^1([0, 1]; \mathbb{R})$ for all $k$-forms $\omega \in C^\infty(\Omega; \Lambda^k)$.

Finally let us define $C^1_{pw}(\Omega)$ in the usual way as the space of continuous functions $S \in C^0([0, 1]; I_k(\Omega))$ whose restrictions $S|_{[a_i, a_{i+1}]}$ are continuously differentiable for some partition of the interval $0 = a_0 < \ldots < a_l = 1$. Note that this requires left and right differentiability at each $a_i$. We will use this space in order to simplify constructions of actual curves. All the proofs can also be made to work using $C^1$ by “slowing down” around the kinks $a_i$ in order to guarantee $\dot{S}(a_i) = 0$.

Let us close this section with an example:

**Example 3.11.** Let $\Omega \subset \mathbb{R}^2$ be open, bounded. Consider the vectorfield $v : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x, 1)$ and the family of integer rectifiable 1-currents $S(t)$ given by integration along the Manifold $M_t = [0, e^t] \times \{t\}$. We want to show that change of $S(t)$ is exactly given Lie derivative in direction $v$. For this let $t_0 \in \mathbb{R}$ and $\omega = fx + gdy$ a 1-form.

First of all, let us have a look at $\frac{\partial S(t_0)}{\partial t}$. Here we have

$$ \lim_{h \to 0} \frac{S(t_0 + h)(\omega) - S(t_0)(\omega)}{h} = \lim_{h \to 0} \frac{1}{h} \left( \int_{M_{t_0+h}} (fx + gdy) - \int_{M_{t_0}} (fx + gdy) \right) $$
3.3. CURVES OF CURRENTS

\[
\begin{aligned}
&= \lim_{h \to 0} \frac{1}{h} \left( \int_{M_{t_0} + h} f \, dx + \int_{M_{t_0}} g \, dy - \int_{M_{t_0}} f \, dx - \int_{M_{t_0}} g \, dy \right) \\
&= \lim_{h \to 0} \frac{1}{h} \left( \int_{M_{t_0} + h} f \, dx + 0 - \int_{M_{t_0}} f \, dx + 0 \right) \\
&= \lim_{h \to 0} \frac{1}{h} \left( \int_0^{e_{t_0}} f(s, t_0 + h) \, ds + \int_0^{e_{t_0} + h} f(s, t_0 + h) \, ds - \int_0^{e_{t_0}} f(s, t_0) \, ds \right) \\
&= \int_0^{e_{t_0}} \lim_{h \to 0} \frac{f(s, t_0 + h) - f(s, t_0)}{h} \, ds + \lim_{h \to 0} \frac{1}{h} \int_0^{e_{t_0} + h} f(s, t_0 + h) \, ds \\
&= \int_0^{e_{t_0}} \frac{\partial f}{\partial y}(s, t_0) \, ds + \lim_{h \to 0} \frac{e_{t_0} + h - e_{t_0}}{h} f(\xi, t_0 + h) \\
&= \int_{M_{t_0}} \frac{\partial f}{\partial y} \, dx + e_{t_0} f(e_{t_0}, t_0)
\end{aligned}
\]

where \( t_0 \leq \xi_h \leq t_0 + h \) is from the mean value theorem and the limits can be interchanged due to the smoothness of every function involved. Furthermore, note that we could get rid of all terms involving \( dy \) since we always integrated in the \( x \)-direction.

On the other hand let us now look at \( \mathcal{L}_v S(t_0) \). Here

\[
\mathcal{L}_v S(t_0) \langle \omega \rangle = - \int_{M_{t_0}} i_v d\omega + di_v \omega \\
= - \int_{M_{t_0}} i_v \left( \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) \, dx \wedge dy + d(xf + g) \\
= - \int_{M_{t_0}} \left( \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) (-dx + xdy) - \int_{\partial M_{t_0}} (xf + g) \\
= \int_{M_{t_0}} \frac{\partial g}{\partial x} \, dx - \int_{M_{t_0}} \frac{\partial f}{\partial y} \, dx - \int_{\partial M_{t_0}} xf - \int_{M_{t_0}} \frac{\partial g}{\partial x} \, dx \\
= - \int_{M_{t_0}} \frac{\partial f}{\partial y} \, dx - e_{t_0} f(e_{t_0}, t_0)
\]

So in fact \( \frac{\partial S(t_0)}{\partial t} \langle \omega \rangle + \mathcal{L}_v S(t_0) \langle \omega \rangle = 0 \). ■
Chapter 4

A short introduction to Cartesian currents

The purpose of this chapter is to give a very brief introduction to the concept of Cartesian currents. As this is only a short overview, for further information, we refer the reader to the original papers [GMS89a, GMS89b, GMS89c] and the subsequent book [GMS98] by Giaquinta, Modica and Soucek.

4.1 Bubbling and discontinuities

A common feature of many problems in the calculus of variations, as well as in related fields, is the occurrence of discontinuities, either as an integral feature of the problem, or more often as the result of applying some sort of limiting procedure. A common example of the former is the study of fractures in elasticity, where instead of a continuous deformation of a reference configuration, a certain set of discontinuities is allowed, albeit penalized by some sort of "tearing energy". A somewhat canonical example of the latter is the bubbling phenomenon in harmonic maps. Here we try to minimize the Dirichlet-energy in a fixed topological class, for example determined by a fixed mapping degree. While it is easy to find a minimizing sequence, converging in a sense for which the energy is lower semicontinuous, this is not enough to preserve the topology. As this shall prove a fundamental motivation we will elaborate a bit on this:

Example 4.1 (Bubbling). Let \( \Omega = B_1(0) \subset \mathbb{R}^2 \). Let \( N = (0,0,1) \in S^2 \) be the “north pole”. We consider maps \( u : \Omega \to \mathbb{S}^2 \subset \mathbb{R}^3 \) such that \( u(x) = N \) for \( x \in \partial \Omega \).

Assuming sufficient regularity, we can define the mapping degree on such maps using

\[
\text{deg}(u) := \frac{1}{4\pi} \int_{\Omega} u \cdot (\partial_1 u \times \partial_2 u) dx.
\]
4.1. BUBBLING AND DISCONTINUITIES

This can also be understood as the integral of the pullback of the volume-form of $S^2$, since
\[ \frac{1}{4\pi} u \cdot (\partial_1 u \times \partial_2 u) dx_1 \wedge dx_2 = u^# \omega_{S^2} \]
where $\omega_{S^2}$ is the canonical volume form of the sphere (with mass normalized to $1$). It can also be understood as $\frac{1}{4\pi}$ times the signed area of $S^2$ covered by $u$. In fact another interpretation of $\text{deg} \, u$ is taking the pushforward of the current $[\Omega]$ by $u$ and testing it with the volume-form.

The important point in this is that $\text{deg} \, u$ will always have integral values. A proof can be found in any text on differential topology.\(^1\) It is also easy to see that under a sufficiently strong sense of convergence for the functions $u$, the map $\text{deg}$ is continuous and thus will stay constant.

Thus it seems reasonable to consider the problem of minimizing the Dirichlet-energy
\[ E(u) := \frac{1}{2} \int_{\Omega} |\nabla u|^2 \, dx \]
under the constraint that $\text{deg} \, u = 1$. Using Young’s, as well as Cauchy’s inequality, we have
\[ \frac{1}{2} |\nabla u|^2 = \frac{1}{2} \left( |\partial_1 u|^2 + |\partial_2 u|^2 \right) \geq |\partial_1 u| |\partial_2 u| = |u| |\partial_1 u| |\partial_2 u| \geq u \cdot \partial_1 u \times \partial_2 u \]
and integrating over $\Omega$ we then get the lower bound
\[ E(u) \geq \int_{\Omega} u \cdot \partial_1 u \times \partial_2 u \, dx = 4\pi \text{deg} \, u = 4\pi. \]

Next we can try to obtain functions whose energy is close to the lower bound. A careful look at the above derivation reveals the conditions under which equality holds. For Young’s inequality, this is the case when $|\partial_1 u| = |\partial_2 u|$. On the other hand, Cauchy’s inequality results in equality if $\partial_1 u$ is perpendicular to $\partial_2 u$, with correct orientation. Taken together, this tells us, that we have equality if and only if $u$ is a conformal, orientation preserving map. The canonical example of such a map is the stereographic map,
\[ v : \mathbb{R}^2 \to S^2; \, x \mapsto \left( \frac{x}{1 + |x|^2}, \frac{|x|^2 - 1}{1 + |x|^2} \right) \]

\(^1\)Or indeed gleaned from the fact that $T := u_{#} [\Omega]$ has to be integer rectifiable and, since $u|_{\partial \Omega}$ is constant, $\partial T = 0$. But then by the constancy theorem B.13, we have $T = k [S^2]$ for some $k \in \mathbb{Z}$ and $\text{deg} \, u = k$.\]
the inverse of the stereographic projection, which indeed has an energy of $4\pi$, when considered on the whole of $\mathbb{R}^2$. However, its restriction to $\Omega$ does not observe the boundary-conditions, as the value $N$ is only taken at infinity, so we will have to try to modify it a bit.

The crucial ingredient for this is a specific property of the Dirichlet-energy in two dimensions, the scaling invariance. Take $\lambda > 0$ and define $v_\lambda(x) := v(\lambda x)$. Then

$$
\frac{1}{2} \int_{\mathbb{R}^2} |\nabla v_\lambda(x)|^2 \, dx = \frac{1}{2} \int_{\mathbb{R}^2} \lambda^2 |\nabla v(\lambda x)|^2 \, dx = \frac{1}{2} \int_{\mathbb{R}^2} |\nabla v|^2 \, dx.
$$

Since we know that $v$ takes values close to $N$ for large $|x|$, we can now use this to rather explicitly construct a minimizing sequence. Fix $\lambda > 0$ large and $\varepsilon > 0$ small. Then we define

$$
u_{\lambda,\varepsilon}(x) := \begin{cases} v_\lambda(x) & \text{for } |x| < 1 - \varepsilon \\ v_\lambda(x\psi_\varepsilon(|x|)) & \text{for } |x| > 1 - \varepsilon \end{cases}
$$

where the monotonous function $\psi_\varepsilon : [1 - \varepsilon, 1) \to \mathbb{R}$ is chosen in such a way that $\psi_\varepsilon(1 - \varepsilon) = 1$ and $\lim_{r \to \infty} \psi_\varepsilon(r) = \infty$. Using this, it is not hard to see, that $u_{\lambda,\varepsilon}$ satisfies the boundary conditions and $\deg u_{\lambda,\varepsilon} = 1$. Furthermore the energy splits into

$$
E(u_{\lambda,\varepsilon}) = \int_{B_{1-\varepsilon}(0)} \frac{1}{2} |\nabla v_\lambda|^2 \, dx + \int_{B_1 \setminus B_{1-\varepsilon}(0)} \frac{1}{2} |\nabla (v_\lambda(x\psi_\varepsilon(|x|)))|^2 \, dx
$$

When sending $\lambda \to \infty$ and $\varepsilon \to 0$, the first integral will converge to the energy of $v$ on the full space which is $4\pi$, while the second integral will vanish.

This shows the optimality of our lower bound. However what about a minimizer? As the Dirichlet-energy is well known to be lower-semicontinuous in the weak $W^{1,2}$ topology, it is reasonable to check the sequence we just constructed for convergence and consider its limit.

As $v$ is close to $N$ for large $|x|$, we can immediately derive uniform convergence of $v_{\lambda,\varepsilon}$ to the constant $N$ on any annulus $B_1(0) \setminus B_\rho(0)$, for $\rho > 0$. Thus we get the pointwise limit $N$ for all $x$ except 0. Furthermore we can derive $L^2$ convergence and by the usual compactness arguments, even convergence in $W^{1,2}$. Yet what has to bother us is that the limit function $N$ obviously has $\deg N = 0$.

As we said in the beginning, one interpretation of the mapping degree is by using the area of the target manifold that is covered using the map. Since we started with a degree of one, we started with a full cover of the sphere, which somehow vanished in the limit. The question here is, where to? As evidenced by the uniform convergence away from 0, it tends to concentrate in
the interior. Indeed, if we consider the definition of the degree, we see that for any $r > 0$

$$\int_{B_r(0)} u_{\lambda,\varepsilon} \cdot \partial_1 u_{\lambda,\varepsilon} \times \partial_2 u_{\lambda,\varepsilon} \, dx \to 4\pi$$

for $\lambda \to \infty, \varepsilon \to 0$, or in other words $u_{\lambda,\varepsilon} \partial_1 u_{\lambda,\varepsilon} \times \partial_2 u_{\lambda,\varepsilon} \, dx \to 4\pi \delta_0$ in the sense of distributions.

So the covering of the sphere concentrates into a single point and thus vanishes when viewed in a classical sense. This phenomenon is called bubbling and is a common problem when studying harmonic maps.

\[
\begin{align*}
\text{(a)} & \quad \text{Figure 4.1: The phenomenon of bubbling, as illustrated in the simple toy model for } u : I \subset \mathbb{R} \to S^1. \text{ Most of the target manifold is covered from an increasingly smaller part of the domain, which shrinks to a point in the limit. The result can be seen as a vertical part of the graph.} \\
\text{(b)} & \quad \text{The rest of this chapter will now be devoted to explaining the Cartesian currents, which are one possible approach to capturing this phenomenon by generalizing the concept of functions.}
\end{align*}
\]

\subsection{4.2 Graphs and vertical parts}

For the rest of the section let $\Omega \subset \mathbb{R}^n$ and $\mathcal{M}$ a $N$-dimensional manifold, for simplicity either compact or $\mathcal{M} = \mathbb{R}^N$. We will usually denote the coordinates of $\Omega$ by $x_1, \ldots, x_n$ and (local) coordinates of $\mathcal{M}$ by $y_1, \ldots, y_N$.

The main idea behind Cartesian currents is somewhat similar to that of Young measures. Think of a function $u : \Omega \to \mathcal{M}$. What is done in both cases in some way is promoting the co-domain $\mathcal{M}$ from a set in which a single value is taken to something which is more domain-like. In the case of Young measures, this is done in a pointwise fashion. Instead of having a value $u(x) \in \mathcal{M}$ at a point $x$, we generalize $u(x)$ to a probability measure $\mu_x$ on $\mathcal{M}$, which represents all the values taken by $u(x)$.\footnote{In this situation, the term “probability measure” is solely used in the technical sense of $\mu_x$ being an unsigned measure of total mass one. In general it should not be understood as the probability of $u$ taking certain values in $x$ but rather as $u$ taking all those values at the same time, each at a certain extent as specified by the measure.} A classical function then is represented by Dirac measures, setting $\mu_x = \delta_{u(x)}$. Different spaces of Young measures then can
be created by taking the set of classical functions interpreted that way and closing it in some reasonable topology. In that way, a Young measure is often said to be generated by a sequence of functions. (See for example [KP91])

In Cartesian currents a more holistic approach is taken. For Young measures, the measures $\mu_x$ may in principle depend arbitrarily on $x$. Any additional coherence conditions are enforced by the generating sequence. Cartesian currents try to solve this problem by looking at the graphs of functions instead and closing those under some suitable convergence of currents. For this we define some notation:

**Definition 4.2.** Let $u : \Omega \to \mathcal{M}$ be reasonably smooth.\(^3\) Then we define the graph of $u$ as the current

$$\mathcal{G}(u) := (id, u)_# [\Omega] \in I_n(\Omega \times \mathcal{M}).$$

Immediately we can observe (without proof) some properties of graphs that will stay invariant under convergence in the sense of currents.

**Lemma 4.3.** Let $u : \Omega \to \mathcal{M}$ be continuously differentiable. Then $\mathcal{G}(u) \in I_n(\Omega \times \mathcal{M})$ has the following properties:

1. $\partial \mathcal{G}(u) = 0$, when treated as a current in $\Omega \times \mathcal{M}$.
2. $\text{supp} \partial \mathcal{G}(u) \subset (\partial \Omega) \times \mathcal{M}$, when treated as a current in $\mathbb{R}^n \times \mathcal{M}$. Here we need to assume that $u$ can be suitably extended onto $\partial \Omega$ and $\partial \Omega$ is of sufficient regularity.
3. $\pi_x # \mathcal{G}(u) = [\Omega]$. Here $\pi_x : \Omega \times \mathcal{M} \to \Omega$ is the orthogonal projection on the first component.
4. The signed measure

$$\phi \mapsto \mathcal{G}(u) \langle \phi dx^1 \wedge ... \wedge dx^n \rangle$$

defined on $\Omega \times \mathcal{M}$ is non-negative.

**Remark 4.4.** These properties will later be used to define Cartesian currents. However the last property may call for some additional explanation. Another way to interpret it, is as a condition on the orientation of the graph. Of course it does not need to be flat, in fact it may oscillate and get arbitrarily close to being vertical. Yet it will never “flip over” and thus reverse its orientation. For a graph this trivially true, yet it will also need to hold for our extension to Cartesian currents.

Additionally, in conjunction with the third condition, this also prevents the graph from having multiple values in a non vertical part. As an example, let

\(^3\)For example $C^1$ or some Sobolev space which guarantees that $\det Du$ is integrable such as $W^{1,n}$.\)
4.2. GRAPHS AND VERTICAL PARTS

Let $u_1$, $u_2$, and $u_3$ be different functions. To have some current such as $T := G(u_1) + G(u_2)$, is obviously not desirable, as this would represent taking the values of two functions at the same time. Indeed, projected down to the first component, we have $\pi_x \# T = 2[\Omega]$. Yet this condition is not enough, as there still is a workaround. Take $T := G(u_1) + G(u_2) - G(u_3)$. Then $\pi_x \# T = [\Omega]$. However in this case the last condition is violated, as $\phi \mapsto -G(u_3)(\phi dx^1 \wedge \ldots \wedge dx^n)$ is a negative measure.

\[ \begin{array}{ll}
  y \in \mathcal{M} & y \in \mathcal{M} \\
  x \in \Omega & x \in \Omega \\
  \end{array} \]

Figure 4.2: Two currents that are clearly not graphs, (a) does not fulfill $\pi_x \# G(u) = [\Omega]$, while (b) does so, yet there $G(u)(\phi dx^1 \wedge \ldots \wedge dx^n)$ is negative in some places.

The following is a way to translate the concept of the $L^p$-norm to currents in a way that coincides with the definition on graphs. Note that while the space of currents is linear, this linearity is not the linearity of functions, as we have already seen in our last example $G(u_1) + G(u_2)$ is not the graph of a function. Thus this $L^p$-norm cannot be used to define an $L^p$-distance between functions. Its main purpose instead is to establish bounds in order to prevent sequences of currents from moving towards infinity. Because of this, the $L^p$-norm is only needed when we are dealing with the case $\mathcal{M} = \mathbb{R}^n$, as similar bounds are trivial if the target $\mathcal{M}$ is a compact manifold.

What we would like to do is formally write

\[ \|T\|_{L^p} := T(\phi dx^1 \wedge \ldots \wedge dx^n). \]

Yet the problem here is that $|y|^p dx^1 \wedge \ldots \wedge dx^n$ does not have compact support and the above expression is thus not well defined. However we can use the following approximation:

**Definition 4.5.** Let $T \in I_n(\Omega \times \mathbb{R}^N)$ such that the signed measure $\phi \mapsto T(\phi dx^1 \wedge \ldots \wedge dx^n)$ is non-negative. Then we define

\[ \|T\|_{L^p} := \lim_{k \to \infty} T(\psi_k(x,y) |y|^p dx^1 \wedge \ldots \wedge dx^n), \]

\[ \]
where the smooth functions $\psi_k : \Omega \times M \to [0,1]$ have compact support and are monotonically increasing in $k$ for each point and $\lim_{k \to \infty} \psi_k(x,y) = 1$ for all $x,y$.

Note that the $L^p$ norm is well defined by the monotone convergence theorem. For this, the non-negative sign of the measure is strictly necessary. Otherwise the exact value of $\|T\|_{L^p}$ will not only depend on the choice of $\psi_k$, but may also be undefined in some cases. As we will only deal with currents, for which $\|T\|_{L^p}$ is well defined, we will sometimes by abuse of notation denote

$$\|T\|_{L^p}^p = T\langle |y|^p \, dx^1 \wedge \ldots \wedge dx^n \rangle$$

without the limit.

We can now define Cartesian currents using these properties.

**Definition 4.6.** Let $T \in I_n(\Omega \times M)$. Then we say that $T$ is a Cartesian current and write $T \in \text{cart}(\Omega; M)$ if all of the following conditions hold true:

1. $\mathcal{M}(T) < \infty$.
2. $\partial T = 0$.
3. $\pi_x # T = \llbracket \Omega \rrbracket$.
4. The measure $\phi \mapsto T\langle \phi \, dx^1 \wedge \ldots \wedge dx^n \rangle$ is non-negative.
5. If $\mathcal{M} = \mathbb{R}^N$, then additionally $\|T\|_{L^1} < \infty$.

**Remark 4.7.** Our definition for manifold-valued Cartesian currents slightly differs from the definition given in [GMS89a]. What they do is to define $\mathbb{R}^L$-valued Cartesian currents similar to how we did and then define $T \in \text{cart}(\Omega; M)$ for embedded manifolds $\mathcal{M} \subset \mathbb{R}^L$ by requiring $T \in \text{cart}(\Omega; \mathbb{R}^L)$ and $\text{supp} T \subset \Omega \times M$. As any smooth manifold can be embedded, the definitions are equivalent and the choice is up to personal preference. In this case, we want to focus on having an intrinsic definition. ■

**Proposition 4.8.** Let $(T_k)_{k \in \mathbb{N}} \subset \text{cart}(\Omega; M)$ and $T_k \to T \in I_n(\Omega \times M)$. Then by the (semi-)continuity of the quantities involved

1. $\mathcal{M}(T) \leq \liminf_{k \to \infty} \mathcal{M}(T_k)$.
2. $\partial T = \lim_{k \to \infty} \partial T_k = 0$.
3. $\pi_x # T = \lim_{k \to \infty} \pi_x # T_k = \llbracket \Omega \rrbracket$.
4. The measure $\phi \mapsto T\langle \phi \, dx^1 \wedge \ldots \wedge dx^n \rangle = \lim_{k \to \infty} T_k\langle \phi \, dx^1 \wedge \ldots \wedge dx^n \rangle$ is non-negative.
5. $\|T\|_{L^p} \leq \liminf_{k \to \infty} \|T_k\|_{L^p}$.

As a result, if $\mathcal{M}(T_k)$ and $\|T_k\|_{L^1}$ are uniformly bounded, then $T \in \text{cart}(\Omega; M)$. 
4.2. GRAPHS AND VERTICAL PARTS

Proof. (i) is well known. (ii) and (iii) are a result of the fact that $\partial$ and the pushforward are continuous. (iv) is easy to verify as the convergence of currents implies a convergence of measures. This leaves (v). For any $\psi : \Omega \times \mathcal{M} \to [0, 1]$ we have from the convergence of currents

$$T(\psi \cdot |y|^p \, dx^1 \wedge ... \wedge dx^n) = \lim_{k \to \infty} T_k(\psi \cdot |y|^p \, dx^1 \wedge ... \wedge dx^n)$$

and as the measure is non-negative and $\psi \leq 1$ we have

$$T_k(\psi \cdot |y|^p \, dx^1 \wedge ... \wedge dx^n) \leq \|T_k\|_{L^p}.$$  

Then this bound holds for all $\psi$ and thus for $\|T\|_{L^p}$.

In an analogous way to the convergence of integer-rectifiable currents, the preceding theorem then motivates the following definition.

**Definition 4.9.** Let $(T_k)_{k \in \mathbb{N}} \subset \text{cart}(\Omega; \mathcal{M})$. We say that $T_k \to T$ in $\text{cart}(\Omega)$ if $T_k \to T$ in the sense of currents and $\mathcal{M}(T_k)$ and $\|T_k\|_{L^p}$ are bounded.

An immediate result of Proposition 4.8 then is the corresponding variant of the compactness closure theorem:

**Corollary 4.10** (Compactness-closure). Let $M_0 > 0, C_0 > 0$. Then the set

$$\{T \in \text{cart}(\Omega; \mathcal{M}) \mid \mathcal{M}(T) \leq M_0, \|T\|_{L^p} \leq C_0\}$$

is sequentially compact in $\text{cart}(\Omega; \mathcal{M})$.

Having defined a way to relax the notion of a function, we can now return to our initial example and observe in what way we converge to a new limit.

**Example 4.11.** We start again with the rescaled inverse of the stereographic projection, $v_\lambda : B_1(0) \to S^2; x \mapsto v(\lambda x)$. This time, we will not bother with the precise way in which the boundary values are obtained. Instead we are more interested in the limit, especially in what happens in the middle. Consider the family of currents $\mathcal{G}(v_\lambda)$. It is not hard to establish a bound on the mass. Since we have a graph, we get using Hölder’s inequality

$$\mathcal{M}(\mathcal{G}(v_\lambda)) = \int_{B_1} \sqrt{1 + |Dv_\lambda|^2} \, dx \leq \left( \int_{B_1} 1 + |Dv_\lambda|^2 \, dx \right)^{1/2} < C < \infty.$$  

As a direct consequence at least for a subsequence, we have convergence against a limit current $T \in \text{cart}(\Omega; S^2)$. In fact it is easy to see that the restriction to points away from 0 converges to the expected pointwise limit, that is for any $\varepsilon > 0$, we have

$$\mathcal{G}(v_\lambda) \upharpoonright \Omega \setminus B_\varepsilon(0) \to \mathcal{G}(N) \upharpoonright \Omega \setminus B_\varepsilon(0)$$
in the sense of \text{cart}(\Omega; \mathbb{S}^2).

However on \( B_{\varepsilon}(0) \), something different happens. We can construct a purely vertical 2-form\(^5\) that does not depend on \( x \) via the pullback of a form on the sphere

\[
\omega := \pi_y^\# (\psi(y)\omega_{\mathbb{S}^2}) \in C^\infty(\Omega \times \mathbb{S}^2; \wedge^2).
\]

In light of our considerations about mapping degree in the beginning of this chapter, testing with this form (where we tacitly ignore the issue of compact support for now) reveals that

\[
\mathcal{G}(v_\lambda) \langle \omega \rangle = \mathcal{G}(v_\lambda) \left( \pi_y^\# (\psi(y)\omega_{\mathbb{S}^2}) \right) = \pi_y^\# \mathcal{G}(v_\lambda) \langle \psi(y)\omega_{\mathbb{S}^2} \rangle \to [\mathbb{S}^2] \langle \psi(y)\omega_{\mathbb{S}^2} \rangle.
\]

So the limit \( T \) should satisfy \( \pi_y^\# T = [\mathbb{S}^2] \). Furthermore, it should do so for any restriction to a small ball \( B_{\varepsilon}(0) \).

In tandem, both conditions only leave a single possible limit:

\[
T := \mathcal{G}(N) + [\{0\} \times \mathbb{S}^2]
\]

Indeed, it is not hard to verify that \( \mathcal{G}(v_\lambda) \to T \) for \( \lambda \to \infty \), using variants of the above considerations.

Now what does this limit signify? As we have seen in the beginning, \( v_\lambda \) tries to cover the whole sphere. In the limit this cover gets completely squeezed into the single point \( 0 \). There it forms a vertical part, which in a way is the central distinguishing feature of Cartesian currents.

In fact this directly motivates the next proposition:

**Proposition 4.12** ([GMS89a, Sec. 2, Thm. 5]). Let \( T \in \text{cart}(\Omega; \mathcal{M}) \). Then \( T \) has a decomposition into a functional and a vertical part, that is

\[
T = \mathcal{G}(u_T) + V_T
\]

where \( u_T : \Omega \to \mathcal{M} \) and

\[
V_T \langle \phi(x, y)dx^1 \wedge ... \wedge dx^n \rangle = 0
\]

for all functions \( \phi \in C^\infty_0(\Omega \times \mathcal{M}; \mathbb{R}) \).

### 4.3 Tangents and derivatives

For now, we have focused on how the concept of a function translates into a current. Yet in nearly any practical problems the derivatives of a function

\(^5\)An \( n \)-covector \( \xi \) will be called vertical if \( \langle \xi, dx^1 \wedge ... \wedge dx^n \rangle = 0 \). Correspondingly, an \( n \)-form \( \omega \) will be called vertical if it is vertical at all points.
are even more important. Even in our starting example, we were concerned with the Dirichlet-energy, which of course is an integral over the square of the derivative.

The usual way to deal with the derivative is to treat it as just another function and to see in which way it converges if necessary in some relaxation. However in Cartesian currents the approach is slightly different. The reason for this is, that the derivative is in a way nicely encoded in the current in the form of the (approximate) tangential space. On the other hand, requiring existence of this tangential space directly results in some conditions on the derivative, which we until now omitted by just requiring that $G(u)$ is integer rectifiable.

First of all consider testing the current $G(u)$ with forms of the type
$$
\omega = \phi(x,y)dx^\alpha \wedge dy^\beta
$$
where $|\alpha| + |\beta| = n$ and where $|\phi(x,y)| \leq 1$. Then
$$
G(u)\langle \omega \rangle = (id,u)_\# [\Omega] \langle \omega \rangle = [\Omega] \langle (id,u)^\# \omega \rangle
$$
$$
= \int_\Omega \phi(x,u(x))dx^\alpha \wedge du^\beta(x)
$$
$$
= \int_\Omega \phi(x,u(x))dx^\alpha \wedge \frac{\partial u^\beta_1}{\partial \gamma_1}dx^{\gamma_1} \wedge ... \wedge \frac{\partial u^\beta_l}{\partial \gamma_l}dx^{\gamma_l}
$$
$$
= \int_\Omega \phi(x,u(x))\sigma(\alpha,\alpha^c)M^\beta_\alpha(Du)dx
$$
where $M^\beta_\alpha(Du)$ is the sub-determinant of the matrix $Du(x)$ for the indices $\alpha^c := \{i \in \{1,\ldots,n\} \mid i \notin \alpha\}$ and $\beta$ and $\sigma(\alpha,\alpha^c)$ is just the sign such that $dx^\alpha \wedge dx^{\alpha^c} = \sigma(\alpha,\alpha^c)dx^1 \wedge ... \wedge dx^n$.

Now the left hand side of this equation is bounded by the mass of $G(u)$ and on the right hand side we can choose $\phi(x,u(x))$ arbitrarily close to $\pm 1$, matching the sign of the other terms. In the limit, this results in
$$
\left\| M^\beta_\alpha(Du) \right\|_{L^1} \leq \mathcal{M}(G(u)).
$$

Thus it makes sense to require a certain integrability of subdeterminants for the functional part of a Cartesian current. Furthermore, in light of the fact that we are often dealing with problems, which are somewhat derived from Sobolev-norms, it is sensible to require a higher integrability. If we consider $u \in W^{1,p}(\Omega; M)$ for $p \geq n$ then the subdeterminants of order $k$ are naturally integrable with exponent $\frac{n}{k}$. The other way round, if we say consider functions $u \in W^{1,2}(\Omega; M)$ where $\Omega \subset \mathbb{R}^3$, then the subdeterminants of order 1 (that is the derivatives themselves) are in $L^2$, those of order 2 are only in $L^1$ and the full determinants of order 3 are only integrable in $L^1$ if we require $\mathcal{M}(G(u)) < \infty$.

To use this integrability in defining a practical class of currents we first need to translate this into a property defined on currents. Consider the derivation
CHAPTER 4. INTRODUCTION TO CARTESIAN CURRENTS

at the beginning of this chapter. If we treat

\[ P(\phi, u) := \int_{\Omega} \phi(x, u(x)) \sigma(\alpha, \alpha^c) M_{\alpha^c}^\beta(Du) \, dx \]

as a sort of dual pairing between \( \phi \) and \( u \), then the \( p \)-norm of \( M_{\alpha^c}^\beta(Du) \) can be given by

\[ \left\| M_{\alpha^c}^\beta(Du) \right\|_{L^p} := \sup\{ P(\phi, u) \mid \| \phi(\cdot, u(\cdot)) \|_{L^q} \leq 1 \} \]

where, as usual, \( \frac{1}{p} + \frac{1}{q} = 1 \). To generalize this, we use the following definition:

**Definition 4.13.** Let \( p = (p_0, \ldots, p_n) \) and let \( q_k \) be the dual exponent for each \( p_k \), that is \( \frac{1}{p_k} + \frac{1}{q_k} = 1 \). Then we define\(^6\)

\[ \left\| \omega \right\|_{M_k^p} := \left( \int_{\Omega} \sup_y \left( \sum_{|\alpha| = n-k, |\beta| = k} |\omega_{\alpha\beta}(x, y)|^2 \right)^{q_k/2} \right)^{1/q_k} \]

where \( \omega = \sum_{|\alpha| + |\beta| = n} \omega_{\alpha\beta} dx^\alpha \wedge dy^\beta \) as usual and

\[ \left\| \omega \right\|_{M_k^p} := \max_{k \in \{0, \ldots, n\}} \left\| \omega \right\|_{M_k^q}. \]

This then yields the dual norm

\[ \left\| T \right\|_{M_k^p} := \sup\{ T(\omega) \mid \left\| \omega \right\|_{M_k^q} \leq 1 \}. \]

By the considerations above it is easy to derive that \( \| G(u) \|_{M_k^p} \) is finite if and only each \( k \)th sub-determinant of \( Du \) is in \( L^p \). This then motivates the definition of the appropriate space of Cartesian currents:

**Definition 4.14.** Let \( p = (p_0, \ldots, p_n) \). Then we define

\[ \text{cart}^p(\Omega; \mathbb{R}^N) := \{ T \in \text{cart}(\Omega; \mathbb{R}^N) \mid \left\| T \right\|_{M_k^p} < \infty \}. \]

**Remark 4.15.** In [GMS89a], Giaquinta, Modica and Soucek also define the classes \( \text{Cart}(\Omega; \mathcal{M}) \) as well as \( \text{Cart}^p(\Omega; \mathcal{M}) \), each time as the closure of the set \( \{ G(u) \mid u \in C^\infty(\Omega; \mathcal{M}) \} \) in \( \text{cart}(\Omega; \mathcal{M}) \) and \( \text{cart}^p(\Omega; \mathcal{M}) \) respectively. In some cases, depending on \( \Omega \) and \( \mathcal{M} \), as well as \( p \), the spaces \( \text{cart}^p(\Omega; \mathcal{M}) \) and \( \text{Cart}^p(\Omega; \mathcal{M}) \) coincide or differ. This of course is related to the question if smooth maps are dense in \( W^{1,p}(\Omega; \mathcal{M}) \), something which is trivially true if

\(^6\)Again, for the sake of readability we deviate a bit from the notation used in [GMS89a], where \( \left\| \omega \right\|_{M_k^q} \) is denoted by \( \left\| \omega \right\|_{M_k^{p_k}} \) instead. Since we will only work with derived results, this should not pose a problem.
4.4. ENERGIES AND POLYCONVEXITY

$p > n$, but for example fails for the hedgehog-singularity $x \mapsto \frac{x}{|x|}$ in $\mathbb{R}^3$, which is in $W^{1,2}(B_1; S^2)$ yet cannot be smoothly approximated (See [SU82, Bet90]). For a quite general overview of the interplay of continuity and Sobolev spaces we also refer to a series of papers by Hang and Lin [HL01, HL03a, HL03b, HL05].

Also of note is the class of diffeomorphisms, where the Cartesian current $T$ has an inverse $T^{-1}$, given by swapping the role of $x$ and $y$, which is a Cartesian current in itself. These currents are of interest for the study of cavitation in elasticity. However, we will not go further in this direction.

Finally, in order to reduces the amount of indices, we define the appropriate space for our example-problem:

**Definition 4.16.** Now let us consider maps from $\Omega \subset \mathbb{R}^2$ to $S^2$. We then define

$$\text{cart}^{2,1}(\Omega; S^2) := \{ T \in \text{cart}^{2,2,1}(\Omega; \mathbb{R}^3) | \text{supp } T \subset \Omega \times S^2 \}.$$

**Example 4.17.** Now, taking a look back at our example, a finite Dirichlet energy implies square-integrability of the derivative. So it makes sense to consider a class of similar Cartesian currents, in this case the class $\text{cart}^{2,1}(\Omega; S^2)$. We will find out, how to properly translate the energy in the next section. For now, we note that for those currents the structure theorem can be slightly improved (see [GMS98, II.4.1.2 Thm 1]). Assume that $T \in \text{cart}^{2,1}(\Omega; S^2)$ and of the form $T = \mathcal{G}(u_T) + V_T$ where $V_T$ is the vertical part. Then from the integrability we get $u_T \in W^{1,2}(\Omega; S^2)$. Furthermore, while we are in the critical exponent, where Sobolev functions are not necessarily continuous, it can be shown that the restriction to values in $S^2$ results in $\partial \mathcal{G}(u_T) = 0$. As a consequence the vertical part $V_T$ has the form

$$V_T = \sum_{i \in I} d_i \{ a_i \} \times S^2$$

where $I$ is a finite (due to the bounded mass) set of indices, $d_i \in \mathbb{Z}$ and $a_i \in \Omega$. This necessarily includes our purported limit $\mathcal{G}(N) + \{ \{0\} \times S^2 \}$.

### 4.4 Energies and polyconvexity

Now that we know how the derivatives of the function relate to the corresponding Cartesian current, we can have another look at the Dirichlet energy. The question now is how to properly extend this energy to Cartesian currents. As we have seen, any Cartesian current has a functional part (Proposition 4.12). So in this sense the Dirichlet energy is well defined for Cartesian currents. However we still have some freedom in the choice of how to treat the vertical part, as we may include it in some way in the energy. As it turns out, we need
to exercise this choice in order to keep our energy lower-semicontinuous. The key to this is a classical concept of the calculus of variations, the notion of polyconvexity. (See [Bal77])

Many of the energies in the calculus of variations can be written as an integral of the form

$$E(u) = \int_{\Omega} f(Du(x))dx.$$ 

Usually such energies have some notion of coercivity, which in this case means that a bound on the energy implies a bound on the $W^{1,p}$-norm. As a consequence, what we usually can expect for sequences of bounded energy is only weak convergence in $W^{1,p}$. In order to obtain lower semicontinuity in this topology, continuity of the integrand is not enough, as weak convergence does not imply pointwise convergence. For a weakly convergent sequence, $Du_k(x)$ may instead oscillate between different values with the limit being somewhere in between. Thus the classical condition for lower semicontinuity of $E$ is convexity of $f$.

For simpler energies this is enough, yet there are many common energies not covered. For many of those energies, especially those rooted in some geometric consideration we instead need to rely on the notion of polyconvexity. Assuming that we have coercivity for a large enough $p$, we do not only get weak convergence of $Du$ but of all the subdeterminants of $Du$ as well. (See also [GMS89c, Müll88, Müll89]) So if we can find a convex function $g$ with

$$f(Du) = g(Du, M(Du)),$$

where $M(Du)$ is the set of all subdeterminants, then again the energy is lower semicontinuous. In this case the integrand $f$ is called polyconvex.

Now as we have seen in the last section, the subdeterminants of $Du$ occur quite naturally in the study of Cartesian currents, as they represent the different components of the tangential space of $G(u)$. In fact we can make the observation that there is a one-to-one correspondence between gradients $Du$ and suitably normalized simple, non-vertical $n$-vectors, given by

$$\xi_{Du(x)} := (\partial_{x_1} + \partial_{x_1}u) \wedge ... \wedge (\partial_{x_n} + \partial_{x_n}u).$$

7The integrand may also depend on $x$ and $u(x)$, yet those are not very important in the oncoming discussion.

8The usual way to normalize an $n$-vector is writing $\xi = av_1 \wedge ... \wedge v_n$ for some $a \in \mathbb{R}$ and orthonormal vectors $v_i$ and then defining $\|\xi\| := |a|$. However when dealing with Cartesian currents it is useful to deviate a bit and to use the structure of the problem, where each tangential vector at $(x,u(x))$ has distinct components in $T_x\Omega$ and in $T_{u(x)}M$. Thus we normalize an $n$-vector by requiring its projection onto $T_x\Omega$ to have norm 1 instead. In other words $\xi$ is normalized if $|\langle \xi, \partial_{x_1} \wedge ... \wedge \partial_{x_n} \rangle| = 1$. Of course this is only possible if $\xi$ is non-vertical.
Thus any function \( f \) of the derivative \( Du \) directly corresponds to a function on those normalized, simple, non-vertical \( n \)-vectors.

In order to introduce convexity, we need to start with extending \( \tilde{f} \) onto a convex domain. For this we note (see [GMS98, II.1.2.1 Thm. 6]) that the convex hull of our normalized, simple, non-vertical \( n \)-vectors is given by

\[
\Lambda_1 := \{ \xi \in \wedge_n T(\Omega \times \mathcal{M}) : \langle \xi, \partial_{x_1} \wedge \ldots \wedge \partial_{x_n} \rangle = 1 \}.
\]

Now we are interested in convex functions \( g : \Lambda_1 \to \mathbb{R} \) where \( g(\xi_{Du}) = f(Du) \). However even in rather classical cases, those functions need not be unique as the following example shows:

**Example 4.18.** Consider the Dirichlet-energy for functions \( u : \Omega \to \mathbb{R}^2 \) where \( \Omega \subset \mathbb{R}^2 \). Then our integrand is given by \( f(Du) = \frac{1}{2} |Du|^2 \). Each 2-vector \( \xi \in \Lambda_1 \) can be written as

\[
\xi = \partial_{x_1} \wedge \partial_{x_2} - \sum_{i \in \{1,2\}} a_{i1} \partial_{x_2} \wedge \partial_{y_i} + \sum_{i \in \{1,2\}} a_{i2} \partial_{x_1} \wedge \partial_{y_i} + b \partial_{y_1} \wedge \partial_{y_2}.
\]

If \( \xi \) is the 2-vector associated to \( Du \), then \( b = \det Du \) and \( a_{ij} = \frac{\partial u_i}{\partial x_j} \), which motivates the choice of signs, and the interchanging of indices above. A reasonable guess would now be to choose

\[
g_1(\xi) := \frac{1}{2} |A|^2,
\]

where \( A \) is the matrix given by the \( a_{ij} \). Yet equally valid is the choice of

\[
g_2(\xi) := \frac{1}{2} |A|^2 + b - \det A.
\]

Both \( g_1 \) and \( g_2 \) coincide with \( f \) for simple 2-vectors and both are convex. ■

A reasonable way to circumvent this problem of ambiguous extensions is to take the maximum of all such convex functions, that is

\[
Pf(\xi) := \sup \{ g(\xi) : g : \Lambda_1 \to \mathbb{R} \text{ with } g(\xi_A) \leq f(A) \forall A \in \mathbb{R}^{N \times n}, g \text{ is convex} \}.
\]

This is called the polyconvex envelope of \( f \) and \( f \) is called polyconvex if \( Pf(\xi_A) = f(A) \) for all \( A \in \mathbb{R}^{N \times n} \).

As we have noted in the previous section, the tangential spaces of Cartesian currents can possibly be vertical, so we still need to extend the integrand to the vertical \( n \)-vectors. Since these are characterized by \( \langle \xi, \partial_{x_1} \wedge \ldots \wedge \partial_{x_n} \rangle = 0 \), we need to do so in two steps. First any function \( g : \Lambda_1 \to \mathbb{R} \) can be extended one-homogeneously\(^9\) to

\[
\Lambda_+ := \{ \xi \in \wedge_n T_{(x,y)}(\Omega \times \mathcal{M}) : \langle \xi, \partial_{x_1} \wedge \ldots \wedge \partial_{x_n} \rangle > 0 \}.
\]

\(^9\)That is such that \( g(\alpha \xi) = \alpha g(\xi) \) for all \( \alpha > 0 \), which in this case means

\[
g(\xi) = \langle \xi, \partial_{x_1} \wedge \ldots \wedge \partial_{x_n} \rangle g \left( \frac{\xi}{\langle \xi, \partial_{x_1} \wedge \ldots \wedge \partial_{x_n} \rangle} \right).
\]
This does not disturb the convexity. But then the boundary of $\Lambda_+$ is exactly given by the vertical $n$-vectors. So as a second step, we extend $g$ lower semicontinuously onto the set

$$\Lambda_0 := \{ \xi \in \wedge_n T_{(x,y)}(\Omega \times \mathcal{M}) \mid \langle \xi, \partial_{x_1} \wedge ... \wedge \partial_{x_n} \rangle = 0 \}$$

by setting $g(\xi) := \liminf_{\nu \to \xi, \nu \in \Lambda_+} g(\nu)$ for all $\xi \in \Lambda_0$.

This process results in a convex, lower semicontinuous function defined on all $n$-vectors $\xi$ for which $\langle \xi, \partial_{x_1} \wedge ... \wedge \partial_{x_n} \rangle \geq 0$, which includes all $n$-vectors that can occur as tangential spaces for Cartesian currents. In a single definition we have:

**Definition 4.19.** Let $f : \mathbb{R}^{N \times n} \to \mathbb{R}$. Then the (unique) function $g : \Lambda_0 \cup \Lambda_+ \to \mathbb{R}$ such that

- $g|_{\Lambda_1}$ is the polyconvex envelope of $f$
- $g|_{\Lambda_+}$ is one-homogeneous
- $g(\xi) := \liminf_{\nu \to \xi, \nu \in \Lambda_+} g(\nu)$ for all $\xi \in \Lambda_0$

is called the lower semicontinuous polyconvex extension of $f$. If $f$ is polyconvex then this extension also satisfies $f(A) = g(\xi_A)$.

Using the properties generated by this construction, in concert with the observation that convergence of Cartesian currents results in a kind of weak convergence of the tangential spaces, we can motivate the following fundamental theorem:

**Theorem 4.20 ([GMS98, II.1.3.1 Thm. 1]).** Let $f : \mathbb{R}^{N \times n} \to \mathbb{R}$ and $g$ its polyconvex extension. We define the corresponding functionals by

$$\mathcal{F}(u) := \int_{\Omega} f(Du)dx$$

for $u \in W^{1,n}(\Omega; \mathcal{M})$ and

$$\tilde{\mathcal{F}}(T) := \int g(\xi)d\|T\|$$

for $T \in \text{cart}(\Omega; \mathcal{M})$ where $\xi$ denotes the tangential $n$-vector of $T$. Then $\mathcal{G}$ is lower semicontinuous in $\text{cart}(\Omega; \mathcal{M})$ and

$$\tilde{\mathcal{F}}(\mathcal{G}(u)) = \mathcal{F}(u)$$

if $f$ is polyconvex.
Remark 4.21. What we have discussed so far, is the case functionals for of Cartesian currents representing functions $u : \Omega \to \mathbb{R}^N$, where the integrand $f$ did not depend on either $x \in \Omega$ or $u(x)$. However not only are dependencies of this kind common, but they are in fact unavoidable if we want to consider the case where $u$ is taking values in a manifold $\mathcal{M}$. In this case the derivatives $\partial_x u$ are tangential vectors in $T_{u(x)} \mathcal{M}$ and, as this space changes with $u(x)$, the dependence of $u(x)$ is thus inevitable. So we need to consider the case of generalizing $F : u \mapsto \int_{\Omega} f(x, u(x), Du(x)) dx,$ where

$$f : \Omega \times T\mathcal{M} \to \mathbb{R}$$

and $T\mathcal{M}$ is understood as the set of tuples $(p, v)$, with $p \in \mathcal{M}$ and $v \in T_p \mathcal{M}$.

To do so, we just remark that all of the previous discussion was local, in the sense that we dealt with $\mathbb{R}^N \times n$ as the $n$-th power of the tangential space $T_y \mathbb{R}^N$, where $y = u(x)$ was not further specified. Therefore we define the parametric polyconvex lower semicontinuous extension of $f$ as the function

$$g : \{(x, p, \xi) \mid x \in \Omega, p \in \mathcal{M}, \xi \in \bigwedge^n T_{(x,p)}(\Omega \times \mathcal{M})\} \to \mathbb{R}$$

such that $g(x, p, .) : \bigwedge^n T_{(x,p)}(\Omega \times \mathcal{M}) \to \mathbb{R}$ is the lower semicontinuous extension of $f(x, p, .) : (T_p \mathcal{M})^n \to \mathbb{R}$. Assuming a reasonably smooth dependence of $f$ on $x, p$, the resulting functional on $\text{cart}(\Omega; \mathcal{M})$ will then still be lower semicontinuous. For the details, we will just refer to [GMS98].

Now let us get back to our example.

Example 4.22 (see [GMS98, II.1.2.4 Prop. 5]). Now let us try to find the correct extension for the Dirichlet energy. Let us start with considering functions $u : \mathbb{R}^2 \to \mathbb{R}^2$.

Before we begin, we note the following useful characterization (see [GMS98, II.1.2.1 Thm 7]): Let $f : \mathbb{R}^{n \times N} \to \mathbb{R}$ such that there exists an $n$-vector $\nu$ for which $f(A) \geq \langle \xi, A \rangle$ for all $A$. Then the polyconvex envelope of $f$ is given by

$$g(\xi) := \sup\{\phi(\xi) \mid \phi : \Lambda_1 \to \mathbb{R} \text{ affine and } \phi(\xi_A) \leq f(A)\forall A\}.$$

As in the last example, we have

$$\xi = \partial_{x_1} \wedge \partial_{x_2} - \sum_{i \in \{1,2\}} a_{i1} \partial_{x_2} \wedge \partial_{y_i} + \sum_{i \in \{1,2\}} a_{i2} \partial_{x_1} \wedge \partial_{y_i} + b \partial_{y_1} \wedge \partial_{y_2}$$

where if $\xi$ is the 2-vector associated to $Du$, then $b = \det Du$ and $a_{ij} = \frac{\partial u_i}{\partial x_j}$. 

CHAPTER 4. INTRODUCTION TO CARTESIAN CURRENTS

So any affine function $\phi : \Lambda_1 \to \mathbb{R}$ can be written as

$$\phi(\xi) := \phi(A, b) := \sum_{i,j \in \{1, 2\}} \alpha_{ij} a_{ij} + \beta b + \gamma$$

where $\alpha, \beta$ and $\gamma$ uniquely identify $\phi$. For now, let us fix those variables. Any such function has to satisfy

$$\phi(A, \det A) \leq f(A) = \frac{1}{2} |A|^2$$

for all $A \in \mathbb{R}^{2 \times 2}$, which can be rewritten into

$$\sum_{i,j \in \{1, 2\}} \alpha_{ij} a_{ij} + \gamma \leq \frac{1}{2} |A|^2 - \beta \det A \quad (4.1)$$

As we are in two dimensions, the right hand side is quadratic and thus for any $\lambda \in \mathbb{R}$ we have the scaling

$$\lambda \sum_{i,j \in \{1, 2\}} \alpha_{ij} a_{ij} + \gamma \leq \lambda^2 \left( \frac{1}{2} |A|^2 - \beta \det A \right).$$

Rescaling and sending $\lambda \to \infty$ then yields

$$\frac{1}{2} |A|^2 - \beta \det A \geq 0.$$  

We also remember that in two dimensions $\det A \leq \frac{1}{2} |A|^2$ holds for all $A \in \mathbb{R}^{2 \times 2}$. From this, using the previous equation, we can then derive $|\beta| \leq 1$.

Without loss of generality for each fixed $\alpha, \beta$, we can assume $\gamma$ to take its maximal value. Then since the right hand side in 4.1 is quadratic and the left hand side affine, for each $A \in \mathbb{R}^{2 \times 2}$ there are $\alpha, \beta$ and $\gamma$ such that equality holds at $A$, that is

$$\sum_{i,j \in \{1, 2\}} \alpha_{ij} a_{ij} + \gamma = \frac{1}{2} |A|^2 - \beta \det A$$

and thus $\phi(A, b) = \frac{1}{2} |A|^2 - \beta \det A + \beta b$. On the other hand, for arbitrary $\alpha, \beta$ and $\gamma$, the inequality 4.1 yields

$$\phi(A, b) \leq \frac{1}{2} |A|^2 - \beta \det A + \beta b.$$  

So the supremum over all $\phi$ and thus $g$ is given by

$$g(\xi) = \sup_{|\beta| \leq 1} \left( \frac{1}{2} |A|^2 + \beta(b - \det A) \right) = \frac{1}{2} |A|^2 + |b - \det A|.$$
As the next step, we extend this homogeneously onto \( \Lambda^+ \), which results in

\[
g(\xi) = \langle \xi, \partial_{x_1} \wedge \partial_{x_2} \rangle g \left( \frac{\xi}{\langle \xi, \partial_{x_1} \wedge \partial_{x_2} \rangle} \right) = \frac{1}{2 \langle \xi, \partial_{x_1} \wedge \partial_{x_2} \rangle} |A|^2 + \frac{1}{\langle \xi, \partial_{x_1} \wedge \partial_{x_2} \rangle} \det A - b \langle \xi, \partial_{x_1} \wedge \partial_{x_2} \rangle |A| \]

Then finally we can consider the limit and extend onto \( \Lambda_0 \). Let \( \xi \in \Lambda_0 \) be given by

\[
\xi = - \sum_{i \in \{1,2\}} a_i \partial_{x_2} \wedge \partial_y + \sum_{i \in \{1,2\}} a_i \partial_{x_1} \wedge \partial_y + b \partial_{y_1} \wedge \partial_{y_2}
\]

similar to before. Consider \( \nu \to \xi \) with \( \nu \in \Lambda^+ \). If \( A \neq 0 \), then the first term in \( g \) will diverge to infinity, no matter how \( \nu \) approaches \( \xi \). Thus in this case \( g(\xi) = \infty \). On the other hand, if \( A = 0 \), then we can set \( \nu = \varepsilon \partial_{x_1} \wedge \partial_{x_2} + b \partial_{y_1} \wedge \partial_{y_2} \) and send \( \varepsilon \to 0 \). With some further reasoning this can be shown to be the optimal choice and thus in this case \( g(\xi) = |b| \) or in total

\[
g(\xi) = \begin{cases} 
\frac{1}{2 \langle \xi, \partial_{x_1} \wedge \partial_{x_2} \rangle} |A|^2 + \frac{1}{\langle \xi, \partial_{x_1} \wedge \partial_{x_2} \rangle} \det A \quad & \text{for } \xi \in \Lambda^+ \\
|b| \quad & \text{for } \xi \in \Lambda_0 \text{ and } A = 0 \\
\infty \quad & \text{otherwise.}
\end{cases}
\]

with \( A \) and \( b \) as above.

This technical example exposes an interesting property of the Dirichlet energy in two dimensions. In this case Cartesian currents allow a vertical part to occur, however only if it is vertical in two directions. In this case the vertical part will contribute to the energy by an amount equal to its mass. This can be distilled in the following result:

**Proposition 4.23** ([GMS98, II.1.2.4 Prop. 15]). Let \( \Omega \subset \mathbb{R}^n \) and \( \mathcal{M} \) be \( N \)-dimensional. The polyconvex lower semicontinuous extension of the Dirichlet energy for currents \( T \in \text{cart}(\Omega; \mathcal{M}) \) is only finite if they can be decomposed in the form

\[
T = \mathcal{G}(u_T) + V_T
\]

where \( u_T \in W^{1,2}(\Omega; \mathcal{M}) \) and \( V_T \) is such that all measures of the forms \( \phi \mapsto V_T(\phi dx^1 \wedge ... \wedge dx^n) \) and \( \phi \mapsto V_T(\phi dx^1 \wedge ... \wedge dx^i \wedge dy^j) \) vanish.\(^{10}\) If additionally \( n = 2 \) or \( N = 2 \), then the extension is given by

\[
E(T) = \frac{1}{2} \int_\Omega |u_T|^2 \, dx + \mathcal{M}(V_T).
\]

\(^{10}\)This can be understood as \( V_T \) being somewhat “doubly vertical”. An equivalent condition is that almost every approximate tangential space of \( V_T \) includes at least two independent vectors from \( \{0\} \times T\mathcal{M} \).
Using this, there are some topological considerations to be made:

**Remark 4.24.** Let $\Omega \subset \mathbb{R}^2$. Consider the case of Cartesian currents $T \in \text{cart}^{2,1}(\Omega; \mathbb{R}^N)$ and thus with finite Dirichlet energy as in the previous proposition. For any function $u : \Omega \to \mathbb{R}^N$ with finite Dirichlet-energy, we have that $\partial G(u) = 0$, as it can be approximated smoothly in $W^{1,2}$ and thus in the sense of Cartesian currents. But then since $\partial T = 0$ and $\partial G(u_T) = 0$, the vertical part $V_T$ also cannot have any boundary.

On the other hand, it is an easy consequence of “doubly vertical” property in the previous lemma, that for a two-dimensional domain, $V_T$ decomposes into the form

$$V_T = \sum_{i \in I} \| \{ a_i \} \| \times S_i$$

where $a_i \in \Omega$ and the $S_i$ are currents on $\mathbb{R}^N$. But then since $V_T$ does not have a boundary, neither can the currents $S_i$. As a result each $S_i$ must be generated by closed (In fact even compact, as their mass is bounded) manifolds.

Now consider the case $N = 2$. There are no closed two-dimensional submanifolds of $\mathbb{R}^2$. As a result, all Cartesian currents are represented by graphs.

Similarly consider the case of Cartesian currents $T \in \text{cart}^{2,1}(\Omega; \mathbb{S}^2)$. The only closed two-dimensional submanifold of $\mathbb{S}^2$ is $\mathbb{S}^2$ itself. In other words we can decompose $T$ into

$$T = G(u) + \sum_{i \in I} \alpha_i \left[ \{ a_i \} \times \mathbb{S}^2 \right].$$

We already noted in our example that bubbling results in vertical copies of $\mathbb{S}^2$. We now see that this is in fact the only possible singularity that can happen.

Finally, now that we have settled lower semicontinuity, let us give some remarks on coercivity.

**Definition 4.25.** Let $\mathcal{A} \subset \text{cart}(\Omega; \mathcal{M})$, where $\mathcal{M}$ is a compact manifold and $E : \mathcal{A} \to \mathbb{R}$. Then $E$ is called coercive if all sub-level sets are bounded in mass, that is if

$$\sup \{ M(T) \mid T \in \mathcal{A}, E(T) < E_0 \} < \infty \text{ for all } E_0 \in \mathbb{R}$$

In the case that $\mathcal{M} = \mathbb{R}^N$, we similarly require

$$\sup \left\{ M(T) + \| T \|_p \mid T \in \mathcal{A}, E(T) < E_0 \right\} < \infty \text{ for all } E_0 \in \mathbb{R}$$

for some $p$.

---

11The dimension 2 comes into play here, as it allows us to control all subdeterminants, since $\det Du \leq \frac{1}{2} |\nabla u|^2$.\footnote{The dimension 2 comes into play here, as it allows us to control all subdeterminants, since $\det Du \leq \frac{1}{2} |\nabla u|^2$.}
4.5. DEGREE AND TOPOLOGY

As an immediate consequence of the compactness-closure theorem we then get

**Corollary 4.26.** Let \( \mathcal{A} \subset \text{cart}(\Omega; \mathcal{M}) \) and \( E : \mathcal{A} \to \mathbb{R} \) coercive. Then any bounded sequence has a converging subsequence.

**Example 4.27.** Let us again consider the Dirichlet-energy extended to Cartesian currents and check for coercivity. We can derive this directly by considering the two parts independently. For the functional part we have

\[
\mathcal{M}(\mathcal{G}(u)) = \int_{\Omega} \sqrt{1 + \vert \nabla u \vert^2} \, dx \leq \int_{\Omega} dx + \int_{\Omega} \vert \nabla u \vert \, dx \leq |\Omega| + \sqrt{|\Omega|} \int_{\Omega} \vert \nabla u \vert^2 \, dx
\]

while the vertical part is bounded directly. Thus \( E : \text{cart}(\Omega; \mathbb{S}^2) \to \mathbb{R} \) is coercive.

Another possibility here is to consider a sufficient local condition for coercivity. If the polyconvex lower semicontinuous extension \( g \) of the integrand \( f \) satisfies

\[
g(\xi) \geq c \Vert \xi \Vert,
\]

for all simple \( n \)-vectors \( \xi \), then \( E(T) \geq c \mathcal{M}(T) \) with the same constant. In our case, this is trivial if \( \xi \in \Lambda_0 \). With some further calculation, the same can be derived for \( \xi \in \Lambda_+ \).

**4.5 Degree and topology**

Now we refer back to our example one last and final time. The initial occasion, for which we needed Cartesian currents in the first place was the problem with topology. In the example we have already seen how the topology is somewhat conserved in the vertical part. This can be conveniently formalized and generalized to a plethora of situations by similar considerations. We will hint at this by giving a possible definition of a mapping degree for Cartesian currents.

**Definition 4.28.** Let \( \Omega \subset \mathbb{R}^n \) with a sufficiently regular boundary and \( \mathcal{M} \) an \( n \)-dimensional manifold. Let \( T \in \text{cart}(\Omega; \mathcal{M}) \). Consider \( \pi_{y \#} T \). By the constancy theorem B.13 this can be written as

\[
\pi_{y \#} T(\omega) = \int_{\mathcal{M}} d(y) \omega(y)
\]

where \( d(y) \) is constant on any connected component of \( \mathcal{M} \setminus \text{supp} \partial \pi_{y \#} T \). Then \( \deg(T; y) := d(y) \).

**Lemma 4.29.** Let \( \Omega, \mathcal{M} \) as before and \( u \in C^0(\Omega; \mathcal{M}) \cap C^1(\Omega; \mathcal{M}) \). Then \( \deg(u; y) = \deg(\mathcal{G}(u); y) \) for all \( y \notin u(\partial \Omega) \).
CHAPTER 4. INTRODUCTION TO CARTESIAN CURRENTS

Figure 4.3: A Cartesian current and its projection on the image space. The multiplicity of this projection then defines the mapping degree.

Proof. Let \( \varepsilon > 0 \) such that \( B_\varepsilon(y) \subset M \setminus u(\partial \Omega) \) and \( \omega \in C^\infty(M; \wedge^n) \) such that \( \text{supp } \omega \subset B_\varepsilon(y) \) and \( \int_M \omega = 1 \). Then per definition \( \deg(u; y) = \int_\Omega u^\# \omega \). On the other hand since \( \pi_y \circ (id, u) = u \) we have

\[
\int_\Omega (id, u)^\# \pi_y \# \omega = \langle \pi_y \# \omega, \rangle \langle G(u) \# \omega \rangle = G(u) \langle \omega \rangle
\]

which is equal to \( \deg(G(u), y) \) as the multiplicity of \( (\pi_y)\# G(u) \) is constant in \( B_\varepsilon(y) \) and thus \( \text{supp } \omega \).

Example 4.30. Let us now take a final look at our example, to see what we knew all along. As we only considered functions \( u : \Omega \to S^2 \) with boundary value \( N \), we get that \( \partial \pi_y \# T = 0 \) for any current \( T = G(u) \), as well as any limit of such currents in \( \text{cart}(\Omega; S^2) \). As a direct result \( \deg(T; y) \) is independent of \( y \) and we simply write \( \deg(T) = \deg(T; y) \).

Now, since we started with a slightly modified stereographic map, we have a full cover of the \( S^2 \) and thus \( \deg(T_k) = 1 \). Now we consider the limit. Here the functional part is the constant function \( N \), which does not contribute to any cover and thus has degree 0. Yet there is also the vertical part \( V = \{0\} \times S^2 \). Here \( (\pi_y)\#(V) = \{S^2\} \) and thus \( \deg(T) = 1 \) is preserved in the limit. As a direct consequence, we have thus shown that the problem of minimizing the Dirichlet-Energy with a prescribed mapping degree has a solution in the Cartesian currents, something which is not the case for functions.\(^\text{12}\)

This preservation of the degree in the limit is something that actually holds for Cartesian currents in general:

Proposition 4.31. Let \( \Omega, M \) as before and \( (T_k)_{k \in \mathbb{N}} \subset \text{cart}(\Omega; M) \) with \( T_k \to T \) in \( \text{cart}(\Omega; M) \). Take \( y \in M \) such that \( \text{dist}(y, \text{supp } \partial \pi_y \# T_k) > \varepsilon \) for some \( \varepsilon > 0 \) independent of \( k \in \mathbb{N} \). Then

\[
\deg(y, T) = \lim_{k \to \infty} \deg(y, T_k).
\]

\(^\text{12}\)Any minimizer will be smooth according to Morrey [Mor48] and any smooth function satisfying the Euler-Lagrange equation has to be conformal. Yet any such map with a constant boundary already needs to be constant. (See Lemaire [Lem78] for a complete argument) Note that the constant boundary is the crucial point here, as Brezis and Coron showed in [BC83] that for any non-constant boundary value there are always local minima of different mapping degrees. (Even though not necessarily of any prescribed degree)
Proof. The proof is similar in idea to similar proofs in the classical case. Let $\omega$ be an $n$-form on $\mathcal{M}$ with support in $B_{\varepsilon}(y)$ and $\int_{\mathcal{M}} \omega = 1$. Then

$$\pi_y \# T_k \langle \omega \rangle = \deg(y, T_k)$$

and by continuity

$$\lim_{k \to \infty} \pi_y \# T_k \langle \omega \rangle = \pi_y \# T \langle \omega \rangle = \deg(y, T)$$

as $\text{dist}(y, \text{supp} \partial \pi_y \# T) > \varepsilon$ has to hold as well. \qed
Chapter 5

Length and Wasserstein-distance

5.1 Length of a curve and a generalized Wasserstein distance

Using the preparation from Chapter 3, we can now define a generalized version of the Wasserstein $p$-distance:

**Definition 5.1.** Let $\mathcal{A} \subset \mathcal{I}_k(\Omega)$ be a closed set of currents, $T_0$ and $T_1$ in $\mathcal{A}$. Let $S \in C^1_{pw}([0,1]; \mathcal{A})$ be such that $S(0) = T_0$ and $S(1) = T_1$. Then we define the $p$-length of the homotopy $S$ by

$$L_p(S) := \int_0^1 \inf \left\{ \int |v|^p \, d\|S(t)\| \mid v \in C^1_{pw}(\text{supp } S(t); \mathbb{R}^n), \dot{S}(t) + \mathcal{L}_v S(t) = 0 \right\} dt.$$

If additionally $\mathcal{A} \subset \hat{\mathcal{I}}_k(\Omega)$, the generalized $p$-Wasserstein distance between $T_0$ and $T_1$ is given by

$$W_{p,\mathcal{A}}(T_0, T_1) := \inf \{ L_p(S) \mid S \in C^1_{pw}([0,1]; \mathcal{A}), S(0) = T_0, S(1) = T_1 \}.$$

Note that, as per the usual definition, the infimum in the integral is equal to infinity if there is no corresponding vector field. For curves of finite length this may still occur at some times, but a vector field will still exist for almost all $t$.

**Lemma 5.2.** Let $\hat{T} \in \mathcal{A} \subset \hat{\mathcal{I}}_k(\Omega)$ be a fixed current and define

$$\mathcal{I} := \{ T \in \hat{\mathcal{I}}_k(\Omega) \mid W_{p,\mathcal{A}}(T - \hat{T}) < \infty \}.$$

Then $W_{p,\mathcal{A}}(\cdot, \cdot)$ is a pseudo-metric on $\mathcal{I}$. 

48
5.1. LENGTH AND GENERALIZED WASSERSTEIN DISTANCE

Figure 5.1: Illustration of generalized Wasserstein distance.

Proof. Per definition we have \( \mathcal{W}_{p,A}(.,. \geq 0 \). By using a constant homotopy \( S(t) = T \), we derive \( \mathcal{W}_{p,A}(T,T) = 0 \). If we reverse the homotopy by using \( \tilde{S}(t) = S(1-t) \) we can show the symmetry \( \mathcal{W}_{p,A}(T_0,T_1) = \mathcal{W}_{p,A}(T_1,T_0) \).

Finally, if we have \( S_1, S_2 \in C^{1}_{pw}([0,1]; \dot{I}_k(\Omega)) \) with \( S_1(1) = S_2(0) \) we can define
\[
S : [0,1] \to \dot{I}_k(\Omega); t \mapsto \begin{cases} 
S_1(2t) & \text{for } t \leq 1/2 \\
S_2(2t-1) & \text{for } t \geq 1/2 
\end{cases}
\]
From this it is easy to derive
\[
L_p(S) = L_p(S_1) + L_p(S_2)
\]
which then proves the triangle inequality after the infimum is taken. \( \square \)

In fact \( \mathcal{W}_{p,A}(.,.) \) is actually a metric on \( \mathcal{I} \) if either \( p = 1 \) or \( A \) has a mass bound, which will be a direct consequence of Theorem 7.1 and the fact that \( \dot{\mathcal{F}} \) already was shown to be a metric. However the final property, that \( \mathcal{W}_{p,A}(T_0,T_1) = 0 \) implies \( T_0 = T_1 \), is a bit harder to show alone.

Remark 5.3. Allowing only boundary-less currents in \( A \) is more or less a choice of convenience. An equivalent condition would be to take a set of currents which all have the same boundary. As we will see in Theorem 10.2 the distance will degenerate if we allow the boundary to change. For this we would have to modify the distance according to Definition 10.3.

However if \( A \) is a set with currents of fixed boundary, then \( \tilde{A} := A - \tilde{T} \) for some fixed \( \tilde{T} \in \mathcal{A} \) is a set of boundaryless currents, which will behave exactly the same for all our purposes.

Without a bound on the mass of the currents in \( A \), this does not hold for \( p > 1 \), as the generalized Wasserstein distance gets degenerate, since it is possible to exchange mass of the current against the length of the vector field. This is best seen in the following example, which is in the spirit of Michor and Mumford [MM05, MM06].

1
Example 5.4. Let $T_0 := \{0\}$ and $T_1 := \{1\}$ be currents in $\dot{I}_0(\mathbb{R})$. Define

$$S_k(t) := \sum_{i=0}^{k} \left[ \left\{ \frac{2i + t}{2k + 1} \right\} \right] - \sum_{i=1}^{k} \left[ \left\{ \frac{2i - t}{2k + 1} \right\} \right].$$

Then for $t = 0$ and $t = 1$ all terms except for the first, or last respectively, cancel, so $S_k(0) = T_0$ and $S_k(1) = T_1$. Furthermore for $t \in (0, 1)$ the corresponding vector field can be read of directly, we have $\dot{S}(t) + \mathcal{L}_v S(t) = 0$, where

$$v(t)(x) := \begin{cases} \frac{1}{2k+1} & \text{for } x = \frac{2i + t}{2k+1}, i \in \{0, \ldots, k\} \\ -\frac{1}{2k+1} & \text{for } x = \frac{2i - t}{2k+1}, i \in \{1, \ldots, k\}. \end{cases}$$

Thus we have

$$\int |v|^p \, d\|S(t)\| = (2k + 1) \frac{1}{(2k + 1)^p}$$

and

$$\mathcal{W}_{p,T_k(\Omega)}(T_0, T_1) \leq \lim_{k \to \infty} \sqrt[p]{(2k + 1) \frac{1}{(2k + 1)^p}} = \lim_{k \to \infty} (2k + 1)^{1/p - 1} \to 0$$

for $p > 1$. \[\]
5.1. LENGTH AND GENERALIZED WASSERSTEIN DISTANCE

While this example only covers the 0-dimensional case, we remark that the idea behind this is universal, as for any admissible curve $S \in C^1_{pw}([0, 1]; \dot{I}_k(\Omega))$, we can also similarly define

$$\tilde{S}(t) := \sum_{i=0}^{k} S \left( \frac{2i + t}{2k + 1} \right) - \sum_{i=1}^{k} S \left( \frac{2i - t}{2k + 1} \right)$$

which exhibits exactly the same phenomenon.

Similarly, the mass also comes into play when comparing distances for different exponents:

**Lemma 5.5.** Let $A \subset \dot{I}_k(\Omega)$, and $T_0$ and $T_1$ in $A$ and $q < p$. If $\mathcal{M}(T) \leq M_0$ for all $T \in A$ then

$$W_{q,A}(T_0, T_1) \leq M_0^{\frac{p-q}{p}} \cdot W_{p,A}(T_0, T_1)$$

**Proof.** For a fixed $\varepsilon > 0$, let $S \in C^1_{pw}([0, 1]; A)$ such that $S(0) = T_0$ and $S(1) = T_1$ and $L_q(S) < \mathcal{W}_{A,q}(T_0, T_1) + \varepsilon$. Note that using Hölder’s inequality with $1 = \frac{q}{p} + \frac{p-q}{p}$ we have for any $T \in A$ and $v : \text{supp} \, A \to \mathbb{R}^n$:

$$\int |v|^q \, d\|T\| \leq \left( \int |v|^p \, d\|T\| \right)^{\frac{q}{p}} \left( \int |1| \, d\|T\| \right)^{\frac{p-q}{p}}$$

$$= \left( \int |v|^p \, d\|T\| \right)^{\frac{q}{p}} (\mathcal{M}(T))^{\frac{p-q}{p}}$$

using this we can then estimate

$$L_p(S)^q = \int_0^1 \inf \left\{ \int |v|^q \, d\|S(t)\| \big| \dot{S} + \mathcal{L}_v S = 0 \right\} \, dt$$

$$\leq \int_0^1 \inf \left\{ \left( \int |v|^p \, d\|S(t)\| \right)^{\frac{q}{p}} (\mathcal{M}(S(t)))^{\frac{p-q}{p}} \big| \dot{S} + \mathcal{L}_v S = 0 \right\} \, dt$$

$$\leq M_0^{\frac{p-q}{p}} \int_0^1 \inf \left\{ \left( \int |v|^p \, d\|S(t)\| \right)^{\frac{q}{p}} \big| \dot{S} + \mathcal{L}_v S = 0 \right\} \, dt$$

and finally using Jensen’s inequality

$$\leq M_0^{\frac{p-q}{p}} \left( \int_0^1 \inf \left\{ \int |v|^p \, d\|S(t)\| \big| \dot{S} + \mathcal{L}_v S = 0 \right\} \, dt \right)^{\frac{q}{p}} = M_0^{\frac{p-q}{p}} L_p(S)^q \Box$$

**Remark 5.6.** Continuing the discussion from Remark 2.12, if we look at the classical Wasserstein distance, we also have a dual formulation (see for example [Vil09, Chp.5]) given by

$$W_1(\mu, \nu) = \sup_{\|\nabla \phi\|_{\infty} \leq 1} \left[ \int \phi \, d\mu - \int \phi \, d\nu \right].$$
It should come as no surprise that the notions of duality for $\mathcal{W}_1$ and $\tilde{F}$ are identical when they overlap for $k = 0$. However, we cannot assume a dual formulation of the Wasserstein-distance for integer rectifiable currents of higher dimension or for the generalized $p$-distance.

### 5.2 A vertical Wasserstein distance for Cartesian currents

In principle, it is possible to directly apply the definition of a generalized Wasserstein distance to Cartesian currents. In practice however, it is better to further utilize the structure of graphs. The obvious idea here is to reduce the admissible space $\mathcal{A}$ to a subspace of Cartesian currents. This also allows us to use near minimizers for the purpose of interpolation. However there is a further possible improvement using the product-structure of the underlying space. As the classical $L^p$-distance is only concerned with the difference between function values for the same fixed point $x \in \Omega$, we will achieve a similar behavior by limiting movement to vertical directions.

![Figure 5.3: Illustration of generalized vertical Wasserstein distance.](image)

**Definition 5.7.** Let $p \geq 1$, $\mathcal{A} \subset \text{cart}(\Omega; \mathcal{M})$ a closed subset, $T_0, T_1 \in \mathcal{A}$. Let $S \in C^1_{pw}([0, 1]; \mathcal{A})$ be such that $S(0) = T_0$ and $S(1) = T_1$. Then we define the $p$-length of the homotopy $S$ by

$$
\mathcal{L}_{p, \text{vert}}(S) := \sqrt{\int_0^1 \inf \left\{ \int_\Omega |v|^p \, dx \mid v \in C^1_{pw}(\Omega; T_x \mathcal{M}) \text{ s.t. } \dot{S}(t) + \mathcal{L}_{(0,v)} S(t) = 0 \right\} dt}
$$

and the generalized $p$-Wasserstein distance between $T_0$ and $T_1$ by

$$
\mathcal{W}_{p, \mathcal{A}, \text{vert}}(T_0, T_1) := \inf \left\{ \mathcal{L}_{p, \text{vert}}(S) \mid S \in C^1_{pw}([0, 1]; \mathcal{A}), S(0) = T_0, S(1) = T_1 \right\}.
$$

Again, as per the usual definition, the infimum in the integral for the length is equal to infinity, if there is no corresponding vector field. We will write $\mathcal{W}_{p, \text{vert}} := \mathcal{W}_{p, \text{cart}(\Omega; \mathcal{M}), \text{vert}}$ if we assume $\mathcal{A}$ to be the set of all admissible currents.
5.2. A VERTICAL WASSERSTEIN DISTANCE

We also want to remark that

\[
\int_{\Omega} |v|^p \, dx = \int |v|^p \, d\pi_x \# S(t)
\]

which further ties this definition back to the definition of the generalized Wasserstein distance in the last section.

**Lemma 5.8.** If \( A \subset \text{cart}(\Omega; \mathcal{M}) \) and \( \mathcal{W}_{p,A,\text{vert}}(T_0, T_1) < \infty \) for any \( T_0, T_1 \in A \), then \( \mathcal{W}_{p,A,\text{vert}} \) defines a pseudo-metric on \( A \) for all \( p \geq 1 \).

**Proof.** This is completely similar to the corresponding proof in Lemma 5.2: Per definition we have \( \mathcal{W}_{p,A,\text{vert}}(\ldots) \geq 0 \). By using a constant homotopy \( S(t) = T \), we derive \( \mathcal{W}_{p,A,\text{vert}}(T, T) = 0 \). If we reverse the homotopy by using \( \tilde{S}(t) = S(1 - t) \) we can show the symmetry \( \mathcal{W}_{p,A,\text{vert}}(T_0, T_1) = \mathcal{W}_{p,A,\text{vert}}(T_1, T_0) \). Finally, if we have \( S_1, S_2 \in C^1_{\text{pw}}([0,1]; A) \) with \( S_1(1) = S_2(0) \) we can define

\[
S : [0,1] \rightarrow A; t \mapsto \begin{cases} S_1(2t) & \text{for } t \leq 1/2 \\ S_2(2t - 1) & \text{for } t \geq 1/2 \end{cases}
\]

From this it is easy to derive

\[
L_p(S) = L_p(S_1) + L_p(S_2)
\]

which then proves the triangle inequality after the infimum is taken. \( \square \)

We note that \( \mathcal{W}_{p,A,\text{vert}} \) actually defines a metric in many cases, depending on the the space \( A \). Again, what is still missing is a proof, that \( \mathcal{W}_{p,A,\text{vert}}(T_0, T_1) = 0 \) implies \( T_0 = T_1 \). For this we will need some lower bounds, as in Proposition 8.9. The finiteness condition is necessary as \( A \) may include Cartesian currents of different degrees which cannot be connected by a continuous curve.

Next let us look a bit more at the significance of the vector field by which our current is moved and its interplay with the derivative of the functional part of a Cartesian current.

**Lemma 5.9.** Let \( u_t : \Omega \rightarrow \mathbb{R}^N \) a differentiable family of functions such that \( G(u_t) \in \text{cart}(\Omega; \mathbb{R}^N) \) and the corresponding family of Cartesian currents \( S(t) := G(u_t) \) is differentiable and for almost all \( t \) there are vector fields \( v_t \) such that \( S(t) + L_{(0,v_t)} S(t) = 0 \). Then the time derivative \( \mu_t : \phi \mapsto \dot{S}(t) \langle \phi ydx \rangle \) is absolutely continuous and \( \mu_t = \dot{u}_t dx = v_t dx \).

**Proof.** By [GMS98, I.4.2.3, Thm 1], for each \( t \), the measure given by \( \phi \mapsto S(t) \langle \phi(x) y_t dx \rangle \) is absolutely continuous and is the same as \( (u_t)_t dx \). Now we consider

\[
\dot{S}(t) \langle \phi(x) y_t dx \rangle = \lim_{h \rightarrow 0} \frac{S(t + h) \langle \phi(x) y_t dx \rangle - S(t) \langle \phi(x) y_t dx \rangle}{h}
\]
which is also well defined for characteristic functions $\phi = X_A$. Thus we have $\dot{S}(t)\langle X_A(x) y_i \rangle = 0$ if $\mathcal{H}^n(A) = 0$ since $\int_A (u_{t+h}(x) - u_t(x))_i \, dx = 0$ in this case. Thus $\phi \mapsto \dot{S}(t)\langle \phi(x) y_i \rangle$ is absolutely continuous.

Additionally we derive

$$
\dot{S}(t)\langle \phi(x) y_i \rangle = S(t)\langle \mathcal{L}_{(0,v_t)} (\phi(x) y_i \rangle \\
= S(t)\langle i_{(0,v_t)} d (\phi(x) y_i \rangle + i_{(0,v_t)} (\phi(x) y_i \rangle \\
$$

where the second term vanishes since $(0, v_t) \perp (e_i, 0)$ for all $i \in \{1, \ldots, n\}$ implies $i_{(0,v_t)} \, dx = 0$ and the first term simplifies to

$$
= S(t)\langle i_{(0,v_t)} (\phi(x) dy^i \wedge dx) \rangle \\
$$

and again using $(0, v_t) \perp (e_i, 0)$

$$
= S(t)\langle \phi(x) (v_t(x))_i \rangle \, dx.
$$

Since the differential form does not depend on $y$, we can project on the $x$ component and finally end up with

$$
\dot{S}(t)\langle \phi(x) v_t(x)_i \rangle \, dx.
$$

which implies $\mu_t = v_t(x) \, dx$. But then since the difference quotient $\frac{u_{t+h} - u_t}{h}$ converges to $v_t$ in the sense of measures, it also follows that $v_t = \dot{u}$ almost everywhere.

**Remark 5.10.** We have seen in Example 5.4 that the generalized Wasserstein distance tends to degenerate for $p > 1$ if it is not restricted to a smaller set of currents. The solution there was to introduce a mass bound. For Cartesian currents this problem does not occur directly, as adding multiple graphs with opposite orientation leads to leaving the space of Cartesian currents immediately.

One could try to work in the direction of the original counterexample in [MM05, MM06], by considering graphs of highly oscillating functions. However even this approach does not work due to our restriction on vertical vector fields.

In fact we will see in Chapter 8 that the $L^p$-distance between the function parts of two Cartesian currents is a lower bound on their $W_{p, \text{vert}}$-distance.
5.2. A VERTICAL WASSERSTEIN DISTANCE

\[
\begin{bmatrix}
\{a\} \times S^2
\end{bmatrix}
\]

\[S_1\]

\[
\begin{bmatrix}
\{b\} \times S^2
\end{bmatrix}
\]

\[S_3\]

\[S_2\]

Figure 5.4: Illustration of degenerate vertical distance. In the first step we use “reverse bubbling” to turn the vertical part into the graph of a function as well as to create additional vertical parts. We then move the graph a small distance in the second step and finally reform the bubbles in the third. The additional vertical parts then annihilate, leaving only one last part. A careful analysis shows that this can be done creating arbitrary small length.

Instead, with Cartesian currents, the problem lies in the vertical parts, where we can mimic the problematic behavior from before as the following example shows:

**Example 5.11.** We consider Cartesian currents from \(\Omega \subset \mathbb{R}^2\) to \(S^2\). We assume that \(a, b \in \Omega\) as well as the connecting line \([a, b] \subset \Omega\). We will show that the two currents \(T_0 := G((0, 0, 1)) + \{a\} \times S^2\) and \(T_1 := G((0, 0, 1)) + \{b\} \times S^2\) actually have a \(W_{2,\text{vert}}\)-distance of 0.

The approach for this will be quite similar to Example 5.4. We will create a number of oppositely signed vertical parts \(\pm \{x_i\} \times S^2\) at intermediate points \(x_i \in [a, b]\) and then move them. Here the situation is made slightly more complicated by the fact that we are only allowed to move in a vertical direction. To counter this, we will need to “unravel” the bubbles which are represented by the vertical parts.

Take a fixed smooth map \(h : B_1(0) \to S^2\) such that \(h|_{\partial B_1(0)} = (0, 0, 1)\) and \(\deg h = 1\), that is when counted with orientation, each point of \(S^2\) is covered by \(h\) once. A good candidate here is the stereographic map, rescaled from \(\mathbb{R}^2\) to \(B_1(0)\) by a smooth function like sending the radius \(r \mapsto \tan(2r/\pi)\). However the precise function does not really matter here. We assume \(h\) to be extended by \((0, 0, 1)\) on \(\mathbb{R}^2\). Note that for \(h_r : x \mapsto h(x/r)\) and \(r \to 0\) we have \(G(h_r) \to G((0, 0, 1)) + \{0\} \times S^2\) in the sense of Cartesian currents, which is the classical bubbling phenomenon.

Now pick \(k \in \mathbb{N}\) and \(\varepsilon > 0\) such that \(\varepsilon < \frac{|a-b|}{k}\) and \(B_{\varepsilon}(x) \subset \Omega\) for all \(x \in [a, b]\).
Let $x_i := a + i \frac{b-a}{k}$. We define

$$u_t(x) := \begin{cases} h \left( \frac{x-x_i}{\varepsilon t} \right) & \text{for } x \in B_{\varepsilon}(x_i), i \in \{0, \ldots, k-1\} \\ (0,0,1) & \text{otherwise} \end{cases}$$

for $t \in (0,1]$, which just gives a bubble unfurling around each point $x_i$. We then define the first part of our homotopy as

$$S_1(t) := \begin{cases} \mathcal{G}((0,0,1)) + \{a \} \times S^2 & \text{for } t = 0 \\ \mathcal{G}(u_t) - \sum_{i=1}^{k-1} \{x_i \} \times S^2 & \text{otherwise.} \end{cases}$$

A short calculation gives us, using Lemma 5.9

$$\mathcal{L}_{2,\text{vert}}(S_1)^2 = \int_0^1 \int_{\Omega} \left| \frac{\partial u_t}{\partial t} \right|^2 \, dx \, dt$$

$$= \int_0^1 \int_{B_{\varepsilon}(0)} \left| \frac{\partial}{\partial t} h \left( \frac{x}{\varepsilon t} \right) \right|^2 \, dx \, dt$$

$$= \int_0^1 \int_{B_{\varepsilon}(0)} \frac{x}{\varepsilon t^2} \cdot \nabla h \left( \frac{x}{\varepsilon t} \right)^2 \, dx \, dt$$

$$= \int_0^1 k \varepsilon^2 \int_{B_{\varepsilon}(0)} \left| \frac{x}{\varepsilon t} \cdot \nabla h \left( \frac{x}{\varepsilon t} \right) \right|^2 \frac{1}{\varepsilon^2 t^2} \, dx \, dt$$

$$= \int_0^1 k \varepsilon^2 \int_{B_{1}(0)} |y \cdot \nabla h(y)|^2 \, dy \, dt$$

$$= C k \varepsilon^2.$$ 

where $C$ only depends on the choice of $h$. As a second step we move those unfurled bubbles to the next position, using

$$v_t(x) := \begin{cases} h \left( \frac{x-(tx_{i+1}+(1-t)x_i)}{\varepsilon} \right) & \text{for } x \in B_{\varepsilon}(tx_{i+1}+(1-t)x_i) \\ (0,0,1) & \text{otherwise} \end{cases}$$

to define

$$S_2(t) := \mathcal{G}(v_t) - \sum_{i=1}^{k-1} \{x_i \} \times S^2.$$ 

Then again, we can calculate the distance

$$\mathcal{L}_{2,\text{vert}}(S_2)^2 = \int_0^1 \int_{\Omega} \left| \frac{\partial v_t}{\partial t} \right|^2 \, dx \, dt$$

$$= \int_0^1 \int_{\Omega} \left| \frac{\partial}{\partial t} h \left( \frac{x-(tx_{i+1}+(1-t)x_i)}{\varepsilon} \right) \right|^2 \, dx \, dt$$
5.2. A VERTICAL WASSERSTEIN DISTANCE

\[
= \int_0^1 \int_\Omega \left| \frac{x_i - x_{i+1}}{\varepsilon} \cdot \nabla h \left( \frac{x - (tx_{i+1} + (1-t)x_i)}{\varepsilon} \right) \right|^2 \, dx \, dt
\]
\[
= \int_0^1 \int_\Omega \left| (x_i - x_{i+1}) \cdot \nabla h \left( \frac{x - (tx_{i+1} + (1-t)x_i)}{\varepsilon} \right) \right|^2 \frac{1}{\varepsilon^2} \, dx \, dt
\]
\[
= \int_0^1 k \int_{B_1(0)} \frac{1}{k^2} \left| (a - b) \cdot \nabla h \left( \frac{y}{\varepsilon} \right) \right|^2 \, dy \, dt
\]
\[
= C \frac{1}{k}
\]

where \( C \) only depends on the choice of \( h \) and is proportional to \( |a - b|^2 \).

Finally we can close all bubbles again using a reverse version of \( v_t \):

\[
w_t(x) := \begin{cases} 
  h \left( \frac{x - x_i}{\varepsilon(1-t)} \right) & \text{for } x \in B_{\varepsilon}(x_i), i \in \{1,...,k\} \\
  (0,0,1) & \text{otherwise}
\end{cases}
\]

and

\[
S_3(t) := \begin{cases} 
  G(2t) - \sum_{i=1}^{k-1} \left[ \{x_i\} \times S^2 \right] & \text{for } t \in [0,1) \\
  \left[ \{b\} \times S^2 \right] & \text{for } t = 1
\end{cases}
\]

with \( \mathcal{L}_{2,\text{vert}}(S_3)^2 = C k \varepsilon^2 \) by a similar calculation. If we combine those three homotopies into a single curve \( S \), we then end up with

\[
\mathcal{L}_{2,\text{vert}}(S) \leq \tilde{C} \sqrt{k \varepsilon^2 + \frac{1}{k}} \to 0 \text{ for } \varepsilon = \frac{|a - b|}{2k}, k \to \infty
\]

and thus \( W_{2,\text{vert}}(T_0, T_1) = 0 \).

**Remark 5.12.** Note that if we restrict the previous example to a single bubble, that is the case \( k = 1 \), which can be done by a bound on the mass of the current, then the expended energy is proportional to \( |a - b| \). In fact the energy is bounded from below by \( 4\pi |a - b| \).
Chapter 6

The trace of a curve

A recurring theme in this work will be the idea that moving a \( k \)-dimensional object traces out a \( k + 1 \)-dimensional volume in the same way as a moving point traces out a curve. This is a rather natural principle which has been used in many different variants under the heading of homotopy formula. Most relevant for us will be the homotopy formula for currents, given by

\[
\partial h_\#(I \times T) = h_\#(I \times \partial T) + h(1, .)_\#T - h(0, .)_\#T
\]

where \( T \in D_k(\Omega) \) and \( h : I \times \Omega \to \Omega \) is a homotopy.\(^1\) Note that in this case \( t \mapsto h(t, .)_\# \) defines a curve of currents. So it is natural to ask, if there is a homotopy formula for curves of currents. We will see in this chapter that there is in fact something quite similar.

6.1 The spatial and the space-time trace

To be more precise, we observe that a sufficiently regular curve of \( k \)-dimensional currents always generates a \( k + 1 \)-dimensional current by the following definition:

**Definition 6.1.** Let \( I \subset \mathbb{R} \) be an interval and \( S : I \to I_k(\Omega) \) an absolutely continuous map such that for almost all times there is a vector-field \( v(t) : \Omega \to \mathbb{R}^n \) with \( \dot{S}(t) + \mathcal{L}_{v(t)} S(t) = 0 \). Then we define the trace of \( S \) as the \( k + 1 \)-current given by

\[
S_\#(I) : \omega \mapsto \int_I S(t) \langle i_{v(t)} \omega \rangle \, dt, \quad \omega \in C^\infty(\Omega; \wedge^{k+1}).
\]

\(^1\)See [dR55, dR84, Chp.14] or [FF60] for a version with additional bounds on the masses of currents involved.
(Note the abuse of notation.) We similarly define the space-time trace of \( S \) as the trace of \((id, S) : I \to \mathbf{I}_k(\mathbb{R} \times \Omega); t \mapsto \{t\} \times S(t)\) given by
\[
(id, S)_\#(I) : \omega \mapsto \int _I \{t\} \times S(t) \langle i_1, v(t) \rangle \omega \, dt, \quad \omega \in C^\infty(I \times \Omega; \wedge^{k+1}).
\]

Figure 6.1: The spatial and space-time traces of the curves \( S_1 : t \mapsto \{0, 1\} \times \{t\} \) and \( S_2 : t \mapsto \{0, 1\} \times \{t^2\} \). Note that both have the same trace in space, but that their space-time trace is different.

**Remark 6.2.** It is well known that the trace of an ordinary curve does not have a unique parametrization, so we can expect our traces to show a similar flexibility. We will see this in Lemma 6.4 However there are two additional phenomena that can occur:

The first one is related to the fact that our curves are not simple and can pass over the same point multiple times.\(^2\) Combining this with the orientation allows those currents to both stack up to a multiple as well as for cancellation to occur.

\(^2\)In this context there seems to be no useful equivalent of “simple” for curves of currents. Just having \( S(t) \neq \pm S(s) \) for \( s \neq t \) does not really prevent overlap of parts of currents and something like \( \mathrm{supp} S(t) \cap \mathrm{supp} S(s) = \emptyset \) is far too restrictive in situations where only a part of the current is in motion.
While the first phenomenon occurs in a similar way for ordinary oriented curves, the second is specific to our (higher dimensional) situation. Since for a fixed time, our curves are not located at a single point in space, we are able to do a nonuniform reparametrization. What this means is that instead of a simple diffeomorphism $\phi : I \to I$, we chose a smooth family of diffeomorphisms $\phi_x : I \to I$ that depends on the position in space to reparametrize the curve of currents. As giving a readable and correct definition of the resulting current is a bit complicated, so we appeal to intuition via a simple example. Let

$$\psi : [0, 1] \times [0, 1] \to [0, 1] \times [0, 1]; (t, x) \mapsto \psi(t, x)$$

be a diffeomorphism that preserves the boundary. Then $t \mapsto \psi(t, \cdot) \in [0, 1]^{2}$ is a curve of currents that will always have the same trace as $t \mapsto [\{t\}] \times S(t)$, namely $[0, 1]^{2}$. The corresponding family of diffeomorphisms is given by $\phi_{(x,y)}(t) := \psi(t, x)$.

Avoiding those problems is one of the motivations behind the space-time trace. Here, we are always able to recover the exact parametrization by simply slicing the trace along a specific constant time. (See Corollary 6.6)

As the notation suggests, our definition of space-time trace should be the same as the trace of a curve of currents given by $t \mapsto J_{\{t\}} \times S(t)$. This is the essence of the next proposition.

**Proposition 6.3.** Let $S : I \to I_k(\Omega)$ an absolutely continuous curve with corresponding vector-field $v(t)$. Then the curve $R : t \mapsto [\{t\}] \times S(t); I \to I_k(I \times \Omega)$ has a corresponding vector-field $(1, v(t))$, that is

$$\dot{R}(t) + L_{(1, v(t))} R(t) = 0$$

for almost all $t$.

**Proof.** Let $\omega \in C^\infty(I \times \Omega; \land^k)$. Then we can write $\omega = dt \wedge \alpha + \beta$ where $\alpha|_{t=t_0} \in C^\infty(\Omega; \land^{k-1})$ and $\beta|_{t=t_0} \in C^\infty(\Omega; \land^k)$ for each $t_0 \in I$, separating $\omega$ into the terms that include $dt$ and those that do not. Now we fix a time $t_0 \in I$ for which $v(t_0)$ exists and note that

$$\mathcal{L}_{(1, v(t_0))} (dt \wedge \alpha) = \left( \text{di}_{(1, v(t_0))} + i_{(1, v(t_0))} d \right) (dt \wedge \alpha)$$

$$= d(\text{di}_{(1, v(t_0))} dt \wedge \alpha) - dt \wedge d(i_{(1, v(t_0))} \alpha) - i_{(1, v(t_0))} (dt \wedge d\alpha)$$

$$= d\alpha - dt \wedge (\text{di}_{(1, v(t_0))} \alpha) - i_{(1, v(t_0))} (dt \wedge d\alpha) - dt \wedge (i_{(1, v(t_0))} d\alpha)$$

$$= -dt \wedge \left( \mathcal{L}_{(1, v(t_0))} \alpha \right).$$
However the tangential space of \{t_0\} \times S(t_0) is orthogonal to the \(t\)-direction in every point, so \([\{t_0\}] \times S(t_0)\langle \frac{\partial}{\partial t} \wedge (L_{(1,v(t_0))\alpha}) \rangle = 0\) and thus
\[
[\{t_0\}] \times S(t_0)\langle L_{(1,v(t_0))\omega} \rangle = [\{t_0\}] \times S(t_0)\langle L_{(1,v(t_0))\beta} \rangle.
\]

Next we simplify the difference-quotient, using the same reasoning.
\[
\frac{[\{t_0 + h\}] \times S(t_0 + h) - [\{t_0\}] \times S(t_0)}{h} = S(t_0 + h)\langle \beta|_{t = t_0 + h} \rangle - S(t_0)\langle \beta|_{t = t_0} \rangle
\]
\[
= S(t_0 + h)\langle \beta|_{t = t_0 + h} - \beta|_{t = t_0} \rangle + \frac{S(t_0 + h) - S(t_0)}{h} \langle \beta|_{t = t_0} \rangle
\]
\[
\rightarrow S(t_0)\langle \frac{\partial \beta}{\partial t}|_{t = t_0} \rangle + \hat{S}(t_0)\langle \beta|_{t = t_0} \rangle
\]
\[
= S(t_0)\langle L_{v(t)}(\beta|_{t = t_0}) \rangle = S(t_0)\langle \frac{\partial \beta}{\partial t}|_{t = t_0} + L_{v(t)}(\beta|_{t = t_0}) \rangle
\]

Note that the convergence of \(S(t_0 + h)\langle \frac{\beta|_{t = t_0 + h} - \beta|_{t = t_0}}{h} \rangle\) is not a problem since \(\beta\) is continuously differentiable and thus the difference quotient converges uniformly.

Finally we calculate (Here we use \(d_x\) to denote the exterior derivative on \(\Omega\), and note that \(i_{v(t)}\) is defined for forms on \(\Omega\) as is evident by the context.)
\[
L_{(1,v(t_0))}\beta = di_{(1,v(t_0))}\beta + i_{(1,v(t_0))}d\beta
\]
\[
= d\left(i_{(1,v(t_0))}\beta\right) + i_{(1,v(t_0))}\left(dt \wedge \frac{\partial \beta}{\partial t}|_{t = t_0} + d_x \beta|_{t = t_0}\right)
\]
\[
= d_x i_{v(t_0)} \beta|_{t = t_0} + dt \wedge \frac{\partial (i_{(1,v(t_0))}\beta)}{\partial t}|_{t = t_0}
\]
\[
+ i_{(1,v(t_0))}dt \wedge \frac{\partial \beta}{\partial t}|_{t = t_0} - dt \wedge i_{v(t_0)}\frac{\partial \beta}{\partial t}|_{t = t_0} + i_{v(t_0)}d_x \beta|_{t = t_0}
\]
\[
= L_{v(t_0)} \beta|_{t = t_0} + \frac{\partial \beta}{\partial t}|_{t = t_0} + dt \wedge \left(\frac{\partial (i_{(1,v(t_0))}\beta)}{\partial t}|_{t = t_0} - i_{v(t_0)}\frac{\partial \beta}{\partial t}|_{t = t_0}\right)
\]

where because of the orthogonality of \(dt\) the last terms again do not contribute, resulting in
\[
\frac{[\{t_0\}] \times S(t_0)\langle L_{(1,v(t_0))}\beta \rangle}{[\{t_0\}] \times S(t_0)\langle \beta|_{t = t_0} \rangle = S(t_0)\langle L_{v(t)} \beta|_{t = t_0} + \frac{\partial \beta}{\partial t}|_{t = t_0} \rangle
\]
and so
\[
\hat{R}(t_0)\langle \omega \rangle = S(t_0)\langle \frac{\partial \beta}{\partial t}|_{t = t_0} + L_{v(t)} \beta|_{t = t_0} \rangle
\]
\[ = R(t_0)\langle \mathcal{L}_{(1,v(t_0))}\beta \rangle = R(t_0)\langle \mathcal{L}_{(1,v(t_0))}\omega \rangle \]

which closes our chain of reasoning. \(\square\)

Let us close this section with some miscellaneous observations which will be useful later on.

**Lemma 6.4.** Let \(I, J \subset \mathbb{R}\) be closed intervals and \(S : J \to I_k(\Omega)\) an absolutely continuous curve with corresponding vector-field \(v(t)\). Furthermore let \(\phi : I \to J\) be a diffeomorphism with orientation given by \(\text{sgn } \dot{\phi} \in \{-1, 1\}\).

Then
\[
(S \circ \phi)_{\#}(I) = \left(\text{sgn } \dot{\phi}\right) S_{\#}(J)
\]

and
\[
(\phi, id)_{\#}((id, S) \circ \phi)_{\#}(I) = \left(\text{sgn } \dot{\phi}\right) (id, S)_{\#}(J).
\]

**Proof.** First we note that a simple application of the chain rule yields
\[
\frac{\partial (S \circ \phi)}{\partial t}(t)\langle \omega \rangle = \frac{\partial \phi}{\partial t} S(\phi(t)) \langle \omega \rangle
\]
for all \(\omega \in C^\infty(\Omega; \wedge^k)\) and thus by linearity of the Lie-derivative
\[
\frac{\partial (S \circ \phi)}{\partial t}(t) + \mathcal{L}_{\frac{\partial \phi}{\partial t} v(\phi(t))} S(\phi(t)) = 0
\]
for all \(t \in I\) for which \((v \circ \phi)(t)\) is defined. Then we have for \(\omega \in C^\infty(\Omega; \wedge^k)\), using linearity and the transformation formula for one-dimensional integrals
\[
(S \circ \phi)_{\#}(I)\langle \omega \rangle = \int_I (S \circ \phi)(t) \left\langle i_{\frac{\partial \phi}{\partial t} v(\phi(t))} \omega \right\rangle dt
\]
\[
= \int_I (S \circ \phi)(t) \langle i_{v(\phi(t))} \omega \rangle \frac{\partial \phi}{\partial t} dt
\]
\[
= \text{sgn } \dot{\phi} \int_J S(t) \langle i_{v(t)} \omega \rangle dt = \text{sgn } \dot{\phi} S_{\#}J.
\]

This proves our first claim.

For the second claim we similarly get that
\[
\frac{\partial ((id, S) \circ \phi)}{\partial t}(t) + \mathcal{L}_{\frac{\partial \phi}{\partial t} (1,v(\phi(t)))} S(\phi(t)) = 0
\]
and
\[
(\phi, id)_{\#}((id, S) \circ \phi)_{\#}(I)\langle \omega \rangle
\]
\[
= ((id, S) \circ \phi)_{\#}(I) \left\langle (\phi, id)^{\#} \omega \right\rangle
\]

Now since treated as a current in \( \mathbb{I}_k(\Omega) \) an absolutely continuous curve with corresponding vector-field \( v(t) \). Then

\[
\partial S_\#(I) = S(b) - S(a) + (\partial S)_\#(I)
\]

and

\[
\partial(id,S)_\#(I) = \llbracket \{b\} \rrbracket \times S(b) - \llbracket \{a\} \rrbracket \times S(a) + (id,\partial S)_\#(I)
\]

treated as a current in \( J \times \Omega \) for \( I \subset J \).

Proof. Using the definition immediately yields

\[
\partial S_\#(I) \langle \omega \rangle = S_\#(I) \langle d\omega \rangle
\]

\[
= \int_a^b S(t) \langle i_{v(t)} d\omega \rangle dt
\]

\[
= \int_a^b S(t) \langle \mathcal{L}_v(t) \omega \rangle - S(t) \langle di_{v(t)} \omega \rangle dt
\]

\[
= \int_a^b \dot{S}(t) \langle \omega \rangle - \partial S(t) \langle i_{v(t)} \omega \rangle dt
\]

\[
= S(b) \langle \omega \rangle - S(a) \langle \omega \rangle - \int_a^b \partial S(t) \langle i_{v(t)} \omega \rangle dt.
\]

Now since \( d \) and \( \mathcal{L}_v \) commute for differential forms, \( \partial \) and \( \mathcal{L}_v \) in turn commute for currents. So we get \( \frac{d}{dt} \partial S(t) + \mathcal{L}_v \partial S(t) = 0 \) and thus the last term is indeed the trace of the boundary.

The proof for \( \partial(id,S)_\#(I) \) then follows from the fact that we can apply the result to \( \dot{S} := (id,S) \) to get

\[
\partial(id,S)_\#(I) = \partial \dot{S}_\#(I)
\]

\[
= \dot{S}(b) - \dot{S}(a) + (\partial \dot{S})_\#(I)
\]

\[
= \llbracket \{b\} \rrbracket \times S(b) - \llbracket \{a\} \rrbracket \times S(a) + (id,\partial S)_\#(I).
\]

\[ \square \]

Corollary 6.6. Let \( I = [a,b] \) be a closed interval and \( S : I \to \mathbb{I}_k(\Omega) \) an absolutely continuous curve with corresponding vector-field \( v(t) \), such that \( \partial S \) is independent of \( t \). Then for any \( t_0 \in I \) we have

\[
\langle (id,S)_\#(I), t < t_0 \rangle = \partial(id,S)_\#([a,t_0]) = \llbracket \{t_0\} \rrbracket \times S(t_0)
\]

in \( I \times \Omega \). Here \( \langle .,t < t_0 \rangle \) denotes slicing along a constant time.
Chapter 6. The Trace of a Curve

Lemma 6.7. Let $I$ be a closed interval and $S : I \to I_k(\Omega)$ an absolutely continuous curve with corresponding vector-field $v(t)$. Let furthermore $\pi : I \times \Omega \to \Omega; (t, x) \mapsto x$ be the projection on the space component. Then

$$\pi\#((id, S)_\#(I)) = S\#(I).$$

Proof. For any $k$-form $\omega \in C^\infty(\Omega; \wedge^k)$ we have that per definition

$$\pi\#((id, S)_\#(I))\langle \omega \rangle = ((id, S)_\#(I))\langle \pi\#\omega \rangle.$$

Now note that a calculation as in Proposition 3.5 gives us

$$i(1,v(t))\pi\#\omega = \pi\#i\pi\#(1,v(t))\omega = \pi\#iv(t)\omega.$$

But then

$$\int \{t\} \times S(t)(i(1,v(t))\pi\#\omega) = \int \{t\} \times S(t)(\pi\#iv(t)\omega) = \pi\#(\int \{t\} \times S(t))\langle iv(t)\omega \rangle = S\langle iv(t)\rangle.$$

Thus integrating over $t$ gives the desired result. \qed

6.2 Length and area of the trace

An important observation about the trace is its relation to the 1-length of the corresponding curve:

Lemma 6.8. Let $I$ be a closed interval and $S : I \to I_k(\Omega)$ an absolutely continuous curve with corresponding vector-field $v(t)$. Then

$$\mathcal{M}(S\#(I)) \leq \int_I \int |v| d\|S(t)\| dt = L_1(S)$$

and

$$\mathcal{M}(id, S)_\#(I) \leq \int_I \int \sqrt{1 + |v|^2} d\|S(t)\| dt$$

where in the last part we have equality if $v$ is orthogonal to $S(t)$ for almost all $t$ and $\mathcal{H}^k$-almost all $x \in \text{supp} S(t)$.

Proof. For the first statement take $\omega$ with $\|\omega(x)\| \leq 1$ for all $x$. Then clearly we have $\|i_v(\omega(x))\| \leq |v(t, x)|$ and so

$$S(t)\langle i_v\omega \rangle \leq \int |v(t, x)| d\|S(t)\|(x).$$

Now integrating in $t$ we get

$$S\#(I)\langle \omega \rangle \leq \int_0^1 \int |v(t, x)| d\|S(t)\|(x) dt.$$
and thus $\mathcal{M}(T) \leq L_1(S)$ by taking the supremum.

The second inequality then follows from the first by using the curve $(id, S) : t \mapsto \{(t)\} \times S(t)$. For equality, we note that if $v$ is orthogonal to $S(t)$ almost everywhere and $\xi(t,.)$ is the unit $k$-vector associated with $S(t)$, then there is a unit $k + 1$-form $\omega(t,x)$, such that $\langle \xi, i_{(1,v)} \omega \rangle = \langle (1,v) \wedge \xi, \omega \rangle = \langle (1,v) \rangle$, which is as regular as $v$ and $\xi$ are. The usual approximation by smooth forms then yields the equality, as they are dense. □

In light of this we also note that there is something similar to the arc-length parametrization for curves.

**Proposition 6.9.** Let $I = [a,b]$ be a closed interval and $S : I \to I_k(\Omega)$ an absolutely continuous curve with corresponding vector-field $v(t)$ and length $L_p(S) < \infty$. Furthermore let $\dot{S}(t) \neq 0$ for almost all $t \in I$. Then there is a (unique), positive oriented diffeomorphism $\phi : I \to I$, such that the reparametrization $R := S \circ \phi$ satisfies

$$L_p(R|_{[a,t]}) = |t - a| L_p(R) \quad \forall t \in I$$

and

$$L_p(R) \leq L_p(S \circ \psi)$$

for all diffeomorphisms $\psi : I \to I$. If $p = 1$, then the last inequality is replaced by the equality $L_1(R) = L_1(S \circ \psi)$.

**Proof.** The proof is similar to the arc-length parametrization for regular curves. We consider the continuous function

$$l_S : I \to I; \quad t \mapsto a + C \int_a^t \left( \int |v|^p d\|S(s)\| \right)^{\frac{1}{p-1}} ds$$

for a $C > 0$. By the fundamental theorem, $l_S$ is differentiable with

$$\frac{\partial l_S}{\partial t}(t) = C \left( \int |v|^p d\|S(t)\| \right)^{\frac{1}{p-1}}$$

almost everywhere. If $\dot{S}(t) \neq 0$ and $v(t)$ exists, then $v(t) \neq 0$. Thus $\frac{\partial l_S}{\partial t}(t) > 0$ for almost all $t$ and so $l_S$ is strictly increasing. Furthermore, we can choose $C$ such that $l_S$ is bijective.

This allows us to define $\psi : I \to I; s \mapsto (l_S)^{-1}(s)$.

Now let $\phi : I \to I$ be a positive oriented diffeomorphism. A calculation similar to those done in Lemma 6.4 reveals that $v\phi$ is the corresponding vector field for $S \circ \phi$ and thus

$$L_p(S \circ \phi) = \int_I |\phi(t)|^p \int |v(\phi(t))|^p d\|S(\phi(t))\| dt$$
As a first consequence, for $p = 1$ this is equal to $L_p(S)$. Secondly, if we use $\phi = \psi$, we note that the integrand is constant. Finally we note that if $\int |v(t)|^p \, d\|S(t)\|$ is constant, then the problem defaults to optimizing

$$
\int I \left| \frac{\partial (\phi^{-1})}{\partial t} (t) \right|^{1-p} \, dt
$$

under the constraint that $\int \frac{\partial (\phi^{-1})}{\partial t} \, dt = |I|$. A quick calculation then reveals that if $p > 1$, this is only minimal if $\phi$ is constant.

As seen in the next example however, even the arc-length parametrization is not necessarily optimal along curves with the same trace, even when there is no annihilation.

**Example 6.10.** Consider two curves, $\phi_1, \phi_2 : [0, 1] \to \mathbb{R}^2$ where

$$
\phi_1(t) := \begin{cases} 
(0, 2t), & 0 \leq t < 1/2 \\
(0, 1/2 + t), & 1/2 \leq t \leq 1
\end{cases} \quad \phi_2(t) := \begin{cases} 
(1, t), & 0 \leq t < 1/2 \\
(1, 2t - 1/2), & 1/2 \leq t \leq 1
\end{cases}
$$

Both curves are continuous and piecewise smooth, so we can consider the curve of currents given by

$$
S(t) := \left[ \{ \phi_1(t) \} \right] + \left[ \{ \phi_2(t) \} \right].
$$

A short calculation immediately reveals that $\dot{S}(t) + L_v S(t) = 0$ with

$$
v(t)(\phi_1(t)) = \begin{cases} 
(0, 2), & 0 \leq t < 1/2 \\
(0, 1), & 1/2 \leq t \leq 1
\end{cases} \quad v(t)(\phi_2(t)) = \begin{cases} 
(0, 1), & 0 \leq t < 1/2 \\
(0, 2), & 1/2 \leq t \leq 1
\end{cases}
$$

and thus $\int |v|^p \, d\|S(t)\| = 1^p + 2^p$ which is constant, so $S$ is parametrized by arc-length. We can also immediately see that

$$
S_\#([0, 1]) = \left[ \{0\} \times [0, 3/2] \right] + \left[ \{1\} \times [0, 3/2] \right].
$$

Yet for $p > 1$, the curve

$$
R(t) := \left[ \{(0, 3/2t)\} \right] + \left[ \{(1, 3/2t)\} \right]
$$

has the same trace and

$$
L_p(R) = \psi (3/2)^p + (3/2)^p < \psi (3/2)^p + 1^p = L_p(S)
$$

if $p > 1$.  

\[\square\]
6.3 Rectifiability

Hand in hand with the previous observation about the length of the trace, we have the following deeper observation about its rectifiability:

**Theorem 6.11.** Let $I$ be a closed interval and $S : I \to I_k(\Omega)$ an absolutely continuous curve with corresponding vector-field $v(t)$, such that $L_1(S) < \infty$. Then $S_#(I) \in I_{k+1}(\Omega)$ and $(id, S)_#(I) \in I_{k+1}(\mathbb{R} \times \Omega)$.

**Proof.** We will show the rectifiability of $(id, S)_#(I)$. The rectifiability of $S_#(I)$ then is an immediate consequence of Lemma 6.7 using Lemma B.12.

We proceed by induction over $k$. For this assume the statement is true for $k - 1$-currents, or that $k = 0$. We want to employ the slicing lemma B.16.

Pick $a := (a_0, a) \in I \times \Omega$. In order to prove the proposition we need to show that

$$\partial((id, S)_#(I) \ll B_r(a))$$

is rectifiable for almost all $r > 0$, then $(id, S)_#(I)$ is rectifiable.

For this we fix $a$ and $r$ and define

$$R := \partial((id, S)_#(I) \ll B_r(a))$$

and note that

$$R(\omega) = \int_{a_0-r}^{a_0+r} \left\{ \{t\} \times (S(t) \ll B_{\tilde{r}(t)}(a)) \langle i_{(1,v(t))} \omega \rangle dt \right\}$$

where $\tilde{r}(t) := \sqrt{r^2 - (a_0 - t)^2}$.

We want to approximate this current by currents we know to be rectifiable. Without loss of generality, assume by moving the domain, that $a_0 = 0$ and $a = 0$. For $m \in \mathbb{N}$, $j \in \mathbb{Z}$ with $|j| \leq m$ define $r^{(m)}(j) := \sqrt{r^2 - (j/m)^2}$ and

$$R_m : \omega \mapsto \sum_{j=-m}^{m-1} \left[ \left\{ \left\{ \{t\} \times (S(rj/m) \ll B_{r^{(m)}(j)}(0) \setminus B_{r^{(m)}(j)}(0)) \langle i_{(1,v(t))} \omega \rangle dt \right\} \times \partial \left( S(t) \ll B_{r^{(m)}(j)}(0) \right) \} \right]$$

The first term in the sum is rectifiable since $S(rj/m)$ is integer rectifiable. The second term vanishes in the case of $k = 0$. If $k > 0$ then we know that

$$\partial \left( S(t) \ll B_{r^{(m)}(j)}(0) \right) \in I_{k-1}(\Omega)$$
Figure 6.2: Main idea of the proof of Proposition 6.11: (a) We construct \((id, S)_{\#}(I)\) by integrating over the currents \([\{t\}] \times S(t)\). (b) We restrict this current to \(B_r(a)\). (c) To show rectifiability we approximate \(\partial((id, S)_{\#}(I) \cup B_r(a))\) by rectifiable currents \(R_m\). (d) This results in the error \(\tilde{T}_m\), which can be controlled.

Furthermore since Lie derivative and exterior derivative commute for differential forms, a short calculation reveals that

\[
\frac{d}{dt} \partial \left( S(t) \cup B_{r_{j+1}}(0) \right) = -\mathcal{L}_v(t) \partial \left( S(t) \cup B_{r_j}(0) \right).
\]

But then the second term is rectifiable by induction over \(k\). Thus in total, \(R_m\) is rectifiable.

Now we show that \(R_m \to R\), at least for a subsequence. For this, define

\[
\tilde{T}_m := \sum_{j=-m}^{m-1} \int_{r_{j/m}}^{r_{(j+1)/m}} \{t\} \times \left[ S(t) \cup \left( B_{r_j}(t) \setminus B_{r_j}(0) \right) \right] \langle i_{(1,v(t))} \omega \rangle dt
\]

Then \(\partial \tilde{T}_m = R - R_m\) and we have

\[
\mathcal{M}(\tilde{T}_m) \leq \sum_{j=-m}^{m-1} \int_{r_{j/m}}^{r_{(j+1)/m}} \int_{B_{r_j}(t) \setminus B_{r_j}(0)} \sqrt{1 + |v(t)|^2} dS(t) |dt|
\]
6.3. **RECTIFIABILITY**

\[ =: \int_{-r}^{r} f_{r}^{(m)}(t) \, dt \]

where \( f_{r}^{(m)}(t) := \int_{B_{r}(t)} B_{\tilde{r}}^{j(m)}(0) \sqrt{1 + |v(t)|^2} \|S(t)\| \) for the matching \( j \). Now for any \( m \in \mathbb{N} \), the function \( f_{r}^{(m)}(t) \) is non-negative and bounded by the integrable function \( t \mapsto \int \sqrt{1 + |v(t)|^2} \|S(t)\| \). Since \( 0 \leq f_{r}^{(2m)}(t) \leq f_{r}^{(m)}(t) \), a limit exists for the subsequence \( m = 2^p \). So by the bounded convergence theorem we have for \( f_{r}(t) := \lim_{p \to \infty} f_{r}^{(2^p)}(t) \) that

\[ \lim_{p \to \infty} F(R - R_{2^p}) \leq \lim_{p \to \infty} M(T_{2^p}) \leq \int_{-r}^{r} f_{r}(t) \, dt. \]

We want to show that \( \int_{-r}^{r} f_{r}(t) \, dt = 0 \) for almost all \( r \). For this note that \( f_{r}(t) > 0 \) implies

\[ \mathcal{M}(S(t) \vartriangle \partial B_{\tilde{r}}(t)(0)) > 0 \]

which can only happen for a countable number of \( r > 0 \) for each fixed \( t \) since \( S(t) \) is bounded in mass. But then by Fubini (\( f \) is measurable in \( t \) as a limit of measurable functions and in \( r \) since the integral over \( r \) is 0) we have

\[ \int_{0}^{\infty} \int_{-r}^{r} f_{r}(t) \, dt \, dr = \int_{-r}^{r} \int_{0}^{\infty} f_{r}(t) \, dr \, dt = 0 \]

and so \( \int_{-r}^{r} f_{r}(t) \, dt = 0 \) for almost all \( r > 0 \).

Hence \( R_{2^p} \to R \) for almost all \( r > 0 \), which means that \( R \) is rectifiable for almost all \( r > 0 \). Therefore \( (id, S)_{\#}(I) \) is rectifiable via the slicing lemma. \( \square \)

As an immediate consequence of the preceding results we get the first half of the equality between homogeneous Flat norm and the \( W_{1} \)-metric.

**Corollary 6.12.** Let \( T_{0}, T_{1} \in \hat{I}_{k}(\Omega) \). Then we have

\[ \hat{F}(T_{1} - T_{0}) \leq W_{1}(T_{0}, T_{1}). \]

**Proof.** Let \( \varepsilon > 0 \) and \( S : [0, 1] \to \hat{I}_{k}(\Omega) \) be a curve with \( L_{1}(S) \leq W_{1}(T_{0}, T_{1}) + \varepsilon \). Then by Lemma 6.5 we know that \( \partial S_{\#}(I) = T_{1} - T_{0} \) and by Theorem 6.11 \( S_{\#}(I) \) is rectifiable. Thus \( S_{\#}(I) \) is admissible in the definition of \( \hat{F}(T_{1} - T_{0}) \). But then using Lemma 6.8 we get

\[ \hat{F}(T_{1} - T_{0}) \leq \mathcal{M}(S_{\#}(I)) \leq L_{1}(S) \leq W_{1}(T_{0}, T_{1}) + \varepsilon \]

and the result follows from sending \( \varepsilon \) to zero. \( \square \)
Remark 6.13. Finally, now that we know that \((id, S)\#(I)\) is rectifiable, we can also identify the corresponding tangential space. Assume that

\[
S(t)\langle \omega \rangle = \int_{\Omega} \theta_t(x) \langle \xi_t(x), \omega(x) \rangle \, d\mathcal{H}^k(x)
\]

where \(\theta_t : \Omega \to \mathbb{Z}\) is the corresponding density function and \(\xi_t\) the simple unit \(k\)-vector corresponding to the local tangential plane. Then we can calculate

\[
(id, S)\#(I)\langle \omega \rangle = \int_I \int_{\Omega} \theta_t(x) \langle \xi_t(x), i(1, v(t, x))\omega(x) \rangle \, d\mathcal{H}^k(x)dt
\]

\[
= \int_I \int_{\Omega} \theta_t(x) \langle \xi_t(x) \wedge (1, v(t, x)), \omega(x) \rangle \, d\mathcal{H}^k(x)dt.
\]

From this we get that \(\xi_t(x) \wedge (1, v(t, x))\) is a multiple of the simple unit \(k\)-vector corresponding to the local tangential plane of \((id, S)\#(I)\). If \(v\) is orthogonal to \(\xi_t\), then we even get

\[
\int_I \int_{\Omega} \theta_t(x) \langle \xi_t(x) \wedge (1, v(t, x)), \omega(x) \rangle \, d\mathcal{H}^k(x)dt = \int_I \int_{\Omega} \theta_t(x) \sqrt{1 + |v|^2} \frac{\xi_t(x) \wedge (1, v(t, x))}{\sqrt{1 + |v|^2}}, \omega(x) \rangle \, d\mathcal{H}^k(x)dt.
\]

where \(\frac{\xi_t(x) \wedge (1, v(t, x))}{\sqrt{1 + |v|^2}}\) is an unit \(k + 1\)-vector. Thus in this case we have

\[
\theta_t(x) \sqrt{1 + |v|^2} \, d\mathcal{H}^k(x)dt = \theta(t, x) d\mathcal{H}^{k+1}(t, x)
\]

with \(\theta(t, x) \in \mathbb{Z}\).

Similar calculations can be done for \(S\#(I)\), however the resulting formulas are not that useful, since the tangential plane of \(S\#(I)\) can be generated by several \(\xi_t \wedge v\) for different times \(t\).
Figure 6.3: The principal idea in Remark 6.13. The tangential $k + 1$-vector $\xi$ of $(id, S)_{\#}(I)$ at $(t, x)$ is composed of the tangential $k$-vector $\xi_t$ of $S(t)$ at $x$ and the corresponding vector field $v$ at $(t, x)$. (here $k = 1$)
Chapter 7

An equivalence theorem

7.1 Homogeneous flat metric equals Wasserstein 1-distance

In the previous chapter, we have already seen that the homogeneous Flat metric can be bounded by the $W_1$-distance (Corollary 6.12). The aim of this section is to improve this result by showing that they are in fact equivalent. In the following we will write $W_1(.,.)$ for $W_{1,h_k}(.,.)$.

**Theorem 7.1.** Let $T_0$ and $T_1$ in $\dot{I}_k(\Omega)$. Then

$$\dot{F}(T_1 - T_0) = W_1(T_0, T_1).$$

The proof of the missing direction of this theorem is a bit more involved.

For this, we start with $T_0$ and $T_1 \in \dot{I}_k(\Omega)$ and $S \in I_{k+1}(\Omega)$ such that $\partial S = T_1 - T_0$. We now need to reverse of the construction that lead us to traces and find a matching homotopy $\tilde{S} \in C^1_{pw}([0, 1], \dot{I}_k(\Omega))$, such that $L_1(\tilde{S}) \leq \mathcal{M}(S)$. We will do so in three steps. First we will take a look at the simpler case where $S$ is a Lipschitz chain. This is found in Lemma 7.3. In the second step, the bulk of which is Lemma 7.5, we will use the deformation theorem to approximate an arbitrary $S$ by Lipschitz chains and some additional parts that are due to homotopies, which we can handle as well. However the deformation theorem then will result an additional multiplicative constant. So in the final step, we will then reduce this constant to one by showing that we only need to apply this lemma on an arbitrarily small part of $S$.

But first of all let us give another lemma that is a variant of the homotopy Lemma B.8, adapted to our purposes.

72
Lemma 7.2. Let $I = [0, 1]$, $h : I \times \Omega \to \mathbb{R}^n$ continuously differentiable, $T \in I_k(\Omega)$, $T_0 = h(0, \cdot)_\ast T$, $T_1 = h(1, \cdot)_\ast T$ and $S : I \to I_k(\Omega)$, $t \mapsto h(t, \cdot)_\ast T$.

Then $S_\#(\int I) = h_\#(\int I \times T)$ and thus 

$$
\mathcal{M}(h_\#(\int I \times T)) \leq L_1(S).
$$

Proof. First, we need to find a good vector field $v$ such that $\dot{S} + \mathcal{L}_v S = 0$. For this we note that according to Lemma 3.4 and the definition of the Lie-derivative

$$
\lim_{t \to t_0} \frac{S(t)\langle \omega \rangle - S(t_0)\langle \omega \rangle}{t - t_0} = \lim_{t \to t_0} T\left(\frac{h(t, \cdot)_\# \omega - h(t_0, \cdot)_\# \omega}{t - t_0}\right) = T\left(\mathcal{L}_{\dot{h}} h(t_0, \cdot)_\# \omega\right).
$$

So we can use $\frac{\partial h}{\partial t}(t, \cdot)$ as an associated (albeit not optimal) vector field.

We want to show that $S_\#(\int I) = h_\#(\int I \times T)$. For this, we note that for any $\omega \in C^\infty(\mathbb{R}^n; \wedge k)$, any point $x \in \Omega$, $t \in I$ and any simple $k$-vector $\xi = \xi_1 \wedge \ldots \wedge \xi_k$ at $x$ we have

$$
[h(t, \cdot)_\# \iota_{\frac{\partial h}{\partial t}(t, \cdot)}(\omega)](x) \langle \xi \rangle
= [\iota_{\frac{\partial h}{\partial t}(t, \cdot)}(h(t, x))] \left(\frac{\partial h(t, \cdot)}{\partial \xi_1}(x), \ldots, \frac{\partial h(t, \cdot)}{\partial \xi_k}(x)\right)
= \omega(h(t, x)) \left(\frac{\partial h}{\partial t}(t, x), \frac{\partial h}{\partial (0, \xi_1)}(t, x), \ldots, \frac{\partial h}{\partial (0, \xi_k)}(t, x)\right)
= [h_\# \omega](t, x) \langle \partial_t \otimes \xi(0, \xi_1), \ldots, 0, \xi_k) \rangle = [h_\# \omega](t, x) \langle \partial_t \otimes \xi \rangle.
$$

Using this we get

$$
S_\#(\int I) \langle \omega \rangle = \int_0^1 h(t, \cdot)_\# T \iota_{\frac{\partial h}{\partial t}(t, \cdot)}(\omega) \, dt
= \int_0^1 T \iota_{\frac{\partial h}{\partial t}(t, \cdot)}(\omega) \, dt
= \int_0^1 \int [h(t, \cdot)_\# \iota_{\frac{\partial h}{\partial t}(t, \cdot)}(\omega)](h(t, x)) \langle \xi(x) \rangle \, d\|T\| \langle x \rangle \, dt
= \int_0^1 \int [h_\# \omega](t, x) \langle \partial_t \otimes \xi(x) \rangle \, d\|T\| \langle x \rangle \, dt
= h_\#(\int I \times T) \langle \omega \rangle
$$

where $\xi$ is the $\mathcal{H}^k$-a.e. defined simple $k$-vector field associated with $T$.

The inequality $\mathcal{M}(h_\#(\int I \times T)) \leq L_1(S)$ then is a direct consequence of Lemma 6.8. \qed
Lemma 7.3. Let \( T_0, T_1 \in \dot{I}_k(\Omega) \), \( S \in I_{k+1}(\Omega) \) such that \( T_1 - T_0 = \partial S \) and \( S = f_\# P \), where \( f : \mathbb{R}^m \to \Omega \) is Lipschitz for some \( m \geq k + 1 \) and \( P \) is an integral polyhedral chain, that is \( P = \sum_{j \in J} \alpha_j \llbracket P_j \rrbracket \) with \( J \) finite, \( (\alpha_j)_{j \in J} \subset \mathbb{Z} \) and each \( P_j \) being an oriented \( k + 1 \)-simplex. Furthermore we require a weak notion of injectivity by assuming that \( H_{k+1}(P_i \cap P_j) = 0 \) and \( H_{k+1}(f(P_i) \cap f(P_j)) = 0 \) \( \forall i,j \in J, i \neq j \) that is, the simplices and their images may intersect but they may never overlap substantially. Then we have

\[
\mathcal{W}_1(T_0, T_1) \leq \mathcal{M}(S)
\]

Proof. By abuse of notation let us for now write \( P_j = \llbracket P_j \rrbracket \). We prove the lemma by induction over the size of \( J \). If \( J = \emptyset \), then \( S = 0 \) and thus \( T_0 = T_1 \) so the statement is trivial. Now let \( j \in J \). We need to show that

\[
\mathcal{W}_1(T_1, T_1 - \partial \alpha_j f_\# P_j) \leq \mathcal{M}(f_\# P_j)
\]  

(7.1)

since then we can use the triangle inequality to estimate by induction

\[
\mathcal{W}_1(T_1, T_0) \leq \mathcal{W}_1(T_1, T_1 - \partial \alpha_j f_\# P_j) + \mathcal{W}_1(T_1 - \partial \alpha_j f_\# P_j, T_0)
\]

\[
\leq \mathcal{M}(\alpha_j f_\# P_j) + \mathcal{M}(S - \alpha_j f_\# P_j)
\]

\[
= \mathcal{M}(S)
\]

where the last equality is a consequence of the injectivity condition.

Now to show (7.1). Without loss of generality, we can assume that

\[
P_j = \left\{ x \in \mathbb{R}^{k+1} \mid x_1 + \ldots + x_{k+1} \leq 1 \land x_i \geq 0 \forall i \in \{1, \ldots, k+1\} \right\}
\]

by prepending an additional Lipschitz map that regularizes the simplex and since we can ignore any additional dimensions because we will only deal with \( f|_{P_j} \) anyway. Furthermore we assume \( \alpha_j = 1 \), which is just a constant factor for all the formulas. Finally, we will take \( T_1 - T_0 = \partial f_\# P_j \) to simplify notation even more.

Now define

\[
P(t) := \{ x \in P_j \mid x_1 \leq t \}
\]

\[
Q(t) := \{ x \in P_j \mid x_1 = t \}
\]

\[
S(t) := T_1 - \partial f_\# \llbracket P(t) \rrbracket.
\]

Then \( S(0) = T_1 \) and \( S(1) = T_1 - \partial f_\# \llbracket P_j \rrbracket = T_0 \). A calculation similar to Lemma 7.2 reveals that

\[
\dot{S}(t) = \mathcal{L} \frac{\partial f_\#}{\partial x_1} Q(t) \quad \forall t \in (0, 1)
\]
7.1. **FLAT METRIC EQUALS 1-DISTANCE**

Moreover by Lemma 3.7 and its following remarks we can replace $\frac{\partial f}{\partial x_1}$ with $(\frac{\partial f}{\partial x_1})^\perp$, its part tangential to $f_\#Q(t)$ on an arbitrary large part of $f_\#Q(t)$. Then since $(\frac{\partial f}{\partial x_1})^\perp$ is defined to be orthogonal to all $\frac{\partial f}{\partial x_j}$ for $j > 1$, we have similar to the calculation in Lemma 7.2

$$(J_{x,f})^2 = \left|\det(Df^T)Df\right| = \left|\left(\frac{\partial f}{\partial x_1}\right)^\perp\right|^2 \left|\det A^T A\right| = \left|\left(\frac{\partial f}{\partial x_1}\right)^\perp\right|^2 (J_{x,f})^2,$$

where $A = \left(\frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_j}\right)$. Thus by applying the area formula B.1 once only on $Q(t)$ and then again in the other direction on all of $P_j$, we get

$$W_1(T_0, T_1) \leq L_1(S) = \int_0^1 \int \left|\left(\frac{\partial f}{\partial x_1}\right)^\perp\right| \|f_\#Q(t)\|dt$$

$$= \int_0^1 \int \left|\left(\frac{\partial f}{\partial x_1}\right)^\perp\right| \sqrt{|\det A^T A|} \|Q(t)\|dt$$

$$= \int_0^1 \|f_\#P_j\|(x) = \mathcal{M}(f_\#P_j)$$

Here we use the injectivity condition to assume that the multiplicity function $N(\cdot)$ used in the area formula is equal to 1 almost everywhere and hence can be ignored as a factor.

By applying the preceding lemma with $f = id$, we get the same statement in the special case of polyhedral chains:

**Corollary 7.4.** Let $T_0, T_1 \in \tilde{I}_k(\Omega)$ and $S \in I_{k+1}(\Omega)$ such that $T_1 - T_0 = \partial S$. If $S$ is an integral polyhedral chain, then

$$W_1(T_0, T_1) \leq \mathcal{M}(S).$$
Now we will use this in the second step by combining everything with the deformation theorem.

**Lemma 7.5.** Let $T_0, T_1 \in \mathcal{I}_k(\Omega)$. Then there is a constant $C$ independent of $T_0$ and $T_1$ such that

$$W_1(T_0, T_1) \leq C\hat{F}(T_1 - T_0).$$

Proof. Take $S \in \mathcal{I}_{k+1}(\Omega)$ with the usual properties that $T_0 - T_1 = \partial S$ and $\mathcal{M}(S) = \hat{F}(T_1 - T_0)$. Now we carefully need to apply the general construction used the deformation theorem [Fed69, 4.2.9] (In light of Remark 2.10) on $S$ to obtain the decomposition

$$S = P + Q + \partial A$$

(7.2)

where $P$ is an integer polyhedral chain consisting of $k+1$ cells of some triangulation, and $Q$ is constructed by pushing $\partial S$ onto the $k$-cells of this triangulation. We want to show that

$$W_1(T_0, T_1) \leq W_1(T_0, T_0 - \partial P) + W_1(T_0 - \partial P, T_1)$$

$$\leq c\mathcal{M}(S) + c\delta\mathcal{M}(\partial S)$$

$$\leq C\mathcal{M}(S) = C\hat{F}(T_0 - T_1)$$

where we use the triangle inequality, the bounds from the deformation theorem B.19 and finally chose $\delta \approx \frac{\mathcal{M}(S)}{\mathcal{M}(\partial S)}$ in order absorb the second term into our
constant. Note that $c$ does not depend on $\delta$ and so $C$ is independent of $S$. The missing step then is (*). For this, first note that (7.2) implies

$$T_0 - T_1 = \partial S = \partial P + \partial Q$$

and thus that $T_0 - (T_0 - \partial P) = \partial P$ and $(T_0 - \partial P) - T_1 = \partial Q$. So

$$W_1(T_0, T_0 - \partial P) \leq M(P) \leq cM(S)$$

is a direct consequence of Corollary 7.4.

To estimate the second term we need to have a closer look at how $Q$ is actually constructed. We create $P$ by pushing $S$ onto the $k + 1$ skeleton via a sequence of homotopies $h_i$. Then in order to match up the boundaries we look at how $\partial S$ is deformed by the same homotopies and trace this movement as $Q_i$. But by applying Lemma 7.2 to those homotopies, we get that the mass of $Q_i$ is bounded by the length of $t \mapsto h_i(t, .) \# \partial S$, which in turn is exactly the quantity estimated in the proof of the deformation theorem. (see for example [Fed69, 4.2.9]) Thus we end up with

$$W_1(T_0 - \partial P, T_1) \leq c\delta M(\partial S),$$

which proves the theorem.

Proof of Theorem 7.1. As mentioned before, one direction of the theorem is Corollary 6.12. For the other let $\varepsilon > 0$ and let $S$ be such that $M(S) = \mathcal{F}(T_1 - T_0)$, $T_1 - T_0 = \partial S$. According to [Mor08, Thm. 4.4] (see also [Fed69, 4.1.28]) there is a polyhedral chain $P$ and a Lipschitz function $f$ such that $M(S - f\#P) < \varepsilon$. Now using

$$\partial f\#P = T_1 - (T_1 - \partial f\#P)$$

and

$$\partial(S - f\#P) = (T_1 - \partial f\#P) - T_0$$

we can estimate by Lemma 7.3 (Note that one can always choose $f$ and $P$ in a way that does not violate our weak injectivity condition) and Lemma 7.5 respectively

$$W_1(T_0, T_1) \leq W_1(T_1, T_1 - \partial f\#P) + W_1(T_1 - \partial f\#P, T_0)$$

$$= M(f\#P) + C M(S - f\#P) < M(S) + C\varepsilon$$

which for $\varepsilon \to 0$ proves the theorem.
7.2 Generalizations, corollaries and remarks

We have already seen in Chapter 5 that we need a bound on the mass of a current in order to avoid degeneracy. The construction in the last section however is incompatible with this. Still we get the following corollary of the proof of Theorem 7.1:

Corollary 7.6. Let $A_c := \{ T \in \dot{I}_k(\Omega) \mid M(T) < c \}$, $T_0$ and $T_1 \in \dot{I}_k(\Omega)$. If $\tilde{\mathcal{F}}(T_0, T_1) < \infty$ then there is a $c \in \mathbb{R}$ such that $\mathcal{W}_{p, A_c}(T_0, T_1) < \infty$ for all $p \in [1, \infty)$.

Proof. This lemma is a result of following the construction in the proof of Lemma 7.5. Any homotopies used therein only use bounded vector fields $v(t)$, which are therefore $L^p$-integrable. Furthermore the mass always stays bounded, so the corollary holds.

Finally let us extend our previous remark (2.12) on dualities with an observation on the Wasserstein distance.

Remark 7.7. As a continuation of Remark 2.12, if we look at the Wasserstein distance, we also have a dual formulation (see for example [Vil09, chp.5]), given by

$$W_1(\mu, \nu) = \sup_{\|\nabla \phi\|_{\infty} \leq 1} \left[ \int \phi d\mu - \int \phi d\nu \right].$$

It should come as no surprise that $\tilde{\mathcal{F}}$ and this notion are identical when they overlap for $k = 0$. But as $\mathcal{F}$ and $\tilde{\mathcal{F}}$ differ, we cannot assume a dual formulation of the Wasserstein-distance for integer rectifiable currents of higher dimension. The same also holds for the generalized $p$-distance.
Chapter 8

Comparing the vertical Wasserstein distance with the $L^p$ distance

8.1 The vector space structure in relation to $L^p$

The main goal of this chapter will be the comparison of vertical Wasserstein distance and the $L^p$ distance. However first, let us reflect a bit more on the different vector space structures involved. So even though our final goal will be maps to manifolds, for now let us first focus on functions whose image is in the space $\mathbb{R}^N$.

In this case the space of functions has the usual vector space structure. If we then go to the corresponding Cartesian currents, we find that as currents, of course they also have a vector space structure, yet one which is completely different. It is easy to see that $\lambda G(u) \neq G(\lambda u)$ and $G(u) + G(v) \neq G(u + v)$. In fact $\lambda G(u)$ (except for $\lambda = 1$) and $G(u) + G(v)$ are not even Cartesian currents. Instead of this naive approach however, we can do the following:

**Remark 8.1.** Let $u, v$ such that $G(u), G(v) \in \text{cart}(\Omega; \mathbb{R}^N)$. Then

1. For any $\lambda \in \mathbb{R}$, we have $G(\lambda u) = (m_\lambda) \# G(u)$, where $m_\lambda : (x, y) \mapsto (x, \lambda y)$.

2. If $v$ is sufficiently smooth, then $G(u + v) = (a_v) \# G(u)$, where $a_v : (x, y) \mapsto (x, y + v(x))$.

While those operations only seem to complicate a simple situation, we can however see that the Wasserstein distance behaves as expected when used in conjunction with them.

**Lemma 8.2.** Let $T_0, T_1 \in \text{cart}(\Omega; \mathbb{R}^N)$. Then we have:
1. For any \( \lambda \in \mathbb{R} \),

\[
W_{p,\text{vert}}((m_\lambda)_\# T_0, (m_\lambda)_\# T_1) = |\lambda| W_{p,\text{vert}}(T_0, T_1),
\]

where as above \( m_\lambda : (x, y) \mapsto (x, \lambda y) \).

2. If \( u : \Omega \to \mathbb{R}^N \) is sufficiently smooth, then

\[
W_{p,\text{vert}}((a_u)_\# T_0, (a_u)_\# T_1) = W_{p,\text{vert}}(T_0, T_1),
\]

where as above \( a_u : (x, y) \mapsto (x, y + u(x)) \).

Proof. 1. Let \( S \in C^1_{pw}([0, 1]; \text{cart}(\Omega; \mathcal{M})) \) be a homotopy such that \( S(0) = T_0 \) and \( S(1) = T_1 \). Define

\[
\tilde{S} : t \mapsto (m_\lambda)_\# S(t)
\]

Then \( \tilde{S}(0) = (m_\lambda)_\# T_0 \) and \( \tilde{S}(1) = (m_\lambda)_\# T_1 \). Furthermore, for any \( t \in [0, 1] \) for which there is a vector field \( v : \Omega \to \mathbb{R}^N \) such that \( \dot{S}(t) + L_{(0, v)}S(t) = 0 \), using Proposition 3.5 there also holds

\[
\frac{\partial}{\partial t} \tilde{S}(t) + L_{(0, v)}(m_\lambda)_\# \tilde{S}(t) = (m_\lambda)_\# \left( \dot{S}(t) + L_{(0, v)}S(t) \right) = 0
\]

and similarly if \( \frac{\partial}{\partial t} \tilde{S}(t) + L_{(0, w)}\tilde{S}(t) = 0 \) for some vector field \( \Omega \to \mathbb{R}^N \), then \( \dot{\tilde{S}}(t) + L_{(0, w/\lambda)}S(t) = 0 \).

But then we have

\[
\int_0^1 \inf \left\{ \int_\Omega |v|^p \, dx \, \middle| \, v \in C^1_{pw}(\Omega; \mathbb{R}^N), \frac{\partial}{\partial t} \tilde{S}(t) + L_{(0, w(t))}\tilde{S}(t) = 0 \right\} dt
\]

\[
= \int_0^1 \left\{ \int_\Omega \left| \frac{\lambda}{n} \right|^p \, dx \, \middle| \, v \in C^1_{pw}(\Omega; \mathbb{R}^N), \dot{\tilde{S}}(t) + L_{(0, w)}S(t) = 0 \right\} dt
\]

or in other words

\[
L_{p,\text{vert}}(\tilde{S}) = |\lambda| L_{p,\text{vert}}(S)
\]

Taking the infimum over all \( S \) we get

\[
W_{p,\text{vert}}((m_\lambda)_\# T_0, (m_\lambda)_\# T_1) \leq \lambda W_{p,\text{vert}}(T_0, T_1).
\]

If \( \lambda = 0 \) both sides have to be 0 and this already is our result. If \( \lambda \neq 0 \) we simply take \( \lambda = \lambda^{-1} \) to obtain the opposite inequality.
2. Similar to the first part let us assume that \( S \in C^1_{pw}([0,1]; \text{cart}(\Omega;\mathcal{M})) \) is a homotopy with \( S(0) = T_0 \) and \( S(1) = T_1 \). Now let

\[
\tilde{S} : t \mapsto (a_u)_\# S(t)
\]

Then \( \tilde{S}(0) = (a_u)_\# T_0 \) and \( \tilde{S}(1) = (a_u)_\# T_1 \). Also as above we have

\[
\dot{\tilde{S}}(t) + \mathcal{L}_{(0,v)}\tilde{S}(t) = (a_u)_\# \dot{S}(t) + \mathcal{L}_{(a_u)_\# (0,v)}(a_u)_\# S(t)
= (a_u)_\# \left( \dot{S}(t) + \mathcal{L}_{(0,v)} S(t) \right) = 0.
\]

From this we can again easily conclude that \( L_{p,\text{vert}}(\tilde{S}) = L_{p,\text{vert}}(S) \) and thus taking the infimum yields

\[
W_{p,\text{vert}}((a_u)_\# T_0, (a_u)_\# T_1) = W_{p,\text{vert}}(T_0, T_1).
\]

8.2 Upper bounds

The basic intuition on why \( W_{p,\text{vert}} \) corresponds to the \( L^p \) distance is rather simple, if one first looks at the case of smooth functions. If we take some homotopy \( S(t) = G(u(t,x)) \) with \( \dot{S} + \mathcal{L}_{(0,v)} S = 0 \), then according to Lemma 5.9 the vector field \( v \) is nothing more than the derivative \( \frac{\partial u}{\partial t} (t,x) \). The problem is to minimize the integral over \( |v|^p \) among all such homotopies with fixed end points. For a fixed \( x \), this translates to minimizing \( \int_0^1 \left| \frac{\partial u}{\partial t} (t,x) \right|^p dt \) with respect to \( u(0,x) = u_0(x) \) and \( u(1,x) = u_1(x) \).

It is well known that the solutions to this are the minimal geodesics between \( u_0(x) \) and \( u_1(x) \). For them we end up with

\[
\int_0^1 \left| \frac{\partial u}{\partial t} (t,x) \right|^p dt = d(u_0(x), u_1(x))^p
\]

which we then just would have to integrate over \( x \) to finish the proof. There is however a slight difficulty in this. One of the fundamental properties of Cartesian currents is that they enforce a kind of coherence. While they allow for jumps, any such discontinuity will be remembered in vertical parts. And even if we start and end with smooth functions \( u_0 \) and \( u_1 \), there is no reason to assume that the geodesics between \( u_0(x) \) and \( u_1(x) \) will depend smoothly on \( x \). So we will need to tread carefully there in order to avoid additional vertical parts. In fact, as will be seen in Example 8.11, this turns out to be the crucial difference and the feature that makes \( W_{p,\text{vert}} \) more suited to dealing with topological singularities than the \( L^p \)-distance.

For now, however, let us focus a bit on the smooth case, to show upper bounds on \( W_{p,\text{vert}} \) by explicitly constructing \( S(t) \).
Lemma 8.3. Let $u_0, u_1 \in C^1(\Omega; \mathcal{M})$ be functions such that $G(u_0), G(u_1) \in \text{cart}(\Omega; \mathcal{M})$. Let $u(t, x) : [0, 1] \times \Omega \to \mathcal{M}$ be a continuously differentiable function such that $u(0, \cdot) = u_0$ and $u(1, \cdot) = u_1$. Then

$$W_{p, \text{vert}}(G(u_0), G(u_1)) \leq \left( \int_{\Omega} \int_0^1 \left| \frac{\partial u}{\partial t}(t, x) \right|^p \, dt \, dx \right)^{1/p}. $$

Proof. Let $S(t) := G(u_t)$. Then according to Lemma 5.9 we get

$$\dot{S}(t) \omega = -\mathcal{L}(0, \partial u/\partial t) S(t) \omega.$$ 

This identifies $\partial u/\partial t$ as a vector field by which $S$ is moved and thus we can gather from the definitions that

$$W_{p, \text{vert}}(G(u_0), G(u_1))^p \leq L_{p, \text{vert}}(S)^p \leq \int_0^1 \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^p \, dx \, dt. \quad \square$$

Corollary 8.4. Let $u_0, u_1$ be continuous functions such that both $G(u_0)$ and $G(u_1) \in \text{cart}(\Omega; \mathcal{M})$. If there is a unique shortest geodesic between $u_0(x)$ and $u_1(x)$ for all $x \in \Omega$, then

$$W_{p, \text{vert}}(G(u_0), G(u_1)) \leq d_{L^p}(u_0, u_1).$$

In particular if $\mathcal{M} = \mathbb{R}^N$, then we always have

$$W_{p, \text{vert}}(G(u_0), G(u_1)) \leq \|u_1 - u_0\|_{L^p}.$$ 

Proof. The first statement is a direct consequence of Lemma 8.3. From differential geometry we know, that if there exists a unique shortest geodesic $\gamma_{a,b} : [0,1] \to \mathcal{M}$ between two points $a$ and $b$, then in a small neighborhood, $\gamma_{a,b}$ depends smoothly on $a$ and $b$. So since $u_0$ and $u_1$ are continuous, $\gamma_{u_0(x),u_1(x)}$ is continuous in $x$ as well, which is all we need. We then apply Jensen’s inequality to see that

$$\int_0^1 |\dot{\gamma}(t)|^p \, dt = \left( \int_0^1 |\dot{\gamma}(t)| \, dt \right)^p = d(\gamma(0), \gamma(1))^p$$

for any shortest geodesic $\gamma$. Here equality holds since $|\dot{\gamma}|$ is constant.

The second statement is shown by construction. In this case, the geodesics are given explicitly by

$$u : (t, x) \mapsto tu_1(x) + (1 - t)u_0(x). \quad \square$$

Remark 8.5. The condition of needing a unique geodesic in the preceding corollary can of course easily be replaced by having a family of geodesics $\gamma_x : [0,1] \to \mathcal{M}$ with $\gamma_x(0) = u_0(x)$ and $\gamma_x(1) = u_1(x)$, which smoothly (or
even just continuously) depends on the point $x$. This gives a slightly weaker condition.

It is however not possible to extend this to a discontinuous family. As soon as it is possible for the shortest geodesic to “jump”, we get into the coherence issues mentioned beforehand. Again the problem is best seen in Example 8.11. ■

8.3 Lower bounds

The construction of a lower bound is a bit more involved. Since we take the infimum over all homotopies, we cannot assume that all currents correspond to smoothly varying graphs of functions anymore. However, we know that Cartesian currents can be separated into the graph of a function and a vertical part. In the end, it turns out that we can ignore the vertical parts of the current, since they do not contribute to the distance. Still, we need to be a bit careful, since even the graph part is not varying smoothly but may jump in some places. The main idea here is to show that those situations are limited to a set of measure zero.

Since the case of Cartesian currents in $\mathbb{R}^n$ is somewhat simpler, we will start there.

Lemma 8.6. Let $u_0, u_1$ be continuous functions such that $\mathcal{G}(u_0), \mathcal{G}(u_1) \in \text{cart}(\Omega; \mathcal{M})$ where $\mathcal{M} \subset \mathbb{R}^N$. Then

$$\|u_1 - u_0\|_{L^p} \leq W_{p,\text{vert}}(\mathcal{G}(u_0), \mathcal{G}(u_1))$$

where $u_0$ and $u_1$ are taken as functions from $\Omega$ to $\mathbb{R}^N$.

Proof. We observe that according to [GMS98, I.4.2.3, Thm 1]

$$\int_{\Omega} ((u_1(x))_i - (u_0(x))_i) \phi(x) dx = \mathcal{G}(u_1)\langle y_i \phi(x) dx \rangle - \mathcal{G}(u_0)\langle y_i \phi(x) dx \rangle.$$

Now let $S(t) \in C^1_{\text{pw}}([0, 1]; \text{cart}(\Omega; \mathcal{M}))$ be a homotopy such that $S(0) = \mathcal{G}(u_0)$ and $S(1) = \mathcal{G}(u_1)$. According to Proposition 4.12, for all $t$ there is a function $u(t)$ and a vertical current $V(t)$ such that $S(t) = \mathcal{G}(u(t)) + V_t$. Furthermore assume that there is the usual vector field $v(t)$ such that $S(t) + \mathcal{L}_{(0,v(t))}S(t) = 0$ for almost all $t$. From the fact that $i_{(0,v(t))}dx = 0$ and $d(\phi(x) dx) = 0$, due to $dx^k \wedge dx = 0$, we conclude that

$$\mathcal{L}_{(0,v(t))}y_i \phi(x) dx = di_{(0,v(t))} (y_i \phi(x) dx) + i_{(0,v(t))}d (y_i \phi(x) dx) = i_{(0,v(t))} (\phi(x)dy_i \wedge dx) = v_i(t)(x)\phi(x) dx.$$

So using this, we end up with

$$\mathcal{G}(u_1)\langle y_i \phi(x) dx \rangle - \mathcal{G}(u_0)\langle y_i \phi(x) dx \rangle$$
\[ \frac{\partial}{\partial x} (u(t)i(x)) - (u_0)i(x) = \int_0^t v_i(t,x) \, dt. \]

If we now look at the absolute value we get that
\[ |u_1(x) - u_0(x)| = \left| \int_0^1 v(t,x) \, dt \right| \leq \int_0^1 |v(t,x)| \, dt \]
and then using Jensen’s inequality this results in
\[ |u_1(x) - u_0(x)|^p = \left( \int_0^1 |v(t,x)| \, dt \right)^p \leq \int_0^1 |v(t,x)|^p \, dt. \]

Integrating this over \( x \) we end up with
\[ \int_\Omega |u_1(x) - u_0(x)|^p \, dx \leq \int_\Omega \int_0^1 |v(t,x)|^p \, dt \, dx \]
which taking the infimum over all choices of \( v \) and over all homotopies \( S \) yields
\[ \|u_1 - u_0\|_{L^p} \leq \mathcal{W}_{p,\text{vert}}(\mathcal{G}(0), \mathcal{G}(u_1)). \]

Note that equality in the previous section can only be expected if \( v(\cdot,x) \) is constant almost everywhere, for almost all \( x \). This is exactly the case for a linear homotopy between functions in \( \mathbb{R}^N \). Thus we have:
Corollary 8.7. Let \( u_0, u_1 \) be continuous functions such that both \( G(u_0) \) and \( G(u_1) \) belong to \( \text{cart}(\Omega; \mathbb{R}^N) \). Then
\[
\|u_1 - u_0\|_{L^p} = W_{p,\text{vert}}(G(u_0), G(u_1)).
\]

Proof. This is a direct consequence of Lemma 8.6 and Corollary 8.4.

Before we get to the actual case of a lower bound for the case where the image is a manifold, we will need the following technical lemma to approximate the geodesic distance independent of the actual starting point.

Lemma 8.8. Let \( M \) be a compact, embedded manifold. For any \( \varepsilon > 0 \) there exists a \( \delta = \delta(M, \varepsilon) > 0 \) such that
\[
\inf_{(a_i) \in P_\delta(x,y)} \sum_{i=1}^k |a_i - a_{i-1}| > d(x, y) - \varepsilon
\]
for all \( x, y \in M \). Here
\[
P_\delta(x, y) = \{(a_0, ..., a_k) \mid \forall 1 \leq i \leq k : a_i \in M, |a_i - a_{i-1}| < \delta, a_0 = x, a_k = y\}
\]
denotes the set of approximate paths between \( x \) and \( y \).

Proof. Let us fix \( x_0, y_0 \in M \). As a possible characterization of geodesic distance, we have
\[
d(x_0, y_0) = \lim_{\delta \to 0} \inf_{(a_i) \in P_\delta(x_0,y_0)} \sum_{i=1}^k |a_i - a_{i-1}|.
\]

Therefore there is \( \delta_0 = \delta_0(x_0, y_0) > 0 \) such that
\[
\inf_{(a_i) \in P_{\delta_0}(x_0,y_0)} \sum_{i=1}^k |a_i - a_{i-1}| > d(x_0, y_0) - \varepsilon/2.
\]

Let \( x, y \in M \) such that \( |x - x_0| < \min(\varepsilon/4, \delta_0) \) and \( |y - y_0| < \min(\varepsilon/4, \delta_0) \). Then for any \( (a_i) \in \{0, ..., k\} \in P_{\delta_0}(x, y) \), we have \( (x_0, a_0, ..., a_k, y_0) \in P_{\delta_0}(x_0, y_0) \) and so
\[
\sum_{i=1}^k |a_i - a_{i-1}| = \sum_{i=1}^k |a_i - a_{i-1}| + |x_0 - x| + |y_0 - y| - |x_0 - x| - |y_0 - y| > d(x, y) - \varepsilon/2 - |x_0 - x| - |y_0 - y| > d(x, y) - \varepsilon.
\]
Taking the infimum over $P_{\delta_0}(x, y)$ we get
\[
\inf_{(a_i)_{i \in P_{\delta_0}}(x, y)} \sum_{i=1}^{k} |a_i - a_{i-1}| > d(x, y) - \varepsilon
\]
for all $x \in B_{\min(\varepsilon/4, \delta_0)}(x_0)$ and $y \in B_{\min(\varepsilon/4, \delta_0)}(y_0)$. Varying $x_0$ and $y_0$ now gives us an open cover of the compact set $\mathcal{M} \times \mathcal{M}$, from which we take a finite sub-cover with centers $(x_i, y_i)_{i \in \{1, \ldots, k\}}$. Then $\delta = \min_{i \in \{1, \ldots, k\}} \delta_0(x_i, y_i)$ has the required property.

Proposition 8.9. Let $\mathcal{M}$ be compact and $u_0, u_1$ be continuous functions such that both $G(u_0)$ and $G(u_1) \in \text{cart}(\Omega; \mathcal{M})$. Then we can estimate the (geodesic) $L^p$ distance by
\[
d_{L^p}(u_1, u_0) \leq W_{p,\text{vert}}(G(u_0), G(u_1)).
\]

Proof. We prove this by contradiction. Assume that there is a curve of currents $S \in C_{pw}([0,1]; \text{cart}(\Omega; \mathcal{M}))$ such that $L^p_{\text{vert}}(S) < d_{L^p}(u_0, u_1)$. Define the vector field $v_t(x)$ such that $\dot{S}(t) + L(0, v_t)(t) = 0$ and still
\[
\int_0^1 \int_\Omega |v_t|^p \, dx \, dt < d_{L^p}(u_0, u_1)^p = \int_\Omega d(u_0(x), u_1(x))^p \, dx.
\]
Then there is an $\varepsilon > 0$ such that the set of "bad" points
\[
B := \left\{ x \in \Omega \left| \left( \int_0^1 |v_t(x)|^p \, dt \right)^{1/p} < d(u_0(x), u_1(x)) - \varepsilon \right. \right\}
\]
has positive measure.

Let $\delta > 0$ such that
\[
\inf \left\{ \sum_{i=1}^{k} |a_i - a_{i-1}| \left| (a_i)_{i \in P_{\delta}(u_0(x), u_1(x))} \right| > d(u_0(x), u_1(x)) - \varepsilon \right. \right\}
\]
for all $x \in B$, as given by Lemma 8.8.

We note that for almost all $x \in B$ there is a minimal number $N = N(x) \in \mathbb{N}$ such that
\[
\int_{i/2^N}^{(i+1)/2^N} |v(t, x)|^p \, dt < \delta^p \quad \text{for all } i \in \{0, \ldots, 2^N - 1\}.
\]

Furthermore, since $v(t, x)$ is a measurable function, for any $N \in \mathbb{N}$, $i \in \{0, \ldots, 2^N - 1\}$ the set
\[
G_{i, N} := \left\{ x \in \Omega \left| \int_{i/2^N}^{(i+1)/2^N} |v(t, x)|^p \, dt < \delta^p \right. \right\}
\]
is measurable. Since for any $N_0 \in \mathbb{N}$ we have

$$
\{x \in \Omega \mid N(x) \leq N_0\} = \bigcup_{i=0}^{2^{N_0} - 1} G_{i,N_i}
$$

the function $N(x)$ is measurable. Thus since $N$ is finite almost everywhere, for any $\varepsilon_0 > 0$ there is an $N_0 \in \mathbb{N}$ such that

$$
\mathcal{H}^n(\{x \in \Omega \mid N(x) > N_0\}) < \varepsilon_0.
$$

Define $G := \{x \in \Omega \mid N(x) \leq N_0\}$ as the set of good points. For each $i \in \{0, ..., 2^{N_0} - 1\}$ define $t_i := i/2^N$. Then the current $S(t_i)$ has the form

$$
S(t_i) = G(u_i) + V(t_i).
$$

From Lemma 8.6 we know that for almost all $x \in \Omega$ we have

$$
|u_i(x) - u_{i+1}(x)|^p \leq \int_{t_i}^{t_{i+1}} |v(t,x)|^p \, dt.
$$

But then for any $x \in G$, we have

$$
|u_i(x) - u_{i+1}(x)| < \delta
$$

for all $i \in \{0, ..., 2^{N_0} - 1\}$ and so

$$
(u_0(x), u_{t_1}(x), ..., u_1(x)) \in P_\delta(u_0(x), u_1(x))
$$

for all $x \in G$ and thus

$$
d(u_0(x), u_1(x)) - \varepsilon < \sum_{i=1}^{2^N} |u_{t_i} - u_{t_{i-1}}|
$$

Now since $B$ has positive measure, $G$ can be chosen so large that there is an $x \in B \cap G$. This, however, gives a contradiction as we then have

$$
d(u_0(x), u_1(x)) - \varepsilon < \sum_{i=1}^{2^N} |u_{t_i} - u_{t_{i-1}}| \\
\leq \left(\sum_{i=1}^{2^N} |u_{t_i} - u_{t_{i-1}}|^p\right)^{1/p} \\
\leq \left(\sum_{i=1}^{2^N} \int_{t_i}^{t_{i+1}} |v(t,x)|^p \, dt\right)^{1/p} \\
< d(u_0(x), u_1(x)) - \varepsilon.
$$
Theorem 8.10. Let $u_0, u_1$ be continuous functions such that $G(u_0), G(u_1) \in \text{cart}(\Omega; M)$. If the shortest geodesic between $u_0(x)$ and $u_1(x)$ depends continuously on $x$, then

$$W_{p,\text{vert}}(G(u_0), G(u_1)) = d_{L^p}(u_0, u_1).$$

Proof. This is a direct consequence of Corollary 8.4 for the upper bound and Proposition 8.9 for the lower bound.

As the last part of this section, let us give an example that shows that the two notions of distance still differ if the two functions are not connected by small geodesics. While an obvious idea would be generating an example by just adding a vertical part to a fixed function, there are more illustrative examples that do not involve vertical parts:

![Illustration of Example 8.11](image)

Figure 8.1: Illustration of Example 8.11. In both cases, we unroll $S^1$ in the vertical direction, i.e. the top of each illustration continues at the bottom. The distance between the two functions is illustrated by the hatched area. In (a) we consider the shortest distance for each $x$ individually, while in (b) the direction in which distance is measured needs to be coherent.

Example 8.11. We consider functions from $I := [-1, 1]$ to $S^1$. Let $u_0$ be the constant function $u_0 : x \to (1, 0)$ and consider

$$u_1 : x \mapsto \begin{cases} (\cos(\phi(x)), \sin(\phi(x))) & \text{for } |x| < \varepsilon \\ (\varepsilon, 0) & \text{otherwise.} \end{cases}$$

Here $\phi$ is a smooth function with $\phi(-\varepsilon) = 0$, $\phi(\varepsilon) = 2\pi$. We can also easily assume all of $u_1$ to be smooth. In the limit of $\varepsilon \to 0$ we then have $G(u_1) \to G(u_0) + \{0\} \times [S^1]$.

Now we can see that the shortest distance between $u_0(x)$ and $u_1(x)$ is given by

$$d(u_0(x), u_1(x)) = \text{dist}(\phi(x), 2\pi\mathbb{Z})$$

for $|x| < \varepsilon$. 


and 0 otherwise. But then
\[
     d_{L^p}(u_0 - u_1) = \left( \int_{-\varepsilon}^{0} \text{dist}(\phi(x), 2\pi \mathbb{Z})^p dx \right)^{1/p} \leq (2\varepsilon \pi^p)^{1/p} = (2\varepsilon)^{1/p} \pi.
\]

So for small \(\varepsilon\), the distance between the two functions is rather small. Contrast this to the Wasserstein distance. A possible homotopy here is given by the function
\[
     u_t : x \mapsto \begin{cases} 
     (1,0) & \text{for } x \leq -\varepsilon \\
     (\cos(t\phi(x)), \sin(t\phi(x))) & \text{for } |x| < \varepsilon \\
     (\cos(2\pi t), \sin(2\pi t)) & \text{for } x \geq \varepsilon.
     \end{cases}
\]

But then for the corresponding homotopy of currents \(S(t) := G(u_t)\) we end up with \(\dot{S}(t) = -L(0,v_t)S(t)\) with
\[
     v_t(x) = \frac{\partial u_t}{\partial t} = \begin{cases} 
     (0,0) & \text{for } x \leq -\varepsilon \\
     \phi(x)(-\sin(t\phi(x)), \cos(t\phi(x))) & \text{for } |x| < \varepsilon \\
     (-\sin(2\pi t), \cos(2\pi t)) & \text{for } x \geq \varepsilon.
     \end{cases}
\]

This leads to
\[
     L_{p,\text{vert}}(S) = \left( \int_{-\varepsilon}^{\varepsilon} |\phi(x)|^p dx + \int_{\varepsilon}^{1} (2\pi)^p dx \right)^{1/p} > (1 - 2\varepsilon)^{1/p} 2\pi
\]

which for small \(\varepsilon\) converges to \(2\pi\).

A nice way to see that this actually is the optimal choice of \(S\) is the following: Assume that there is a homotopy \(S\) with \(S(0) = G(u_0), S(1) = G(u_1)\) and \(L_{p,\text{vert}}(S) < (1 - \varepsilon)^{1/p} 2\pi\). Let \(v(t,\cdot)\) be the corresponding vector field. Then there is a small \(\delta > 0\), an interval \([a_1, b_1] \subset [-1,-\varepsilon]\) and an interval \([a_2, b_2] \subset [\varepsilon, 1]\) such that \(\int_0^1 |v|^p dt < (2\pi - \delta)^p\) for all \(x \in G_1 \cup G_2\).

What this means, in light of the other results in this section, is that on the good intervals \(G_1\) and \(G_2\), the current \(S(t)\) has no support close to \(I \times (-1,0)\), since this would require moving a distance of \(\pi\) to get there and then a second movement of \(\pi\) to get back to \((1,0)\).

But then the endpoints are essentially fixed. What is left now is degree theory. It is now easy to modify \(S(t)\) such that \(S(t) \perp (-1,a_1) \times \mathbb{S}^1 = G((1,0))\) and \(S(t) \perp (b_2,1) \times \mathbb{S}^1 = G((1,0))\). But then the degrees of \(S(0)\) and \(S(1)\) do not match, as \(\pi_y \# S(t)\) is continuous and without boundary, yet \(\pi_y \# S(0) = \mathbb{S}^1 \neq 0 = \pi_y \# S(1)\), which is a contradiction. ■
Chapter 9

Limits of curves, lower semicontinuity

9.1 Nonexistence of minimizers

Since our distances are all defined via an infimum over the length of curves, an immediate question should be about the existence of a minimizing curve. Sadly, in most cases the answer turns out to be that there is none. The general problem is related to the degeneracy issues we have seen in Example 5.4 and Example 5.11.

Even in the case \( p = 1 \) where this does not occur, the situation is unclear. We know that \( \bar{F}(T_0 - T_1) \) has a minimizer in the form of a current \( S \) such that \( \partial S = T_0 - T_1 \), yet we are only able to approximate \( S \) by a homotopy up to an arbitrary small error in length (see the proof of Theorem 7.1). We will however see the existence of a minimizer in this case, under the additional condition of a mass bound in Corollary 9.6.

Other than that, only cases we know the existence of a minimizer for the vertical Wasserstein distance are the trivial ones. For example any \( T \in \text{cart}^{2,1}(\Omega; \mathbb{R}^2) \) for \( \Omega \subset \mathbb{R}^2 \), the current is always of the form \( T = \mathcal{G}(u) \) for \( u \in W^{1,2}(\Omega; \mathbb{R}^2) \) (see Remark 4.24). But then as a result, using the reasoning in Chapter 8, the distance \( \mathcal{W}_{p,\text{cart}^{2,1, \text{vert}}} \) is identical to the \( L^p \) distance and the minimizing homotopies are linear interpolations.

For most of the other cases there are no minimizers, since the length-functional is not lower-semicontinuous, as the following example illustrates in the non-vertical case.

\textbf{Example 9.1.} Fix \( a \geq 0 \) and consider the function

\[
  f(x) := \begin{cases} 
  ax & \text{for } 0 \leq x \leq \frac{1}{2} \\
  a - ax & \text{for } \frac{1}{2} \leq x \leq 1 
  \end{cases}
\]
Then a short calculation reveals that the family of currents \( S : t \mapsto G(f + t) \)
is moved by the vector field\(^1\)
\[
\begin{aligned}
   v(t, x, y) := \\
   \begin{cases}
   (-a, 1) \frac{1}{1 + a^2} & \text{for } 0 \leq x < \frac{1}{2} \\
   (a, 1) \frac{1}{1 + a^2} & \text{for } \frac{1}{2} < x \leq 1.
   \end{cases}
\end{aligned}
\]
Since the length of the vector field is constant and \( \mathcal{M}(G(f)) = \sqrt{1 + a^2} \), we get
\[
\int |v|^p \, d\|S(t)\| = \sqrt{1 + a^2} \frac{1}{\sqrt{1 + a^2}} = \frac{1}{\sqrt{1 + a^{2p-1}}}.
\]
Now let us treat \( f \) as a periodic function and define a sequence of rescaled versions
\[
   f_k : [0, 1] \to \mathbb{R}; x \to f(kx)
\]
and \( S_k : t \mapsto G(f_k + t) \). Then still \( \mathcal{M}(G(f)) = \sqrt{1 + a^2} \) and for a similarly rescaled vector field \( v_k \), we get again that
\[
\int |v_k|^p \, d\|S_k(t)\| = \frac{1}{\sqrt{1 + a^{2p-1}}}.
\]
Yet, when looking at the convergence for \( k \to \infty \), we notice that \( S_k(t) \to S(t) = G(t) \). The corresponding normal vector field in this case is \( v = (0, 1) \), yet
\[
\int |v|^p \, d\|S(t)\| = 1 > \frac{1}{\sqrt{1 + a^{2p-1}}} = \lim_{k \to \infty} \inf \int |v_k|^p \, d\|S_k(t)\|
\]
for \( p > 1 \). Thus the length of a curve is not lower semicontinuous. \( \blacksquare \)

**Remark 9.2.** In the previous example, when going over to the limit, the mass of the current decreases. However, in order to compensate, the length of the vector field increases by the same amount. In the case of \( p = 1 \), this is expected, as the traces of \( S_k \) and \( S \) have the same area, which according to Lemma 6.8 should be equal to their \( L_1 \)-length as there is no form of overlap.

\(^1\)Note that we do not consider the vertical Wasserstein distance, which would forbid movement in a non-vertical direction. The notation as a graph here is just for convenience.
On the other hand for $p > 1$, the vector field is counted with a different exponent. Thus the $p$-length of $S_k$ decreases with increasing $a$, while the $p$-length of $S$ stays the same. In the extreme, when sending $a \to \infty$, the $p$-length converges to 0 as the mass of $S_k(t)$ diverges. This again is the degeneracy phenomenon discussed in Example 5.4.

It is possible to construct a similar example for the vertical Wasserstein distance, yet since the degeneracy there stems from the vertical parts, this is a bit more involved.

Note also that while one would expect the vector field of the limit $v$ to be some kind of limit of the vector fields $v_k$, this is clearly not the case. While its direction is a sort of average of the oscillating directions of $v_k$, the length of $v$ corresponds neither to the length of $v_k$ nor to the length of the average. In fact, since the vector field is only well defined on the support of $S_k$, the situation is even more complicated, as the changes in domain along the sequence $S_k$ would be hard to incorporate into a notion of convergence.

9.2 Limits of curves

Normally a the process of finding a minimizer would involve the direct method, that is finding a converging minimizing sequence. Indeed this step holds true, as the following proposition shows.

**Proposition 9.3.** Let $I = [a,b] \subset \mathbb{R}$ a closed interval, $p \geq 1$, $A \subset I_k(\Omega)$ closed and bounded in mass. Let $(S_i)_{i \in \mathbb{N}} \subset C_{pw}^1(I;A)$ be a sequence with a uniform “Lipschitz bound”, that is there is $L > 0$ such that

$$L_p(S_i|_{[s,t]}) \leq L|s-t| \text{ for all } s,t \in I, i \in \mathbb{N}.$$  

Then there is a converging subsequence (not relabeled) $(S_k)_{k \in \mathbb{N}}$ and a limit $S \in C^0(I;A)$ such that $S_k \to S$ uniformly in $\mathcal{F}$.

**Proof.** Using Lemma 5.5 we know that a similar Lipschitz bound holds in the form of

$$W_{1,A}(S_i(s),S_i(t)) \leq L|s-t| \text{ for all } s,t \in I, i \in \mathbb{N}.$$  

As the distance is defined by an infimum over a larger set, we furthermore have

$$W_1(S_i(s),S_i(t)) \leq W_{1,A}(S_i(s),S_i(t)).$$  

Finally by Theorem 7.1 we have

$$W_1(S_i(s),S_i(t)) = \mathcal{F}(S_i(s) - S_i(t))$$  

which results in all $S_i$ being equi-Lipschitz continuous.

But then since $A$ is compact in the $\mathcal{F}$-topology, the existence of a converging subsequence is the result of the Arzela-Ascoli theorem. 

$\square$
Note that the ‘Lipschitz bound’ holds along minimizing sequences, as we can always reparametrize by arc length. Furthermore, note the absence of any bound on $L_p(S)$, which would run contrary to Example 9.1. In fact we will even have to prove some sort of differentiability for $S$ first.

On a basic level this is a result of the uniform Lipschitz bound, which results in $S$ being Lipschitz in the $\mathcal{F}$ topology and thus differentiable almost everywhere. However we can show even more, as it is possible to recover a vector field for which $\dot{S}(t) + L_v(S)(t) = 0$ by reversing the argument used in Remark 6.13.

**Lemma 9.4.** Let $I = [a,b] \subset \mathbb{R}$ a closed interval, $p \geq 1$, $A \subset I_k(\Omega)$ closed and bounded in mass. Let $(S_i)_{i \in \mathbb{N}} \subset C^1_{pw}(I;A)$ be a sequence with a uniform “$L_p$-Lipschitz bound” and $S_k \to S$ as in Proposition 9.3. Then for almost all $t \in I$ there exists a piecewise continuous vector field $v(t)$ such that

$$\dot{S}(t) + L_v(S)(t) = 0.$$

![Figure 9.2](image)

**Figure 9.2:** The general approach in Lemma 9.4. While we have uniform convergence of the continuous curves $S_i$ to a limit curve $S$ (top left to bottom left), the corresponding vector fields need to be recovered by a detour using the traces of the individual $S_i$ (top right), which converge to a limit $\tilde{S}$ (bottom right) as well. This limit then relates to $S$ via slicing and we can recover a vector field using the corresponding tangential spaces.
**Proof.** We know that for all \( t \in I \), the current \( S(t) \) is rectifiable. Let us denote the simple unit \( k \)-vector that corresponds to the tangential plane by \( \xi_t \), defined almost everywhere on \( \text{supp} \ S(t) \). Furthermore we know that \( (id, S_i) \# (I) \) is rectifiable by Theorem 6.11 and \( M \left( (id, S_i) \# (I) \right) \leq L_1(S_i) \leq L |a - b| \) by Lemma 6.8. Here using the uniform bounds we have on the lengths \( L_p(S_i) \) and the masses \( M(S_i(t)) \), the constant \( L \) can be chosen independently of \( i \). Then by the compactness-closure theorem, up to a subsequence \( (id, S_i) \# (I) \) converges to \( \tilde{S} \). Using Corollary 6.6 and the continuity of slicing, we additionally get \( \langle \tilde{S}, t < t_0 \rangle = S(t_0) \) for almost all \( t_0 \).

Now let \( \xi \) be the unit \( k + 1 \)-vector corresponding to the tangential plane of \( \tilde{S} \). Using [GMS98, I.2.2.5 Lemma 1] we find that for \( H^1 \) almost all \( t \) and \( H^k \) almost all \( x \in \text{supp} S(t) \) the space spanned by \( \xi_t(x) \) is a subspace of the space spanned by \( \xi \), that is, there is an unit vector \( w(t, x) \) such that

\[
\xi(t, x) = w(t, x) \wedge \xi_t(t, x).
\]

This can be done in such a way, that the tangential space spanned by \( \xi(t, x) \) is the space spanned by \( \xi_t(x) \) together with the projection of \((1, 0, ..., 0)\) onto the space spanned by \( \xi(t, x) \). As a results \( w \) has a component in the \( t \)-direction.\(^2\) This will allow us to write \( w(t, x) = (1, v(t, x)) \), where \( v \) is the actual desired vector field.

Furthermore \( \tilde{S} \) can be seen as the union of countably many \( C^1 \) manifolds of dimension \( k + 1 \) and \( \xi \) is continuous on those manifolds. As a further consequence of [GMS98, I.2.2.5 Lemma 1], then \( w \) and \( \xi_t \) are continuous on those manifolds as well. In particular then for almost all \( t \in I \), \( w \) is continuous \( H^k \)-almost everywhere on \( S(t) \). Now fix \( \omega \in C^\infty(\Omega; \wedge^k) \), treated as a \( k \)-form in \( I \times \Omega \), constant in time. Reversing the calculation in Remark 6.13 we get

\[
\frac{S(t + h) - S(t)}{h} \langle \omega \rangle = \frac{1}{h} \partial \left( \tilde{S} \left[ t, t + h \right] \times \Omega \right) \langle \omega \rangle
\]

\[
= \frac{1}{h} \int_{[t,t+h] \times \Omega} \langle d\omega, \xi \rangle \ d\left\| \tilde{S} \right\|
\]

\[
= \frac{1}{h} \int_{[t,t+h] \times \Omega} \langle d\omega, w \wedge \xi_t \rangle \ d\left\| \tilde{S} \right\|
\]

\[
= \frac{1}{h} \int_{[t,t+h] \times \Omega} \langle i_w d\omega, \xi_t \rangle \ d\left\| \tilde{S} \right\|
\]

\[
= \frac{1}{h} \int_t^{t+h} \int_{u_0}^1 \langle i_w d\omega, \xi_t \rangle \ d\left\| S(t) \right\| \ dt
\]

\(^2\)This is however not true everywhere. The tangential space of \( \tilde{S} \) can be vertical at some points, yet for almost all times this may only happen on a \( H^k \) zero set. This is a Sard-like property of the slicing construction.
9.2. LIMITS OF CURVES

\[
\rightarrow \int \frac{1}{w_0} \langle i_w \mathrm{d} \omega, \xi_t \rangle \mathrm{d} \|S(t)\| = S(t) \langle \mathcal{L}_{\frac{w}{w_0}} \omega \rangle
\]

where \( w_0 \) is the time component of \( w \) and thus \( \frac{w}{w_0} = (1, v) \).

We have thus identified \( \tilde{S} = (\text{id}, S)_\#(I) \).

**Lemma 9.5.** Let \( I = [a, b] \subset \mathbb{R} \) a closed interval, \( A \subset I_k(\Omega) \) closed and bounded in mass. Let \( (S_i)_{i \in \mathbb{N}} \subset C_{pw}^1(I; A) \) be a sequence of curves and \( S_i \rightarrow S \) as in Proposition 9.3. Let \( v_i \) be the corresponding perpendicular vector fields for \( S_i \) and \( v \) the perpendicular vector field for \( S \). Then for any open subset \( U \subset \Omega \) and any interval \( J \subset I \) we have

\[
\int_J \int_U |v| \mathrm{d} \|S\| \mathrm{d} t \leq \liminf_{i \rightarrow \infty} \int_J \int_U |v_i| \mathrm{d} \|S_i\| \mathrm{d} t.
\]

In particular we have \( L_1(S) \leq \liminf_{i \rightarrow \infty} L_1(S_i) \).

**Proof.** We can always assume that \( U = \Omega \) by considering the restrictions \( S_i \mathbb{1}_U \) instead of the actual currents. Now fix \( r > 0 \). Let us define new curves by

\[
R_i : \frac{1}{r} J \rightarrow A; s \mapsto S_i(rs),
\]

\[
R : \frac{1}{r} J \rightarrow A; s \mapsto S(rs).
\]

Then by linearity of the time- and Lie-derivative we get

\[
\dot{R}(s) + \mathcal{L}_v R(s) = r\dot{S}(rs) + r\mathcal{L}_v S(rs) = 0
\]

and thus the corresponding vector fields are \( rv_i \) and \( rv \). Furthermore since we only rescaled the time direction by a constant factor, we still have \( R_i \rightarrow R \) uniformly and \( (\text{id}, R_i)_\#(J) \rightarrow (\text{id}, R)_\#(J) \). Now applying Lemma 6.8 we end up with

\[
\int_J \int_U |v| \mathrm{d} \|S(t)\| \mathrm{d} t = \int_J \int_U r |v| \mathrm{d} \|S(rs)\| \mathrm{d} s
\]

\[
< \int_J \int_U \sqrt{1 + |rv|^2} \mathrm{d} \|S(rs)\| \mathrm{d} s
\]

\[
= \mathcal{M}
\]

\[
\leq \liminf_{i \rightarrow \infty} \mathcal{M}
\]

\[
= \liminf_{i \rightarrow \infty} \int_J \int_U \sqrt{1 + |rv|^2} \mathrm{d} \|S_i(rs)\| \mathrm{d} s
\]
\[ = \liminf_{i \to \infty} \int_J \int r \sqrt{\frac{1}{r^2} + |v_i|^2} d\|S_i(r)s\| \, ds \]

\[ = \liminf_{i \to \infty} \int_J \int \sqrt{\frac{1}{r^2} + |v_i|^2} d\|S_i(t)\| \, dt \]

\[ \leq \liminf_{i \to \infty} \frac{1}{r} \int_J \int d\|S_i(t)\| \, dt + \int_J \int |v_i| d\|S_i(t)\| \, dt \]

\[ = \frac{|J|}{r} M_0 + \liminf_{i \to \infty} \int_J \int |v_i| d\|S_i(t)\| \, dt \]

Here \( M_0 \) is the mass bound for \( A \) and \(|J|\) the length of the interval. Now sending \( r \to \infty \) finishes the proof.

Note that we again took the detour of an additional time component, only to lessen its import by rescaling it. The idea here is again to prevent annihilation of area generated at different times. This becomes apparent in Lemma 6.8 where there is only an inequality between the integral over \( v \) and \( \mathcal{M}(S_{\#}(I)) \), while the proof needs equality in order to use the estimate in both directions.

A direct consequence of the preceding now is:

**Corollary 9.6.** Let \( A \subset I_k(\Omega) \) be closed and bounded in mass. Then there exist geodesics in the \( W_{1,A} \)-distance, that is, for any \( T_0, T_1 \in A \) the functional \( L_1 \) has a minimizer in the class

\[ \{ S \in C^{1}_{pw}(I; A) \mid S(0) = T_0, S(1) = T_1 \}. \]

**Remark 9.7.** We can now try to apply the same idea to the vertical Wasserstein distance. Using exactly the same arguments, we will again recover a vector field \( w \) with nonzero time component almost everywhere. What we actually want in the end is a vector field of the form \((1,0,v)\), without horizontal movement. Instead we only get a vector field in the form of \((1,\tilde{v},v)\).

As we have seen in Lemma 3.7, the Lie derivative stays invariant under the addition of tangential vector fields. So, as long as we are in the graph part of the current, there will always be a tangential vector to \( S(t) \) of the form \((\tilde{v},u)\), so we can use the vector field \((1,0,v-u)\). The problematic situation only occurs in the vertical part of our current.

This is best seen by having another look at Example 5.11. This time, we will fix \( k = 1 \) and consider the limit \( \varepsilon \to 0 \). What happens is that we slightly unravel a bubble at the point \( a \), move it as a function to the point \( b \) and then reform it. Due to the fixed \( k \), we will always have bounds on mass and the length of our curve will never degenerate. Yet if we take the limit \( \varepsilon \to 0 \), we clearly end up with the limiting curve

\[ S(t) := G((0,0,1)) + \|\{tb + (1-t)a\}\| \times S^2. \]
Then \( \dot{S}(t) + \mathcal{L}_v S(t) = 0 \) for

\[
v(t, x, y) := \begin{cases} 
    b - a & \text{for } x = tb + (1 - t)a, y \in S^2 \\
    0 & \text{otherwise}
\end{cases}
\]

which cannot be replaced by a vertical vector field. ■

The solution to this problem would be to replace our distance by a relaxed distance, which takes those limits into consideration.

## 9.3 Lower semicontinuity

As Example 9.1 showed, sadly in general there is no lower semicontinuity of the length of a curve apart from the case of \( p = 1 \). However, what we expect is that there is still lower semicontinuity of the Wasserstein distance in the following form:

**Conjecture 9.8** (Lower semicontinuity). Let \( p \geq 1 \), \( A \subset I_k(\Omega) \) closed and bounded in mass. Furthermore let \( (R_i)_{i \in \mathbb{N}}, (T_i)_{i \in \mathbb{N}} \subset A \) with \( R_i \to R \in A \) and \( T_i \to T \in A \). Then

\[
W_{p, A}(R, T) \leq \liminf_{i \to \infty} W_{p, A}(R_i, T_i).
\]

The reason, why we expect this lower semicontinuity, is the same reason, for which we know that there is no lower semicontinuity of the length functional. As we have seen in the last section, it is beneficial to add a small oscillation that increases the mass and decreases the length of the vector-field by the same factor. Then since the length of a curve roughly scales with the length of the vector-field to the power \( p \) and yet only linearly in the mass, a relative increase in mass by the factor \( r \) will thus decrease the length of the curve by a factor of \( r^{1-p} \). (This has of course to be understood somewhat locally, as this exchange is more beneficial where the vector-field is comparatively large and useless for parts of the current which stay still.)

This is the only way in which the lack of mass conservation should affect us negatively. Consider the following (non-rigorous) argument: Take any sequence of curves \( S_i \to S \) in the usual sense of Proposition 9.3. Then as the mass is lower semicontinuous, we know that the mass of the trace does not increase and thus

\[
\mathcal{M}(S\#(I)) \leq \liminf_{i \to \infty} \mathcal{M}(S_i\#(I))
\]

for any interval \( I \). On the other hand for a short interval, the mass of the trace can be estimated (see Lemma 6.8) by \( \mathcal{M}(S\#(I)) \approx \int_I \int |v| \, d||S(t)|| \, dt \). Thus if we assume \( |v| \) and \( |v_i| \) to be constant (which locally is close to the truth, that is, if we restrict ourselves to a small region of the actual domain), then
roughly $\|S(t)\| \cdot |v| \leq \liminf_{i \to \infty} \|S_i(t)\| |v_i|$ So if the decrease in mass is given by $\mathcal{M}(S(t)) \approx r^{-1} \lim_{i \to \infty} \mathcal{M}(S_i(t))$ with $r \geq 1$, then in the limit the vector field can only increase by the same amount, that is at most $|v| \approx r \lim_{i \to \infty} |v_i|$. As a result we have

$$\int |v|^p d\|S(t)\| \leq r^{p-1} \lim_{i \to \infty} \int |v_i|^p d\|S(t)\|$$

which is the factor which we got from the examples in the last section.

But then on the other hand if we know this, we could repeat the construction of oscillations. Assume now that the $S_i$ are curves with $S_i(0) = R_i$ and $S_i(1) = T_i$ with $R_i$ and $T_i$ as in the conjecture and $L_p(S_i) = W_{p, A}(R_i, T_i) + o(1)$. Then we can use Proposition 9.3 to find a limiting curve $S$ with $S(0) = R$ and $S(1) = T$. By the above consideration we would know that (assuming again $|v|$ to be roughly constant) $L_p(S) \leq r^{p-1} \liminf_{i \to \infty} L_p(S_i)$, where $r$ is the decrease in mass. On the other hand, being allowed an additional factor of $r$ in the mass, we know by the example how to construct a curve $\tilde{S}$ with $L_p(\tilde{S}) = r^{-p} L_p(S) + o(1)$ and identical boundary data. This would allow us to conclude lower semicontinuity.

Of course this idea oversimplifies things somewhat and it is hard to make this argument completely rigorous. Some of the details are easily fixed. We have for example already seen that while Lemma 6.8 only yields an inequality, we can get closer to an equality by using the space-time trace instead. However other problems are more difficult. The most worrisome detail that still eludes us, is how to define the relative decrease in mass $r$, which in a full proof would be a local quantity.

### 9.4 Relaxed Energy

Finally, let us wrap up this discussion by showing that assuming the previous conjecture, there are in fact minimizers for a relaxed energy functional, which is somewhat related to our counterexamples.

We already noted in Proposition 9.3 that a converging sequence of curves has a limit with a corresponding vector field according to Lemma 9.4. The only problem now is that for $p > 1$ the limit curve has a larger length. A simple approach would be to redefine the length of a curve similar to the length of a rectifiable curve on the interval $I = [a, b]$:

$$\tilde{L}_{p, A}(S) := \lim_{\delta \to 0} \inf \left\{ \sum_{i=1}^{l} W_{p, A}(S(t_i), S(t_{i-1})) \left| (t_0, ..., t_l) \in P_\delta \right. \right\}$$

where $P_\delta$ is the set of all partitions of $[a, b]$ with steps smaller than $\delta$, that is $P_\delta := \{(t_0, ..., t_l) \mid a = t_0 < t_1 < ... < t_l = b, |t_i - t_{i-1}| < \delta \}$. 
This definition of length of course still results in the same notion of distance:

**Lemma 9.9.**

\[ W_{p,A}(T_0, T_1) = \inf \left\{ \tilde{L}_{p,A} \mid S \in C([0,1]; A), T_0 = S(0), T_1 = S(1) \right\} \]

**Proof.** Let \( S : I \to A \) be a curve, \( \delta > 0 \) and \((t_0, ..., t_l) \in P_\delta \) a partition. By the triangle inequality

\[ W_{p,A}(T_0, T_1) \leq W_{p,A}(S(t_0), S(t_1)) + ... + W_{p,A}(S(t_{l-1}), S(t_l)) \]

so taking the infimum over all partitions, sending \( \delta \to 0 \) and taking the infimum over all curves yields

\[ W_{p,A}(T_0, T_1) \leq \inf \left\{ \tilde{L}_{p,A} \mid S \in C([0,1]; A), T_0 = S(0), T_1 = S(1) \right\}. \]

Now fix \( \varepsilon > 0 \). Then there exists a curve \( S \) such that

\[ L_p(S) < W_{p,A}(T_0, T_1) + \varepsilon. \]

Additionally, we can assume \( S \) to be parametrized by arc-length. Thus for all \( s,t \in I, s < t \):

\[ W_{p,A}(S(s), S(t)) \leq L_p(S_{|[s,t]}) < (W_{p,A}(T_0, T_1) + \varepsilon) (t - s) \]

Then for any \( \delta > 0 \) and any partition \((t_0, ..., t_l) \in P_\delta \) we have

\[ \sum_{i=1}^{l} W_{p,A}(S(t_i), S(t_{i-1})) < W_{p,A}(T_0, T_1) + \varepsilon \]

and so

\[ \tilde{L}_{p,A}(S) < W_{p,A}(T_0, T_1) + \varepsilon. \]

Sending \( \varepsilon \to 0 \) then results in the missing inequality. \( \square \)

On the other hand however, using this relaxed energy would result in lower semicontinuity and thus in existence of a minimizer.

**Lemma 9.10.** Let \( (S_k)_{k \in \mathbb{N}} \subset A \) equicontinuous in \( W_{p,A} \) and \( S_k \to S \) in the sense of Proposition 9.3. Then

\[ \tilde{L}_{p,A}(S) \leq \liminf_{k \to \infty} \tilde{L}_{p,A}(S_k). \]

**Proof.** Let

\[ I(\delta) := \inf \left\{ \sum_{i=1}^{l} W_{p,A}(S(t_i), S(t_{i-1})) \mid (t_0, ..., t_l) \in P_\delta \right\} \]
Due to the triangle inequality, we can assume that by bounded and thus constant after choice of a subsequence. So and thus 

Now fix \( \delta > 0 \). For any \( k \in \mathbb{N} \) there exists \( (t_0^{(k)} , ..., t_i^{(k)}) \in P_\delta \) such that

\[
I_k(\delta) \geq \sum_{i=1}^{l_k} W_{p,A} \left(S \left( t_i^{(k)} \right), S \left( t_{i-1}^{(k)} \right) \right) - \frac{1}{k}
\]

and thus

\[
\liminf_{k \to \infty} I_k(\delta) \geq \liminf_{k \to \infty} \sum_{i=1}^{l_k} W_{p,A} \left(S \left( t_i^{(k)} \right), S \left( t_{i-1}^{(k)} \right) \right).
\]

Due to the triangle inequality, we can assume that \( |t_i^{(k)} - t_{i+2}^{(k)}| > \delta \) for all \( k \in \mathbb{N}, 0 \leq i < l_k - 2 \), otherwise we could replace \( (t_0^{(k)} , ..., t_i^{(k)}, t_{i+1}^{(k)}, t_{i+2}^{(k)}, ..., t_{l_k}^{(k)}) \) by \( (t_0^{(k)}, ..., t_i^{(k)} , t_i^{(k)} - \delta, ..., t_{l_k}^{(k)}) \in P_\delta \) while decreasing the sum. But then \( l_k \) is bounded and thus constant after choice of a subsequence. So \( l_k = l \) and after another choice of subsequence \( (t_0^{(k)} , ..., t_i^{(k)}) \) converges to \( (t_0 , ..., t_l) \in P_{2\delta} \).

Now using that all \( S_k \) have bounded length and are equicontinuous, we have that

\[
\liminf_{k \to \infty} \sum_{i=1}^{l_k} W_{p,A} \left(S_k \left( t_i^{(k)} \right), S_k \left( t_{i-1}^{(k)} \right) \right)
\]

\[
= \liminf_{k \to \infty} \sum_{i=1}^{l} W_{p,A} \left(S_k \left( t_i \right), S_k \left( t_{i-1} \right) \right)
\]

\[
\geq \inf \left\{ \liminf_{k \to \infty} \sum_{i=1}^{l} W_{p,A} \left(S_k \left( t_i \right), S_k \left( t_{i-1} \right) \right) \left| (t_0 , ..., t_l) \in P_{2\delta} \right. \right\}
\]

where the last step needs the lower-semicontinuity of \( W_{p,A}. \) Thus

\[
I(2\delta) \leq \liminf_{k \to \infty} I_k(\delta).
\]

Finally, since \( P_{\delta'} \subset P_\delta \) for \( \delta' \leq \delta \), \( I \) and \( I_K \) are both decreasing functions. Thus

\[
\liminf_{k \to \infty} \tilde{L}_{p,A} (S_k) = \liminf_{k \to \infty} I_k(\delta) \geq \liminf_{k \to \infty} I_k(\delta_0) \geq I(2\delta_0) \xrightarrow{\delta_0 \to 0} \tilde{L}_{p,A} (S) \]

\( \square \)
9.4. RELAXED ENERGY

As a direct consequence we get the existence of $\tilde{L}_{p,A}$-geodesics.

**Corollary 9.11.** Let $T_0, T_1 \in A$, $W_{p,A}(T_0, T_1) < \infty$. Then there exists a curve $S : [0, 1] \to A$ such that

$$\tilde{L}_{p,A}(S) = W_{p,A}(T_0, T_1).$$

**Proof.** Take a minimizing sequence for $\tilde{L}_{p,A}$. This can be done equicontinuous by reparametrizing. Then according to Proposition 9.3, there exists a converging subsequence, which will converge to a minimizer according to Lemma 9.10.

A next step of study at this point would then be to analyze in which way $\tilde{L}_{p,A}$ is related to a local norm, that is if it can be written in the form

$$\tilde{L}_{p,A}(S) = \int_0^1 F(S(t), \dot{S}(t))dt$$

which would give $A$ a structure similar to that of a Finsler manifold. Here $F$ plays the role of $\int |v|^p d\|S(t)\|$, where $v$ is of course determined by by $\dot{S}$. It would be reasonable to expect something similar in this case, except with some possible scaling to account for the excess mass. One could then continue to do the same for Cartesian currents, which may be of special interest, as by the reasoning of Remark 9.7 some movement in non-vertical directions needs to be allowed. All of this does however depend on the lower semicontinuity, so we will not continue this discussion for now.
Chapter 10

Including the boundary

As mentioned in the beginning, we focused on the special case of boundary-less currents as they are commonly occurring in interesting applications such as the Cartesian currents and since they simplify a lot of the proofs. It should however not pose a problem to extend the previous results to the case of currents with boundaries. The aim of this chapter is to present some of the underlying ideas.

10.1 A generalized Wasserstein distance for currents with boundary

Let us first have another look at the conservation of multiplicity. When dealing with smooth vector fields,

\[ \mathcal{L}_v T(\omega) = -T(\mathcal{L}_v \omega) \]

makes sense no matter if \( T \) has a boundary or not. We might not be able to simplify this to \( -T(\imath_v \omega) \) anymore, but this will only add an additional term. We can even note the practical identity

\[ \partial \mathcal{L}_v T(\omega) = -T(\mathcal{L}_v d\omega) = T(d\mathcal{L}_v \omega) = \mathcal{L}_v \partial T(\omega). \]

The problems however start to occur, when we are trying to minimize the \( L^p \) norm of this vector field. We have seen in Lemma 3.7 that the optimal vector field is normal to the manifold. When \( T \) has no boundary we can always choose this minimizer. Yet when \( \partial T \) exists, a normal vector field will only allow \( \partial T \) to move in such a normal direction, it will be impossible to extend \( T \) in a tangential direction instead. To complicate things more, even in the case where \( \partial T \) moves in such a direction, the infimum of the \( L^p \)-norm of all possible vector fields will still be the tangential part, as we are still able vary the vector field in the interior as the following example illustrates.
Example 10.1. Let $\Omega = \mathbb{R}^2$ and let $T(t) = \left[ [0, 1+t] \times \{t\} \right]$ for $t \in [0,1]$. A reasonable idea would be to define a parametrization

$$\phi_t : [0,1] \to \mathbb{R}^2; s \mapsto ((1+t)s,t).$$

Then $T(t) = \phi_t \# \left[ [0,1] \right]$ and a corresponding vector field is given by

$$v := \frac{\partial}{\partial t} \phi_t \circ \phi_t^{-1}(x,y) = \left( \frac{x}{1+t}, 1 \right)$$

for all $(x,y) \in \text{supp} T(t)$. Yet other vector fields are equally possible. By the reasoning of Lemma 3.7 we can modify $v$ by an arbitrary vector field tangential to $\text{supp} T(t)$. To see this explicitly take

$$w : \text{supp} T(t) \to \mathbb{R}^2; (x,y) \to (\psi(x),0)$$

where $\psi \in C^\infty_0([0,1+t])$. Now for $\omega := f(x,y)dx + g(x,y)dy$ consider

$$T(t)\langle \mathcal{L}_w \omega \rangle = T(t)\left( \text{d} (f(x,y)\psi(x)) + i_w \left( \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) \text{d} x \wedge \text{d} y \right)$$

$$= \partial T(t) \langle f(x,y)\psi(x) \rangle + T(t) \left( \left( \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) \psi \text{d} y \right)$$

$$= 0.$$

Here the first term is zero since $\psi$ vanishes on $\text{supp} \partial T$ and the second term is zero since $T(t)$ only extends in $x$-direction and thus any integration over $\text{d} y$ vanishes. Hence

$$\dot{T} + \mathcal{L}_{v+w} T = \dot{T} + \mathcal{L}_v T = 0$$

Yet if we try to minimize to minimize

$$\int |v + w|^p \| T(t) \| = \int_0^{1+t} \left| \left( \frac{x}{1+t} + \psi(x), 1 \right) \right|^p \text{d} x$$

we run into a problem. The optimal choice here would of course be $\psi_0(x) := \frac{-x}{1+t}$, yet this is nonzero at $1+t$ and thus not allowed. However we can approximate $\psi_0$ arbitrarily close in $L^p$ by using $\psi_\varepsilon(x) := \psi(x)\lambda_\varepsilon(x)$ where

$$\lambda_\varepsilon(x) = \begin{cases} 1 & \text{for } x < 1+t - \varepsilon \\ 0 & \text{for } 1+t - \frac{\varepsilon}{2} < x \\ \text{smooth} & \text{in between} \end{cases}$$

So we see that the infimum is not obtained. \hfill \blacksquare
As a direct consequence, we cannot use the minimal vector field anymore. However when trying to define a generalized Wasserstein distance in the same way as before, there is even a bigger problem, as we cannot see the movement of the boundary. In fact we have the following degeneracy:

**Theorem 10.2.** Let $T_0, T_1 \in I_k(\Omega)$. Then there is a curve $T : [0, 1] \to I_k(\Omega)$ with $T(0) = T_0$ and $T(1) = T_1$ such that

$$\int_0^1 \inf \left\{ \int |v|^p \, dt : \|T(t)\| = 0 \right\} \, dt = 0.$$

**Proof.** We show that there exists such a curve linking $T_0$ to 0. The result then follows from taking a similar curve from 0 to $T_1$. Furthermore we restrict ourselves to the case where $T_0$ is included in the image of a $C^1$ manifold diffeomorphic to the $k$-dimensional unit ball. The general case then follows by covering $T_0$ with such manifolds and applying the curves one by one.

Let $\phi : B_1(0) \to \Omega$ be the chart for the manifold with $supp T_0 \subset \phi(B_1(0))$. Then we define the following family of diffeomorphisms

$$r_t : \phi(B_1(0)) \to \phi(B_{1-t}(0)), \, x \mapsto \phi((1-t)\phi^{-1}(x))$$

and the curve

$$T : (0, 1) \to I_k(\Omega), \, t \to (r_t)_#T_0$$

which has the trivial limits $\lim_{t \to 0} T(t) = T_0$ and $\lim_{t \to 1} T(t) = 0$.

Now a possible vector field immediately follows from the definition of the diffeomorphism, setting

$$v(t,x) = \frac{\partial r_t}{\partial t}(r_t^{-1}(x))$$

for all $x \in supp T(t)$. This vector field is not good enough to show the theorem, however by our previous deliberations it is not hard to see how to improve
it. Since all $T(t)$ lie in the same $k$-dimensional manifold, $v$ is pointing in a tangential direction for all interior points of $T_t$. So in those points $v = 0$ would be optimal. We just need to be careful about not changing $v$ on $\partial T(t)$, since there $v = 0$ is not admissible.

Since $T(t)$ is integer rectifiable, $\mathcal{H}^{k-1}(\text{supp } \partial T(t))$ is bounded. But then for any $\varepsilon > 0$ there exists a smooth function

$$
\psi_\varepsilon : \phi(B_1(0)) \to [0, 1]
$$

with $\psi_\varepsilon = 1$ on $\partial T(t)$ and $\mathcal{H}^k(\text{supp } \psi_\varepsilon) < \varepsilon$, which can be constructed in the following way:

From the definition of the Hausdorff-measure, for any $\delta > 0$ there exists a finite set of points $(x_i)_{i \in J}$ such that $\text{supp } \partial T(t) \subset \bigcup_{i \in J} B_\delta(x_i)$ and

$$
\sum_{i \in J} (\text{diam } B_\delta(x_i))^{k-1} < c \mathcal{H}^{k-1}(\text{supp } \partial T(t))
$$

where $c$ is independent of $\delta$. Using the usual constructions, then there exists a smooth function $\psi$ which is equal to 1 on $\bigcup_{i \in J} B_\delta(x_i)$ and has support in $\bigcup_{i \in J} B_{2\delta}(x_i)$. However then

$$
\mathcal{H}^k(\text{supp } \psi \cup \phi(B_1)) \leq \mathcal{H}^k\left( \bigcup_{i \in J} B_{2\delta}(x_i) \right)
$$

$$
\leq c \sum_{i \in J} (\text{diam } B_{2\delta}(x_i))^k
$$

$$
\leq c2\delta \sum_{i \in J} (\text{diam } B_{2\delta}(x_i))^{k-1} < c\delta \mathcal{H}^{k-1}(\text{supp } \partial T(t)) < \varepsilon
$$

if $\delta$ is chosen small enough.

Now we define $v_\varepsilon(t, .) = \psi_\varepsilon v(t, .)$ which by construction is equal to $v$ on $\partial T(t)$ and constructed so that $v - v_\varepsilon$ is tangential to $T(t)$ since $v$ is tangential to $T(t)$. So $\dot{T} + \mathcal{L}_{v_\varepsilon} T = 0$. Yet

$$
\int |v_\varepsilon|^p \, d\|T(t)\| \leq \mathcal{H}^k(\text{supp } \psi_\varepsilon) \sup_{x \in \phi(B_1)} |v(x)|^p \to 0
$$

for $\varepsilon \to 0$.

A lesson from this is that we cannot ignore the movement of the boundary. Instead we should treat it in the same way as the movement of the interior. Thus we get
Figure 10.2: The general idea of Theorem 10.2. Instead of moving a current \( T_0 \) to its new position \( T_1 \) (a), we shrink \( T_0 \) along itself and create \( T_1 \) in a reverse fashion (b). This way we can use a vector field \( v \) that vanishes apart from an arbitrary small set close to \( \partial T(t) \).

**Definition 10.3.** Let \( A \subset I_k(\Omega) \) be a closed set of currents, \( T_0 \) and \( T_1 \) in \( A \). Let \( S \in C^1_{pw}([0,1]; A) \) be such that \( S(0) = T_0 \) and \( S(1) = T_1 \). Then we define the \( p \)-length with boundary of the homotopy \( S \) by

\[
L_{p,\partial}(S) := (L_p(S)^p + L_p(\partial S)^p)^{\frac{1}{p}}
\]

and the generalized \( p \)-Wasserstein distance with boundary between \( T_0 \) and \( T_1 \) by

\[
W_{p,\partial,A}(T_0, T_1) := \inf \{ L_{p,\partial}(S) \mid S \in C^1_{pw}([0,1]; A), S(0) = T_0, S(1) = T_1 \}.
\]

In this definition of \( L_{p,\partial} \) we treated \( S \) and \( \partial S \) as different entities altogether. However it should be noted that it is indeed possible to consider a single vectorfield moving both of them which will result in an equivalent definition.

**Lemma 10.4.** Let \( I \) be an interval and \( S \in C^1_{pw}(I; A) \) then for all \( t \in I \) we have

\[
\inf \left\{ \int |v|^p \, d\|S(t)\| + \int |v|^p \, d\|\partial S(t)\| \bigg| \dot{S}(t) + L_v S(t) = 0 \right\} \\
= \inf \left\{ \int |v|^p \, d\|S(t)\| \bigg| \dot{S}(t) + L_v S(t) = 0 \right\} \\
+ \inf \left\{ \int |w|^p \, d\|\partial S(t)\| \bigg| \dot{\partial S}(t) + L_w \partial S(t) = 0 \right\},
\]

where the infima are taken over all piecewise continuously differentiable functions. As a direct consequence

\[
L_{p,\partial}(S)^p = \int_0^1 \inf \left\{ \int |v|^p \, d\|S\| + \int |v|^p \, d\|\partial S\| \bigg| \dot{S} + L_v S = 0 \right\} dt.
\]
10.2. CORRESPONDENCE WITH THE FLAT METRIC

Proof. As the infima on the right hand side are taken separately, taking \( v = w \) shows that the left hand side is not larger than the right hand side of the equation. For the other direction we note that any vector field \( v \) transporting \( S \) also transports \( \partial S \). Since any tangential vector of \( \partial S(t) \) is also a tangential vector of \( S(t) \), we can thus assume that \( v \) is tangential to \( \partial S(t) \) by Lemma 3.7 and thus already optimal for the first term. But then the two sides of the equation coincide.

The second part of the lemma then follows from integrating over \([0, 1] \). □

Remark 10.5. At this point we can revisit the situation in Theorem 10.2. If we take the curve that was constructed there and consider its \( L_{p, \partial} \) length instead, this will no longer be zero. While as before, the integral over the bulk of \( T(t) \) does not contribute to the length, the integral over \( \partial T \) now results in a positive distance. In fact the curve \( \partial T : t \mapsto \partial T(t) \) now is an admissible curve for connecting the boundary-less currents \( T_0 \) and 0. If we now additionally require that \( T(t) \in \mathcal{A} \) for all \( t \in [0, 1] \), where \( \mathcal{A} \) has an uniform bound on the mass of all boundaries, we even get

\[
L_{p, \partial}(T) = L_p(\partial T) \geq \mathcal{W}_{p, \partial} \mathcal{A}(T_0, 0) \geq c\mathcal{W}_{1, \partial} \mathcal{A}(T_0, 0) \geq c\mathcal{M}(T_0),
\]

by using Lemma 5.5 and Corollary 6.12. □

10.2 Correspondence with the Flat metric

We can now try to adapt some of what we did in the preceding chapters to the case of Wasserstein distance with boundary. Let us start with the trace of a curve. Assume that \( S : [0, 1] \mapsto I_k(\Omega) \) is piecewise differentiable with corresponding vectorfield \( v \). Then we still define, as before

\[
S_\#([0, 1]) \langle \omega \rangle = \int_0^1 S(t) \langle i_v(t) \omega \rangle dt.
\]

Note that Lemma 6.5 now tells us that

\[
\partial S_\#([0, 1]) = S(1) - S(0) + (\partial S)_\#([0, 1]).
\]

Then in light of Lemma 6.8, we have

\[
\mathcal{M}(S_\#([0, 1])) + \mathcal{M}\left( (\partial S)_\#([0, 1]) \right) \leq L_1(S) + L_1(\partial S) = L_{1, \partial}(S).
\]

Furthermore Theorem 6.11 tells us that \((\partial S)_\#([0, 1])\) is integer rectifiable and the theorem can be easily extended to show that then also \( S_\#([0, 1]) \in I_k(\Omega) \) in order to get
Corollary 10.6. Let $T_0, T_1$ in $I_k(\Omega)$. Then

$$\mathcal{F}(T_1 - T_0) \leq W_{1,0}(T_0, T_1).$$

Proof. As said before, we only need to check rectifiability of $S_\#([0,1]) \in I_k(\Omega)$. We can proceed as in the proof of Theorem 6.11 and show rectifiability of the space-time trace instead. The only difference is that when trying to apply the slicing lemma, the additional term $\partial(id, S)_\#([0,1]) \L B_r(a)$ appears. Yet as $d$ and pullback commute, so do $\partial$ and pushforward and so this term is equivalent to $(id, \partial S)_\#([0,1]) \L B_r(a)$ and thus rectifiable by induction.

The remaining question now is, if the reasoning in Chapter 7 can be extended similarly to show equivalence in the form of

Theorem 10.7. Let $T_0, T_1$ in $I_k(\Omega)$. Then

$$\mathcal{F}(T_1 - T_0) = W_{1,0}(T_0, T_1).$$

Proof. We will just sketch the differences to the boundary-less case. Maybe the most important detail here is replicating Lemma 7.5. Assume that there are $S \in I_{k+1}(\Omega)$ and $R \in I_k(\Omega)$ such that $\partial S = T_1 - T_0 + R$.

We apply the deformation theorem to $S, T_0, T_1$ and $R$ at the same time. This results in a homotopy $h$ transforming all four of those currents into polyhedral chains $P_S, P_{T_0}, P_{T_1}$ and $P_R$, whose masses are bounded by the original currents and which are chosen such that $\partial P_S = P_{T_1} - P_{T_0} + P_R$. It is also easy to see that the $L_{1,0}$-lengths of $t \mapsto h(t,.)_\#T_0$ and $t \mapsto h(t,.)_\#T_1$ are bounded by a constant (depending on dimensions and the $T_i$) times the skeleton-size $\varepsilon$ and thus can be chosen arbitrarily small.

Then the equivalent of Lemma 7.5 follows from an appropriate variant of Lemma 7.3 whose proof is nearly identical apart from having to take a more careful look at the boundary.

Using those estimates, the final proof then is identical to the one of Theorem 7.1.
Part II

Gradient Flows in spaces of currents
Chapter 11

Minimizing movements

11.1 Gradient flows

We will now shift our attention to the theory of gradient flows. The goal here is to establish an existence result and some properties for flows in the space of (Cartesian) currents. Before we begin, let us start by a somewhat abstract treatment.

In general, a gradient flow is a solution to an equation of the form

\[ \dot{u}(t) + \nabla E(u(t)) = 0 \]

where \( u \) is a time-dependent function and \( \nabla E \) is the gradient of an energy. The function \( u \) usually takes its values in a metric space \( \mathcal{A} \), which is usually a vector space or a submanifold thereof. Thus at each time, this equation lives on the tangential space at the point \( u(t) \). The choice of metric is important, as \( \nabla E \) is defined as the dual of the co-tangential quantity \( DE(u(t)) \) with respect to the metric.

A simple example of a gradient flow is the heat flow. Here we take \( \mathcal{A} := C^\infty(T^n; \mathbb{R}^N) \) to be the space of smooth functions and

\[
E(u(t)) := \frac{1}{2} \int_{T^n} \left| \frac{\partial u}{\partial x}(t) \right|^2 dx
\]

as the Dirichlet-energy. Now if we choose the \( L^2 \)-norm to generate our metric, we end up with

\[
\langle \nabla E(u(t)), \phi \rangle_{L^2} := DE(u(t))\langle \phi \rangle
= \left. \frac{\partial}{\partial s} \right|_{s=0} E(u(t) + s\phi) = \int_{T^n} \nabla u(t) \nabla \phi dx
\]

1The choice of the flat torus \( T^n \) as domain should not be seen as important at this point, it will only allow us to ignore potential issues with boundaries and integrability.
\[ -\int_{\mathbb{T}^n} \Delta u(t) \phi \, dx = \langle -\Delta u, \phi \rangle_{L^2} \]

and thus can identify
\[ \nabla E(u(t)) = -\Delta u(t). \]

This calculation then leads to the well known equation for the heat flow
\[ \dot{u}(t) = \Delta u(t). \]

Note that all three choices, of space, metric and energy, are important and any change will lead to a different equation. This is obvious for a different choice of energy, but also true for the other two. If we use the \( \dot{W}^{1,2} \)-seminorm instead of the \( L^2 \)-norm, our calculation of \( \nabla E \) changes to
\[ DE(u(t))(\phi) = \int_{\mathbb{T}^n} \nabla u(t) \nabla \phi \, dx = \langle u, \phi \rangle_{\dot{W}^{1,2}} \]

and thus
\[ \nabla E(u(t)) = u(t). \]

Similarly, if we try to restrict our space to functions with values on a manifold, for example \( C^\infty(\mathbb{T}^n; \mathbb{S}^N) \), this also changes the derivation. In this case, we are only able to test \( DE(u(t)) \) with functions that are tangential at \( u(t) \) and as a tangential quantity, \( \nabla E(u(t)) \) has to be tangential as well. After some calculations, it then turns out, that thus
\[ \nabla E(u(t)) = -\Delta u(t) - |\nabla u(t)|^2 u(t) \]

which results in the harmonic map heat flow
\[ \dot{u}(t) = \Delta u(t) + |\nabla u(t)|^2 u(t). \]

### 11.2 Abstract gradient flows

While it is easy to give a definition of a gradient flow in the case of an Hilbert-space, the concept can be extended into the much more general setting of metric spaces. To do so however, it is necessary to replace some of the definitions, as even just defining a gradient requires identification of the space with its dual. The following concepts go back to De Giorgi and coauthors [dGMT80] as well as Ambrosio [Amb90]. A detailed discussion can be found in the book [AGS08].

The general idea here is that it is possible to characterize a gradient flow using only the absolute values of the quantities involved, as a curve of maximal slope. This means that in a local sense, the gradient flow is characterized by two facts, that firstly, the speed of the curve corresponds to the speed of the change in
energy along the curve and that secondly, among all curves satisfying the first condition, the decrease of energy is the fastest. Both of those conditions can be formulated as

\[ \| \dot{u} \| = \| \nabla E(u) \| \]

\[ \frac{d}{dt} E(u(t)) = \nabla E(u)(\dot{u}) = -\| \nabla E(u) \| \| \dot{u} \|. \]

Using Young’s inequality, the two conditions can even further be combined into a single condition, since

\[ -\| \nabla E(u) \| \| \dot{u} \| = -\frac{1}{2} \| \nabla E(u) \|^2 - \frac{1}{2} \| \dot{u} \|^2 \]

if and only if \( \| \nabla E(u) \| = \| \dot{u} \| \). Thus we conclude that \( u \) is a gradient flow if and only if

\[ \frac{d}{dt} E(u(t)) = -\frac{1}{2} \| \nabla E(u) \|^2 - \frac{1}{2} \| \dot{u} \|^2 \]

In order to turn this into a definition, the next step then is to identify a suitable generalization of \( \| \dot{u} \| \) and \( \| \nabla E(u) \| \) in metric spaces. This is done via the next two definitions:

**Definition 11.1** ([AGS08, 1.1.2], [Amb90]). Let \((X,d)\) be a closed metric space, \(I \subset \mathbb{R}\) an interval. Let \( u : I \to X \) be an absolutely continuous curve. Then the metric derivative of \( u \) is given by

\[ \| \dot{u} \| (t) := \lim_{s \to t} \frac{d(u(s), u(t))}{|s - t|}. \]

This limit exists for almost every \( t \in I \).

**Definition 11.2** ([AGS08, 1.2.1, 1.2.2]). Let \((X,d)\) be a closed metric space, \( E : X \to \mathbb{R} \) an energy functional. Then \( g_E : [0, \infty] \to \mathbb{R} \) is called a strong upper gradient of \( E \) if \( g_E \circ u \) is Borel measurable for every absolutely continuous curve \( u : I \to X \), and for all those curves we have

\[ |(E \circ u)(t) - (E \circ u)(s)| \leq \int_s^t (g_E \circ u)(r) \| \dot{u} \|(r) dr \quad \forall s, t \in I \]

where \( \| \dot{u} \| \) is the metric derivative from the last definition.

Similarly \( g_E : [0, \infty] \to \mathbb{R} \) is called a weak upper gradient if for each absolutely continuous curve \( u : I \to X \) such that \( t \mapsto (g_E \circ u)(t) \| \dot{u} \|(t) \in L^1(I; \mathbb{R}) \) and \( E \circ u \) is of bounded variation, we have

\[ \left| \frac{d}{dt} (g_E \circ u)(t) \right| \leq g \circ u(t) \| \dot{u} \|(t) \quad \text{for a.a. } t \in I. \]
Together with the reasoning in the beginning of this section, this motivates the following abstract generalization of a gradient flow:

**Definition 11.3** ([AGS08, 1.3.2], see also [dGMT80]). Let \((X, d)\) be a closed metric space, \(E: X \to \mathbb{R}\) an energy functional. A locally absolutely continuous map \(u: I \to X\) is called a curve of maximal slope of \(E\) if

\[
\frac{d}{dt}(E \circ u)(t) \leq -\frac{1}{2} \|\dot{u}\|(t)^2 - \frac{1}{2} (g_E \circ u)(t)^2 \quad \text{for a.a. } t \in I.
\]

Here \(g_E\) is a weak upper gradient of \(E\).

### 11.3 Minimizing movements

In the previous section, we have identified the key ingredients for a gradient flow. In this section, we will describe a general strategy to prove existence of a gradient flow for a wide variety of those ingredients, which was first described by de Giorgi [dG93] and advanced by Ambrosio [Amb95].

A well-known method for approximating solutions to ordinary differential equations is the backwards or implicit Euler scheme. This method can be easily adapted to gradient flows. If the function at our current time step is \(u_k\), then the next time step should be a solution \(u_{k+1}\) of

\[
\frac{u_k - u_{k+1}}{h} = -\nabla E(u_{k+1}).
\]

One then readily sees that this is nothing more than the Euler-Lagrange equation of the functional

\[
F_k(u) := \frac{1}{2h} \|u_k - u\|^2 + E(u).
\]

However nothing in this functional requires the Hilbert-space setting, since it can be equally formulated in terms of the metric as

\[
F_k(u) := \frac{1}{2h} d(u_k, u)^2 + E(u).
\]

Now assuming sufficient coercivity and the usual lower semicontinuity of \(E\), it is not hard to verify that \(F_k\) has a minimizer. The minimizer \(u_{k+1}\) of this functional then is successively used to define the next functional \(F_{k+1}\) and so on. From this, we can construct a piecewise constant approximate solution to our gradient flow problem.

This general scheme is known under the name minimizing movements. What is left to do is to show that those approximate solutions converge if the step size \(h\) is sent to 0 in order to show an existence-result for the corresponding gradient flow. This has been done in several special cases such as the mean...
CHAPTER 11. MINIMIZING MOVEMENTS

Figure 11.1: An abstract view of the minimizing movement functional.

curvature flow [LS95, MSS16]. Furthermore there is a rather general theory by Ambrosio, Gigli and Savaré (see [AGS08], also [AS07]) which yields an existence result for a large class of problems. They arrive at the following result:

**Theorem 11.4** ([AGS08, 2.0.6,2.2.3,2.3.1]). Let $(X,d)$ be a metric space and $\sigma$ a Hausdorff-topology on $X$, weaker than the induced topology and such that $d$ is lower-semicontinuous in $\sigma$. Let $E : X \to \mathbb{R}$ be an energy functional. Furthermore assume that

1. The functional $E$ is $\sigma$-lower semicontinuous on bounded subsets of $X$.
2. The functional $E$ is coercive. That is $\inf_{u \in X} \frac{1}{2h} d(u,u_0)^2 + E(u) > -\infty$ for some $h > 0$ and $u_0 \in X$.\(^2\)
3. Every bounded subset of $X$ which is contained in a sub-level of $E$ is relatively sequentially compact in the $\sigma$ topology.

Then the minimizing movement scheme is well defined and the approximate solutions converge to a curve of maximal slope.

\(^2\)Using the triangle inequality and the fact that $d(\ldots) > 0$, this condition then directly implies $\inf_{u \in X} \frac{1}{2h} d(u, \tilde{u})^2 + E(u) > -\infty$ for all $\tilde{u} \in X$ and $\tilde{h} < h$. 
Chapter 12

The case of currents

After the overview of the general situation in the last chapter, we will now focus on the special case of (Cartesian) currents. While this situation is already covered by the general existence theorem 11.4, we will see that we can utilize the compactness properties of currents to show convergence of the minimizing movement scheme in a stronger sense. For now we will again deal with the boundary-less case. (See also Remark 12.16)

12.1 The minimizing movement iteration for currents

Let us now take a look at how the minimizing movement iterations works on currents. Here the underlying topology $\sigma$ as in Theorem 11.4 is the usual convergence of currents. This gives us access to the compactness properties that we need. However in order for a set of currents to be relatively compact, we additionally need a bound on the mass. To achieve this, we will slightly extend the required coercivity of our energy.

Definition 12.1. Let $A \subset D_k(\Omega)$. A functional $E : A \to \mathbb{R}$ is called coercive if all sub-level-sets are bounded in mass, that is

$$\sup\{M(T) \mid T \in A, E(T) < E_0\} < \infty \quad \text{for all } E_0 \in \mathbb{R}.$$  

We note that this is similar to the notion of coercivity for Cartesian currents (see Definition 4.25).

Remark 12.2. We also note that this not only results in the relevant compactness property, but also implies the coercivity of Theorem 11.4 for lower-semicontinuous energies, as any sub-level set then is compact and thus has a minimizer with finite energy. Therefore $E$ is then bounded from below. ■

This is all we need to show that the minimizing movement iteration is well defined for currents:
Proposition 12.3. Let $\mathcal{A} \subset D_k(\Omega)$ be a closed set of currents, $d : \mathcal{A} \times \mathcal{A} \to \mathbb{R}_{\geq 0}$ a metric and $E : \mathcal{A} \to \mathbb{R}$ a coercive energy functional. If $d$ and $E$ are lower-semicontinuous with respect to convergence of currents, then the functional

$$F^{(h)}_{T_0} : T \mapsto \frac{1}{2h} d(T, T_0)^2 + E(T)$$

has a minimizer for all $T_0 \in \mathcal{A}$ and all $h > 0$. In other words, the minimizing movement iteration is well defined.

Proof. We proceed by the direct method as usual. Fix $h > 0$, $T_0 \in \mathcal{A}$. Let $(T_i)_{i \in \mathbb{N}} \subset \mathcal{A}$ be a minimizing sequence, that is

$$\lim_{i \to \infty} F^{(h)}_{T_0}(T_i) = \inf_{T \in \mathcal{A}} F^{(h)}_{T_0}(T).$$

Since $\frac{1}{2h} d(T, T_0)^2$ is non-negative, we know that

$$\lim_{i \to \infty} E(T_i) \leq \lim_{i \to \infty} F^{(h)}_{T_0}(T_i) = \inf_{T \in \mathcal{A}} F^{(h)}_{T_0}(T) < \infty$$

and thus the energy is bounded along the sequence. But then by coercivity of $E$, we know that

$$\sup_{i \in \mathbb{N}} \mathcal{M}(T_i) < \infty$$

and thus there is a converging subsequence $T_i \to T \in \mathcal{A}$ (as usual not relabeled). But then by the lower semicontinuity of $d$ and $E$, we get

$$F^{(h)}_{T_0}(T) \leq \liminf_{i \to \infty} F^{(h)}_{T_0}(T_i) = \inf_{T \in \mathcal{A}} F^{(h)}_{T_0}(T).$$

Thus we have found our minimizer. \qed

Remark 12.4. It should be noted that this minimizer does not have to be unique. As should be expected even the resulting gradient flow problem will not necessarily have a unique solution. This is a consequence of the fact that there is no real gradient anymore and as result there may be more than one direction of steepest descent.

To finish this section, we will highlight a few additional properties of this iteration. The first of which is fundamental for our convergence theorem.

Lemma 12.5. Assume the conditions of Proposition 12.3 and denote a minimizer of $F^{(h)}_{T_0}$ by $T_1$. Then $E(T_1) \leq E(T_0)$ with equality only if $T_1 = T_0$ and

$$d(T_0, T_1) \leq \sqrt{2h} \sqrt{E(T_0) - E(T_1)}.$$
Proof. First we calculate

\[ F_{T_0}^{(h)}(T_0) = \frac{1}{2h} d(T_0, T_0)^2 + E(T_0) = E(T_0). \]

but then since \( T_1 \) is the minimizer, we have

\[ \frac{1}{2h} d(T_0, T_1)^2 + E(T_1) = F_{T_0}^{(h)}(T_1) \leq F_{T_0}^{(h)}(T_0) = E(T_0). \]

Then the first inequality is a direct consequence of the fact that \( \frac{1}{2h} d(T_0, T_1)^2 \geq 0 \) and the second inequality results from solving the equation for \( d(T_0, T_1) \).

Lemma 12.6. Assume the conditions of Proposition 12.3 and denote a minimizer of \( F_{T_0}^{(h)} \) by \( T_h \). Then \( T_h \to T_0 \) in \( d \) and thus also in the topology of currents.

Proof. From the previous lemma we know that \( E(T_h) \leq E(T_0) \). Furthermore we know that the energy is bounded from below by some \( E_- \). But then

\[ d(T_0, T_h) \leq \sqrt{2h(2h)} \leq \sqrt{2h(2h)} \leq \sqrt{2h(2h)} \to 0. \]

12.2 Space-time solution

Let us now elaborate on how we will improve our notion of a solution. The general idea here will be inspired by an observation similar to the idea behind Cartesian currents. Let \( S : I \to \mathbb{I}_k(\Omega) \) be a solution to some gradient-flow problem. Then as discussed in Chapter 6, assuming enough regularity, the graph of \( S \) can also be interpreted as a rectifiable \( k + 1 \)-current \( \tilde{S} \in \mathbb{I}_{k+1}(I \times \Omega) \). We will then call \( \tilde{S} \) a space-time solution to the gradient flow problem.

Also, as we did when considering the space-time traces, given the current \( \tilde{S} \), we are able to recover the function \( S \) by slicing along constant times:

\[ \tilde{S}, t < t_0 := \partial \left( \tilde{S} \cup (-\infty, t_0) \times \Omega \right) = \{ t_0 \} \times S(t_0) \]

This will be of special importance, since we will not generate our space-time solution out of the graph of a function but rather as the limit of the space-time analogue of step functions.

Furthermore note, that in Lemma 9.4 we were using the approximate tangential space of space-time currents to recover the vector fields associated with the derivative of the corresponding curve. Especially in the case of Cartesian currents, when combining this with the identification of the vector field as a
derivative of the functional part in Lemma 5.9, this will allow us to derive some additional results which would not have been available in the abstract setting.

However, before we get to finding actual space-time solutions we want to define the space-time analogue of piecewise constant functions. For this, we need an additional condition on the metric, which is important in the convergence theorem.

**Definition 12.7.** Let $\mathcal{A} \subset I_k(\Omega)$, together with a metric $d$ on $\mathcal{A}$. We say that $d$ is compatible with $\hat{\mathcal{F}}$ if there is a constant $C_0$ such that

$$\hat{\mathcal{F}}(T_0, T_1) \leq C_0 d(T_0, T_1)$$

for all $T_0, T_1 \in \mathcal{A}$. Similarly we say $d$ is compatible with $\mathcal{F}, \mathcal{F}$ and so on, if the corresponding condition holds.

**Remark 12.8.** Note that this implies that the topology of currents is weaker than the topology induced by $d$, which was one of the conditions in Theorem 11.4. In a way, this is the appropriate variant of this condition, since our topology has a metric. However, strictly speaking this is a stronger requirement than having a weaker topology, as the constant involved is uniform. ■

Now for any compatible metric, we are able to define a useful notion of piecewise constant space-time functions.

**Definition 12.9.** Let $\mathcal{A} \subset I_k(\Omega)$ closed, let $E : \mathcal{A} \to \mathbb{R}$ be an energy and $d$ a metric which is compatible with $\hat{\mathcal{F}}$. Let $h > 0$, $I = (0, \tau)$ be an interval. If $\mathcal{A}, d$ and $E$ satisfy the conditions of Proposition 12.3, then the minimizing movement iteration has a sequence of minimizers we denote by $(T_i^{(h)})_{i \in \mathbb{N}}$.

For this sequence we define the corresponding piecewise constant space-time current $S^{(h)} \in I_{k+1}(I \times \Omega)$ by

$$S^{(h)} := \frac{\tau}{h} \sum_{i=0}^{\lfloor \frac{\tau}{h} \rfloor} \int [h_i, h(i + 1)] \times T_i^{(h)} + \frac{\tau}{h} \sum_{i=1}^{\lfloor \frac{\tau}{h} \rfloor} \int \{h_i\} \times V_i^{(h)}.$$

Here $V_i^{(h)} \in I_{k+1}(\Omega)$ is defined as a minimizing current for $\hat{\mathcal{F}}(T_i^{(h)} - T_{i-1}^{(h)})$, that is

$$\partial V_i^{(h)} = T_i^{(h)} - T_{i-1}^{(h)} \text{ and } \mathcal{M}(V_i^{(h)}) = \hat{\mathcal{F}}(T_i^{(h)} - T_{i-1}^{(h)}).$$

We immediately can see the following properties:

---

Note that this implies that the topology of currents is weaker than the topology induced by $d$, which was one of the conditions in Theorem 11.4. In a way, this is the appropriate variant of this condition, since our topology has a metric. However, strictly speaking this is a stronger requirement than having a weaker topology, as the constant involved is uniform. ■

Now for any compatible metric, we are able to define a useful notion of piecewise constant space-time functions.

**Definition 12.9.** Let $\mathcal{A} \subset I_k(\Omega)$ closed, let $E : \mathcal{A} \to \mathbb{R}$ be an energy and $d$ a metric which is compatible with $\hat{\mathcal{F}}$. Let $h > 0$, $I = (0, \tau)$ be an interval. If $\mathcal{A}, d$ and $E$ satisfy the conditions of Proposition 12.3, then the minimizing movement iteration has a sequence of minimizers we denote by $(T_i^{(h)})_{i \in \mathbb{N}}$.

For this sequence we define the corresponding piecewise constant space-time current $S^{(h)} \in I_{k+1}(I \times \Omega)$ by

$$S^{(h)} := \frac{\tau}{h} \sum_{i=0}^{\lfloor \frac{\tau}{h} \rfloor} \int [h_i, h(i + 1)] \times T_i^{(h)} + \frac{\tau}{h} \sum_{i=1}^{\lfloor \frac{\tau}{h} \rfloor} \int \{h_i\} \times V_i^{(h)}.$$

Here $V_i^{(h)} \in I_{k+1}(\Omega)$ is defined as a minimizing current for $\hat{\mathcal{F}}(T_i^{(h)} - T_{i-1}^{(h)})$, that is

$$\partial V_i^{(h)} = T_i^{(h)} - T_{i-1}^{(h)} \text{ and } \mathcal{M}(V_i^{(h)}) = \hat{\mathcal{F}}(T_i^{(h)} - T_{i-1}^{(h)}).$$

We immediately can see the following properties:

---

$^1$Here and in the following we clandestinely assume that $\frac{\tau}{h}$ is an integer. Otherwise, we would always have to round down and add a last shortened summand. This does not impact the general concept but will needlessly complicate the notation, especially in light of the fact that we are free to choose sub-sequences $h_k \to 0$ in the next chapter anyway.
Figure 12.1: The construction of space-time solutions and Theorem 12.12. (a) We consider the graph of the piecewise constant approximation. (b) We add vertical parts in order to remove the boundary. In this way we construct the current \( S(h) \) in Definition 12.9. (c) Now if \( h \) gets smaller, the total mass of \( S(h) \) is still uniformly bounded and thus (d) the current will converge to a space-time solution \( S \).

**Lemma 12.10.** Let \( S(h) \) as in the preceding definition. Then

1. \( S(h) \) is well defined as a current in \( I_{k+1}(I \times \Omega) \)
2. \( \partial S(h) = 0 \)
3. \( \langle S(h), t < t_0 \rangle = \partial(S(h)_L(-\infty, t) \times \Omega) = \| \{t_0\} \| \times T_i(h) \) where \( i \) is such that \( t_0 \in (h_i, h(i + 1)] \).

**Proof.**

1. Since we require \( d \) to be compatible with \( \hat{F} \), we know that the “connecting” currents \( V_i(h) \) exist by Lemma 2.6. Now \( S(h) \) is a finite sum of integer rectifiable currents in \( I \times \Omega \). Thus \( S(h) \) is itself in \( I_k(\Omega) \).

2. The boundaries of the horizontal parts are given by

\[
\partial [\{h_i, h(i + 1)\}] \times T_i(h) = [\{h(i + 1)\}] \times T_i(h) - [\{h_i\}] \times T_i(h),
\]
except for the first and last index $i$, where the first (or respectively the second) term vanishes since it lies on the boundary of $I \times \Omega$. Similarly, the boundaries of the vertical parts are given by

$$\partial \llbracket \{hi\} \rrbracket \times V_i^{(h)} = \llbracket \{hi\} \rrbracket \times T_i^{(h)} - \llbracket \{hi\} \rrbracket \times T_{i-1}^{(h)}.$$ 

Summing over $i$, all those terms cancel.

3. The first equality is a direct consequence of $\partial S(h) = 0$. For the second one, we notice that all the terms $\llbracket \{hi\} \rrbracket \times V_i^{(h)}$ either do not intersect $(-\infty, t_0] \times \Omega$ or are completely in the interior. Thus they do not contribute. The same is true for all the horizontal parts $\llbracket \{hi, h(i+1)\} \rrbracket \times T_i^{(h)}$, except where $hi < t_0 \leq h(i+1)$. But then

$$\left(\llbracket \{hi, h(i+1)\} \rrbracket \times T_i^{(h)}, t < t_0\right) = \llbracket \{t_0\} \rrbracket \times T_i^{(h)}.$$ 

Remark 12.11. It is possible to extend the definition to $\tau = \infty$, that is the interval $I = (0, \infty)$. In this case $S^{(h)}$ is only locally integer rectifiable as there is no global mass bound unless $M(T_i^{(h)}) \to 0$ sufficiently fast, since

$$M\left(\sum_{i=0}^{\infty} \llbracket \{hi, h(i+1)\} \rrbracket \times T_i^{(h)}\right) = \sum_{i=0}^{\infty} h M(T_i^{(h)}),$$

which does not necessarily converge. (However there is in fact a global mass bound on the vertical parts, which we will see in the next section.)

While it is completely possible to work with locally integrable currents to directly derive the same theory for $\tau = \infty$, we will restrain ourselves to finite intervals for now, as this case is easier to handle.

12.3 The convergence theorem for currents

Let us now state and proof the basic convergence theorem:

**Theorem 12.12** (Convergence theorem, basic version). Let $A \subset I_k(\Omega)$ be closed. Let $E : A \to \mathbb{R}$ be coercive and lower-semicontinuous. Let $d$ be a metric that is lower-semicontinuous and compatible with $\mathcal{F}$. Then for any initial value $T_0$ and any final time $\tau$, there is a sequence of step-sizes $h_j \to 0$ such that the space-time step functions $S^{(h_j)}$ converge to a space-time solution $S$ in $[0, \tau] \times \Omega$.

**Proof.** As mentioned above, let us initially pick the sequence $h_j := \tau / j$, so that $\tau / h_j$ always is an integer. Again as before we use $(T^{(h_j)}_i)_{i \in \mathbb{N}}$ to denote the minimizing movement iteration for a fixed $h_j$. Such a sequence exists by
Proposition 12.3. The proof now is an application of the compactness-closure theorem. For this, we need to show that the mass of the $S^{(h_j)}$ stays bounded.

For the horizontal components, we note that $E\left(T^{(h_j)}_i\right) \leq E(T_0)$ holds for all step sizes and steps by Lemma 12.5. Thus we are always in the sub-level set \(\{T \in \mathcal{A} | E(T) \leq E(T_0)\}\) and so by coercivity, there exists a uniform bound \(\mathcal{M}\left(T^{(h_j)}_i\right) \leq M_0\) on the mass.

Then the estimate is simple, as
\[
\mathcal{M}\left(\left\|h_ji, h_j(i+1)\right\| \times T^{(h_j)}_i\right) = h_j \mathcal{M}\left(T^{(h_j)}_i\right) \leq h_j M_0
\]
and thus
\[
\mathcal{M}\left(\sum_{i=0}^{\tau/h_j} \left\|h_ji, h_j(i+1)\right\| \times T^{(h_j)}_i\right) \leq \sum_{i=0}^{\tau/h_j} h_j M_0 = \tau M_0.
\]

For the vertical parts however we need to bound the total $\dot{\mathcal{F}}$-distance summed over all the steps, independently of $h_j$. Using Lemma 12.5, we have
\[
d\left(T^{(h_j)}_{i+1}, T^{(h_j)}_i\right)^2 \leq 2h_j \left(E\left(T^{(h_j)}_i\right) - E\left(T^{(h_j)}_{i+1}\right)\right).
\]

Using this and the compatibility-condition, we then end up with
\[
\mathcal{M}\left(\sum_{i=1}^{\tau/h_j} \{h_ji\} \times V^{(h_j)}_i\right) \leq C_0 \sum_{i=1}^{\tau/h_j} 2h_j \sqrt{E\left(T^{(h_j)}_i\right) - E\left(T^{(h_j)}_{i+1}\right)}
\]
and using the Hölder-inequality
\[
\leq C_0 \sqrt{\sum_{i=1}^{\tau/h_j} 2h_j \sum_{i=1}^{\tau/h_j} E\left(T^{(h_j)}_i\right) - E\left(T^{(h_j)}_{i+1}\right)}
\]
\[
\leq C_0 \sqrt{2\tau \left(E(T_0) - \inf_{\mathcal{A}} E\right)}.
\]

But then, since $\mathcal{M}(S^{(h_j)})$ is thus uniformly bounded and $\partial S^{(h)} = 0$, there exists a converging subsequence by Corollary 2.11.

Remark 12.13. According to Lemma 12.10 (3.) we know that
\[
\left\langle S^{(h_k)}, t < t_0 \right\rangle = [t_0] \times T^{(h_k)}_{i_k}
\]
for a matching $i_k$. But then the continuity of slicing implies that
\[
\left\langle S^{(h_k)}, t < t_0 \right\rangle \to \langle S, t > t_0 \rangle \text{ for } k \to \infty.
\]
Projecting this along the $t$-axis, we obtain

$$T^{(h_k)}_{i_k} \rightarrow \pi_{x#}(S, t < t_0).$$

So in fact the piecewise approximation converges pointwise against

$$T : t \mapsto \pi_{x#}(S, t < t_0).$$

As a direct consequence of this observation we get:

**Corollary 12.14.** Let $S$ be the resulting current in Theorem 12.12. Then the map

$$T : I \rightarrow I_k(\Omega); t_0 \mapsto \langle \pi_x \rangle_{S, t < t_0}$$

is a curve of maximal slope for $E$ in $(d, A)$.

**Proof.** Studying the proof of Theorem 11.4 more closely (see [AGS08, Chp. 2]), the curve of maximal slope is constructed as the limit of our piecewise constant curves. Since we identified $\pi_{x#}(S, t < t_0)$ as the same limit in the preceding remark, the corollary then follows.

Now that we have convergence, we can try to establish some additional properties of the solution

**Lemma 12.15.** Let $S$ be the resulting current in Theorem 12.12, $T(t) := \pi_{x#}(S, t < t_0)$. Then the following holds:

1. $E(T(t))$ is monotonously decreasing in $t$.
2. $T$ is $1/2$-Hölder continuous in $d$ with a constant of $\sqrt{2(E_0 - \inf_A E)}$.
3. $S$ has not vertical parts in time. That is for any $k$-form $\omega \in C^\infty(\Omega; \wedge^k)$, the measure on $(0, \tau)$ defined by

$$\phi \mapsto S(\phi dt \wedge \omega)$$

is absolutely continuous.

**Proof.** 1. This is a consequence of [AGS08], Thm 2.3.3, especially in light of Lemma 3.1.2. The direct calculation however is a bit arduous.

2. Pick $s, t \in I$, $s > t$. Fix $h > 0$, such that $h << s - t$. Now let $s \in [hi, h(i+1))$ and $t \in [hj, h(j+1))$. Now from Lemma 12.5 we get, as in the proof of Theorem 12.12 that

$$d\left(T_i^{(h)}, T_j^{(h)}\right) \leq \sum_{l=i}^{j-1} d\left(T_l^{(h)}, T_{l+1}^{(h)}\right)$$
\[
\leq \sum_{l=i}^{j-1} 2h \sqrt{E\left(T_{l+1}^{(h)}\right) - E\left(T_{l}^{(h)}\right)}
\]
\[
\leq \sqrt{\sum_{l=i}^{j-1} 2h \sum_{l=i}^{j-1} E\left(T_{l+1}^{(h)}\right) - E\left(T_{l}^{(h)}\right)}
\]
\[
\leq \sqrt{(s-t) + 2h \sqrt{2 \left(E_0 - \inf_{A} E\right)}}
\]

Taking the limit \( h \to 0 \) we thus get, using the lower-semicontinuity of \( d \)

\[
d(T(t), T(s)) \leq \sqrt{s-t} \sqrt{2 \left(E_0 - \inf_{A} E\right)}
\]

and thus the Hölder-continuity.

3. Let \( \omega \in C^\infty(\Omega; \wedge^k) \) fixed and define \( \mu : \phi \mapsto S(\phi dt \wedge \omega) \). Now pick \( s, t \in I, s > t \). Note that from the mass estimates used in Theorem 12.12, we get

\[
\mathcal{M}(S \mathbb{L} (t, s) \times \Omega) \leq M_0(s - t) + C_0 \sqrt{s-t} \sqrt{2 \left(E_0 - \inf_{A} E\right)}
\]

similarly to the last point. But then \( \|dt \wedge \omega\| \leq \|\omega\| \) we get

\[
\mu((t, s)) \leq \|\omega\| \mathcal{M}(S \mathbb{L} (t, s) \times \Omega) \to 0
\]

uniformly for \( |s - t| \to 0 \). Thus \( \mu \) is absolutely continuous.

Remark 12.16. While we again have focused the case of a boundary-less current, as before, this was only in order to simplify the exposition. It is not hard to see the appropriate changes that need to be made. The coercivity of course needs to include the mass of the boundary in addition to the mass of the actual current. In the same vein, we need compatibility with \( \mathcal{F} \) instead of \( \mathcal{F} \). Then we can repeat the previous calculations with an additional boundary term, which will behave in much of the same way as the rest. The resulting space-time solution will now of course have a boundary, and the time slices of this in turn will describe the evolution of the actual boundary.

12.4 The convergence theorem in the Cartesian case

The subclass of currents in which we are most interested, is the class of Cartesian currents. As they are integer rectifiable, of course the convergence theorem applies to them as well. However, using their additional structure we
can in fact derive an improved version. The main idea here is a current in
\( T \in \text{cart}(\Omega; \mathbb{R}^N) \) consists of the graph of a function \( u_T : \Omega \rightarrow \mathbb{R}^n \) and some
additional vertical parts, that is it has the form

\[
T = \mathcal{G}(u_T) + V_T.
\]

But then a function \( S : I \rightarrow \text{cart}(\Omega; \mathbb{R}^N) \) consequently gives rise to a corre-
sponding function

\[
u_S : I \times \Omega \rightarrow \mathbb{R}^N, (t, x) \rightarrow u_{S(t)}(x).
\]

Treating the vertical parts in the same way, one can assume that
\( S \) then gives
rise to a Cartesian current in \( \text{cart}(I \times \Omega; \mathbb{R}^N) \), given sufficient regularity.

This is exactly what happens in the convergence theorem. If we work on the
class of Cartesian currents, we will always end up with a Cartesian current in
space time, as the following theorem shows.

**Theorem 12.17** (Convergence theorem, Cartesian currents version). Let \( A \subset \text{cart}(\Omega; \mathbb{R}^N) \) be closed, bounded in mass by \( M_0 \) and with \( \| \|_{L^1} < L_0 \). Let \( E : A \rightarrow \mathbb{R} \) be coercive and lower-semicontinuous. Let \( d \) be a metric that is compatible with \( \dot{F} \). Then for any final time \( \tau \), there is a sequence of step-sizes \( h_k \rightarrow 0 \) such that the space-time step functions \( S^{(h_k)} \) converge to a space-time solution \( S \) in
\([0, \tau] \times \Omega \), where \( S \in \text{cart}(I \times \Omega; \mathbb{R}^N) \).

**Proof.** Since \( \text{cart}(\Omega; \mathbb{R}^N) \) is a subset of \( I_k(\Omega \times \mathbb{R}^N) \), this theorem is just a special case of Theorem 12.12. The only additional thing we need to show is that we indeed have \( S \in \text{cart}(I \times \Omega; \mathbb{R}^N) \). To do so, we need to verify the
conditions in Definition 4.6.

1. \( \mathcal{M}(S) < \infty \) is a direct consequence of \( S \in I_k(\Omega \times \mathbb{R}^n) \). For \( \| S \|_{L^1} \), we note that \( \| S \|_{L^1} \leq \liminf_{k \rightarrow \infty} \| S^{(h_k)} \|_{L^1} \) and that

\[
\| S^{(h_k)} \|_{L^1} = \sum_{i=0}^{\tau/h_k} \| T_i^{(h_k)} \|_{L^1} \leq \tau L_0
\]

2. \( \partial S = 0 \) holds as before.

3. By continuity

\[
(\pi_{t,x})_# S = \lim_{k \rightarrow \infty} (\pi_{t,x})_# S^{(h_k)}
\]

\[
= \lim_{k \rightarrow \infty} \sum_{i=0}^{\tau/h_k} \left( (h_k i, h_k (i + 1)) \times S^{(h_k)} \right)
\]

\[
= \lim_{k \rightarrow \infty} \sum_{i=0}^{\tau/h_k} [h_k i, h_k (i + 1)] \times [\Omega]
\]

\[
= \lim_{k \rightarrow \infty} [I \times \Omega] = [I \times \Omega]
\]
4. Assume that the signed measure $\phi \mapsto S\langle \phi dt \wedge dx^1 \wedge ... \wedge dx^n \rangle$ is negative at some point. In other words, there is a non-negative function $\phi : I \times \Omega \times \mathbb{R}^n$ with compact support such that $S\langle \phi dt \wedge dx^1 \wedge ... \wedge dx^n \rangle < 0$.

Then by the convergence of $S^{(h_k)}$, there is a $k \in \mathbb{N}$ large enough such that $S^{(h_k)}\langle \phi dt \wedge dx^1 \wedge ... \wedge dx^n \rangle < 0$. So we know that

$$S^{(h_k)}\langle \phi dt \wedge dx^1 \wedge ... \wedge dx^n \rangle$$

$$= \sum_{i=0}^{\tau/h_k} \left( [[h_k, h_k(i + 1)] \times S^{(h_k)} \langle \phi dt \wedge dx^1 \wedge ... \wedge dx^n \rangle < 0ight)$$

and thus there is at least one negative summand. However then

$$\left( [[h_k, h_k(i + 1)] \times S^{(h_k)} \langle \phi dt \wedge dx^1 \wedge ... \wedge dx^n \rangle \right)$$

$$= \int_{h_k}^{h_k(i+1)} T^{(h_k)}_i \langle \phi|_{t=s} dx^1 \wedge ... \wedge dx^n \rangle ds < 0$$

so there must be $s \in [h_k, h_k(i + 1)]$ for which

$$T^{(h_k)}_i \langle \phi|_{t=s} dx^1 \wedge ... \wedge dx^n \rangle < 0$$

which is a contradiction since $\phi|_{t=s}$ is non-negative with compact support in $\Omega \times \mathbb{R}^n$ and $T^{(h_k)}_i \in \text{cart}(\Omega; \mathbb{R}^n)$ implies that the corresponding measure is non-negative. 

\[ \square \]

**Remark 12.18.** The theorem also holds for $\text{cart}(\Omega; \mathcal{M})$. In this case, the $L^1$ bound is trivial, as, assuming that by translation $0 \in \mathcal{M}$, we have the universal bound $\|T\|_{L^1} \leq |\Omega| \text{diam } \mathcal{M}$. The resulting space-time current will be in $\text{cart}(I \times \Omega; \mathcal{M})$. 

\[ \blacksquare \]
Chapter 13

Gradient flows in the generalized Wasserstein distance

In this chapter we will try to sketch how to apply the considerations from the previous chapter to our generalized Wasserstein distance, specifically $W_2$, and remark on some possible applications. This of course will assume lower semicontinuity as stated in Conjecture 9.8. The idea here is similar to Lemma 9.4, as again the part of the tangential space that extends in time direction allows us to extract a vector field, which will be the vector field along which the current is actually moved. The estimates for the Wasserstein distance will then allow us to bound the $L^2$ norm of this vector field. But first we need to modify the minimizing movement scheme in order achieve better estimates.

13.1 Minimizing movements for the 2-Wasserstein distance

It is of course possible to simply apply the convergence theorem of Chapter 12 to the generalized Wasserstein directly. This leads to:

**Corollary 13.1.** Let $\mathcal{A} \subset I_k(\Omega)$ closed and bounded in mass. Let $E : \mathcal{A} \to \mathbb{R}$ be a lower semicontinuous energy functional. Then for any time interval $[0, \tau]$ there exists a space-time solution to the gradient flow problem for $E$ and $(\mathcal{A}, W_{2,\mathcal{A}})$ for all initial values $T_0 \in \mathcal{A}$.

**Proof.** In order to apply Theorem 12.12, we need to check the missing conditions. Since the space $\mathcal{A}$ is bounded in mass, all sub-level sets of $E$ are trivially bounded in mass, so $E$ is necessarily coercive. The metric $W_2$ is lower semicontinuous in the topology of currents assuming Conjecture 9.8. Finally
the compatibility condition results from Lemma 5.5 which together with the mass bound allows us to estimate against the $W_1$ distance, which then in turn gives us the $\mathcal{F}$ bound we need by applying Corollary 6.12.

There are however some improvements to this, which make use of the specific structure of our generalized Wasserstein distance. First of all, even while there are no geodesics, the Wasserstein distance comes directly equipped with connecting curves of currents. Thus it makes sense to change from a piecewise constant approximation to a series of curves.

**Definition 13.2.** Let $\mathcal{A} \subset I_k(\Omega)$ closed and let $E : \mathcal{A} \to \mathbb{R}$ be an energy. Let $h > 0$, $\varepsilon > 0$, $I = (0, \tau)$ be an interval.¹ If $\mathcal{A}$, $\mathcal{W}_2$ and $E$ satisfy the conditions of Proposition 12.3, then the minimizing movement iteration has a sequence of minimizers we denote by $(T^{(h)}_i)_{i \in \mathbb{N}}$. For this sequence we define an $\varepsilon$-approximate space-time current $S^{(h)}_\varepsilon \in I_{k+1}(I \times \Omega)$ by

$$S^{(h)}_\varepsilon := (id, T^{(h)}_\varepsilon)\#(I).$$

Here $T^{(h)}_\varepsilon : [0, \tau] \to \mathcal{A}$ is a curve such that $T^{(h)}_\varepsilon(ih) = T^{(h)}_i$ for all $i \in \{0, ..., \tau/h\}$

and

$$L_2\left(\left|T^{(h)}_\varepsilon|_{[i-1,i)h]}\right| \leq \mathcal{W}_2\left(T^{(h)}_i, T^{(h)}_{i-1}\right) + h\varepsilon \text{ for all } i \in \{1, ..., \tau/h\}.$$

The curve $T^{(h)}_\varepsilon$ is by no means unique, however its existence is guaranteed by definition and in the end we are only interested in the limit $\varepsilon \to 0$, $h \to 0$. For this limit we note the following corresponding variant of Theorem 12.12:

**Proposition 13.3.** Let $\mathcal{A} \subset I_k(\Omega)$ closed, $T_0 \in \mathcal{A}$ and let $E : \mathcal{A} \to \mathbb{R}$ be an energy. Let $I = (0, \tau)$ be an interval. Then for any $h > 0$, $\varepsilon > 0$, we have

$$\mathcal{M}\left(S^{(h)}_\varepsilon\right) \leq \tau(M_0 + C_0\varepsilon) + C_0\sqrt{2\tau E(T_0) - \inf_{\mathcal{A}} E}$$

where $M_0$ is the corresponding mass bound for $\{T \in \mathcal{A} \mid E(T) \leq E(T_0)\}$ and $C_0$ depends only on $M_0$. As a result any sequence $h_j \to 0$, $\varepsilon_j \to 0$ has subsequence, for which $S^{(h_j)}_{\varepsilon_j} \to S$.

**Proof.** The necessary estimate is found in Lemma 6.8. From there we get that

$$\mathcal{M}\left(S^{(h)}_\varepsilon\right) \leq \int_I \int \sqrt{1 + |\nu|^2} d\left\|T^{(h)}_\varepsilon(t)\right\| dt$$

¹As in the last chapter, we assume $\xi$ to be an integer in order to simplify notation.
where $v$ is the usual corresponding vector field such that $\dot{T}_\varepsilon^{(h)} + \mathcal{L}_{v(t)} T_\varepsilon^{(h)} = 0$.

But then we get

$$\leq \int_1 \int 1 + |v| d\left\|T_\varepsilon^{(h)}(t)\right\| dt$$

and using Lemma 5.5

$$\leq \tau M_0 + L_1 \left(T_\varepsilon^{(h)}\right) \leq \tau M_0 + C_0 L_2 \left(T_\varepsilon^{(h)}\right)$$

$$\leq \tau M_0 + C_0 \sum_{i=1}^{\tau/h} \left(\mathcal{W}_2 \left(T_\varepsilon^{(h)}(i), T_\varepsilon^{(h)}(i+1)\right) + \varepsilon h\right)$$

$$\leq \tau M_0 + C_0 \sum_{i=1}^{\tau/h} \sqrt{2\varepsilon} \sqrt{E(T_\varepsilon^{(h)}(i)) - E(T_\varepsilon^{(h)}(i+1)) + C_0 \tau \varepsilon}$$

$$\leq \tau(M_0 + C_0 \varepsilon) + C_0 \sqrt{2\tau} \sqrt{E(T_0) - \inf_A E}$$

by our usual estimates. Since $\partial S_\varepsilon^{(h)} = 0$ the rest of the proposition then is a result of the compactness closure theorem.

**Remark 13.4.** We have already remarked that in general we cannot expect any uniqueness of the limit. However it should be noted that this variant also does not really add any further non-uniqueness. What we are doing for constructing $S_\varepsilon^{(h)}$ is advancing both in time as well as along the curve $T_\varepsilon^{(h)}$. What we could do just as well is alternate between the two, that is advance time from $h(i)$ to $h(i+1)$ and then fix the time variable while we advance from $T_\varepsilon^{(h)}(i)$ to $T_\varepsilon^{(h)}(h(i+1))$. This would result in a construction similar to $S^{(h)}$ from the previous chapter (albeit not necessarily using the shortest connecting vertical part).

Both ways of doing so are obviously homotopic to each other, as transforming between them is similar an inner variation (compare also Lemma 6.4). Furthermore at each time slice, this homotopy only needs to move $T_\varepsilon^{(h)}(i)$ to $T_\varepsilon^{(h)}(t)$ for $t \in [h(i), h(i+1)]$, so its length will converge to $0$ for $h \to 0$. As a result both will converge to the same limit.

The next improvement we need to make is again related to the reoccurring problem of degenerative distance (see Example 5.4 and 5.11). We noted in Example 9.1 that any excess mass leads to a shortening of distances. While the generalized Wasserstein distance allows setting a mass bound by restricting to the subset $A$, this ignores the interplay between energy and mass.

One of the basic properties of our energy we require is the coercivity, so any bound on energy results in a mass bound. Furthermore one of the fundamental properties of the gradient flow is the monotonicity of energy with time. So
while the energy and thus the mass of the intermediate points $T_i^{(h)}$ will decrease, in our basic scheme, the curves connecting $T_i^{(h)}$ to $T_{i-1}^{(h)}$ will not respect these improved bounds. Luckily this can be easily incorporated as well:

**Proposition 13.5.** Let $A \subset I_k(\Omega)$ closed and $E$ be a coercive, lower semi-continuous energy. For any $E_0 \in \mathbb{R}$ define the set

$$A_{E_0} := \{ T \in A | E(T) \leq E_0 \}.$$ 

Then the corrected minimizing movement iteration, given by

$$T_i^{(h)} := \arg \min \left\{ \frac{1}{2h} \mathcal{W}_{2, A_{E_i}} \left( T, T_i^{(h)} \right) + E(T) \bigg| T \in A_{E_i} \right\}$$

with $E_i := E \left( T_i^{(h)} \right)$, is well defined.

**Proof.** The only thing that really changed compared to the previous version of this in Proposition 12.3 is the added energy bound. We already know that $\mathcal{W}_{2, A}$ is lower-semicontinuous and compatible with $\bar{F}$. Since $A$ is closed and $E$ is lower-semicontinuous, $A_{E_0}$ is closed for all $E_0 \in \mathbb{R}$ as well. Finally $T_i^{(h)} \in A_{E_0}$, trivially, so the minimization problem is not empty. Thus there is a minimizer. \hfill \qed

**Remark 13.6.** The main reason for applying Lemma 9.4 we can now use the construction of the previous section to recover a corresponding vector field. This is especially relevant in the case of Cartesian currents, where the vector field corresponds to the derivative of the functional part (See Lemma 5.9). One should however keep in mind that the recovered vector field need not longer be completely vertical, see Remark 9.7.

**13.2 Possible applications**

Finally let us sketch two possible applications of all of this.
CHAPTER 13. GRADIENT FLOWS IN $W_2$

Figure 13.2: A possible problem with excess energy during interpolation. (a) While the energy of the intermediate points $T^{(h)}_t$ is decreasing, (b) this is not true for the interpolating curves. To avoid this, (c) we restrict the energy to its initial value in each step.

The first one was already mentioned several times. We utilized the two-dimensional harmonic map heat flow as a kind of sounding board throughout the text and comparing what we did to the general result by Struwe [Str85], most things fit nicely. We have already seen that the generalized Wasserstein $2$-distance closely fits the $L^2$-distance which is the metric used in this flow. In the two dimensional situation the currents only have vertical parts in the form of full spheres, which are the result of bubbling and take up an energy of $\|S^2\| = 4\pi$, which is exactly the amount that is lost. Apart from the bubbling (which in the classical setting is considered as a special case and which is kind of natural in the setting of currents) we get the same existence for all times.

There is however a small possible difference. While in the classical approach any bubble will completely vanish, bubbles in the Cartesian currents only “hide” in the vertical part. As a result, in our situation, they can possibly reappear in a reverse bubble (see [Top02, Top11]). In a way then, we can think of the bubbles as not vanishing but simply shrinking to a smaller scale.

A place to proceed to from here is of course the harmonic map heat flow for higher dimensions. In this situation, singularities can possibly occur more often and in different variants. An interesting case to study here would be maps from $\Omega \subset \mathbb{R}^3$ to $S^2$. Here we can have “open singularities”, like the hedgehog $\frac{x}{|x|}$, (see also [Lin87]) which will persist in time (in fact the hedgehog is a critical point of the Dirichlet energy) instead of only occurring at fixed times. Furthermore if we want to consider it as a Cartesian current, we have to note that the graph of the hedgehog has a boundary, which needs to be closed. Due to topology, the only possible way to achieve this, is to form a kind of vertical tube $[0, a] \times S^2$, which connects the singularity to a point $a$ on the boundary (or another singularity). Since this vertical part has a mass proportional to the length of such tubes, it contributes to the energy. This increase in energy also occurs when approximating maps with such singularities by smooth maps. The optimal approximation then tries to minimize the length
of all such tubes, in what is called “minimal connections” (See [BCL86]).\footnote{A phenomenon, which can incidentally be treated as an optimal transport problem, where point-masses generated by singularities need to either be moved towards the boundary, or in case of different orientations towards each other.} In a Cartesian currents-framework these minimal connections would then play an actual role in the evolution, as one way of reducing there energy of course is by moving connected singularities closer to each other.

The second possible application is unrelated to Cartesian currents. We already noted the mean curvature flow as a classic example of the minimizing movements-scheme. In fact this is even further related to what we are doing. The metric in this case (see [ATW, LS95]) is in principle given by the term

\[ d(\partial A, \partial B)^2 := \int_{A \Delta B} \text{dist}(x, \partial A) dx \]

This is surprisingly close to our definition of a generalized Wasserstein 2-distance. If one treats $\partial A$ and $\partial B$ as currents, then in order to move one into the other, one needs to traverse all of the symmetric difference $A \Delta B$. Furthermore if we need to move a distance of $r$, then for currents, the corresponding vectorfield will be of magnitude $r$. Thus the Wasserstein 2-distance will roughly scale with $\sqrt{r^2} = r$. On the other hand $\text{dist}(x, \partial A)$ will be $r$ on average and the domain will scale linearly with $r$ as well. Thus $d(\partial A, \partial B) \approx \sqrt{r \cdot r} = r$.

Hence we conjecture that the minimizing movements-scheme using the generalized Wasserstein distance $W_2$ as metric and the mass $\mathcal{M}$ as energy, will converge to the mean curvature flow.

The mean curvature flow has already been analyzed from the viewpoint of currents by Ilmanen [Ilm94], though using completely different methods. It would be interesting to see if his results could be recovered by this different approach. Furthermore any such result would easily extend to the case of volume preserving mean curvature flow (See [MSS16]), by considering only currents from the set \{ $T \in I_{n-1}(\Omega) \mid T = \partial [U], U \subset \Omega, |U| = M_0$, $\Omega \subset \mathbb{R}^n$, which is closed.\footnote{While mass in general is only lower-semicontinuous for currents, it can only be annihilated when surfaces with different orientation or multiplicity meet. All sets $[\Omega]$ however have the same orientation and multiplicity. Another way to see this is by noting that convergence in this set implies convergence of the characteristic functions $\chi_U$ in the sense of measures.}
Appendices
Appendix A

Relevant facts from differential geometry

Definition A.1 (Manifold). We say that a second countable Hausdorff-space $\mathcal{M}$ is a $n$-dimensional Manifold, if $\mathcal{M}$ is locally homeomorphic to $\mathbb{R}^n$. We will call each local homeomorphism $\phi : U \subset \mathbb{R}^n \rightarrow V \subset \mathcal{M}$ a coordinate chart and the set of those charts atlas. If $\phi_2^{-1} \circ \phi_1 \in C^1(U_1;U_2)$ (or $C^\infty(U_1;U_2), C^{0,1}(U_1;U_2)$ and so on) for all charts $\phi_1, \phi_2$ in the atlas, we call $\mathcal{M}$ a continuously differentiable (or smooth, Lipschitz and so on) Manifold. If the $\det D(\phi_2^{-1} \circ \phi_1) > 0$ for all coordinate charts, we call the atlas orientable. A manifold is called orientable, if it has an orientable atlas.

Note that any open subset of $\mathbb{R}^n$ is an $n$-dimensional manifold.

Definition A.2 (Submanifold). Let $\mathcal{M}$ be a $m$-dimensional Manifold and $\mathcal{N} \subset \mathcal{M}$ a $n$-dimensional manifold, both at least continuously differentiable. Then we call $\mathcal{N}$ a submanifold of $\mathcal{M}$ if one of the following equivalent conditions hold:

1. For each $x \in \mathcal{N}$ there exists a coordinate chart $\phi$ of $\mathcal{M}$ around $x$ with $\phi(0) = x$ such that $\phi^{-1}(V \cap \mathcal{N}) = U \cap \mathbb{R}^n$ and $\phi|_{U \cap \mathbb{R}^n}$ is a coordinate chart of $\mathcal{N}$.

2. For each $x \in \mathcal{N}$ there exists a neighborhood $U$ in $\mathcal{M}$ and a continuously differentiable map $u : U \rightarrow \mathbb{R}^{n-m}$ such that $u^{-1}(0) = \mathcal{N} \cap U$ and $Du(x)$ has full rank.

If $\mathcal{M}$ an $\mathcal{N}$ are only $C^{0,1}$, then condition (i) is still applicable and will work as a definition.

Remark A.3 (Manifold with boundary). There are different ways to treat manifolds with boundaries, which we will use somewhat interchangeably, depending on our needs.
1. The direct way: Let $\mathcal{M}$ be an $n$-dimensional manifold, $\partial \mathcal{M}$ a set such that $\partial \mathcal{M} \cap \mathcal{M} = \emptyset$, $\partial \mathcal{M} \cup \mathcal{M}$ is also second countable and Hausdorff and for each $x \in \partial \mathcal{M}$ there is a local homeomorphism $\phi : U \subset \mathbb{R}_{\geq 0} \times \mathbb{R}^{n-1} \to V \subset \partial \mathcal{M} \cup \mathcal{M}$ such that $\phi^{-1}(V \cap \partial \mathcal{M}) = \{0\} \times \mathbb{R}^{n-1}$.

Note that as an immediate result $\partial \mathcal{M}$ is an $n-1$-dimensional manifold.

2. The embedded way: Let $\mathcal{M}$ be an $n$-dimensional manifold and $\tilde{\mathcal{M}}$ be an $n$-dimensional manifold such that $\mathcal{M} \subset \tilde{\mathcal{M}}$ as a submanifold and $\mathcal{M}$ is locally compact. Then $\partial \mathcal{M} \subset \tilde{\mathcal{M}}$ is well defined. If furthermore $\partial \mathcal{M}$ also is an $n-1$-dimensional manifold, we call $\mathcal{M}$ a manifold with boundary.

Definition A.4 (Maps between manifolds). Let $\mathcal{M}$, $\mathcal{N}$ be manifolds, $U \subset \mathcal{M}$ open, $f : U \to \mathcal{N}$ continuous. We say that $f$ is of class $C^k$ if $\mathcal{M}$, $\mathcal{N}$ are of class $C^k$ and for each $x \in U$ there are coordinate charts $\phi$ of $\mathcal{M}$ around $x$ and $\psi$ of $\mathcal{N}$ around $f(x)$ such that $\psi^{-1} \circ f \circ \phi$ is of class $C^k$.

Definition A.5 (Tangential space). Let $\mathcal{M}$ be an $n$-dimensional $C^1$-manifold, $x \in \mathcal{M}$. Then the tangential space at $x$ can be defined by the following condition: Let $C := \{\gamma \in C^1([-\epsilon, \epsilon]; \mathcal{M}) \, | \, \gamma(0) = x\}$. Then

$$T_x\mathcal{M} := C/\sim$$

where $\gamma_1 \sim \gamma_2$ if $\frac{\partial}{\partial t}\phi^{-1} \circ \gamma_1(0) = \frac{\partial}{\partial t}\phi^{-1} \circ \gamma_2(0)$ for any coordinate chart around $x$. In fact if this condition holds for any coordinate chart, it immediately holds for all of them.

The tangential space of $\mathcal{M}$ then is given as the disjoint union $\bigcup_{x \in \mathcal{M}} \{x\} \times T_x\mathcal{M}$. If $\mathcal{M}$ is only Lipschitz, we can find that $T_x\mathcal{M}$ still is well defined for almost all $x \in \mathcal{M}$, however here the notion of approximate tangential space is more appropriate (For the details on that we refer to [Mor08, Fed69]).

A function $v : \mathcal{M} \to T\mathcal{M}$ such that $v(x) \in T_x\mathcal{M}$ is called a tangential vector field.

Definition A.6 (Pushforward of vectors). Let $\mathcal{M}, \mathcal{N}$ be continuously differentiable manifold, $U \subset \mathcal{M}$ and $f : U \to \mathcal{N}$ continuously differentiable. Then for each $x \in U$ we define the pushforward

$$f_\# : T_x\mathcal{M} \to T_{f(x)}\mathcal{N}; v \mapsto \frac{d}{dt}(f \circ \alpha)(0)$$

where $\alpha \in C^1([-\epsilon, \epsilon]; \mathcal{M})$, $\alpha(0) = x$, $\frac{d}{dt}\alpha(0) = v$.

If $f$ is injective and $v$ is a tangential vector field, then the pushforward is similarly defined and $f_\#v$ can be considered a tangential vector field on $f(U)$. 
Definition A.7 (Riemannian manifold). A continuously differentiable manifold \( \mathcal{M} \) is called a Riemannian manifold, if it has an additional Riemannian metric, that is for every point \( x \in \mathcal{M} \) there is a symmetric positive definite bi-linear form

\[
g(x) : T_{x} \mathcal{M} \times T_{x} \mathcal{M} \to \mathbb{R}; (u, v) \to g(x)(u, v)
\]

depending smoothly on \( x \) when expressed in coordinates. Note that a Riemannian metric should not be confused with a metric in the sense of a distance function.

For subsets of \( \mathbb{R}^{n} \) there is a natural Riemannian metric given by the scalar product

\[
g(x)(u, v) = \langle u, v \rangle_{\mathbb{R}^{n}}.
\]

Lemma A.8. Let \( \mathcal{M} \subset \mathcal{N} \) be a submanifold and \( x \in \mathcal{M} \). Then

1. There is a canonical embedding \( T_{x} \mathcal{M} \subset T_{x} \mathcal{N} \).
2. If \( g_{\mathcal{N}} \) is a Riemannian metric on \( \mathcal{N} \), then the restriction

\[
g_{\mathcal{M}}(x) := g_{\mathcal{N}}(x)|_{T_{x} \mathcal{M} \times T_{x} \mathcal{M}}
\]

is a Riemannian metric on \( \mathcal{M} \).

Definition A.9 (Isometry). Let \( \mathcal{M}, \mathcal{N} \) be Riemannian manifolds, \( U \subset \mathcal{M} \) open and \( f : U \to \mathcal{N} \) continuously differentiable. Then \( f \) is called an isometry if for all \( x \in U, u, v \in T_{x} \mathcal{M} \) we have

\[
g_{\mathcal{M}}(x)(u, v) = g_{\mathcal{N}}(x)(f_{\#}u, f_{\#}v).
\]

Theorem A.10 (Nash-embedding theorem). Let \( \mathcal{M} \) be an \( n \)-dimensional Riemannian manifold. Then there exists an \( N \in \mathbb{N} \) only depending on \( n \) such that there is an isometric embedding (an injective isometry)

\[
f : \mathcal{M} \to \mathbb{R}^{N}.
\]

This allows us to always consider manifolds embedded in \( \mathbb{R}^{N} \). However there the viewpoint of intrinsic geometry still has many uses, so we will try to stick to it for as long as possible.

Definition A.11 (k-vectors, exterior product). Let \( V \) be an \( n \)-dimensional vector space. Then the set of (alternating) \( k \)-vectors is given by

\[
\wedge^{k} V := \left\{ \sum_{\sigma \in \pi_{k}} \text{sgn}(\sigma) v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)} \bigg| v_{1}, \ldots, v_{k} \in V \right\} \subset \underbrace{V \otimes \ldots \otimes V}_{k\text{-times}}
\]
where $\pi_k$ is the set of permutations of the set $\{1,\ldots,k\}$. Per convention we identify $\wedge^0 V := \mathbb{R}$. We define the exterior product of vectors by

$$\wedge : V \times V \to \wedge^2 V; \ (v, w) \mapsto v \wedge w := v \otimes w - w \otimes v.$$ 

This can be extended to

$$\wedge : \wedge^k V \times \wedge^l V \to \wedge^{k+l} V; \ \ (\xi, \eta) \mapsto \xi \wedge \eta := \frac{1}{k!l!} \sum_{\sigma \in \pi_n} \text{sgn}(\sigma) \xi_{\sigma(1)} \ldots \sigma(k) \otimes \eta_{\sigma(k+1)} \ldots \sigma(k+l).$$

Finally, if $e_1, \ldots, e_n$ forms a basis of $V$, then a basis of $\wedge^k V$ is given by $(e_{i_1} \wedge \ldots \wedge e_{i_k})_{i_1 < \ldots < i_k}$.

**Remark A.12.**

1. If we apply the previous definition to a dual space $V^*$ instead of $V$, then the $\wedge^k V^*$ can be understood as the set of alternating $k$-linear forms (often just called $k$-forms). The space $\wedge^k V^*$ can be identified as the dual space of $\wedge^k V$.

2. If $\xi \in \wedge^k V$ can be written as $\xi = v_1 \wedge \ldots \wedge v_k$ for vectors $v_1, \ldots, v_k \in V$, then it is called simple. This is not true in general as the example $e_1 \wedge e_2 + e_3 \wedge e_4 \in \wedge^2 \mathbb{R}^4$ shows.

3. Each simple non-zero $k$-vector $\xi = v_1 \wedge \ldots \wedge v_k$ generates a $k$-dimensional linear subspace $W_\xi$ of $V$ spanned by $v_1, \ldots, v_k$, which is independent of the decomposition $v_1, \ldots, v_k$ of $\xi$. In turn, given a $k$-dimensional subspace $W$ spanned by $v_1, \ldots, v_k$, we can define a corresponding $k$-vector by $\xi = v_1 \wedge \ldots \wedge v_k$. This $k$-vector then generates $W$ and is unique up to a factor. To be more precise if the change of basis of $W$ between $v_1, \ldots, v_k$ and $w_1, \ldots, w_k$ is given by the matrix $A \in \mathbb{R}^{k \times k}$, then

$$w_1 \wedge \ldots \wedge w_k = \det A v_1 \wedge \ldots \wedge v_k.$$ 

Thus if $W$ is oriented and possesses a scalar product, we uniquely identify $W$ with the $k$-vector $\xi_W = e_1 \wedge \ldots \wedge e_k$, where $e_1, \ldots, e_k$ forms an oriented orthonormal basis of $W$. We will often make use of this when dealing with tangential spaces of embedded manifolds.

**Definition A.13 (k-form).** Let $\mathcal{M}$ be a smooth Manifold. A differentiable $k$-form on $\mathcal{M}$ is a smooth function

$$\omega : \mathcal{M} \to \wedge^k T^* \mathcal{M}$$

by which we mean $\omega(x) \in \wedge^k (T_x \mathcal{M})^*$. We will then write $\omega \in C^\infty (\mathcal{M}; \wedge^k)$. The spaces $\omega \in C^l (\mathcal{M}; \wedge^k)$ are similarly defined. From now on we will use “$k$-form” as a shorthand for differentiable $k$-form.
Assume that there is a local coordinate chart \( \psi : U \to \mathcal{M} \). Then the standard basis \( e_i \) on \( \mathbb{R}^n \) induces a basis \( \partial_{x_i} := \psi^* e_i \) on \( T_{\psi(x)} \mathcal{M} \) for all \( x \in U \). We then denote the corresponding dual basis of \( T_{\psi(x)}^* \mathcal{M} \) by \( dx^i \).

**Definition A.14 (exterior derivative).** Let \( \mathcal{M} \) be a differentiable Manifold and \( f \in C^1(\mathcal{M}) \). We define the exterior derivative of \( f \) by

\[
    df : v \mapsto \frac{\partial f}{\partial v}.
\]

Now assume a local coordinate system on \( \mathcal{M} \). If \( \omega = \sum_{1 \leq i_1 < \ldots < i_k \leq n} f_I dx^I \) locally, then we define

\[
    d\omega = \sum_{1 \leq i_1 < \ldots < i_k \leq n} (df_I) dx^I.
\]

This definition is independent of the choice of coordinates.

**Lemma A.15.** The following calculation rules are easily derived:

1. \( d^2 = 0 \)
2. \( d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta \).

**Definition A.16 (pullback).** Let \( \mathcal{M}, \mathcal{N} \) be differentiable manifolds, \( f : \mathcal{M} \to \mathcal{N} \) continuously differentiable. Then we define

\[
    f^\# \omega(x) : T_{f(x)} \mathcal{N}^k \to \mathbb{R}; (v_1, \ldots, v_k) \mapsto \omega(f(x)) \langle f^\# v_1, \ldots, f^\# v_k \rangle
\]

for any \( \omega \in C^\infty(\mathcal{N}; \wedge^k) \).

**Remark A.17.** Note that exterior derivative and pushforward commute. In fact up to a constant factor, it is possible to characterize the exterior derivative as the only operator from \( k \) to \( k-1 \)-forms that commutes with pushforwards. (See [Pal59])

**Definition A.18 (Contraction).** Let \( \mathcal{M} \) be a manifold, \( \omega \in C^\infty(\mathcal{M}; \wedge^k) \) and \( v \) a tangential vector field on \( \mathcal{M} \). Then we define

\[
    \iota_v \omega(x) : (w_1, \ldots, w_{k-1}) \mapsto \omega(x) \langle v(x), w_1, \ldots, w_{k-1} \rangle.
\]

**Definition A.19 (Lie-derivative).** Let \( \mathcal{M} \) be a differentiable manifold, \( v \) a \( C^1 \) tangential vector field on \( \mathcal{M} \). Then the solutions to the ODE

\[
    \frac{\partial \phi_t(x)}{\partial t} = v(\phi_t(x)), \quad \phi_0(x) = x
\]

at least locally and for small \( t \) induce a \( C^1 \) map \( \phi_t : \mathcal{M} \to \mathcal{M} \).

Now let \( \omega \in C^\infty(\mathcal{M}; \wedge^k) \). Then the Lie-derivative of \( \omega \) in direction \( v \) is given by

\[
    \mathcal{L}_v \omega := \lim_{t \to 0} \frac{\phi_t^\# \omega - \omega}{t}.
\]
Lemma A.20 (Cartan’s magic formula). Let \( v \) be a continuously differentiable tangential vector field and \( \omega \in C^\infty(\Omega; \Lambda^k) \). Then

\[
\mathcal{L}_v \omega = di_v \omega + i_v d\omega.
\]

Definition A.21 (Integration along manifolds). Let \( U \subset \mathbb{R}^n \) be an open and bounded set, \( \omega \in C^\infty(U; \Lambda^n) \). Then \( \omega(x) = f(x)dx^1 \wedge ... \wedge dx^n \) and we define

\[
\int_U \omega = \int_U f \, dx.
\]

Let \( M \) be an orientable \( n \)-dimensional manifold, \( \omega \in C^\infty(M; \Lambda^n) \). Let \( \{(\phi_i)_{i \in I}\} \) be a locally finite atlas with \( \phi_i : U_i \mapsto M \) and \( \{(\psi_i)_{i \in I}\} \subset C^1(M; \mathbb{R}) \) such that \( \text{supp} \psi \subset \phi_i(U_i) \) and \( \sum_{i \in I} \psi_i = 1 \). Then we define

\[
\int_M \omega := \sum_{i \in I} \int_{U_i} \phi_i^\#(\psi_i \omega).
\]

Theorem A.22 (Stokes theorem). Let \( M \) be a \( n \)-dimensional manifold with boundary, \( \omega \in C^\infty(M; \Lambda^{n-1}) \). Then

\[
\int_{\partial M} \omega = \int_M d\omega.
\]

Definition A.23. Let \( M \) be a Riemannian manifold and \( I \subset \mathbb{R} \) be an interval. For any continuous, piecewise differentiable curve \( \gamma : I \to M \), we define its length by

\[
L(\gamma) := \int_I |\dot{\gamma}(t)| \, dt.
\]

It is easy to see that \( L(\gamma) \) is invariant under reparametrization. The geodesic distance between \( x, y \in M \) is given by

\[
d(x, y) := \inf \{ L(\gamma) \mid \gamma \in C^1([0,1]; M), \gamma(0) = x, \gamma(1) = y \}.
\]

A curve \( \gamma \) is called a (minimal) geodesic connecting \( x \) and \( y \) if it is a minimizer of this functional and \( |\dot{\gamma}| \) is independent of \( t \).

Proposition A.24. Let \( M \) be a compact manifold. Then for any two points \( x, y \in M \) there is a geodesic connecting \( x \) and \( y \).
Appendix B

Relevant facts from geometric measure theory

This section only serves the purpose of repeating a few rather well known facts and definitions that are needed later on. Any reader already familiar with Geometric Measure Theory should be able to skip this section and may only need to use it as reference for notation. The standard references for Geometric measure theory are [Fed69, Sim83] and [Mor08], although it can be enlightening to look into the original paper [FF60] for some of the theorems and into the original book about currents by de Rham [dR55] (or the translation [dR84]).

In the following, the space of currents in $\Omega$ will be given by the dual space of smooth differential forms with compact support, that is

$$D_k(\Omega) := \left(C_0^\infty(\Omega; \wedge^k)\right)^*.$$

If $M$ is an oriented $C^1$ manifold, we then denote the corresponding current by

$$[M] : \omega \mapsto \int_M \omega.$$

**Theorem B.1** (Area-Formula [Fed69, 3.2.3],[Sim83, 8.2],[Mor08, 3.7]). Let $f : \mathbb{R}^m \to \mathbb{R}^n$ Lipschitz, $m \leq n$.

1. If $A$ is $\mathcal{L}^m$-measurable, then

$$\int_A J_m f(x) d\mathcal{L}^m x = \int_{\mathbb{R}^n} N(f | A, y) d\mathcal{H}^m(y).$$

2. If $u$ is $\mathcal{L}^m$-integrable, then

$$\int_A u(x) J_m f(x) d\mathcal{L}^m x = \int_{\mathbb{R}^n} \sum_{x \in f^{-1}(y)} u(x) d\mathcal{H}^m(y).$$
where \( J_m f(x) = \sqrt{\det(Df(x)Df(x)^T)} \) is the Jacobian determinant.

**Definition B.2** ([Mor08, p.39], [Sim83, 26.3], [Fed69, 4.1.7]). Let \( T \in D_k(\Omega) \) be a current. Then its boundary \( \partial T \in I_{k-1}(\Omega) \) is defined by the relation

\[
\partial T(\omega) := T(d\omega).
\]

**Definition B.3** ([Mor08, p.41], [Sim83, 26.4], [FF60, 2.4], [Fed69, 4.1.7]). On each \( k \)-form \( \omega \in C^\infty(\Omega; \wedge^k) \) we can define a local norm (the co-mass) by

\[
\|\omega(x)\| = \sup\{\langle \omega(x), \xi \rangle | \xi \text{ is a simple } k\text{-vector and } |\xi| \leq 1\}
\]

where the simple \( k \)-vectors are those of the form \( \xi = ce_1 \wedge ... \wedge e_k \), with \( (e_i)_{i \in \{1, ..., k\}} \) orthonormal and \( |\xi| = c \). The mass of a current \( T \in D_k(\Omega) \) then is given by

\[
\mathcal{M}(T) := \sup\{T(\omega) | \omega \in C^\infty(\Omega; \wedge^k), \|\omega(x)\| \leq 1\}.
\]

A current \( T \) is called a normal current if \( \mathcal{M}(T) < \infty \) and \( \mathcal{M}(\partial T) < \infty \).

**Lemma B.4.** The mass is a lower semi-continuous function. More precisely, for any sequence \((T_i)_{i \in \mathbb{N}} \subset D_k(\Omega) \) with \( T_i \to T \in D_k(\Omega) \) we have

\[
\liminf_{i \to \infty} \mathcal{M}(T_i) \geq \mathcal{M}(T).
\]

**Definition B.5** (Currents representable by integration [Fed69, 4.1.5], [Mor08, 4.3B]). Let \( T \in D_k(\Omega) \). We call \( T \) representable by integration if there exists a locally finite Borel regular measure \( \mu \) and a \( \mu \)-integrable function \( \xi(x) \) of simple unit \( k \)-vectors such that

\[
T(\omega) = \int_\Omega \langle \omega(x), \xi(x) \rangle \, d\mu \quad \forall \omega \in C^\infty(\Omega; \wedge^k).
\]

We will then write \( \|T\| := \mu \) and \( \vec{T} := \xi \).

**Remark B.6.** Note that if \( T \) is representable by integration, we are no longer restricted to testing \( T \) with smooth differential forms. Instead we can use forms with arbitrary \( \mu \)-integrable coefficients. Note that independent of \( \mu \), this includes all bounded Borel-measurable functions. We can then define the restriction to a subset \( U \subset \Omega \) by

\[
T\mathring{\downarrow} U(\omega) := T(X_U \cdot \omega).
\]

**Definition B.7** ([FF60, 2.5],[Fed69, 4.1.14]). Let \( f : U \to \Omega \) be a smooth map, \( T \) a \( k \)-current. Then the pushforward is defined by

\[
f_#T(\omega) = T(f^#\omega) \quad \text{for all } \omega \in C^\infty(\Omega; \wedge^k).
\]
Furthermore if \( f : U \to \Omega \) is a Lipschitz map, and \( T \) a normal current, the pushforward is defined by
\[
T_f = \lim_{i \to \infty} (f_i)_\#T
\]
where the \( f_i \) are smooth and \( f_i \to f \) uniformly with uniformly bounded derivatives.

**Lemma B.8** (Homotopy Lemma [FF60, 2.7], [Sim83, 26.22]). Let \( T \) be a \( k \)-current, \( f, g : U \to \Omega \) and \( h : \mathbb{R} \times U \to \Omega \) smooth such that \( h(0,.) = f \) and \( h(1,.) = g \). Then
\[
\partial h_\#([0,1] \times T) + h_\#([0,1] \times \partial T) = g_\#(T) - f_\#(T).
\]

**Definition B.9** (polyhedral chains [Fed69, 4.1.22], [Mor08, p.41], [Sim83, ]). We call \( T \in \mathcal{D}_k(\Omega) \) a polyhedral chain if there is a finite number of oriented \( k \)-polyhedrons \( (P_i)_{i \in \{1,\ldots,l\}} \) such that
\[
T := \sum_{i=1}^l [P_i].
\]
We call \( T \in \mathcal{D}_k(\Omega) \) a Lipschitz chain if additionally there are Lipschitz maps \( (P_i)_{i \in \{1,\ldots,l\}} \) such that
\[
T := \sum_{i=1}^l f_i_\# [P_i].
\]
If the \( f_i \) are continuously differentiable, we will call \( T \) a \( C^1 \)-chain, if they are smooth, we will call \( T \) a polyhedral \( C^\infty \)-chain, and so on.

**Definition B.10** (Integer rectifiable currents (integral currents) [FF60, 3.7], [Fed69, 4.1.24], [Mor08, p.41], [Sim83, 27.1]). A current \( T \in \mathcal{D}_k(\Omega) \) is called integer rectifiable or integral current if one of the following (equivalent) conditions holds

1. There are countably many oriented \( C^1 \)-sub-manifolds \( (M_i)_{i \in \mathbb{N}} \) and integer weights \( (\alpha_i)_{i \in \mathbb{N}} \) such that
\[
T = \sum_{i \in \mathbb{N}} \alpha_i [M_i]
\]
and \( \mathcal{M}(T) < \infty \), \( \mathcal{M}(\partial T) < \infty \).
2. There are \( k \)-rectifiable sets \( (M_i)_{i \in \mathbb{N}} \) and integer weights \( (\alpha_i)_{i \in \mathbb{N}} \) such that
\[
T(\omega) = \sum_{i \in \mathbb{N}} \int_{M_i} \alpha_i \langle \omega(x), \xi(x) \rangle \, d\mathcal{H}^k(x)
\]
where \( \xi(x) \) is the approximate tangential unit \( k \)-vector of \( M_i \) at \( x \). This is to say \( T \) is representable by integration with measure \( \sum_{i \in \mathbb{N}} \alpha_i \mathcal{H}^k \upharpoonright M_i \).
3. There is a sequence of polyhedral Lipschitz chains \((T_i)_{i \in \mathbb{N}}\) such that
\[
\lim_{i \to \infty} \mathcal{M}(T - T_i) = 0.
\]

4. There is a sequence of polyhedral \(C^1\) chains \((T_i)_{i \in \mathbb{N}}\) such that
\[
\lim_{i \to \infty} \mathcal{M}(T - T_i) = 0.
\]

We then write \(T \in I_k(\Omega)\).

Remark B.11. The definition used by Federer is (3), while most other authors use (2). The conditions (1) and (4) are restatements of (2) and (3) respectively, using the fact that rectifiable sets can be approximated in \(\mathcal{H}^k\) by \(C^1\)-diffeomorphisms (see also [Mor08, 3.11], [Fed69, 3.2.18]).

Lemma B.12 ([GMS98, p.149]). Let \(\Omega \subset \mathbb{R}^n\), \(f : \overline{\Omega} \to \mathbb{R}^m\) be a Lipschitz map and \(T \in I_k(\Omega)\). Then \(f_\#T \in I_k(\mathbb{R}^n)\).

Theorem B.13 (Constancy theorem [Fed69, 4.1.31],[Mor08, 4.9]). Let \(\mathcal{M} \subset \Omega\) be a \(k\)-dimensional connected oriented \(C^1\) submanifold with boundary and \(T \in I_k(\Omega)\) such that \(\mathcal{M}(T) < \infty\), \(\mathcal{M}(\partial T) < \infty\), \(\text{supp} \ T \subset \overline{\mathcal{M}}\) and \(\text{supp} \partial T \subset \partial \mathcal{M}\). Then \(T = k [\Omega]\) for some \(k \in \mathbb{R}\). Furthermore if \(\partial T\) is integer rectifiable then \(T\) is as well.

Definition B.14 (Flat metric [Fed69, 4.1.24],[Mor08, 4.3],[Sim83, §31]). Let \(T \in I_k(\Omega)\). Then the flat metric of \(T\) is given by
\[
F(T) = \sup \{ \mathcal{M}(S) + \mathcal{M}(R) \mid T = \partial S + R, S \in I_{k+1}(\Omega), R \in I_k(\Omega) \}.
\]

Lemma B.15 ([Sim83, 31.2]). Let \((T_i)_{i \in \mathbb{N}} \subset I_k(\Omega), T \in I_k(\Omega)\) such that \(\mathcal{M}(T_i)\) and \(\mathcal{M}(\partial T_i)\) are uniformly bounded. Then
\[
T_i \to T \quad \text{if and only if} \quad F(T_i - T) \to 0
\]

Lemma B.16 (Slicing Lemma [Fed69, 4.2.15],[Whi89]). Let \(T \in D_k(\Omega)\) be a normal \(k\)-current, \(\partial T = 0\). Then \(T\) is integer rectifiable if and only if for any \(a \in \Omega\) the current
\[
\partial [T \res B_r(a)]
\]
is integer rectifiable for almost all \(r > 0\).

Theorem B.17 (Compactness closure [FF60, 8.12, 8.13]). Let \((T_i)_{i \in \mathbb{N}} \subset I_k(\Omega)\) such that \(\mathcal{M}(T_i)\) and \(\mathcal{M}(\partial T_i)\) are uniformly bounded. Then there exists \(T \in I_k(\Omega)\) and a sub-sequence \((T_i)_{i \in \mathbb{N}}\) (not relabeled) such that
\[
T_i \to T.
\]

In other words, for all \(C > 0\) the set \(\{ T \in I_k(\Omega) \mid \mathcal{M}(T) \leq C, \mathcal{M}(\partial T) \leq C \}\) is sequentially compact.
Theorem B.18 (Isoperimetric inequality [Fed69, 4.2.10], [FF60, Cor. 6.3], [Mor08, 5.3]). Let $T \in I_k(\mathbb{R}^n)$ with $\partial T = 0$. Then there exists $S \in I_{k+1}(\mathbb{R}^n)$ such that $\partial S = T$ and

$$\mathcal{M}(S)^{k/(k+1)} \leq C \mathcal{M}(T)$$

where $C = 2n^{2k+2}$.

Theorem B.19 (Deformation theorem [Fed69, 4.2.9], [FF60, Thm. 5.5]). Let $T \in D_k(\mathbb{R}^n)$ be a normal current and $\varepsilon > 0$. Then there exist $P, Q \in D_k(\mathbb{R}^n)$, $S \in D_{k+1}(\mathbb{R}^n)$ such that for $C = 2n^{2k+2}$ the following holds:

1. $T = P + Q + \partial S$
2. If $k \neq 0$ then
   $$\mathcal{M}(P)/\varepsilon^k \leq C \left( \mathcal{M}(T)/\varepsilon^k + \mathcal{M}(\partial T)/\varepsilon^{k-1} \right)$$
   $$\mathcal{M}(\partial P)/\varepsilon^{k-1} \leq C \mathcal{M}(\partial T)/\varepsilon^{k-1}$$
   $$\mathcal{M}(Q)/\varepsilon^k \leq C \mathcal{M}(\partial T)/\varepsilon^{k-1}$$
   $$\mathcal{M}(S)/\varepsilon^{k+1} \leq C \mathcal{M}(T)/\varepsilon^k$$
3. If $k = 0$ then $\mathcal{M}(P) \leq \mathcal{M}(T)$, $Q = 0$, $\mathcal{M}(S)/\varepsilon \leq \mathcal{M}(T)$
4. $\text{supp } P \cup \text{supp } S \subset \overline{B_{2n\varepsilon}}(\text{supp } T)$
   $\text{supp } \partial P \cup \text{supp } Q \subset \overline{B_{2n\varepsilon}}(\text{supp } \partial T)$, if $k \neq 0$.
5. $\text{supp } P \subset W_k$ and if $k \neq 0$ also $\text{supp } \partial P \subset W_{k-1}$, where $W_k$ is a regular $k$-skeleton of size $\varepsilon$. In other words, $P$ is a polyhedral chain.
6. If $T \in I_k(\mathbb{R}^n)$, then $P, Q \in I_k(\mathbb{R}^n)$, $S \in I_{k+1}(\mathbb{R}^n)$.
7. If $\partial T \in I_{k-1}(\mathbb{R}^n)$ then $Q \in I_k(\mathbb{R}^n)$.
8. If $\partial T$ is a polyhedral chain, then so is $Q$.
9. If $T$ is a polyhedral chain, then so is $S$.
10. There is a sequence of homotopies $h_i$ such that $Q = \sum_{i=1}^{n-k} Q_i$ where we have $Q_1 = h_1((0,1] \times \partial T)$ and $Q_i = h_i(\#([0,1] \times (T + \sum_{j=1}^{i-1} \partial Q_j))$.
    Furthermore $H^k(\text{supp } Q_i \cap \text{supp } Q_j) = 0$ for all $i \neq j$.

Remark B.20. We can always subdivide $\mathbb{R}^n$ into a regular grid of hypercubes with side-length $\varepsilon$. Then the $k$-dimensional boundaries of those cubes form a regular $k$-skeleton $W_k$ of size $\varepsilon$ as mentioned in the deformation theorem. One can also write

$$W_k := \{ x \in \mathbb{R}^n \mid x_i \in Q \varepsilon \mathbb{Z}^n + a \text{ for } n-k \text{ indices } i \in \{1, \ldots, n\} \}$$

where $Q$ is a fixed orthogonal matrix and $a \in \mathbb{R}^n$ a fixed point. It is easy to see, that $W_k$ is unique up to rotations and translations. Since the proof of the deformation theorem is constructive, we are actually free to choose $W_k$, although the canonical choice usually is $Q = I$, $a = 0$. ■
Main index

Arc length reparametrization, 65
Area-formula, 140
Atlas, 134
boundary, 141
Bubbling, 26, 55
Cartan’s formula, 16, 139
Cartesian currents, 32
chain, 142
Lipschitz, 142
polyhedral, 142
co-mass, 141
Coercive
on currents, 115
coercive
Cartesian Current, 44
Compactness-closure theorem, 143
boundary-less, 15
Compatible
metric, 118
Constancy theorem, 143
Contraction, 138
Convergence Theorem
Cartesian version, 124
Convergence theorem
Basic version, 120
Coordinate chart, 134
current
chain, 142
integer rectifiable, 142
Currents
Graph, 30
currents
continuous curve, 22
representable by integration, 141
Curve of maximal slope, 113
existence, 114
Deformation theorem
boundary free, 14
on manifolds, 14
defformation theorem, 144
Degenerate distance
generalized Wasserstein, 50
vertical Wasserstein, 55
Differential form, 137
dualities
homogeneous flat norm, 15
Wasserstein, 78
Exterior derivative, 138
Exterior product, 136
Flat metric, 143
homogeneous, 9
flat norm, 15
homogeneous, 15
Gradient flow
curve of maximal slope, 113
Hilbert space, 110
Homogeneous flat metric, 9
homogeneous flat norm, 15
homotopy lemma, 142
curve generation, 72
integral current, 142
Isometry, 136
isoperimetric inequality, 144

$k$-skeleton, 144
k-vector, 136

Lie derivative, 16, 138

Manifold, 134
  Riemannian, 136
  submanifold, 134
  with boundary, 134
mass, 141

Metric
  compatible, 118
  Riemannian, 136

Metric derivative, 112

Minimizing movements, 113

Parametrization
  by arc length, 65

Projection
  of trace, 64

Pullback, 138

Pushforward
  Vector, 135
  pushforward
    of currents, 141

Reparametrization
  effect on trace, 62
  Riemannian metric, 136
Simple forms, 137

slicing lemma, 143

Solution
  space-time, 117
Strong upper gradient, 112

Tangential space, 135

Theorem
  compactness closure, 143
  boundary-less, 15
  Convergence, 120, 124
  deformation
    boundary free, 14
  Nash-embedding, 136
  Stokes, 139

theorem
  deformation, 144
  equivalence, 72
  rectifiability of traces, 67

trace
  boundary formula, 63
  curve of currents, 58
  rectifiability theorem, 67
  space-time, 58

Upper gradient, 112

Wasserstein distance
  generalized, 48
  finiteness, 78
  vertical, 52
  with boundary, 106
Weak upper gradient, 112
Symbol index

$C^0(I; I_k(\Omega))$, 22
$C^1_{pw}(\text{supp} T; \mathbb{R}^n)$, 17
$\text{cart}^2(\Omega; S^2)$, 37
$\text{cart}(\Omega; M)$, 32
$\text{cart}^p(\Omega; \mathbb{R}^N)$, 36

d, 138
$\partial T$, 141

$F(T)$, 143
$f_\#T$, 141
$f_\#v$, 135
$f_\#\omega$, 138
$F(T)$, 15
$\tilde{F}(T)$, 9
$\tilde{F}(T)$, 15

g_E, 112
$\mathcal{G}(u)$, 30

$(id, S)_\#(I)$, 58
$I_k$, 142

$i_v\omega$, 138

$L_{p,\partial}(S)$, 106
$L_p(S)$, 48
$L_{p,\text{vert}}(S)$, 52
$\mathcal{L}_vT$, 16
$\mathcal{L}_e\omega$, 138

$\|T\|_{M_p}$, 36
$\|\omega\|_{M_T}$, 36
$M(T)$, 141

$\|\omega(x)\|$, 141

$S_\#(I)$, 58
$T_2 M$, 135

$|\dot{u}|(t)$, 112

$W_{p,\partial}(T_0, T_1)$, 106
$W_p(T_0, T_1)$, 48
$W_{p,\text{vert}}(T_0, T_1)$, 52
Bibliography


