2015

Geometric gradient flow in the space of smooth embeddings

2015

Geometric gradient flow in the space of smooth embeddings

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

GEOMETRIC GRADIENT FLOW IN SPACE OF SMOOTH EMBEDDINGS

by

DARA GOLD

B.S. Mathematics, Georgetown University, 2010

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2015

Approved by

First Reader ____
Steve Rosenberg, PhD
Professor of Mathematics

Second Reader ____
Gene Wayne, PhD
Professor of Mathematics

Third Reader ____
David Fried, PhD
Professor of Mathematics

Acknowledgments

Thank you very much to Steve Rosenberg for his patience and guidance over the past two years. I have been extremely appreciative of the time he has spent with me in developing this thesis and for his feedback in guiding my work.

A special thank you also to Gene Wayne for his input and suggestions in applying the Cauchy-Kovalevskaya Theorem in the fifth chapter.

GEOMETRIC GRADIENT FLOW IN SPACE OF SMOOTH EMBEDDINGS

(Order No.)

DARA GOLD

Boston University, Graduate School of Arts and Sciences, 2015

Major Professor: Steve Rosenberg, PhD, Professor of Mathematics

ABSTRACT

Given an embedding $\phi : M \to \mathbb{R}^N$ of a closed, compact manifold into N-dimensional Euclidean space, we aim to perform gradient flow of a penalty function $P : \mathrm{Emb}(M, \mathbb{R}^N) \to \mathbb{R}$ in the space $\mathrm{Emb}(M, \mathbb{R}^N)$ to find an ideal manifold embedding. We study the computation of the gradient for a penalty function that contains both a curvature and distance term. We also find a lower bound for how long an embedding $\phi(M)$ will remain in the space of embeddings when moving in a fixed, normal gradient direction. Finally, we study the distance penalty function in a special case in which we can prove short time existence of the flow using the Cauchy-Kovalevskaya Theorem.

Contents

Chapter 1

Introduction

A standard problem in data analysis and machine learning is determining how to approximate a finite fixed set of points in Euclidean space $\mathbb{R}^N$ with a k-dimensional manifold. One approach is to fit the set with a manifold that passes through every point, but as is more commonly seen with points in the plane, this approach has problems of "over-fitting." Additionally, such a manifold will have high curvature. Another approach is to use a manifold with no curvature (i.e. a plane in $\mathbb{R}^3$) which intuitively will increase the distance from the manifold to the set of points. This prompts the question of what is meant by an ideal manifold embedding.

To be more precise, given a fixed set of points $S = \{x_i\}_{i \in I} \in \mathbb{R}^N$, we want to approximate S with an embedded k-dimensional manifold. The manifold will be denoted M, and its embedded image will be $\phi(M) \in \mathbb{R}^N$ for some smooth embedding $\phi \in \mathrm{Emb}(M, \mathbb{R}^N)$. A best approximation to S, or an ideal manifold embedding, will minimize a combination of the curvature of $\phi(M)$ and its distance from the set S. Therefore evaluating the quality of a given embedding ϕ requires a penalty function P, which will be a function of ϕ and will incorporate both a curvature and a distance term. A ϕ, if it exists, that minimizes this function will be our optimal embedding. We examine the penalty function given by

$$P(\phi) = \int_M |R(\phi(m))|^2 \mathrm{dvol}_M + \sum_i \int_M d^2(\phi(m), x_i) \mathrm{dvol}_M$$

where $R(\phi(m))$ is the Riemann curvature tensor evaluated at $\phi(m) \in \phi(M)$ and $d(\phi(m), x_i)$ is the Riemannian distance from $\phi(m)$ to the closest point $x_i \in S$ (m is a point in the initial manifold M). Here the curvature and distance terms are weighted equally but coefficients can be added to emphasize one feature over the other. It should be noted that for points in $\phi(M)$ with more than one closest point in S, the gradient of the penalty function can have problems of smoothness, but we are assuming for now that no such points exist. In order to minimize this function we must compute the negative gradient (vector) $Z_\phi = -\nabla P_\phi$ of the penalty function at a point ϕ, and flow in this direction from ϕ towards a critical point (a critical embedding in the space of embeddings). Although negative gradient flow guarantees movement towards a critical point, we leave it as a future direction to determine if the critical point is in fact a local/ global minimum or another type of critical point.

Chapters 2 and 3

Chapters 2 and 3 investigate the curvature penalty term P_c only (the first term in $P(\phi)$). The gradient of the distance penalty term is considered in detail in Chapter 5. To compute the gradient of the curvature term $Z_c = \nabla P_c$ at an embedding $\phi_0 \in \mathrm{Emb}(M, \mathbb{R}^N)$, we fix a direction vector $\vec{X} \in T_{\phi_0}\mathrm{Emb}(M, \mathbb{R}^N)$ (note that $\vec{X}$ is a vector field along $\phi(M) \in \mathbb{R}^N$). We know that the gradient is governed by the equation:

$$d(P_c)_{\phi_0}(\vec{X}) = \langle \nabla P_c(\phi_0), \vec{X} \rangle = \int_M \nabla P_c(\phi_0) \cdot \vec{X} \mathrm{dvol}_M$$

where the integrand uses the Euclidean dot product in $\mathbb{R}^N$. (Note that this is the L^2 gradient, as the standard L^2 inner product is used in the computation. Also the dvol volume element is a function of the embedding ϕ because it uses the metric g which is the induced metric of $\mathbb{R}^N$ onto $\phi(M)$). Additionally, we know that for a parameterized curve in $\phi_s \in \mathrm{Emb}(M, \mathbb{R}^N)$, where s belongs to an interval around 0 such that $\phi_s(0) = \phi_0$ and

$\frac{\mathrm{d}\phi_s}{\mathrm{ds}}|_{s=0} = \vec{X}$, we have

$$d(P_c)_{\phi_0}(\vec{X}) = \frac{\mathrm{d}P_c}{\mathrm{ds}}|_{s=0} = \frac{\mathrm{d}}{\mathrm{ds}}|_{s=0} \int_M |R(\phi_s(m))|^2 \mathrm{dvol}_M$$

$$= \int_M \frac{\mathrm{d}}{\mathrm{ds}}|_{s=0} |R(\phi_s(m))|^2 \mathrm{dvol}_M + \int_M |R(\phi_s(m))|^2 \frac{\mathrm{d}}{\mathrm{ds}}|_{s=0} \mathrm{dvol}_M \quad (1)$$

If we can arrange the results of this computation into the form $\int_M f \cdot \vec{X} \mathrm{dvol}_M$ then we can conclude that $\nabla P_c = f$. This computation is handled in Chapter 2. While known techniques can be used for the second term in (1), the first term and in particular the integrand $\frac{\mathrm{d}}{\mathrm{ds}}|_{s=0}|R(\phi(m))|^2$ had not been explicitly computed prior to this.

Adding the assumption that the initial manifold M has a boundary gives the next set of computations in Chapter 3 for the boundary terms of the final gradient expression derived in Chapter 2.

Chapter 4

In carrying out a gradient flow analysis in the space of smooth embeddings $\mathrm{Emb}(M, \mathbb{R}^N)$, we are ultimately interested in how long continuous flow exists, how to determine if we are approaching a global minimum ϕ (as opposed to a local minimum), time estimates for the flow and other related questions. These questions are complicated by the fact that $\mathrm{Emb}(M, \mathbb{R}^N)$ is open in the space of all smooth maps $C^\infty(M, \mathbb{R}^N)$ (considered in the Fréchet topology). This suggests that gradient flow can flow "out of the space" of embeddings by, for example, two image points coinciding (which would make the embedding no longer injective). From a programming standpoint though, gradient flow will be done in discretized time steps. This means that the gradient vector will be explicitly computed at a point $\phi \in \mathrm{Emb}(M, \mathbb{R}^N)$. $\phi(M)$ will then be allowed to travel in the direction of this fixed vector field for a fixed short period time after which a new gradient vector will be computed etc. In the second project we quantify the short period of time for which $\phi(M)$ can move in its fixed gradient direction while remaining an embedding in $\mathbb{R}^N$ (this requires checking immersion and injective conditions). This work is based on Milnor's set up of focal points and manifolds in Euclidean space in *Morse Theory* [5]. It is important to note that his work depends on the vector field's being normal to $\phi(M)$ at every point. For this reason, Chapter 4 will start with a proof of a condition on the penalty function that will guarantee its gradient being normal to $\phi(M)$ (namely that the penalty function is diffeomorphism invariant).

This project uses applications of the ϵ-neighborhood theorem, a quantitative version of the inverse function theorem and Milnor's work on critical and focal points.

Chapter 5

Chapter 5 investigates the distance penalty term $P_d = \int_M d^2(\phi(m), x_i)\mathrm{dvol}_M$ from the original penalty function. Rosenberg's former work treated the computation of the gradient in three different ways (holding the volume form constant, projecting the gradient onto normal directions and varying the volume form). We want to investigate existence of flow in the case of embedding a circle into $\mathbb{R}^2$, where the set S consists of the origin only. Each of the three computations gives rise to a different coupled PDE ($\langle \frac{\mathrm{d}\phi^1}{\mathrm{dt}}, \frac{\mathrm{d}\phi^2}{\mathrm{dt}} \rangle = -\mathrm{grad} P_d(\phi)$). To be consistent with the way we compute the gradient of the curvature term, we are first interested in the case of the varying volume form. Unfortunately this leads to a second order, nonlinear PDE to which we can't apply any standard PDE techniques. However when we consider the distance term with fixed volume element dvol and project its gradient onto normal directions we get a first order, analytic nonlinear PDE to which we can apply the Cauchy-Kovalevskaya theorem to conclude short time existence of the flow. The project rigorously applies the proof of the theorem to our set up and extracts a lower bound on existence for a short time solution.

The final component of the project is the extension of the example to all simple closed curves in $\mathbb{R}^2$.

A discussion of future directions is included in Concluding remarks.

Note: Chapter 2 consists of Steve Rosenberg's notes written before my contributions to this project. My contribution begins on page 21 with the general computation of the adjoint operator on 2-tensors. Appendix C, or the computation of the distance penalty gradient terms is also from Steve's previous work.

Chapter 2

Gradient Computation

As stated in the introduction, we are interested in a penalty function that involves both a distance and a curvature term:

$$P(\phi) = \int_M |R(\phi(m))|^2 \mathrm{dvol}_M + \int_M d^2(\phi(m), x_i) \mathrm{dvol}_M$$

In this chapter (based on computations done by Steve Rosenberg) we compute the gradient of only the curvature term:

$$P_c(\phi) = \int_M |R(\phi(m))|^2 \mathrm{dvol}_M$$

Important Note: Throughout Chapters 2 and 3, integrals will be written (for ease of notation) as above, taken over M with the volume form also being denoted on M. It is crucial to note that what is actually meant is

$$P_c(\phi) = \int_{\phi(M)} |R(\phi(m))|^2 \mathrm{dvol}_{g_\phi}$$

where g_ϕ is the metric induced from $\mathbb{R}^N$ onto the embedded image, $\phi(M)$. Let $\phi : M \to \mathbb{R}^N$ be an embedding of a fixed closed manifold M into $\mathbb{R}^N$. $N = \phi(M)$ has the induced metric g_ϕ from $\mathbb{R}^N$, so this induces a metric on M. We want to vary the embedding in the direction of a vector field X along $\phi(M)$, so $X_m \in T_{\phi(m)}\mathbb{R}^N$ for $m \in M$. Note, with this set up, the pull back of the tangent bundle $T\mathbb{R}^N|_{\phi(M)}$ on $\phi(M) \subset \mathbb{R}^N$ is given by:

$$\phi^* T\mathbb{R}^N|_{m\in M} = T_{\phi(m)}\mathbb{R}^N$$

(Thus $X \in \Gamma(\phi^* T\mathbb{R}^N)$.) We want to understand the variation of the total curvature function

$$P_c(\phi) = \int_M |R(\phi(m))|^2 \mathrm{dvol}_M = \int_M R^{ijkl} R_{ijkl}\ \mathrm{dvol}_M$$

in the direction X. In particular, we want the gradient of this function on the set of all embeddings $\mathrm{Emb}(M, \mathbb{R}^N)$, namely a formula of the form

$$\delta_X P_c(\phi) = dP_{c_\phi}(X) = \int_M \langle X, Z\rangle \mathrm{dvol}_M, \quad i.e.\ \nabla P_c(\phi) = Z \in \Gamma(\phi^* T\mathbb{R}^N). \qquad (2.1)$$

Note that $\langle X, Z\rangle = X \cdot Z$ is just the dot product of vectors in $\mathbb{R}^N$, so usually we'll just write $X \cdot Z$.

The tangent space $T_\phi \mathrm{Emb}(M, \mathbb{R}^N) = \Gamma(\phi^* T\mathbb{R}^N)$ in some Sobolev or Fréchet topology has L^2 inner product

$$g'_\phi(Y, Y') = \langle Y, Y'\rangle_\phi = \int_M Y_{\phi(m)} \cdot Y'_{\phi(m)} \mathrm{dvol}_M(m).$$

We consider the function

$$P_c : \mathrm{Emb}(M, \mathbb{R}^N) \to \mathbb{R},\ P_c(\phi) = \int_M |R_\phi|^2 \mathrm{dvol}_M, \qquad (2.2)$$

where R_ϕ is the curvature tensor of g_ϕ, and the norm $|R_\phi|^2$ is calculated in the metric g_ϕ. We want to write

$$dP_{c_\phi(X)} = g'_\phi(X, Z_\phi) = \int_M X \cdot Z_\phi\ \mathrm{dvol}_M$$

for some $Z_\phi \in T_\phi \mathrm{Emb}(M, \mathbb{R}^N)$, in which case $Z_\phi = \mathrm{grad}(C)$ is the gradient vector field for

C.

To make the setup precise, we want to do all the calculations on M. Now $\phi^* g_\phi$ is a metric on M, which we just denote by g_ϕ. Similarly, we can pull R_ϕ back to

$$m \in M \mapsto R^i{}_{jkl}(\phi(m))\phi_*^{-1}\partial_i \otimes \phi^* dx^j \otimes \phi^* dx^k \otimes \phi^* dx^l,$$

which we just denote by R_ϕ. In this notation, $|R_\phi|^2 = |R_\phi|^2_{g_\phi}$ is the same in the old or the new notation, so the integral in (2.2) is unchanged.

Let $T_M = TM \otimes T^*M^{\otimes 3}$, so (the pullback) $R_\phi \in \Gamma(T_M)$. Let $\mathcal{M} = M \times \mathrm{Emb}(M, \mathbb{R}^N) \to \mathbb{R}^N$. As ϕ varies, this produces a section $R \in \Gamma(\pi^* T_M \to \mathcal{M})$, where $\pi : \mathcal{M} \to M$ is the projection.

$\mathcal{M}$ has the metric

$$h_{(m,\phi)} = \begin{pmatrix} g_\phi(m) & \\ & g'_\phi \end{pmatrix},$$

which is in block form but is not a product metric. The associated LC connection ∇^h on $\mathcal{M}$ is given by the six term formula

$$\begin{aligned} \langle \nabla^h_{\tilde{X}} \tilde{Y}, \tilde{Z} \rangle &= \tilde{X}\langle \tilde{Y}, \tilde{Z}\rangle_h + \tilde{Y}\langle \tilde{X}, \tilde{Z}\rangle_h - \tilde{Z}\langle \tilde{X}, \tilde{Y}\rangle_h \\ &\quad + \langle [\tilde{X}, \tilde{Y}], \tilde{Z}\rangle_h + \langle [\tilde{Z}, \tilde{X}], \tilde{Y}\rangle_h - \langle [\tilde{Y}, \tilde{Z}], \tilde{X}\rangle_h \end{aligned} \tag{2.3}$$

for $\tilde{X}, \tilde{Y}, \tilde{Z} \in T\mathcal{M}$. Writing $\tilde{X} = X_1 + X_2$ with $X_1 \in TM, X_2 \in T\mathrm{Emb}(M, \mathbb{R}^N)$ and similarly for $\tilde{Y}, \tilde{Z}$, we see that for each $m \in M, \phi \in \mathrm{Emb}(M, \mathbb{R}^N)$,

$$\nabla^h|_{T(M \times \{\phi\})} = \nabla^\phi, \ \nabla^h|_{T(\{m\} \times \mathrm{Emb}(M, \mathbb{R}^N))} = \nabla',$$

where ∇' is the LC connection associated to g'. ∇^h extends to tensor bundles over $\mathcal{M}$; in particular, expressions like $\nabla^h R = \nabla^h R(m, \phi)$ make sense. This will be used starting in (2.11).

2.1 The basic variational formula

To begin the gradient computation, we have

$$\delta_X P_c(\phi) = \int_M \delta_X \langle R, R\rangle \text{dvol} + \int_M |R|^2 \delta_X \text{dvol}. \tag{2.4}$$

Note: δ_X is explained in Appendix A.

2.2 The Computation Terms

The second term on the RHS of (2.4) is treated in **Summary dvol** in (2.8). The first term on the RHS of (2.4) is treated starting with (2.12). It involves five terms: the first term on the RHS of (2.12) breaks into six terms, which are treated in **Summary AI − AVI** and whose computations are covered in sections **2.5** and **2.6**. The other four terms on the RHS of (2.12) are treated in **Summary CII − CV** in section **2.4**.

2.3 The term $\delta_X$dvol in (2.4)

As in [4, p.7] let ω be the one-form on $\phi(M)$ defined by $\omega(Y) = X \cdot Y$ for $Y \in T(\phi(M))$. By pullback, we can consider ω to be a one-form on M. Then

$$\delta_X \text{dvol} = -X \cdot \text{Tr II dvol} + d * \omega,$$

where $*$ is the Hodge star, $d = d_M$ and Tr II is defined in Appendix A. Therefore

$$\begin{aligned}\int_M |R|^2 \delta_X \text{dvol} &= \int_M |R|^2(-X \cdot \text{Tr II dvol} + d * \omega) \\ &= \int_M (-X \cdot |R|^2 \text{Tr II}) \text{dvol} - \int_M d|R|^2 \wedge *\omega \\ &= \int_M (-X \cdot |R|^2 \text{Tr II}) \text{dvol} - \int_M \langle d|R|^2, \omega\rangle \text{dvol}.\end{aligned} \tag{2.5}$$

Here by Stokes' Theorem, $\int_M |R|^2 d*\omega = -\int_M d|R|^2 \wedge *\omega$ and $\int_M \alpha \wedge *\beta = \int_M \langle \alpha, \beta\rangle \mathrm{dvol}$, which is valid on any Riemannian manifold. Focusing on the last term in (2.5), we get

$$d|R|^2 = d\langle R, R\rangle = 2\langle \nabla R, R\rangle, \tag{2.6}$$

where we consider $R \in \Gamma(T_M)$. Note that $\langle \nabla R, R\rangle \in \Omega^1(M)$. Also, the standard index lowering isomorphism $\alpha : T_x\bar{M} \to T_x^*\bar{M}$ on any Riemannian manifold $\bar{M}$ is given by $X \mapsto X^\flat = \langle X, \cdot\rangle$, so for our case $\bar{M} = \mathbb{R}^N$, we get $\omega = X^\flat$. Thus the last term in (2.5) is

$$\begin{aligned} \int_M \langle d|R|^2, \omega\rangle \mathrm{dvol} &= 2\int_M \left\langle \langle \nabla R, R\rangle, X^\flat \right\rangle \mathrm{dvol} \\ &= 2\int_M \left\langle \langle \nabla R, R\rangle^\sharp, X \right\rangle \mathrm{dvol}, \end{aligned} \tag{2.7}$$

where $\alpha^{-1}(c) = c^\sharp$. Combining (2.5) – (2.7) gives

Summary dvol:

$$\int_M |R|^2 \delta_X \mathrm{dvol} = \int_M \left\langle -|R|^2 \mathrm{Tr}\ \mathrm{II} - 2\langle \nabla R, R\rangle^\sharp, X \right\rangle \mathrm{dvol}. \tag{2.8}$$

2.4 The term $\delta_X\langle R, R\rangle$ in (2.4), part I.

Preliminary calculation: Let (x^i) be local coordinates on $U' \subset M$. In detail, we have a diffeomorphism $\alpha : U \to U', U \subset \mathbb{R}^k$. Then ϕ induces $\phi \circ \alpha$, and tangent spaces to $\phi(U')$ have basis $(\phi \circ \alpha)_*(\partial/\partial x^i)$, where $\partial/\partial x^i$ are the standard tangent vectors in $\mathbb{R}^k$. We have

$$(\phi \circ \alpha)_* \left(\frac{\partial}{\partial x^i}\right) = \frac{\partial(\phi \circ \alpha)}{\partial x^i} = \frac{\partial \phi}{\partial x^i} = \phi_i,$$

where the last two terms are usual definitions (i.e. derivatives are taken component wise in the ϕ vector). Thus in these coordinates,

$$g_{ij} = g_{ij,\phi} = g_\phi(\partial_i, \partial_j) = \phi^* g(\partial_i, \partial_j) = g(\phi_*\partial_i, \phi_*\partial_j) = \phi_*\partial_i \cdot \phi_*\partial_j = \phi_i \cdot \phi_j. \tag{2.9}$$

We remind ourselves that ϕ_i means $(\partial/\partial x^i)(\phi \circ \alpha)$.

For $\varepsilon \approx 0$, $\phi_\varepsilon : M \to \mathbb{R}^N$, $\phi_\varepsilon(m) = \phi(m) + \varepsilon X_{\phi(m)}$ is also an embedding, so $N_\varepsilon = \phi_\varepsilon(M)$ is still diffeomorphic to M. Thus if $\{x^i\}$ are local coordinates on M near m, the tangent space $T_{\phi_\varepsilon(m)} N_\varepsilon$ is spanned by

$$\left\{ \frac{\partial \phi_\varepsilon}{\partial x^i} \right\} = \left\{ \frac{\partial \phi}{\partial x^i} + \varepsilon \frac{\partial X}{\partial x^i} \right\}.$$

Thus in these coordinates

$$\begin{aligned} \delta_X g_{ij} &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \left(\frac{\partial \phi_\varepsilon}{\partial x^i} \cdot \frac{\partial \phi_\varepsilon}{\partial x^j} \right) = \frac{\partial X}{\partial x^i} \cdot \frac{\partial \phi}{\partial x^j} + \frac{\partial X}{\partial x^j} \cdot \frac{\partial \phi}{\partial x^i} \\ &\stackrel{\text{def}}{=} X_i \cdot \phi_j + X_j \cdot \phi_i, \end{aligned} \tag{2.10}$$

where $\cdot$ is the dot product of vectors in $\mathbb{R}^N$.

Now we begin the calculation of the first term on the RHS of (2.4).

For the first term on the RHS of (2.4), $R = R_\phi$ is a tensor on $\mathcal{M}$ involving only M coordinates but depending on ϕ. Thus

$$\delta_X \langle R, R \rangle_g = \delta_X \langle R, R \rangle_{g_\phi} = X \langle R, R \rangle_h = 2 \langle \nabla^h_X R, R \rangle_h. \tag{2.11}$$

Note that X is in $T_\phi \mathrm{Emb}(M, \mathbb{R}^N)$. The point is that (2.11) avoids differentiating the metric – there is no term of the form $\delta_X g$. Now

$$\begin{aligned} \nabla^h_X R &= (\delta_X R^i{}_{jkl}) \partial_i \otimes dx^j \otimes dx^k \otimes dx^l + R^i{}_{jkl} (\nabla^h_X \partial_i) \otimes dx^j \otimes dx^k \otimes dx^l \\ &\quad + R^i{}_{jkl} \partial_i \otimes (\nabla^h_X dx^j) \otimes dx^k \otimes dx^l + R^i{}_{jkl} \partial_i \otimes dx^j \otimes (\nabla^h_X dx^k) \otimes dx^l \end{aligned}$$

$$+R^i{}_{jkl}\partial_i \otimes dx^j \otimes dx^k \otimes (\nabla^h_X dx^l). \tag{2.12}$$

We will compute the first term $\langle(\delta_X R^i{}_{jkl})\partial_i \otimes dx^j \otimes dx^k \otimes dx^l, R\rangle$ in sections **2.5** and **2.6**, which are summarized in **Summary AI − AVI.** Now, we will compute the other four terms on the RHS, which are summarized in **Summary CII- CV**.

To compute the second term on the RHS of (2.12), we have

$$\begin{aligned} 2\langle\nabla^h_X\partial_i, \partial_a\rangle_h &= X\langle\partial_i, \partial_a\rangle_h + \partial_i\langle X, \partial_a\rangle_h - \partial_a\langle X, \partial_i\rangle_h \\ &\quad +\langle[X, \partial_i], \partial_a\rangle_h + \langle[\partial_a, X], \partial_i\rangle_h - \langle[\partial_i, \partial_a], X\rangle_h \qquad (2.13) \\ &= X\langle\partial_i, \partial_a\rangle_h = X\langle\partial_i, \partial_a\rangle_g = X(g_{ia}) \\ &= X_i \cdot \phi_a + X_a \cdot \phi_i, \end{aligned}$$

where we have used (i) $\langle X, \partial_i\rangle_h = \langle X, \partial_a\rangle_h = 0$, (ii) $[X, \partial_i] = [X, \partial_a] = 0$ (since we can extend $X \in T_\phi \mathrm{Emb}(M, \mathbb{R}^N)$ to a vector field on $\mathcal{M}$ which is independent of $m \in M$), (iii) $[\partial_i, \partial_a] = 0$, and (iv) (2.10) for the final line. Thus the contribution of the second term on the RHS of (2.12) to (2.11) is[1]

$$\begin{aligned} &\langle 2R^i{}_{jkl}(\nabla^h_X\partial_i) \otimes dx^j \otimes dx^k \otimes dx^l, R^a{}_{bcd}\partial_a \otimes dx^b \otimes dx^c \otimes dx^d\rangle_g \\ &= R^i{}_{jkl}(X_i \cdot \phi_a + X_a \cdot \phi_i)R^a{}_{bcd}g^{jb}g^{kc}g^{ld} \\ &= R_{rjkl}g^{ri}X_i \cdot \phi_a R^a{}_{bcd}g^{jb}g^{kc}g^{ld} + R^i{}_{jkl}X_a \cdot \phi_i g^{ar}R_{rbcd}g^{jb}g^{kc}g^{ld} \\ &= \langle dX, \langle R, \langle d\phi, R\rangle_1\rangle_{234}\rangle + \langle dX, \langle\langle d\phi, R\rangle_1, R\rangle_{234}\rangle \\ &= 2\langle dX, \langle R, \langle d\phi, R\rangle_1\rangle_{234}\rangle \\ &= 2\langle X, -\mathrm{div}(\langle R, \langle d\phi, R\rangle_1\rangle_{234})\rangle, \end{aligned}$$

where $d\phi = \nabla\phi$ just as a matter of notation/convenience, and in the last line we have switched R from a (4,0) tensor to a (3,1) tensor. To review the notation: $\langle d\phi, R\rangle_1$ means

[1] Here we use $\langle \mathrm{grad} f, Y\rangle = \langle f, -\mathrm{div} Y\rangle$ for a function f and a vector field Y.

the pairing of $d\phi$ with the first = upper index in R, leaving a (3,0) tensor; $\langle R, \langle d\phi, R\rangle_1\rangle_{234}$ is the pairing of this (3,0) tensor with the last three indices of the (4,0) tensor R, leaving a (1,0) tensor; in the last line, we consider $\langle R, \langle d\phi, R\rangle_1\rangle_{234}$ as a (0,1) tensor (with one upper index i).

Recall that for a vector field $Y = Y^r\partial_r$,

$$\text{div } Y = \frac{1}{\sqrt{\det(g)}}\partial_r\Big(\sqrt{\det(g)}Y^r\Big)$$

Summary CII: The second term on the RHS of (2.12) contributes

$$\begin{aligned} -2\langle X, \text{div}(\langle R, \langle d\phi, R\rangle_1\rangle_{234})\rangle &= -2\Big\langle X, \frac{1}{\sqrt{\det(g)}}\partial_r\Big(\sqrt{\det(g)}R^r{}_{jkl}\phi_a R^a{}_{bcd}g^{jb}g^{kc}g^{ld}\Big)\Big\rangle \\ &= -2\Big\langle X, \frac{1}{\sqrt{\det(g)}}\partial_r\Big(\sqrt{\det(g)}R^r{}_{jkl}\phi_i R^{ijkl}\Big)\Big\rangle \end{aligned}$$

to (2.11).

For the third term on the RHS of (4.3), we note that

$$\langle\nabla^h_X dx^j, \partial_i\rangle + \langle dx^j, \nabla^h_X\partial_i\rangle = X\langle dx^j, \partial_i\rangle = 0 \Longrightarrow \langle\nabla^h_X dx^j, \partial_i\rangle = -\langle dx^j, \nabla^h_X\partial_i\rangle.$$

For $\nabla^h_X\partial_i = A^k_i\partial_k$, we have

$$A^k_i g_{kb} = \langle\nabla^h_X\partial_i, \partial_b\rangle = \frac{1}{2}(X_i\cdot\phi_b + X_b\cdot\phi_i) \Longrightarrow A^k_i = \frac{1}{2}g^{bk}(X_i\cdot\phi_b + X_b\cdot\phi_i),$$

so

$$\begin{aligned} \nabla^h_X dx^j &= \langle\nabla^h_X dx^j, \partial_r\rangle dx^r = -\langle dx^j, \nabla^h_X\partial_r\rangle dx^r \\ &= -\langle dx^j, \frac{1}{2}g^{sk}(X_r\cdot\phi_s + X_s\cdot\phi_r)\partial_k\rangle dx^r \\ &= -\frac{1}{2}g^{sj}(X_r\cdot\phi_s + X_s\cdot\phi_r)dx^r. \end{aligned}$$

Thus the contribution of the third term on the RHS of (2.12) to (2.11) is

$$\begin{aligned}
&\langle 2R^i{}_{jkl}\partial_i \otimes (\nabla^h_X dx^j) \otimes dx^k \otimes dx^l, R^a{}_{bcd}\partial_a \otimes dx^b \otimes dx^c \otimes dx^d \rangle_g \\
&= R^i{}_{jkl} g_{ia}(-g^{sj}(X_r \cdot \phi_s + X_s \cdot \phi_r)) g^{rb} R^a{}_{bcd} g^{kc} g^{dl} \\
&= \langle dX, -\langle\langle d\phi, R\rangle_2, R\rangle_{134}\rangle + \langle dX, -\langle\langle d\phi, R\rangle_2, R\rangle_{134}\rangle \\
&= -2\langle X, -\mathrm{div}(\langle\langle d\phi, R\rangle_2, R\rangle_{134})\rangle.
\end{aligned}$$

Summary CIII: The third term on the RHS of (2.12) contributes

$$2\langle X, \mathrm{div}(\langle R, \langle \mathrm{grad}(\phi), R\rangle_2\rangle_{134})\rangle$$

$$\begin{aligned}
&= 2\left\langle X, \frac{1}{\sqrt{\det(g)}}\partial_r\left(\sqrt{\det(g)} R^i{}_{jkl} g_{ia} g^{sj}\phi_s g^{rb} R^a{}_{bcd} g^{kc} g^{ld}\right)\right\rangle \\
&= 2\left\langle X, \frac{1}{\sqrt{\det(g)}}\partial_r\left(\sqrt{\det(g)} R^{ij}{}_{kl}\phi_j R_i{}^{rkl}\right)\right\rangle
\end{aligned}$$

to (2.11).

The fourth and fifth terms on the RHS of (2.12) are similar to the previous term.

Summary CIV + CV: The fourth and fifth terms on the RHS of (2.12) contribute

$$\begin{aligned}
&2\langle X, \mathrm{div}(\langle R, \langle \mathrm{grad}(\phi), R\rangle_3\rangle_{124}) + \mathrm{div}(\langle R, \langle \mathrm{grad}(\phi), R\rangle_4\rangle_{123})\rangle \\
&= 2\left\langle X, \frac{1}{\sqrt{\det(g)}}\partial_r\left(\sqrt{\det(g)} R^i{}_j{}^k{}_l \phi_k R_i{}^{jrl}\right)\right\rangle \\
&\quad +2\left\langle X, \frac{1}{\sqrt{\det(g)}}\partial_r\left(\sqrt{\det(g)} R^i{}_{jk}{}^l \phi_l R_i{}^{jkr}\right)\right\rangle
\end{aligned}$$

to (2.11).

We show that the terms integrated against X in CII, CIII, CIV, CV are all the same, by showing that the curvature expressions are equal.

Lemma 1. $R^r{}_{jkl}\phi_i R^{ijkl} = R^{ij}{}_{kl}\phi_j R_i{}^{rkl} = R^i{}_j{}^k{}_l\phi_k R_i{}^{jrl} = R^i{}_{jk}{}^l\phi_l R_i{}^{jkr}$.

Proof. The second, third and fourth curvature terms in the Lemma equal the first curvature term, since

$$\begin{aligned}
R^{ij}{}_{kl}\phi_j R_i{}^{rkl} &= R^{ji}{}_{kl}\phi_i R_j{}^{rkl} = R^{ij}{}_{kl}\phi_i R^r{}_j{}^{kl} = R^{ijkl}\phi_i R^r{}_{jkl},\\
R^i{}_j{}^k{}_l\phi_k R_i{}^{jrl} &= R^k{}_l{}^i{}_j\phi_i R_k{}^{lrj} = R^i{}_j{}^k{}_l\phi_i R^{rj}{}_k{}^l = R^{ijkl}\phi_i R^r{}_{jkl},\\
R^i{}_{jk}{}^l\phi_l R_i{}^{jkr} &= R^k{}_{lj}{}^i\phi_i R_k{}^{ljr} = R^i{}_j{}^k{}_l\phi_i R^{rj}{}_k{}^l = R^{ijkl}\phi_i R^r{}_{jkl}.
\end{aligned}$$

□

Noting that the CII contribution comes with a minus sign, we get:

Summary CII − CV: The second, third, fourth and fifth terms on the RHS of (2.12) contribute

$$4\left\langle X, \frac{1}{\sqrt{\det(g)}}\partial_r\left(\sqrt{\det(g)}R^r{}_{jkl}\phi_i R^{ijkl}\right)\right\rangle$$

to (2.11).

This concludes the contribution of $\delta_X\langle R, R\rangle$, except for the first term in (2.12), which will be treated next.

2.5 The term $\delta_X\langle R, R\rangle$ in (2.4), part II.

We next compute $\delta_X R^i{}_{jkl}$ from (2.12). The useful formula is the Corollary given in (2.25), and the coordinate free version is Thm. 1.

To compute $\delta_X R$, we certainly need to understand $\delta_X\nabla$. The difference of connections on M is tensorial, so $\delta_X \nabla \in \Lambda^1(M, Hom(TM))$ is also tensorial. We have

$$(\delta_X \nabla)_{\partial_k} \partial_l = \delta_X(\nabla_{\partial_k} \partial_l) = (\delta_X \Gamma^r_{kl} \partial_r) = \dot{\Gamma}^r_{kl} \partial_r$$

for ease of notation (where $\dot{\Gamma}$ refers to a derivative in the X direction). As before, for the vectors $X, \phi \in \mathbb{R}^N$, set

$$X_a = \left(\dots, \frac{\partial X^i}{\partial x^a}, \dots \right), \phi_b = \left(\dots, \frac{\partial \phi^i}{\partial x^b}, \dots \right),$$

as vectors in $\mathbb{R}^N$. For $\delta_X g_{ij} = \dot{g_{ij}}$, (2.10) becomes

$$\delta_X g_{ij} \dot{g}_{ij} = X_i \cdot \phi_j + X_j \cdot \phi_i. \tag{2.14}$$

Then

$$\begin{aligned} 2\dot{\Gamma}^r_{kl} &= \delta_X [g^{rm}(\partial_l g_{mk} + \partial_k g_{ml} - \partial_m g_{kl})] \\ &= -g^{ra} \dot{g}_{ab} g^{bm}(\partial_l g_{mk} + \partial_k g_{ml} - \partial_m g_{kl}) + g^{rm}(\partial_l \dot{g}_{mk} + \partial_k \dot{g}_{ml} - \partial_m \dot{g}_{kl}) \\ &= -2g^{ra}(X_a \cdot \phi_b + X_b \cdot \phi_a)\Gamma^b_{lk} \qquad (2.15) \\ &\quad +2g^{ra}(X_a \cdot \phi_{kl} + X_{kl} \cdot \phi_a), \end{aligned}$$

where $(A^{-1})^{\cdot} = -A^{-1}\dot{A}A^{-1}$ for an invertible matrix A.

More notation: X is an $\mathbb{R}^N$-valued function on M, so it has a vector-valued Hessian

$$\text{Hess } X = \nabla dX \in \Gamma(T^*M \otimes T^*M)^N.$$

$d\phi$ is a vector-valued one-form on M, and so has the corresponding $\mathbb{R}^N$-valued vector field $(d\phi)^\sharp = \text{grad}(\phi)$.

Definition 1.

$$\operatorname{grad}(\phi)\cdot \text{Hess } X \in \Gamma(TM\otimes T^*M\otimes T^*M) = \Gamma(Hom(TM\otimes TM, TM))$$

is the tensor product of $\operatorname{grad}(\phi) = (d\phi)^\sharp$ *with* Hess X *followed by taking the dot product in the* $\mathbb{R}^n$ *variables.*

Lemma 2. *For* $A, B \in T_mM$*,*

$$\begin{aligned}(\delta_X\nabla)_A B &= (\operatorname{grad}(\phi)\cdot \text{Hess } X + \operatorname{grad}(X)\cdot \text{Hess } \phi)(A,B)\\ &= \operatorname{grad}(\phi)\cdot(\text{Hess } X(A,B)) + \operatorname{grad}(X)\cdot(\text{Hess } \phi(A,B)),\end{aligned}$$

where on the left hand side B *is extended arbitrarily to a vector field near* m*.*

Note that the RHS is the dot product of an $\mathbb{R}^N$-valued tangent vector with an $\mathbb{R}^N$-valued function, and so is a tangent vector to M.

Proof. We have

$$\begin{aligned}\operatorname{grad}(\phi)\cdot(\text{Hess } X(\partial_k,\partial_l)) &= g^{ra}\phi_a\partial_r\cdot \nabla dX(\partial_k,\partial_l)\\ &= g^{ra}\phi_a\partial_r\cdot(\partial_k dX - X_b\Gamma^b_{ks}dx^s)(\partial_l)\\ &= g^{ra}\phi_a\partial_r\cdot(X_{kl} - X_b\Gamma^b_{kl})\\ &= [g^{ra}\phi_a\cdot(X_{kl} - X_b\Gamma^b_{kl})]\partial_r\end{aligned}$$

This equals two of the terms on the RHS of (2.15). Similarly, the other two terms in (2.15) equal $\operatorname{grad}(X)\cdot(\text{Hess } \phi(\partial_k,\partial_l))$. Thus

$$(\delta_X\nabla)_{\partial_k}\partial_l = \operatorname{grad}(\phi)\cdot(\text{Hess } X(\partial_k,\partial_l)) + \operatorname{grad}(X)\cdot(\text{Hess } \phi(\partial_k,\partial_l))$$

Since both sides of this equation are tensorial, the same expression holds for general tangent vectors A, B.

□

We now calculate the variation of the curvature tensor in invariant terms.
Our conventions:

$$R(\partial_k, \partial_l)\partial_j = (\nabla_k\nabla_l - \nabla_l\nabla_k)\partial_j = R^i{}_{jkl}\partial_i, \text{with } R^i{}_{jkl} = \partial_k\Gamma^i_{lj} - (k \leftrightarrow l) + \Gamma^r_{lj}\Gamma^i_{kr}\partial_i - (k \leftrightarrow l).$$

Theorem 1. *We have*

$$\begin{aligned} \delta_X R(\partial_k, \partial_l)\partial_j &= \Big(\text{Hess } X(\cdot, \partial_k)^\sharp \cdot \text{Hess } \phi(\partial_l, \partial_j) - (k \leftrightarrow l) \\ &\quad -\text{grad}(X) \cdot \langle \text{grad}(\phi), R(\partial_k, \partial_l)\partial_j\rangle\Big) + (X \leftrightarrow \phi). \end{aligned} \tag{2.16}$$

Here $(k \leftrightarrow l)$ refers to the previous term, and $(X \leftrightarrow \phi)$ refers to the previous three terms.

Proof. $\delta_X R^i{}_{jkl}$ is the i^{th} component of

$$\begin{aligned} (\delta_X R)(\partial_k, \partial_l)\partial_j &= \delta_X(\nabla_k\nabla_l\partial_j - (k \leftrightarrow l) - \nabla_{[\partial_k, \partial_l]}\partial_j) \\ &= (\delta_X\nabla)_k\nabla_l\partial_j + \nabla_k(\delta_X\nabla)_l\partial_j - (k \leftrightarrow l), \end{aligned}$$

since $[\partial_k, \partial_l] = 0$. By Lemma 2,

$$\begin{aligned} (\delta_X R)(\partial_k, \partial_l)\partial_j &= \text{grad}(\phi) \cdot \text{Hess } X(\partial_k, \nabla_{\partial_l}\partial_j) + \text{grad}(X) \cdot \text{Hess } \phi(\partial_k, \nabla_{\partial_l}\partial_j) \\ &\quad +\nabla_k(\text{grad}(\phi) \cdot \text{Hess } X(\partial_l, \partial_j)) + \nabla_k(\text{grad}(X) \cdot \text{Hess } \phi(\partial_l, \partial_j)) \\ &\quad -(k \leftrightarrow l). \end{aligned} \tag{2.17}$$

(Here $(k \leftrightarrow l)$ refers to the previous four terms.)

We now work in normal coordinates at a point, where $g_{ab} = \delta_{ab}, \partial_i g_{ab} = \Gamma^s_{lj} = 0$ and e.g. Hess $X(\partial_k, \partial_l) = X_{kl}$. The second term on the RHS of (2.17) contains

$$\text{Hess } \phi(\partial_k, \nabla_{\partial_l}\partial_j) = \Gamma^s_{lj}\text{Hess } \phi(\partial_k, \partial_s) = 0$$

at this point. Thus the second and sixth term vanish. By symmetry in ϕ, X, the first and fifth terms also vanish.

The fourth term in normal coordinates equals

$$\nabla_k(\text{grad}(X) \cdot \text{Hess } \phi(\partial_l, \partial_j)) = \nabla_k(g^{ra}X_a\partial_r)\phi_{lj} + \text{grad}(X) \cdot \partial_k(\nabla d\phi(\partial_l, \partial_j)). \tag{2.18}$$

For the first term on the RHS of (2.18), only $g^{ra}X_{ak}\phi_{lj}\partial_r$ survives, which we allow ourselves to write as $X_{ak}\phi_{lj}\partial_a$. The last term in (2.18) is

$$\begin{aligned} X_a\partial_a \cdot \partial_k(\nabla d\phi(\partial_l, \partial_j)) &= X_a\partial_a(\partial_k(\nabla(\phi_b dx^b)(\partial_l, \partial_j))) \qquad (2.19) \\ &= X_a\partial_a(\partial_k(\partial_s\phi_b dx^b \otimes dx^s(\partial_l, \partial_j) - \phi_b\Gamma^b_{cd}dx^c \otimes dx^d(\partial_l, \partial_j))) \\ &= [X_a(\phi_{kjl} - \phi_b\partial_k\Gamma^b_{lj})]\partial_a. \end{aligned}$$

Thus the fourth term and the eighth term in (2.17) equal

$$\begin{aligned} &[X_{ak}\phi_{lj} + X_a\phi_{kjl} - X_a\phi_b\partial_k\Gamma^b_{lj} - (k \leftrightarrow l)]\partial_a \\ &= [X_{ak}\phi_{lj} - X_{al}\phi_{kj} - X_a\phi_b R^b{}_{jkl}]\partial_a. \end{aligned} \tag{2.20}$$

The RHS of (2.20) in more invariant terms is

$$[\text{Hess } X(\partial_a, \partial_k)\cdot\text{Hess } \phi(\partial_l, \partial_j) - \text{Hess } X(\partial_a, \partial_l)\cdot\text{Hess } \phi(\partial_k, \partial_j)]\partial_a - \text{grad}(X)\cdot\phi_b R^b{}_{jkl}. \tag{2.21}$$

Note that

$$\phi_b R^b{}_{jkl} = d\phi(R(\partial_k, \partial_l)\partial_j) = \langle\text{grad}(\phi), R(\partial_k, \partial_l)\partial_j\rangle, \tag{2.22}$$

and Hess $X(\partial_r, \partial_k) \cdot$ Hess $\phi(\partial_l, \partial_j) g^{ra}\partial_a$ is dual to the one-form $A \mapsto$ Hess $X(A, \partial_k) \cdot$ Hess $\phi(\partial_l, \partial_j)$. Thus the first term in (2.21) is

$$\begin{aligned} \text{Hess } X(\partial_r, \partial_k) \cdot \text{Hess } \phi(\partial_l, \partial_j) g^{ra}\partial_a &= [\text{Hess } X(\cdot, \partial_k) \cdot \text{Hess } \phi(\partial_l, \partial_j))]^\sharp \\ &= \text{Hess } X(\cdot, \partial_k)^\sharp \cdot \text{Hess } \phi(\partial_l, \partial_j). \end{aligned} \tag{2.23}$$

Therefore, (2.21) becomes

$$\text{Hess } X(\cdot, \partial_k)^\sharp \cdot \text{Hess } \phi(\partial_l, \partial_j) - (k \leftrightarrow l) - \text{grad}(X)\langle \text{grad}(\phi), R(\partial_k, \partial_l)\partial_j \rangle. \tag{2.24}$$

The third and seventh term give the same as (2.24) with $(X \leftrightarrow \phi)$. This proves (2.16).

□

Here is the variation of the curvature in (non-normal) coordinates. In local coordinates,

Corollary 1.

$$\begin{aligned} \delta_X R^i{}_{jkl} &= \left(g^{ia}(X_{ak} + X_r \Gamma^r_{ak})(\phi_{lj} + \phi_s \Gamma^s_{lj}) - (k \leftrightarrow l) - g^{ia} X_a \phi_b R^b{}_{jkl} \right) \\ &\quad + (X \leftrightarrow \phi). \end{aligned} \tag{2.25}$$

Proof. For a one-form ω, $\omega^\sharp = g^{ia}\omega(\partial_a)\partial_i$. Thus

$$\text{Hess } X(\cdot, \partial_k)^\sharp \cdot \text{Hess } \phi(\partial_l, \partial_j) = [g^{ia}\text{Hess } X(\partial_a, \partial_k) \cdot \text{Hess } \phi(\partial_l, \partial_j)]\partial_i.$$

Also, Hess $X = \nabla dX = \nabla(X_r dx^r) = X_{rp} dx^r \otimes dx^p + X_s \Gamma^s_{rp} dx^r \otimes dx^p$, so

$$g^{ia}\text{Hess } X(\partial_a, \partial_k) \cdot \text{Hess } \phi(\partial_l, \partial_j)) = g^{ia}(X_{ak} + X_r \Gamma^r_{ak})(\phi_{lj} + \phi_s \Gamma^s_{lj}).$$

The third term on the RHS of (2.25) is in (2.22). □

2.6 The term $\delta_X\langle R, R\rangle$ in (2.12), part III.

The final term we need to compute is

$$(\text{A}) = (\delta_X R^i{}_{jkl}) R_i{}^{jkl},$$

which comes from the first term on the RHS of (2.12). Recall that we want to write

$$(A) = \langle X, Z\rangle \tag{2.26}$$

for some $\mathbb{R}^N$-valued vector Z.

By Theorem 1, (A) involves six pieces, the first of which is

$$\mathbf{AI} = \langle g^{ir}\text{Hess }(X)(\partial_r, \partial_k)\cdot\text{Hess }(\phi)(\partial_l, \partial_j)\partial_i \otimes dx^k \otimes dx^l \otimes dx^j, R^a{}_{bcd}\partial_a \otimes dx^b \otimes dx^c \otimes dx^d\rangle,$$

which equals

$$\langle\text{Hess }(X)(\partial_i, \partial_k)\cdot\text{Hess }(\phi)(\partial_l, \partial_j)dx^i \otimes dx^k \otimes dx^l \otimes dx^j, R_{abcd}dx^a \otimes dx^b \otimes dx^c \otimes dx^d\rangle, \tag{2.27}$$

which equals

$$\begin{aligned}
&\text{Hess }(X)(\partial_i, \partial_k)\cdot\text{Hess }(\phi)(\partial_l, \partial_j)g^{ia}g^{kb}g^{lc}g^{jd}R_{abcd} \\
&= \text{Hess }(X)(\partial_i, \partial_k)\cdot\text{Hess }(\phi)(\partial_l, \partial_j)R^{iklj} \\
&= \langle\text{Hess }(X)(\partial_i, \partial_k)dx^i \otimes dx^k, \text{Hess }(\phi)(\partial_l, \partial_j)R_{abcd}g^{cl}g^{dj}dx^a \otimes dx^b\rangle \qquad (2.28) \\
&= \langle\text{Hess }(X)(\partial_i, \partial_k)dx^i \otimes dx^k, \text{Hess }(\phi)(\partial_l, \partial_j)R_{ab}{}^{lj}dx^a \otimes dx^b\rangle \\
&= \langle\text{Hess }(X), \langle\text{Hess }(\phi), R\rangle_{34}\rangle,
\end{aligned}$$

where the last line is defined by the line above it.

Denoting $\langle\text{Hess }(X), \langle\text{Hess }(\phi), R\rangle_{34}\rangle$ by (AI), we have (omitting the integration over

M)

$$\begin{aligned}
\text{(AI)} &= \langle \nabla dX, \langle \text{Hess }(\phi), R\rangle_{34}\rangle \\
&= \langle X, \delta\nabla^{*}\langle \text{Hess }(\phi), R\rangle_{34}\rangle \\
&= \langle X, -\text{div}[(\nabla^{*}\langle \text{Hess }(\phi), R\rangle_{34})^{\sharp}]\rangle
\end{aligned} \tag{2.29}$$

To write the last line above in coordinates we first look at the computation of ∇^* on a general two tensor:

Letting $\omega = \omega_i dx^i$ and $\beta = \beta_{ab} dx^a \otimes dx^b$ we have

$$\begin{aligned}
\langle \nabla\omega, \beta\rangle &= \left\langle \nabla(\omega_i dx^i), \beta_{ab} dx^a \otimes dx^b \right\rangle \\
&= \int \left\langle d(\omega_i) \otimes dx^i + \omega_i \otimes \nabla dx^i, \beta_{ab} dx^a \otimes dx^b \right\rangle \text{dvol} \\
&= \int \left\langle \frac{\partial \omega_i}{\partial x_j} dx^j \otimes dx^i - \Gamma^i_{jk}\omega_i dx^j \otimes dx^k, \beta_{ab} dx^a \otimes dx^b \right\rangle \text{dvol} \\
&= \int \left\langle \text{grad}\omega_i, \beta_{ab} g^{ma} g^{ib} \partial_m \right\rangle \text{dvol} - \int \omega_i \Gamma^i_{jk} g^{ja} g^{kb} \beta_{ab} \text{dvol} \\
&= -\int \left\langle \omega, g_{ez}\text{div}(\beta_{ab} g^{ma} g^{eb} \partial_m) dx^z \right\rangle \text{dvol} - \int \omega_i \Gamma^i_{jk} g^{ja} g^{kb} \beta_{ab} \text{dvol} \\
&= -\int \langle \omega, g_{ez}\text{div}(\beta^{me}\partial_m) dx^z \rangle \text{dvol} - \int \omega_i \Gamma^i_{jk} \beta^{jk} \text{dvol} \\
&= -\int \langle \omega, g_{ez}\text{div}(\beta^{me}\partial_m) dx^z \rangle \text{dvol} - \int \left\langle \omega, \Gamma^u_{jk}\beta^{jk} g_{uq} dx^q \right\rangle \text{dvol} \\
&= \int \left\langle \omega, [-g_{ez}\text{div}(\beta^{me}\partial_m) - \Gamma^u_{jk}\beta^{jk} g_{uz}] dx^z \right\rangle \text{dvol}
\end{aligned}$$

And therefore for a general 2-tensor $\beta_{ab} dx^a \otimes dx^b$ we have

$$\nabla^*\beta = [-g_{ez}\text{div}(\beta^{me}\partial_m) - \Gamma^u_{jk}\beta^{jk} g_{uz}] dx^z$$

Returning to the computation of the last line in (2.34) we can start by writing $\langle \text{Hess }(\phi), R\rangle_{34}$

in coordinates:

$$\begin{aligned}
&\left\langle \text{Hess}(\phi)(\partial_l,\partial_j)dx^l\otimes dx^j, R_{abcd}dx^a\otimes dx^b\otimes dx^c\otimes dx^d\right\rangle_{34}\\
&= \text{Hess}(\phi)(\partial_l,\partial_j)R_{abcd}g^{lc}g^{jd}dx^a\otimes dx^b\\
&= \text{Hess}(\phi)(\partial_l,\partial_j)R_{ab}{}^{lj}dx^a\otimes dx^b
\end{aligned}$$

$$\Rightarrow \beta^{me} = \text{Hess}(\phi)(\partial_l,\partial_j)R^{melj}$$

Applying ∇^* to this 2-tensor we have:

$$\begin{aligned}
\nabla^*\langle \text{Hess}(\phi),R\rangle_{34} &= [-g_{ez}\text{div}(\text{Hess}(\phi)(\partial_l,\partial_j)R^{melj}\partial_m) - \Gamma^u_{nk}\text{Hess}(\phi)(\partial_l,\partial_j)R^{nklj}g_{uz}]dx^z\\
&= [-g_{ez}\text{div}([\phi_{lj}-\phi_r\Gamma^r_{lj}]R^{melj}\partial_m) - \Gamma^u_{nk}[\phi_{lj}-\phi_r\Gamma^r_{lj}]R^{nklj}g_{uz}]dx^z\\
&= [-g_{ea}\frac{1}{\sqrt{\text{det}g}}\partial_m([\phi_{lj}-\phi_r\Gamma^r_{lj}]R^{melj}\sqrt{\text{det}g}) -\\
\ldots\ & \Gamma^u_{nk}[\phi_{lj}-\phi_r\Gamma^r_{lj}]R^{nklj}g_{ua}]dx^a\\
&= [-g_{ea}\frac{1}{\sqrt{\text{det}g}}([\phi_{lj}-\phi_r\Gamma^r_{lj}]\partial_m(R^{melj}\sqrt{\text{det}g}))\\
&\quad -\frac{1}{\sqrt{\text{det}g}}\Big((\phi_{ljm}-\phi_{rm}\Gamma^r_{lj}-\phi_r\partial_m(\Gamma^r_{lj}))R^m{}_a{}^{lj}\sqrt{\text{det}g}\Big)\\
&\quad -\Gamma^u_{nk}[\phi_{lj}-\phi_r\Gamma^r_{lj}]R^{nklj}g_{ua}]dx^a
\end{aligned}$$

Thus

$$\begin{aligned}
[\nabla^*\langle \text{Hess}(\phi),R\rangle_{34}]^\sharp &= g^{az}[-g_{ea}\frac{1}{\sqrt{\text{det}g}}([\phi_{lj}-\phi_r\Gamma^r_{lj}]\partial_m(R^{melj}\sqrt{\text{det}g}))\\
&\quad -\frac{1}{\sqrt{\text{det}g}}\Big((\phi_{ljm}-\phi_{rm}\Gamma^r_{lj}-\phi_r\partial_m(\Gamma^r_{lj}))R^m{}_a{}^{lj}\sqrt{\text{det}g}\Big)\\
&\quad -\Gamma^u_{nk}[\phi_{lj}-\phi_r\Gamma^r_{lj}]R^{nklj}g_{ua}]\partial_z \qquad (2.30)
\end{aligned}$$

And finally

$$-\text{div}\left([\nabla^*\langle \text{Hess}(\phi),R\rangle_{34}]^\sharp\right) = \frac{1}{\sqrt{\text{det}g}}\partial_z\Big(g^{az}[g_{ea}\frac{1}{\sqrt{\text{det}g}}([\phi_{lj}-\phi_r\Gamma^r_{lj}]\partial_m(R^{melj}\sqrt{\text{det}g}))$$

$$
\begin{aligned}
& +\frac{1}{\sqrt{\det g}}\Big((\phi_{ljm} - \phi_{rm}\Gamma^r_{lj} - \phi_r\partial_m(\Gamma^r_{lj}))R^{m\ \, lj}_{\ a}\sqrt{\det g}\Big) \\
& +\Gamma^u_{nk}[\phi_{lj} - \phi_r\Gamma^r_{lj}]R^{nklj}g_{ua}]\sqrt{\det g}\Big) \\
= \;& \frac{1}{\sqrt{\det g}}\partial_z\Big([\frac{1}{\sqrt{\det g}}([\phi_{lj} - \phi_r\Gamma^r_{lj}]\partial_m(R^{mzlj}\sqrt{\det g})) \\
& \Big((\phi_{ljm} - \phi_{rm}\Gamma^r_{lj} - \phi_r\partial_m(\Gamma^r_{lj}))R^{mzlj}\Big) \qquad (2.31) \\
& +\Gamma^z_{nk}[\phi_{lj} - \phi_r\Gamma^r_{lj}]R^{nklj}]\sqrt{\det g}\Big)
\end{aligned}
$$

Summary AI: The contribution of (AI) to the RHS of (2.26) is given by

$$
\begin{aligned}
\Big\langle X, & \frac{1}{\sqrt{\det g}}\partial_z\Big(\Big[\frac{1}{\sqrt{\det g}}([\phi_{lj} - \phi_r\Gamma^r_{lj}]\partial_m(R^{mzlj}\sqrt{\det g})) \\
& + \Big((\phi_{ljm} - \phi_{rm}\Gamma^r_{lj} - \phi_r\partial_m(\Gamma^r_{lj}))R^{mzlj}\Big) \\
& + \Gamma^z_{nk}[\phi_{lj} - \phi_r\Gamma^r_{lj}]R^{nklj}\Big]\sqrt{\det g}\Big)\Big\rangle
\end{aligned}
$$

The second term in (A) is

$$
\begin{aligned}
\text{(AII)} \;=\; & \langle \text{Hess }(X)(\partial_i, \partial_l) \cdot \text{Hess }(\phi)(\partial_k, \partial_j) dx^i \otimes dx^k \otimes dx^l \otimes dx^j, \\
& R_{abcd} dx^a \otimes dx^b \otimes dx^c \otimes dx^d \rangle \qquad (2.32)
\end{aligned}
$$

where k and l have been switched from their positions in AI.

Lemma 3. *When taking the inner product of two 4-tensors, switching two indices i and j in one is equivalent to switching the same two indices, in the other.*

Proof. Let $A = A_{iklj} dx^i \otimes dx^k \otimes dx^l \otimes dx^j$ and $R = R_{abcd} dx^a \otimes dx^b \otimes dx^c \otimes dx^d$. Then we have (without loss of generality):

$$
\begin{aligned}
\langle s^{ij}(A), R\rangle \;=\; & \Big\langle s^{ij}(A_{iklj} dx^i \otimes dx^k \otimes dx^l \otimes dx^j), R_{abcd} dx^a \otimes dx^b \otimes dx^c \otimes dx^d \Big\rangle \\
=\; & \Big\langle A_{iklj} dx^j \otimes dx^k \otimes dx^l \otimes dx^i, R_{abcd} dx^a \otimes dx^b \otimes dx^c \otimes dx^d \Big\rangle \\
=\; & A_{iklj} R_{abcd} g^{ja} g^{kb} g^{lc} g^{id}
\end{aligned}
$$

$$
\begin{aligned}
&= A_{iklj}R_{dcba}g^{jd}g^{kb}g^{lc}g^{ia} \\
&= \left\langle A_{iklj}dx^i \otimes dx^k \otimes dx^l \otimes dx^j, R_{dbca}dx^a \otimes dx^b \otimes dx^c \otimes dx^d \right\rangle \\
&= \left\langle A, s^{ij}(R) \right\rangle
\end{aligned}
$$

□

and the same proof will apply to any choice of indices i and j.

Because AII is constructed from AI by switching the k and l indices in the first entry of (2.16), this is equivalent to switching the middle indices in the curvature tensor in AI's final form. However, there is a caveat to this approach. We know the curvature tensor $R = R_{abcd}dx^a \otimes dx^b \otimes dx^c \otimes dx^d$ satisfies the following properties:

$$R_{abcd} = -R_{bacd},\ R_{abcd} = -R_{abdc},\ R_{abcd} = R_{cdab} \tag{2.33}$$

However, performing one of these operations followed by a switch in R's indices is not equivalent to merely switching R's indices because:

$$R_{abcd} = -R_{bacd} \xrightarrow{23} -R_{bcad}$$

is not the same as:

$$R_{abcd} \xrightarrow{23} R_{acbd}$$

because $-R_{bcad} \neq R_{acbd}$

In our case this means that in order to apply Lemma 3 in the construction of AII, we cannot first use (2.33) on the entries in AI.

With this restriction in mind, we can construct AII simply by switching the second and third indices on the curvature tensor terms as they appear in AI. Therefore:

Summary AII: The contribution of (AII) to the RHS of (2.26) is given by

$$\Big\langle X, \frac{1}{\sqrt{\det g}}\partial_z\bigg(\bigg[\frac{1}{\sqrt{\det g}}([\phi_{lj} - \phi_r\Gamma^r_{lj}]\partial_m(R^{mlzj}\sqrt{\det g})) + \Big((\phi_{ljm} - \phi_{rm}\Gamma^r_{lj} - \phi_r\partial_m(\Gamma^r_{lj}))R^{mlzj}\Big) + \Gamma^z_{nk}[\phi_{lj} - \phi_r\Gamma^r_{lj}]R^{nlkj}\bigg]\sqrt{\det g}\bigg)\Big\rangle.$$

By Theorem 1, the third term in (A) is

$$\begin{aligned}
\text{(AIII)} &= -\langle \text{grad}(X)\langle \text{grad}(\phi), R^i{}_{jkl}\partial_i\rangle dx^j \otimes dx^k \otimes dx^l, R^a{}_{bcd}\partial_a \otimes dx^b \otimes dx^c \otimes dx^d\rangle \\
&= -\langle \text{grad}(X)\langle g^{rs}\phi_s\partial_r, R^i{}_{jkl}\partial_i\rangle dx^j \otimes dx^k \otimes dx^l, R^a{}_{bcd}\partial_a \otimes dx^b \otimes dx^c \otimes dx^d\rangle \\
&= -\langle \text{grad}(X), \phi_i R^i{}_{jkl} R^a{}_{bcd} g^{jb} g^{kc} g^{ld}\partial_a\rangle \\
&= -\langle \text{grad}(X), \phi_i R^i{}_{jkl} R^{ajkl}\partial_a\rangle \\
&= \langle X, \text{div}(\phi_i R^i{}_{jkl} R^{ajkl}\partial_a)\rangle \qquad (2.34) \\
&= \Big\langle X, \phi_i R^i{}_{jkl} R^{ajkl}\frac{1}{\sqrt{\det g}}\partial_a\sqrt{\det g} + \partial_a(\phi_i R^i{}_{jkl} R^{ajkl})\Big\rangle.
\end{aligned}$$

Summary AIII: The contribution of (AIII) to the RHS of (2.26) is given by

$$\text{(AIII)} = \Big\langle X, \phi_i R^i{}_{jkl} R^{ajkl}\frac{1}{\sqrt{\det g}}\partial_a\sqrt{\det g} + \partial_a(\phi_i R^i{}_{jkl} R^{ajkl})\Big\rangle. \qquad (2.35)$$

The term (AIV) in Theorem 1 equals (AI) with some indices switched. Specifically, we must have $(i \leftrightarrow l)$ and $(k \leftrightarrow j)$. Using Lemma 3 we know that switching these two pairs of indices is equivalent to switching both the first and third, as well as the second and fourth indices of the curvature tensor terms in AI. However:

$$R_{abcd} \xrightarrow{s^{13}} R_{cbad} \xrightarrow{s^{24}} R_{cdab} = R_{abcd}$$

where the last equality is from (2.33). Therefore AIV's contribution is identically equal to AI's and we have:

Summary AIV: The contribution of (AIV) to the RHS of (2.26) is given by

$$\Big\langle X, \frac{1}{\sqrt{\det g}}\partial_z\Big(\Big[\frac{1}{\sqrt{\det g}}([\phi_{lj}-\phi_r\Gamma^r_{lj}]\partial_m(R^{mzlj}\sqrt{\det g}))$$
$$+\Big((\phi_{ljm}-\phi_{rm}\Gamma^r_{lj}-\phi_r\partial_m(\Gamma^r_{lj}))R^{mzlj}\Big)$$
$$+\Gamma^z_{nk}[\phi_{lj}-\phi_r\Gamma^r_{lj}]R^{nklj}\Big]\sqrt{\det g}\Big)\Big\rangle.$$

The fifth term (AV) in Theorem 1 is

$$\langle \text{Hess }(\phi)(\partial_i,\partial_l)\cdot \text{Hess }(X)(\partial_k,\partial_j)dx^i\otimes dx^k\otimes dx^l\otimes dx^j, R_{abcd}dx^a\otimes dx^b\otimes dx^c\otimes dx^d\rangle$$

which amounts to a reordering of the original $iklj$ indices to $kjil$ and an analogous switch on R's indices takes $abcd$ to $bdac$. However by the (2.33) identities we know that $R_{bdac}=R_{acbd}$ which is the same reordering that we performed to construct AII from AI. Therefore, AV's contribution is identically equal to AII's and we have:

Summary AV: The contribution of (AV) to the RHS of (2.26) is given by the inner product of X with AV which equals

$$\Big\langle X, \frac{1}{\sqrt{\det g}}\partial_z\Big(\Big[\frac{1}{\sqrt{\det g}}([\phi_{lj}-\phi_r\Gamma^r_{lj}]\partial_m(R^{mlzj}\sqrt{\det g}))$$
$$+\Big((\phi_{ljm}-\phi_{rm}\Gamma^r_{lj}-\phi_r\partial_m(\Gamma^r_{lj}))R^{mlzj}\Big)$$
$$+\Gamma^z_{nk}[\phi_{lj}-\phi_r\Gamma^r_{lj}]R^{nlkj}\Big]\sqrt{\det g}\Big)\Big\rangle.$$

For the sixth term in (A), we note that the fourth line in (2.34) equals $X_a\phi_iR^i{}_{jkl}R^{ajkl}$.

The sixth term in (A) equals

$$
\begin{aligned}
&-\langle \operatorname{grad}(\phi)\langle \operatorname{grad}(X), R(\partial_k,\partial_l)\partial_j\rangle R\rangle \\
&= \langle g^{rs}\phi_r\partial_s \otimes X_i R^i{}_{jkl} dx^j \otimes dx^k \otimes dx^l, R^a{}_{bcd}\partial_a \otimes dx^b \otimes dx^c \otimes dx^d\rangle \\
&= g^{rs}\phi_r g_{sa} X_i R^i{}_{jkl} R^a{}_{bcd} g^{jb} g^{kc} g^{ld} \\
&= g^{rs}\phi_r g_{sa} X_i R^i{}_{jkl} R^{ajkl} \qquad (2.36) \\
&= \phi_i X_a R^a{}_{jkl} R^{ijkl} \\
&= \phi_i X_a R^{ajkl} R^i{}_{jkl} \\
&= \text{(AIII)}
\end{aligned}
$$

Thus the contribution from (AIII) and (AVI) are equal.

Summary AIII + AVI: The contribution of the third term (AIII) and the sixth term (AVI) to the RHS of (2.26) are equal, so by Summary AIII these two terms contribute

$$
\text{(AIII)} + \text{(AVI)} = 2\left\langle X, \phi_i R^i{}_{jkl} R^{ajkl} \frac{1}{\sqrt{\det g}} \partial_a \sqrt{\det g} + \partial_a(\phi_i R^i{}_{jkl} R^{ajkl})\right\rangle. \qquad (2.37)
$$

2.7 The gradient of $C(\phi) = \int_M |R|^2$dvol

We can now combine **Summary dvol, Summary AI – AVI, Summary CII – CV** to produce the gradient of $P_c(\phi) = \int_M |R|^2$dvol.

Proposition 1. *Let $\phi = (\phi^1, \ldots, \phi^N) : M \to \mathbb{R}^N$ be an embedding. The gradient vector field for $\int_{\phi(M)} |R|^2$ dvol at $\phi \in \operatorname{Emb}(M, \mathbb{R}^N)$ is the $\mathbb{R}^N$-valued vector field $(Z^1, \ldots, Z^N)$ on M with α component*

$$
Z^\alpha = 4\frac{1}{\sqrt{\det(g)}}\partial_r\left(\sqrt{\det(g)} R^r{}_{jkl} \phi^\alpha_i R^{ijkl}\right) \qquad \text{(CII – CV)}
$$

$$+2\frac{1}{\sqrt{\det(g)}}\partial_z\Bigg(\Bigg[\frac{1}{\sqrt{\det(g)}}([\phi^\alpha_{lj}-\phi^\alpha_r\Gamma^r_{lj}]\partial_m(R^{mzlj}\sqrt{\det(g)}))$$
$$+\Big((\phi^\alpha_{ljm}-\phi^\alpha_{rm}\Gamma^r_{lj}-\phi^\alpha_r\partial_m(\Gamma^r_{lj}))R^{mzlj}\Big)$$
$$+\Gamma^z_{nk}[\phi^\alpha_{lj}-\phi^\alpha_r\Gamma^r_{lj}]R^{nklj}\Bigg]\sqrt{\det(g)}\Bigg) \qquad (\text{AI}+\text{AIV})$$
$$-2\frac{1}{\sqrt{\det(g)}}\partial_z\Bigg(\Bigg[\frac{1}{\sqrt{\det(g)}}([\phi^\alpha_{lj}-\phi^\alpha_r\Gamma^r_{lj}]\partial_m(R^{mlzj}\sqrt{\det(g)}))$$
$$+\big((\phi^\alpha_{ljm}-\phi^\alpha_{rm}\Gamma^r_{lj}-\phi^\alpha_r\partial_m(\Gamma^r_{lj}))R^{mlzj}\big)$$
$$+\Gamma^z_{nk}[\phi^\alpha_{lj}-\phi^\alpha_r\Gamma^r_{lj}]R^{nlkj}\Bigg]\sqrt{\det(g)}\Bigg) \qquad (\text{AII}+\text{AV})$$
$$+2\phi^\alpha_i R^i{}_{jkl}R^{ajkl}\frac{1}{\sqrt{\det g}}\partial_a\sqrt{\det g}+2\partial_a(\phi^\alpha_i R^i{}_{jkl}R^{ajkl}) \qquad (\text{AIII})+(\text{AVI})$$
$$-|R|^2(\text{Tr II})^\alpha-2\langle\nabla R,R\rangle^{\sharp,\alpha}. \qquad (\text{dvol})$$

We can simplify Prop. 1 by combining (AI + AIV) and (AII + AV) using the Bianchi identity:

$$R^{mzlj}-R^{mlzj}=R^{mzlj}+R^{mljz}=-R^{mjzl}=R^{mjlz}.$$

This gives:

$$Z^\alpha=4\frac{1}{\sqrt{\det(g)}}\partial_r\left(\sqrt{\det(g)}R^r{}_{jkl}\phi^\alpha_i R^{ijkl}\right) \qquad (\text{CII}-\text{CV})$$
$$+2\frac{1}{\sqrt{\det(g)}}\partial_z\Bigg(\Bigg[\frac{1}{\sqrt{\det(g)}}([\phi^\alpha_{lj}-\phi^\alpha_r\Gamma^r_{lj}]\partial_m(R^{mjlz}\sqrt{\det(g)}))$$
$$+\left(\phi^\alpha_{ljm}-\phi^\alpha_{rm}\Gamma^r_{lj}-\phi^\alpha_r\partial_m(\Gamma^r_{lj})\right)R^{mjlz}$$
$$+\Gamma^z_{nk}[\phi^\alpha_{lj}-\phi^\alpha_r\Gamma^r_{lj}]R^{njlk}\Bigg]\sqrt{\det(g)}\Bigg) \qquad (\text{AI}+\text{AII}+\text{AIV}+\text{AV})$$
$$+2\phi^\alpha_i R^i{}_{jkl}R^{ajkl}\frac{1}{\sqrt{\det g}}\partial_a\sqrt{\det g}+2\partial_a(\phi^\alpha_i R^i{}_{jkl}R^{ajkl}) \qquad (\text{AIII})+(\text{AVI})$$
$$-|R|^2(\text{Tr II})^\alpha-2\langle\nabla R,R\rangle^{\sharp,\alpha}. \qquad (\text{dvol})$$

This can be further simplified. Let δ denote the term in the line (CII – CV) above without the coefficient 4. Similarly, let $\beta+\gamma$ denote the two terms on the line (AIII) + (AVI) with the index a replaced with r and without the coefficient 2. Since $R^r{}_{jkl}R^{ijkl} = R^i{}_{jkl}R^{rjkl}$, we get $\delta = \beta + \gamma$. Thus the two lines (CII – CV), (AIII) + (AVI) combine to give 6δ. Thus the previous equation gives:

Theorem 2. *Let $\phi = (\phi^1, \ldots, \phi^N) : M \to \mathbb{R}^N$ be an embedding. The gradient vector field for $\int_{\phi(M)} |R|^2$ dvol at $\phi \in \mathrm{Emb}(M, \mathbb{R}^N)$ is the $\mathbb{R}^N$-valued vector field $(Z^1, \ldots, Z^N)$ on M with α component*

$$
\begin{aligned}
Z^\alpha = {} & 6\frac{1}{\sqrt{\det(g)}}\partial_r\left(\sqrt{\det(g)}R^r{}_{jkl}\phi_i^\alpha R^{ijkl}\right) && \text{(CII – CV)}\\
& +2\frac{1}{\sqrt{\det(g)}}\partial_z\left(\left[\frac{1}{\sqrt{\det(g)}}([\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]\partial_m(R^{mjlz}\sqrt{\det(g)}))\right.\right.\\
& + \left(\phi^\alpha_{ljm} - \phi^\alpha_{rm}\Gamma^r_{lj} - \phi^\alpha_r\partial_m(\Gamma^r_{lj})\right)R^{mjlz}\\
& \left.\left.+\Gamma^z_{nk}[\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]R^{njlk}\right]\sqrt{\det(g)}\right) && \text{(AI + AII + AIV + AV)}\\
& -|R|^2(\mathrm{Tr\ II})^\alpha - 2\langle\nabla R, R\rangle^{\sharp,\alpha}. && \text{(dvol)}
\end{aligned}
$$

Chapter 3

Boundary Terms

The computation of the gradient of the curvature penalty term $\nabla P_c = \vec{Z}$ in Chapter 2 assumes the initial manifold M doesn't have a boundary. In the case of embedding a manifold with boundary (e.g. embedding a solid ball into $\mathbb{R}^N$), the following gives the computation of all missing boundary terms from the second chapter. Following the structure of Theorem 2 (end of Chapter 2) we need to compute boundary terms for terms **dvol**, **CII -CV**, and **AI-AVI**. Note that although terms **AIII** and **AVI** were combined in the final form of Theorem 2, these boundary terms are treated explicitly in this chapter.

1. dvol Boundary Term:

(original boundaryless expression given in (2.8))

Coordinate-free expression:

From Stokes' Theorem we have

$$\int_M d(|R|^2 \wedge *\omega) = \int_{\partial M} |R|^2 \wedge *\omega = \int_M d|R|^2 \wedge *\omega + \int_M |R|^2 \wedge d*\omega$$

$$\Rightarrow \int_M |R|^2 \wedge d*\omega = \int_{\partial M} |R|^2 \wedge *\omega - \int_M d|R|^2 \wedge *\omega$$

$$\Rightarrow \int_{\partial M} |R|^2 \wedge *\omega$$

is the final boundary term.

In coordinates:

We need to compute $*\omega$ where ω is a one-form on $\phi(M)$ that acts as follows:

$$\omega(Y) = \langle X, Y\rangle_{\mathbb{R}^N} = X \cdot Y_{\mathbb{R}^N} = (\text{Proj}_{T_{\phi(m)}\phi(M)} X) \cdot Y_{\mathbb{R}^N} = \langle \text{Proj} X, Y\rangle_{g_\phi}$$

where $X \epsilon T_{\phi(m)}\mathbb{R}^N$ and $Y \epsilon T_{\phi(m)}\phi(M)$ and 'Proj' is the $\mathbb{R}^N$ orthogonal projection of $T_{\phi(m)}\mathbb{R}^N$ to $T_{\phi(m)}\phi(M)$.

We will denote X's projection onto the tangent space of $\phi(M)$ by $\tilde{X}$. Because we are using the induced metric on $T_{\phi(m)}\phi(M)$, the inner product of vectors is the same as the Euclidean dot product. In the following computations it should be noted that in some cases $\tilde{X}$'s indices range from 1 to k (when treated in $T_{\phi(m)}\phi(M)$) and in others from 1 to N (when treated in $T_{\phi(m)}\mathbb{R}^N$).

In coordinates we have:

$$\omega = g_{ij}\tilde{X}^i dx^j$$

and

$$*dx^j = \sqrt{\det g_M}(-1)^{j-1} dx^1 \wedge \cdots \wedge \hat{dx^j} \wedge \cdots \wedge dx^k$$

$$\Rightarrow *\omega = *(g_{ij}\tilde{X}^i dx^j) = g_{ij}\tilde{X}^i(*dx^j) = g_{ij}\tilde{X}^i(\sqrt{\det g_M}(-1)^{j-1} dx^1 \wedge \cdots \wedge \hat{dx^j} \wedge \cdots \wedge dx^k)$$

On ∂M we have $dx^k = 0$ so pulling $*\omega$ onto the boundary gives:

$$\begin{aligned} i^*(*\omega) &= i^*(g_{ij}\tilde{X}^i\sqrt{\det g_M}(-1)^{j-1} dx^1 \wedge \cdots \wedge \hat{dx^j} \wedge \cdots \wedge dx^k) \\ &= g_{ik}\tilde{X}^i(\sqrt{\det g_M}(-1)^{k-1} dx^1 \wedge \cdots \wedge dx^{k-1}) \end{aligned} \tag{3.1}$$

Substituting (3.1) into the original boundary term $\int_{\partial M} |R|^2 \wedge *\omega$ we have:

$$\int_{\partial M} |R|^2 \wedge (g_{ik}\tilde{X}^i\sqrt{\det g_M}(-1)^{k-1} dx^1 \wedge \cdots \wedge dx^{k-1})$$

$$= \int_{\partial M} |R|^2 g_{ij} g^{jr} g_{rk} \tilde{X}^i \sqrt{\det g_M} (-1)^{k-1} dx^1 \wedge \cdots \wedge dx^{k-1}$$

$$= \int_{\partial M} |R|^2 g_{ij} g^{jr} g_{rk} \tilde{X}^i \frac{\sqrt{\det g_M}}{\sqrt{\det g_{\partial M}}} (-1)^{k-1} \mathrm{dvol}_{g_{\partial M}}$$

$$= \int_{\partial M} \left\langle \tilde{X}, |R|^2 \frac{\sqrt{\det g_M}}{\sqrt{\det g_{\partial M}}} (-1)^{k-1} g^{lr} g_{rk} \partial_l \right\rangle_{g_\phi} \mathrm{dvol}_{\partial M}$$

$$= \int_{\partial M} \left\langle \tilde{X}, |R|^2 \frac{\sqrt{\det g_M}}{\sqrt{\det g_{\partial M}}} (-1)^{k-1} \partial_k \right\rangle_{g_\phi} \mathrm{dvol}_{\partial M} = \int_{\partial M} \tilde{X} \cdot (|R|^2 \frac{\sqrt{\det g_M}}{\sqrt{\det g_{\partial M}}} (-1)^{k-1} \partial_k) \mathrm{dvol}_{\partial M}$$

$$= \int_{\partial M} X \cdot (|R|^2 \frac{\sqrt{\det g_M}}{\sqrt{\det g_{\partial M}}} (-1)^{k-1} \partial_k) \mathrm{dvol}_{\partial M}$$

But seeing that $(-1)^{k-1}\partial_k$ is the inward pointing normal vector on ∂M which we denote as ν we have:

$$\textbf{Final boundary term:} \int_{\partial M} X \cdot (|R|^2 \frac{\sqrt{\det g_M}}{\sqrt{\det g_{\partial M}}} \nu) \mathrm{dvol}_{\partial M}$$

2. CII boundary term:

Coordinate-free Expression:

Note, we are treating the term $2\langle dX, \langle R, \langle d\phi, R\rangle_1 \rangle_{234}\rangle$ on page 12, where we know:

$$\langle dX, \langle R, \langle d\phi, R\rangle_1 \rangle_{234}\rangle = \langle \mathrm{grad} X, \langle R, \langle d\phi, R\rangle_1 \rangle^\sharp_{234}\rangle$$

Furthermore, by the divergence theorem on Riemannian manifolds we have (for general X, T):

$$\int_M \sum_\alpha^N \langle \mathrm{grad} X^\alpha, T^\alpha \rangle \, \mathrm{dvol}_g = -\sum_\alpha \int_M X^\alpha \cdot \mathrm{div} T^\alpha \mathrm{dvol}_g + \sum_\alpha \int_{\partial M} X^\alpha \, \langle T^\alpha, \nu \rangle_g \, \mathrm{dvol}_{\partial M}$$

where ν is the inward pointing normal vector on ∂M. By the fact that in our case

$$T^\alpha = \langle R, \langle d\phi^\alpha, R\rangle_1\rangle^\sharp_{234} = \langle \mathrm{grad}\phi^\alpha, \langle R, R\rangle_{234}\rangle^\sharp$$

substituting this into the boundary expression above we get the coordinate-free boundary term:

$$2\sum_\alpha \int_{\partial M} X^\alpha \left\langle \langle \mathrm{grad}\phi^\alpha, \langle R, R\rangle_{234}\rangle^\sharp, \nu \right\rangle \mathrm{dvol}_g$$

where

$$\langle R, R\rangle_{234} = R^r{}_{jkl} R^a_{bcd} g^{jb} g^{kc} g^{ld} \partial_r \otimes \partial_a$$

In coordinates:

Expanding the boundary term from above gives:

$$\sum_\alpha \int_{\partial M} X^\alpha \langle T^\alpha, \nu\rangle_g \mathrm{dvol}_{g_{\partial M}} = \sum_\alpha \int_{\partial M} X^\alpha \left\langle T^{i,\alpha}\partial_i, \nu^j \partial_j \right\rangle_g \mathrm{dvol}_{\partial M}$$

$$= \int_{\partial M} \sum_\alpha X^\alpha g_{ij} T^{i,\alpha} \nu^j \mathrm{dvol}_{\partial M} \qquad (3.2)$$

where, using the fact that $T^\alpha = T^{i,\alpha}\partial_i = g^{ri}T^\alpha_r \partial_i$, in the last term above we have:

$$g_{ij} T^{i,\alpha} \nu^j = g_{ij} g^{ri} T^\alpha_r \nu^j = T^\alpha_j \nu^j$$

$$\Rightarrow (3.2) = \int_{\partial M} X \cdot (g_{ij} T^{i,l} \nu^j \partial_l)_{(N)} \mathrm{dvol}_{\partial M} = \int_{\partial M} X \cdot T^l_j \nu^j \partial_l \mathrm{dvol}_{\partial M} \qquad (3.3)$$

Letting $S^a_j = R_{jqkl}R^a_{bcd}g^{qb}g^{kc}g^{ld}$ we have

$$T^\alpha_j \nu^j = S^a_j \phi^\alpha_a \nu^j = R_{jqkl}R^a_{bcd}g^{qb}g^{kc}g^{ld}\phi^\alpha_a \nu^j$$

$$\Rightarrow (3.3) = \int_{\partial M} X \cdot (R_{jqkl}R^a_{bcd}g^{qb}g^{kc}g^{ld}\phi_a \nu^j)\mathrm{dvol}_{\partial M}$$

where $\phi_a = (\phi^1_a, \cdots, \phi^N_a)$

Final CII boundary term: $2\int_{\partial M} X \cdot (R_{jqkl}R^a_{bcd}g^{qb}g^{kc}g^{ld}\phi_a \nu^j)\mathrm{dvol}_{\partial M}$

3. CIII boundary term:

Working from the top of page 14 in Chapter 2:

In coordinates:

$$\left\langle 2R^i_{jkl}\partial_i \otimes (\nabla^h_X dx^j) \otimes dx^k \otimes dx^l, R^a_{bcd}\partial_a \otimes dx^b \otimes dx^c \otimes dx^d \right\rangle_g$$

$$= R^i_{jkl}g_{ia}(-g^{sj}(X_r \cdot \phi_s + X_s \cdot \phi_r))g^{rb}R^a_{bcd}g^{kc}g^{dl}$$

$$= -\sum_{\alpha=1}^{N} R^i_{jkl}R^a_{bcd}g_{ia}g^{sj}X^\alpha_r \phi^\alpha_s g^{rb}g^{kc}g^{ld} - \sum_{\alpha=1}^{N} R^i_{jkl}R^a_{bcd}g_{ia}g^{sj}X^\alpha_s \phi^\alpha_r g^{rb}g^{kc}g^{ld} \quad (3.4)$$

Just looking at the first sum in the line above and letting

$$S^s_b = R_{ajkl}R^a_{bcd}g^{sj}g^{kc}g^{ld}$$

$$T^\alpha_b = S^s_b \phi^\alpha_s$$

and recalling that the k-component vector T^α is defined as:

$$T^\alpha = T^{r,\alpha}\partial_r = g^{br}T^\alpha_b \partial_r$$

we have:

$$-\sum_{\alpha=1}^{N} R^i_{jkl} R^a_{bcd} g_{ia} g^{sj} X^\alpha_r \phi^\alpha_s g^{rb} g^{kc} g^{ld} = -\sum_{\alpha=1}^{N} R_{ajkl} R^a_{bcd} g^{sj} X^\alpha_r \phi^\alpha_s g^{rb} g^{kc} g^{ld} =$$

$$-\sum_{\alpha=1}^{N} X^\alpha_r g^{rb} T^\alpha_b = -\sum_{\alpha=1}^{N} X^\alpha_r T^{r,\alpha} = -\sum_{\alpha=1}^{N} g_{pi} g^{ir} X^\alpha_r T^{p,\alpha}$$

$$= -\sum_{\alpha=1}^{N} \left\langle g^{ir} X^\alpha_r \partial_i, T^{p,\alpha} \partial_p \right\rangle = -\sum_{\alpha=1}^{N} \langle \mathrm{grad} X^\alpha, T^\alpha \rangle$$

Integrating the last term gives:

$$-\int_M \sum_{\alpha=1}^{N} \langle \mathrm{grad} X^\alpha, T^\alpha \rangle \, \mathrm{dvol}_g = \sum_\alpha \int_M X^\alpha \cdot \mathrm{div} T^\alpha \mathrm{dvol}_g - \sum_\alpha \int_{\partial M} X^\alpha \, \langle T^\alpha, \nu \rangle_g \, \mathrm{dvol}_{g_{\partial M}}$$

where ν is the inward pointing normal vector on ∂M. Expanding the boundary term gives:

$$-\sum_\alpha \int_{\partial M} X^\alpha \, \langle T^\alpha, \nu \rangle_g \, \mathrm{dvol}_{g_{\partial M}} = -\sum_\alpha \int_{\partial M} X^\alpha \left\langle T^{p,\alpha} \partial_p, \nu^i \partial_i \right\rangle_g \mathrm{dvol}_{\partial M}$$

$$= -\sum_\alpha \int_{\partial M} X^\alpha g_{pi} T^{p,\alpha} \nu^i \mathrm{dvol}_{\partial M} \tag{3.5}$$

Recall that $g_{pi} T^{p,\alpha} \nu^i = g_{pi} g^{bp} T^\alpha_b \nu^i = T^\alpha_i \nu^i$ so we have:

$$\Rightarrow (3.5) = -\int_{\partial M} X \cdot g_{pi} T^{p,\alpha} \nu^i \partial_\alpha \mathrm{dvol}_{\partial M} = -\int_{\partial M} X \cdot (T^\alpha_i \nu^i \partial_\alpha) \mathrm{dvol}_{\partial M} \tag{3.6}$$

and because $T^\alpha_i \nu^i = S^s_i \phi^\alpha_s \nu^i = R_{ajkl} R^a_{icd} g^{sj} g^{kc} g^{ld} \phi^\alpha_s \nu^i$

$$\Rightarrow (3.6) = -\int_{\partial M} X \cdot R_{ajkl} R^a_{icd} g^{sj} g^{kc} g^{ld} \phi_s \ni^i \, \mathrm{dvol}_{\partial_M}$$

where $\phi_s = (\phi^1_s, \cdots, \phi^N_s)$.

Final CIII boundary term: $-2 \int_{\partial M} X \cdot R_{ajkl} R^a_{bcd} g^{sj} g^{kc} g^{ld} \phi_s \nu^b \mathrm{dvol}_{\partial_M}$

(where factor of 2 is from two terms in (3.4))

Coordinate-free Expression: $-2\sum_\alpha \int_{\partial M} X^\alpha \left\langle \langle \mathrm{grad}\phi^\alpha, \langle R, R\rangle\rangle^\sharp, \nu\right\rangle \mathrm{dvol}_g$

Using the definition of T^α_b above we can define the one-form $T^\alpha_b dx^b$ on $\phi(M)$ for which T^α is the corresponding vector field. If we define

$$\langle R, R\rangle = S^s_b \partial_s \otimes dx^b = R_{ajkl} R^a_{bcd} g^{sj} g^{kc} g^{ld} \partial_s \otimes dx^b$$

then

$$T^\alpha_b dx^b = R_{ajkl} R^a_{bcd} g^{sj} g^{kc} g^{ld} \phi^\alpha_s dx^b = (\langle R, R\rangle)^s_b \phi^\alpha_s dx^b$$

$$= \left\langle g^{ij} \phi^\alpha_i \partial_j, R_{ajkl} R^a_{bcd} g^{sj} g^{kc} g^{ld} \partial_s \right\rangle dx^b = \langle \mathrm{grad}\phi^\alpha, \langle R, R\rangle\rangle_b \, dx^b$$

$$= \langle \mathrm{grad}\phi^\alpha, \langle R, R\rangle\rangle$$

$$\Rightarrow T^\alpha = \langle \mathrm{grad}\phi^\alpha, \langle R, R\rangle\rangle^\sharp$$

and substituting into the boundary term gives:

$$-2\sum_\alpha \int_{\partial M} X^\alpha \langle T^\alpha, \nu\rangle_g \, \mathrm{dvol}_{g_{\partial M}} = -2\sum_\alpha \int_{\partial M} X^\alpha \left\langle \langle \mathrm{grad}\phi^\alpha, \langle R, R\rangle_b\rangle^\sharp, \nu\right\rangle \mathrm{dvol}_g$$

4. CIV boundary term:

In coordinates: A similar computation to term **CIII** gives:

$-2\int_{\partial M} X \cdot R_{ajkl} R^a_{bcd} g^{jb} g^{ld} g^{sk} \phi_s \nu^c \mathrm{dvol}_{\partial_M}$

Coordinate-free Expression: $-2\sum_\alpha \int_{\partial M} X^\alpha \left\langle \langle \mathrm{grad}\phi^\alpha, \langle R, R\rangle\rangle^\sharp, \nu\right\rangle \mathrm{dvol}_g$

with

$$\langle R,R\rangle = R_{ajkl}R^a_{bcd}g^{jb}g^{ld}g^{sk}\partial_s \otimes dx^c$$

5. CV boundary term:

In coordinates: A similar computation to term **CIII** gives:

$-2\int_{\partial M} X \cdot R_{ajkl}R^a_{bcd}g^{jb}g^{kc}g^{sl}\phi_s\nu^d \mathrm{dvol}_{\partial_M}$

Coordinate-free Expression: $-2\sum_\alpha \int_{\partial M} X^\alpha \left\langle \langle \mathrm{grad}\phi^\alpha, \langle R,R\rangle\rangle^\sharp, \nu \right\rangle \mathrm{dvol}_g$

with

$$\langle R,R\rangle = R_{ajkl}R^a_{bcd}g^{jb}g^{kc}g^{sl}\partial_s \otimes dx^d$$

Final Contribution of CII-CV Boundary Terms

Note: CIV and CV boundary terms are equal by the following computation:

It suffices to show that $R_{ajkl}R^a{}_{bcd}g^{jb}g^{ld}g^{sk}\nu^c = R_{ajkl}R^a{}_{bcd}g^{jb}g^{kc}g^{sl}\nu^d$ from the integrands of CIV and CV respectively:

$$R_{ajkl}R^a{}_{bcd}g^{jb}g^{ld}g^{sk}\nu^c = R_{aj}{}^s{}_l R^{aj}{}_c{}^l\nu^c = R_{ajl}{}^s R^{ajl}{}_c\nu^c \overset{c\to d}{=} R_{ajl}{}^s R^{ajl}{}_d\nu^d$$

$$\overset{l\to k}{=} R_{ajk}{}^s R^{ajk}{}_d\nu^d = R_{ajkl}R^a{}_{bcd}g^{jb}g^{kc}g^{sl}\nu^d$$

as required.

Similar computations show CII = CIII = CIV = CV (in boundary terms disregarding coefficients for the moment). The coefficient on the CII boundary term is 2 but the coefficient on CIII- CV's boundary terms is -2, giving a final contribution from the CII - CV boundary terms of:

$$-4\int_{\partial M} X\cdot(R_{jqkl}R^a{}_{bcd}g^{qb}g^{kc}g^{ld}\nu^j(\phi_a))\mathrm{dvol}_{\partial M}$$

$$= -4\left\langle X, R_{jqkl}R^a{}_{bcd}g^{qb}g^{kc}g^{ld}\nu^j(\phi_a)\right\rangle$$

where the above expression is taken from the CII boundary term.

Therefore the αth component of the C terms' gradient vector field is:

$$Y^\alpha(C) = -4R_{jqkl}R^a{}_{bcd}g^{qb}g^{kc}g^{ld}\phi_a^\alpha\nu^j$$

6. AI boundary terms:

In (2.29) we had the equality:

$$\langle\nabla dX, \langle\mathrm{Hess}(\phi), R\rangle_{34}\rangle = \langle X, \delta\nabla^*\langle\mathrm{Hess}(\phi), R\rangle_{34}\rangle$$

Two boundary terms emerge in this equality, the first of which comes from applying the ∇^* adjoint operation and the second from applying δ.

6a. First AI boundary term:

We will first compute the general formula for the ∇^* operator on a 2-tensor and then apply it to our case.

In coordinates:

In general, for a co-vector γ and 2-tensor β we have $\langle\nabla\gamma,\beta\rangle = \langle\gamma,\nabla^*\beta\rangle$. With $\gamma = \gamma_i dx^i$ and $\beta = \beta_{ab}dx^a\otimes dx^b$ we have:

$$\langle\nabla\gamma,\beta\rangle = \int_M \left\langle\nabla(\gamma_i dx^i), \beta_{ab}dx^a\otimes dx^b\right\rangle\mathrm{dvol}$$

$$= \int_M \left\langle d(\gamma_i)\otimes dx^i + \gamma_i\nabla(dx^i), \beta_{ab}dx^a\otimes dx^b\right\rangle\mathrm{dvol}$$

where we will only get a boundary term taking derivatives off of γ_i in the first term. Expanding the first term in the sum above gives:

$$\int_M \left\langle d(\gamma_i)\otimes dx^i, \beta_{ab}dx^a \otimes dx^b \right\rangle \text{dvol} = \int_M \left\langle \partial_q(\gamma_i)dx^q \otimes dx^i, \beta_{ab}dx^a \otimes dx^b \right\rangle \text{dvol}$$

$$= \int_M \partial_q(\gamma_i)g^{qa}g^{ib}\beta_{ab}\text{dvol} = \int_M \left\langle \text{grad}(\gamma_i), \beta_{ab}g^{ma}g^{ib}\partial_m \right\rangle \text{dvol}$$

$$= -\int_M \gamma_i \text{div}(\beta_{ab}g^{ma}g^{ib}\partial_m)\text{dvol} + \int_{\partial M} \gamma_i \left\langle \beta_{ab}g^{ma}g^{ib}\partial_m, \nu \right\rangle \text{dvol}_{\partial M}$$

$$= -\int_M \left\langle \gamma_i dx^i, g_{ez}\text{div}(\beta_{ab}g^{ma}g^{eb}\partial_m)dx^z \right\rangle \text{dvol} + \int_{\partial M} \gamma_i \left\langle \beta_{ab}g^{ma}g^{ib}\partial_m, \nu \right\rangle \text{dvol}_{\partial M}$$

$$= -\int_M \left\langle \gamma, g_{ez}\text{div}(\beta_{ab}g^{ma}g^{eb}\partial_m)dx^z \right\rangle \text{dvol} + \int_{\partial M} \gamma_i \left\langle \beta_{ab}g^{ma}g^{ib}\partial_m, \nu \right\rangle \text{dvol}_{\partial M}$$

giving us a final boundary term in coordinates for general co-vector γ and 2-tensor β by the second term in the line above.

Adapting the general formula above for ∇^* to this case we have that γ is dX and β is $\langle \text{Hess}(\phi), R\rangle_{34} = \text{Hess}(\phi)(\partial_l, \partial_j){R_{ab}}^{lj}dx^a \otimes dx^b \Rightarrow \beta_{ab} = \text{Hess}(\phi)(\partial_l, \partial_j){R_{ab}}^{lj} = [\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}]{R_{ab}}^{lj}$. The boundary term from above becomes:

$$\int_{\partial M} \gamma_i \left\langle \beta_{ab}g^{ma}g^{ib}\partial_m, \nu \right\rangle \text{dvol}_{\partial M} = \sum_\alpha \int_{\partial M} X^\alpha_i \left\langle [\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]{R_{ab}}^{lj}g^{ma}g^{ib}\partial_m, \nu \right\rangle \text{dvol}_{\partial M}$$

$$= \sum_\alpha \int_{\partial M} \left\langle \text{grad}X^\alpha, \left\langle [\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]{R_{ab}}^{lj}g^{ma}g^{pb}\partial_m, \nu \right\rangle \partial_p \right\rangle \text{dvol}_{\partial M}$$

$$= -\sum_\alpha \int_{\partial M} X^\alpha \text{div}(\left\langle [\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]{R_{ab}}^{lj}g^{ma}g^{pb}\partial_m, \nu \right\rangle \partial_p)\text{dvol}_{\partial M}$$

$$= \int_{\partial M} X \cdot Y \text{dvol}_{\partial M}$$

where $Y^\alpha = -\text{div}(\left\langle [\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]{R_{ab}}^{lj}g^{ma}g^{pb}\partial_m, \nu \right\rangle \partial_p)$

6b. 2nd AI boundary term:

The second AI boundary term also comes from the (2.29) equality:

$$\langle \nabla dX, \langle \mathrm{Hess}(\phi), R\rangle_{34}\rangle = \langle X, \delta\nabla^* \langle \mathrm{Hess}(\phi), R\rangle_{34}\rangle$$

this time from the application of the δ adjoint operator.

In general for a function f we have:

$$\langle df, \omega\rangle = \int_M \langle df, \omega\rangle\, \mathrm{dvol}_g = \int_M \left\langle \triangledown f, \alpha^{-1}(\omega)\right\rangle_g \mathrm{dvol}_g$$

$$= -\int_M f \cdot\ \mathrm{div}(\alpha^{-1}(\omega))\mathrm{dvol}_g + \int_{\partial M} f \left\langle \alpha^{-1}(\omega), \nu\right\rangle \mathrm{dvol}_{\partial M}$$

$$= \int_M f\delta(\omega)\mathrm{dvol}_g + \int_{\partial M} f \left\langle \alpha^{-1}(\omega), \nu\right\rangle \mathrm{dvol}_{\partial M}$$

$$= \langle f, \delta(\omega)\rangle + \text{boundary term}$$

In our case f is the αth component of X and ω is the αth component of $\nabla^* \langle \mathrm{Hess}(\phi), R\rangle_{34}$. Substituting these into the boundary term and summing over α gives the coordinate-free expression above.

In coordinates:

$$\int_{\partial M} f \left\langle \alpha^{-1}(\omega), \nu\right\rangle \mathrm{dvol}_{\partial M} = \sum_\alpha \int_{\partial M} X^\alpha \left\langle \alpha^{-1}(\nabla^* \langle \mathrm{Hess}(\phi), R\rangle_{34}), \nu\right\rangle^\alpha \mathrm{dvol}_{\partial M}$$

$$= \sum_\alpha \int_{\partial M} X^\alpha \left\langle [\nabla^* \langle \mathrm{Hess}(\phi), R\rangle_{34}]^\sharp, \nu\right\rangle^\alpha \mathrm{dvol}_{\partial M} = \int_{\partial M} X \cdot Y$$

where

$$Y^\alpha = \left\langle [\nabla^* \langle \mathrm{Hess}(\phi), R\rangle_{34}]^\sharp, \nu\right\rangle$$

$$= \left\langle g^{az}\left[- g_{ea}\frac{1}{\sqrt{\det g}}([\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]\partial_m(R^{melj}\sqrt{\det g}))\right.\right.$$

$$\left.\left. -\frac{1}{\sqrt{\det g}}\left((\phi^\alpha_{ljm} - \phi^\alpha_r m\Gamma^r_{lj} - \phi^\alpha_r\partial_m(\Gamma^r_{lj}))R^{m\ \ lj}_{\ a}\sqrt{\det g}\right) - \Gamma^u_{nk}[\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]R^{nklj}g_{ua}\right]\partial_z, \nu\right\rangle_g$$

7a. 1st AII boundary term

In coordinates:

By the argument given in Lemma 3 (Chapter 2) we know that AII's boundary terms can be constructed from AI's by switching the middle two indices on the curvature tensor R terms. Therefore the contribution of the first AII boundary term is:

$$-\int_{\partial M} X \cdot Y \mathrm{dvol}_{\partial M}$$

where $Y^\alpha = -\mathrm{div}(\left\langle[\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]R_{a\ b}^{\ l\ j}g^{ma}g^{pb}\partial_m, \nu\right\rangle\partial_p)$ and the minus sign is from the AII term being subtracted in Theorem 1 (Chapter 2), flipping the sign of the boundary component.

7b. 2nd AII boundary term:

Similarly the second AII boundary term is constructed by switching the curvature R term indices on the second AI boundary term. The contribution of this term is:

$$-\int_{\partial M} X \cdot Y$$

where

$$Y^\alpha = \left\langle [\nabla^* \left\langle \mathrm{Hess}(\phi), R\right\rangle_{34}]^\sharp, \nu\right\rangle$$

$$= \left\langle g^{az}\left[- g_{ea}\frac{1}{\sqrt{\det g}}([\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]\partial_m(R^{mlej}\sqrt{\det g}))\right.\right.$$

$$-\frac{1}{\sqrt{\det g}}\left((\phi^\alpha_{ljm} - \phi^\alpha_r m\Gamma^r_{lj} - \phi^\alpha_r \partial_m(\Gamma^r_{lj}))R^{ml}{}_a{}^j\sqrt{\det g}\right) - \Gamma^u_{nk}[\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]R^{nlkj}g_{ua}\Big]\partial_z, \nu\Big\rangle_g$$

and we again have the sign flipped on the boundary component.

8. AIII and AVI boundary term:

In the computation of AIII in (2.39) (Chapter 2), a boundary term emerges in the equality:

$$-\left\langle \text{grad}(X), \phi_i R^i{}_{jkl}R^{ajkl}\partial_a\right\rangle = \left\langle X, \text{div}(\phi_i R^i{}_{jkl}R^{ajkl}\partial_a)\right\rangle$$

In particular we have:

$$-\left\langle \text{grad}(X), \phi_i R^i{}_{jkl}R^{ajkl}\partial_a\right\rangle = -\int_M \left\langle \text{grad}(X), \phi_i R^i{}_{jkl}R^{ajkl}\partial_a\right\rangle \text{dvol} =$$

$$-\left(-\sum_\alpha \int_M X^\alpha \text{div}(\phi^\alpha_i R^i{}_{jkl}R^{ajkl}\partial_a)\text{dvol} + \sum_\alpha \int_{\partial M} X^\alpha \left\langle \phi^\alpha_i R^i{}_{jkl}R^{ajkl}\partial_a, \nu\right\rangle \text{dvol}_{\partial M}\right)$$

giving a boundary term contribution from AIII and AVI of:

$$-2\int X\cdot Y \text{dvol}_{\partial M}$$

where $Y^\alpha = \left\langle \phi^\alpha_i R^i{}_{jkl}R^{ajkl}\partial_a, \nu\right\rangle$

9a. 1st AIV boundary term:

By the argument given after (2.40), AIV's first boundary term is the same as AI's first boundary term:

$$\int_{\partial M} X\cdot Y \text{dvol}_{\partial M}$$

where $Y^\alpha = -\text{div}\left(\left\langle [\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]R_{ab}{}^{lj}g^{ma}g^{pb}\partial_m, \nu\right\rangle \partial_p\right)$

9b. 2nd AIV boundary term:

Similarly the 2nd AIV boundary term is the same as the second AI boundary term:

$$\int_{\partial M} X \cdot Y$$

where

$$Y^\alpha = \left\langle [\nabla^* \langle \mathrm{Hess}(\phi), R\rangle_{34}]^\sharp, \nu \right\rangle$$

$$= \Bigg\langle g^{az}\Bigg[- g_{ea}\frac{1}{\sqrt{\det g}}([\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}]\partial_m(R^{melj}\sqrt{\det g}))$$

$$-\frac{1}{\sqrt{\det g}}\Big((\phi^\alpha_{ljm} - \phi^\alpha_r m\Gamma^r_{lj} - \phi^\alpha_r \partial_m(\Gamma^r_{lj}))R^{m\ \ lj}_{\ a}\sqrt{\det g}\Big) - \Gamma^u_{nk}[\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]R^{nklj}g_{ua}\Bigg]\partial_z, \nu\Bigg\rangle$$

10a. 1st AV boundary term:

The first AV boundary term is the same as the first AII boundary term (see argument given after (2.40)):

$$-\int_{\partial M} X \cdot Y \mathrm{dvol}_{\partial M}$$

where $Y^\alpha = -\mathrm{div}(\left\langle [\phi^\alpha_{lj} - \phi^\alpha_r\Gamma^r_{lj}]R_{a\ b}^{\ l\ j}g^{ma}g^{pb}\partial_m, \nu\right\rangle \partial_p)$

10b. 2nd AV boundary term:

Similarly, the second AV boundary term is the same as the second AII boundary term:

$$-\int_{\partial M} X \cdot Y$$

where

$$Y^\alpha = \left\langle [\nabla^* \langle \mathrm{Hess}(\phi), R\rangle_{34}]^\sharp, \nu \right\rangle$$

$$= \left\langle g^{az} \left[- g_{ea} \frac{1}{\sqrt{\det g}} ([\phi^{\alpha}_{lj} - \phi^{\alpha}_{r} \Gamma^{r}_{lj}] \partial_m (R^{mlej} \sqrt{\det g})) \right. \right.$$

$$\left. \left. - \frac{1}{\sqrt{\det g}} \left((\phi^{\alpha}_{ljm} - \phi^{\alpha}_{r} m \Gamma^{r}_{lj} - \phi^{\alpha}_{r} \partial_m (\Gamma^{r}_{lj})) R^{ml\ \ j}_{\ \ \ a} \sqrt{\det g} \right) - \Gamma^{u}_{nk} [\phi^{\alpha}_{lj} - \phi^{\alpha}_{r} \Gamma^{r}_{lj}] R^{nlkj} g_{ua} \right] \partial_z, \nu \right\rangle$$

Final Contribution of the A Boundary Terms

The α-th component of the A terms' boundary gradient vector field is:

$$Y^\alpha(A) = -2\text{div}(\Big\langle [\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] R_{ab}{}^{lj} g^{ma} g^{pb} \partial_m, \nu \Big\rangle \, \partial_p)$$

$$+2\Big\langle g^{az} \Big[-g_{ea} \frac{1}{\sqrt{\text{det}g}} ([\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] \partial_m (R^{melj} \sqrt{\text{det}g}))$$

$$-\frac{1}{\sqrt{\text{det}g}} \Big((\phi^\alpha_{ljm} - \phi^\alpha_r m \Gamma^r_{lj} - \phi^\alpha_r \partial_m (\Gamma^r_{lj})) R^m{}_a{}^{lj} \sqrt{\text{det}g} \Big) - \Gamma^u_{nk} [\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] R^{nklj} g_{ua} \Big] \partial_z, \nu \Big\rangle$$

$$+2\text{div}(\Big\langle [\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] R_a{}^l{}_b{}^j g^{ma} g^{pb} \partial_m, \nu \Big\rangle \, \partial_p)$$

$$-2\Big\langle g^{az} \Big[-g_{ea} \frac{1}{\sqrt{\text{det}g}} ([\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] \partial_m (R^{mlej} \sqrt{\text{det}g}))$$

$$-\frac{1}{\sqrt{\text{det}g}} \Big((\phi^\alpha_{ljm} - \phi^\alpha_r m \Gamma^r_{lj} - \phi^\alpha_r \partial_m (\Gamma^r_{lj})) R^{ml}{}_a{}^j \sqrt{\text{det}g} \Big) - \Gamma^u_{nk} [\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] R^{nlkj} g_{ua} \Big] \partial_z, \nu \Big\rangle$$

$$-2 \left\langle \phi^\alpha_i R^i{}_{jkl} R^{ajkl} \partial_a, \nu \right\rangle$$

Proposition 2. *Let $\phi = (\phi^1, \ldots, \phi^N) : M \to \mathbb{R}^N$ be an embedding where M is closed with boundary. The gradient vector field for $\int_{\phi(M)} |R|^2 \, dvol$ at $\phi \in \text{Emb}(M, \mathbb{R}^N)$ has the $\mathbb{R}^N$-valued vector field $(Z^1, \ldots, Z^N)$ given in Theorem 2 and the $\mathbb{R}^N$ valued boundary terms $(Y^1, \ldots, Y^N)$ with α component:*

$$Y^\alpha = (|R|^2 \frac{\sqrt{\det g_M}}{\sqrt{\det g_{\partial M}}} \nu)^\alpha - 4 R_{jqkl} R^a{}_{bcd} g^{qb} g^{kc} g^{ld} \phi^\alpha_a \nu^j$$

$$-2\,div(\Big\langle [\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] R_{ab}{}^{lj} g^{ma} g^{pb} \partial_m, \nu \Big\rangle \, \partial_p)$$

$$+2\Big\langle g^{az} \Big[-g_{ea} \frac{1}{\sqrt{detg}} ([\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] \partial_m (R^{melj} \sqrt{detg}))$$

$$-\frac{1}{\sqrt{detg}} \Big((\phi^\alpha_{ljm} - \phi^\alpha_r m \Gamma^r_{lj} - \phi^\alpha_r \partial_m (\Gamma^r_{lj})) R^m{}_a{}^{lj} \sqrt{detg} \Big) - \Gamma^u_{nk} [\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] R^{nklj} g_{ua} \Big] \partial_z, \nu \Big\rangle$$

$$+2\,div(\Big\langle [\phi^\alpha_{lj} - \phi^\alpha_r \Gamma^r_{lj}] R_a{}^l{}_b{}^j g^{ma} g^{pb} \partial_m, \nu \Big\rangle \, \partial_p)$$

$$-2\Big\langle g^{az}\Big[-g_{ea}\frac{1}{\sqrt{detg}}([\phi^{\alpha}_{lj}-\phi^{\alpha}_{r}\Gamma^{r}_{lj}]\partial_m(R^{mlej}\sqrt{detg}))$$

$$-\frac{1}{\sqrt{detg}}\left((\phi^{\alpha}_{ljm}-\phi^{\alpha}_{r}m\Gamma^{r}_{lj}-\phi^{\alpha}_{r}\partial_m(\Gamma^{r}_{lj}))R^{ml}{}_{a}{}^{j}\sqrt{detg}\right)-\Gamma^{u}_{nk}[\phi^{\alpha}_{lj}-\phi^{\alpha}_{r}\Gamma^{r}_{lj}]R^{nlkj}g_{ua}\Big]\partial_z,\nu\Big\rangle$$

$$-2\left\langle\phi^{\alpha}_{i}R^{i}{}_{jkl}R^{ajkl}\partial_a,\nu\right\rangle$$

Chapter 4

Normal Gradient Flow and Estimate for Flow in Fixed Direction

In this chapter we discuss the case of a gradient vector field being normal to $\phi(M)$ at every point (M is assumed to be closed). We first characterize when the gradient vector field ∇P_ϕ is normal; namely that the penalty function $P : \mathrm{Emb}(M, \mathbb{R}^N) \to \mathbb{R}$ is invariant under diffeomorphism of $\phi(M)$. We then give an explicit estimate, in terms of $\phi(M)$, for how long the embedding ϕ can flow in a fixed, normal gradient direction and remain in the space of embeddings. It is important to note that $\mathrm{Emb}(M, \mathbb{R}^N)$ is open in the space of all maps from M to $\mathbb{R}^N$ in both the C^k and C^∞ topologies.

4.1 Condition for Normal Gradient Vector Field

In this section, we prove an infinite dimensional analogue of the standard finite dimensional result that gradient vectors are perpendicular to level surfaces.

In the following theorem we use the gradient of the penalty function ∇P, which is defined with respect to the L^2 inner product on $T_\phi C^\infty(M, \mathbb{R}^N)$. For $X \in T_\phi \mathrm{Emb}(M, \mathbb{R}^N)$, the gradient is characterized by

$$dP_\phi(X) = \langle \nabla P_\phi, X \rangle = \int_{\phi(M)} \nabla P \cdot X \mathrm{dvol}$$

where the volume form is induced from $\mathbb{R}^N$ and we are using the Euclidean dot product. ∇P being pointwise normal to $\phi(M)$ means that $\nabla P_{\phi(m)} \cdot X_{\phi(m)} = 0$ for all $\phi(m) \in \phi(M)$.

Theorem 3. *For a penalty function* $P : \text{Emb}(M, \mathbb{R}^N) \to \mathbb{R}$*, and fixed* $\phi \in \text{Emb}(M, \mathbb{R}^N)$*, the gradient* ∇P *will be normal to* $\phi(M)$ *for each* $m \in M$ *if and only if* P *is invariant under diffeomorphisms* $\alpha : \phi(M) \to \phi(M)$*, that are in the path component of the identity in Diff(*$\phi(M)$*), i.e.* $P(\alpha(\phi(M))) = P(\phi(M))$*.*

Proof. ($\Leftarrow$) Assume $P(\alpha(\phi(M))) = P(\phi(M))$ where $\alpha : \phi(M) \to \phi(M)$ is a diffeomorphism that is generated from the flow of a time independent vector field on $\phi(M)$. We know that $\nabla P_\phi \perp_{L^2} X_\phi$ for all X_ϕ that are tangent to the level set containing $\phi \in \text{Emb}(M, \mathbb{R}^N)$.

Claim: All vector fields $Y_\phi \in \Gamma(T\phi(M))$ lie tangent to the level set of $\phi \in \text{Emb}(M, \mathbb{R}^N)$.

Proof of Claim: For $Y_\phi \in \Gamma(T\phi(M))$ we have an associated flow along $\phi(M)$ given by

$$\alpha_{Y,t} : \phi(M) \to \phi_t(M)$$

$$\alpha_{Y,t}(\phi(m)) = \phi_t(m)$$

and $\dot{\alpha_{Y,t}}\phi(m) = Y_{\phi(m)}$. Furthermore $\alpha_{Y,t} : \phi(M) \to \phi_t(M)$ is a diffeomorphism for all t. Therefore we can say

$$D_\phi P(Y) = \frac{\mathrm{d}}{\mathrm{dt}}|_{t=0} P(\phi_t) = \frac{\mathrm{d}}{\mathrm{dt}}|_{t=0} P(\alpha_{Y,t}(\phi)) = 0$$

where we have used the assumption and the fact that $\alpha_t(\phi) = \phi_t$.

We conclude that $\nabla P_\phi \perp_{L^2} Y_\phi$ for all vector fields $Y_\phi \in \Gamma(T\phi(M))$ and therefore that Y_ϕ lies tangent to ϕ's level set in $\text{Emb}(M, \mathbb{R}^N)$. We now need to show that $\nabla P_\phi(\phi(m)) \perp Y_\phi(\phi(m))$ pointwise.

Fix $\phi(m_0) \in \phi(M)$ and a vector $Q(\phi(m_0)) \in T_{\phi(m_0)}\phi(M)$. Choose a sequence of smooth functions $f_{\epsilon_k} : \phi(M) \to \mathbb{R}$ such that $\int_{\phi(M)} f_{\epsilon_k} \text{dvol} = 1$, $\text{supp} f_{\epsilon_k} \subset B_{\epsilon_k}(\phi(m_0)) \cap \phi(M)$ and $\epsilon_k \to 0$ (Here, $B_{\epsilon_k}(\phi(m_0))$ is the ball of radius ϵ_k centered at $\phi(m_0)$). Define vector fields

Y_{ϵ_k} on $\phi(M)$ by

$$Y_{\epsilon_k}(\phi(m)) = f_{\epsilon_k}(\phi(m)) \cdot Q(\phi(m_0)).$$

Then we have

$$\begin{aligned} 0 &= \lim_{\epsilon_k \to 0} \langle \nabla P_\phi, Y_{\epsilon_k} \rangle = \lim_{\epsilon_k \to 0} \langle \nabla P_\phi, f_{\epsilon_k} \cdot Q(\phi(m_0)) \rangle \\ &= \lim_{\epsilon_k \to 0} \int_{\phi(M)} \nabla P_\phi(\phi(m)) \cdot f_{\epsilon_k} Q(\phi(m_0)) = \nabla P_\phi \phi(m_0) \cdot Q(\phi(m_0)). \end{aligned}$$

Therefore $\nabla P_\phi \perp Y_\phi$ pointwise.

($\Rightarrow$) Assume that $\nabla P_{\phi(m)} \perp \phi(M)$ for all $\phi(m) \in \phi(M)$. This is equivalent to saying $\nabla P_{\phi(m)} \perp Y_{\phi(m)}$ at each point $\phi(m) \in \phi(M)$ for all vector fields $Y \in \Gamma(T\phi(M))$. This gives

$$\frac{\mathrm{d}}{\mathrm{dt}}|_{t=0} P(\phi_t) = 0, \quad \dot{\phi}_t|_{t=0} = Y,$$

which means that moving in the direction of the flow $\alpha_{Y,T}$ generated by a fixed vector field Y is equivalent to moving along a level set in $\mathrm{Emb}(M, \mathbb{R}^N)$. Because flows generated in this way are diffeomorphisms of $\phi(M)$ we can conclude that

$$P(\alpha_{Y,t}(\phi(M))) = P(\phi(M))$$

for all α, t, Y.

$\square$

4.2 An Estimate for Flows in Normal Gradient Directions

The above result gives a condition for determining if the gradient vector field generated by a penalty function is normal at every point in $\phi(M)$. In the case where this is true, we would next like to consider how far $\phi(M)$ can move in a fixed normal gradient direction while remaining an embedding. The next set of results gives an explicit estimate for the lower bound of this flow.

3.1 **Notation and Definitions**

- ϵ is the size of a neighborhood around $\phi(M)$ in which each point has a unique closest point in $\phi(M)$. The existence of this neighborhood for M closed is guaranteed by the ϵ-Neighborhood Theorem [3]. It is given explicitly in Lemma 6 in the proof of Theorem 5.
- We will use two sets of coordinates on $\mathbb{R}^N$. Standard coordinates will be denoted $(x^1, \ldots, x^N)$. We will also be representing points in $\phi(M)$ and in a small neighborhood around $\phi(M)$ as elements of the normal bundle $N\phi(M)$. In coordinates they will be given as $(q^1, \ldots, q^k, r^1, \ldots, r^{N-k})$ where the first k components are manifold coordinates and the last $N-k$ are coordinates for the normal space. These will be referred to as normal coordinates. For $q \in \phi(M)$, its representation is $(q^1, \ldots, q^k, 0, \ldots, 0)$. For $w = (q^1, \ldots, q^k, r^1, \ldots, r^{N-k})$ inside a small neighborhood of $\phi(M)$, $q = (q^1 \cdots q^k, 0 \cdots 0)$ is w's closest point in $\phi(M)$ and $(0, \cdots, 0, r^1, \cdots r^{N-k}) = w - q \in N\phi(M)$.
- A vector in $N_q\phi(M)$ will be denoted as either $t\vec{v}(q)$ (where $\vec{v}$ is unit length- this notation is generally used when referring to a fixed normal vector field) or as $r^i w_i(q)$ where the vectors $\{w_i\}$ are an orthonormal spanning set of the normal space at q. There are $N-k$ vectors $\{w_i\}$, each with N coordinates (note this can only be done locally).
- For $\phi(M) \subset \mathbb{R}^N$, the map $E : N\phi(M) \to \mathbb{R}^N$ acts by $E(q, r) = q + r$ (sending points to the end of perpendicular vectors in the normal bundle over $\phi(M)$). It is given explicitly by:

$$\begin{aligned} E((q^1, \cdots, q^k, r^1, \cdots, r^{n-k})) &= (x^1(q) + r^i w_i^1(q), \cdots, x^N(q) + r^i w_i^N(q)) \\ &= (\phi^1(q) + r^i w_i^1(q), \cdots, \phi^N(q) + r^i w_i^N(q)), \end{aligned}$$

where the domain is in normal coordinates and the range is in standard coordinates. Points $e = q_e + r_e$ for which the Jacobian of the E map isn't full rank (at the point (q_e, r_e)) are defined as 'focal points.' [5]

• The inclusion map $\phi(M) \to \mathbb{R}^N$ takes points $(q^1, \cdots, q^k) \mapsto (x^1(\vec{q}), \cdots, x^N(\vec{q}))$. It is a standard result that the first fundamental form in the direction $\vec{v}$ is the matrix with entries $(g_{ij}) = \left(\frac{\partial \vec{x}}{\partial q^i} \cdot \frac{\partial \vec{x}}{\partial q^j}\right)$ (Euclidean dot product) and the second fundamental form is the matrix with entries $(\vec{v} \cdot \vec{l}_{ij})$ where $\vec{l}_{ij}$ is the normal component of the vector $\frac{\partial^2 \vec{x}}{\partial q^i \partial q^j}$ and $\vec{v}$ is a fixed unit normal vector field. • In choosing coordinates that make the first fundamental form the identity matrix, the eiqenvalues $p_1, \cdots, p_k$ of the second fundamental form are called the 'principal curvatures' at $q = \phi(m) \in \phi(M)$. Considering the normal line $l = q + t\vec{v}$ extending from $q \in \phi(M)$ ($\vec{v}$ is a fixed unit normal vector at q) we have the proposition [5, p. 34]:

Proposition 3. *The focal points of* $([\phi(M)], q)$ *along* l *are precisely the points* $q + p_i^{-1}\vec{v}$, *where* $1 \leq i \leq k, p_i \neq 0$

• $K = \max_{\phi(m) \in \phi(M)} p_{\phi(m)}$ where $p_{\phi(m)}$ is the largest eigenvalue of $(\vec{r}_{\phi(m)} \cdot l_{ij})$ evaluated at $q = \phi(m) \in \phi(M)$ and r is a unit length normal vector in the normal bundle $N_{\phi(m)}\phi(M)$.

• δ is chosen such that for $d_{\mathbb{R}^N}(x, y) < \delta$ $(x, y \in \phi(M))$ we know $x + r_x^i w_i(x) \neq y + r_y^i(y)$ for $|r_x| < \delta$ and $|r_y| < \delta$. It is defined explicitly after the proof of Lemma 6.

Note: The next two theorems are stated in terms of unit length normal vector fields on $\phi(M)$. The Euler class of the normal bundle is the obstruction to the existence of such a vector field. If this class is nonzero, we apply the theorem to vector fields where each vector has length at most one.

Theorem 4. *Let* $\vec{v}$ *be a normal vector field of length at most one along* $\phi(M) \subset \mathbb{R}^N$ *and* ϵ *be as defined above.* $\phi_t(M) = \{\phi(m) + t\vec{v} : m \in M\}$ *is immersed in* $\mathbb{R}^N$ *for* $t < \epsilon$.

Proof. We want to show that the map $M \to \phi_t(M)$ is an immersion for $t < \epsilon$, but because $\phi(M)$ is assumed to be embedded in $\mathbb{R}^N$ it suffices to show that the map $F : \phi(M) \to \phi_t(M)$ (where for $q \in \phi(M)$, $F(q) = q + t\vec{v}(q)$) is an immersion. We want to consider $\phi_t(M)$ as sitting in an open subset of $\mathbb{R}^N$ that we can identify with the normal bundle over $\phi(M)$. In particular, the ϵ-Neighborhood Theorem [3] gives that on a compact, manifold without boundary in $\mathbb{R}^N$, $\phi(M)$ in our case, there exists a sufficiently small ϵ such that for each point w in Y^ϵ– the set of points in $\mathbb{R}^N$ a distance less than ϵ from the manifold– there is a unique closest point q in $\phi(M)$. Furthermore $w - q \in N_q\phi(M)$. We can diffeomorphically identify points in Y^ϵ with elements in $N\phi(M)$ as follows:

$$w \mapsto (w - q)_q$$

where q is w's unique closest point in $\phi(M)$. When considering the case of our fixed vector field $t\vec{V}$ along $\phi(M)$ as a section of the normal bundle we get the following coordinate representation of this section:

$$\phi(m) + t\vec{v}(\phi(m)) \mapsto (q^1, \cdots, q^k, tv^1(q), \cdots, tv^{N-k}(q))$$

where now the vector components are function of q. Therefore the map:

$$F : \phi(M) \to \phi_t(M) \subset Y^\epsilon$$

has the normal coordinate representation:

$$(q^1, \cdots, q^k) \mapsto (q^1, \cdots, q^k, tv^1(q), \cdots, tv^{N-k}(q))$$

the differential of which is given by:

$$DF(q) = \begin{pmatrix} \frac{\partial q^1(q)}{\partial q^1} & \cdots & \frac{\partial q^1(q)}{\partial q^k} \\ \vdots & & \vdots \\ \frac{\partial (tv)^{n-k}(q)}{\partial q^1} & \cdots & \frac{\partial (tv)^{n-k}(q)}{\partial q^k} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 1 \\ \vdots & & \vdots \\ \frac{\partial (tv)^{n-k}(q)}{\partial q^1} & \cdots & \frac{\partial (tv)^{n-k}(q)}{\partial q^k} \end{pmatrix}$$

which has rank k, showing that the map taking $\phi(M) \to \phi_t(M)$ is an immersion for $t < \epsilon$. □

Next, we would like to show that ϕ_t is injective, which along with its being an immersion (Theorem 4) and the assumption that M is compact is enough to conclude that ϕ_t is an embedding. While Theorem 4 showed that ϕ_t is an immersion for $t \leq \epsilon$, Theorem 5 will show injectivity for $t \leq t^*$. Lemma 5 (included in the proof of Theorem 5) shows that $t^* \leq \epsilon$. Therefore the final theorem showing ϕ_t is an embedding is on the interval $t \leq t^*$.

The statement of Theorem 5 uses the new value δ which is defined explicitly after the proof of Lemma 6. Recall that δ is chosen such that for $d_{\mathbb{R}^N}(x, y) < \delta$ $(x, y \in \phi(M))$ we know $x + r_x^i w_i(x) \neq y + r_y^i w_i(y)$ for $|r_x| < \delta, |r_y| < \delta$ (note that the definition applies to general vectors in $N\phi(M)$, as opposed to the fixed normal vector field $\vec{v}$).

Theorem 5. *Let $\vec{v}$ be a normal vector field of length at most one along $\phi(M) \subset \mathbb{R}^N$ Let $t^* = min\{K^{-1}, \delta/3\}$. Then $\phi_t : M \to \mathbb{R}^N$ given by $m \mapsto \phi(m) + tv(\vec{\phi(m)})$ is an embedding for $t \leq t^*$.*

Proof. It should be noted that we are interested in the injectivity of the map $\phi_t : M \to \mathbb{R}^N$ defined above, but because $\phi(M)$ is embedded in $\mathbb{R}^N$ it suffices to show that $F : \phi(M) \to \phi_t(M)$ is injective for $t \leq t^*$.

To view F as a map acting on open subsets of $\mathbb{R}^N$ we define the function H_t from $Y^{\epsilon - t} \to Y^\epsilon$, the set of points a distance $\epsilon - t$ and ϵ from $\phi(M)$ in $\mathbb{R}^N$ respectively. Setting

$\pi : Y^\epsilon \to \phi(M)$ with $\pi(w)$ the closest point in $\phi(M)$ to w we can define:

$$H_t(w) = w + t\vec{v}_{\pi(w)}.$$

Note that $H_t|_{\phi(M)} = F$.

We continue the proof with a series of Lemmas.

Lemma 4. *$DH_t(q_0)$ is invertible for $w = q_0 \in \phi(M)$*

Proof. For $H_t : Y^{\epsilon - t} \to Y^\epsilon$ via $w \mapsto w + t\vec{v}(\pi(w))$ its normal coordinate representation (explained in proof of Theorem 4) is given by:

$$(q^1, \cdots, q^k, r^1, \cdots, r^{N-k}) \mapsto (q^1, \cdots, q^k, r^1 + tv^1(\pi(q)), \cdots, r^{n-k} + tv^{N-k}(\pi(q)))$$

where it should be noted that the r^i's are independent of coordinates but the $v^i(q)$'s are the coordinates for the fixed vector field along $\phi(M)$ which depend on q. For $w = q_0 \in \phi(M)$ the differential of the H_t map (taken in coordinates) is given by:

$$DH_t(w)$$

$$= \begin{pmatrix} \frac{\partial q^1(\vec{q},0)}{\partial q^1} & \cdots & \frac{\partial q^1(\vec{q},0)}{\partial q^k} & \frac{\partial q^1(\vec{q},0)}{\partial r^1} & \cdots \\ \vdots & & & & \\ \vdots & & & & \\ \frac{\partial q^k(\vec{q},0)}{\partial q^1} & \cdots & \frac{\partial q^k(\vec{q},0)}{\partial q^k} & \frac{\partial q^k(\vec{q},0)}{\partial r^1} & \cdots \\ \frac{\partial (r^1 + tv^1(q))(\vec{q},0)}{\partial q^1} & \cdots & \frac{\partial (r^1 + tv^1(q))(\vec{q},0)}{\partial q^k} & \frac{\partial (r^1 + tv^1(q))(\vec{q},0)}{\partial r^1} & \cdots \\ \vdots & & & & \vdots \\ \frac{\partial (r^{n-k} + tv^{n-k}(q))(\vec{q},0)}{\partial q^1} & \cdots & \frac{\partial (r^{n-k} + tv^{n-k}(q))(\vec{q},0)}{\partial q^k} & \frac{\partial (r^{n-k} + tv^{n-k}(q))(\vec{q},0)}{\partial r^1} & \cdots \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ \frac{\partial(tv^1(q))(\vec{q},0)}{\partial q^1} & \cdots & \frac{\partial(tv^1(q))(\vec{q},0)}{\partial q^k} & 1 & \cdots & 0 \\ \vdots & & & & & \vdots \\ \frac{\partial(tv^{n-k}(q))(\vec{q},0)}{\partial q^1} & \cdots & \frac{\partial(tv^{n-k}(q))(\vec{q},0)}{\partial q^k} & 0 & \cdots & 1 \end{pmatrix}$$

This matrix is invertible for all t so we can conclude that there exists a ball $B_{\delta_{H_t}^{q_0}}$ of radius $\delta_{H_t}^{q_0}$ around q_0, on which H_t is a diffeomorphism. □

Let $\delta_{H_t} = \min_{q_0} \delta_{H_t}^{q_0}$. Although DH_t is invertible for all time (the size of the neighborhood will change according to t), we must have $t < \epsilon$ for H_t to be defined. Therefore t is less than ϵ and we can say: For $x, y \in \phi(M)$ with $d_{\mathbb{R}^N}(x,y) < \delta_{H_t}$, we have $x + t\vec{v}(x) \neq y + t\vec{v}(y)$ for $t < \epsilon$, and we can show injectivity:

Lemma 5. $H_t|_{\phi(M)}$ *is injective for* $t < t^* = min\{\epsilon, \frac{\delta_{H_t}}{3}\}$.

Proof. Assume instead that there exists some $x, y \in \phi(M)$ such that $x + tv\vec{(x)} = y + tv\vec{(y)}$ and $t < t^*$. We know by assumption that $d_{\mathbb{R}^N}(x,y) > \delta_{H_t}$. Therefore:

$$\begin{aligned} \delta_{H_t} < d_{\mathbb{R}^N}(x,y) &= |x-y| \\ &= |x - (x + t\vec{v}(x)) + (x + t\vec{v}(x)) - y| \\ &= |x - (x + t\vec{v}(x)) + (y + t\vec{v}(y)) - y| \\ &\leq |x - (x + t\vec{v}(x))| + |(y + t\vec{v}(y)) - y| \\ &= |t\vec{v}(x)| + |t\vec{v}(y)| = 2|t| < 2|t^*| \\ &\leq 2\delta_{H_t}/3 \end{aligned}$$

which is a contradiction. □

We now must compute ϵ (the size of the neighborhood around $\phi(M)$ within which each point has a unique closest point in $\phi(M)$). Lemma 6 again uses δ which is defined explicitly following the proof. Recall: δ is chosen such that for $d_{\mathbb{R}^N}(x,y) < \delta$ $(x, y \in \phi(M))$ we know $x + r_x^i w_i(x) \neq y + r_y^i w_i(y)$ for $|r_x| < \delta$ and $|r_y| < \delta$

Lemma 6. *We may take* $\epsilon = min\{K^{-1}, \delta/3\}$ *where* $|r_x| < \delta$ *and* $|r_y| < \delta$.

Proof. Suppose there exists $w \in Y^\epsilon$ such that there are two closest points $x, y \in \phi(M)$. Then we can write $w = x + r_x^i w_i(x) = y + r_y^i w_i(y)$ where $|r_x| < \epsilon$ and $|r_y| < \epsilon$. We know by assumption that $d_{\mathbb{R}^N}(x,y) > \delta$ and we have a similar proof as in Lemma 5:

$$\begin{aligned}
\delta < d_{\mathbb{R}^N}(x,y) &= |x - y| \\
&= |x - (x + r_x^i w_i(x)) + (x + r_x^i w_i(x)) - y| \\
&= |x - (x + r_x^i w_i(x)) + (y + r_y^i w_i(y)) - y| \\
&\leq |x - (x + r_x^i w_i(x))| + |(y + r_y^i w_i(y)) - y| \\
&= |r_x^i w_i(x)| + |r_y^i w_i(y)| \\
&= |r_x| + |r_y| < 2\epsilon \leq 2\delta/3
\end{aligned}$$

which is a contradiction. □

We will obtain δ in the following way: Recall $E : N\phi(M) \to \mathbb{R}^N$ acts on points in the normal bundle over $\phi(M)$ by $(q, r) \mapsto q + r$. Here we will be considering the compact subset of $N\phi(M)$ which consists of vectors $\vec{r}$ such that $|r| \leq .999K^{-1}$. In coordinates, recall E is given by:

$$\begin{aligned}
E((q^1, \cdots, q^k, r^1, \cdots, r^{n-k})) &= (x^1(q) + r^i w_i^1(q), \cdots, x^N(q) + r^i w_i^N(q)) \\
&= (\phi^1(q) + r^i w_i^1(q), \cdots, \phi^N(q) + r^i w_i^N(q)).
\end{aligned}$$

Fix $q_0 = (q_0^1, \cdots, q_0^k, 0, \cdots, 0) \in \phi(M)$. For a point (q_0, r_0) in the fiber over q_0 we know that $DE(q_0, r_0)$ is invertible (see proof of Proposition 3) and therefore there is a ball of radius $\delta_{(q_0,r_0)}$ around (q_0, r_0) on which E is a diffeomorphism. Because the fiber over q_0 is compact, we can let $\delta_{q_0} = \min_{r_0} \ \delta_{(q_0,r_0)} > 0$.

Consider the set

$$A_{q_0} = \{q \in \phi(M) \ : d_{\mathbb{R}^N}\,(q, q_0) < \delta_{q_0}/2\}.$$

Then E is a diffeomorphism on the subset of $N\phi(M)$ given in normal coordinates by

$$B_{q_0} = \{(q^1, \cdots, q^k, r^1, \cdots, r^{n-k})| \ |r| < \delta_{q_0}/2, (q^1, \cdots, q^k, 0, \cdots, 0) \in A_{q_0}\}$$

as follows: For $(q_1, r_1) \in B_{q_0}$:

$$\begin{aligned} |(q_1, r_1) - (q_0, 0)| &= |(q_1, r_1) - (q_1, 0) + (q_1, 0) - (q_0, 0)| \\ &< |(q_1, r_1) - (q_1, 0)| + |(q_1, 0) - (q_0, 0)| \\ &= |r_1| + |(q_1, 0) - (q_0, 0)| \\ &< \delta_{q_0}/2 + \delta_{q_0}/2 = \delta_{q_0}. \end{aligned}$$

Therefore for $(q_1, r_1), (q_2, r_2) \in B_{q_0}$ $\big((q_1, r_1) \neq (q_2, r_2)\big)$ we know $(q_1, 0), (q_2, 0) \in A_{q_0}$ and $E((q_1, r_1)) = q_1 + r_1^i w_i \neq q_2 + r_2^1 w_1 = E((q_2, r_2))$.

We let

$$\delta = \inf_{q_0} \ \delta_{q_0}/2.$$

We can now say that for $x, y \in \phi(M)$ and $d_{\mathbb{R}^N}(x, y) < \delta$ we have $x + r_x^i w_i(x) \neq y + r_y^i w_i(y)$ for $|r_x| < \delta$ and $|r_y| < \delta$ by construction.

It remains to compute $\delta_{(q_0,r_0)}$ explicitly, from which we can get δ with the method described above (Recall, $\delta_{(q_0,r_0)}$ is the radius around (q_0, r_0) on which E is a diffeomorphism). We will compute $\delta_{(q_0,r_0)}$ using a quantitative version of the Implicit Function Theorem (adapted to the Inverse Function Theorem case), given as a proposition below. The formulation of the theorem, along with its proof is in Appendix B.

For $G \in C^1(\mathbb{R}^{2N}, \mathbb{R}^N)$, let $(q_0, y_0) \in \mathbb{R}^{2N}$ satisfy $G(q_0, y_0) = 0$. For fixed $\gamma > 0$ let $V_\gamma = \{(q, y) \in \mathbb{R}^{2N} : |q - q_0| \le \gamma, |y - y_0| \le \gamma\}$. In the case where $G(q, y) = E(q) - y$, the following theorem is the adaptation of the Implicit Function Theorem to the Inverse Function Theorem (here the matrix norm $||A||$ is the sup norm over the entries):

Proposition 4. *Assume that $\partial_q G(q_0, y_0)$ is invertible and choose $\delta^0 > 0$ such that $sup_{(q,y)\in V_{\delta^0}} ||1 - [\partial_q G(q_0, y_0)]^{-1}\partial_q G(q, y)|| \le 1/2$. Let $B_{\delta^0} = sup_{(q,y)\in V_{\delta^0}} ||\partial_y G(q, y)||$ and $M = ||\partial_q G(q_0, y_0)^{-1}||$. Let $\delta_1 = (2MB_{\delta^0})^{-1}\delta^0$ and $\Gamma_{\delta_1} = \{y \in \mathbb{R}^m : ||y - y_0|| < \delta_1\}$. Then in the case that $G(q, y) = E(q) - y$, the solutions to $G(q, y) = 0 (\Rightarrow E(q) = y)$ in the set $\{(q, y) : ||q - q_0|| < \delta^0, ||y - y_0|| < \delta_1\}$ are given by $(E^{-1}(y), y)$. Alternatively, E is a diffeomorphism on $E^{-1}(B_{\delta_1}(y_0)) \cap B_{\delta^0}(q_0)$.*

We will apply the proposition to $E : N\phi(M) \to \mathbb{R}^N$. Specifically, in applying the proposition we have $((q_0, r_0), y_0)$ as a base point (as opposed to simply writing (q, y) as in the proposition statement, we will write $((q, r), y)$ to emphasize use of normal coordinates), we have $G((q, r), y) = E(q, r) - y$ and $G((q_0, r_0), y_0) = 0$ ($\Rightarrow E((q_0, r_0)) = y_0$). Therefore:

$$\partial_{(q,r)} G((q_0, r_0), y_0) = DE(q_0, r_0) = \begin{pmatrix} \frac{\partial\phi^1(q_0,r_0)}{\partial q^1} + r^i \frac{\partial w_i^1(q_0,r_0)}{\partial q^1} & \cdots & w_{n-k}^1(q_0) \\ \vdots & & \vdots \\ \frac{\partial\phi^N(q_0,r_0)}{\partial q^1} + r^i \frac{\partial w_i^N(q_0,r_0)}{\partial q^1} & \cdots & w_{n-k}^N(q_0) \end{pmatrix}$$

which is invertible for $|r| < K^{-1}$ as required by the proposition's assumption. Again, our goal is to get a $\delta_{(q_0,r_0)}$ neighborhood around (q_0, r_0) on which E is a diffeomorphism. Following the proposition's steps we have:

Step 1:

$$B_{\delta^0_{(q_0,r_0)}} = \sup_{((q,r),y)\in V_{\delta^0_{(q_0,r_0)}}} ||\partial_y G((q, r), y)||$$

$$= \sup_{((q,r),y)\in V_{\delta^0_{(q_0,r_0)}}} ||\partial_y(E(q,r)-y)||$$

$$= \sup_{((q,r),y)\in V_{\delta^0_{(q_0,r_0)}}} \left\| \begin{pmatrix} -1 & 0 & 0 \\ \vdots & & \vdots \\ 0 & & -1 \end{pmatrix} \right\| = 1,$$

Step 2:

$$M = ||\partial_{(q,r)}G((q_0,r_0),y_0)^{-1}|| = ||DE(q_0,r_0)^{-1}||.$$

Using Cramer's rule and the matrix adjugate to invert $DE(q_0,r_0)$, we have

$$(DE(q_0,r_0)^{-1})_{(j,z)} = \frac{1}{\det(DE(q_0,r_0))}(-1)^{(z+j)}DE(q_0,r_0)^*_{(j,z)}$$

where $DE(q_0,r_0)^*_{(j,z)}$ is the (j,z)th minor of $DE(q_0,r_0)$, or the determinant of the $(n-1)\times(n-1)$ matrix constructed by deleting the jth row and zth column of $DE(q_0,r_0)$, which gives an explicit way to compute M above.

Step 3:

We want to compute $\delta^0_{(q_0,r_0)}$ such that $\sup_{((q,r),y)\in V_{\delta^0_{(q_0,r_0)}}} ||1-[DE(q_0,r_0)]^{-1}DE(q,r)|| \leq 1/2$. Since this expression doesn't rely on y, we need $\delta^0_{(q_0,r_0)}$ such that for $|(q,r)| < \delta^0_{(q_0,r_0)} \Rightarrow ||1-[DE(q_0,r_0)]^{-1}DE(q,r)|| \leq 1/2$. To do this we can consider a first order Taylor series expansion of $DE(q,r)$ around (q_0,r_0). (Note: the j index in the second matrix below refers to coordinates in $\mathbb{R}^N$, not an exponent.) We have:

$$DE(q,r)$$
$$= \begin{pmatrix} \frac{\partial\phi^1(q_0,r_0)}{\partial q^1} + r^i\frac{\partial w_i^1(q_0,r_0)}{\partial q^1} & \cdots & w^1_{N-k}(q_0) \\ \vdots & & \vdots \\ \frac{\partial\phi^N(q_0,r_0)}{\partial q^1} + r^i\frac{\partial w_i^N(q_0,r_0)}{\partial q^1} & \cdots & w^N_{N-k}(q_0) \end{pmatrix}$$

$$+\begin{pmatrix} \sum_{j=1}^{N} R_j^{(1,1)}(q,r)(z-z_o)^j & \cdots & \sum_{j=1}^{N} R_j^{(1,N)}(q,r)(z-z_o)^j \\ \vdots & & \vdots \\ \sum_{j=1}^{N} R_j^{(N,1)}(q,r)(z-z_o)^j & \cdots & \sum_{j=1}^{N} R_j^{(N,N)}(q,r)(z-z_o)^j \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial \phi^1(q_0)}{\partial q^1} + r_0^i \frac{\partial w_i^1(q_0)}{\partial q^1} & \cdots & w_{N-k}^1(q_0) \\ \vdots & & \vdots \\ \frac{\partial \phi^N(q_0)}{\partial q^1} + r_0^i \frac{\partial w_i^N(q_0)}{\partial q^1} & \cdots & w_{N-k}^N(q_0) \end{pmatrix}$$

$$+\begin{pmatrix} \sum_{j=1}^{N} R_j^{(1,1)}(q,r)(z-z_o)^j & \cdots & \sum_{j=1}^{N} R_j^{(1,N)}(q,r)(z-z_o)^j \\ \vdots & & \vdots \\ \sum_{j=1}^{N} R_j^{(N,1)}(q,r)(z-z_o)^j & \cdots & \sum_{j=1}^{N} R_j^{(N,N)}(q,r)(z-z_o)^j \end{pmatrix}$$

where $z - z_o = (q^1 - q_0^1, \cdots, q^k - q_0^k, r^1 - r_0^1, \cdots, r^{n-k} - r_0^{N-k})$. We have a uniform bound on the error term given by:

$$|R_j^{(l,m)}(q,r)| \leq \max\left\{ \left| \frac{\partial f_l^m((q,r))}{\partial z^j} \right| : 1 \leq j \leq N, r \leq .999K^{-1}, q \in \phi(M) \right\} \stackrel{\text{def}}{=} G^{(m,l)}$$

For (l,m) with $1 \leq l \leq N$ and $1 \leq m \leq k$, $f_l^m = \frac{\partial \phi^m(q)}{\partial q^l} + r^i \frac{\partial w_i^m(q)}{\partial q^l}$. For (l,m) with $1 \leq l \leq N$ and $k+1 \leq m \leq N$, $f_l^m = w_m^l(q)$.

Plugging the above sum for $DE(q,r)$ in the expression $||1 - [DE(q_0,r_0)]^{-1}DE(q,r)||$ we see that the first term cancels with the identity matrix and we are left with:

$$\left\| [DE(q_0,r_0)]^{-1} \begin{pmatrix} \sum_{j=1}^{N} R_j^{(1,1)}(q,r)(z-z_o)^j & \cdots & \sum_{j=1}^{N} R_j^{(1,N)}(q,r)(z-z_o)^j \\ \vdots & & \vdots \\ \sum_{j=1}^{N} R_j^{(N,1)}(q,r)(z-z_o)^j & \cdots & \sum_{j=1}^{N} R_j^{(N,N)}(q,r)(z-z_o)^j \end{pmatrix} \right\|$$

$$= \left\| \begin{pmatrix} ([DE(q_0,r_0)]^{-1})_{(1,p)} \sum_{j=1}^{N} R_j^{(p,1)}(q,r)(z-z_o)^j & \cdots \\ \vdots & \\ ([DE(q_0,r_0)]^{-1})_{(N,p)} \sum_{j=1}^{N} R_j^{(p,1)}(q,r)(z-z_o)^j & \cdots \end{pmatrix} \right.$$

$$\left. \cdots \begin{pmatrix} ([DE(q_0,r_0)]^{-1})_{(1,p)} \sum_{j=1}^{N} R_j^{(p,N)}(q,r)(z-z_o)^j \\ \vdots \\ ([DE(q_0,r_0)]^{-1})_{(N,p)} \sum_{j=1}^{N} R_j^{(p,N)}(q,r)(z-z_o)^j \end{pmatrix} \right\|$$

$$\leq \left\| \begin{pmatrix} ([DE(q_0,r_0)]^{-1})_{(1,p)} \delta^0_{(q_0,r_0)} \sum_{j=1}^{N} R_j^{(p,1)}(q,r) & \cdots \\ \vdots & \\ ([DE(q_0,r_0)]^{-1})_{(N,p)} \delta^0_{(q_0,r_0)} \sum_{j=1}^{N} R_j^{(p,1)}(q,r) & \cdots \end{pmatrix} \right.$$

$$\left. \cdots \begin{pmatrix} ([DE(q_0,r_0)]^{-1})_{(1,p)} \delta^0_{(q_0,r_0)} \sum_{j=1}^{N} R_j^{(p,N)}(q,r) \\ \vdots \\ ([DE(q_0,r_0)]^{-1})_{(N,p)} \delta^0_{(q_0,r_0)} \sum_{j=1}^{N} R_j^{(p,N)}(q,r) \end{pmatrix} \right\|$$

$$\leq \left\| \begin{pmatrix} ([DE(q_0,r_0)]^{-1})_{(1,p)} \delta^0_{(q_0,r_0)} NG^{(p,1)} & \cdots & ([DE(q_0,r_0)]^{-1})_{(1,p)} \delta^0_{(q_0,r_0)} NG^{(p,N)} \\ \vdots & & \vdots \\ ([DE(q_0,r_0)]^{-1})_{(N,p)} \delta^0_{(q_0,r_0)} NG^{(p,1)} & \cdots & ([DE(q_0,r_0)]^{-1})_{(N,p)} \delta^0_{(q_0,r_0)} NG^{(p,N)} \end{pmatrix} \right\| \tag{4.1}$$

Letting $\delta^0_{(q_0,r_0)} = \frac{1}{2\max\limits_{(l,m)}([DE(q_0,r_0)]^{-1})_{(l,p)} NG^{(p,m)}}$ we have that the last term in (4.1) does not exceed 1/2, as each entry has absolute value less than 1/2 by construction.

Step 4: Now that we have a value for $\delta^0_{(q_0,r_0)}$ we can compute $\delta^1_{(q_0,r_0)}$ as in the statement of Proposition 4 by:

$$\delta^1_{(q_0,r_0)} = (2MB_{\delta^0_{(q_0,r_0)}})^{-1} \delta^0_{(q_0,r_0)} = (2M)^{-1} \delta^0_{(q_0,r_0)}$$

where the last equality is from Step 1 and M is computed in Step 2.

Step 5: By Proposition 4 we know E is a diffeomorphism on

$$P_{(q_0,r_0)} = E^{-1}(B_{\delta^1_{(q_0,r_0)}}(y_0)) \cap B_{\delta^0_{(q_0,r_0)}}(q_0,r_0).$$

In particular, we need a ball of radius $\delta_{(q_0,r_0)}$ around (q_0,r_0) on which E is a diffeomorphism. First, we need a $\delta^3_{(q_0,r_0)}$ such that for

$$|(q,r)-(q_0,r_0)| < \delta^3_{(q_0,r_0)} \Rightarrow |E(q,r)-E(q_0,r_0)| = |E(q,r)-y_0| < \delta^1_{(q_0,r_0)}$$

We can again compute this $\delta^3_{(q_0,r_0)}$ using a Taylor series expansion of E around (q_0,r_0). We have

$$E(q,r) = E(q_0,r_0) + \big(\sum_j R^1_j(q,r)((q,r)-(q_0,r_0))^j, \cdots, \sum_j R^N_j(q,r)((q,r)-(q_0,r_0))^j\big)$$

where we have bounds on the error terms given by:

$$|R^p_j(q,r)| \le \max\left\{\left|\frac{\partial(\phi^p + r^i w^p_i)(q,r)}{\partial z^j}\right| : 1 \le j \le N, q \in \phi(M), r \le .999K^{-1}\right\} \overset{\text{def}}{=} G^p$$

Then we have

$$\begin{aligned}
&|E(q,r)-E(q_0,r_0)|^2 \\
&\quad = |\big(\sum_j R^1_j(q,r)((q,r)-(q_0,r_0))^j, \cdots, \sum_j R^N_j(q,r)((q,r)-(q_0,r_0))^j\big)|^2 \\
&\quad = \sum_{p=1}^N (\sum_j R^p_j(q,r)((q,r)-(q_0,r_0))^j)^2 = \sum_{p=1}^N |\sum_j R^p_j(q,r)((q,r)-(q_0,r_0))^j|^2 \\
&\quad \le \sum_{p=1}^N \sum_j |R^p_j(q,r)((q,r)-(q_0,r_0))^j|^2 \le \sum_{p=1}^N \sum_j |G^p((q,r)-(q_0,r_0))^j|^2
\end{aligned}$$

$$\leq \sum_{p=1}^{N}\sum_{j}|G^p\delta^3_{(q_0,r_0)}|^2 = (\delta^3_{(q_0,r_0)})^2\sum_{p=1}^{N}\sum_{j}|G^p|^2 = N(\delta^3_{(q_0,r_0)})^2\sum_{p=1}^{N}|G^p|^2.$$

Therefore

$$|E(q,r) - E(q_0,r_0)| \leq \delta^3_{(q_0,r_0)}\sqrt{N\sum_{p=1}^{N}|G^p|^2},$$

and letting $\delta^3_{(q_0,r_0)} = \delta^1_{(q_0,r_0)}/\left(\sqrt{N\sum_{p=1}^{N}|G^p|^2}\right)$ gives the required radius. We finally set $\delta_{(q_0,r_0)} = \min\{\delta^3_{(q_0,r_0)}, \delta^0_{(q_0,r_0)}\}$

Returning to the statement in Lemma 5, we had: $H_t|_{\phi(M)}$ is injective for $t < t^* = \min\{\epsilon, \frac{\delta_{H_t}}{3}\} = \min\{K^{-1}, \delta/3, \frac{\delta_{H_t}}{3}\}$. By definition we know that for $x, y \in \phi(M)$ and $d_{\mathbb{R}^N}(x,y) < \delta$ we have $x + r^i_x w_i(x) \neq y + r^i_y w_i(y)$ (where $|r_x| < \delta$ and $|r_y| < \delta$). However, we also have that for $x, y \in \phi(M)$ satisfying $d_{\mathbb{R}^N}(x,y) < \delta_{H_t}$, $x + t\vec{v}(x) \neq y + t\vec{v}(y)$ for $t < \epsilon < \delta$. Therefore we can say that $\delta < \delta_{H_t}$. This is because our specific vector field $t\vec{v}$ gives a particular set of r^i's at each point, allowing for a larger diffeormorphic neighborhood around the base point than a neighborhood that works for all set of r^i's. Therefore we have $H_t|_{\phi(M)}$ is injective for $t < t^* = \min\{\epsilon, \frac{\delta_{H_t}}{3}\} = \min\{K^{-1}, \frac{\delta}{3}, \frac{\delta_{H_t}}{3}\} = \min\{K^{-1}, \frac{\delta}{3}\}$ as required.

We have shown that ϕ_t is an injective immersion for $t \leq t^*$ (by the fact that $t^* \leq \epsilon$). Since M is compact ϕ_t is an embedding.

This concludes the proof of Theorem 5. $\square$

Chapter 5

Distance Penalty Function: A Special Case

We would like to study gradient flow in the case of a simple example, where we can explicitly compute the flow, or at least determine its existence with the use of standard PDE techniques. We will embed a circle into $\mathbb{R}^2$ and examine its flow in approximating the fixed set of data points consisting only of the origin.

Furthermore, because of the computational complexity of the curvature term, this chapter will only treat the distance penalty term, restated below. In particular, we will consider three variations of the computation of its gradient vector field and their resulting flows. It should be noted that the second case is a projection of the first case onto normal directions and is therefore no longer the gradient, but is of interest in light of the findings in Chapter 4.

Let $M = S^1$, $N = 2$ and the collection of fixed points to be approximated ($S = \{x_i\}$ in the Introduction) consist only of the origin. With this set up, we are considering negative gradient flow in the space $\mathrm{Emb}(S^1, \mathbb{R}^2)$. The initial embedding ϕ_0 will be a circle centered on the y-axis:

$$\phi_0(\theta) = (\cos(\theta), \sin(\theta) - k)$$

We would like to look at the negative gradient flow of the distance penalty function

$$P_d(\phi) = \int_{S^1} d^2(\phi(m), x_i) \mathrm{dvol}_{S^1}$$

under three variations of the gradient computation. The computation of the gradient in the following three cases is given in Appendix C.

5.1 CASE 1: Holding the volume element constant in the gradient computation

Holding the volume element constant means we are considering the volume form on S^1 as opposed to the induced volumed form on $\phi(S^1)$. Treating the volume in this way for the gradient calculation gives a gradient flow of (See Appendix C):

$$\mathrm{grad}P(\phi(m)) = 2\phi(m)$$

which in general is non-normal to points in $\phi(S^1)$. With the set up and initial conditions above we have the system:

$$\frac{\mathrm{d}}{\mathrm{dt}}\phi_t(\theta) = -\mathrm{grad}P(\phi) = -2\phi_t(\theta)$$

$$\phi_0(\theta) = (\cos(\theta), \sin(\theta) - k)$$

for fixed $k \in \mathbb{Z}$. Consider the flow

$$\phi_t(\theta) = (\cos(\theta)e^{-2t}, (\sin(\theta) - k)e^{-2t})$$

Then we have

$$\frac{\mathrm{d}}{\mathrm{dt}}\phi_t(\theta) = (-2\cos(\theta)e^{-2t}, -2(\sin(\theta) - k)e^{-2t})) = -2\phi_t(\theta)$$

with the required initial condition.

NOTE: It is clear that the flow $\phi(\theta)$ exists for all time but we need to check that it is also an embedding for all time. We want to see if there exists a time t where:

$$\phi_t(\theta_1) = (\cos(\theta_1)e^{-2t}, (\sin(\theta_1) - k)e^{-2t}) = (\cos(\theta_2)e^{-2t}, (\sin(\theta_2) - k)e^{-2t}) = \phi_t(\theta_2)$$

For this to be the case we would have:

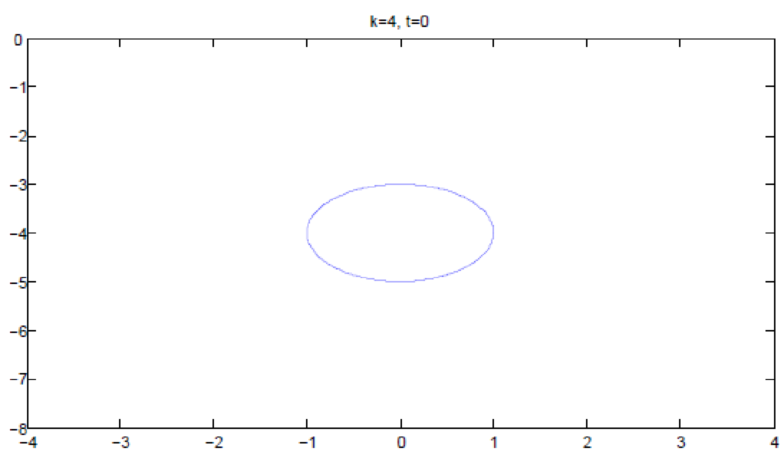
$$\cos(\theta_1) = \cos(\theta_2) \Rightarrow \theta_1 = \pm\theta_2$$

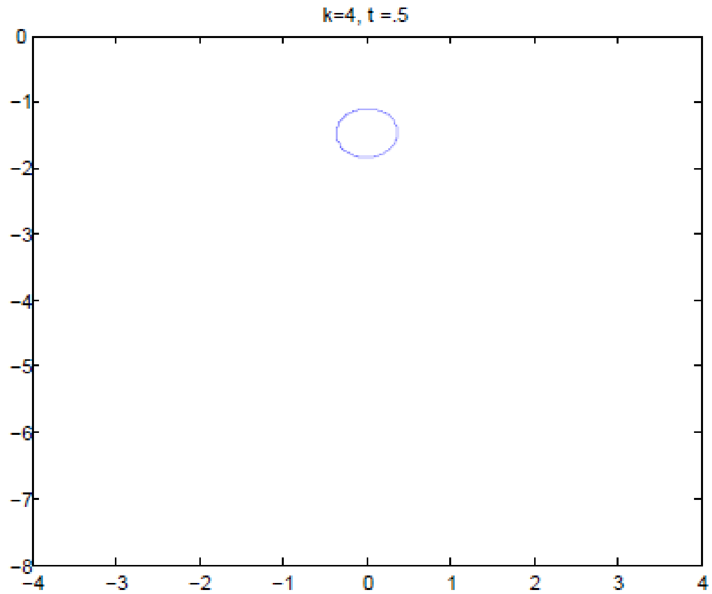
and

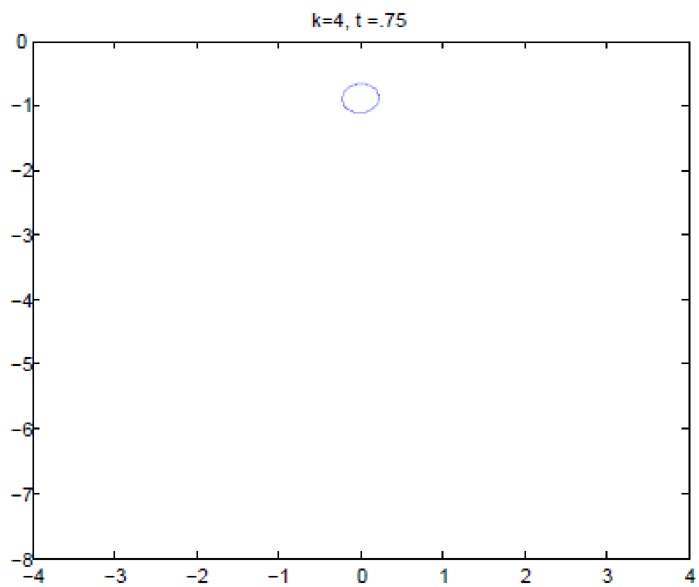
$$\sin(\theta_1) = \sin(\theta_2)$$

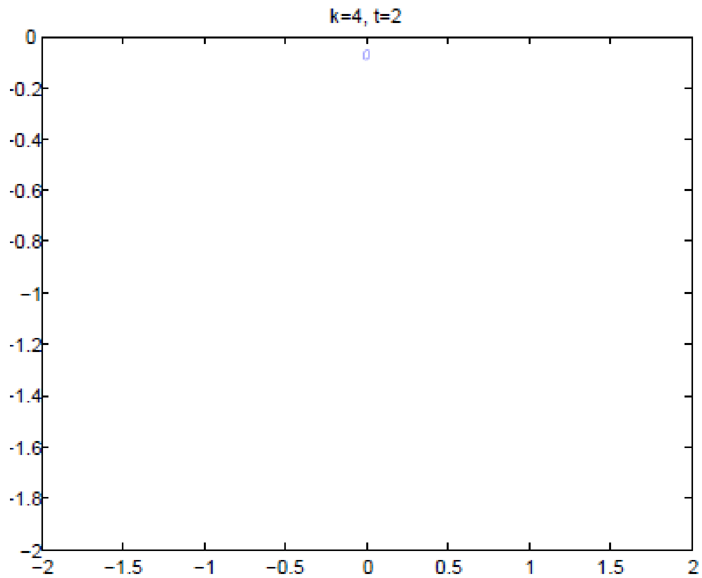
which with the restriction about means $\theta_1 = \theta_2$. Therefore the flow $\phi_t(\theta)$ is an embedding for all time.

Below we can see the progression of the initial embedding at four times (with k set to 4):




k=4, t =.5
0
-1
-2
-3
-4
-5
-6
-7
-8
−4
−3
−2
−1
0
1
2
3
4


k=4, t =.75
0
-1
-2
-3
-4
-5
-6
-7
-8
−4
−3
−2
−1
0
1
2
3
4



5.2 CASE 2: Projecting Case 1 Gradient Flow Onto Normal Directions

In this case we will study the projection of Case 1's gradient vector field onto normal directions. While this new vector field on $\phi(S^1)$ will no longer be gradient flow, it is computationally worthwhile to study from a PDE perspective, especially in light of Chapter 4's findings. The gradient flow from Case 1 given by

$$\mathrm{grad}P(\phi)(m) = 2\phi(m)$$

is in general not normal to point in $\phi(M)$. Adding in the term (see Appendix C):

$$-2\sum_{r,l=1}^{k}\sum_{j=1}^{N} g^{rl}\phi(m)^j \frac{\partial\phi(m)^j}{\partial y^l}\partial_r$$

will result in a vector field that projects the gradient flow's component functions onto their normal directions. It should be noted that the term above is for a general embedding into $\mathbb{R}^N$ and will be adapted to our case below.

With the same initial embedding and set up as in Case 1, we would like to write the new gradient equation as an explicit coupled system of PDE's. In other words we want to write $\frac{\partial\phi(t,\theta)}{\partial t} = -\text{grad}P(\phi)$ as a two component function of ϕ^1, ϕ^2 and their derivatives of θ and t.

Term 1:

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

by the setup.

Term 2: The manifold coordinates (r, l indices) consist only of θ in our case and $j = 1, 2$. Therefore the term reduces to:

$$-2\sum_{r,l=1}^{k}\sum_{j=1}^{N} g^{rl}\phi(m)^j \frac{\partial\phi(m)^j}{\partial y^l}\partial_r = -2(g^{\theta\theta}\phi^1\frac{\partial\phi^1}{\partial\theta} + g^{\theta\theta}\phi^2\frac{\partial\phi^2}{\partial\theta})\frac{\partial}{\partial\theta} \tag{5.1}$$

We need to write $\frac{\partial}{\partial\theta}$ as a two component vector in $\mathbb{R}^2$, where we are specifically considering the pushforward of $\frac{\partial}{\partial\theta}$ under the map $\phi : S^1 \to \mathbb{R}^2$. Therefore

$$\phi_*(\frac{\partial}{\partial\theta}) = \frac{\partial\phi^1}{\partial\theta}\partial_x + \frac{\partial\phi^2}{\partial\theta}\partial_y$$

Note that: $g^{\theta\theta} = 1/g_{\theta\theta}$, where

$$g_{\theta,\theta} = \frac{\partial\phi}{\partial\theta}\cdot\frac{\partial\phi}{\partial\theta} = (\frac{\partial\phi^1}{\partial\theta})^2 + (\frac{\partial\phi^2}{\partial\theta})^2 = |\frac{\partial\phi}{\partial\theta}|^2$$

and therefore $g^{\theta\theta} = 1/|\frac{\partial\phi}{\partial\theta}|^2$. Substituting back into (5.1) we have:

$$-2(g^{\theta\theta}\phi^1\frac{\partial\phi^1}{\partial\theta} + g^{\theta\theta}\phi^2\frac{\partial\phi^2}{\partial\theta})\frac{\partial}{\partial\theta} = \left(\frac{-2\phi^1}{|\frac{\partial\phi}{\partial\theta}|^2}\frac{\partial\phi^1}{\partial\theta} + \frac{-2\phi^2}{|\frac{\partial\phi}{\partial\theta}|^2}\frac{\partial\phi^2}{\partial\theta}\right)\left(\frac{\partial\phi^1}{\partial\theta}, \frac{\partial\phi^2}{\partial\theta}\right)$$

$$= \left(\frac{-2\phi^1(\frac{\partial \phi^1}{\partial \theta})^2 - 2\phi^2 \frac{\partial \phi^1}{\partial \theta}\frac{\partial \phi^2}{\partial \theta}}{|\frac{\partial \phi}{\partial \theta}|^2}, \frac{-2\phi^1 \frac{\partial \phi^1}{\partial \theta}\frac{\partial \phi^2}{\partial \theta} - 2\phi^2(\frac{\partial \phi^2}{\partial \theta})^2}{|\frac{\partial \phi}{\partial \theta}|^2} \right)$$

Putting the two terms together we have a final coupled system:

$$\frac{\partial \phi(t,\theta)}{\partial t} = (\frac{\partial \phi^1(t,\theta)}{\partial t}, \frac{\partial \phi^2(t,\theta)}{\partial t})$$

$$= - \left(2\phi^1 + \frac{-2\phi^1(\frac{\partial \phi^1}{\partial \theta})^2 - 2\phi^2 \frac{\partial \phi^1}{\partial \theta}\frac{\partial \phi^2}{\partial \theta}}{|\frac{\partial \phi}{\partial \theta}|^2}, 2\phi^2 + \frac{-2\phi^1 \frac{\partial \phi^1}{\partial \theta}\frac{\partial \phi^2}{\partial \theta} - 2\phi^2(\frac{\partial \phi^2}{\partial \theta})^2}{|\frac{\partial \phi}{\partial \theta}|^2} \right)$$

Each component in the vector field has the physical interpretation of projecting $-2\phi^i$ onto the normal direction at $\phi(m) \in \phi(M)$, making the gradient flow normal at every point. We can represent this system in matrix form:

$$\begin{pmatrix} \frac{\partial \phi^1}{\partial t} \\ \frac{\partial \phi^2}{\partial t} \end{pmatrix} = \frac{2}{|\frac{\partial \phi}{\partial \theta}|^2} \begin{pmatrix} \phi^1 \frac{\partial \phi^1}{\partial \theta} & \phi^2 \frac{\partial \phi^1}{\partial \theta} \\ \phi^1 \frac{\partial \phi^2}{\partial \theta} & \phi^2 \frac{\partial \phi^2}{\partial \theta} \end{pmatrix} \begin{pmatrix} \frac{\partial \phi^1}{\partial \theta} \\ \frac{\partial \phi^2}{\partial \theta} \end{pmatrix} - 2 \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix}$$

$$\phi(\theta, 0) = (\phi^1(\theta, 0), \phi^2(\theta, 0)) = (\cos(\theta), \sin(\theta) - k)$$

This is a first order fully nonlinear system in (ϕ^1, ϕ^2), as the first order term is squared. We would like to apply the Cauchy-Kovalevskaya theorem [1] to the system, which gives a power series solution with a radius of convergence r in the (θ, t) domain (r to be determined). The full text of the theorem, as written in Evans' *Partial Differential Equations* [1], is given in Appendix D. Obtaining a value for r will give us a time interval on which the solution is defined. The theorem requires the system to contain analytic functions of $\vec{\phi}$, have analytic Cauchy data specified on an analytic, noncharacteristic hypersurface and be quasilinear. The hypersurface Γ in our case is the 1-dimensional hyperplane in $\mathbb{R}^2$ given by $t = 0$ on which we have the above boundary conditions. The hypersurface is noncharacteristic for the PDE if the coefficient of the highest order derivative in t (in our case the coefficient of $\frac{\partial\phi}{\partial t}$) is nonzero for all values of its arguments. Because the coefficient of this term is 1 in our case, this condition is trivially satisfied. We need to make two adjustments to the system to transform it to a quasilinear system with analytic Cauchy data that is identically 0 on Γ.

Step 1: Transforming to Quasilinear System

Folland's treatment of the Cauchy-Kovalevskaya theorem in *Introduction to Partial Differential Equations* [2] contains a method for transforming a fully nonlinear system to a quasilinear system, after which we will revert back to the set up of Evans' proof. The goal is to end up with the system in the form:

$$\partial_t \vec{Y} = A_\theta(\theta, t, \vec{Y})\partial_\theta \vec{Y} + B(\theta, t, \vec{Y})$$

$$\vec{Y}(\theta, 0) = \vec{\Phi}(\theta)$$

where Y, B and Φ are vector valued, A is matrix valued and where the system is now quasilinear in Y's entries. The full text of the transformation method is given in Appendix D. In our case, α (the multi-index in the spatial variables used in Folland) is a single index (our only spatial variable is θ), j is the order of derivatives in t and $k = 1$ (first order system). This gives the following entries for $\vec{Y}$:

$$\vec{Y} = \begin{pmatrix} \vec{y_{00}} \\ \vec{y_{10}} \\ \vec{y_{01}} \end{pmatrix} = \begin{pmatrix} \phi^1 \\ \phi^2 \\ \frac{\partial \phi^1}{\partial \theta} \\ \frac{\partial \phi^2}{\partial \theta} \\ \frac{\partial \phi^1}{\partial t} \\ \frac{\partial \phi^2}{\partial t} \end{pmatrix}$$

which will make the system above six dimensional:

$$\begin{pmatrix} \partial_t \phi^1 \\ \partial_t \phi^2 \\ \partial^2_{\theta t} \phi^1 \\ \partial^2_{\theta t} \phi^2 \\ \partial^2_{tt} \phi^1 \\ \partial^2_{tt} \phi^2 \end{pmatrix} = A_\theta(\theta, t, \vec{Y}) \begin{pmatrix} \partial_\theta \phi^1 \\ \partial_\theta \phi^2 \\ \partial^2_{\theta\theta} \phi^1 \\ \partial^2_{\theta\theta} \phi^2 \\ \partial^2_{\theta t} \phi^1 \\ \partial^2_{\theta t} \phi^2 \end{pmatrix} + B(\theta, t, \vec{Y})$$

with initial conditions:

$$y_{00}(\theta, 0) = \phi(\theta, t) = (\cos\theta, \sin\theta - k)$$

$$y_{10}(\theta, 0) = \frac{\partial \phi}{\partial \theta}(\theta, 0) = \frac{\partial \phi_0}{\partial \theta}(\theta) = (-\sin\theta, \cos\theta)$$

$$y_{01}(\theta, 0) = \frac{\partial \phi}{\partial t}(\theta, 0) = \left(-2\phi_0^1 + \frac{2\phi_0^1(\frac{\partial \phi_0^1}{\partial \theta})^2 + 2\phi_0^2 \frac{\partial \phi_0^1}{\partial \theta}\frac{\partial \phi_0^2}{\partial \theta}}{|\frac{\partial \phi_0}{\partial \theta}|^2}, -2\phi_0^2 + \frac{2\phi_0^1 \frac{\partial \phi_0^1}{\partial \theta}\frac{\partial \phi_0^2}{\partial \theta} + 2\phi_0^2(\frac{\partial \phi_0^2}{\partial \theta})^2}{|\frac{\partial \phi_0}{\partial \theta}|^2} \right)$$

$$= (-2\cos\theta + 2\cos\theta(-\sin\theta)^2 + 2(\sin\theta - k)(-\sin\theta)(\cos\theta),$$

$$-2(\sin\theta - k) + 2\cos\theta(-\sin\theta)(\cos\theta) + 2(\sin\theta - k)(\cos\theta)^2)$$

$$= (-2\cos\theta + 2\sin^2\theta\cos\theta - 2\sin^2\theta\cos\theta + 2k\sin\theta\cos\theta,$$

$$-2\sin\theta + 2k - 2\sin\theta\cos^2\theta + 2\sin\theta\cos^2\theta - 2k\cos^2\theta)$$

$$= \big(-2\cos\theta + 2k\sin\theta\cos\theta, -2\sin\theta + 2k - 2k\cos^2\theta\big)$$

We need to compute the $A_\theta(\theta, t, \vec{Y})$ matrix and the $B(\theta, t, \vec{Y})$ vector, which require computing $\frac{\partial^2\phi^1}{\partial\theta\partial t}, \frac{\partial^2\phi^2}{\partial\theta\partial t}, \frac{\partial^2\phi^1}{\partial t^2}$ and $\frac{\partial^2\phi^2}{\partial t^2}$ (the first two components of $\partial_t\vec{Y}$, or $\langle\frac{\partial\phi^1}{\partial t}, \frac{\partial\phi^2}{\partial t}\rangle$, are given by the original system):

$$1) \qquad \frac{\partial^2\phi^1}{\partial\theta\partial t} = \partial_\theta\bigg(-2\phi^1 + \frac{2\phi^1(\frac{\partial\phi^1}{\partial\theta})^2 + 2\phi^2\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2}\bigg)$$

$$= -2\frac{\partial\phi^1}{\partial\theta} + \frac{2|\frac{\partial\phi}{\partial\theta}|^2\big(\frac{\partial\phi^1}{\partial\theta}(\frac{\partial\phi^1}{\partial\theta})^2 + 2\phi^1\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2}\big) - 2\phi^1(\frac{\partial\phi^1}{\partial\theta})^2\big(2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2} + 2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta^2}\big)}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$+\frac{2|\frac{\partial\phi}{\partial\theta}|^2\big(\frac{\partial\phi^2}{\partial\theta}\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta} + \phi^2(\frac{\partial^2\phi^1}{\partial^2\theta}\frac{\partial\phi^2}{\partial\theta} + \frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^2}{\partial^2\theta})\big) - 2\phi^2\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta}(2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2} + 2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta^2})}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$= -2\frac{\partial\phi^1}{\partial\theta} + \frac{2|\frac{\partial\phi}{\partial\theta}|^2\big((\frac{\partial\phi^1}{\partial\theta})^3 + 2\phi^1\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2}\big) - 2\phi^1(\frac{\partial\phi^1}{\partial\theta})^2\big(2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2} + 2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta^2}\big)}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$+\frac{2|\frac{\partial\phi}{\partial\theta}|^2\big((\frac{\partial\phi^2}{\partial\theta})^2\frac{\partial\phi^1}{\partial\theta} + \phi^2(\frac{\partial^2\phi^1}{\partial^2\theta}\frac{\partial\phi^2}{\partial\theta} + \frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^2}{\partial^2\theta})\big) - 2\phi^2\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta}(2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2} + 2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta^2})}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$2) \qquad \frac{\partial^2\phi^2}{\partial\theta\partial t} = \partial_\theta\bigg(-2\phi^2 + \frac{2\phi^1\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta} + 2\phi^2(\frac{\partial\phi^2}{\partial\theta})^2}{|\frac{\partial\phi}{\partial\theta}|^2}\bigg)$$

$$= -2\frac{\partial\phi^2}{\partial\theta} + \frac{2|\frac{\partial\phi}{\partial\theta}|^2\big((\frac{\partial\phi^2}{\partial\theta})^3 + 2\phi^2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta^2}\big) - 2\phi^2(\frac{\partial\phi^2}{\partial\theta})^2\big(2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta^2} + 2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2}\big)}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$+\frac{2|\frac{\partial\phi}{\partial\theta}|^2\big((\frac{\partial\phi^1}{\partial\theta})^2\frac{\partial\phi^2}{\partial\theta} + \phi^1(\frac{\partial^2\phi^2}{\partial^2\theta}\frac{\partial\phi^1}{\partial\theta} + \frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2})\big) - 2\phi^1\frac{\partial\phi^2}{\partial\theta}\frac{\partial\phi^1}{\partial\theta}(2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta^2} + 2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta^2})}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$3) \qquad \frac{\partial^2\phi^1}{\partial t^2} = \partial_t\bigg(-2\phi^1 + \frac{2\phi^1(\frac{\partial\phi^1}{\partial\theta})^2 + 2\phi^2\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2}\bigg)$$

$$= -2\frac{\partial\phi^1}{\partial t} + \frac{2|\frac{\partial\phi}{\partial\theta}|^2\big(\frac{\partial\phi^1}{\partial t}(\frac{\partial\phi^1}{\partial\theta})^2 + 2\phi^1\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta\partial t}\big) - 2\phi^1(\frac{\partial\phi^1}{\partial\theta})^2\big(2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta\partial t} + 2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta\partial t}\big)}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$+\frac{2|\frac{\partial\phi}{\partial\theta}|^2\big(\frac{\partial\phi^2}{\partial t}\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta} + \phi^2(\frac{\partial^2\phi^1}{\partial\theta\partial t}\frac{\partial\phi^2}{\partial\theta} + \frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta\partial t})\big) - 2\phi^2\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta}(2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta\partial t} + 2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta\partial t})}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$4) \qquad \frac{\partial^2\phi^1}{\partial t^2} = \partial_t\bigg(-2\phi^2 + \frac{2\phi^1\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta} + 2\phi^2(\frac{\partial\phi^2}{\partial\theta})^2}{|\frac{\partial\phi}{\partial\theta}|^2}\bigg)$$

$$= -2\frac{\partial\phi^2}{\partial t} + \frac{2|\frac{\partial\phi}{\partial\theta}|^2\big(\frac{\partial\phi^2}{\partial t}(\frac{\partial\phi^2}{\partial\theta})^2 + 2\phi^2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta\partial t}\big) - 2\phi^2(\frac{\partial\phi^2}{\partial\theta})^2\big(2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta\partial t} + 2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta\partial t}\big)}{|\frac{\partial\phi}{\partial\theta}|^4}$$

$$+\frac{2|\frac{\partial\phi}{\partial\theta}|^2\left(\frac{\partial\phi^1}{\partial t}\frac{\partial\phi^2}{\partial\theta}\frac{\partial\phi^1}{\partial\theta} + \phi^1(\frac{\partial^2\phi^2}{\partial\theta\partial t}\frac{\partial\phi^1}{\partial\theta} + \frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta\partial t})\right) - 2\phi^1\frac{\partial\phi^2}{\partial\theta}\frac{\partial\phi^1}{\partial\theta}(2\frac{\partial\phi^2}{\partial\theta}\frac{\partial^2\phi^2}{\partial\theta\partial t} + 2\frac{\partial\phi^1}{\partial\theta}\frac{\partial^2\phi^1}{\partial\theta\partial t})}{|\frac{\partial\phi}{\partial\theta}|^4}$$

Our matrix system then becomes:

$$\begin{pmatrix} \partial_t\phi^1 \\ \partial_t\phi^2 \\ \partial^2_{\theta t}\phi^1 \\ \partial^2_{\theta t}\phi^2 \\ \partial^2_{tt}\phi^1 \\ \partial^2_{tt}\phi^2 \end{pmatrix} = A_\theta(\theta, t, \vec{Y}) \begin{pmatrix} \partial_\theta\phi^1 \\ \partial_\theta\phi^2 \\ \partial^2_{\theta\theta}\phi^1 \\ \partial^2_{\theta\theta}\phi^2 \\ \partial^2_{\theta t}\phi^1 \\ \partial^2_{\theta t}\phi^2 \end{pmatrix} + B(\theta, t, \vec{Y})$$

$$= A(\theta, t, Y) \begin{pmatrix} \partial_\theta\phi^1 \\ \partial_\theta\phi^2 \\ \partial^2_{\theta\theta}\phi^1 \\ \partial^2_{\theta\theta}\phi^2 \\ \partial^2_{\theta t}\phi^1 \\ \partial^2_{\theta t}\phi^2 \end{pmatrix} + \begin{pmatrix} -2\phi^1 \\ -2\phi^2 \\ 0 \\ 0 \\ -2\frac{\partial\phi^1}{\partial t} \\ -2\frac{\partial\phi^2}{\partial t} \end{pmatrix}$$

The rows of the A matrix are given by:

$$A_{1,i} = \left(\begin{array}{cccccc} \frac{2\phi^1 \frac{\partial\phi^1}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} & \frac{2\phi^2 \frac{\partial\phi^1}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} & 0 & 0 & 0 & 0 \end{array} \right)$$

$$A_{2,i} = \left(\begin{array}{cccccc} \frac{2\phi^1 \frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} & \frac{2\phi^2 \frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} & 0 & 0 & 0 & 0 \end{array} \right)$$

$$A_{3,i} = \left(\begin{array}{ccc} -2 + \frac{2(\frac{\partial\phi^1}{\partial\theta})^2 + 2(\frac{\partial\phi^2}{\partial\theta})^2}{|\frac{\partial\phi}{\partial\theta}|^2} & 0 & \frac{4\phi^1 \frac{\partial\phi^1}{\partial\theta} + 2\phi^2 \frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} + \frac{-4\phi^1(\frac{\partial\phi^1}{\partial\theta})^3 - 4\phi^2(\frac{\partial\phi^1}{\partial\theta})^2 \frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^4} \end{array} \right)$$

$$... \left(\begin{array}{ccc} \frac{-4\phi^1(\frac{\partial\phi^1}{\partial\theta})^2 \frac{\partial\phi^2}{\partial\theta} - 4\phi^2 \frac{\partial\phi^1}{\partial\theta}(\frac{\partial\phi^2}{\partial\theta})^2}{|\frac{\partial\phi}{\partial\theta}|^4} + \frac{2\phi^2 \frac{\partial\phi^1}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} & 0 & 0 \end{array} \right)$$

$$A_{4,i} = \left(\begin{array}{ccc} 0 & -2 + \frac{2(\frac{\partial\phi^2}{\partial\theta})^2 + 2(\frac{\partial\phi^1}{\partial\theta})^2}{|\frac{\partial\phi^2}{\partial\theta}|^2} & \frac{-4\phi^2(\frac{\partial\phi^2}{\partial\theta})^2 \frac{\partial\phi^1}{\partial\theta} - 4\phi^1 \frac{\partial\phi^2}{\partial\theta}(\frac{\partial\phi^1}{\partial\theta})^2}{|\frac{\partial\phi}{\partial\theta}|^4} + \frac{2\phi^1 \frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} \end{array} \right)$$

$$... \left(\begin{array}{ccc} \frac{4\phi^2 \frac{\partial\phi^2}{\partial\theta} + 2\phi^1 \frac{\partial\phi^1}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} + \frac{-4\phi^2(\frac{\partial\phi^2}{\partial\theta})^3 - 4\phi^1(\frac{\partial\phi^2}{\partial\theta})^2 \frac{\partial\phi^1}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^4} & 0 & 0 \end{array} \right)$$

$$A_{5,i} = \left(\begin{array}{ccc} \frac{2\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^1}{\partial t} + 2\frac{\partial\phi^2}{\partial\theta}\frac{\partial\phi^2}{\partial t}}{|\frac{\partial\phi}{\partial\theta}|^2} & 0 & 0 \end{array} \right)$$

$$... \left(\begin{array}{ccc} 0 & \frac{4\phi^1 \frac{\partial\phi^1}{\partial\theta} + 2\phi^2 \frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} + \frac{-4\phi^1(\frac{\partial\phi^1}{\partial\theta})^3 - 4\phi^2(\frac{\partial\phi^1}{\partial\theta})^2 \frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^4} & \frac{-4\phi^1(\frac{\partial\phi^1}{\partial\theta})^2 \frac{\partial\phi^2}{\partial\theta} - 4\phi^2 \frac{\partial\phi^1}{\partial\theta}(\frac{\partial\phi^2}{\partial\theta})^2}{|\frac{\partial\phi}{\partial\theta}|^4} + \frac{2\phi^2 \frac{\partial\phi^1}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} \end{array} \right)$$

$$A_{6,i} = \left(\begin{array}{ccc} 0 & \frac{2\frac{\partial\phi^2}{\partial\theta}\frac{\partial\phi^2}{\partial t} + 2\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^1}{\partial t}}{|\frac{\partial\phi^2}{\partial\theta}|^2} & 0 \end{array} \right)$$

$$... \left(\begin{array}{cc} \frac{-4\phi^2(\frac{\partial\phi^2}{\partial\theta})^2 \frac{\partial\phi^1}{\partial\theta} - 4\phi^1 \frac{\partial\phi^2}{\partial\theta}(\frac{\partial\phi^1}{\partial\theta})^2}{|\frac{\partial\phi}{\partial\theta}|^4} + \frac{2\phi^1 \frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} & \frac{4\phi^2 \frac{\partial\phi^2}{\partial\theta} + 2\phi^1 \frac{\partial\phi^1}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2} + \frac{-4\phi^2(\frac{\partial\phi^2}{\partial\theta})^3 - 4\phi^1(\frac{\partial\phi^2}{\partial\theta})^2 \frac{\partial\phi^1}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^4} \end{array} \right)$$

Now our system is first order quasilinear in the entries of $\vec{Y}$ (i.e. the components of the A matrix and B vector are explicit analytic functions in $\vec{Y}$'s entries).

Step 2: Transforming System to have Analytic Cauchy Data that is Identically 0 on Γ

Cauchy data for a k-th order quasilinear system consists of derivatives up to order $k-1$ of the solution defined on the boundary Γ in the normal direction (to Γ at each $x \in \Gamma$). Evans' [1] definition of Cauchy data is given in Appendix D. In our case $\Gamma = t = 0$, but because $k = 1$, the Cauchy data consists only of initial data for the system defined at $t = 0$. The set up requires Cauchy data to be identically 0 on Γ, or $\vec{Y}|_{t=0} = 0$, but we know from the computation of the vector of initial conditions above that this is not the case. We can adjust the system by defining

$$\Psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \\ \psi^5 \\ \psi^6 \end{pmatrix} = Y - Y_0 = \begin{pmatrix} \phi^1 - \phi^1_0 \\ \phi^2 - \phi^2_0 \\ \frac{\partial \phi^1}{\partial \theta} - \frac{\partial \phi^1}{\partial \theta}|_{t=0} \\ \frac{\partial \phi^2}{\partial \theta} - \frac{\partial \phi^2}{\partial \theta}|_{t=0} \\ \frac{\partial \phi^1}{\partial t} - \frac{\partial \phi^1}{\partial t}|_{t=0} \\ \frac{\partial \phi^2}{\partial t} - \frac{\partial \phi^1}{\partial t}|_{t=0} \end{pmatrix} = \begin{pmatrix} \phi^1 - \cos\theta \\ \phi^2 - (\sin\theta - k) \\ \frac{\partial \phi^1}{\partial \theta} + \sin\theta \\ \frac{\partial \phi^2}{\partial \theta} - \cos\theta \\ \frac{\partial \phi^1}{\partial t} - (-2\cos\theta + 2k\sin\theta\cos\theta) \\ \frac{\partial \phi^2}{\partial t} - (-2\sin\theta + 2k - 2k\cos^2\theta) \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \phi^1 \\ \phi^2 \\ \frac{\partial \phi^1}{\partial \theta} \\ \frac{\partial \phi^2}{\partial \theta} \\ \frac{\partial \phi^1}{\partial t} \\ \frac{\partial \phi^2}{\partial t} \end{pmatrix} = \begin{pmatrix} \psi^1 + \cos\theta \\ \psi^2 + \sin\theta - k \\ \psi^3 - \sin\theta \\ \psi^4 + \cos\theta \\ \psi^5 - 2\cos\theta + 2k\sin\theta\cos\theta \\ \psi^6 - 2\sin\theta + 2k - 2k\cos^2\theta \end{pmatrix}$$

With this substitution in the system from Step 1, it is clear that we will have a quasilinear first order system in the entries of Ψ with Cauchy data that is identically zero on Γ ($\Psi_{t=0} = 0$). Note: to align with Evans' proof we have $n = 2$ (number of domain variables ($(x_1, x_2) = (\theta, t)$), $m = 6$ (number of Ψ components) and $\Psi = 0$ when $t = 0$ for all values of θ. With this second adjustment our system is:

$$\partial_t \vec{\Psi} = \tilde{A}_\theta(\theta, t, \vec{\Psi}) \partial_\theta \vec{\Psi} + \tilde{B}(\theta, t, \vec{\Psi})$$

$$\vec{\Psi}(\theta, 0) = 0$$

where the $\tilde{A}$ matrix and $\tilde{B}$ vector contain analytic functions in the entries of Ψ, θ and t.

Step 3: Applying Cauchy-Kovalevskaya Theorem

Our system now satisfies the conditions of the Cauchy-Kovalevskaya Theorem as written in Evans, where the radius of convergence r in the (θ, t) domain ((15) in Appendix D) will be extracted from the proof. The first part of the proof demonstrates how to compute each term of the power series solution:

$$\vec{\Psi} = \sum_\alpha \Psi_\alpha (\theta, t)^\alpha = \sum_\alpha \frac{D^\alpha \psi(0)}{\alpha!} (\theta, t)^\alpha$$

Note: this will technically be a six component vector solution, each component of which is a power series. We are only interested in the first two components. Step 2 of Evans' proof (Appendix D) finds a radius of convergence on which every function in both the $\tilde{A}$ matrix and $\tilde{B}$ vector converges. Consider the entry $\tilde{A}_{11}$ under the change of variables made in Step 2 above:

$$\tilde{A}_{11} = \frac{2\phi^1 \frac{\partial \phi^1}{\partial \theta}}{|\frac{\partial \phi}{\partial \theta}|^2} = \frac{2(\psi^1 + \cos\theta)(\psi^3 - \sin\theta)}{(\psi^3 - \sin\theta)^2 + (\psi^4 + \cos\theta)^2}$$

$$= \frac{2(\psi^1\psi^3 - \psi^1 \sin\theta + \psi^3 \cos\theta - \cos\theta \sin\theta)}{(\psi^3 - \sin\theta)^2 + (\psi^4 + \cos\theta)^2}$$

Because this is a rational function in the components of Ψ with trigonometric coefficients, it is analytic for all ψ, θ in its domain. To determine a $|\psi|$ range on which the denominator is nonzero we have:

$$(\psi^3 - \sin\theta)^2 + (\psi^4 + \cos\theta)^2 = (\psi^3)^2 - 2\psi^3 \sin\theta + \sin^2\theta + (\psi^4)^2 + 2\psi^4 \cos\theta + \cos^2\theta$$

$$= 1 + (\psi^3)^2 + (\psi^4)^2 + 2(\psi^4 \cos\theta - \psi^3 \sin\theta)$$

This will be greater than zero when $1 + (\psi^3)^2 + (\psi^4)^2 > -2(\psi^4 \cos\theta - \psi^3 \sin\theta)$ and in particular when $1 > -2(\psi^4 \cos\theta - \psi^3 \sin\theta)$. But because $-4|\psi| < 2(\psi^4 \cos\theta - \psi^3 \sin\theta) \Rightarrow -2(\psi^4 \cos\theta - \psi^3 \sin\theta) < 4|\psi|$, fixing $0 < |\psi| < \frac{1}{4}$ guarantees this condition.
We can see that all entries of $\tilde{A}$ and $\tilde{B}$ are similar rational functions with either $|\frac{\partial\phi}{\partial\theta}|^2$ or $|\frac{\partial\phi}{\partial\theta}|^4$ in the denominator and therefore their power series also converge on the same $|\psi|$ range. In particular we can say the component functions are analytic when $|\Psi| < \frac{1}{4}$.

To mimic Step 2 of Evans' proof, we can also say the component functions of the $\tilde{A}$ matrix are analytic when $|\psi| + |\theta| < \frac{1}{4} = s$. But in reality the component functions are analytic whenever $|\psi| < 1/4$, for all θ. Steps 4 through 6 in Evans' proof show that the power series solution converges by constructing a new system (with a $\tilde{A}^*$ matrix and $\tilde{B}^*$ vector in our setup), whose power series solution Ψ^* converges and "majorizes" the original. A definition for "majorize" is given in Appendix D. This proves that the original power series solution converges by Lemma (i) (Appendix D- first part of Lemma entitled Majorants)).

To get a system that majorizes the original, we know that the power series expansion of all component functions $\tilde{A}_{ij}, \tilde{B}_q$ for $(0 < i, j, q < 6)$ converge for (Ψ, θ) such that $|(\Psi, \theta)| < \frac{s}{2} = \frac{1}{8}$. This is by the fact that

$$|(\Psi, \theta)| < \frac{s}{2} \Rightarrow s \geq 2|(\Psi, \theta)| = 2\sqrt{(\psi^1)^2 + \cdots + (\psi^6)^2 + \theta^2} \geq \sqrt{(\psi^1)^2 + \cdots + (\psi^6)^2} + |\theta|$$

$$\geq |\Psi| + |\theta|$$

Again this is done to exactly match Evans' set up in Lemma (ii) in Appendix D. Choosing r such that $0 \leq r\sqrt{7} \leq \frac{1}{8}$ (or $r \leq \frac{1}{8\sqrt{7}}$) we can apply Lemma (ii) to each component function of the system to get a majorizing system:

$$\partial_t \vec{\Psi}^* = \tilde{A}^*_\theta(\theta, t, \vec{\Psi})\partial_\theta \vec{\Psi}^* + \tilde{B}^*(\theta, t, \vec{\Psi})$$

where each component of both the $\tilde{A}^*$ matrix and $\tilde{B}^*$ vector is:

$$\tilde{A}^*_{ij} = \tilde{B}^*_q = \frac{Cr}{r - (\psi^1 + \cdots + \psi^6) - (\theta)}$$

for $0 \leq i, j, q \leq 6$. Note that this system will majorize the original for $|(\psi, \theta)| \leq \frac{r}{\sqrt{7}}$ so in particular it will majorize the original for $|\Psi| + |\theta| < \frac{r}{\sqrt{7}}$ ($|\theta|$ can range between 0 and $\frac{r}{\sqrt{7}}$). This uses the fact that $|(\Psi, \theta)| < |\Psi| + |\theta|$, written to match the last line of Evans 243, according to the proof of Lemma (ii). Additionally C is chosen to be

$$C = \max\{C_{ij}, C_q\}$$

$0 \leq i, j, q \leq 6$ where the C_{ij}, C_q's are chosen as in the proof of Lemma (ii) for each of the component functions in $\tilde{A}^*$ and $\tilde{B}^*$. The majorizing system is then (by Evans page 244):

$$\partial_t \vec{\Psi}^* = \frac{Cr}{r - (\psi^{1*} + \cdots + \psi^{*6}) - \theta}\Big(\sum_l \Psi^{l*}_\theta + 1\Big)$$

for $|(\theta, t)| < \frac{r}{\sqrt{7}}$ with solution:

$$\Psi^* = v^*(1, 1, 1, 1, 1, 1)$$

$$v^*(\theta, t) = \frac{1}{12}\left(\frac{r}{\sqrt{7}} - \theta - \sqrt{(\frac{r}{\sqrt{7}} - \theta)^2 - 24C\frac{r}{\sqrt{7}}t}\right)$$

which is analytic (for $|(\theta,t)| < \frac{r}{\sqrt{7}}$) and for (θ,t) such that the quantity under the square root is positive. This gives an additional constraint on (θ,t):

$$(\frac{r}{\sqrt{7}}-\theta)^2 > 24C\frac{r}{\sqrt{7}}t \Rightarrow t < \frac{\sqrt{7}(\frac{r}{\sqrt{7}}-\theta)^2}{24C}$$

Because $0 < |\theta| < \frac{r}{\sqrt{7}}$, t approaches 0 as $|\theta|$ approaches $\frac{r}{\sqrt{7}}$. Taking a compact interval around $\theta = 0$ gives a region on which solutions exist for positive time. For example we can restrict the interval to $|\theta| < \frac{r}{7}$, with the smallest value of t occurring at $\theta = \frac{r}{7}$. This gives a lower bound for t in this compact interval around $\theta = 0$. We can get a similar interval for all $\theta \in [0, 2\pi]$, and because S^1 is compact we can take the minimum time interval across the circle for short time existence of the solution everywhere.
We have therefore shown with the application of the Cauchy-Kovalevskaya Theorem that for the vector field constructed through projecting the Case 1 gradient onto normal directions, a solution for the flow (towards the origin) exists for short time.

Step 4: Special Case k=0

Recall that the original system

$$\frac{\partial\phi(t,\theta)}{\partial t} = \left(\frac{\partial\phi^1(t,\theta)}{\partial t}, \frac{\partial\phi^2(t,\theta)}{\partial t}\right)$$

$$= \left(-2\phi^1 + \frac{2\phi^1(\frac{\partial\phi^1}{\partial\theta})^2 + 2\phi^2\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta}}{|\frac{\partial\phi}{\partial\theta}|^2}, -2\phi^2 + \frac{2\phi^1\frac{\partial\phi^1}{\partial\theta}\frac{\partial\phi^2}{\partial\theta} + 2\phi^2(\frac{\partial\phi^2}{\partial\theta})^2}{|\frac{\partial\phi}{\partial\theta}|^2}\right)$$

had initial conditions: $\phi(\theta,0) = (\cos\theta, \sin\theta - k)$. Setting $k = 0$ means our initial embedding is the unit circle centered at the origin.

We can compute the solution by computing the power series given in the Cauchy-

Kovalevskaya theorem. We will compute:

$$(\phi^1(\theta,t),\phi^2(\theta,t)) = (\sum_{\alpha}\phi^1_{\alpha}(\theta,t)^{\alpha},\sum_{\alpha}\phi^2_{\alpha}(\theta,t)^{\alpha})$$

where $\phi^i_{\alpha} = \frac{D^{\alpha}\phi^1(0)}{\alpha!}$. We can explicitly compute the ϕ^1 series as follows:

To have Cauchy data that is identically zero when $t = 0$ we make a similar adjustment as in Step 2 above. We can define:

$$(\psi^1,\psi^2) = (\phi^1 - \phi^1|_{t=0}, \phi^2 - \phi^2|_{t=0}) = (\phi^1 - \cos\theta, \phi^2 - \sin\theta)$$

and we will therefore be computing Taylor coefficients in ψ around $(\theta, t) = (0, 0)$.

$\alpha = (n, 0)$ **terms (**n **derivatives in** θ **and 0 in** t**):**

$$\frac{D^{(n,0)}\psi^1(0)}{n!} = \frac{\partial^n\phi^1}{\partial\theta^n}(0,0) - \frac{\partial^n\phi^1_0}{\partial\theta^n}(0) = 0$$

$\alpha = (n, 1)$ **terms (**n **derivatives in** θ **and 1 in** t**):**

$$\frac{D^{(0,1)}\psi^1(0)}{1!} = \frac{\partial\phi^1}{\partial t}(0,0) - \frac{\partial\phi^1_0}{\partial t}(0) = \frac{\partial\phi^1}{\partial t}(0,0) = -2\phi^1_0(0) = -2\cos(\theta)|_0$$

Note: If we just consider summing the power series solution over the $\alpha = (n, 1)$ terms we have:

$$\sum_n \frac{D^{(n,1)}\psi^1(0)}{n!1!}\theta^n t = \frac{t}{1!}\sum_n \frac{D^{(n,1)}\psi^1(0)}{n!}\theta^n = \frac{t}{1!}\sum_{n=0}^{\infty} -2\frac{D^{(n)}(\cos\theta)}{n!}(0)\theta^n$$

where the summation is the Taylor series around $\theta = 0$ of $-2\cos\theta$ so that the last term above is simply equal to $\frac{t}{1!}(-2\cos\theta)$.

$\alpha = (n, 2)$ terms (n derivatives in θ and 2 in t):

$$\frac{D^{(0,2)}\psi^1(0)}{2!} = \frac{1}{2!}\big(\frac{\partial^2\phi^1}{\partial^2 t}(0,0) - \frac{\partial^2\phi_0^1}{\partial^2 t}(0)\big) = \frac{1}{2!}\big(\frac{\partial^2\phi^1}{\partial^2 t}(0,0)\big) = \frac{-2}{2!}\frac{\partial\phi^1}{\partial t}(0,0)$$
$$= \frac{-2}{2!}(-2\cos(\theta)|_0) = \frac{4}{2!}\cos\theta$$

Again, if summing the power series solution only over the $\alpha = (n, 2)$ terms we have:

$$\sum_n \frac{D^{(n,2)}\psi^1(0)}{n!2!}\theta^n t^2 = \frac{t^2}{2!}\sum_n \frac{D^{(n,2)}\psi^1(0)}{n!}\theta^n = \frac{t^2}{2!}\sum_{n=0}^{\infty} 4\frac{D^{(n)}(\cos\theta)}{n!}(0)\theta^n$$

where the last summation is the Taylor series around $\theta = 0$ of $4\cos\theta$ so the right hand term equals $\frac{t^2}{2!}(4\cos\theta)$.

It is clear from here that summing the power series solution over all $\alpha = (n, m)$ is:

$$\psi^1(\theta, t) = \sum_\alpha \frac{D^\alpha\phi^1(0)}{\alpha!} = \sum_{m=1}^{\infty}\sum_{n=0}^{\infty}\frac{D^{(n,m)}\phi^1(0)}{n!m!} = \sum_{m=1}^{\infty}\frac{t^m}{m!}(-2)^m\cos\theta$$

$$= \cos\theta\sum_{m=1}^{\infty}\frac{(-2t)^m}{t!} = (e^{-2t} - 1)\cos\theta$$

$$\Rightarrow \psi^1(\theta, t) = e^{-2t}\cos\theta - \cos\theta = \phi^1 - \phi^1_{t=0} = \phi^1 - \cos\theta$$

$$\Rightarrow \phi^1(\theta, t) = e^{-2t}\cos\theta$$

.

We can similarly compute that $\phi^2(\theta, t) = e^{-2t}\sin\theta$ giving a final solution to the system:

$$\phi(\theta, t) = (e^{-2t}\cos\theta, e^{-2t}\sin\theta)$$

Note that this solution matches the solution in Case 1 (holding the volume element constant in gradient computation) with $k = 0$, which suggests that the system in this special case reduces to $\phi(\theta,t) = (-2\phi^1, -2\phi^2)$ for all t. In this case the correction term that projects $(-2\phi^1, -2\phi^2)$ onto normal directions vanishes because the vector field is already normal to the initial embedding when $k = 0$.

5.3 CASE 3: Gradient flow with varying volume element

Finally, we want to compute the gradient of the distance penalty function:

$$P(\phi) = \int_{\phi(S^1)} d^2(\phi(m), x_i)\mathrm{dvol}_{\phi(S^1)}$$

now allowing the volume element to vary. This means that the volume form is the induced volume form on $\phi(S^1)$. The gradient vector field for the general case is then given by (see Appendix C):

$$\mathrm{grad}P(\phi)(m) = 2\phi(m) - d^2(\phi(m), x_i)\mathrm{Tr}II - 2\sum_{r,l=1}^{k}\sum_{j=1}^{N} g^{rl}\phi(m)^j \frac{\partial \phi(m)^j}{\partial y^l}\partial_r$$

where the first and third term were adapted to our example in Case 2 and the middle term becomes:

Middle Term:

$$d^2(\phi(m), x_i)\mathrm{Tr}II = |\phi|^2 B_{\phi(m)}(\partial_\theta, \partial_\theta) = |\phi|^2 \left(\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta\right)^N$$

where $\left(\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta\right)^N$ refers to taking the normal component of the given vector field $\left(\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta\right)^N$ and $B : T_{\phi(m)}\phi(S^1) \times T_{\phi(m)}\phi(S^1) \to N_{\phi(m)}\phi(S^1)$ is the trace of the second fundamental form (bilinear) (See Appendix A for definition). Because the trace acts on an orthonormal field (and our $\partial_\theta = \phi_*(\partial_\theta^{S^1})$) we must actually compute:

$$|\phi|^2(\nabla^{\mathbb{R}^2}_{\frac{\partial_\theta}{|\partial_\theta|}} \frac{\partial_\theta}{|\partial_\theta|})^N = |\phi|^2(\frac{1}{|\partial_\theta|}\nabla^{\mathbb{R}^2}_{\partial_\theta}\frac{\partial_\theta}{|\partial_\theta|})^N = |\phi|^2\left(\frac{1}{|\partial_\theta|}[\frac{1}{|\partial_\theta|}\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta + d(\frac{1}{|\partial_\theta|})(\partial_\theta)\partial_\theta]\right)^N \quad (5.2)$$

$$= |\phi|^2\left(\frac{1}{|\partial_\theta|^2}\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta\right)^N = |\phi|^2\left(\frac{1}{|\frac{\partial\phi}{\partial\theta}|^2}\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta\right)^N$$

where the second to last term above is from the fact that we are projecting a tangential vector into the normal space. Considering now only $\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta$ we have:

$$\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta = \nabla^{\mathbb{R}^2}_{\partial_\theta}\left(\frac{\partial\phi^1}{\partial\theta}\partial_x + \frac{\partial\phi^2}{\partial\theta}\partial_y\right)$$

$$= \frac{\partial\phi^1}{\partial\theta}\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_x + d(\frac{\partial\phi^1}{\partial\theta})(\partial_\theta)\partial_x + \frac{\partial\phi^2}{\partial\theta}\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_y + d(\frac{\partial\phi^2}{\partial\theta})(\partial_\theta)\partial_y$$

$$= d(\frac{\partial\phi^1}{\partial\theta})(\partial_\theta)\partial_x + d(\frac{\partial\phi^2}{\partial\theta})(\partial_\theta)\partial_y \quad (5.3)$$

(the last equality is by the fact that $\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_x = \nabla^{\mathbb{R}^2}_{\left(\frac{\partial\phi^1}{\partial\theta}\partial_x + \frac{\partial\phi^2}{\partial\theta}\partial_y\right)}\partial_x = \frac{\partial\phi^1}{\partial\theta}\nabla^{\mathbb{R}^2}_{\partial_x}\partial_x + \frac{\partial\phi^2}{\partial\theta}\nabla^{\mathbb{R}^2}_{\partial_y}\partial_x = 0$ under the $\mathbb{R}^2$ metric.) Continuing from (5.3) we have:

$$d(\frac{\partial\phi^1}{\partial\theta})(\partial_\theta)\partial_x + d(\frac{\partial\phi^2}{\partial\theta})(\partial_\theta)\partial_y = \partial_\theta(\frac{\partial\phi^1}{\partial\theta})d\theta(\partial_\theta)\partial_x + \partial_\theta(\frac{\partial\phi^2}{\partial\theta})d\theta(\partial_\theta)\partial_y$$

$$= \frac{\partial^2\phi^1}{\partial\theta^2}\partial_x + \frac{\partial^2\phi^2}{\partial\theta^2}\partial_y$$

Substituting back into (5.2) for the final term we have:

$$|\phi|^2\left(\frac{1}{|\frac{\partial\phi}{\partial\theta}|^2}\nabla^{\mathbb{R}^2}_{\partial_\theta}\partial_\theta\right)^N = |\phi|\left(\frac{1}{|\frac{\partial\phi}{\partial\theta}|^2}(\frac{\partial^2\phi^1}{\partial\theta^2}\partial_x + \frac{\partial^2\phi^2}{\partial\theta^2}\partial_y)\right)^N$$

$$= |\phi|^2\mathrm{Proj}_{\mathrm{Normal}}\left(\frac{1}{|\frac{\partial\phi}{\partial\theta}|^2}\frac{\partial^2\phi^1}{\partial\theta^2}, \frac{1}{|\frac{\partial\phi}{\partial\theta}|^2}\frac{\partial^2\phi^2}{\partial\theta^2}\right)$$

$$= |\phi|^2 \frac{\left(\frac{1}{|\frac{\partial \phi}{\partial \theta}|^2}\frac{\partial^2 \phi^1}{\partial \theta^2}, \frac{1}{|\frac{\partial \phi}{\partial \theta}|^2}\frac{\partial^2 \phi^2}{\partial \theta^2}\right) \cdot \left(-\frac{\partial \phi^2}{\partial \theta}, \frac{\partial \phi^1}{\partial \theta}\right)}{\left(-\frac{\partial \phi^2}{\partial \theta}, \frac{\partial \phi^1}{\partial \theta}\right) \cdot \left(-\frac{\partial \phi^2}{\partial \theta}, \frac{\partial \phi^1}{\partial \theta}\right)} \left(-\frac{\partial \phi^2}{\partial \theta}, \frac{\partial \phi^1}{\partial \theta}\right)$$

$$= |\phi|^2 \frac{-\frac{1}{|\frac{\partial \phi}{\partial \theta}|^2}\frac{\partial^2 \phi^1}{\partial \theta^2}\frac{\partial \phi^2}{\partial \theta} + \frac{1}{|\frac{\partial \phi}{\partial \theta}|^2}\frac{\partial^2 \phi^2}{\partial \theta^2}\frac{\partial \phi^1}{\partial \theta}}{(\frac{\partial \phi^2}{\partial \theta})^2 + (\frac{\partial \phi^1}{\partial \theta})^2} \left(-\frac{\partial \phi^2}{\partial \theta}, \frac{\partial \phi^1}{\partial \theta}\right)$$

$$= |\phi|^2 \left(\frac{\frac{\partial^2 \phi^1}{\partial \theta^2}(\frac{\partial \phi^2}{\partial \theta})^2 - \frac{\partial^2 \phi^2}{\partial \theta^2}\frac{\partial \phi^1}{\partial \theta}\frac{\partial \phi^2}{\partial \theta}}{|\frac{\partial \phi}{\partial \theta}|^4}, \frac{-\frac{\partial^2 \phi^1}{\partial \theta^2}\frac{\partial \phi^2}{\partial \theta}\frac{\partial \phi^1}{\partial \theta} + \frac{\partial^2 \phi^2}{\partial \theta^2}(\frac{\partial \phi^1}{\partial \theta})^2}{|\frac{\partial \phi}{\partial \theta}|^4}\right)$$

Recall that the gradient equation is:

$$\frac{\partial \phi(t,\theta)}{\partial t} = \left(\frac{\partial \phi^1(t,\theta)}{\partial t}, \frac{\partial \phi^2(t,\theta)}{\partial t}\right) = -\text{grad}P(\phi)$$

Combining with the first and third term from Case 2 we have a final gradient vector field:

$$\frac{\partial \phi(t,\theta)}{\partial t} = \langle\frac{\partial \phi^1(t,\theta)}{\partial t}, \frac{\partial \phi^2(t,\theta)}{\partial t}\rangle = -\text{grad}P(\phi)$$

$$-\Bigg(2\phi^1 - |\phi|^2 \frac{\frac{\partial^2 \phi^1}{\partial \theta^2}(\frac{\partial \phi^2}{\partial \theta})^2 - \frac{\partial^2 \phi^2}{\partial \theta^2}\frac{\partial \phi^1}{\partial \theta}\frac{\partial \phi^2}{\partial \theta}}{|\frac{\partial \phi}{\partial \theta}|^4} + \frac{-2\phi^1(\frac{\partial \phi^1}{\partial \theta})^2 - 2\phi^2\frac{\partial \phi^1}{\partial \theta}\frac{\partial \phi^2}{\partial \theta}}{|\frac{\partial \phi}{\partial \theta}|^2},$$

$$2\phi^2 - |\phi|^2 \frac{-\frac{\partial^2 \phi^1}{\partial \theta^2}\frac{\partial \phi^2}{\partial \theta}\frac{\partial \phi^1}{\partial \theta} + \frac{\partial^2 \phi^2}{\partial \theta^2}(\frac{\partial \phi^1}{\partial \theta})^2}{|\frac{\partial \phi}{\partial \theta}|^4} + \frac{-2\phi^1\frac{\partial \phi^1}{\partial \theta}\frac{\partial \phi^2}{\partial \theta} - 2\phi^2(\frac{\partial \phi^2}{\partial \theta})^2}{|\frac{\partial \phi}{\partial \theta}|^2}\Bigg)$$

This is a second order nonlinear PDE with initial conditions given on a noncharacteristic hypersurface, which prevents the direct application of the Cauchy-Kovalevskaya theorem. It is left as a future direction of the project to investigate applying advanced PDE techniques to this second order system.

5.4 Generalizing to Simple Closed Curves

Instead of using the initial condition of embedding a circle centered on the y axis, we would like to consider the initial condition of starting with any simple closed curve. We will assume the curve is regular, so its tangent vector will never vanish. More precisely we want to investigate the same set up under the three cases of gradient computations with initial condition:

$$\phi_0(\theta) = (\phi^1(\theta, 0), \phi^2(\theta, 0))$$

where $\theta \in [0, 2\pi]$ and $(\phi^1(0,0), \phi^2(0,0)) = (\phi^1(2\pi, 0), \phi^2(2\pi, 0))$, but where this property does not hold for any other $a \neq b \in [0, 2\pi]$. Note that in the general case we are still embedding a one dimensional manifold M into $\mathbb{R}^2$ and denoting M's coordinate chart parameter as θ. Therefore $\frac{\partial}{\partial\theta}$'s pushforward under ϕ is still

$$\phi_*(\frac{\partial}{\partial\theta}) = \frac{\partial\phi^1}{\partial\theta}\partial_x + \frac{\partial\phi^2}{\partial\theta}\partial_y$$

Case 1: Holding the volume element constant in the gradient computation

Because the computation of the gradient terms does not depend on choice of initial conditions, holding the volume element constant still gives $\mathrm{grad}(\phi(m)) = 2(\phi(m) - x_i)$ and with the same set up as above:

$$\frac{\mathrm{d}}{\mathrm{dt}}\phi_t(\theta) = -\mathrm{grad}P(\phi) = -2\phi_t(\theta)$$

$$\phi_0(\theta) = (\phi^1(\theta, 0), \phi^2(\theta, 0)$$

so that the flow

$$\phi_t(\theta) = (\phi^1(\theta, 0)e^{-2t}, \phi^2(\theta, 0)e^{-2t})$$

satisfies the gradient expression.

Case 2: Projection of Case 1 Onto Normal Directions gradient written as a matrix system is the same as before, now with arbitrary initial conditions:

$$\begin{pmatrix} \frac{\partial \phi^1}{\partial t} \\ \frac{\partial \phi^2}{\partial t} \end{pmatrix} = \frac{2}{|\frac{\partial \phi}{\partial \theta}|^2} \begin{pmatrix} \phi^1 \frac{\partial \phi^1}{\partial \theta} & \phi^2 \frac{\partial \phi^1}{\partial \theta} \\ \phi^1 \frac{\partial \phi^2}{\partial \theta} & \phi^2 \frac{\partial \phi^2}{\partial \theta} \end{pmatrix} \begin{pmatrix} \frac{\partial \phi^1}{\partial \theta} \\ \frac{\partial \phi^2}{\partial \theta} \end{pmatrix} - 2 \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix}$$

$$\phi(\theta, 0) = (\phi^1(\theta, 0), \phi^2(\theta, 0))$$

We made two major transformations to the system to get a form to which we could apply the Cauchy-Kovalevskaya theorem. Changing the system into a quasilinear system did not rely on choice of initial conditions (y_{00}, y_{10}, y_{01} initial data can be left as arbitrary with no effect on the process), nor did the change of variables performed to get analytic Cauchy data. In Step 3 of this case, applying the theorem required getting a radius of convergence for the entries of the $\tilde{A}$ and $\tilde{B}$ matrices, and more specifically finding a $|\psi|$ range on which each entry's denominators were non-zero. This computation was tailored to the specific initial data but can be generalized as follows:

Note that for general initial conditions we still use the same change of variables to obtain the $\tilde{A}, \tilde{B}$ matrices:

$$\Psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \\ \psi^5 \\ \psi^6 \end{pmatrix} = Y - Y_0 = \begin{pmatrix} \phi^1 - \phi^1_0 \\ \phi^2 - \phi^2_0 \\ \frac{\partial \phi^1}{\partial \theta} - \frac{\partial \phi^1}{\partial \theta}|_{t=0} \\ \frac{\partial \phi^2}{\partial \theta} - \frac{\partial \phi^2}{\partial \theta}|_{t=0} \\ \frac{\partial \phi^1}{\partial t} - \frac{\partial \phi^1}{\partial t}|_{t=0} \\ \frac{\partial \phi^2}{\partial t} - \frac{\partial \phi^1}{\partial t}|_{t=0} \end{pmatrix}$$

and therefore the denominators of $\tilde{A}$ and $\tilde{B}$'s entries are:

$$|\frac{\partial \phi(\theta)}{\partial \theta}|^2 = (\frac{\partial \phi^1(\theta)}{\partial \theta})^2 + (\frac{\partial \phi^2(\theta)}{\partial \theta})^2 = (\psi^3 + \frac{\partial \phi^1_0(\theta)}{\partial \theta})^2 + (\psi^4 + \frac{\partial \phi^2_0(\theta)}{\partial \theta})^2$$

$$= (\psi^3)^2 + (\psi^4)^2 + (\frac{\partial\phi_0^1(\theta)}{\partial\theta})^2 + (\frac{\partial\phi_0^2(\theta)}{\partial\theta})^2 + 2(\psi^3\frac{\partial\phi_0^1(\theta)}{\partial\theta} + \psi^4\frac{\partial\phi_0^2(\theta)}{\partial\theta})$$

which is greater that zero when

$$-2(\psi^3\frac{\partial\phi_0^1(\theta)}{\partial\theta} + \psi^4\frac{\partial\phi_0^2(\theta)}{\partial\theta}) < (\psi^3)^2 + (\psi^4)^2 + (\frac{\partial\phi_0^1(\theta)}{\partial\theta})^2 + (\frac{\partial\phi_0^2(\theta)}{\partial\theta})^2$$

Let $K_1 = \inf_\theta|\frac{\partial\phi_0(\theta)}{\partial\theta}|$ and $K_2 = \sup_\theta\{|\frac{\partial\phi_0^1(\theta)}{\partial\theta}|, |\frac{\partial\phi_0^2(\theta)}{\partial\theta}|\}$ and consider $|\psi| < \frac{K_1}{4K_2}$. Then we have:

$$4K_2|\psi| < K_1 \Rightarrow -2(\psi^3\frac{\partial\phi_0^1(\theta)}{\partial\theta} + \psi^4\frac{\partial\phi_0^2(\theta)}{\partial\theta}) < K_1$$

$$< (\frac{\partial\phi_0^1(\theta)}{\partial\theta})^2 + (\frac{\partial\phi_0^2(\theta)}{\partial\theta})^2 < (\psi^3)^2 + (\psi^4)^2 + (\frac{\partial\phi_0^1(\theta)}{\partial\theta})^2 + (\frac{\partial\phi_0^2(\theta)}{\partial\theta})^2$$

as required, where the inequality after the follows from

$$-2(\psi^3\frac{\partial\phi_0^1(\theta)}{\partial\theta} + \psi^4\frac{\partial\phi_0^2(\theta)}{\partial\theta}) < 2|\psi^3\frac{\partial\phi_0^1(\theta)}{\partial\theta} + \psi^4\frac{\partial\phi_0^2(\theta)}{\partial\theta}|$$

$$< 2(|\psi||\frac{\partial\phi_0^1(\theta)}{\partial\theta}| + |\psi||\frac{\partial\phi_0^2(\theta)}{\partial\theta}|) < 4|\psi|K_2 < K_1$$

With a $|\psi|$ range on which the entries of $\tilde{A}$ and $\tilde{B}$ are analytic, the rest of the proof goes through the same way and we can conclude existence of short time flow. NOTE: $K_1 > 0$ by the assumption that the curve is regular.

Case 3: Varying volume form

The same problems (mentioned in Case 3 of the previous example) arise in applying known PDE techniques to a second order nonlinear PDE.

Chapter 6

Conclusions

The work in this dissertation can be continued in a variety of future directions. One viable open problem is to prove existence of short time gradient flow for the full penalty function P (including both its distance and curvature terms):

$$P(\phi) = \int_M |R(\phi(m))|^2 \mathrm{dvol}_M + \sum_i \int_M d^2(\phi(m), x_i) \mathrm{dvol}_M$$

We examined the existence of the gradient flow for the distance term for a simple example in Chapter 5. However, recall that the gradient of the curvature term is given at the end of Chapter 2. It is a fourth order nonlinear equation, to which many standard PDE techniques do not apply. One may be able to start by showing the existence of the short time flow for a simple example and generalizing from there.

Similarly, one can attempt to derive a lower bound for gradient flow for the distance penalty term for the general case, as opposed to only studying the example given in Chapter 5.

Another set of open questions arises when considering approximating more than one point in $\mathbb{R}^N$ (recall that the simple example in Chapter 5 studied flow towards the origin). Issues of discontinuity can arise in the computation of the gradient vector field when points on $\phi(M)$ have more than one closest point in the fixed data set in $\mathbb{R}^N$.

There are several big-picture questions to be addressed as well, such as deciding on the most appropriate topology on $\mathrm{Emb}(M, \mathbb{R}^N)$ as well as determining how to evaluate if negative gradient flow moves towards a local or global minimum or another type of critical point.

Appendix A

Geometry Terms and Definitions

- **The Riemannian Curvature Tensor R:**

The Riemannian curvature tensor is used to determine when a neighborhood of a Riemannian manifold, M admits a distortion free map to $\mathbb{R}^N$ with the standard metric. In particular, when $R = 0$ in a neighborhood of a point $m \in M$, a (possibly different) neighborhood of m will have such a distortion free map. The curvature tensor is given by:

$$R = R^i{}_{jkl} \partial_{x^i} \otimes dx^j \otimes dx^k \otimes dx^l \in T_1^3$$

(where T_1^3 denotes curvature tensor of type (1,3) and $R^i{}_{jkl}$ is defined below).

- **Christoffel Symbol:**

Christoffel symbols are a central component of the integrability condition for when a manifold admits a distortion free map to $\mathbb{R}^N$ (referred to in definition above). It is given by:

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} \left(\frac{\partial g_{kl}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^l} \right)$$

- **The Symbol $R^i{}_{jkl}$:**

$$R^i{}_{jkl} = \frac{\partial \Gamma^i_{jk}}{\partial x^l} - \frac{\partial \Gamma^i_{jl}}{\partial x^k} + \Gamma^s_{jk} \Gamma^i_{sl} - \Gamma^s_{jl} \Gamma^i_{sk}$$

- **The Symbol** δ_X:

Recall that for penalty function $P : \text{Emb}(M, \mathbb{R}^N) \to \mathbb{R}$, we derived the gradient vector field

$Z = \nabla P(\phi_0)$ at a fixed embedding ϕ_0 in the direction X, by the following:

$$\frac{d}{ds}|_{s=0} P(\phi_s) = D_{\vec{X}} P(\phi_0) = \langle X, \nabla P(\phi_0) \rangle$$

where $X \in T_{\phi_0}\text{Emb}(M, \mathbb{R}^N)$ and ϕ_s is a parameterized curve in $\text{Emb}(M, \mathbb{R}^N)$. The tangent vector of the curve ϕ_s at ϕ_0 is X. In the left hand term, $\frac{d}{ds}$ can be replaced with the symbol δ_X, which denotes taking a derivative with respect to s in the X direction.

- **Levi-Civita connection:** The Levi-Civita connection is the generalization to Riemannian manifolds (M, g) of taking the derivative of a vector field in TM in the neighborhood of a point with respect to another fixed vector field. Its formal definition is given by [6]:

Definition 2. *Let (M, g) be a Riemannian manifold. Define $\nabla : TM \otimes \Gamma(TM) \to \Gamma(TM), (x \otimes X) \to \nabla_v X$, by the conditions:*
(i) $\nabla_{\partial_j} \partial_k = \Gamma^i_{jk} \partial_i$
(ii) $\nabla_{\lambda v} X = \lambda \nabla_v X$ *and* $\nabla_{v+w} X = \nabla_v X + \nabla_w X$
(iii) $\nabla_v (fX) = f \nabla_v X + df(v) X$*, for all smooth* $f : M \to R$

- For a k-dimensional manifold embedded in $\mathbb{R}^N$, the inclusion map $\phi(M) \to \mathbb{R}^N$ takes points $(q^1, \cdots, q^k) \mapsto (x^1(\vec{q}), \cdots, x^N(\vec{q}))$. It is a standard result that the first fundamental form is the matrix with entries $(g_{ij}) = \left(\frac{\partial \vec{x}}{\partial q^i} \cdot \frac{\partial \vec{x}}{\partial q^j}\right)$ (Euclidean dot product) and the second fundamental form in the direction $\vec{v}$ is the matrix with entries $(\vec{v} \cdot \vec{l}_{ij})$ where $\vec{l}_{ij}$ is the normal component of the vector $\frac{\partial^2 \vec{x}}{\partial q^i \partial q^j}$.

- **Trace of Second Fundmental Form, TrII:**

The trace of the second fundamental form, denoted TrII at a point $q \in \phi(M)$ is given by:

$$\text{Tr}II = \sum_{i=1}^{k} B(e_i, e_i)$$

where $\{e_i\}_i$ are basis vectors of $T_q\phi(M)$ and B_q is a symmetric bilinear map of $T_q\phi(M) \to$

$N_q\phi(M)$ that acts by:

$$B_{X,Y} = (\nabla_X^{\mathbb{R}^N} Y)^{\perp}$$

for any $X, Y \in T_q\phi(M)$, where $Z^{\perp}$ refers to projecting Z onto its normal component and where $\nabla^{\mathbb{R}^N}$ is the Levi-Cevita connection in $\mathbb{R}^N$.

Appendix B

This quantitative version of the Implicit Function theorem and its variation of standard proof techniques is by Carlangelo Liverani:

Let $n, m \in \mathbb{N}$ and $F \in C^1(\mathbb{R}^{n+m}, \mathbb{R}^m)$ and let $(x_0, \lambda_0) \in \mathbb{R}^m \times \mathbb{R}^n$ such that $F(x_0, \lambda_0) = 0$. For each $\delta > 0$ let $V_\delta = \{(x, \lambda) \in \mathbb{R}^{m+n} : ||x - x_0|| \le \delta, ||\lambda - \lambda_0|| \le \delta\}$.

Quantitative Implicit Function Theorem:

Theorem 6. (Quantitative Implicit Function Theorem) *Assume that* $\partial_x F(x_0, \lambda_0)$ *is invertible and choose* $\delta > 0$ *such that* $sup_{(x,\lambda)\in V_\delta} ||1 - [\partial_x F(x_0, \lambda_0)]^{-1} \partial_x F(x, \lambda)|| \le 1/2$. *Let* $B_\delta = sup_{(x,\lambda)\in V_\delta} ||\partial_\lambda F(x, \lambda)||$ *and* $M = ||\partial_x F(x_0, \lambda_0)^{-1}||$. *Set* $\delta_1 = (2MB_\delta)^{-1}\delta$ *and* $\Gamma_{\delta_1} = \{\lambda \in \mathbb{R}^n : \ ||\lambda - \lambda_0|| < \delta_1\}$. *Then there exists* $g \in C^1(\Gamma_{\delta_1}, \mathbb{R}^m)$ *such that all the solutions of the equation* $F(x, \lambda) = 0$ *in the set* $\{(x, \lambda) : \ ||\lambda - \lambda_0|| < \delta_1, ||x - x_0|| < \delta\}$ *are given by* $(g(\lambda), \lambda)$. *In addition,* $\partial_\lambda g(\lambda) = -(\partial_x F(g(\lambda), \lambda))^{-1} \partial_\lambda F(g(\lambda), \lambda)$

Proof: Set $A(x, \lambda) = \partial_x F(x, \lambda), M = ||A(x_0, \lambda_0)^{-1}||$.

We want to solve the equation $F(x, \lambda) = 0$. Let λ be such that $||\lambda - \lambda_0|| < \delta_1 \le \delta$. Consider $U_\delta = \{x \in \mathbb{R}^m : ||x - x_0|| \le \delta\}$ and the function $\Omega : U_\delta \to \mathbb{R}^m$ defined by

$$\Omega_\lambda(x) = x - A(x_0, \lambda_0)^{-1} F(x, \lambda).$$

For $x \in U(\lambda), F(x, \lambda) = 0$ is equivalent to $x = \Omega_\lambda(x)$.

Next,

$$||\Omega_\lambda(x_0) - \Omega_{\lambda_0}(x_0)|| \le M||F(x_0, \lambda)|| \le MB_\delta\delta_1$$

In addition, $||\partial_x \Omega_\lambda|| = ||1 - A(x_0, \lambda_0)^{-1} A(x, \lambda)|| \leq 1/2$. Thus

$$||\Omega_\lambda(x) - x_0|| \leq \frac{1}{2}||x - x_0|| + ||\Omega_\lambda(x_0) - x_0|| \leq \frac{1}{2}||x - x_0|| + MB_\delta \delta_1 \leq \delta$$

The existence of $x \in U_\delta$ such that $\Omega_\lambda(x) = x$ follows by the Fixed Point Theorem. We have therefore obtained a function $g : \Gamma_{\delta_1} = \{\lambda : ||\lambda - \lambda_0|| \leq \delta_1\} \to \mathbb{R}^m$ such that $F(g(\lambda), \lambda) = 0$.

It remains to prove regularity. Let $\lambda, \lambda' \in \Gamma_{\delta_1}$. From above we have

$$||g(\lambda) - g(\lambda')|| \leq \frac{1}{2}||g(\lambda) - g(\lambda')|| + MB_\delta|\lambda - \lambda'|$$

This yields the Lipschitz continuity of the function g. To obtain the differentiability we note that, by the differentiability of F and the above Lipschitz continuity of g, for $h \in \mathbb{R}^n$ small enough,

$$||F(g(\lambda + h), \lambda + h) - F(g(\lambda), \lambda) + \partial_x F[g(\lambda + h) - g(\lambda), h] + \partial_\lambda F(g(h), h)|| = o(||h||)$$

Since $F(g(\lambda + h), \lambda + h) = F(g(\lambda), \lambda) = 0$ we have

$$\lim_{h \to 0} ||h||^{-1} ||g(\lambda + h) - g(\lambda) + [\partial_x F(g(h), h)]^{-1} \partial_\lambda F(g(h), h)|| = 0,$$

which concludes the proof.

Appendix C

Computation of the distance penalty function under three variations (computations by Steve Rosenberg):

Recall, the distance penalty function is given by:

$$P_d(\phi) = \sum_i \int_M d^2(\phi(m), x_i) \mathrm{dvol}_M$$

(where $S = \{x_i\}_i$ is the set of fixed points we are approximating in $\mathbb{R}^N$. To compute the gradient of this function at a point $\phi_0 \in C^\infty(M, \mathbb{R}^N)$, we can consider a fixed vector $X \in T_{\phi_0} C^\infty(M, \mathbb{R}^N)$ and we know:

$$d(P_d)_{\phi_0}(\vec{X}) = \langle \nabla P_d(\phi_0), \vec{X} \rangle = \int_M \nabla P_d(\phi_0) \cdot \vec{X} \mathrm{dvol}_M$$

Additionally, we know that for a parameterized curve in $\phi(s) \in \mathrm{C}^\infty(M, \mathbb{R}^N)$, where s belongs to an interval around 0 such that $\phi(0) = \phi_0$ and $\frac{\mathrm{d}\phi}{\mathrm{ds}}|_{s=0} = \vec{X}$, we have

$$d(P_d)_{\phi_0}(\vec{X}) = \frac{\mathrm{d}P_d}{\mathrm{ds}}|_{s=0} = \frac{\mathrm{d}}{\mathrm{ds}}|_{s=0} \int_M d^2(\phi(m), x_i) \mathrm{dvol}_M$$

$$= \delta_X \int_M d^2(\phi(m), x_i) \mathrm{dvol}_M \tag{6.1}$$

If we can arrange the results of this computation into the form $\int_M f \cdot \vec{X} \mathrm{dvol}_M$ then we can conclude that $\nabla P_d = f$.

Case 1: Hold Volume Form Constant

Note: If there is more than one data point in S then all of the following terms should be summed over i:

$$\delta_X P_d(\phi) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} p_2(\phi+\epsilon X) = \int_M \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \Big[(\phi(y)^k + \epsilon X^k - x_i^k)^2 \text{ dvol}(y)\Big]$$

$$= 2\int_M \sum_k (\phi(y)^k - x_i^k)X^k \text{ dvol}(y) = 2\int_M (\phi(y) - x_i)\cdot X \text{ dvol}(y)$$

$$\Rightarrow \nabla P_d(\phi(y)) = 2(\phi(y) - x_i)$$

Case 2: Projecting Case 1 Onto Normal Directions

The gradient computed in Case 1 is not necessarily normal to $\phi(M)$ at all points. However, consider the modified gradient flow $\nabla\hat{P}_d$ from Case 1 given by:

$$\nabla\hat{P}_d(\phi(y)) = 2(\phi(y) - x_i) - 2\sum_{r,l=1}^{k}\sum_{j=1}^{N} g^{rl}(\phi(y)^j - x_i^j)\frac{\partial\phi(y)^j}{\partial y^l}\partial_r \tag{6.2}$$

The computation of the second term in this expression is made explicit in Case 3. For now, set $\phi(y) = v, u = \phi(y) - x_i$. Then $\partial v/\partial y^l = \partial_l$ in this notation. The two terms above satisfy:

$$(u - g^{rl}(u\cdot\partial_l)\partial_r)\cdot\partial_s = u\partial_s - g^{rl}(u\cdot\partial_l)g_{rs} = u\cdot\partial_s - u\cdot\partial_s = 0$$

Therefore letting $\nabla\hat{P}_d$ be the expression in (6.l) gives a vector field that is normal at every point in $\phi(M)$, but is not technically a gradient flow. Its translation into a first order quasilinear PDE in Chapter 5 makes it a worthwhile object of study because we are able to apply the Cauchy- Kovalevskaya Theorem.

Case 3: Varying Volume Form

We now compute the full gradient ∇P_d of the distance penalty function with varying volume form. In this scenario, the volume form uses the induced metric from $\mathbb{R}^N$:

$$\delta_X P_d(\phi) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} P_d(\phi + \epsilon X) = \int_M \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \Big[(\phi(y)^k + \epsilon X^k - x_i^k)^2 \text{ dvol}(y)\Big]$$

$$= 2\int_M \sum_k (\phi(y)^k - x_i^k)X^k \text{ dvol}(y) + \int_M d^2(\phi(y), x_i)\delta_X \text{dvol}(y)$$

$$= 2\int_M (\phi(y) - x_i)\cdot X \text{ dvol}(y) + \int_M d^2(\phi(y), x_i)(-X \cdot \text{Tr II dvol}(y) + d * \omega)$$

$$= 2\int_M (\phi(y) - x_i)\cdot X \text{ dvol}(y) - \int_M d^2(\phi(y), x_i)\text{Tr II} \cdot X \text{ dvol}(y)$$

$$- \int_M \text{grad}(d^2(\phi(y), x_i)) \cdot X \text{ dvol}(y)$$

$$= 2\int_M (\phi(y) - x_i)\cdot X \text{ dvol}(y) - \int_M d^2(\phi(y), x_i)\text{Tr II} \cdot X \text{ dvol}(y)$$

$$- \int_M \left(\sum_{r,l=1}^{k=\dim(M)} g^{rl}\partial_l d^2(\phi(y), x_i) \right) \partial_r \cdot X \text{ dvol}(y)$$

$$= 2\int_M (\phi(y) - x_i)\cdot X \text{ dvol}(y) - \int_M d^2(\phi(y), x_i)\text{Tr II} \cdot X \text{ dvol}(y)$$

$$-2\int_M \sum_{r,l} g^{rl} \sum_{j=1}^{N} (\phi(y)^j - x_i^j)\frac{\partial \phi(y)^j}{\partial y^l} \cdot X^r \text{ dvol}(y)$$

Thus

$$\nabla(P_d)_\phi(y) = 2(\phi(y) - x_i) - d^2(\phi(y), x_i)\text{Tr II} - 2\sum_{r,l=1}^{k}\sum_{j=1}^{N} g^{rl}(\phi(y)^j - x_i^j)\frac{\partial \phi(y)^j}{\partial y^l}\partial_r$$

$$\in \Gamma(\phi^* T\mathbb{R}^N) = T_\phi \text{Emb}(M, \mathbb{R}^N)$$

Note that TrII is perpendicular to $\phi(M)$, so by Case 2, ∇P_d is perpendicular to $\phi(M)$.

Appendix D

1) Definition of **Cauchy data**, as given on page 233-234 of Evans [1]:

Given the kth order quasilinear PDE:

$$\sum_{|\alpha|=k} a_\alpha(D^{k-1}u, \ldots, u, x)D^\alpha u + a_0(D^{k-1}u, \ldots, u, x) = 0 \tag{6.3}$$

in some open region $U \subset \mathbb{R}^n$. Let us assume that Γ is a smooth, $(n-1)$-dimensional hypersurface in U, the unit normal to which at any point $x^0 \in \Gamma$ is $\nu(x^0) = \nu = (\nu^1, \ldots, \nu^n)$. Let $g_0, \ldots, g_{k-1} : \Gamma \to \mathbb{R}$ be k given functions. The Cauchy Problem is then to find a function u solving (6.1), subject to the boundary conditions:

$$u = g_0, \frac{\partial u}{\partial \nu} = g_1, \ldots, \frac{\partial^{k-1}}{\partial \nu^{k-1}} = g_{k-1}$$

on Γ. We say that the above equations prescribe the **Cauchy data** $g_0, \ldots, g_{k-1}$ on Γ.

2) Full text of Cauchy-Kovalevskaya Theorem in Evans [1] page 239-244:

4.6.3. Cauchy-Kovalevskaya Theorem.

We turn now to our primary task of building a power series solution for the k^{th}-order quasilinear partial differential equation (1), with analytic Cauchy data (2) specified on an analytic, noncharacteristic hypersurface Γ.

a. Reduction to a first-order system.

We intend to construct a solution u as a power series, but must first transform the boundary-value problem (1), (2) into a more convenient form.

First of all, upon flattening out the boundary by an analytic mapping (as in §4.6.1), we can reduce to the situation that $\Gamma \subset \{x_n = 0\}$. Additionally, by subtracting off appropriate analytic functions, we may assume the Cauchy data are identically zero. Consequently we may assume without loss that our problem reads:

$$(14) \quad \begin{cases} \sum_{|\alpha|=k} a_\alpha(D^{k-1}u, \dots, u, x) D^\alpha u \\ \qquad\qquad + a_0(D^{k-1}u, \dots, u, x) = 0 & \text{for } |x| < r \\ \qquad\qquad u = \frac{\partial u}{\partial x_n} = \dots = \frac{\partial^{k-1} u}{\partial x_n^{k-1}} = 0 & \text{for } |x'| < r, x_n = 0, \end{cases}$$

$r > 0$ to be found. As usual we write $x' = (x_1, \dots, x_{n-1})$.

Finally we transform to a first-order *system*. To do so we introduce the function

$$\mathbf{u} := \left(u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}, \frac{\partial^2 u}{\partial x_1^2}, \dots, \frac{\partial^{k-1} u}{\partial x_n^{k-1}}\right),$$

the components of which are all the partial derivatives of u of order less than k. Let m hereafter denote the number of components of $\mathbf{u}$ by m, so that $\mathbf{u} : \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{u} = (u^1, \dots, u^m)$. Observe from the boundary condition in (14) that $\mathbf{u} = 0$ for $|x'| < r, x_n = 0$.

Now for $k \in \{1, \dots, m-1\}$, we can compute $u^k_{x_n}$ in terms of $\{\mathbf{u}_{x_j}\}_{j=1}^{n-1}$. Furthermore in view of the noncharacteristic condition $a_{(0,\dots,0,k)} \neq 0$ near

0, we can utilize the PDE in (14) also to solve for $u^m_{x_n}$ in terms of $\mathbf{u}$ and $\{\mathbf{u}_{x_j}\}_{j=1}^{n-1}$.

Employing these relations, we can consequently transform (14) into a boundary-value problem for a first-order system for $\mathbf{u}$, the coefficients of which are analytic functions. This system is of the general form:

$$(15)\qquad \begin{cases} \mathbf{u}_{x_n} = \sum_{j=1}^{n-1} \mathbf{B}_j(\mathbf{u}, x')\mathbf{u}_{x_j} + \mathbf{c}(\mathbf{u}, x') & \text{for } |x| < r \\ \mathbf{u} = 0 & \text{for } |x'| < r,\ x_n = 0, \end{cases}$$

where we are given the analytic functions $\mathbf{B}_j : \mathbb{R}^m \times \mathbb{R}^{n-1} \to \mathbb{M}^{m\times m}$ $(j = 1, \dots, n-1)$ and $\mathbf{c} : \mathbb{R}^m \times \mathbb{R}^{n-1} \to \mathbb{R}^m$. We will write $\mathbf{B}_j = ((b_j^{kl}))$ and $\mathbf{c} = (c^1, \dots, c^m)$. Carefully note that we have assumed $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$ do not depend on x_n. We can always reduce to this situation by introducing if necessary a new component u^{m+1} of the unknown $\mathbf{u}$, with $u^{m+1} \equiv x_n$.

In particular, the components of the system of partial differential equations in (15) read

$$(16)\qquad u^k_{x_n} = \sum_{j=1}^{n-1}\sum_{l=1}^{m} b_j^{kl}(\mathbf{u}, x')u^l_{x_j} + c^k(\mathbf{u}, x') \quad (k = 1, \dots, m).$$

b. Power series for solutions.

Having reduced to the special form (15), we can now expand $\mathbf{u}$ into a power series, and, more importantly, verify that this series converges near 0.

THEOREM 2 (Cauchy-Kovalevskaya Theorem). *Assume $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$ are real analytic functions. Then there exist $r > 0$ and a real analytic function*

$$(17)\qquad \mathbf{u} = \sum_{\alpha} \mathbf{u}_\alpha x^\alpha$$

solving the boundary-value problem (15).

Proof. 1. We must compute the coefficients

$$(18)\qquad \mathbf{u}_\alpha = \frac{D^\alpha \mathbf{u}(0)}{\alpha!}$$

in terms of $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$, and then show that the power series (18) so obtained in fact converges if $|x| < r$ and $r > 0$ is small enough.

2. As the functions $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$ are analytic, we can write

$$\mathbf{B}_j(z,x') = \sum_{\gamma,\delta} \mathbf{B}_{j,\gamma,\delta} z^\gamma x^\delta \quad (j = 1,\dots,n-1) \tag{19}$$

and

$$\mathbf{c}(z,x') = \sum_{\gamma,\delta} \mathbf{c}_{\gamma,\delta} z^\gamma x^\delta, \tag{20}$$

these power series convergent if $|z| + |x'| < s$ for some small $s > 0$. Thus

$$\mathbf{B}_{j,\gamma,\delta} = \frac{D_z^\gamma D_x^\delta \mathbf{B}_j(0,0)}{(\gamma+\delta)!}, \ \mathbf{c}_{\gamma,\delta} = \frac{D_z^\gamma D_x^\delta \mathbf{c}(0,0)}{(\gamma+\delta)!} \tag{21}$$

for $j = 1,\dots,n-1$ and all multiindices γ,δ.

3. Since $\mathbf{u} \equiv 0$ on $\{x_n = 0\}$, we have

$$\mathbf{u}_\alpha = \frac{D^\alpha \mathbf{u}(0)}{\alpha!} = 0 \quad \text{for all multiindices } \alpha \text{ with } \alpha_n = 0. \tag{22}$$

Now fix $i \in \{1,\dots,n-1\}$ and differentiate (16) with respect to x_i:

$$u^k_{x_n x_i} = \sum_{j=1}^{n-1}\sum_{l=1}^{m}\left(b_j^{kl} u^l_{x_i x_j} + b^{kl}_{j,x_i} u^l_{x_j} + \sum_{p=1}^m b^{kl}_{j,z_p} u^p_{x_i} u^l_{x_j}\right) + c^k_{x_i} + \sum_{p=1}^m c^k_{z_p} u^p_{x_i}.$$

In view of (22), we conclude $u^k_{x_n x_i}(0) = c^k_{x_i}(0,0)$.

If α is a multiindex having the form $\alpha = (\alpha_1,\dots,\alpha_{n-1},1) = (\alpha',1)$, we likewise prove by induction that

$$D^\alpha u^k(0) = D^{\alpha'} c^k(0,0).$$

Next suppose $\alpha = (\alpha', 2)$. Then

$$\begin{aligned} D^\alpha u^k &= D^{\alpha'}(u^k_{x_n})_{x_n} \\ &= D^{\alpha'}\left(\sum_{j=1}^{n-1}\sum_{l=1}^m b_j^{kl} u^l_{x_j} + c^k\right)_{x_n} \quad \text{by (16)} \\ &= D^{\alpha'}\left(\sum_{j=1}^{n-1}\sum_{l=1}^m (b_j^{kl} u^l_{x_j x_n} + \sum_{p=1}^m b^{kl}_{j,z_p} u^p_{x_n} u^l_{x_j}) + \sum_{p=1}^m c^k_{z_p} u^p_{x_n}\right). \end{aligned}$$

Thus

$$D^\alpha u^k(0) = D^{\alpha'}\left(\sum_{j=1}^{n-1}\sum_{l=1}^m b_j^{kl} u^l_{x_j x_n} + \sum_{p=1}^m c^k_{z_p} u^p_{x_n}\right)\Bigg|_{x=\mathbf{u}=0}.$$

The expression on the right hand side can be worked out to be a polynomial *with nonnegative coefficients* involving various derivatives of $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$, and the derivatives $D^\beta \mathbf{u}$, where $\beta_n \leq 1$.

More generally, for each multiindex α and each $k \in \{1, \dots, m\}$, we compute

$$D^\alpha u^k(0) = p_\alpha^k(\dots, D_z^\gamma D_x^\delta \mathbf{B}_j, \dots, D_z^\gamma D_x^\delta \mathbf{c}, \dots, D^\beta \mathbf{u}, \dots)\big|_{x=\mathbf{u}=0},$$

where p_α^k denotes some polynomial with nonnegative coefficients.

Recalling (18)–(21), we deduce for each α, k that

$$u_\alpha^k = q_\alpha^k(\dots, \mathbf{B}_{j,\gamma,\delta}, \dots, \mathbf{c}_{\gamma,\delta}, \dots, \mathbf{u}_\beta, \dots), \tag{23}$$

where

(24) $\quad q_\alpha^k$ is a polynomial *with nonnegative coefficients*

and

(25) $\quad \beta_n \leq \alpha_n - 1$ for each multiindex β on the right hand side of (23).

4. Thus far we have merely demonstrated that if there is a smooth solution of (15), then we can compute all of its derivatives at 0 in terms of known quantities. This of course we already know from the discussion in §4.6.1, since the plane $\{x_n = 0\}$ is noncharacteristic.

We now intend to employ (22)–(25) and the *method of majorants* to show the power series (17) actually converges if $|x| < r$ and r is small. For this, let us first suppose

$$\mathbf{B}_j^* \gg \mathbf{B}_j \quad (j = 1, \dots, n-1) \tag{26}$$

and

$$\mathbf{c}^* \gg \mathbf{c}, \tag{27}$$

where

$$\mathbf{B}_j^* := \sum_{\gamma,\delta} \mathbf{B}_{j,\gamma,\delta}^* z^\gamma x^\delta \quad (j = 1, \dots, n-1)$$

and

$$\mathbf{c}^* := \sum_{\gamma,\delta} \mathbf{c}_{\gamma,\delta}^* z^\gamma x^\delta,$$

these power series convergent for $|z| + |x'| < s$. Then

$$0 \le |\mathbf{B}_{j,\gamma,\delta}| \le \mathbf{B}^*_{j,\gamma,\delta},\ \ 0 \le |\mathbf{c}_{\gamma,\delta}| \le \mathbf{c}^*_{\gamma,\delta} \tag{28}$$

for all j, γ, δ.

We consider next the new boundary-value problem

$$\begin{cases} \mathbf{u}^*_{x_n} = \sum_{j=1}^{n-1} \mathbf{B}^*_j(\mathbf{u}^*, x')\mathbf{u}^*_{x_j} + \mathbf{c}^*(\mathbf{u}^*, x') & \text{for } |x| < r \\ \mathbf{u}^* = 0 & \text{for } |x'| < r,\ x_n = 0, \end{cases} \tag{29}$$

and, as above, look for a solution having the form

$$\mathbf{u}^* = \sum_\alpha \mathbf{u}^*_\alpha x^\alpha, \tag{30}$$

where

$$\mathbf{u}^*_\alpha = \frac{D^\alpha \mathbf{u}^*(0)}{\alpha!}. \tag{31}$$

5. We claim

$$0 \le |u^k_\alpha| \le u^{k*}_\alpha \quad \text{for each multiindex } \alpha.$$

The proof is by induction. The general step follows since

$$\begin{aligned} |u^k_\alpha| &= |q^k_\alpha(\dots, \mathbf{B}_{j,\gamma,\delta}, \dots, \mathbf{c}_{\gamma,\delta}, \dots, \mathbf{u}_\beta, \dots)| \quad \text{by (23)} \\ &\le q^k_\alpha(\dots, |\mathbf{B}_{j,\gamma,\delta}|, \dots, |\mathbf{c}_{\gamma,\delta}|, \dots, |\mathbf{u}_\beta|, \dots) \quad \text{by (24)} \\ &\le q^k_\alpha(\dots, \mathbf{B}^*_{j,\gamma,\delta}, \dots, \mathbf{c}^*_{\gamma,\delta}, \dots, \mathbf{u}^*_\beta, \dots) \quad \text{by (24), (28) and induction} \\ &= u^{k*}_\alpha. \end{aligned}$$

Thus

$$\mathbf{u}^* \gg \mathbf{u}, \tag{32}$$

and so it suffices to prove that the power series (30) converges near zero.

6. As demonstrated in the proof of assertion (ii) of the lemma in §4.6.2, if we choose

$$\mathbf{B}^*_j := \frac{Cr}{r - (x_1 + \dots + x_{n-1}) - (z_1 + \dots + z_m)} \begin{pmatrix} 1 & \dots & 1 \\ & \ddots & \\ 1 & & 1 \end{pmatrix}$$

for $j = 1, \dots, n-1$, and

$$\mathbf{c}^* := \frac{Cr}{r - (x_1 + \cdots + x_{n-1}) - (z_1 + \cdots + z_m)}(1, \dots, 1),$$

then (26), (27) will hold if C is large enough, $r > 0$ is small enough, and $|x'| + |z| < r$.

Hence the problem (29) reads

$$\begin{cases} \mathbf{u}^*_{x_n} = \dfrac{Cr}{r - (x_1 + \cdots + x_{n-1}) - (u^{1*} + \cdots + u^{m*})} \left(\sum_{j,l} \mathbf{u}^{l*}_{x_j} + 1\right) & \text{for } |x| < r \\ \mathbf{u}^* = 0 & \text{for } |x'| < r,\ x_n = 0. \end{cases}$$

However, this problem has an explicit solution, namely

$$\mathbf{u}^* = v^*(1, \dots, 1), \tag{33}$$

for

$$\begin{aligned} v^*(x) := &\frac{1}{mn}(r - (x_1 + \cdots + x_{n-1}) \\ &- [(r - (x_1 + \cdots + x_{n-1}))^2 - 2mnCrx_n]^{1/2}). \end{aligned} \tag{34}$$

This expression is analytic for $|x| < r$, provided $r > 0$ is sufficiently small. Thus $\mathbf{u}^*$ defined by (33) necessarily has the form (30), (31), the power series (30) converging for $|x| < r$. As $\mathbf{u}^* \gg \mathbf{u}$, the power series (17) converges as well for $|x| < r$.

This defines the analytic function $\mathbf{u}$ near 0. Since the Taylor expansions of the analytic functions $\mathbf{u}_{x_n}$ and $\sum_{j=1}^{n-1} \mathbf{B}_j(\mathbf{u}, x)\mathbf{u}_{x_j} + \mathbf{c}(\mathbf{u}, x)$ agree at 0, they agree as well throughout the region $|x| < r$. □

Remark. The Cauchy-Kovalevskaya Theorem is valid also for fully nonlinear, analytic PDE: see Folland [**F1**] □

3) Definition of **majorize** and Lemmas (i) and (ii) from Evans [1]:

DEFINITION. *Let*

$$f = \sum_\alpha f_\alpha x^\alpha, \; g = \sum_\alpha g_\alpha x^\alpha$$

be two power series. We say g majorizes *f, written*

$$g \gg f,$$

provided

$$g_\alpha \geq |f_\alpha| \quad \textit{for all multiindices } \alpha.$$

LEMMA (Majorants).

(i) *If $g \gg f$ and g converges for $|x| < r$, then f also converges for $|x| < r$.*

(ii) *If $f = \sum_\alpha f_\alpha x^\alpha$ converges for $|x| < r$ and $0 < s\sqrt{n} < r$, then f has a majorant for $|x| < s\sqrt{n}$.*

Proof. 1. To verify assertion (i), we check

$$\sum_\alpha |f_\alpha x^\alpha| \leq \sum_\alpha g_\alpha |x_1|^{\alpha_1} \cdots |x_n|^{\alpha_n} < \infty \quad \text{if } |x| < r.$$

2. Let $0 < s\sqrt{n} < r$ and set $y := s(1, \ldots, 1)$. Then $|y| = s\sqrt{n} < r$ and so $\sum_\alpha f_\alpha y^\alpha$ converges. Thus there exists a constant C such that

$$|f_\alpha y^\alpha| \leq C \quad \text{for each multiindex } \alpha.$$

In particular,

$$|f_\alpha| \leq \frac{C}{y_1^{\alpha_1} \cdots y_n^{\alpha_n}} = \frac{C}{s^{|\alpha|}} \leq C \frac{|\alpha|!}{s^{|\alpha|}\alpha!}.$$

But then

$$g(x) := \frac{Cs}{s - (x_1 + \cdots + x_n)} = C \sum_\alpha \frac{|\alpha|!}{s^{|\alpha|}\alpha!} x^\alpha$$

majorizes f for $|x| < s\sqrt{n}$. □

4) Folland's [2] adjustment of the original Cauchy problem to a quasilinear system: After a change of coordinates, we can assume the Cauchy problem takes the form:

$$\partial_t^k u = G(x, t, (\partial_x^\alpha \partial_t^j u)_{|\alpha|+j \leq k, j<k}) \qquad (1.24)$$

$$\partial_t^j u(x, 0) = \phi_j(x) \quad (0 \leq j < k)$$

The main result is the following existence theorem:

Theorem 7. *If $G, \phi_0, \ldots, \phi_{k-1}$ are analytic near the origin, there is a neighborhood of the origin on which the Cauchy problem (1.24) has a unique analytic solution.*

The following theorem gives the method of adjusting (1.24) to a first order, quasilinear system:

(1.31) Theorem.
The Cauchy problem (1.24) is equivalent to the Cauchy problem for a certain first order quasi-linear system,

$$\begin{aligned} \partial_t Y &= \sum_1^{n-1} A_j(x,t,Y)\partial_{x_j} Y + B(x,t,Y), \\ Y(x,0) &= \Phi(x), \end{aligned} \tag{1.32}$$

in the sense that a solution to one problem can be read off from a solution to the other. Here Y, B, and Φ are vector-valued functions, the A_j's are matrix-valued functions, and A_j, B, and Φ are explicitly determined by the functions in (1.24).

Proof: The vector Y is to have components $(y_{\alpha j})$, $0 \leq |\alpha| + j \leq k$. In what follows it is understood that $\partial_x^\alpha \partial_t^j u$ is to be replaced by $y_{\alpha j}$ as an independent variable in G. Also, if α is a nonzero multi-index, $i = i(\alpha)$ will denote the smallest number such that $\alpha_i \neq 0$, and 1_i will denote the multi-index with 1 in the ith place and 0 elsewhere. The first order system to be solved is

$$\begin{aligned} &\text{(a)} \quad \partial_t y_{\alpha j} = y_{\alpha(j+1)} \qquad (|\alpha| + j < k), \\ &\text{(b)} \quad \partial_t y_{\alpha j} = \partial_{x_i} y_{(\alpha-1_i)(j+1)} \qquad (|\alpha| + j = k,\ j < k), \\ &\text{(c)} \quad \partial_t y_{0k} = \frac{\partial G}{\partial t} + \sum_{|\alpha|+j<k} \frac{\partial G}{\partial y_{\alpha j}} y_{\alpha(j+1)} \\ &\qquad\qquad + \sum_{|\alpha|+j=k,\ j<k} \frac{\partial G}{\partial y_{\alpha j}} \partial_{x_i} y_{(\alpha-1_i)(j+1)}, \end{aligned} \tag{1.33}$$

and the initial conditions are

$$a) y_{\alpha j}(x,0) = \partial_x^\alpha \phi_j(x) \quad (j < k)$$

$$b) y_{0k}(x,0) = G(x,0,(\partial_x^\alpha \phi_j(x))_{|\alpha|+j\le k, j<k})$$

Bibliography

[1] Evans, L. C. *Partial Differential Equations.* Providence, RI, 2010.

[2] Folland, G. B. *Introducation to Partial Differential Equations.* Princeton University Press, Princeton, NJ, 1995.

[3] Guillemin, V., and Pollack, A. *Differential Topology.* Prentice Hall, Inc., Englewood Cliffs, NJ, 1974.

[4] H. Blaine Lawson, J. *Lectures on Minimal Submanifolds: Volume I.* Publish or Perish, Inc., Berkeley, CA, 1980.

[5] Milnor, J. *Morse Theory.* Princeton University Press, Princeton, NJ, 1969.

[6] Rosenberg, S. *The Laplacian on a Riemannian Manifold.* Cambridge University Press, 1997.

Curriculum Vitae

[redacted]