The Differential Geometry of Landmark Shape Manifolds: Metrics, Geodesics, and Curvature

by

Mario Micheli

Laurea, Università di Padova, Italy, 1999

M. S., University of California at Berkeley, 2001

Sc. M., Brown University, 2003

A Dissertation submitted in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy
in the Division of Applied Mathematics at Brown University

Providence, Rhode Island

May 2008

This dissertation by Mario Micheli is accepted in its present form
by the Division of Applied Mathematics as satisfying the
dissertation requirement for the degree of Doctor of Philosophy.

Date ______________ ____

David B. Mumford, Director

Recommended to the Graduate Council

Date ______________ ____

Stuart A. Geman, Reader

Date ______________ ____

Peter W. Michor, Reader

Approved by the Graduate Council

Date ______________ ____

Sheila Bonde, Dean of the Graduate School

Curriculum Vitæ

Mario Micheli was born in Rovereto (Trento), Italy, on August 27, 1973. He received the *laurea* degree in Telecommunications Engineering from the University of Padova, Italy, in 1999; as an undergraduate, he was an exchange student at the Université Bordeaux I, France, during the Fall semester of 1997 and at the University of California at Berkeley during the academic year 1998-1999. While at Berkeley he wrote a thesis for his *laurea* degree with the title *A Probabilistic Approach to Three-dimensional Autonomous Navigation*.

In the Fall of 2000 he enrolled into a graduate program at UC Berkeley and received Masters of Science degree in Electrical Engineering in May 2001; while at Berkeley he worked under the supervision of Professors Shankar S. Sastry and Michael I. Jordan, both from the Department of Electrical Engineering and Computer Sciences, and wrote a thesis with the title *Random Sampling of Continuous-time Stochastic Dynamical Systems: Analysis, State Estimation, and Applications.*.

In September of 2001 he began his doctoral work in the field of Applied Mathematics at Brown University, where he was supervised by Professor David B. Mumford. While at Brown he was supported by research assistantships, a teaching fellowship, a Florence Harnish Fellowship, and a dissertation fellowship. He received a Masters of Science degree from Brown University in Applied Mathematics in May 2003.

In September 2008 he will join the Department of Mathematics of the University of California at Los Angeles, as a postdoctoral scholar.

Dedicated to my parents, Margherita and Giuseppe Micheli

Acknowledgements

First and foremost I wish to thank my thesis advisor, David Mumford, for having given me the honor of working with him and for having introduced me to the wonderful topic of shape spaces; his precious teachings will help me for many years to come. Stuart Geman has also been an invaluable resource, and I am very grateful for all of our many helpful discussions. I must add that both David Mumford and Stuart Geman have given me, during my years at Brown University, advice and support that often times went beyond the academic realm; their graciousness and wholeheartedness are truly rare. I also wish to give my special thanks to Peter Michor of the University of Vienna for his remarkable patience and flexibility as reader of this thesis.

I had initially promised myself that I would not attempt to make a list of all the people and friends who have made these many years at Brown University the life-changing experience that it has been—although it would have been an excellent excuse to add a few pages to this thesis. However, I finally decided to (partially) break my promise: some of them have been so caring and supportive in different ways during the very last part of my stay here at Brown that I would have probably not been able to complete my work without them. They are Dzigbodi Agbenyadzie, Yi Cai, Indrek Kulaots, Akil Narayan, Anish Shah, Vito Stella, and Wei-Ying Wong. They truly deserve my gratitude.

Last, but certainly not least, I would like to thank my parents, Margherita and Giuseppe Micheli, for their unconditional support, selfless love, and continuous advice; despite the physical distance separating us they have always been the harbor where to seek refuge in case of need. This work is dedicated to them.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

The study of *shapes* and their similarities is central in computer vision, in that it allows to recognize and classify objects from their representation. One has the interest of defining a *distance* function between shapes, which both expresses the meaning of similarity between shapes for the application and task that one has in mind, and at the same time is mathematically sound and treatable. In recent years the use of differential-geometric techniques for the study of shape deformation has rapidly spread to broad applied fields such as Pattern Analysis and Statistical Methods (e.g. for object recognition, target detection and tracking, classification of biometric data, and automated medical diagnostics).

One of the main ideas in this area has been to use fluid flow notions [**8, 9**], which lead to Riemannian metrics on many deformation related spaces ("shape spaces") [**22, 24, 36, 42, 43, 47**]; the aforementioned distance function is, in these cases, precisely the geodesic distance with respect to such metrics. However, the geometry of these Riemannian manifolds has remained *terra incognita* until very recently, when certain fundamental questions started being addressed [**31, 32, 33, 48**]: for example, the *curvature* of such manifolds has remained completely unknown in most cases. This thesis focuses on the computation of sectional curvature and its implications in one of the simplest shape spaces, which is that of landmark points.

The knowledge of curvature on a Riemannian manifold is essential in that it allows one to infer about the existence of conjugate points, the well-posedness of the problem of computing the implicit mean (and higher statistical moments) of samples on the manifold, and more. Such issues are of fundamental importance since they allow to build templates, i.e. shape classes that represent typical situations in certain applications like the emerging field of computational anatomy [**10, 14, 15,**

18, 22, 34, 35, 46]. For example, templates can used for the identification of structures in Magnetic Resonance Images (MRI) of brains. A template can represent the prototypical structure of a healthy person's brain, or the structure of the brain of someone developing Alzheimer's disease: such templates are matched to the MRI scan of an individual patient, and the geodesic distances between the data and the templates can then be used to formulate a diagnosis on the patient's health. In Medical Imaging statistical analysis is normally performed on the tangent space at the implicit mean (or Karcher mean [**38**]), but the differential-geometric issues mentioned above are too often ignored, which can lead to conspicuous inaccuracies. We will briefly return on these motivating issues in the concluding chapter of this thesis, after the geometric structure of the shape manifold of landmark points has been explored.

The thesis is organized as follows. Chapter 2 formally introduces the energy functional that has to be minimized in order to compute the distance between between two sets (landmarks configurations); most of the chapter is dedicated to proving that such functional can be written as the energy of a path with respect to a Riemannian metric tensor, so that the square root of the minimized energy is, in fact, a geodesic distance. Once the Riemannian structure is established, the geodesic equations are developed (in the Hamiltonian formalism) in Chapter 3, where conservation laws deriving from the translation-invariance and the rotation-invariance of the metric tensor are also explored.

It turns out that the *cometric* tensor (i.e. the inverse of the metric tensor) for the Riemannian manifold of N landmarks in D dimensions can be written as a matrix whose elements depend only on $2D$ of all the ND coordinates of the manifold, making the matrix of partial derivatives of the cometric tensor very sparse. This suggests finding a general formula for the Riemannian curvature tensor and for sectional curvature in terms of the first and second partial derivatives of the *cometric* instead of the metric; Chapter 4 is dedicated precisely to solving this highly nontrivial problem, for a generic n-dimensional Riemannian manifold. In Chapter 5, which is central in this thesis, we apply the formulas developed in the previous one precisely to the

Riemannian metric for the manifold of landmarks; special attention is dedicated to the simple but informative examples of one-dimensional landmarks. In Chapter 6 we study the qualitative dynamics of landmarks, i.e. the geodesic trajectories that solve the equations developed in Chapter 3; in particular, we study the effect of curvature on said dynamics by verifying, for example, the existence of conjugate points in regions of the manifold of positive curvature. Finally, Chapter 7 summarizes our results, discusses the their potential application to the statistical analysis of medical images, and draws the future plans of our research.

CHAPTER 2

The Riemannian Manifold of Landmarks

In this chapter we illustrate how the shape space of landmarks can be endowed with the structure of a Riemannian manifold. The treatment is rigorous, however we skip some technicalities in the mathematical preliminaries section regarding the regularity of diffeomorphisms, limiting ourselves to citing and later using results that the reader can find, for example, in [**43, 47**]. We formulate the distance between two shapes in terms of the average kinetic energy of a velocity field that transports one shape into the other; we then show that such energy can be expressed in the form of the energy of a path with respect to a Riemannian metric tensor. At the end of the chapter we briefly discuss ways of extending the approach to generic shapes, by formulating it in terms of Lie groups of diffeomorphisms acting on shape manifolds.

1. General framework

Let $\mathcal{I}$ be the space of N landmark points in D dimensions, that is, the generic element of $\mathcal{I}$ is given by $I = (x^1, x^2, \ldots, x^N)$, $x^i \in \mathbb{R}^D$, with $x^i \neq x^j$ for $i \neq j$. Our objective is to endow $\mathcal{I}$ with a *distance* function $d : \mathcal{I} \times \mathcal{I} \to \mathbb{R}^+$ that will turn out to be the geodesic distance [**11, 23, 27**] with respect to a Riemannian metric. The idea is to find, among the diffeomorphisms of the plane that map a shape into another, the one that is generated by the velocity field of minimal average "kinetic energy" (to be defined): the distance between the two shapes will be given by the square root of such average energy.

1.1. Mathematical preliminaries. We will start with some definitions and preliminary notions. Let $\mathcal{Q}$ be the set of differentiable landmarks paths, that is:

$$\mathcal{Q} \triangleq \left\{ \left\{ q^i : [0,1] \to \mathbb{R}^D \right\}_{i=1}^N \,\middle|\, q^i(\cdot) \text{ is differentiable, for all } i \right\};$$

we shall indicate the generic element of such set simply with q, so that $q(t) = \{q^1(t), \ldots, q^N(t)\}$, $t \in [0,1]$. Let V be a set of functions $\mathbb{R}^D \to \mathbb{R}^D$ that has the structure of an *admissible* Hilbert space $\big(V, \langle\cdot,\cdot\rangle_V\big)$ (see [**17, 43, 47**], or Appendix A for an essential treatment of such spaces); the most salient property of admissible Hilbert spaces is that V is embedded in $C_0^1(\mathbb{R}^D, \mathbb{R}^D)$ (the *subscript* 0 denotes functions that vanish at infinity [**16**]), i.e. such that for some constant C we have $\|u\|_{1,\infty} \le C\|u\|_V$, for all $u \in V$. For example, V can be chosen to be Sobolev space $H^k(\mathbb{R}^D, \mathbb{R}^D)$ with its norm [**13**]:

$$||u||_V^2 \triangleq \int_{\mathbb{R}^D} \big\langle Lu(x), u(x)\big\rangle_{\mathbb{R}^D}\, dx; \tag{2.1}$$

in the above expression $L = (\mathrm{id} - a^2\Delta)^k$ is a self-adjoint spatial differential operator ($a \in \mathbb{R}$, $k \in \mathbb{N}$ and Δ is the Laplacian) that is applied to each of the D components of the vector field[1] u. By the Sobolev Embedding Theorem [**16**] we have in fact that if $k > \frac{D}{2} + 1$ then V is embedded in $C_0^1(\mathbb{R}^D, \mathbb{R}^D)$. Space $L^1([0,1], V)$ is defined as the set of functions $v : [0,1] \to V : t \mapsto v_t$ such that:

$$||v||_{L^1([0,1],V)} \triangleq \int_0^1 \|v_t\|_V\, dt < \infty\,,$$

while space $L^2([0,1], V)$ is the set of functions $v : [0,1] \to V : t \mapsto v_t$ such that:

$$||v||^2_{L^2([0,1],V)} \triangleq \int_0^1 \|v_t\|_V^2\, dt < \infty\,.$$

[1]If we write $u = (u^1, u^2, \ldots, u^D)$ the following holds when $L = (\mathrm{id} - a^2\Delta)^k$:

$$\int_{\mathbb{R}^D} \big\langle Lu(x), u(x)\big\rangle_{\mathbb{R}^D}\, dx = \int_{\mathbb{R}^D} \sum_{\ell=1}^{D} \sum_{m=0}^{k} \binom{k}{m} a^{2m} \sum_{|\alpha|=m} \big|D^\alpha u^\ell\big|^2\, dx,$$

where we have used the multi-index notation introduced in [**13, 16**]. In particular for $k = 2$ the above expression becomes:

$$\int_{\mathbb{R}^D} \big\langle Lu(x), u(x)\big\rangle_{\mathbb{R}^D}\, dx = \int_{\mathbb{R}^D} \sum_{\ell=1}^{D} \Big\{\big|u^\ell\big|^2 + 2a^2\big\|\nabla u^\ell\big\|^2_{\mathbb{R}^D} + a^4\big\|Hu^\ell\big\|^2_{\mathbb{R}^{D\times D}}\Big\}\, dx,$$

where ∇u^ℓ and Hu^ℓ are, respectively, the gradient and the Hessian matrix of scalar function u^ℓ. We can see that the Sobolev norm is a linear combination of the L^2 norms of a function *and* of its derivatives; parameter a is just a scaling factor.

Set $L^2([0,1],V)$ is a subset of $L^1([0,1],V)$ and is in fact a Hilbert space with inner product $\langle u,v\rangle_{L^2([0,1],V)} \triangleq \int_0^1 \langle u,v\rangle_V \, dt$; we will write the generic element of $L^1([0,1],V)$ or $L^2([0,1],V)$ explicitly as $v_t(x)$, $t\in[0,1]$, $x\in\mathbb{R}^D$.

It is well known from the theory of ordinary differential equations [**7**] that, given a generic vector field $v = v_t(x)$, $t\in[0,1]$, $x\in\mathbb{R}^D$, under some regularity assumptions on v the D-dimensional non-autonomous dynamical system

$$\left\{\begin{array}{rcl} \dot{z} & = & v_t(z) \\ z(t_0) & = & x \end{array}\right. \tag{2.2}$$

has a unique solution of the type $z(t) = \psi(t,t_0,x)$. Let V be an admissible Hilbert space; for any $v \in L^1\big([0,1],V\big)$ we shall define $\varphi^v_{st}(x) \triangleq \psi(t,s,x)$; fixing $t=1$ and $s=0$ we get $\varphi^v(x) \triangleq \varphi^v_{01}(x)$, which is the *diffeomorphism generated* (or *induced*) *by* v. Given an admissible Hilbert space V we will call the set

$$\mathcal{G}_V \triangleq \Big\{\varphi^v : v \in L^1\big([0,1],V\big)\Big\}$$

the *group of diffeomorphisms* generated by V; its name is justified by the result that for any $v\in L^1\big([0,1],V\big)$ the map $\varphi^v : \mathbb{R}^D \to \mathbb{R}^D$ is indeed a diffeomorphism and the set $\mathcal{G}_V$ is a group with respect to the operation of *composition* between functions [**43, 47**]. In the language of manifolds it turns out that $\mathcal{G}_V$ is an infinite-dimensional Lie group and V is its Lie algebra.

1.2. Definition of the distance function. For generic velocity $v\in L^2\big([0,1],V\big)$ and landmark trajectories $q\in\mathcal{Q}$ define the energy

$$E[v,q] \triangleq \int_0^1 \int_{\mathbb{R}^D} \big\langle Lv_t(x), v_t(x)\big\rangle_{\mathbb{R}^D} \, dx\, dt + \lambda \int_0^1 \sum_{i=1}^N \left\| \frac{dq^i}{dt}(t) - v_t\big(q^i(t)\big)\right\|^2_{\mathbb{R}^D} dt. \tag{2.3}$$

We claim that a distance function d on $\mathcal{I}$ between two landmark sets (or shapes) $I = (x^1,x^2,\ldots,x^N)$ and $I' = (y^1,y^2,\ldots,y^N)$ can be defined as

$$d(I,I') \triangleq \inf_{v,q} \Big\{ \sqrt{E[v,q]} : v\in L^2\big([0,1],V\big),\ q\in\mathcal{Q} \text{ with } q(0)=I,\ q(1)=I' \Big\}; \tag{2.4}$$

the main objective of this chapter is in fact to show that the above function is in fact a geodesic distance with respect to a Riemannian metric.

The above infimum is computed over all differentiable landmark paths $q \in \mathcal{Q}$ that satisfy the boundary conditions, and vector fields $v \in L^2\big([0,1], V\big)$. Such fields more or less exactly "transport" (i.e. generate diffeomorphisms that map) the first set of landmarks I into the second one I', depending on the value of smoothing parameter $\lambda \in (0, \infty]$. Such paramenter is a weight between the first term (the aforementioned "kinetic energy", averaged over a unit of time), that measures the smoothness of the vector field that generates the diffeomorphism, and the second term, which measures the exactness of the matching. When $\lambda = \infty$ we have *exact* matching, i.e. the landmark trajectories exactly satisfy the ordinary differential equations $\dot{q}^i = v_t(q^i)$, $i = 1, \ldots, N$ which are obtained by setting the integrands of the second term in the right-hand side of (2.3) equal to zero. When $\lambda < \infty$ we have *regularized* (or *approximate*) matching, i.e. the landmark trajectories "almost" satisfy such ODE, so that the diffeomorphism generated by $v_t(x)$ does not transport I exactly into I'; this allows for the time varying vector field to be smoother, thus resulting in a smaller distance between the two given landmark configurations. For this reason the second term in (2.3) is often referred to as *smoothing term*; its function is to make distance d tolerant of small diffeomorphisms, so that object variations due to *noise* in data are neglected. For the rest of the chapter we shall consider generic values of smoothing parameter λ.

Note that the *smaller* λ is, the smoother will be the vector field and the less exact will be the matching. In fact for $\lambda \simeq 0$ the minimizer (v, q) of energy (2.3) (if it actually exists) has velocity v is close to zero, so that the diffeomorphism φ^v it generates is close to the identity map. This makes the second integral in (2.3) large, but its contribution to $E[v, q]$ is small since λ is almost zero. The corresponding minimizing trajectories q will be (almost) straight lines. We shall return to this discussion further on the chapter, when the dependance on λ of the Riemannian metric tensor of landmarks manifold $\mathcal{I}$ will be clear.

2. Riemannian formulation

We remind the reader that our objective is to show that d defined in (2.4) is a *geodesic distance* on $\mathcal{I}$ with respect to some Riemannian metric, which in fact will depend on the chosen differential operator L and on smoothing parameter λ.

2.1. Minimizing velocity fields and momenta. The following result holds.

Proposition 2.1. *For a fixed* $\bar{q} = \{\bar{q}^i : [0,1] \to \mathbb{R}^D\}_{i=1}^N \in \mathcal{Q}$ *there exists a minimizer with respect to* $v \in L^2([0,1], V)$ *of* $E[v, \bar{q}]$ *and it belongs to the set of vector fields of the general form:*

$$v_t(x) = \sum_{i=1}^{N} \alpha_i(t)\, G\big(x, \bar{q}^i(t)\big), \tag{2.5}$$

where $G : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}$ *is the Green's function (or fundamental solution) of operator* L *in* (2.3) *and coefficients* $\alpha_i : [0,1] \to \mathbb{R}^D$, $i = 1, \ldots, N$ *are continuous functions.*

Note that typically $G(x,y) = \gamma\big(\|x-y\|_{\mathbb{R}^D}\big)$ for some bell-shaped scalar function $\gamma : [0,\infty) \to \mathbb{R}$; function G is sometimes called the *kernel* of space V (in section 4 of this chapter we will briefly return on the topic of kernels). Therefore, for fixed differentiable trajectories $\bar{q}^i(\cdot)$, $i = 1, \ldots, N$, energy $E[v, \bar{q}]$ can be minimized with respect to functions $\alpha_i(\cdot)$, $i = 1, \ldots, N$. We shall indicate with $p_i(\cdot)$, $i = 1, \ldots, N$ the *minimizing* values of coefficients α_i in expression (2.5) and with $v^* = v_t^*(x)$, $t \in [0,1]$, $x \in \mathbb{R}^D$ the resulting *minimizing* velocity field; the coefficients p_i are called *momenta* (once the Riemannian formulation is proven, it will turn that such vectors will be the actual momenta for landmark points, as defined in Hamiltonian mechanics).

Remark. We would like to argue that it makes sense, from a physical point of view, that the minimum-energy velocity field that transports the landmarks along trajectories $\bar{q}$ must be of the form (2.5). In fact what the formula expresses is that the N landmarks are moved around space $\mathbb{R}^D$ by "lumps" of velocity fields centered around each one of the landmarks themselves; such lumps cannot be point-supported, since the global velocity field $v_t(x)$ must minimize a norm of the Sobolev type and therefore be smooth to a certain degree.

Before proving the above proposition and computing the momenta in function of trajectories $\bar{q}$ and their time derivatives, we shall introduce some notation that will allow us to express the results in terms of matrix and vector summations/multiplications. The scalar components of the N landmark trajectories $q^i = \left(q^{i,1}, \ldots, q^{i,D}\right)$, $i = 1, \ldots, N$ can be ordered in an $N \times D$ matrix:

$$\begin{bmatrix} q^{1,1} & q^{1,2} & \cdots & q^{1,D} \\ q^{2,1} & q^{2,2} & \cdots & q^{2,D} \\ q^{3,1} & q^{3,2} & \cdots & q^{3,D} \\ \vdots & \vdots & \ddots & \vdots \\ q^{N,1} & q^{N,2} & \cdots & q^{N,D} \end{bmatrix},$$

where the coordinates of the first, second, ..., N^{th} landmarks lie on the first, first, second, ..., N^{th} rows of the above matrix. In order to keep notation consistent, from now on we shall consider trajectories $q^i(\cdot)$, $i = 1, \ldots, N$ as $1 \times N$ *row* vectors. Instead we will indicate with $q^{(k)}$ the $N \times 1$ *column* vector of the k^{th} components of the set of labeled landmarks:

$$q^{(k)} \triangleq \begin{bmatrix} q^{1,k} \\ q^{2,k} \\ \vdots \\ q^{N,k} \end{bmatrix},$$

for $k = 1, \ldots, D$, while with an abuse of notation we will indicate with q the *superposition* of the above defined $q^{(k)}$'s, i.e.:

$$q \triangleq \begin{bmatrix} q^{(1)} \\ q^{(2)} \\ \vdots \\ q^{(D)} \end{bmatrix}, \tag{2.6}$$

which is a $DN \times 1$ *column* vector (the abuse of notation consists in the fact that we are using the symbol q for both the above vector and the generic element of $\mathcal{Q}$, but the context in which such symbol is used should clarify its meaning).

As far as the scalar components momenta $p_i = \big(p_{i,1}, \ldots, p_{i,D}\big)$, $i = 1, \ldots, N$ are concerned we can order them in a matrix in the same way that we ordered the $q^{i,k}$'s:

$$\begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,D} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,D} \\ p_{3,1} & p_{3,2} & \cdots & p_{3,D} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & p_{N,2} & \cdots & p_{N,D} \end{bmatrix}; \tag{2.7}$$

similarly to what we did before, we will define:

$$p_{(k)} \triangleq \begin{bmatrix} p_{1,k} \\ p_{2,k} \\ \vdots \\ p_{N,k} \end{bmatrix}, \quad \text{for } k = 1, \ldots, D \qquad \text{and} \qquad p \triangleq \begin{bmatrix} p_{(1)} \\ p_{(2)} \\ \vdots \\ p_{(D)} \end{bmatrix},$$

which are *column* vectors of size $N \times 1$ and $DN \times 1$, respectively[2]. In fact, we will also need to order in a matrix the generic coefficients that appear in (2.5) in the same way that the momenta are ordered:

$$\begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,D} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,D} \\ \alpha_{3,1} & \alpha_{3,2} & \cdots & \alpha_{3,D} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{N,1} & \alpha_{N,2} & \cdots & \alpha_{N,D} \end{bmatrix},$$

and vectors $\alpha_{(k)}$, $k = 1, \ldots, D$ and α (which are *column* vectors of size $N \times 1$ and $DN \times 1$, respectively) are defined in an analogous manner. Consistently with the above

[2]We chose *superscript* indices for landmark coordinates and *subscript* indices for momenta because it will turn out that the derivatives of landmark coordinates $\dot{q}$ and momenta p live, repectively, on tangent bundle $T\mathcal{I}$ and on cotangent bundle $T^*\mathcal{I}$ of the Riemannian manifold [**28**].

notation we shall consider velocity fields of the type (2.5) to be *row* vectors. Also, we will define the following $N \times N$ matrix:

$$S(q) \triangleq \begin{bmatrix} G(q^1, q^1) & G(q^1, q^2) & \cdots & G(q^1, q^N) \\ G(q^2, q^1) & G(q^2, q^2) & \cdots & G(q^2, q^N) \\ \vdots & \vdots & \ddots & \vdots \\ G(q^N, q^1) & G(q^N, q^2) & \cdots & G(q^N, q^N) \end{bmatrix} \tag{2.8}$$

where G is the kernel of V; its generic element will be $S^{ij} \triangleq G(q^i, q^j)$, with $i, j = 1, \ldots, N$. Matrix $S(q)$ is definite positive, whence invertible, by Corollary A.5 in Appendix A. Given the above machinery we can proceed to the following proof.

PROOF OF PROPOSITION 2.1. We will use variational calculus techniques. Neglecting the bar above the symbol q and some of the time arguments for notational compactness, for two time-dependent vector fields $v, w \in L^2\big([0,1], V\big)$ and for an arbitrary $\varepsilon \in \mathbb{R}$ we have

$$E[v + \varepsilon w, q] = \int_0^1 \big\langle v_t + \varepsilon w_t, v_t + \varepsilon w_t \big\rangle_V \, dt + \lambda \int_0^1 \sum_{i=1}^N \big\| \dot{q}^i - \big(v_t + \varepsilon w_t\big)(q^i) \big\|^2_{\mathbb{R}^D} \, dt,$$

that is:

$$\begin{aligned} E[v + \varepsilon w, q] \;=\; & \int_0^1 \Big\{ \big\langle v_t, v_t \big\rangle_V + 2\varepsilon \big\langle v_t, w_t \big\rangle_V + \varepsilon^2 \big\langle w_t, w_t \big\rangle_V \Big\} \, dt \\ & + \lambda \int_0^1 \sum_{i=1}^N \Big\{ \big\langle \dot{q}^i - v_t(q^i), \dot{q}^i - v_t(q^i) \big\rangle_{\mathbb{R}^D} - 2\varepsilon \big\langle \dot{q}^i - v_t(q^i), w_t(q^i) \big\rangle_{\mathbb{R}^D} \\ & \qquad + \big\langle w_t(q^i), w_t(q^i) \big\rangle_{\mathbb{R}^D} \Big\} \, dt; \end{aligned}$$

therefore the first variation of the functional is

$$\delta E \triangleq \frac{\partial}{\partial \varepsilon} E[v + \varepsilon w, q] \Big|_{\varepsilon = 0} = 2 \int_0^1 \Big\{ \big\langle v_t, w_t \big\rangle_V - \lambda \sum_{i=1}^N \big\langle \dot{q}^i - v_t(q^i), w_t(q^i) \big\rangle_{\mathbb{R}^D} \Big\} \, dt.$$

By the reproducing property described in Corollary A.4 we have that

$$\big\langle \dot{q}^i - v_t(q^i), w_t \big\rangle_{\mathbb{R}^D} = \big\langle G(\cdot, q^i)\big[\dot{q}^i - v_t(q^i)\big], w_t(q^i) \big\rangle_V$$

for $i = 1, \dots, N$, so that

$$\delta E = 2 \int_0^1 \left\langle v_t - \lambda \sum_{i=1}^N G(\cdot, q^i)\big[\dot{q}^i - v_t(q^i)\big], w_t \right\rangle_V dt.$$

Setting the $\delta E = 0$ for all $w \in L^2\big([0,1], V\big)$ yields:

$$v_t(x) = \sum_{i=1}^N \lambda\big[\dot{q}^i(t) - v_t\big(q^i(t)\big)\big] G\big(x, q^i(t)\big), \qquad t \in [0,1],\ x \in \mathbb{R}^D$$

which is precisely of the form (2.5), having set $\alpha_i(t) = \lambda\big[\dot{q}^i(t) - v_t\big(q^i(t)\big)\big]$. $\square$

The momenta may be computed in the following manner.

PROPOSITION 2.2. *For a fixed* $\bar{q} = \big\{\bar{q}^i : [0,1] \to \mathbb{R}^D\big\}_{i=1}^N \in \mathcal{Q}$ *the minimizer with respect to* $v \in L^2([0,1], V)$ *of* $E[v, \bar{q}]$ *is*

$$v_t^*(x) = \sum_{i=1}^N p_i(t)\, G\big(x, \bar{q}_i(t)\big), \tag{2.9}$$

where the components of momenta (2.7) *are given by:*

$$p_{(k)}(t) = \left(S\big(\bar{q}(t)\big) + \frac{\mathrm{id}}{\lambda}\right)^{-1} \cdot \frac{d}{dt}\bar{q}^{(k)}(t), \tag{2.10}$$

$k = 1, \dots, D$*, where* id *is the* $N \times N$ *identity matrix.*

In order to prove the above result we will need two lemmas.

LEMMA 2.3. *When the velocity field has the form* (2.5)*, its squared norm* $\|v_t\|_V^2$ *at a given time* t *can be expressed as:*

$$\int_{\mathbb{R}^D} \big\langle Lv_t(x), v_t(x)\big\rangle_{\mathbb{R}^D} dx \;=\; \sum_{k=1}^D \alpha_{(k)}^T(t)\, S\big(q(t)\big)\, \alpha_{(k)}(t).$$

PROOF. We are going to use some of the notation and terminology introduced in Appendix A. When the velocity field in form (2.5), i.e. $v_t(x) \;=\; \sum_{i=1}^N \alpha_i\, G(x, q^i)$, we have:

$$\begin{aligned}
\int_{\mathbb{R}^D} \big\langle Lv_t(x), v_t(x)\big\rangle_{\mathbb{R}^D}\, dx \;&=\; \big\langle v_t(x), v_t(x)\big\rangle_V \\
=\;\; \sum_{i,j=1}^{N} \big\langle G(\cdot, q^i)\alpha_i, G(\cdot, q^j)\alpha_j\big\rangle_V \;&\overset{(*)}{=}\; \sum_{i,j=1}^{N} \big\langle \alpha_i, G(q^i, q^j)\alpha_j\big\rangle_{\mathbb{R}^D} \\
=\;\; \sum_{i,j=1}^{N} G(q^i, q^j)\big\langle \alpha_i, \alpha_j\big\rangle_{\mathbb{R}^D} \;&=\; \sum_{k=1}^{D}\sum_{i,j=1}^{N} G(q^i, q^j)\, \alpha_{i,k}\alpha_{j,k} \;=\; \sum_{k=1}^{D} \alpha_{(k)}^T\, S(q)\, \alpha_{(k)},
\end{aligned}$$

where in step $(*)$ we have used Corollary A.4. ☐

LEMMA 2.4. *When the velocity field has the form* (2.5) *the integrand of the smoothing term in energy* (2.3) *can be written as:*

$$\sum_{i=1}^{N} \Big\| \dot q^i(t) - v_t\big(q^i(t)\big)\Big\|^2_{\mathbb{R}^D} = \sum_{k=1}^{D} \Big(S\big(q(t)\big)\, \alpha_{(k)}(t) - \dot q^{(k)}(t)\Big)^T \Big(S\big(q(t)\big)\, \alpha_{(k)}(t) - \dot q^{(k)}(t)\Big).$$

PROOF. We will write the components of the velocity field as $v_t = (v_t^1, \ldots, v_t^D)$. Neglecting the time argument for notational compactness, we have that

$$v_t^k(x) = \sum_{j=1}^{N} \alpha_{j,k}\, G(x, q^j), \qquad \text{whence} \qquad v_t^k(q^i) = \sum_{j=1}^{N} \alpha_{j,k}\, G(q^i, q^j) = \sum_{j=1}^{N} S^{ij}\alpha_{j,k},$$

so that we can write the $N \times 1$ column vector $\dot q^{(k)} - S(q)\alpha_{(k)}$ as

$$\dot q^{(k)} - S(q)\alpha_{(k)} = \begin{bmatrix} \dot q^{1,k} - \sum_{j=1}^{N} S^{1j}\alpha_{j,k} \\ \dot q^{2,k} - \sum_{j=1}^{N} S^{2j}\alpha_{j,k} \\ \vdots \\ \dot q^{N,k} - \sum_{j=1}^{N} S^{Nj}\alpha_{j,k} \end{bmatrix} = \begin{bmatrix} \dot q^{1,k} - v_t^k(q^1) \\ \dot q^{2,k} - v_t^k(q^2) \\ \vdots \\ \dot q^{N,k} - v_t^k(q^N) \end{bmatrix}.$$

Therefore the integrand of the smoothing term in (2.3) can be expressed as:

$$\begin{aligned}
\sum_{i=1}^{N} \Big\| \dot q^i - v_t\big(q^i\big)\Big\|^2_{\mathbb{R}^D} &= \sum_{k=1}^{D}\sum_{i=1}^{N} \Big| \dot q^{i,k} - v_t^k\big(q^i\big)\Big|^2 \;=\; \sum_{k=1}^{D} \Big\| \dot q^{(k)} - S(q)\, \alpha_{(k)}\Big\|^2_{\mathbb{R}^N} \\
=\;\; \sum_{k=1}^{D} \Big\| S(q)\, \alpha_{(k)} - \dot q^{(k)}\Big\|^2_{\mathbb{R}^N} \;&=\; \sum_{k=1}^{D} \Big(S(q)\, \alpha_{(k)} - \dot q^{(k)}\Big)^T \Big(S(q)\, \alpha_{(k)} - \dot q^{(k)}\Big),
\end{aligned}$$

which precisely is what we wanted to prove. ☐

Proof of Proposition 2.2. We will omit the bar above the symbol q and some arguments for notational compactness. For a velocity field of the type (2.5) Lemmas 2.3 and 2.4 imply that, for a velocity field of the type (2.5), the energy (2.3) can be written in the form:

$$E[v,q] = \int_0^1 \sum_{k=1}^D \left\{ \alpha_{(k)}^T S \alpha_{(k)} + \lambda \left(S\alpha_{(k)} - \dot{q}^{(k)} \right)^T \left(S\alpha_{(k)} - \dot{q}^{(k)} \right) \right\} dt.$$

In fact, we will indicate the above functional with $U[\alpha, q]$ since it is now a function of the α coefficients and the landmark trajectories. Once again, we will be using variational principles. For arbitrary functions $\alpha_{(k)}, \beta_{(k)} : [0,1] \to \mathbb{R}^{N\times 1}$, $k = 1, \dots, D$ and any $\varepsilon \in \mathbb{R}$ we have

$$\begin{aligned} U[\alpha + \varepsilon\beta, q] \;\; = \;\; & \int_0^1 \sum_{k=1}^D \Big\{ \big(\alpha_{(k)} + \varepsilon\beta_{(k)}\big)^T S \big(\alpha_{(k)} + \varepsilon\beta_{(k)}\big) \\ & + \lambda \left[S\big(\alpha_{(k)} + \varepsilon\beta_{(k)}\big) - \dot{q}^{(k)} \right]^T \left[S\big(\alpha_{(k)} + \varepsilon\beta_{(k)}\big) - \dot{q}^{(k)} \right] \Big\} \, dt, \end{aligned}$$

that is:

$$\begin{aligned} U[\alpha + \varepsilon\beta, q] = & \int_0^1 \sum_{k=1}^D \Bigg\{ \alpha_{(k)}^T S \alpha_{(k)} + 2\varepsilon\, \alpha_{(k)}^T S \beta_{(k)} + \varepsilon^2 \beta_{(k)}^T S \beta_{(k)} \\ & + \lambda \Big[\big(S\alpha_{(k)} - \dot{q}^{(k)}\big)^T \big(S\alpha_{(k)} - \dot{q}^{(k)}\big) + 2\varepsilon \big(S\alpha_{(k)} - \dot{q}^{(k)}\big)^T S\beta_{(k)} + \varepsilon^2 \beta_{(k)}^T S^T S \beta_{(k)} \Big] \Bigg\} dt. \end{aligned}$$

Therefore the first variation of functional U with respect to α is given by:

$$\delta U \triangleq \frac{\partial}{\partial \epsilon} U[\alpha + \varepsilon\beta, q] \Big|_{\varepsilon = 0} = 2 \int_0^1 \sum_{k=1}^D \left\{ \alpha_{(k)}^T S + \lambda \big(S\alpha_{(k)} - \dot{q}^{(k)} \big)^T S \right\} \beta_{(k)} \, dt,$$

which is equal to zero for any choice of functions $\beta_{(k)}$, $k = 1, \dots, D$ if and only if

$$\alpha_{(k)}^T S + \lambda \big(S\alpha_{(k)} - \dot{q}^{(k)} \big)^T S = 0, \qquad k = 1, \dots, D.$$

Matrix S is positive definite, whence invertible, therefore the above equations are equivalent to

$$\alpha_{(k)} + \lambda \big(S\alpha_{(k)} - \dot{q}^{(k)} \big) = 0, \qquad k = 1, \dots, D$$

which immediately yield:

$$\alpha_{(k)} = \left(S + \frac{\mathrm{id}}{\lambda} \right)^{-1} \dot{q}^{(k)} \qquad k = 1, \dots, D,$$

that are precisely equations (2.10). $\square$

2.2. Main result. It is convenient, at this point, to introduce the $N \times N$ matrix:

$$R(q) \triangleq \left(S(q) + \frac{\mathrm{id}}{\lambda}\right)^{-1}$$

where $S(q)$ was defined in (2.8) and also the $DN \times DN$, block-diagonal matrix

$$g(q) \triangleq \begin{bmatrix} R(q) & 0 & \cdots & 0 \\ 0 & R(q) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R(q) \end{bmatrix} = \begin{bmatrix} S(q) + \frac{\mathrm{id}}{\lambda} & 0 & \cdots & 0 \\ 0 & S(q) + \frac{\mathrm{id}}{\lambda} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S(q) + \frac{\mathrm{id}}{\lambda} \end{bmatrix}^{-1}, \tag{2.11}$$

where the $N \times N$ block $R(q)$ is repeated exactly D times (the zeroes above represent $N \times N$ blocks of zeroes); the choice of symbol g in the definition above is justified, as we shall see in Theorem 2.7, by the fact that (2.11) is precisely the *metric tensor* of Riemannian manifold $\mathcal{I}$, and g is the classic symbol used for such tensor in Differential Geometry. We will need the following technical result.

LEMMA 2.5. *Given two generic sets $\mathcal{S}$ and $\mathcal{W}$, consider a function $f : \mathcal{S} \times \mathcal{W} \to \mathbb{R}$ and assume that, for any fixed $\bar{w} \in \mathcal{W}$, function $f(s, \bar{w})$ has a unique minimizer with respect to $s \in \mathcal{S}$: we shall define $h(\bar{w}) \triangleq \arg\min_{s \in \mathcal{S}} f(s, \bar{w})$. Under these hypotheses,*

$$\inf_{(s,w) \in \mathcal{S} \times \mathcal{W}} f(s, w) = \inf_{w \in \mathcal{W}} f(h(w), w).$$

PROOF. Let $\nu_1 \triangleq \inf_{(s,w) \in \mathcal{S} \times \mathcal{W}} f(s, w)$ and $\nu_2 \triangleq \inf_{w \in \mathcal{W}} f(h(w), w)$. It is obvious that $\nu_1 \leq \nu_2$: in fact we have that $\nu_2 = \inf_{(s,w) \in \mathrm{graph}(h)} f(s, w) \geq \nu_1$, since $\mathrm{graph}(h) \subset U \times V$. We now want to prove that $\nu_1 \geq \nu_2$. But if $\nu_1 < \nu_2$ then by the definition of ν_1 (a number μ is the infumum of a set $\mathcal{T} \subset \mathbb{R}$ if: (a) for all $x \in \mathcal{T}$, we have $x \geq \nu$; (b) for any choice of $\mu' > \mu$, $\exists\, y \in \mathcal{T}$ such that $y < \mu'$) there would exist a pair $(s^*, w^*) \in \mathcal{S} \times \mathcal{W}$ such that $f(s^*, w^*) < \nu_2$. By the definition of h we would have $f(h(w^*), w^*) \leq f(s^*, w^*) < \nu_2$, which is a contradiction of the definition of ν_2. $\square$

For a fixed pair of landmark sets $I = \{x_1, \ldots, x_N\}$ and $I' = \{y_1, \ldots, y_N\}$ we can apply the above Lemma to sets

$$\mathcal{S} = L^2([0,1], V) \quad \text{and} \quad \mathcal{W} = \mathcal{Q}_0 \triangleq \big\{q \in \mathcal{Q} : q(0) = I, q(1) = I'\big\},$$

and to function $f = E[v, q]$. Velocity v^* provided by expression (2.9) in Proposition 2.2 minimizes $E[v, \bar{q}]$ for a fixed trajectory $\bar{q}$, so that $v^* = h(\bar{q})$; in fact, with an abuse of notation, we shall write $v^* = v^*(\bar{q})$. The following consequence holds.

Corollary 2.6. *Fix landmark sets $I = (x^1, \ldots, x^N)$ and $I' = (y^1, \ldots, y^N)$. For an arbitrary $\bar{q} \in \mathcal{Q}_0$ let velocity $v^* = v^*(\bar{q})$ be the minimizer of $E[v, \bar{q}]$ provided by Proposition 2.2. Then the function d defined in* (2.4) *is such that*

$$d(I, I')^2 \triangleq \inf_{(v,\bar{q}) \in L^2([0,1],V) \times \mathcal{Q}_0} E\big[v, \bar{q}\big] = \inf_{\bar{q} \in \mathcal{Q}_0} E\big[v^*(\bar{q}), \bar{q}\big].$$

If there exists a unique $q^* \in \mathcal{Q}_0$ such that $q^* = \arg\min_{\bar{q} \in \mathcal{Q}_0} E\big[v^*(\bar{q}), \bar{q}\big]$ then

$$\arg \min_{(v,\bar{q}) \in L^2([0,1],V) \times \mathcal{Q}_0} E\big[v, \bar{q}\big] = \big(v^*(q^*), q^*\big),$$

where $v^*(q^*)$ is the minimizer with respect to v of energy $E[v, q^*]$, again provided by Proposition 2.2; under these conditions, $d(I, I') = \sqrt{E\big[v^*(q^*), q^*\big]}$.

Proof. First note that the square root and the "inf" can be exchanged since the square root is a monotone increasing function. The first part of the corollary follows directly from Lemma 2.5. If $q^* = \arg\min_{\bar{q} \in \mathcal{Q}_0} E[v(\bar{q}), \bar{q}]$ is well defined then $\inf_{\bar{q} \in \mathcal{Q}_0} E\big[v^*(\bar{q}), \bar{q}\big] = E\big[v^*(q^*), q^*\big]$, so that

$$\big(v^*(q^*), q^*\big) = \arg \inf_{(v,\bar{q}) \in L^2([0,1],V) \times \mathcal{Q}_0} E\big[v, \bar{q}\big]$$

and $d(I, I') = \sqrt{E\big[v^*(q^*), q^*\big]}$. Note that the above infimum is actually achieved by a pair in $L^2([0,1], V) \times \mathcal{Q}_0$, and therefore is a minimum. $\square$

Inserting $v^*(\bar{q})$ into $E[v, \bar{q}]$ yields a new energy expression that only depends on $\bar{q}$, which we shall indicate with $\widetilde{E}[\bar{q}]$. In fact, given the arbitrariness of the choice of trajectory $\bar{q}$ from now on we will *drop the bar* above symbol q so that we can write:

$$\widetilde{E}[q] \triangleq E\big[v^*(q), q\big], \quad q \in \mathcal{Q}$$

where $v^*(q)$ is the minimizer of $E[v, q]$ with respect to v, given by Proposition 2.2. By the above corollary, $d(I, I') = \inf_{q \in \mathcal{Q}_0} \sqrt{\widetilde{E}[q]}$. We are not ready to prove the main result of this chapter; a discussion will follow.

THEOREM 2.7. *For an arbitrary landmark trajectory $q \in \mathcal{Q}$ it turns out that $\widetilde{E}[q]$ has the form:*

$$\widetilde{E}[q] = \int_0^1 \dot{q}(t)^T g\big(q(t)\big)\, \dot{q}(t)\, dt, \tag{2.12}$$

where $\dot{q}$ is the derivative of the $DN \times 1$ column vector q defined by (2.6) *and g is the $DN \times DN$ matrix defined by* (2.11).

PROOF. Let $q \in \mathcal{Q}$ be an arbitrary set of landmark trajectories, and $v^*(q)$ be the minimizer with respect to v of $E[v, q]$, provided by Proposition 2.2. Applying Lemmas 2.3 and 2.4 to formula (2.9) rather than to formula (2.5) allows one to write:

$$\begin{aligned}
\widetilde{E}[q] &= E[v^*(q), q] \\
&= \int_0^1 \int_{\mathbb{R}^D} \big\langle L v_t^*(x), v_t^*(x) \big\rangle_{\mathbb{R}^D}\, dx\, dt + \lambda \int_0^1 \sum_{i=1}^N \left\| \frac{dq^i}{dt}(t) - v_t^*(q^i(t)) \right\|_{\mathbb{R}^D}^2 dt \\
&= \int_0^1 \sum_{k=1}^D p_{(k)}^T\, S(q)\, p_{(k)}\; dt + \lambda \int_0^1 \sum_{k=1}^D \Big(S(q)\, p_{(k)} - \dot{q}^{(k)} \Big)^T \Big(S(q)\, p_{(k)} - \dot{q}^{(k)} \Big)\; dt \\
&= \int_0^1 \sum_{k=1}^D \Big\{ p_{(k)}^T\, S(q)\, p_{(k)} + \lambda \Big(S(q)\, p_{(k)} - \dot{q}^{(k)} \Big)^T \Big(S(q)\, p_{(k)} - \dot{q}^{(k)} \Big) \Big\}\; dt\,.
\end{aligned}$$

Inserting the formula for the momenta $p_{(k)} = \Big(S(q)+\frac{\mathrm{id}}{\lambda}\Big)^{-1}\dot{q}^{(k)}$, provided by (2.10), into the above expression yields[3]

$$\begin{aligned}\widetilde{E}[q] \;=\; & \int_0^1 \sum_{k=1}^{D} \Big\{ (\dot{q}^{(k)})^T \Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-T} S \Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} \dot{q}^{(k)} \\ & +\lambda \Big[S\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1}\dot{q}^{(k)} - \dot{q}^{(k)}\Big]^T \Big[S\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1}\dot{q}^{(k)} - \dot{q}^{(k)}\Big] \Big\}\, dt,\end{aligned}$$

which we may rewrite as

$$\begin{aligned}\widetilde{E}[q] \;=\; & \int_0^1 \sum_{k=1}^{D} (\dot{q}^{(k)})^T \Big\{ \Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} S \Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} \\ & +\lambda \Big[S\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} - \mathrm{id}\Big]^T \Big[S\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} - \mathrm{id}\Big] \Big\} (\dot{q}^{(k)})\; dt. \qquad (2.13)\end{aligned}$$

It is the case that

$$\Big[\mathrm{id} - S\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1}\Big] \cdot \lambda \Big(S+\frac{\mathrm{id}}{\lambda}\Big) \;=\; \Big[\Big(S+\frac{\mathrm{id}}{\lambda}\Big) - S\Big]\lambda \;=\; \mathrm{id},$$

whence

$$\mathrm{id} - S\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} = \frac{1}{\lambda}\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1}.$$

Inserting the above expression into (2.13) yields:

$$\begin{aligned}\widetilde{E}[q] &= \int_0^1 \sum_{k=1}^{D} (\dot{q}^{(k)})^T \Big[\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} S \Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} + \frac{1}{\lambda}\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1}\Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} \Big] \dot{q}^{(k)}\, dt \\ &= \int_0^1 \sum_{k=1}^{D} (\dot{q}^{(k)})^T \Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} \Big[S+\frac{\mathrm{id}}{\lambda}\Big] \Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} \dot{q}^{(k)}\; dt \\ &= \int_0^1 \sum_{k=1}^{D} (\dot{q}^{(k)})^T \Big(S+\frac{\mathrm{id}}{\lambda}\Big)^{-1} \dot{q}^{(k)}\; dt \;=\; \int_0^1 \dot{q}^T g(q)\, \dot{q}\; dt\,,\end{aligned}$$

which is precisely the statement of the theorem. □

[3]We are using the notation $A^{-T} \triangleq \big(A^{-1}\big)^T$; if A is symmetric then $A^{-T} = A^{-1}$.

2.3. Discussion. First and foremost, since matrix $g : \mathbb{R}^{DN} \to \mathbb{R}^{DN\times DN}$ is positive definite for any landmark set we have that expression (2.12) provides a *Riemannian energy function* [**11, 23, 27**]. Therefore the set of landmarks $\mathcal{I}$ is endowed with a Riemannian structure, where g is the *metric tensor*, *geodesic curves* are the extrema of energy functional (2.12), and the *geodesic distance* between two shapes $I, I' \in \mathcal{I}$ is given by $d(I, I') = \inf_{q\in\mathcal{Q}_0} \sqrt{\widetilde{E}[q]}$, where $\mathcal{Q}_0 = \left\{q \in \mathcal{Q} : q(0) = I, q(1) = I'\right\}$. Note that the (infinite-dimensional) diffeomorphism group $\mathcal{G}_V$ and its Lie algebra V have formally "disappeared" from the energy, their information being incorporated into the metric tensor of the Riemannian shape manifold $\mathcal{I}$, of *finite* dimension $n = DN$; the metric tensor in fact depends on the kernel G of space V, and on smoothing parameter λ. In following chapters we will investigate the geometry of the Riemannian structure of $\mathcal{I}$, namely its curvature.

Since metric tensor g is block diagonal with D blocks and the $DN \times 1$ vector q can be partitioned in precisely the D vectors of the components of the landmarks set, at a first look it may seem that the D dimensions of q could be treated in an independent manner from one another, i.e. that we could split enargy (2.12) into D "uncoupled" integrals

$$\tilde{E}[q] = \sum_{k=1}^{D} \int_0^1 \left(\dot{q}^{(k)}\right)^T \left(S(q) + \frac{\mathrm{id}}{\lambda}\right)^{-1} \dot{q}^{(k)}\, dt,$$

and minimize them separately. In fact (while the above formula is correct) this is *not* the case since *all* the components appear in the argument of matrix $S(q) = \left[G(q^i, q^j) + \frac{\delta^{ij}}{\lambda}\right]_{1\le i,j\le N}$.

We should note that for small values of parameter λ the metric tensor g gets close (up to a multiplicative constant) to the $DN \times DN$ identity matrix, i.e. the diagonal elements become far larger than the off-diagonal elements, and they are all approximately given by $1/\lambda$. This means that the metric converges to the Euclidean metric in $\mathbb{R}^{DN}$ for $\lambda \to 0$ and the geodesic curves converge to straight lines, as we anticipated at the end of section 1 of this chapter.

It is well known from the theory of Riemannian manifolds that energies of the type (2.12) may not have a unique minimizer. However, if such length-minimizing geodesic path q^* exists and is numerically computable, by Corollary 2.6 one can also compute the "other half" of the minimizer of energy (2.3), i.e. the velocity field v^* such that $(v^*, q^*) = \arg\min_{v,q} E[v, q]$ by applying formula (2.9) of Proposition 2.2 to trajectory q^*. Consequently, one can also (numerically) compute the corresponding diffeomorphism φ^{v^*} by implementing the system of ordinary differential equations (2.2).

Last, but not least we should remark that the Lagrangian function that corresponds to energy (2.12) is:

$$\mathcal{L}(q, \dot{q}) \;=\; \frac{1}{2}\,\dot{q}^T S(q)\dot{q} \;=\; \frac{1}{2}\sum_{\ell=1}^{D} \left(\dot{q}^{(\ell)}\right)^T \left(S(q) + \frac{\mathrm{id}}{\lambda}\right)^{-1} \dot{q}^{(\ell)}; \tag{2.14}$$

it is easy to see that the *momenta*, which we defined as the coefficients of the velocity field of the form (2.5) that minimizes energy $E[v, q]$ with respect to v, actually coincide with the momenta of landmark points in the sense of classical mechanics [**2**]. In fact in the Hamiltonian formalism the momenta are defined as $p_{i,k} = \frac{\partial \mathcal{L}}{\partial q^{i,k}}(q, \dot{q})$, $i = 1, \ldots, N$, $k = 1, \ldots, D$; that is, in vector notation, $p_{(k)} = \frac{\partial \mathcal{L}}{\partial q^{(k)}}$, $k = 1, \ldots, D$. Applying such definition to (2.14) yields

$$p_{(k)} = \left(S(q) + \frac{\mathrm{id}}{\lambda}\right)^{-1} \dot{q}^{(k)}, \qquad k = 1, \ldots, D$$

which coincide with equations (2.10) of Proposition 2.2. Whence the use of the term momenta is justified. Chapter 3 is dedicated to momenta for landmarks, Hamilton's equations (cogeodesic flow), and conservation laws.

3. Numerical Examples

In this section we briefly illustrate the qualitative behavior of geodesics connecting pairs of landmark shapes $I = (x^1, \ldots, x^N)$ and $I' = (y^1, \ldots, y^N)$ on the plane ($D = 2$). In the figures that follow black dots and circles are, respectively, the initial and final

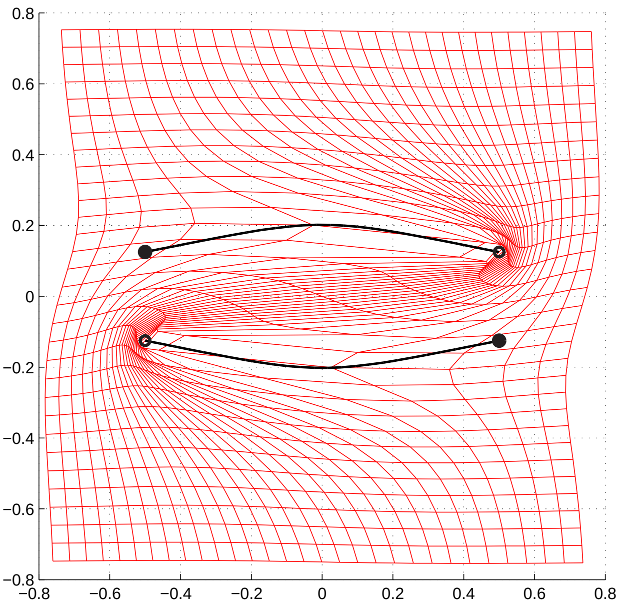


FIGURE 2.1. Geodesic curve for landmarks traveling in *opposite* directions; black dots and circles represent initial and final positions, respectively. The corresponding diffeomorphism φ_{01}^v is also shown.

landmark sets I and I'; the curves represent, in principle[4], the length minimizing ($2N$-dimensional) geodesic between the initial and final configurations. In each of the figures that follow the diffeomorphism φ_{01}^v (induced by the velocity (2.9) that corresponds to the depicted landmark trajectories) is also shown. In all the examples smoothing parameter λ is set to a high (but finite) value.

Figure 2.1 shows a case with $N = 2$ where one landmark must travel from the left to the right and the other must do the opposite. The qualitative behavior of the geodesic $q(t) = \big(q^1(t), q^2(t)\big)$, $t \in [0, 1]$ is such that the two arcs "repel" each other: in fact if the two landmarks traveled too close to each other the vertical derivative of the horizontal component of the velocity field $v_t(x)$ would be large in a neighborhood

[4]The curves were computed a *conjugate gradient* [**6, 39**] descent algorithm on the energy for fixed boundary conditions and it may well be the case that the curves are, in fact, points of local minimum of the functional. An alternative method for solving the boundary value problem is the *geodesic shooting* method [**35**], which we did not implement here.

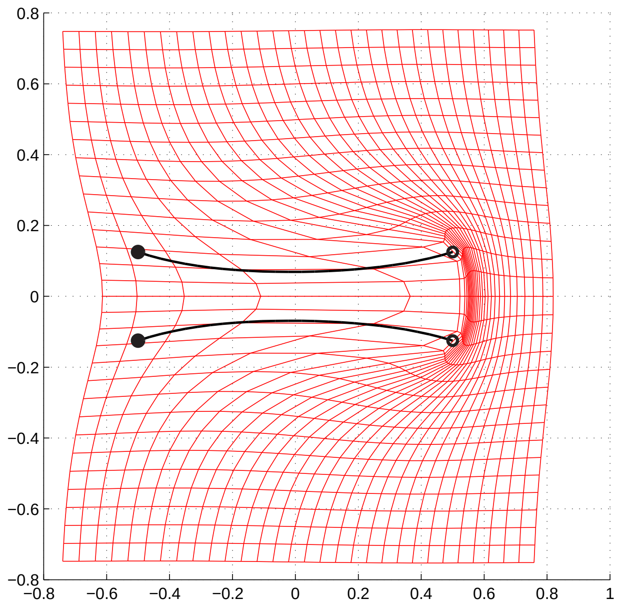


FIGURE 2.2. Geodesic curve for landmarks traveling in *the same* direction; black dots and circles represent initial and final positions, respectively. The corresponding diffeomorphism φ^v_{01} is also shown.

of the origin $x = (0,0)$ and time $t = \frac{1}{2}$, thus giving a strong contribution to the *second* term $\int_0^1 \int_{\mathbb{R}^2} \sum_{\ell=1}^2 \|\nabla v_t^\ell(x)\|^2_{\mathbb{R}^2}\, dx\, dt$ of the Sobolev norm[5] of the velocity.

On the other hand, the opposite happens when the two landmarks must travel in the same direction, as shown in Figure 2.2: the two arcs of the geodesic "attract" each other. The two landmarks tend to "carpool", i.e. to use a velocity field with the smallest possible support in order to minimize the *first* term $\int_0^1 \int_{\mathbb{R}^2} \sum_{\ell=1}^2 |v_t^\ell(x)|^2\, dx\, dt$ of the Sobolev norm of the velocity field.

Finally, Figure 2.3 depicts a somewhat more complex situation where four corners of a square are moved (in this case, $N = 4$). We shall limit ourselves to note that the shape of the four arcs of the (8-dimensional) geodesic curve are mostly determined by the fact that the bottom-right landmark must take the longest journey in a unit of time, whence traveling the fastest and causing the top-right and bottom-left landmarks to initially "pull away" from it. This sudden evasive maneuver causes the

[5]See footnote 1.

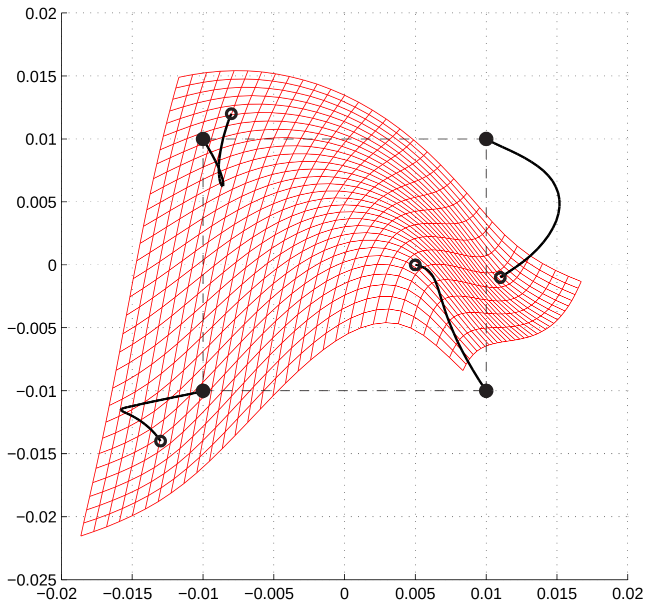


FIGURE 2.3. Deformation of a square; black dots and circles represent initial and final positions, respectively. The corresponding diffeomorphism φ_{01}^{v} is also shown.

remaining (top-left) landmark to initially travel towards the center of the square, but eventually turning back so to be able to reach its designed destination.

4. More on Kernels

The Green's function of differential operator $L = (\mathrm{id} - a^2\Delta)^k$ in $\mathbb{R}^D$ has the form $G(x, y) = \gamma\big(||x - y||_{\mathbb{R}^D}\big)$, with $\gamma : (0, \infty) \to \mathbb{R}$ given by the bell-shaped function [**20**]:

$$\gamma(\varrho) = \frac{2\pi^k}{(k-1)!\,(2\pi a)^{k+\frac{D}{2}}}\,\varrho^{k-\frac{D}{2}}\,K_{k-\frac{D}{2}}\Big(\frac{\varrho}{a}\Big), \tag{2.15}$$

where K_ν is a modified Bessel function [**1**, Ch. 9]; γ can be extended to 0 by continuity. Such functions are referred to as Bessel kernels or Sobolev kernels (since they are the inverse of the differential operators that define the Sobolev norms). One can prove (see Appendix B) that (2.15) is a solution to the ordinary differential equation

$$\gamma'' = \frac{2\nu - 1}{\varrho}\,\gamma' + \frac{1}{a^2}\,\gamma\,, \tag{2.16}$$

where $\nu = k - \frac{D}{2}$, with the appropriate choice of boundary conditions that can be directly derived from (2.15); Propositions B.4 and B.5 describe the behavior of $\gamma(\varrho)$ in a neighborhood of zero.

We should however note that the whole theory we have presented so far would still hold if, instead of starting from an admissible Hilbert space V of the Sobolev kind and then building its kernel G (as the Green's function of the differential operator L that defines the norm on V), we started from a generic continuous, positive definite scalar kernel $G \in L^2(\mathbb{R}^D \times \mathbb{R}^D, \mathbb{R})$; the Mercer-Hilbert-Schmidt theorem [**40, 45, 47**] would then allow us to reconstruct the Hilbert space V (which, in general, will not be of the Sobolev kind) that has G as its reproducing kernel, i.e. such that $\langle G(\cdot, x)\alpha, v\rangle_V = \langle \alpha, v(x)\rangle_{\mathbb{R}^D}$ for any point $x \in \mathbb{R}^D$, vector $\alpha \in \mathbb{R}^D$, and function $v \in V$ (as it is the case for kernels for Sobolev-type admissible Hilbert spaces, see Corollary A.4). The energy to be minimized with respect to $v \in L^2([0,1], V)$ and $q \in \mathcal{Q}$ (with the appropriate boundary conditions) would the general form:

$$E[v,q] \triangleq \int_0^1 \|v_t\|_V^2 \, dt + \lambda \int_0^1 \sum_{i=1}^N \left\| \frac{dq^i}{dt}(t) - v_t\big(q^i(t)\big) \right\|_{\mathbb{R}^D}^2 dt,$$

which of course coincides with (2.3) when the norm on space V is of the Sobolev type (2.1). Proposition 2.2 and Theorem 2.7 would still hold. The discussion on how to construct an admissible space from a kernel goes well beyond the scope of this introductory chapter, and the reader is referred to [**47**] for further details. However, in the future we will use functions γ not given by (2.15) but that are of a form which is easier to manipulate both analytically and numerically, such as Gaussians

$$\gamma(\varrho) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big\{ -\frac{\varrho^2}{2\sigma^2} \Big\}, \tag{2.17}$$

or heavy-tailed Cauchy-type functions

$$\gamma(\varrho) = \frac{1}{1 + a^2\varrho^2}\,. \tag{2.18}$$

We will see in Chapter 5 that the Riemannian curvature tensor and sectional curvature for the landmarks manifold can be expresses in terms of function γ and its first and second derivatives γ' and γ''; therefore, for a specific choice of the kernel, from a computational point of view it is convenient to know how γ and its derivatives are related to each other. In the case of kernels of the Sobolev form, i.e. deriving from an admissible Hilbert space with norm (2.1), such relationship is precisely provided by (2.16); in the case of *Gaussian* type kernels (2.17) the first and second derivatives are related to γ simply by

$$\gamma'(\varrho) = -\frac{\varrho}{\sigma^2}\,\gamma(\varrho) \quad \text{and} \quad \gamma''(\varrho) = \frac{1}{\sigma^2}\Big(\frac{\varrho}{\sigma^2} - 1\Big)\,\gamma(\varrho), \tag{2.19}$$

respectively; on the other hand, in the case of *Cauchy* kernels (2.18) it is easy to get equations

$$\gamma'(\varrho) = -2a^2\varrho\gamma^2(\varrho) \quad \text{and} \quad \gamma''(\varrho) = 8a^4\varrho^2\gamma^3(\varrho) - 2a^2\gamma^2(\varrho). \tag{2.20}$$

We will use these relationships later on in Chapter 5.

5. Further Generalizations

In this final section we shall briefly illustrate how the theory of landmark manifolds that we have presented so far may be thought of as a particular case of a more general approach to shape analysis, based on Lie groups of diffeomorphisms acting on shape manifolds. Such general approach is presented in [**36, 47**] and here we shall limit ourselves to summarizing it, noting that it is applicable to landmarks, images, planar curves, surfaces, and currents [**17**].

5.1. Extensions to generic Shapes. Given a manifold $\mathcal{I}$ of "objects" (e.g. landmark configurations, images, curves, etc.) consider a Lie group $\mathcal{G}$ that acts on $\mathcal{I}$, i.e. a map $\mathcal{G} \times \mathcal{I} \to \mathcal{I} : (g, I) \mapsto g \cdot I$. The group also acts on $\mathcal{A} = \mathcal{G} \times \mathcal{I}$ by the operation $\mathcal{G} \times \mathcal{A} \to \mathcal{A} : (g, (h, I)) \mapsto (gh, g \cdot I)$. Assuming we are given a left-invariant distance $D(\cdot, \cdot)$ on $\mathcal{A}$ (i.e. $D(g{\cdot}a, g{\cdot}a') = D(a, a')$, for all $a, a' \in \mathcal{A}$ and $g \in \mathcal{G}$) this induces a (pseudo-)distance on $\mathcal{I}$ by defining: $d(I, I') = \inf\big\{D\big((g, I), (g, I')\big) : g, g' \in \mathcal{G}\big\}$.

This simple idea can be applied to groups $\mathcal{G}$ of diffeomorphisms $\mathbb{R}^D \to \mathbb{R}^D$ (with group operation $\varphi\psi = \psi \circ \phi$) acting on shape manifolds $\mathcal{I}$, such as sets of D-dimensional landmarks or images (described as scalar functions $I : \mathbb{R}^D \to \mathbb{R}$, with $D = 2$). In fact we can construct a left-invariant metric $\langle\cdot, \cdot\rangle_{(g,I)}$ on $\mathcal{A} = \mathcal{G} \times \mathcal{I}$, that keeps into account both the magnitude of the diffeomorphism generated by vector fields on $\mathcal{G}$ and variations in the structure of objects in $\mathcal{I}$. There are several advantages in designing a left-invariant metric on $\mathcal{G} \times \mathcal{A}$ (and not just on $\mathcal{G}$): for example, the resulting distance d on $\mathcal{I}$ is tolerant to small group actions, so that object variations due to "noise" are neglected; also, we shall see later that in the case of *image matching* the approach allows to variations in the image values themselves, and not just in the geometry via a diffeomorphism of the domain. We should also note that, from a purely mathematical point of view, left-invariance is a *conditio sine qua non* for d to be a distance on $\mathcal{I}$.

The energy of a smooth path $a : [0,1] \mapsto \mathcal{A} : t \mapsto a_t = (g_t, I_t)$, is

$$E[a] \;=\; \int_0^1 \left\| \frac{da_t}{dt} \right\|_{a_t}^2 dt \;=\; \int_0^1 \left\| \left(\frac{dg_t}{dt}, \frac{dI_t}{dt} \right) \right\|_{(g_t, I_t)}^2 dt\,;$$

it turns out that since the metric $\langle \cdot, \cdot \rangle_{(g,I)}$ was constructed to be left-invariant, the geodesic distance on $\mathcal{A}$: $D(b, b') = \big\{ \sqrt{E[a]} : a(0) = b, a(1) = b' \big\}$ is left-invariant with respect to $\mathcal{G}$ actions. By the simple construction above, we have that $d(I, I') = \inf \big\{ D\big((g, I), (g, I')\big) : g, g' \in \mathcal{G} \big\}$ is in fact a distance on $\mathcal{I}$. Using the left-invariance of the metric it is often convenient to rewrite the above energy as follows:

$$\begin{aligned}
E[a] &= \int_0^1 \left\| \big(L_{g_t^{-1}}\big)_* \Big(\frac{dg_t}{dt}, \frac{dI_t}{dt} \Big) \right\|_{\left(g_t^{-1} g_t,\, L_{g_t^{-1}}(I_t)\right)}^2 dt \\
&= \int_0^1 \left\| \Big(\frac{dg_t}{dt} \circ g_t^{-1}, \big(L_{g_t^{-1}}\big)_* \frac{dI_t}{dt} \Big) \right\|_{\left(e,\, g_t^{-1} \cdot I_t\right)}^2 dt \\
&= \int_0^1 \left\| \Big(v_t, \frac{d}{dt} \big(g_s^{-1} \cdot I_t\big)\Big|_{s=t} \Big) \right\|_{(e, J_t)}^2 dt =: E[v, J],
\end{aligned} \tag{2.21}$$

where we have set $v_t := \frac{dg_t}{dt} \circ g_t^{-1} \in T_e\mathcal{G}$ and $J_t := g_t^{-1} \cdot I_t \in \mathcal{I}$ (we have indicated with L_φ both the left action on $\mathcal{G} \times \mathcal{I}$ and on $\mathcal{I}$, and with $(L_\varphi)_*$ the corresponding pushforward map [**28**] on the corresponding tangent spaces). It turns out that the pair of functions (g_t, I_t) is uniquely determined by (v_t, J_t), whence with an abuse of notation we can indicate the above energy as $E[v, J]$. The induced distance on $\mathcal{I}$ is $d(I, I') = \inf \big\{ \sqrt{E[v, J]} : v_t \in T_e\mathcal{G}, J_0 = I, J_1 = I' \big\}$. We will now illustrate how this general approach applies to the manifold of images and the manifold of landmarks.

5.2. Examples. Depending on how the *object space* $\mathcal{I}$ is defined, on how the diffeomorphism group $\mathcal{G}$ acts on it and on how the left-invariant metric on $\mathcal{A} = \mathcal{G} \times \mathcal{I}$ is constructed, the energy (to be minimized) takes different forms. For example, first assume that manifold $\mathcal{I}$ is the set of *scalar images* $I : [0,1]^2 \to \mathbb{R}$, that $\mathcal{G}$ is a group of diffeomorphisms on $[0,1]^2$ (that leave the boundary unchanged), and that the group action is given by composition: $g \cdot I = I \circ g$. Let V be the Lie algebra $T_e\mathcal{G}$ of Lie group $\mathcal{G}$; we will assume that V is an *admissible* Hilbert space (which is obviously also a constraint on the group of diffeomorphisms $\mathcal{G}$ that V generates). If the left-invariant metric is such that

$$\big\|(v,\xi)\big\|^2_{(e,I)} = \|v\|^2_V + \lambda \int_{[0,1]^2} |\xi|^2 \, dx \tag{2.22}$$

for all $I \in \mathcal{I}$, $v \in V = T_e\mathcal{G} \subset C^\infty(\mathbb{R}^2, \mathbb{R}^2)$, and $\xi \in T_I\mathcal{I}$, then energy (2.21) takes the form

$$E[v,J] = \int_0^1 \big\|v_t\big\|^2_V \, dt + \lambda \int_0^1 \int_{[0,1]^2} \Big|\frac{\partial J_t}{\partial t} + \big\langle \nabla_x J, v_t \big\rangle\Big|^2 \, dx \; dt,$$

which in the case of a Sobolev-type admissible Hilbert space V looks like:

$$E[v,J] = \int_0^1 \int_{[0,1]^2} \big\langle Lv_t(x), v_t(x) \big\rangle_{\mathbb{R}^2} \, dx \, dt + \lambda \int_0^1 \int_{[0,1]^2} \Big|\frac{\partial J_t}{\partial t} + \big\langle \nabla_x J, v_t \big\rangle\Big|^2 \, dx \; dt.$$

Note that the integrand in the second term on the right-hand side is the well-known *optical flow constraint equation* [**19**], or *transport equation* for J, so that the first integral penalizes large variations due to the diffeomorphism of the domain, while the second one penalizes *violations* of the optical flow constraint equation, i.e. variations of pixel intensity along the flow. This also allows for matching two images with different luminance values. The induced distance on $\mathcal{I}$ between two images I and I' is given by $d(I, I') = \inf\big\{\sqrt{E[v,J]} : v_t \in T_e\mathcal{G}, \; J_0 = I, \; J_1 = I'\big\}$.

Assume now that $\mathcal{I}$ is the space of *landmark points*, that is, the generic element is given by $I = (x_1, x_2, \ldots, x_N)$, $x_i \in \mathbb{R}^D$, with $x_i \neq x_j$ for $i \neq j$. If $\mathcal{G}$ is a group of diffeomorphisms $\mathbb{R}^D \to \mathbb{R}^D$ that leave the point at infinity unchanged, the group action is $g \cdot I = g^{-1}(I) = (g^{-1}(x_1), \ldots, g^{-1}(x_N))$ and the left-invariant metric is such that

$$\big\|(v,\xi)\big\|^2_{(e,I)} = \big\|v\big\|^2_V + \lambda \sum_{i=1}^{N} \big\|x_i\big\|^2_{\mathbb{R}^D}$$

for all $I \in \mathcal{I}$, $v \in V = T_e\mathcal{G} \subset C^\infty(\mathbb{R}^D, \mathbb{R}^D)$, and $\xi \in T_I\mathcal{I} = \mathbb{R}^{DN}$, then energy (2.21) takes the form

$$E[v,q] = \int_0^1 \int_{\mathbb{R}^D} \big\langle Lv_t(x), v_t(x)\big\rangle_{\mathbb{R}^D}\, dx\, dt + \lambda \int_0^1 \sum_{i=1}^{N} \Big\|\frac{dq_i}{dt}(t) - v_t(q_i(t))\Big\|^2_{\mathbb{R}^d}\, dt,$$

where we have assumed that the admissible Hilbert space is of the Sobolev type; in the above equation we have indicated with $q(t) = (q_1(t), \ldots, q_N(t))$ the set $J_t = g_t^{-1} \cdot I_t = \big(g_t(x_1(t)), \ldots, g_t(x_N(t))\big)$. Once again, the induced distance on $\mathcal{I}$ is given by $d(I, I') = \inf\big\{\sqrt{E[v,q]} \,:\, v_t \in T_e\mathcal{G},\, q(0) = I,\, q(1) = I'\big\}$. Incidentally, note that the above energy coincides with (2.3) and the induced distance $d(I, I')$ coincides with (2.4); that is, this was precisely our starting point at the beginning of this chapter.

CHAPTER 3

Momenta and Conservation Laws

In this chapter we will explore the structure of Hamilton's equations for the Riemannian manifold of landmarks and discuss the corresponding conserved quantities.

1. Hamilton's equations

1.1. Generalities. It is well known from the theories of classical mechanics and variational calculus [**2, 4, 37**] that for a system with n degrees of freedom the n second order Euler-Lagrange equations for an energy of the type

$$F[q] = \int_0^1 \mathcal{L}(q, \dot{q})\ dt,$$

namely

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0 \qquad i = 1, \ldots, n, \tag{3.1}$$

are equivalent to $2n$ first order differential equations, known as Hamilton's equations, where the variables are positions $q = (q^1, \ldots, q^n) \in \mathbb{R}^n$ and momenta $p = (p_1, \ldots, p_n) \in \mathbb{R}^n$. The momenta are in fact defined as the variables

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}^i}(q, \dot{q}), \qquad i = 1, \ldots, n;$$

if the determinant of matrix $\left[\frac{\partial \mathcal{L}}{\partial \dot{q}^i \partial \dot{q}^j}\right]_{i,j=1}^n$ is nonzero then the above system of equations is (locally) invertible and we can write $\dot{q}_i = \dot{q}_i(p, q)$, $i = 1, \ldots, n$, or just $\dot{q} = \dot{q}(p, q)$ in a more compact notation. The Hamiltonian is the scalar function

$$\mathcal{H}(p, q) \triangleq \Big[\sum_{i=1}^n p_i\, \dot{q}_i - \mathcal{L}(q, \dot{q})\Big]_{\dot{q} = \dot{q}(p,q)};$$

as anticipated above, the Euler-Lagrange equations (3.1) turn out to be are equivalent to Hamilton's Equations:

$$\begin{aligned}\dot{q}^i &= \frac{\partial \mathcal{H}}{\partial p_i}(p,q),\\ \dot{p}_i &= -\frac{\partial \mathcal{H}}{\partial q^i}(p,q).\end{aligned}$$

The flow determined by Hamilton's equations is known as the *cogeodesic flow* [**23**]. The Hamiltonian function is always an integral of motion, i.e. it is conserved along the cogeodesic flow, since

$$\frac{d}{dt}\mathcal{H}(p,q) = \sum_{i=1}^{n}\Big\{\frac{\partial \mathcal{H}}{\partial p_i}\dot{p}_i + \frac{\partial \mathcal{H}}{\partial q^i}\dot{q}^i\Big\} = \sum_{i=1}^{n}\Big\{-\frac{\partial \mathcal{H}}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q^i} + \frac{\partial \mathcal{H}}{\partial q^i}\frac{\partial \mathcal{H}}{\partial p_i}\Big\} \equiv 0.$$

This approach is very convenient when the variables live on a Riemannian manifold $\mathcal{M}$ with metric tensor $g(\cdot)$, since in this case the Lagrangian and Hamiltonian functions have a particularly simple form. In fact the Lagrangian is given by:

$$\mathcal{L}(q,\dot{q}) \;=\; \frac{1}{2}\,\dot{q}^T g(q)\dot{q} \;=\; \frac{1}{2}\sum_{i,j=1}^{N} g_{ij}(q)\,\dot{q}_i\,\dot{q}_j;$$

momenta are scalar variables

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}^i} = \sum_{i=1}^{n} g_{ij}(q)\,\dot{q}^j, \quad i = 1,\dots,n, \qquad \text{i.e.} \quad p = g(q)\dot{q} \quad \text{in vector form,} \tag{3.2}$$

and the Hamiltonian function is

$$\mathcal{H}(p,q) \;=\; \frac{1}{2}\,p^T g(q)^{-1} p \;=\; \frac{1}{2}\sum_{i,j=1}^{N} g^{ij}(q)\,p_i\,p_j \tag{3.3}$$

where we have indicated with g^{ij}, $i,j = 1,\dots,n$ the elements of the *inverse* of the metric tensor, $g(q)^{-1}$, also known as the *cometric tensor*; in other words, it is such that $\sum_{j=1}^{n} g^{ij}(q)\,g_{jk}(q) = \delta^i_k$ (Krnonecker's symbol). Hamilton's equations are

$$\dot{q}^i \;=\; \frac{\partial \mathcal{H}}{\partial p_i} \;=\; \sum_{j=1}^{n} g^{ij}(q)\,p_j \tag{3.4}$$

$$\dot{p}_i \;=\; -\frac{\partial \mathcal{H}}{\partial q^i} \;=\; \frac{1}{2}\sum_{j,k=1}^{n} \frac{\partial g^{jk}}{\partial q^i}(q)\,p_j\,p_k \tag{3.5}$$

for $i = 1, \ldots, n$. Note that equation (3.5) is in fact equivalent to definition (3.2). The *geodesic flow* on the tangent bundle $T\mathcal{M}$ is obtained from the cogeodesic flow by the first of Hamilton's equations. Being the Hamiltonian an integral of motion the cogeodesic flow maps the set $E_c \triangleq \{(q,p) \in T^*\mathcal{M} : H(p,q) = c\}$ onto itself for $c \geq 0$, so that the cotangent bundle $T^*\mathcal{M}$ is partitioned into the level sets E_c, $c \geq 0$. In the following we will compute Hamilton's equations in the case of the Riemannian manifold of landmarks and explore the corresponding conservation laws.

1.2. The landmarks manifold case. The metric tensor (2.11) of the Riemannian manifold of landmarks is, in fact, the *inverse* of a matrix (whose non-zero elements are given by the Green's function of a differential operator) which makes the Hamiltonian approach especially convenient since the Hamiltonian function itself (3.3) is expressed in terms of the cometric tensor. A slight complication comes from to the fact that the variables $q^{i,k}$ introduced in Chapter 2 have in fact *two* indices, $i = 1, \ldots, N$ and $k = 1, \ldots, D$, the former indicating the landmark label while the latter refers to the dimensional component of the single landmark; the dimension of the Riemannian manifold $\mathcal{I}$ is $n = DN$.

Equation (2.10) of Chapter 2 can be rewritten as

$$\dot{q}^{(k)}(t) = \left(S\big(q(t)\big) + \frac{\mathrm{id}}{\lambda}\right) p_{(k)}(t), \qquad k = 1, \ldots, D$$

so that (omitting the time argument) the single components are

$$\dot{q}^{i,k} = \sum_{j=1}^{N} \left(G(q^i, q^j) + \frac{\delta^{ij}}{\lambda}\right) p_{j,k}\,, \qquad i = 1, \ldots, N, \quad k = 1, \ldots, D; \tag{3.6}$$

they can also be expressed in the convenient compact form

$$\dot{q}^i \;=\; \sum_{j=1}^{N} \left(G(q^i, q^j) + \frac{\delta^{ij}}{\lambda}\right) p_j\,, \qquad i = 1, \ldots, N, \tag{3.7}$$

where both the left-hand and the right-hand sides are $1 \times D$ *row* vectors. Before proceeding to the derivation of the second set of equations (3.5) we will introduce some useful notation.

Given the block-diagonal nature of the $DN \times DN$ metric tensor g, defined in (2.11), it is convenient to write its generic element as g_{iajb}, with $i, j = 1, \dots, N$ and $a, b = 1, \dots, D$, where:

$i =$ row index *within* a $N \times N$ block; $a =$ index of row block;

$j =$ column index *within* a $N \times N$ block; $b =$ index of column block.

For example, with the above convention in the simple case $N = 3$, $D = 2$ the elements of a tensor q would be ordered as follows:

$$g = \left[\begin{array}{ccc|ccc} g_{1111} & g_{1121} & g_{1131} & g_{1112} & g_{1122} & g_{1132} \\ g_{2111} & g_{2121} & g_{2131} & g_{2112} & g_{2122} & g_{2132} \\ g_{3111} & g_{3121} & g_{3131} & g_{3112} & g_{3122} & g_{3132} \\ \hline g_{1211} & g_{1221} & g_{1231} & g_{1212} & g_{1222} & g_{1232} \\ g_{2211} & g_{2221} & g_{2231} & g_{2212} & g_{2222} & g_{2232} \\ g_{3211} & g_{3221} & g_{3231} & g_{3212} & g_{3222} & g_{3232} \end{array}\right]. \tag{3.8}$$

In general, by Theorem 2.7 the metric tensor for the landmarks manifold is

$$g(q) = \mathrm{diag}\{\underbrace{R(q), \dots, R(q)}_{D \text{ times}}\}$$

where $R(q)$ is the $N \times N$ matrix $\left(S(q) + \frac{\mathrm{id}}{\lambda}\right)^{-1}$ (so that in the above example only the two 3×3 diagonal blocks of matrix (3.8) would be non-zero). Therefore, if we indicate the generic element of $R(q)$ with with $R_{ij}(q)$, $i, j = 1, \dots N$, we have that the elements of metric tensor g can be expressed as

$$g_{iajb}(q) = R_{ij}(q)\, \delta_{ab}\,, \qquad i, j = 1, \dots, N, \quad a, b = 1, \dots, D$$

where δ_{ab} is Kronecker's delta. We can employ analogous notational conventions for the *inverse* of the metric tensor, namely the elements of $g(q)^{-1}$ can be written as

$$g^{iajb}(q) = R^{ij}(q)\, \delta^{ab}\,, \qquad i, j = 1, \dots, N, \quad a, b = 1, \dots, D,$$

where $R^{ij}(q)$ is the generic element of the *inverse* of $R(q)$. Since $R(q)^{-1} = S(q) + \frac{\mathrm{id}}{\lambda}$ we have that $R^{ij}(q) = G(q^i, q^j) + \frac{\delta^{ij}}{\lambda}$, so that the cometric tensor becomes:

$$g^{iajb}(q) = \Big(G(q^i, q^j) + \frac{\delta^{ij}}{\lambda} \Big)\, \delta^{ab}\,, \qquad i,j = 1,\dots,N, \quad a,b = 1,\dots,D.$$

With this notation we may rewrite the first set of differential equations (3.6) as

$$\dot{q}^{i,a} = \sum_{j=1}^{N} \sum_{b=1}^{D} g^{iajb}(q)\, p_{j,b}\,, \qquad i = 1,\dots,N, \quad a = 1,\dots,D,$$

which are formally consistent with (3.4), so that (3.6) are precisely the first set of Hamilton's equations, with $n = DN$.

The notation introduced above will be especially useful in the computation of curvature for landmarks manifolds, which will be done in Chapter 5. In any case we have that the Hamiltonian function of the system can be expressed as:

$$\begin{aligned} \mathcal{H}(p,q) \;&=\; \frac{1}{2}\, p^T g(q)^{-1} p \;=\; \frac{1}{2} \sum_{i,j=1}^{N} \sum_{a,b=1}^{D} g^{iajb}(q)\, p_{i,a}\, p_{j,b} \\ &=\; \frac{1}{2} \sum_{i,j=1}^{N} \sum_{a,b=1}^{D} \Big(G(q^i, q^j) + \frac{\mathrm{id}}{\lambda} \Big)\, \delta^{ab}\, p_{i,a}\, p_{j,b} \;=\; \frac{1}{2} \sum_{i,j=1}^{N} \sum_{a=1}^{D} \Big(G(q^i, q^j) + \frac{\delta^{ij}}{\lambda} \Big)\, p_{i,a}\, p_{j,a}\,, \end{aligned}$$

that is:

$$\boxed{\mathcal{H}(p,q) \;=\; \frac{1}{2} \sum_{i,j=1}^{N} \Big(G(q^i, q^j) + \frac{\delta^{ij}}{\lambda} \Big)\, \langle p_i, p_j \rangle_{\mathbb{R}^D}\,.} \tag{3.9}$$

Note that $G : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R} : (\xi, \eta) \mapsto G(\xi, \eta)$, the chosen Green's function (i.e. the kernel of space V), is a function of $2D$ real arguments:

$$G(\xi, \eta) = G\big(\xi^1, \xi^2, \dots, \xi^D, \eta^1, \eta^2, \dots, \eta^D\big);$$

from now on we shall indicate with $\frac{\partial G}{\partial \xi^\ell}$ its derivative with respect to the ℓ^{th} component of vector $\xi \in \mathbb{R}^D$. Also, we will indicate with $\nabla_\xi G : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D$ the *row* vector:

$$\nabla_\xi G(\xi, \eta) = \left[\;\frac{\partial G}{\partial \xi^1}(\xi, \eta) \quad \frac{\partial G}{\partial \xi^2}(\xi, \eta) \quad \cdots \quad \frac{\partial G}{\partial \xi^D}(\xi, \eta)\; \right].$$

As we said in Chapter 2, Green's function G is of the form $G(\xi,\eta)=\gamma\big(||\xi-\eta||_{\mathbb{R}^D}\big)$, for some $\gamma:[0,\infty)\to\mathbb{R}$; if this is the case, then:

$$\frac{\partial G}{\partial \xi^\ell}(\xi,\eta)=\gamma'\big(||\xi-\eta||_{\mathbb{R}^D}\big)\,\frac{\xi^\ell-\eta^\ell}{||\xi-\eta||_{\mathbb{R}^D}}\,,\qquad \ell=1,\ldots,D. \tag{3.10}$$

We now have all the machinery to state and prove the following result.

PROPOSITION 3.1. *The DN first order ordinary differential equations hold:*

$$\dot{p}_{i,k}=-\sum_{j=1}^{N}\frac{\partial G}{\partial \xi^k}(q^i,q^j)\,\langle p_i,p_j\rangle_{\mathbb{R}^D}\,,\qquad i=1,\ldots,N,\quad k=1,\ldots,D. \tag{3.11}$$

Therefore Hamilton's equations for the Riemannian manifold of landmarks are

$$\left\{\begin{array}{rcl}\dot{q}^i &=& \displaystyle\sum_{j=1}^{N}\left(G(q^i,q^j)+\frac{\delta^{ij}}{\lambda}\right)p_j\\ \dot{p}_i &=& \displaystyle-\sum_{j=1}^{N}\nabla_\xi G(q^i,q^j)\,\langle p_i,p_j\rangle_{\mathbb{R}^D}\end{array}\right.\qquad i=1,\ldots,N. \tag{3.12}$$

PROOF. The first of equations (3.12) is simply given by (3.7), which derives directly from the definition of momenta. The second of Hamilton's equations is:

$$\begin{aligned}\dot{p}_{i,k} &= -\frac{\partial\mathcal{H}}{\partial q^{i,k}}(p,q)=-\frac{1}{2}\sum_{j,\ell=1}^{N}\frac{\partial}{\partial q^{i,k}}\Big(G(q^j,q^\ell)+\frac{\mathrm{id}}{\lambda}\Big)\,\big\langle p_j\,p_\ell\big\rangle_{\mathbb{R}^D}\\ &= -\sum_{j=1}^{N}\frac{\partial G}{\partial\xi^k}(q^i,q^j)\,\big\langle p_i\,p_j\big\rangle_{\mathbb{R}^D}\end{aligned}$$

for $i=1,\ldots,N$ and $k=1,\ldots,D$, which coincides with (3.11). Such equations can be written in the compact form given by the second of equations (3.12). □

2. Conservation laws

As we mentioned above the flow determined by equations (3.12) is called the cogeodesic flow: the geodesic flow on the manifold $\mathcal{I}$ is determined by the solutions $q^i(\cdot)$, $i=1,\ldots,N$ of the above equations. The Hamiltonian function is constant along the solutions of (3.12), so that the cogeodesic flow partitions the cotangent bundle into the level sets of $\mathcal{H}(p,q)$.

The structure of the system of equations (3.12) is such that, in fact, other quantities are conserved along the flow. Before describing such integrals of motion we will quickly introduce some more notation. If $\big(p_i(\cdot), q^i(\cdot)\big)$, $i = 1, \ldots, N$ are solutions of (3.12), let $\varphi_t : \mathbb{R}^D \to \mathbb{R}^D$ be the time-dependent diffeomorphism $\varphi_t(\xi) \triangleq \varphi^v_{0t}(\xi)$ induced by velocity field $v_t(\xi) = \sum_{i=1}^N p_i(t)\, G\big(\xi, q^i(t)\big)$, $t \in [0,1]$, $\xi \in \mathbb{R}^D$. We will denote the components of such diffeomorphism and the velocity field as $\varphi_t(\xi) = \big(\varphi_t^1(\xi), \ldots, \varphi_t^D(\xi)\big)$ and $v_t(\xi) = \big(v_t^1(\xi), \ldots, v_t^D(\xi)\big)$, respectively. Let $D\varphi_t$ be the Jacobian matrix of the diffeomorphism, i.e.

$$D\varphi_t(\xi) \triangleq \begin{bmatrix} \frac{\partial \varphi_t^1}{\partial \xi^1} & \frac{\partial \varphi_t^1}{\partial \xi^2} & \cdots & \frac{\partial \varphi_t^1}{\partial \xi^D} \\ \frac{\partial \varphi_t^2}{\partial \xi^1} & \frac{\partial \varphi_t^2}{\partial \xi^2} & \cdots & \frac{\partial \varphi_t^2}{\partial \xi^D} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi_t^D}{\partial \xi^1} & \frac{\partial \varphi_t^D}{\partial \xi^2} & \cdots & \frac{\partial \varphi_t^D}{\partial \xi^D} \end{bmatrix}. \tag{3.13}$$

Note that $D\varphi_0(\xi) = \mathrm{id}$ (the $D \times D$ identity matrix) for all $\xi \in \mathbb{R}^D$, since $\varphi_0(\xi) = \xi$. The following important result holds.

Proposition 3.2 (Strong Conservation Law). *Assume $\lambda = \infty$ (exact matching problem). Then it is the case that*

$$\sum_{k=1}^D p_{i,k}(t)\, \frac{\partial \varphi_t^k}{\partial \xi^\ell}\big(q_i(0)\big) = p_{i,\ell}(0), \qquad i = 1, \ldots, N, \quad \ell = 1, \ldots, D \tag{3.14}$$

for all $t \in [0,1]$, which may be written in terms of (row) vector and matrix multiplication, as[1]

$$p_i(t) \cdot D\varphi_t\big(q_i(0)\big) = p_i(0), \qquad i = 1, \ldots, N \tag{3.15}$$

for all $t \in [0,1]$.

[1]In differential-geometric, coordinate-free notation the conservation law (3.15) may be written as $\big[\varphi_t\big(q_i(0)\big)\big]^* p_i(t) = p_i(0)$, where the upper star denotes the *pullback* map [**28**] applied to cotangent vector $p_i(t) \in T^*_{q^i(t)}\mathcal{I}$.

Proof. Since $D\varphi_0 = \mathrm{id}$ the proposition holds for $t=0$. We claim that the time derivative of the left-hand side of (3.14) is zero. In fact,

(3.16)

$$\frac{d}{dt}\Big\{ \sum_{k=1}^{D} p_{i,k}(t)\,\frac{\partial \varphi_t^k}{\partial \xi^\ell}\big(q^i(0)\big)\Big\} = \sum_{k=1}^{D} \Big\{ \dot{p}_{i,k}(t)\,\frac{\partial \varphi_t^k}{\partial \xi^\ell}\big(q^i(0)\big) + p_{i,k}(t)\,\frac{\partial}{\partial \xi^\ell}\frac{\partial \varphi_t^k}{\partial t}\big(q^i(0)\big)\Big\};$$

we now want to compute the last term on the right-hand side of the above expression. Since $\varphi_t(\cdot)$ is the diffeomorphism induced by the velocity field $v_t(\cdot)$ we have that $\frac{\partial \varphi_t^k}{\partial t}(\xi) = v_t^k\big(\varphi_t(\xi)\big)$, $k=1,\dots,D$, for all $\xi \in \mathbb{R}^D$. Whence by the chain rule

$$\frac{\partial}{\partial \xi^\ell}\frac{\partial \varphi_t^k}{\partial t}(\xi) = \frac{\partial}{\partial \xi^\ell}\, v_t^k\big(\varphi_t(\xi)\big) = \sum_{m=1}^{D} \frac{\partial v_t^k}{\partial \xi^m}\big(\varphi_t(\xi)\big)\,\frac{\partial \varphi_t^m}{\partial \xi^\ell}(\xi), \qquad \xi \in \mathbb{R}^D.$$

since $v_t^k(x) = \sum_{j=1}^N p_{j,k}(t)\, G\big(x, q^j(t)\big)$, $x \in \mathbb{R}^D$, the above expression becomes

$$\frac{\partial}{\partial \xi^\ell}\frac{\partial \varphi_t^k}{\partial t}(\xi) = \sum_{m=1}^{D}\sum_{j=1}^{N} p_{j,k}(t)\,\frac{\partial G}{\partial \xi^m}\big(\varphi_t(\xi), q^j(t)\big)\,\frac{\partial \varphi_t^m}{\partial \xi^\ell}(\xi), \qquad \xi \in \mathbb{R}^D.$$

In the case of exact matching $\varphi_t\big(q^i(0)\big) = q^i(t)$, so for $\xi = q^i(0)$ we get:

$$\frac{\partial}{\partial \xi^\ell}\frac{\partial \varphi_t^k}{\partial t}\big(q^i(0)\big) = \sum_{m=1}^{D}\sum_{j=1}^{N} p_{j,k}(t)\,\frac{\partial G}{\partial \xi^m}\big(q^i(t), q^j(t)\big)\,\frac{\partial \varphi_t^m}{\partial \xi^\ell}\big(q^i(0)\big). \tag{3.17}$$

On the other hand, by Proposition 3.1,

$$\dot{p}_{i,k}(t) = -\sum_{m=1}^{D}\sum_{j=1}^{N} \frac{\partial G}{\partial \xi^k}\big(q^i(t), q^j(t)\big)\, p_{i,m}(t)\, p_{j,m}(t), \tag{3.18}$$

for $i=1,\dots,N$ and $k=1,\dots,D$. Inserting equations (3.17) and (3.18) into expression (3.16) finally yields:

$$\begin{aligned}
&\frac{d}{dt}\Big\{ \sum_{k=1}^{D} p_{i,k}(t)\,\frac{\partial \varphi_t^k}{\partial \xi^\ell}\big(q^i(0)\big)\Big\} \\
&= \sum_{k=1}^{D}\Big\{ -\sum_{m=1}^{D}\sum_{j=1}^{N} \frac{\partial G}{\partial \xi^k}\big(q^i(t), q^j(t)\big)\, p_{i,m}(t)\, p_{j,m}(t)\,\frac{\partial \varphi_t^k}{\partial \xi^\ell}\big(q^i(0)\big) \\
&\quad + p_{i,k}(t)\sum_{m=1}^{D}\sum_{j=1}^{N} p_{j,k}(t)\,\frac{\partial G}{\partial \xi^m}\big(q^i(t), q^j(t)\big)\,\frac{\partial \varphi_t^m}{\partial \xi^\ell}\big(q^i(0)\big)\Big\} \;\equiv\; 0\,,
\end{aligned}$$

where we have implicitly relabeled some summation indices. □

Since φ_t is a diffeomorphism for all time $t \in [0,1]$ we have that its Jacobian matrix $D\varphi_t$ is *invertible* for all t, which immediately implies the following result.

Corollary 3.3. *Assume $\lambda = \infty$. If $p_i(0) = 0$ for some index (landmark) $i \in \{1, \ldots, N\}$, then the corresponding momentum is such that $p_i(t) \equiv 0$ for all $t \in [0,1]$.*

We should also note the strong conservation law for momenta and the system of differential equations (3.12) are, in fact, equivalent; more precisely, the following proposition holds.

Proposition 3.4. *Assume $\lambda = \infty$. If differential equations* (3.7)*, which are equivalent to the definition of momenta, and conservation laws* (3.15) *hold for all $t \in [0,1]$, then so do differential equations* (3.11)*.*

Proof. Differentiating equations (3.14) with respect to time yields:

$$\begin{aligned} 0 &= \frac{d}{dt}\left\{ \sum_{k=1}^{D} p_{i,k}(t)\, \frac{\partial \varphi_t^k}{\partial \xi^\ell}\big(q^i(0)\big) \right\} \\ &= \sum_{k=1}^{D} \left\{ \dot{p}_{i,k}(t)\, \frac{\partial \varphi_t^k}{\partial \xi^\ell}\big(q^i(0)\big) + p_{i,k}(t)\, \frac{\partial}{\partial \xi^\ell}\frac{\partial \varphi_t^k}{\partial t}\big(q^i(0)\big) \right\}, \qquad \begin{array}{l} i = 1, \ldots, N, \\ \ell = 1, \ldots, D. \end{array} \end{aligned}$$

Inserting equation (3.17) from the proof of Proposition 3.2 into the right-hand side of the above expression yields the equations:

$$\begin{aligned} \sum_{k=1}^{D} \dot{p}_{i,k}(t)\, \frac{\partial \varphi_t^k}{\partial \xi^\ell}\big(q^i(0)\big) &= -\sum_{k=1}^{D} p_{i,k}(t) \sum_{m=1}^{D}\sum_{j=1}^{N} p_{j,k}(t)\, \frac{\partial G}{\partial \xi^m}\big(q^i(t), q^j(t)\big)\, \frac{\partial \varphi_t^m}{\partial \xi^\ell}\big(q^i(0)\big) \\ &= -\sum_{m=1}^{D}\sum_{j=1}^{N} \big\langle p_i(t), p_j(t)\big\rangle_{\mathbb{R}^D}\, \frac{\partial G}{\partial \xi^m}\big(q^i(t), q^j(t)\big)\, \frac{\partial \varphi_t^m}{\partial \xi^\ell}\big(q^i(0)\big) \end{aligned}$$

for $i = 1, \ldots, N$ and $\ell = 1, \ldots, D$. Indicating with $\frac{\partial \varphi_t}{\partial \xi^\ell}$ the ℓ-th *column* of Jacobian matrix (3.13) the above equations may be written in terms of vector multiplications:

$$\dot{p}_i(t) \cdot \frac{\partial \varphi_t}{\partial \xi^\ell}\big(q^i(0)\big) = -\sum_{j=1}^{N} \big\langle p_i(t), p_j(t)\big\rangle_{\mathbb{R}^D}\, \nabla_\xi G\big(q^i(t), q^j(t)\big) \cdot \frac{\partial \varphi_t}{\partial \xi^\ell}\big(q^i(0)\big),$$

which again hold for $i = 1, \ldots, N$ and $\ell = 1, \ldots, D$. Whence:

$$\dot{p}_i(t) \cdot D\varphi_t\big(q^i(0)\big) = -\sum_{j=1}^{N} \big\langle p_i(t), p_j(t) \big\rangle_{\mathbb{R}^D} \nabla_\xi G\big(q^i(t), q^j(t)\big) \cdot D\varphi_t\big(q^i(0)\big),$$

for $i = 1, \ldots, N$; such equations imply the second set of (3.12), and whence (3.11), by the invertibility of the Jacobian matrix $D\varphi_t\big(q^i(0)\big)$. $\square$

At this point one could employ Emmy Noether's Theorem [**2, 4, 37**] and use the symmetries of the metric tensor (specifically, translation- and rotation-invariance) to prove the conservation of linear momentum and angular momentum. We will prove such conservation laws directly, simply by manipulating differential equations (3.12) and the general form of the kernel G and its partial derivatives (3.10).

Proposition 3.5 (Conservation of Linear Momentum, or First Weak Conservation Law). *For any choice of smoothing parameter λ, the quantity*

$$P(p) \triangleq \sum_{i=1}^{N} p_i\,, \tag{3.19}$$

which is a D-dimensional vector, is conserved in time.

Proof. Summing the set of equations (3.11) over index i yields:

$$\begin{aligned}\sum_{i=1}^{N} \dot{p}_{i,k} &= -\sum_{i,j=1}^{N} \frac{\partial G}{\partial \xi^k}(q^i, q^j)\, \langle p_i, p_j \rangle_{\mathbb{R}^D} \\ &= -\sum_{i=1}^{N} \frac{\partial G}{\partial \xi^k}(q^i, q^i)\, \|p_i\|^2_{\mathbb{R}^D} - \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{\partial G}{\partial \xi^k}(q^i, q^j)\, \langle p_i, p_j \rangle_{\mathbb{R}^D}.\end{aligned} \tag{3.20}$$

for any $k = 1, \ldots, D$. Since $\frac{\partial G}{\partial \xi^k}(x, x) = 0$ for all $x \in \mathbb{R}^D$ the first summation in the above expression is identically equal to zero. On the other hand, by (3.10) we have that $\frac{\partial G}{\partial \xi^k}(x, y) = -\frac{\partial G}{\partial \xi^k}(y, x)$ for all $x, y \in \mathbb{R}^D$ with $x \neq y$, therefore the second summation in (3.20) is also zero since all the terms are pairwise opposite. In conclusion, $\sum_{i=1}^{N} \dot{p}_{i,k} \equiv 0$, so that

$$\sum_{i=1}^{N} p_{i,k}(t) = \sum_{i=1}^{N} p_{i,k}(0), \qquad k = 1, \ldots, D, \quad t \in [0, 1].$$

Writing the above equalities in vector form proves the proposition. $\square$

Proposition 3.6 (Conservation of Angular Momentum, or Second Weak Conservation Law). *For any choice of smoothing parameter λ, if $D \geq 2$ the scalar quantities*

$$L_{ab}(p,q) \triangleq \sum_{i=1}^{N} \big(q^{i,a}p_{i,b} - q^{i,b}p_{i,a}\big) = \big\langle q^{(a)}, p_{(b)}\big\rangle_{\mathbb{R}^N} - \big\langle q^{(b)}, p_{(a)}\big\rangle_{\mathbb{R}^N}, \tag{3.21}$$

defined for $a, b = 1, \ldots, D$, with $a < b$, are conserved in time.

Proof. Let $R^{ij}(q) = G(q^i, q^j) + \frac{\delta^{ij}}{\lambda}$. Differentiating (3.21) and using (3.12) yields:

$$\begin{aligned}
\frac{d}{dt}L_{ab} &= \sum_{i=1}^{N} \big\{\dot{q}^{i,a}p_{i,b} - \dot{q}^{i,b}p_{i,a} + q^{i,a}\dot{p}_{i,b} - q^{i,b}\dot{p}_{i,a}\big\} \\
&= \sum_{i=1}^{N} \Big\{ \sum_{j=1}^{N} R^{ij}\big(p_{j,a}p_{i,b} - p_{j,b}p_{i,a}\big) + q^{i,a}\dot{p}_{i,b} - q^{i,b}\dot{p}_{i,a} \Big\} = \sum_{i=1}^{N} \big\{q^{i,a}\dot{p}_{i,b} - q^{i,b}\dot{p}_{i,a}\big\},
\end{aligned}$$

where we have used the symmetry of R^{ij}. Whence

$$\begin{aligned}
\frac{d}{dt}L_{ab} &= \sum_{i=1}^{N} \Big\{ q^{i,a} \sum_{j=1}^{N} \frac{\partial G}{\partial \xi^b}(q^i, q^j)\langle p_i, p_j\rangle_{\mathbb{R}^D} - q^{i,b} \sum_{j=1}^{N} \frac{\partial G}{\partial \xi^a}(q^i, q^j)\langle p_i, p_j\rangle_{\mathbb{R}^D} \Big\} \\
&= \sum_{i,j=1}^{N} \langle p_i, p_j\rangle_{\mathbb{R}^D} \Big\{ q^{i,a}\frac{\partial G}{\partial \xi^b}(q^i, q^j) - q^{i,b}\frac{\partial G}{\partial \xi^a}(q^i, q^j) \Big\},
\end{aligned}$$

and by (3.10):

$$\begin{aligned}
\frac{d}{dt}L_{ab} &= \sum_{i,j=1}^{N} \langle p_i, p_j\rangle_{\mathbb{R}^D} \frac{\gamma'\big(\|q^i - q^j\|_{\mathbb{R}^D}\big)}{\|q^i - q^j\|_{\mathbb{R}^D}} \Big\{ q^{i,a}\big(q^{i,b} - q^{j,b}\big) - q^{i,b}\big(q^{i,a} - q^{j,a}\big) \Big\} \\
&= \sum_{i,j=1}^{N} \langle p_i, p_j\rangle_{\mathbb{R}^D} \frac{\gamma'\big(\|q^i - q^j\|_{\mathbb{R}^D}\big)}{\|q^i - q^j\|_{\mathbb{R}^D}} \Big\{ -q^{i,a}q^{j,b} + q^{i,b}q^{j,a} \Big\},
\end{aligned}$$

which is identically equal to zero by the symmetries of the first two factors. ☐

Note that the linear momentum and angular momentum conservation laws consist, respectively, of D and $\frac{D(D-1)}{2}$ *scalar* conservation laws; taking into the account the fact that the Hamiltonian (which is a scalar) is also conserved, the cogeodesic flow, i.e. the dynamics of (p,q), takes place on a space of dimension $2DN - 1 - D - \frac{D(D-1)}{2}$.

CHAPTER 4

Curvature in terms of the Cometric Tensor

In this chapter we compute a formula for the Riemannian curvature tensor and the sectional curvature for a generic n-dimensional Riemannian manifold $\mathcal{M}$ in terms of the elements of the cometric tensor (i.e. the inverse of metric tensor); also, we will show how sectional curvature can be written as a ratio of quadratic forms on the space of alternating 2-forms on $T_p\mathcal{M}$, which may also be expressed in terms of the cometric. An accessible text on alternating forms is [**21**]. Classic references for differential geometry are, for example, [**11**] and [**27**] (we shall use the notation introduced in the latter); a modern, more advanced text on the topic is [**23**].

1. Motivation

We saw in Chapter 2 that when shape is modeled as a labeled N-tuple of landmarks in D dimensions the corresponding metric tensor, when written as a matrix g, turns out to be the inverse of a positive definite matrix: that is, we may write $g(q) = \Big(\mathrm{diag}\Big\{S(q) + \frac{\mathrm{id}}{\lambda}, \dots, S(q) + \frac{\mathrm{id}}{\lambda}\Big\}\Big)^{-1}$, where the elements of $S(q)$ are computed by evaluating a given Green's function at different locations.

Under these circumstances calculating sectional curvature in the traditional way, i.e. by computing Christoffel symbols and their partial derivatives, turns out to be a formidable task since it involves computing successive derivatives of the inverse of a tensor. Therefore it would be convenient to have access to a formula expressing sectional curvature (and, more in general, the Riemannian curvature tensor) in terms of the derivatives of the *inverse* of g, that are more easily computed.

The spirit of the current chapter is precisely to express geometric quantities for a generic Riemannian manifold $\mathcal{M}$ of dimension n in as functions of the cometic tensor. In particular, we will solve the highly non-trivial problem to provide a formula for the Riemannian curvature tensor and the numerator of sectional curvature (see next section) in terms of g^{ij}, $\frac{\partial}{\partial x^k} g^{ij}$, and $\frac{\partial}{\partial x^\ell \partial x^k} g^{ij}$, with $i, j, k, \ell = 1, \dots, n$. In the last part of this chapter we will express sectional curvature as the ratio of quadratic forms on the space of alternating 2-tensors $\Lambda^2(T_p\mathrm{M})$, which will allow usa to formulate the problem of finding bounds for sectional curvature as a generalized eigenvalue problem. In the next chapter we will apply these formulas to the metric tensor of the landmarks manifolds. For the sake of notational compactness, from now on we shall use the simple symbol ∂_i in lieu of $\frac{\partial}{\partial x^i}$, $i = 1, \dots, n$. Moreover we will employ Einstein's summation convention: that is, an index occurring twice in a product is to be summed from 1 to n; for example, $X^i \partial_i$ is an abbreviation for $\sum_{i=1}^n X^i \partial_i$.

2. Generalities on the Riemannian Curvature Tensor

Suppose that $\mathcal{M}$ is an n-dimensional Riemannian manifold with metric tensor g. If we consider a local chart (U, φ) on the manifold with coordinates $(x^1, \dots, x^n)$ the metric is represented by a positive definite, symmetric matrix

$$\big[g_{ij}(x)\big]_{i,j=1,\dots,n}$$

where the coefficients depend smoothly on $x \in \varphi(U) \subseteq \mathbb{R}^n$. The product of two tangent vectors $X, Y \in T_p\mathcal{M}$, with $X = X^i \partial_i$ and $Y = Y^i \partial_i$, is

$$\langle X, Y \rangle_p = g_{ij}\big(x(p)\big) X^i Y^j \, ;$$

in particular, $g_{ij}(x(p)) = \langle \partial_i, \partial_j \rangle_p$.

Notation. We shall denote the partial derivatives of the elements of tensor g as follows: $g_{ij,k} \triangleq \frac{\partial}{\partial x^k} g_{ij}$ and $g_{ij,k\ell} \triangleq \frac{\partial^2}{\partial x^\ell \partial x^k} g_{ij}$, for $i, j, k, \ell = 1, \dots, n$.

Indicating with $\mathcal{T}(\mathcal{M})$ the space of smooth vector fields on the manifold $\mathcal{M}$, let $\nabla : \mathcal{T}(\mathcal{M}) \times \mathcal{T}(\mathcal{M}) \to \mathcal{T}(\mathcal{M})$ be the Levi-Civita connection of the Riemannian manifold. The Christoffel symbols are defined by $\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k$. It is well known that they have the form: $\Gamma_{ij}^k = \frac{1}{2} g^{k\ell}(g_{i\ell,j} + g_{j\ell,i} - g_{ij,\ell})$. The *Riemannian curvature* *endomorphism* is the map $R : \mathcal{T}(\mathcal{M}) \times \mathcal{T}(\mathcal{M}) \times \mathcal{T}(\mathcal{M}) \to \mathcal{T}(\mathcal{M})$ defined by

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{4.1}$$

In local coordinates $R(\partial_i, \partial_j)\partial_k = R^{\ell}_{ijk}\partial_\ell$, and

$$R_{ijkm} \triangleq \langle R(\partial_i, \partial_j)\partial_k, \partial_m\rangle = g_{m\ell} R^{\ell}_{ijk}.$$

The *Riemannian curvature tensor* acts on vector fields as follows:

$$R(X, Y, Z, W) = \langle R(X, Y)Z, W\rangle \tag{4.2}$$

and in coordinates it is written as $R = R_{ijkm} dx^i \otimes dx^j \otimes dx^k \otimes dx^m$. The Riemannian curvature tensor has a number of symmetries:

$$\begin{aligned} R_{ijk\ell} &= -R_{jik\ell}, & R_{ijk\ell} &= -R_{ij\ell k}, \\ R_{ijk\ell} &= R_{k\ell ij}, & R_{ijk\ell} + R_{jki\ell} + R_{kij\ell} &= 0, \end{aligned} \tag{4.3}$$

the last of which is known as the *first Bianchi identity*.

With the above conventions, the sectional curvature associated to a pair of non-parallel tangent vectors X and Y is given by:[1]

$$\begin{aligned} K(X, Y) &= \frac{R(X, Y, Y, X)}{\|X\|^2\|Y\|^2 - \langle X, Y\rangle^2} = \frac{\langle R(X, Y)Y, X\rangle}{\|X\|^2\|Y\|^2 - \langle X, Y\rangle^2} \\ &= \frac{R_{ijkm} X^i Y^j Y^k X^m}{\|X\|^2\|Y\|^2 - \langle X, Y\rangle^2}\,; \end{aligned} \tag{4.4}$$

[1]The notation described above is the one adopted by Lee [**27**]. Other authors use different sign conventions, however in a way that the definition of sectional curvature eventually agrees in sign with (4.4). For example, Jost [**23**] defines $R(X, Y)Z$ in the same way as above, but then defines the coefficients of the Riemannian curvature tensor as follows: $R_{ijkm} = \langle R(\partial_i, \partial_j)\partial_m, \partial_k\rangle$, i.e. with a sign that is opposite to Lee's convention; however, Jost eventually defines the numerator of sectional curvature as $R_{ijkm} X^i Y^j X^k Y^m$, so that it agrees in sign with (4.4). On the other hand Do Carmo [**11**] defines the Riemannian curvature endomorphism as follows: $R(X, Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z$, that is, with a sign that is opposite to (4.1); the coefficients are then defined formally as in (4.2), $R_{ijkm} = \langle R(\partial_i, \partial_j)\partial_k, \partial_m\rangle$ so that their sign is in fact opposite to our convention (they coincide with Jost's); but then the numerator of $K(X, Y)$ is set to be equal to $R_{ijkm} X^i Y^j X^k Y^m$, so Do Carmo's definition of sectional curvature is eventually consistent with (4.4).

note that the denominator is always positive by the Cauchy-Schwarz inequality.

There are different ways of expressing the Riemannian curvature tensor in terms of the metric tensor g; the following proposition turns out to be useful for our purposes since it provides an expression that does not require to compute derivatives of the Christoffel symbols. The proof that follows is an adaptation from one found in [**29**].

PROPOSITION 4.1. *The following expression holds:*

$$2R_{ijkm} = g_{ik,jm} + g_{jm,ik} - g_{jk,im} - g_{im,jk} + 2\Gamma^r_{ik}\Gamma^s_{jm}g_{rs} - 2\Gamma^r_{jk}\Gamma^s_{im}g_{rs}. \tag{4.5}$$

PROOF. From the definition of the Christoffel symbols, $\nabla_{\partial_i}\partial_j = \Gamma^k_{ij}\partial_k$,

$$\begin{aligned} 2\langle\nabla_{\partial_i}\partial_j, \partial_m\rangle &= 2\langle\Gamma^k_{ij}\partial_k, \partial_m\rangle = 2\Gamma^k_{ij}g_{mk} = g_{mk}g^{k\ell}(g_{i\ell,j} + g_{j\ell,i} - g_{ij,\ell}) \\ &= \delta^\ell_m(g_{i\ell,j} + g_{j\ell,i} - g_{ij,\ell}) = g_{im,j} + g_{jm,i} - g_{ij,m}\,; \end{aligned}$$

an appropriate rearrangement of the indices yields the following expression:

$$2\langle\nabla_{\partial_j}\partial_k, \partial_m\rangle = g_{jm,k} + g_{km,j} - g_{jk,m}\,. \tag{4.6}$$

By the metric compatibility of the connection we have that

$$\partial_i\langle\nabla_{\partial_j}\partial_k, \partial_m\rangle = \langle\nabla_{\partial_i}\nabla_{\partial_j}\partial_k, \partial_m\rangle + \langle\nabla_{\partial_j}\partial_k, \nabla_{\partial_i}\partial_m\rangle,$$

whence, by (4.6),

$$\begin{aligned} 2\langle\nabla_{\partial_i}\nabla_{\partial_j}\partial_k, \partial_m\rangle + 2\langle\nabla_{\partial_j}\partial_k, \nabla_{\partial_i}\partial_m\rangle &= 2\partial_i\langle\nabla_{\partial_j}\partial_k, \partial_m\rangle \\ &= \partial_i(g_{jm,k} + g_{km,j} - g_{jk,m}) = g_{jm,ki} + g_{km,ji} - g_{jk,mi}\,. \end{aligned} \tag{4.7}$$

By switching i and j we also have that

$$2\langle\nabla_{\partial_j}\nabla_{\partial_i}\partial_k, \partial_m\rangle + 2\langle\nabla_{\partial_i}\partial_k, \nabla_{\partial_j}\partial_m\rangle = g_{im,kj} + g_{km,ij} - g_{ik,mj}\,. \tag{4.8}$$

Combining (4.7) and (4.8) yields

$$\begin{aligned} &2\langle\nabla_{\partial_i}\nabla_{\partial_j}\partial_k, \partial_m\rangle - 2\langle\nabla_{\partial_j}\nabla_{\partial_i}\partial_k, \partial_m\rangle \\ &\quad = g_{jm,ki} - g_{jk,mi} - g_{im,kj} + g_{ik,mj} - 2\langle\nabla_{\partial_j}\partial_k, \nabla_{\partial_i}\partial_m\rangle + 2\langle\nabla_{\partial_i}\partial_k, \nabla_{\partial_j}\partial_m\rangle. \end{aligned}$$

But by definition $R(\partial_i, \partial_j)\partial_k = \nabla_{\partial_i}\nabla_{\partial_j}\partial_k - \nabla_{\partial_j}\nabla_{\partial_i}\partial_k$, whence

$$2R_{ijkm} = 2\langle R(\partial_i, \partial_j)\partial_k, \partial_m\rangle = 2\langle \nabla_{\partial_i}\nabla_{\partial_j}\partial_k, \partial_m\rangle - 2\langle \nabla_{\partial_j}\nabla_{\partial_i}\partial_k, \partial_m\rangle,$$

so we have proven that

$$2R_{ijkm} = g_{jm,ki} - g_{jk,mi} - g_{im,kj} + g_{ik,mj} - 2\langle \nabla_{\partial_j}\partial_k, \nabla_{\partial_i}\partial_m\rangle + 2\langle \nabla_{\partial_i}\partial_k, \nabla_{\partial_j}\partial_m\rangle. \tag{4.9}$$

By the definition of the Christoffels,

$$\begin{aligned}
\langle \nabla_{\partial_i}\partial_k, \nabla_{\partial_j}\partial_m\rangle &= \langle \Gamma^r_{ik}\partial_r, \Gamma^s_{jm}\partial_s\rangle &= \Gamma^r_{ik}\Gamma^s_{jm} g_{rs}\,, \\
\langle \nabla_{\partial_j}\partial_k, \nabla_{\partial_i}\partial_m\rangle &= \langle \Gamma^r_{jk}\partial_r, \Gamma^s_{im}\partial_s\rangle &= \Gamma^r_{jk}\Gamma^s_{im} g_{rs}\,.
\end{aligned}$$

Inserting the above expressions into the right-hand side of (4.9) finally yields (4.5). □

Notation. For any pair of tangent vectors $X, Y \in T_p\mathcal{M}$ we shall denote with $\Gamma(X, Y)$ the following vector in $T_p\mathcal{M}$:

$$\Gamma(X, Y) \triangleq \Gamma^k_{ij} X^i Y^j \partial_k\,,$$

where the Γ^k_{ij} are the Christoffel symbols for metric tensor g.

Given the above notation the numerator (and the sign) of sectional curvature $K(X, Y)$ may be computed using the following result.

Proposition 4.2. *The following expressions hold for any pair $X, Y \in T_p\mathcal{M}$:*

$$\begin{aligned}
2R(X, Y, Y, X) &= -(X^iY^j - Y^iX^j) g_{ik,jm} (X^kY^m - Y^kX^m) \\
&\quad + 2\|\Gamma(X, Y)\|^2 - 2\langle \Gamma(X, X), \Gamma(Y, Y)\rangle.
\end{aligned}$$

Proof. We have that

$$\begin{aligned}
&-(X^iY^j - Y^iX^j) g_{ik,jm} (X^kY^m - Y^kX^m) \\
&= -X^iX^kY^jY^m g_{ik,jm} + X^iY^kY^jX^m g_{ik,jm} + Y^iX^kX^jY^m g_{ik,jm} - Y^iY^kX^jX^m g_{ik,jm} \\
&= 2X^iY^kY^jX^m g_{ik,jm} - X^iX^kY^jY^m g_{ik,jm} - Y^iY^kX^jX^m g_{ik,jm} \\
&= X^iY^kY^jX^m g_{ik,jm} + X^mY^jY^kX^i g_{mj,ki} - X^iX^mY^jY^k g_{im,jk} - Y^jY^kX^iX^m g_{jk,im} \\
&= X^iY^jY^kX^m \big(g_{ik,jm} + g_{jm,ik} - g_{im,jk} - g_{jk,im}\big).
\end{aligned}$$

As far as the Christoffel symbols are concerned,

$$\begin{aligned} g_{rs}X^iY^jY^kX^m\Gamma^r_{ik}\Gamma^s_{jm} &= \langle X^iY^k\Gamma^r_{ik}\partial_r, Y^jX^m\Gamma^s_{jm}\partial_s\rangle \\ &= \langle \Gamma(X,Y), \Gamma(X,Y)\rangle \ = \ \|\Gamma(X,Y)\|^2, \end{aligned}$$

and

$$g_{rs}X^iY^jY^kX^m\Gamma^r_{jk}\Gamma^s_{im} \ = \ \langle X^iX^m\Gamma^s_{im}\partial_s, Y^jY^k\Gamma^r_{jk}\partial_r\rangle \ = \ \langle \Gamma(X,X), \Gamma(Y,Y)\rangle.$$

This completes the proof. □

3. The dual Riemannian Curvature Tensor

One of the purposes of the current chapter is to provide a formula for the numerator of sectional curvature (4.4) in terms the elements of the cometric tensor and their derivatives g^{ij}, $g^{ij,}{}_{k}$, and $g^{ij,}{}_{k\ell}$, $i, j, k, \ell = 1, \ldots, n$. The key idea is to define a "dual" curvature tensor by *raising the indices* of the Riemannian curvature tensor defined and described in the previous section. That is, if we define the coefficients of the dual Riemannian curvature tensor as $R^{ursv} \triangleq R_{ijkm}\, g^{iu}g^{jr}g^{ks}g^{mv}$ and for an arbitrary pair of tangent vectors $X = X^i\partial_i$ and $Y = Y^i\partial_i$ we consider the cotangent vectors $X^\flat = X_i dx^i$ and $Y^\flat = Y_i dx^i$, with $X_i = g_{ij}X^j$ and $Y_i = g_{ij}Y^j$ (the "flat" perator $\flat : T_p\mathcal{M} \to T_p^*\mathcal{M}$ lowers the indices of a tangent vector [**21, 28**]), then the numerator of $K(X,Y)$ may be rewritten as:

$$\begin{aligned} \langle R(X,Y)Y, X\rangle &= R_{ijkm}X^iY^jY^kX^m \ = \ R_{ijkm}g^{iu}g^{jr}g^{ks}g^{mv}X_uY_rY_sX_v \\ &= R^{ursv}X_uY_rY_sX_v. \end{aligned} \tag{4.10}$$

In this section we shall factorize the tensor coefficients R_{ijkm} in the following way:

$$R_{ijkm} = g_{iu}\, g_{jr}\, g_{ks}\, g_{mv}\, R^{ursv}, \tag{4.11}$$

and express R^{ursv} in terms of g^{ij} and its first and second partial derivatives. In the section that follows we will compute the action of the dual tensor on the cotangent vectors $(X^\flat, Y^\flat, Y^\flat, X^\flat)$, which will cause drastic simplifications and provide a surprisingly simple formula for the numerator of sectional curvature.

In Proposition 4.1 we proved the following general formula:

$$(4.12)\qquad 2R_{ijkm} = g_{ik,jm} + g_{jm,ik} - g_{jk,im} - g_{im,jk} + 2\Gamma^{\varphi}_{ik}\Gamma^{\psi}_{jm}g_{\varphi\psi} - 2\Gamma^{\varphi}_{jk}\Gamma^{\psi}_{im}g_{\varphi\psi}\,.$$

The key idea is to try to factorize *each* term of the above formula in the way that is suggested by (4.11). We will start with the terms that involve the first derivatives of the metric tensor, i.e. the Christoffel symbols on the right-hand side of (4.12).

3.1. First derivatives. If we write the metric tensor as a matrix we have that $g = Q^{-1}$, so that[2] $\partial_x g = -Q^{-1}\cdot\partial_x Q\cdot Q^{-1}$. In index notation the first partial derivative is given by

$$g_{jm,k} = -g_{jr}\,g^{rs,}_{\ \ \ k}\ g_{sm}\,.$$

By the above formula we may rewrite the Christoffel symbols in (4.12) in a way that is suitable for our purposes.

[2]This expression generalizes the fact that for a scalar differentiable function $f:\mathbb{R}\to\mathbb{R}$ of the form $f(x) = \frac{1}{g(x)}$ its derivative is given by $f'(x) = -\frac{g'(x)}{g^2(x)}$. In general, recall the definition of *differential*: if Ω is an open set in $\mathbb{R}^n$, ϕ maps Ω into $\mathbb{R}^m$, x is in Ω and there exists a linear transformation $A\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ such that $|\phi(x+h)-\phi(x)-Ah| = o(|h|)$, then we say that ϕ is differentiable at x and we write $d\phi(x) = A$ (see [**41**] for details). The definition can be easily extended to matrix-valued functions (and in fact to maps between Banach spaces). If a function f is defined as $f: GL_n(\mathbb{R})\to GL_n(\mathbb{R}) : A\mapsto A^{-1}$ (note that $GL_n(\mathbb{R})$ is an open subset of $\mathbb{R}^{n\times n}$) then its differential computed at a point $A\in GL_n(\mathbb{R})$ is the linear operator $df(A)\in\mathcal{L}\big(GL_n(\mathbb{R}),GL_n(\mathbb{R})\big)$ defined as $df(A)B = -A^{-1}BA^{-1}$, for all $B\in GL_n(\mathbb{R})$; in fact it can be proven that $\big\|f(A+H) - f(A) + A^{-1}HA^{-1}\big\|/\|H\|\to 0$ as $\|H\|\to 0$, for example by using the matrix inversion lemma [**25**]. If f is composed with a differentiable function $h:\mathbb{R}\to GL_n(\mathbb{R})$, i.e. if we define $\psi = f\circ h$, then the differential of ψ computed at a generic point $z\in\mathbb{R}$ is a linear operator $d\psi\in\mathcal{L}\big(\mathbb{R},GL_n(\mathbb{R})\big)$ (simply representable as a matrix) that can be calculated via the chain rule, i.e. $d\psi(z) = df(h(z))dh(z)$. In the case of $f(A) = A^{-1}$ this yields $d\psi(z) = -\big(h(z)\big)^{-1}h'(z)\big(h(z)\big)^{-1}$.

For example,

$$
\begin{aligned}
\Gamma^{\varphi}_{jk} &= \frac{1}{2} g^{\varphi\ell}\big(g_{j\ell,k} + g_{\ell k,j} - g_{jk,\ell}\big) \\
&= \frac{1}{2} g^{\varphi\ell}\big(- g_{jr}\, g^{rs,}{}_{k}\;\; g_{s\ell} - g_{\ell r}\, g^{rs,}{}_{j}\;\; g_{sk} + g_{jr}\, g^{rs,}{}_{\ell}\;\; g_{sk}\big) \\
&= \frac{1}{2}\big(g_{jr}\, g^{\varphi\ell}\, g^{rs,}{}_{\ell}\;\; g_{sk} - g_{jr}\, g^{rs,}{}_{k}\;\; \delta^{\varphi}_{s} - \delta^{\varphi}_{r}\, g^{rs,}{}_{j}\;\; g_{sk}\big) \\
&= \frac{1}{2}\big(g_{jr}\, g^{\varphi\ell}\, g^{rs,}{}_{\ell}\;\; g_{sk} - g_{jr}\, g^{r\varphi,}{}_{k}\;\; - g^{\varphi s,}{}_{j}\;\; g_{sk}\big).
\end{aligned}
$$

We can extract factor $g_{jr}\, g_{sk}$ from the first of the three terms above, but we need to manipulate the other two in order to be able to do the same with them. We have that $g^{r\varphi,}{}_{k} \;\; = g^{r\varphi,}{}_{\ell}\;\; \delta^{\ell}_{k} = g^{r\varphi,}{}_{\ell}\;\; g^{\ell s}\, g_{sk}$ and similarly $g^{\varphi s,}{}_{j}\;\; = g^{\varphi s,}{}_{\ell}\;\; \delta^{\ell}_{j} = g^{\varphi s,}{}_{\ell}\;\; g^{\ell r} g_{jr}$, whence the above may be expressed as

$$
\Gamma^{\varphi}_{jk} = \frac{1}{2} g_{jr}\big(g^{\varphi\ell}\, g^{rs,}{}_{\ell}\;\; - g^{s\ell}\, g^{r\varphi,}{}_{\ell}\;\; - g^{r\ell}\, g^{\varphi s,}{}_{\ell}\;\; \big) g_{sk}\,. \tag{4.13}
$$

As far as the remaining three Christoffel symbols in (4.12) are concerned, by appropriately relabeling the indices we obtain the following:

$$
\begin{aligned}
\Gamma^{\psi}_{im} &= \frac{1}{2} g_{iu}\big(g^{\psi\ell}\, g^{uv,}{}_{\ell}\;\; - g^{v\ell}\, g^{u\psi,}{}_{\ell}\;\; - g^{u\ell}\, g^{\psi v,}{}_{\ell}\;\; \big) g_{vm}\,, \\
\Gamma^{\varphi}_{ik} &= \frac{1}{2} g_{iu}\big(g^{\varphi\ell}\, g^{us,}{}_{\ell}\;\; - g^{s\ell}\, g^{u\varphi,}{}_{\ell}\;\; - g^{u\ell}\, g^{\varphi s,}{}_{\ell}\;\; \big) g_{sk}\,, \\
\Gamma^{\psi}_{jk} &= \frac{1}{2} g_{jr}\big(g^{\psi\ell}\, g^{rv,}{}_{\ell}\;\; - g^{v\ell}\, g^{r\psi,}{}_{\ell}\;\; - g^{r\ell}\, g^{\psi v,}{}_{\ell}\;\; \big) g_{vm}\,.
\end{aligned}
$$

It is convenient, at this point, to define "anti-Christoffel" symbols, or dual Christoffel symbols, in the following manner:[3]

$$
\widetilde{\Gamma}^{rs}_{u} \triangleq \frac{1}{2} g_{u\varphi}\big(g^{s\varphi,}{}_{\xi}\;\; g^{\xi r} + g^{r\varphi,}{}_{\xi}\;\; g^{\xi s} - g^{rs,}{}_{\xi}\;\; g^{\xi\varphi}\big). \tag{4.14}
$$

Notation. If we also define:

$$
g^{ij,k} \triangleq g^{ij,}{}_{\xi}\;\; g^{\xi k} \quad \text{and} \quad g^{ij,k\ell} \triangleq g^{ij,}{}_{\xi\eta}\;\; g^{\xi k} g^{\eta\ell}, \tag{4.15}
$$

[3]Given the form of expression (4.13), perhaps it would make more sense to add a minus sign in front of definition (4.14). In any case this is rather irrelevant for our purposes, since on the right-hand side of (4.12) we have the products of *pairs* of Christoffel symbols so that, in substituting the definition of the dual Christoffels in such products, opposite signs would cancel.

then the dual symbols assume an aspect that is formally analogous to the traditional definition of Christoffel symbols:

$$\widetilde{\Gamma}_{u}^{rs} = \frac{1}{2}\, g_{u\varphi}\big(g^{s\varphi,r} + g^{r\varphi,s} - g^{rs,\varphi}\big).$$

With the above notation and using the obvious fact that $g_{\varphi\psi} = g_{\varphi\rho}\, g^{\rho\sigma} g_{\sigma\psi}$ we may rewrite the last term of the right-hand side of equation (4.12) as follows:

$$\begin{aligned}
-2\,\Gamma_{jk}^{\varphi}\,\Gamma_{im}^{\psi}\, g_{\varphi\psi} = &- \frac{1}{2}\, g_{jr}\, g_{ks}\, g_{iu}\, g_{mv}\, \big(g^{rs,}{}_{\xi}\; g^{\xi\varphi} - g^{r\varphi,}{}_{\xi}\; g^{\xi s} - g^{s\varphi,}{}_{\xi}\; g^{\xi r}\big)\, g_{\varphi\rho} \\
&\cdot g^{\rho\sigma}\, g_{\sigma\psi}\, \big(g^{uv,}{}_{\xi}\; g^{\xi\psi} - g^{u\psi,}{}_{\xi}\; g^{\xi v} - g^{v\psi,}{}_{\xi}\; g^{\xi u}\big) \\
= &-\; g_{iu}\, g_{jr}\, g_{ks}\, g_{mv}\, 2\,\widetilde{\Gamma}_{\rho}^{rs}\,\widetilde{\Gamma}_{\sigma}^{uv}\, g^{\rho\sigma},
\end{aligned}$$

whereas the second-to-last of the same formula becomes

$$2\,\Gamma_{ik}^{\varphi}\,\Gamma_{jm}^{\psi}\, g_{\varphi\psi} = g_{iu}\, g_{jr}\, g_{ks}\, g_{mv}\, 2\,\widetilde{\Gamma}_{\rho}^{us}\,\widetilde{\Gamma}_{\sigma}^{rv}\, g^{\rho\sigma}.$$

Note that we have precisely achieved the desired factorization, as in (4.11).

3.2. Second derivatives. If we write the metric tensor as a matrix $g = Q^{-1}$ its second partial derivative is $\partial_y\partial_x g = Q^{-1}\cdot\big(\partial_y Q\cdot Q^{-1}\cdot\partial_x Q + \partial_x Q\cdot Q^{-1}\cdot\partial_y Q - \partial_y\partial_x Q\big)\cdot Q^{-1}$, or, in index notation,

$$g_{jm,ki} \;=\; g_{jr}\big(g^{r\lambda,}{}_{i}\; g_{\lambda\mu}\, g^{\mu v,}{}_{k} \;+ g^{r\lambda,}{}_{k}\; g_{\lambda\mu}\, g^{\mu v,}{}_{i} \;- g^{rv,}{}_{ki}\;\big) g_{vm}\,;$$

some manipulation is in order so to achieve the factorization that we desire.

We have that:

$$\begin{aligned}
g_{jm,ik} &= g_{jr} g_{mv}\, \delta_i^{\zeta}\delta_k^{\xi}\big(g^{r\lambda,}{}_{\zeta}\; g_{\lambda\mu}\, g^{\mu v,}{}_{\xi} \;+ g^{r\lambda,}{}_{\xi}\; g_{\lambda\mu}\, g^{\mu v,}{}_{\zeta} \;- g^{rv,}{}_{\zeta\xi}\;\big) \\
&= g_{jr} g_{mv}\, g_{iu} g_{ks}\, g^{u\zeta} g^{s\xi}\big(g^{r\lambda,}{}_{\zeta}\; g_{\lambda\mu}\, g^{\mu v,}{}_{\xi} \;+ g^{r\lambda,}{}_{\xi}\; g_{\lambda\mu}\, g^{\mu v,}{}_{\zeta} \;- g^{rv,}{}_{\zeta\xi}\;\big) \\
&= g_{iu} g_{jr} g_{ks} g_{mv}\, \big[\big(g^{r\lambda,}{}_{\zeta}\; g^{\zeta u}\big) g_{\lambda\mu}\, \big(g^{\mu v,}{}_{\xi}\, g^{\xi s}\big) + \big(g^{r\lambda,}{}_{\xi}\, g^{\xi s}\big) g_{\lambda\mu}\, \big(g^{\mu v,}{}_{\zeta}\, g^{\zeta u}\big) - g^{rv,}{}_{\zeta\xi}\, g^{\zeta u} g^{\xi s}\big] \\
&= g_{iu} g_{jr} g_{ks} g_{mv}\, \big(g^{r\lambda,u} g_{\lambda\mu}\, g^{\mu v,s} + g^{r\lambda,s} g_{\lambda\mu}\, g^{\mu v,u} - g^{rv,su}\big),
\end{aligned}$$

where we have used definitions (4.15). Note that this is exactly the factorization (4.11) we needed. So the first four terms on the right-hand side of (4.12) may be rewritten, respectively, as follows:

$$\begin{aligned}
g_{jm,ki} &= g_{iu}g_{jr}g_{ks}g_{mv}\left(g^{r\lambda,u}g_{\lambda\mu}\,g^{\mu v,s} + g^{r\lambda,s}g_{\lambda\mu}\,g^{\mu v,u} - g^{rv,su}\right),\\
g_{ik,jm} &= g_{iu}g_{jr}g_{ks}g_{mv}\left(g^{u\lambda,v}g_{\lambda\mu}\,g^{\mu s,r} + g^{u\lambda,r}g_{\lambda\mu}\,g^{\mu s,v} - g^{su,rv}\right),\\
-g_{jk,im} &= g_{iu}g_{jr}g_{ks}g_{mv}\left(-\,g^{r\lambda,u}g_{\lambda\mu}\,g^{\mu s,v} - g^{r\lambda,v}g_{\lambda\mu}\,g^{\mu s,u} + g^{rs,uv}\right),\\
-g_{im,jk} &= g_{iu}g_{jr}g_{ks}g_{mv}\left(-\,g^{u\lambda,r}g_{\lambda\mu}\,g^{\mu v,s} - g^{u\lambda,s}g_{\lambda\mu}\,g^{\mu v,r} + g^{uv,rs}\right),
\end{aligned}$$

where the last three are obtained just by appropriately rearranging the indices. Again, we have achieved exactly the factorization that we wanted.

3.3. Expressions for the dual Riemannian Curvature Tensor. Inserting the formulas we have computed in the above subsections into the right-hand side of (4.12) and comparing the result with (4.11) yields the following formula:

$$\begin{aligned}
2R^{ursv} = &-g^{rv,us} - g^{us,rv} + g^{rs,uv} + g^{uv,rs} + 2\widetilde{\Gamma}^{rv}_{\rho}\widetilde{\Gamma}^{us}_{\sigma}g^{\rho\sigma} - 2\widetilde{\Gamma}^{rs}_{\rho}\widetilde{\Gamma}^{uv}_{\sigma}g^{\rho\sigma}\\
&+\ g^{r\lambda,u}g_{\lambda\mu}\,g^{\mu v,s} - g^{r\lambda,u}g_{\lambda\mu}\,g^{\mu s,v} + g^{u\lambda,r}g_{\lambda\mu}\,g^{\mu s,v} - g^{u\lambda,r}g_{\lambda\mu}\,g^{\mu v,s}\\
&+\ g^{r\lambda,s}g_{\lambda\mu}\,g^{\mu v,u} + g^{u\lambda,v}g_{\lambda\mu}\,g^{\mu s,r} - g^{r\lambda,v}g_{\lambda\mu}\,g^{\mu s,u} - g^{u\lambda,s}g_{\lambda\mu}\,g^{\mu v,r}.
\end{aligned} \tag{4.16}$$

In other words the structure of R^{ursv} is formally analogous to the one of R_{ijkm} expressed by formula (4.5), *except* for the eight “correction terms” that appear in the last two lines of (4.16). Note, in particular, that the last four terms could be combined with the products of dual Christoffel symbols that appear in the first line. However, inserting the definitions of the Christoffel symbols so that they can be combined with the last four terms above is not that convenient, since the last four terms in (4.16) do not include any of the derivatives of the type $g^{rv,\varphi}$, $g^{us,\varphi}$, $g^{rs,\varphi}$ or $g^{uv,\varphi}$, therefore doing so does not yield any significant simplification: in fact the two products of Christoffel symbols correspond to *eighteen* more elementary terms, only four of which can be combined with the four terms in the third line of the above expression. Also note that the terms in the second line of (4.16) do not combine with anything else but stand by themselves.

We will now introduce some useful notation which will later ease the computation of sectional curvature in terms of the cometric tensor and its partial derivatives. Define $B_{\varphi}^{u,v} \triangleq g_{\varphi\psi}\, g^{\psi u,v}$ and consider the cotangent vectors $B^{u,v} = B_{\varphi}^{u,v} dx^{\varphi} \in T_p^*\mathcal{M}$; in such expressions indices u and v do *not* commute. The cometric, i.e. the inner product of cotangent vectors of this type in the cotangent space $T_p^*\mathcal{M}$ is given by $\langle B^{u,v}, B^{s,r}\rangle = B_{\varphi}^{u,v} B_{\psi}^{s,r} \langle dx^{\varphi}, dx^{\psi}\rangle = B_{\varphi}^{u,v} B_{\psi}^{s,r} g^{\varphi\psi} = g^{\eta u,v} g_{\eta\sigma}\, g^{\sigma s,r}$. Similarly, let $\widetilde{\Gamma}^{uv} = \widetilde{\Gamma}_{\varphi}^{uv}\, dx^{\varphi} \in T_p^*\mathcal{M}$ so that $\langle \widetilde{\Gamma}^{uv}, \widetilde{\Gamma}^{sr}\rangle = \widetilde{\Gamma}_{\varphi}^{uv} \widetilde{\Gamma}_{\psi}^{sr} \langle dx^{\varphi}, dx^{\psi}\rangle = \widetilde{\Gamma}_{\varphi}^{uv} \widetilde{\Gamma}_{\psi}^{sr} g^{\varphi\psi}$. Note that with an abuse of notation we have indicated the cometric with $\langle\cdot,\cdot\rangle$, i.e. using the same symbol that we had used for the metric in previous sections. Also, for any pair of tangent vectors and the corresponding cotangent vectors $X^{\flat} = X_i\, dx^i$, $Y^{\flat} = Y_i\, dx^i$ in $T_p^*\mathcal{M}$ define a new vector $B(X^{\flat}, Y^{\flat}) \in T_p^*\mathcal{M}$ as follows: $B(X^{\flat}, Y^{\flat}) \triangleq X_u Y_v\, B_{\varphi}^{u,v}\, dx^{\varphi}$; again, note that in such definition $X^{\flat}$ and $Y^{\flat}$ do not commute. For a given set of dual Christoffel symbols, in an analogous fashion we define: $\widetilde{\Gamma}(X^{\flat}, Y^{\flat}) \triangleq X_u Y_v\, \widetilde{\Gamma}_{\varphi}^{uv}\, dx^{\varphi}$, in which case $X^{\flat}$ and $Y^{\flat}$ *do* commute.

With the above conventions we have that (4.16) can be rewritten as follows:

$$
\begin{aligned}
(4.17)\quad 2R^{ursv} &= -g^{rv,us} - g^{us,rv} + g^{rs,uv} + g^{uv,rs} + 2\langle \widetilde{\Gamma}^{rv}, \widetilde{\Gamma}^{us}\rangle - 2\langle \widetilde{\Gamma}^{rs}, \widetilde{\Gamma}^{uv}\rangle \\
&\quad + \langle B^{r,u}, B^{v,s}\rangle - \langle B^{r,u}, B^{s,v}\rangle + \langle B^{u,r}, B^{s,v}\rangle - \langle B^{u,r}, B^{v,s}\rangle \\
&\quad + \langle B^{r,s}, B^{v,u}\rangle + \langle B^{u,v}, B^{s,r}\rangle - \langle B^{r,v}, B^{s,u}\rangle - \langle B^{u,s}, B^{v,r}\rangle.
\end{aligned}
$$

Using the above expression one can easily verify that the dual tensor satisfies identities $R^{rusv} = -R^{ursv}$, $R^{urvs} = -R^{ursv}$, $R^{svur} = R^{ursv}$, and $R^{ursv} + R^{usvr} + R^{uvrs} = 0$ for any choice of the indices u, r, s and v, which are analogous to symmetries (4.3) for the regular Riemannian curvature tensor. Equation (4.17) can be manipulated fairly easily to compute the numerator of sectional curvature (4.4). In fact in the next section we will multiply it by the components of cotangent vectors $(X^{\flat}, Y^{\flat}, Y^{\flat}, X^{\flat})$, many simplifications will occur and the numerator of (4.4) will finally take a surprisingly simple and elegant form (in terms of the cometric tensor and its partial derivatives). However, it is actually convenient to have a "full blown" expression for the dual tensor: the reason why this is the case will be made evident is section 5 of the current chapter, where we will find bounds on sectional curvature as the solution of a generalized eigenvalue problem.

Proposition 4.3. *The Riemannian curvature tensor with raised indices R^{ursv} may be written in function of the cometric tensor and its derivatives as follows:*

$$
\begin{aligned}
&(\mathrm{T}_1) \qquad 2R^{ursv} = -\, g^{rv,su} + g^{rs,uv} - g^{us,rv} + g^{uv,rs} \\
&(\mathrm{T}_2) \qquad\qquad -\frac{1}{2}\Big\{ g^{rs,}{}_{\eta}\, g^{\eta\sigma} g^{uv,}{}_{\sigma} - g^{rs,}{}_{\psi}\, \big(g^{\psi u,v} + g^{\psi v,u}\big) - g^{uv,}{}_{\varphi}\, \big(g^{\varphi r,s} + g^{\varphi s,r}\big) \Big\} \\
&(\mathrm{T}_3) \qquad\qquad +\frac{1}{2}\Big\{ g^{us,}{}_{\eta}\, g^{\eta\sigma} g^{vr,}{}_{\sigma} - g^{us,}{}_{\psi}\, \big(g^{\psi v,r} + g^{\psi r,v}\big) - g^{vr,}{}_{\varphi}\, \big(g^{\varphi u,s} + g^{\varphi s,u}\big) \Big\} \\
&(\mathrm{T}_4) \qquad\qquad -\frac{1}{2}\big(g^{\varphi r,s} - g^{\varphi s,r}\big) g_{\varphi\psi} \big(g^{\varphi u,v} - g^{\varphi v,u}\big) \\
&(\mathrm{T}_5) \qquad\qquad +\frac{1}{2}\big(g^{\varphi u,s} - g^{\varphi s,u}\big) g_{\varphi\psi} \big(g^{\varphi r,v} - g^{\varphi v,r}\big) \\
&(\mathrm{T}_6) \qquad\qquad +\big(g^{\varphi u,r} - g^{\varphi r,u}\big) g_{\varphi\psi} \big(g^{\varphi s,v} - g^{\varphi v,s}\big),
\end{aligned}
$$

where, as usual, $g^{ij,k} \triangleq g^{ij,}{}_{\xi}\, g^{\xi k}$ *and* $g^{ij,k\ell} \triangleq g^{ij,}{}_{\xi\eta}\, g^{\xi k} g^{\eta \ell}$.

Remark. From now on, we shall refer to the six terms in the above proposition as T_1, …, T_6. Before proceeding to the proof we should note that we have achieved expressing the dual Riemannian curvature tensor in terms of the cometric tensor; the metric tensor (with "dummy" lower indices φ and ψ) still appears in terms T_4, T_5 and T_6, which will later reduce to only one in the formula for sectional curvature.

Proof. We will expand and recombine the terms in expression (4.17). Clearly the terms involving second derivatives need no manipulation. We can expand the Christoffel symbols as follows:

$$
2\langle \widetilde{\Gamma}^{us}, \widetilde{\Gamma}^{rv} \rangle = \frac{1}{2}\big[\big(g^{\varphi u,s} + g^{\varphi s,u}\big) - g^{us,\varphi}\big] g_{\varphi\eta}\, g^{\eta\sigma}\, g_{\sigma\psi} \big[\big(g^{\psi r,v} + g^{\varphi v,r}\big) - g^{rv,\psi}\big],
$$

that is:

$$2\langle\widetilde{\Gamma}^{us},\widetilde{\Gamma}^{rv}\rangle = \frac{1}{2}\Big\{g^{us,}{}_{\eta}\, g^{\eta\sigma} g^{rv,}{}_{\sigma} - g^{us,}{}_{\psi}\left(g^{\psi r,v} + g^{\psi v,r}\right) - g^{rv,}{}_{\varphi}\left(g^{\varphi u,s} + g^{\varphi s,u}\right)\Big\}$$
$$+ \frac{1}{2}\underbrace{\left(g^{\varphi u,s} + g^{\varphi s,u}\right) g_{\varphi\psi}\left(g^{\psi r,v} + g^{\psi v,r}\right)}_{\langle B^{u,s} + B^{s,u},\, B^{r,v} + B^{v,r}\rangle}.$$

We can now combine the inner product on the right-hand side of the above formula with the last two terms of (4.17) to get:

$$\frac{1}{2}\langle B^{u,s} + B^{s,u}, B^{r,v} + B^{v,r}\rangle - \langle B^{r,v}, B^{s,u}\rangle - \langle B^{u,s}, B^{v,r}\rangle = \frac{1}{2}\langle B^{u,s} - B^{s,u}, B^{r,v} - B^{v,r}\rangle$$
$$= \frac{1}{2}\left(g^{\varphi u,s} - g^{\varphi s,u}\right) g_{\varphi\psi}\left(g^{\psi r,v} - g^{\psi v,r}\right),$$

whence we may conclude that

$$2\langle\widetilde{\Gamma}^{us},\widetilde{\Gamma}^{rv}\rangle - \langle B^{r,v}, B^{s,u}\rangle - \langle B^{u,s}, B^{v,r}\rangle = \mathrm{T}_3 + \mathrm{T}_5.$$

In a completely similar fashion one can prove that:

$$-2\langle\widetilde{\Gamma}^{rs},\widetilde{\Gamma}^{uv}\rangle + \langle B^{r,s}, B^{v,u}\rangle + \langle B^{u,v}, B^{s,r}\rangle = \mathrm{T}_2 + \mathrm{T}_4.$$

As far as the remaining four correction terms in the second line of formula (4.17) it is the case that:

$$\langle B^{r,u}, B^{v,s}\rangle - \langle B^{r,u}, B^{s,v}\rangle + \langle B^{u,r}, B^{s,v}\rangle - \langle B^{u,r}, B^{v,s}\rangle \;=\; \langle B^{r,u} - B^{u,r}, B^{v,s} - B^{s,v}\rangle$$
$$=\; \left(g^{\varphi r,u} - g^{\varphi u,r}\right) g_{\varphi\psi}\left(g^{\psi v,s} - g^{\psi s,v}\right) \;=\; \mathrm{T}_6,$$

and this concludes the proof. □

4. Sectional Curvature in terms of the cometric tensor

The numerator of sectional curvature may be computed from the dual Riemannian curvature tensor by employing (4.10). Instead of using the "full-blown" expression provided by Proposition 4.3 we will work on the more compact formula (4.17), which will make the procedure somewhat smoother and more elegant. Multiplying by the components of cotangent vectors $(X^\flat, Y^\flat, Y^\flat, X^\flat)$ yields:

$$\begin{aligned}
2R^{ursv}X_uY_rY_sX_v \;=\; & -\; Y_rX_vg^{rv,us}X_uY_s - X_uY_sg^{us,rv}Y_rX_v \\
& +\; Y_rY_sg^{rs,uv}X_uX_v + X_uX_vg^{uv,rs}Y_rY_s \\
& +\; 2Y_rX_v\langle\widetilde{\Gamma}^{rv}, \widetilde{\Gamma}^{us}\rangle X_uY_s - \; 2Y_rY_s\langle\widetilde{\Gamma}^{rs}, \widetilde{\Gamma}^{uv}\rangle X_uX_v \\
& +\; Y_rX_u\langle B^{r,u}, B^{v,s}\rangle X_vY_s - Y_rX_u\langle B^{r,u}, B^{s,v}\rangle Y_sX_v \\
& +\; X_uY_r\langle B^{u,r}, B^{s,v}\rangle Y_sX_v - X_uY_r\langle B^{u,r}, B^{v,s}\rangle X_vY_s \\
& +\; Y_rY_s\langle B^{r,s}, B^{v,u}\rangle X_vX_u - Y_rX_v\langle B^{r,v}, B^{s,u}\rangle Y_sX_u \\
& +\; X_uX_v\langle B^{u,v}, B^{s,r}\rangle Y_sY_r - X_uY_s\langle B^{u,s}, B^{v,r}\rangle X_vY_r,
\end{aligned}$$

that is,

$$\begin{aligned}
2R^{ursv}X_uY_rY_sX_v \;=\; & Y_rY_sg^{rs,uv}X_uX_v + X_uX_vg^{uv,rs}Y_rY_s - 2X_uY_sg^{us,rv}Y_rX_v \\
& +\; 2Y_rX_v\langle\widetilde{\Gamma}^{rv}, \widetilde{\Gamma}^{us}\rangle X_uY_s - \; 2Y_rY_s\langle\widetilde{\Gamma}^{rs}, \widetilde{\Gamma}^{uv}\rangle X_uX_v \\
& +\; 2X_uX_v\langle B^{u,v}, B^{s,r}\rangle Y_sY_r + 2X_uY_r\langle B^{u,r}, B^{s,v}\rangle Y_sX_v \\
& -\; 2X_uY_s\langle B^{u,s}, B^{v,r}\rangle X_vY_r - 2Y_sX_u\langle B^{s,u}, B^{r,v}\rangle Y_rX_v\,.
\end{aligned}$$

Proceeding in a way that is completely analogous to the the proof of Proposition 4.2 we may rewrite the above expression as follows:

$$\begin{aligned}
2R^{ursv}X_uY_rY_sX_v \;=\; & (X_uY_r - Y_uX_r)\, g^{su,rv}(X_sY_v - Y_sX_v) \\
& +\; 2\|\widetilde{\Gamma}(X^\flat, Y^\flat)\|^2 - 2\langle\widetilde{\Gamma}(X^\flat, X^\flat), \widetilde{\Gamma}(Y^\flat, Y^\flat)\rangle \\
& +\; 2X_uX_v\langle B^{u,v}, B^{s,r}\rangle Y_sY_r + 2X_uY_r\langle B^{u,r}, B^{s,v}\rangle Y_sX_v \\
& -\; 2X_uY_s\langle B^{u,s}, B^{v,r}\rangle X_vY_r - 2Y_sX_u\langle B^{s,u}, B^{r,v}\rangle Y_rX_v\,,
\end{aligned} \tag{4.19}$$

so that the "correction terms" have been reduced in number, with respect to expressions (4.16) and (4.17), from eight to four. No further simplification is possible, unless we insert the definition of dual Christoffel symbols into the above equation. That is exactly what we are going to do next; now that the number of terms has almost been halved, the computation is not going to be unbearably complicated.

THEOREM 4.4. *For an arbitrary pair of vectors $X = X^i\partial_i$ and $Y = Y^i\partial_i$ in $T_p\mathcal{M}$ the numerator of sectional curvature* (4.4) *at point p may be written as:*

$$2R^{ursv}X_uY_rY_sX_v$$
$$= \big(X_uY_r - Y_uX_r\big)\Big(g^{su,rv} - \frac{1}{4}\,g^{us,}_{\ \ \ \varphi}\;g^{rv,\varphi} + g^{us,}_{\ \ \ \varphi}\;g^{\varphi r,v} - \frac{3}{2}\,g^{u\psi,r}g_{\psi\xi}\,g^{\xi s,v}\Big)\big(X_sY_v - Y_sX_v\big),$$

where $X_i = g_{ij}X^j$, $Y_i = g_{ij}Y^j$ and, as usual, $g^{ij,k} \triangleq g^{ij,}_{\ \ \ \xi}\;g^{\xi k}$ and $g^{ij,k\ell} \triangleq g^{ij,}_{\ \ \ \xi\eta}\;g^{\xi k}g^{\eta\ell}$.

PROOF. We shall split the right-hand side of (4.19) into three terms,

$$\Phi_1 \triangleq (X_uY_r - Y_uX_r)\,g^{su,rv}(X_sY_v - Y_sX_v),$$
$$\Phi_2 \triangleq 2\|\widetilde{\Gamma}(X^\flat, Y^\flat)\|^2 - 2\langle\widetilde{\Gamma}(X^\flat, X^\flat), \widetilde{\Gamma}(Y^\flat, Y^\flat)\rangle,$$

and

$$\Phi_3 \triangleq 2X_uX_v\langle B^{u,v}, B^{s,r}\rangle Y_sY_r + 2X_uY_r\langle B^{u,r}, B^{s,v}\rangle Y_sX_v$$
$$-\ \ 2X_uY_s\langle B^{u,s}, B^{v,r}\rangle X_vY_r - 2Y_sX_u\langle B^{s,u}, B^{r,v}\rangle Y_rX_v\,,$$

so that the second derivatives of the inverse of the metric tensor are in term Φ_1, the terms that concern the Christoffel symbols are contained in term Φ_2, whereas term Φ_3 consists of the “correction” previously discussed. Obviously we have that $2R^{ursv}X_uY_rY_sX_v = \Phi_1 + \Phi_2 + \Phi_3$. Term Φ_1 is in a form that cannot be further simplified. We shall manipulate the other two terms, Φ_2 and Φ_3, and then combine them; this will yield the surprisingly simple final form for the numerator of sectional curvature as a function of the derivatives of the inverse of the metric tensor reported in the statement of the theorem.

Inserting the definition of the dual Christoffel symbols into Φ_2 yields:

$$\Phi_2 \;=\; \frac{1}{2}X_uY_r\Big[\big(g^{u\psi,r} + g^{r\psi,u}\big) - g^{ur,\psi}\Big]g_{\psi\varphi}\,g^{\varphi\eta}g_{\eta\xi}\Big[\big(g^{\xi s,v} + g^{\xi v,s}\big) - g^{sv,\xi}\Big]Y_sX_v$$
$$-\,\frac{1}{2}X_uX_v\Big[\big(g^{u\psi,v} + g^{v\psi,u}\big) - g^{uv,\psi}\Big]g_{\psi\varphi}\,g^{\varphi\eta}g_{\eta\xi}\Big[\big(g^{\xi r,s} + g^{\xi s,r}\big) - g^{rs,\xi}\Big]Y_rY_s,$$

that is

$$\Phi_2 = +\frac{1}{2} X_u Y_r \Big[g^{ur,}{}_{\varphi}\, g^{\varphi\eta} g^{sv,}{}_{\eta} - g^{ur,}{}_{\xi} \big(g^{\xi s,v} + g^{\xi v,s}\big) - g^{ur,}{}_{\psi} \big(g^{\psi u,r} + g^{\psi r,u}\big) \Big] Y_s X_v \tag{4.20a}$$

$$- \frac{1}{2} X_u X_v \Big[g^{uv,}{}_{\varphi}\, g^{\varphi\eta} g^{rs,}{}_{\eta} - g^{uv,}{}_{\xi} \big(g^{\xi r,s} + g^{\xi s,r}\big) - g^{rs,}{}_{\psi} \big(g^{\psi u,v} + g^{\psi v,u}\big) \Big] Y_r Y_s \tag{4.20b}$$

$$+ \frac{1}{2} X_u Y_r \big(g^{\psi u,r} + g^{\psi r,u}\big) g_{\psi\xi} \big(g^{\xi s,v} + g^{\xi v,s}\big) X_v Y_s \tag{4.20c}$$

$$- \frac{1}{2} X_u X_v \big(g^{\psi u,v} + g^{\psi v,u}\big) g_{\psi\xi} \big(g^{\xi r,s} + g^{\xi s,r}\big) Y_r Y_s \,. \tag{4.20d}$$

The last two terms of (4.20a) may be written (following the multiplication by $X_u Y_r Y_s X_v$) as

$$\begin{aligned} X_u Y_r g^{ur,}{}_{\xi} \big(g^{\xi s,v} + g^{\xi v,s}\big) Y_s X_v + Y_s X_v g^{sv,}{}_{\xi} \big(g^{\xi u,r} + g^{\xi r,u}\big) X_u Y_r \\ = 2\, X_u Y_r g^{ur,}{}_{\xi} \big(g^{\xi s,v} + g^{\xi v,s}\big) Y_s X_v \,, \end{aligned}$$

whereas the second term of (4.20b) may be manipulated as follows:

$$X_u X_v\, g^{uv,}{}_{\xi} \big(g^{\xi r,s} + g^{\xi s,r}\big) Y_r Y_s = 2\, X_u X_v\, g^{uv,}{}_{\xi}\, g^{\xi r,s} Y_r Y_s$$

and analogously for the third term of (4.20b):

$$Y_r Y_s\, g^{rs,}{}_{\xi} \big(g^{\xi u,v} + g^{\xi v,u}\big) X_u X_v = 2\, Y_r Y_s\, g^{rs,}{}_{\xi}\, g^{\xi u,v} X_u X_v \,.$$

Furthermore, the multiplication in line (4.20c) can be restated in the following terms:

$$\begin{aligned} X_u Y_r \big(g^{\psi u,r} + g^{\psi r,u}\big) g_{\psi\xi} \big(g^{\xi s,v} + g^{\xi v,s}\big) X_v Y_s \\ = \big(X_u Y_r + Y_u X_r\big) g^{\psi r,u} g_{\psi\xi}\, g^{\xi s,v} \big(X_v Y_s + Y_v X_s\big), \end{aligned}$$

and the one in line (4.20d) can be simplified as:

$$X_u X_v \big(g^{\psi u,v} + g^{\psi v,u}\big) g_{\psi\xi} \big(g^{\xi r,s} + g^{\xi s,r}\big) Y_r Y_s = 4\, X_u X_v\, g^{\psi u,v} g_{\psi\xi}\, g^{\xi r,s}\, Y_r Y_s \,.$$

So Φ_2 can be re-expressed as follows:

$$\begin{aligned} \Phi_2 = X_u Y_r \Big[\frac{1}{2} g^{ur,}{}_{\varphi}\, g^{\varphi\eta} g^{sv,}{}_{\eta} - g^{ur,}{}_{\xi} \big(g^{\xi s,v} + g^{\xi v,s}\big) \Big] Y_s X_v \\ - X_u X_v \Big[\frac{1}{2} g^{uv,}{}_{\varphi}\, g^{\varphi\eta} g^{rs,}{}_{\eta} - g^{uv,}{}_{\xi}\, g^{\xi r,s} - g^{rs,}{}_{\xi}\, g^{\xi u,v} \Big] Y_r Y_s \\ + \frac{1}{2} \big(X_u Y_r + Y_u X_r\big) g^{\psi r,u} g_{\psi\xi}\, g^{\xi s,v} \big(X_v Y_s + Y_v X_s\big) - 2\, X_u X_v\, g^{\psi u,v} g_{\psi\xi}\, g^{\xi r,s}\, Y_r Y_s \,. \end{aligned}$$

Further simplifications are actually possible as far as the first two terms of the above expression are concerned. In fact, first note that:

$$
\begin{aligned}
-\frac{1}{4}\big(X_u Y_r - Y_u X_r\big) g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, \big(X_s Y_v - Y_s X_v\big) \;&= \\
= \; -\frac{1}{4}\big(&X_u Y_r\, g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, X_s Y_v - Y_u X_r\, g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, X_s Y_v \\
&- X_u Y_r\, g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, Y_s X_v + Y_u X_r\, g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, Y_s X_v\big) \\
= \; -\frac{1}{4}\big(&X_u X_s\, g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, Y_r Y_v - Y_u X_s\, g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, X_r Y_v \\
&- X_u Y_s\, g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, Y_r X_v + Y_u Y_s\, g^{us,}{}_{\varphi}\, g^{\varphi\xi} g^{rv,}{}_{\xi}\, X_r X_v\big) \\
= \; \frac{1}{2} X_u Y_r\, g^{ur,}{}_{\varphi}\, & g^{\varphi\xi} g^{sv,}{}_{\xi}\, X_s Y_v - \frac{1}{2} X_u X_v\, g^{uv,}{}_{\varphi}\, g^{\varphi\xi} g^{rs,}{}_{\xi}\, Y_r Y_s\,.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\big(X_u Y_r - Y_u X_r\big) g^{us,}{}_{\varphi}\, g^{\varphi r,v}\big(X_s Y_v - Y_s X_v\big) \\
&= \; X_u Y_r g^{us,}{}_{\varphi}\, g^{\varphi r,v} X_s Y_v - Y_u X_r g^{us,}{}_{\varphi}\, g^{\varphi r,v} X_s Y_v \\
&\quad\; - X_u Y_r g^{us,}{}_{\varphi}\, g^{\varphi r,v} Y_s X_v + Y_u X_r g^{us,}{}_{\varphi}\, g^{\varphi r,v} Y_s X_v \\
&= \; X_u X_s g^{us,}{}_{\varphi}\, g^{\varphi r,v} Y_r Y_v - Y_u X_s g^{us,}{}_{\varphi}\, g^{\varphi r,v} X_r Y_v \\
&\quad\; - X_u Y_s g^{us,}{}_{\varphi}\, g^{\varphi r,v} Y_r X_v + Y_u Y_s g^{us,}{}_{\varphi}\, g^{\varphi r,v} X_r X_v \\
&= \; X_u Y_v \Big[- g^{ur,}{}_{\xi}\, \big(g^{\xi s,v} + g^{\xi v,s}\big)\Big] Y_s X_v - X_u X_v \Big[- g^{uv,}{}_{\xi}\, g^{\xi r,s} - g^{rs,}{}_{\xi}\, g^{\xi u,v}\Big] Y_r Y_s\,,
\end{aligned}
$$

so that Φ_2 can finally be written as

$$
\begin{aligned}
\Phi_2 =& \big(X_u Y_r - Y_u X_r\big)\Big(- \frac{1}{4}\, g^{us,}{}_{\varphi}\, g^{rv,\varphi} + g^{us,}{}_{\varphi}\, g^{\varphi r,v}\Big)\big(X_s Y_v - Y_s X_v\big) \\
&+ \frac{1}{2}\big(X_u Y_r + Y_u X_r\big) g^{\psi r,u} g_{\psi\xi}\, g^{\xi s,v}\big(X_v Y_s + Y_v X_s\big) - 2\, X_u X_v\, g^{\psi u,v} g_{\psi\xi}\, g^{\xi r,s}\, Y_r Y_s\,.
\end{aligned}
$$

As far as term Φ_3 is concerned, we shall simply rewrite it as

$$
\begin{aligned}
\Phi_3 \;=\; & 2X_u X_v g^{u\psi,v} g_{\psi\xi}\, g^{s\xi,r} Y_s Y_r + 2X_u Y_r g^{u\psi,r} g_{\psi\xi}\, g^{\xi s,v} Y_s X_v \\
& -\; 2X_u Y_s g^{u\psi,s} g_{\psi\xi}\, g^{v\xi,r} X_v Y_r - 2Y_s X_u g^{s\psi,u} g_{\psi\xi}\, g^{r\xi,v} Y_r X_v\,.
\end{aligned}
$$

Combining such formula with the latest expression for Φ_2 gives:

$$\begin{aligned}\Phi_2+\Phi_3 &= \big(X_uY_r-Y_uX_r\big)\Big(-\frac{1}{4}\,g^{us,}{}_{\varphi}\,g^{rv,\varphi}+g^{us,}{}_{\varphi}\,g^{\varphi r,v}\Big)\big(X_sY_v-Y_sX_v\big)\\ &+\ \frac{1}{2}\big(X_uY_r+Y_uX_r\big)g^{\psi r,u}g_{\psi\xi}\,g^{\xi s,v}\big(X_vY_s+Y_vX_s\big)\\ &+\ 2X_uY_rg^{u\psi,r}g_{\psi\xi}\,g^{\xi s,v}Y_sX_v\\ &-\ 2X_uY_sg^{u\psi,s}g_{\psi\xi}\,g^{v\xi,r}X_vY_r-2Y_sX_ug^{s\psi,u}g_{\psi\xi}\,g^{r\xi,v}Y_rX_v\,.\end{aligned}$$

Performing the multiplication on the second line of the right-hand side and recombining the resulting terms yields:

$$\begin{aligned}\Phi_2+\Phi_3 &= \big(X_uY_r-Y_uX_r\big)\Big(-\frac{1}{4}\,g^{us,}{}_{\varphi}\,g^{rv,\varphi}+g^{us,}{}_{\varphi}\,g^{\varphi r,v}\Big)\big(X_sY_v-Y_sX_v\big)\\ &+\ 3X_uY_rg^{u\psi,r}g_{\psi\xi}\,g^{\xi s,v}Y_sX_v\\ &-\ \frac{3}{2}X_uY_sg^{u\psi,s}g_{\psi\xi}\,g^{v\xi,r}X_vY_r-\frac{3}{2}Y_sX_ug^{s\psi,u}g_{\psi\xi}\,g^{r\xi,v}Y_rX_v\,.\end{aligned}$$

But we have that

$$\begin{aligned}&-\frac{3}{2}\big(X_uY_r-Y_uX_r\big)g^{u\psi,r}g_{\psi\xi}\,g^{\xi s,v}\big(X_sY_v-Y_sX_v\big)\\ &\quad=\ 3X_uY_rg^{u\psi,r}g_{\psi\xi}\,g^{\xi s,v}Y_sX_v-\frac{3}{2}X_uY_sg^{u\psi,s}g_{\psi\xi}\,g^{v\xi,r}X_vY_r-\frac{3}{2}Y_sX_ug^{s\psi,u}g_{\psi\xi}\,g^{r\xi,v}Y_rX_v\,,\end{aligned}$$

therefore:

$$\begin{aligned}&2R^{ursv}X_uY_rY_sX_v\ =\ \Phi_1+\Phi_2+\Phi_3\ =\\ &\quad=\big(X_uY_r-Y_uX_r\big)\left(g^{su,rv}-\frac{1}{4}\,g^{us,}{}_{\varphi}\,g^{rv,\varphi}+g^{us,}{}_{\varphi}\,g^{\varphi r,v}-\frac{3}{2}\,g^{u\psi,r}g_{\psi\xi}\,g^{\xi s,v}\right)\big(X_sY_v-Y_sX_v\big),\end{aligned}$$

which is precisely what we wanted to prove. □

Remark. We have expressed the numerator of sectional curvature in terms of the cometric tensor and its derivatives; in the formula provided by Theorem 4.4 the only term in the middle factor that depends on the metric tensor (with lower indices) is the fourth one. We should note that the formula was later verified by professor Peter W. Michor of the University of Vienna, who provided an alternative proof [**30**] with index-free notation.

5. Bounds on Sectional Curvature

In this section we show how sectional curvature for an n-dimensional manifold $\mathcal{M}$ can be written as the ratio of quadratic forms on the space of alternating 2-forms $\Lambda^2(T_p\mathcal{M})$. This allows to formulate the the problem of finding bounds for sectional curvature as a generalized eigenvalue problem. In the case of three-dimensional manifolds (such as the manifold of three landmarks in one dimension) these bounds are actually achieved by computable pairs of tangent vectors. Following the spirit of the chapter we will express the results and formulas in terms of the cometric tensor and the dual Riemannian curvature tensor. These results, as well as those achieved in the previous sections, will be used in the next chapter to compute and plot the sectional curvature of landmark manifolds. We will indicate with $\Lambda^k(V)$ the linear space of alternating k-forms on a linear space V (e.g. $\Lambda^1(V) = V^*$), and with $\Sigma^k(V)$ the set of symmetric k-tensors on space V; for reference, see [**21**] and [**28**].

The following proposition holds:

Proposition 4.5. *For an arbitrary pair $X, Y \in T_p\mathcal{M}$ define $\omega = X^\flat \wedge Y^\flat \in \Lambda^2(T_p\mathcal{M})$, i.e. $\omega = \sum_{u<r}(X_u Y_r - Y_u X_r)\, dx^u \wedge dx^r$. We can write sectional curvature at point $p \in \mathcal{M}$ as follows:*

$$K(X, Y) = \frac{\varrho(\omega, \omega)}{\gamma(\omega, \omega)}\,,$$

where:

- *ϱ is a symmetric bilinear form on $\Lambda^2(T_p\mathcal{M})$, i.e. $\varrho \in \Sigma^2\big(\Lambda^2(T_p\mathcal{M})\big)$, defined as follows:*[4]

  $$\varrho(\eta, \xi) \;=\; -\,\eta_{ur}\,\frac{R^{ursv}}{4}\,\xi_{sv}$$

  for any pair $\eta, \xi \in \Lambda^2(T_p\mathcal{M})$. In the above definition we intend the following: $\eta_{ur} = \eta\big(\frac{\partial}{\partial x^u}, \frac{\partial}{\partial x^r}\big)$, i.e. $\eta = \sum_{u<r} \eta_{ur}\, dx^u \wedge dx^r$, and similarly for ξ.

[4]The minus signs follows from the sign conventions for the Riemannian curvature tensor that we adopted, following Lee's notation [**27**]; see footnote 1 of this chapter.

- *γ is a symmetric bilinear form on $\Lambda^2(T_p\mathcal{M})$, i.e. $\gamma \in \Sigma^2\big(\Lambda^2(T_p\mathcal{M})\big)$, defined as follows:*

$$\gamma(\eta,\xi) \;=\; \eta_{ur}\,\frac{G^{ursv}}{4}\,\xi_{sv}$$

for any pair $\eta,\xi \in \Lambda^2(T_p\mathcal{M})$, with $G^{ursv} = g^{us}g^{rv} - g^{uv}g^{rs}$. Note that G^{ursv} has the same symmetries as the dual Riemannian curvature tensor R^{ursv}.

Proof. Sectional curvature associated to the 2-plane spanned by tangent vectors X and Y may be written as (4.4): $K(X,Y) = \frac{X^iY^jR_{ijk\ell}Y^kX^\ell}{\|X\|^2\|Y\|^2-\langle X,Y\rangle^2}$. We have that $X^\flat = X_u dx^u$ and $Y^\flat = Y_r dx^r$, so that $\omega = X^\flat \wedge Y^\flat = X_uY_r\,dx^u \wedge dx^r = \sum_{u<r}(X_uY_r - X_rY_u)\,dx^u \wedge dx^r$, by the skew-symmetry of the wedge product. By definition, $\varrho(\eta,\xi) = -\frac{1}{4}\eta_{ur}R^{ursv}\xi_{sv}$, whence:

$$\begin{aligned}
\varrho(\omega,\omega) &= -\frac{1}{4}(X_uY_r - X_rY_u)R^{ursv}(X_sY_v - X_vY_s)\\
&= -\frac{1}{4}\big(X_uY_rR^{ursv}X_sY_v - X_uY_rR^{ursv}X_vY_s - X_rY_uR^{ursv}X_sY_v + X_rY_uR^{ursv}X_vY_s\big)\\
&= -\frac{1}{4}\cdot 4\,X_uY_rR^{ursv}X_sY_v \;=\; X_uY_rR^{ursv}Y_sX_v\,,
\end{aligned}$$

where we have used the well-knows symmetries of R^{ursv} multiple times. As far as the denominator is concerned, first note that $G^{ursv} = g^{us}g^{rv} - g^{uv}g^{rs}$ has the same symmetries as the dual Riemannian curvature tensor R^{ursv}, since

$$\begin{aligned}
G^{rusv} &= g^{rs}g^{uv} - g^{rv}g^{us} &= -G^{ursv},\\
G^{urvs} &= g^{uv}g^{rs} - g^{us}g^{rv} &= -G^{ursv},\\
G^{svur} &= g^{su}g^{vr} - g^{sr}g^{vu} &= G^{ursv}.
\end{aligned}$$

By definition $\gamma(\eta,\xi) = \frac{1}{4}\eta_{ur}G^{ursv}\xi_{sv}$, whence:

$$\begin{aligned}
\gamma(\omega,\omega) &= \frac{1}{4}(X_uY_r - X_rY_u)G^{ursv}(X_sY_v - X_vY_s) \;=\; X_uY_rG^{ursv}X_sY_v\\
&= X_uY_r(g^{us}g^{rv} - g^{uv}g^{rs})X_sY_v \;=\; X_ug^{us}X_s\,Y_rg^{rv}Y_v - X_ug^{uv}Y_v\,Y_rg^{rs}X_s\\
&= \|X\|^2\|Y\|^2 - \langle X,Y\rangle^2,
\end{aligned}$$

since we have, for example, that $X_ug^{us}X_s = X_ug^{u\xi}g_{\xi\psi}g^{\psi s}X_s = X^\xi g_{\xi\psi}X^\psi = \|X\|^2$. $\square$

Proposition 4.6. *Bilinear forms ϱ and γ may be expressed as:*

$$\varrho(\eta,\xi) \;=\; -\sum_{u<r}\eta_{ur}\sum_{s<v}R^{ursv}\,\xi_{sv} \quad \textit{and} \quad \gamma(\eta,\xi) \;=\; \sum_{u<r}\eta_{ur}\sum_{s<v}G^{ursv}\,\xi_{sv}\,,$$

for any $\eta,\xi\in\Lambda^2(T_p\mathcal{M})$.

Proof. It follows from $\eta_{ru} = -\eta_{ur}$, $\xi_{ru} = -\xi_{ur}$ and the symmetries the two tensors of G^{ursv} and R^{ursv}. □

Corollary 4.7. *Sectional curvature may be expressed as follows:*

$$K(X,Y) = \frac{-\sum_{u<r}\omega_{ur}\sum_{s<v}R^{ursv}\,\omega_{sv}}{\sum_{\bar u<\bar r}\omega_{\bar u\bar r}\sum_{\bar s<\bar v}G^{\bar u\bar r\bar s\bar v}\,\omega_{\bar s\bar v}}\,, \tag{4.21}$$

where $\omega_{ur} = X_uY_r - X_rY_u$ for any pair of indices u,r.

We also introduce the linear operators:

$$\widetilde{R} : \Lambda^2(T_p\mathcal{M}) \to \Lambda^2(T_p^*M) \tag{4.22}$$

$$: \sum_{s<v}\omega_{sv}\,dx^s\wedge dx^v \mapsto \sum_{u<r}w^{ur}\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r}, \qquad \text{with} \qquad w^{ur} = \sum_{s<v}R^{ursv}\omega_{sv}\,,$$

and

$$\widetilde{G} : \Lambda^2(T_p\mathcal{M}) \to \Lambda^2(T_p^*M) \tag{4.23}$$

$$: \sum_{s<v}\omega_{sv}\,dx^s\wedge dx^v \mapsto \sum_{u<r}w^{ur}\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r}, \qquad \text{with} \qquad w^{ur} = \sum_{s<v}G^{ursv}\omega_{sv}\,.$$

Proposition 4.8. *Linear operator $\widetilde{G}$ is invertible. In fact, if $w = \widetilde{G}\omega$ for some arbitrary $\omega\in\Lambda^2(T_p\mathcal{M})$ then $\omega_{sv} = \sum_{u<r}H_{svur}w^{ur}$, with $H_{svur} = g_{su}g_{vr} - g_{sr}g_{vu}$.*

Proof. The proof consists of the following straightforward computation:

$$\begin{aligned}
\sum_{u<r}H_{\bar s\bar v ur}w^{ur} \;&=\; \frac{1}{2}\,H_{\bar s\bar v ur}\,w^{ur} \;=\; \frac{1}{2}\,H_{\bar s\bar v ur}\sum_{s<v}G^{ursv}\,\omega_{sv}\\
&= \frac{1}{4}\,H_{\bar s\bar v ur}\,G^{ursv}\,\omega_{sv} \;=\; \frac{1}{4}\,(g_{\bar s u}g_{\bar v r} - g_{\bar s r}g_{\bar v u})\,(g^{us}g^{rv} - g^{uv}g^{rs})\,\omega_{sv}\\
&= \frac{1}{4}\,(g_{\bar s u}g_{\bar v r}g^{us}g^{rv} - g_{\bar s u}g_{\bar v r}g^{uv}g^{rs} - g_{\bar s r}g_{\bar v u}g^{us}g^{rv} + g_{\bar s r}g_{\bar v u}g^{uv}g^{rs})\,\omega_{sv}\\
&= \frac{1}{4}(\delta^s_{\bar s}\delta^v_{\bar v} - \delta^v_{\bar s}\delta^s_{\bar v} - \delta^v_{\bar s}\delta^s_{\bar v} + \delta^s_{\bar s}\delta^v_{\bar v})\omega_{sv} \;=\; \frac{1}{2}(\delta^s_{\bar s}\delta^v_{\bar v} - \delta^v_{\bar s}\delta^s_{\bar v})\omega_{sv} \;=\; \frac{1}{2}(\omega_{\bar s\bar v} - \omega_{\bar v\bar s}) \;=\; \omega_{\bar s\bar v},
\end{aligned}$$

which, by the arbitrariness of ω, proves the proposition. □

With the above machinery, we can formulate the problem of finding bounds on sectional curvature as follows:

if: $-\sum_{s<v} R^{ursv}\omega_{sv} = \sigma \sum_{s<v} G^{ursv}\omega_{sv}$ for some $\sigma \in \mathbb{R}$, i.e. if $-\widetilde{R}\omega = \sigma\widetilde{G}\omega$, i.e. if $\sigma \in \mathbb{R}$ is a *generalized* eigenvalue [**26**] of the pair $(-\widetilde{R}, \widetilde{G})$ with generalized eigenvector ω,

then: $K(X,Y) = \sigma$.

Remark. Since $\widetilde{G}$ is invertible by Proposition 4.8 the generalized spectrum of the pair $(-\widetilde{R}, \widetilde{G})$ coincides with the ordinary spectrum of the linear transformation $\Lambda^2(T_p\mathcal{M}) \to \Lambda^2(T_p\mathcal{M}) : \omega \mapsto -\widetilde{G}^{-1}\widetilde{R}\,\omega$.

Given a chart around a point p a basis for $\Lambda^2(T_p\mathcal{M})$ is $\left\{dx^s \wedge dx^v : s < v\right\}$, while a basis for $\Lambda^2(T_p^*\mathcal{M})$ is $\left\{\frac{\partial}{\partial x^s} \wedge \frac{\partial}{\partial x^v} : s < v\right\}$; both spaces have, in fact, dimension $\binom{n}{2}$. We may express linear transformations $\widetilde{R}$ and $\widetilde{G}$ with respect to such bases with matrices, whose elements are given precisely by coefficients R^{ursv} and G^{ursv}, respectively. A way to write down matrix $\mathbf{G} \triangleq \left[G^{ursv}\right]_{\substack{u<r\\ s<v}}$ is the following:

$$\mathbf{G} = \left[\begin{array}{cccc|ccc|c}
G^{1212} & G^{1213} & \cdots & G^{121n} & G^{1223} & \cdots & G^{122n} & \cdots \\
G^{1312} & G^{1313} & \cdots & G^{131n} & G^{1323} & \cdots & G^{132n} & \cdots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \cdots \\
G^{1n12} & G^{1n13} & \cdots & G^{1n1n} & G^{1n23} & \cdots & G^{1n2n} & \cdots \\
\hline
G^{2312} & G^{2313} & \cdots & G^{231n} & G^{2323} & \cdots & G^{232n} & \cdots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \cdots \\
G^{2n12} & G^{2n13} & \cdots & G^{2n1n} & G^{2n23} & \cdots & G^{2n2n} & \cdots \\
\hline
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right] \in \mathbb{R}^{\binom{n}{2}\times\binom{n}{2}};$$

for example when $n = \dim\mathcal{M} = 3$ we get:

$$\mathbf{G} = \left[\begin{array}{cc|c}
G^{1212} & G^{1213} & G^{1223} \\
G^{1312} & G^{1313} & G^{1323} \\
\hline
G^{2312} & G^{2313} & G^{2323}
\end{array}\right].$$

Similar arguments hold for matrix $-\mathbf{R} \triangleq \left[-R^{ursv}\right]_{\substack{u<r\\ s<v}}$. Note that the generic *diagonal* element of $-\mathbf{R}$ is $-R^{urur} = R^{urru}$.

For completeness, we will state and prove the following fairly obvious fact, on the relationship between operators $\widetilde{R}$, $\widetilde{G}$ and the elements of the corresponding tensors.

PROPOSITION 4.9. *It is the case that*

$$\begin{aligned} R^{ijk\ell} &= \left[\widetilde{R}(dx^k \wedge dx^\ell)\right](dx^i, dx^j) \\ \textit{and} \quad G^{ijk\ell} &= \left[\widetilde{G}(dx^k \wedge dx^\ell)\right](dx^i, dx^j), \end{aligned}$$

for any choice of indices i, j, k and ℓ.

PROOF. We shall prove the above proposition for $G^{ijk\ell}$. For the sake of computation, it is convenient to write $\widetilde{G}\omega$, with $\omega \in \Lambda^2(T_p\mathcal{M})$, as a summation over *all* indices, as follows:

$$\begin{aligned} \widetilde{G}\omega &= \sum_{u<r}\Big(\sum_{s<v} G^{ursv}\omega_{sv}\Big)\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r} = \frac{1}{2}\sum_{u<r}\left(G^{ursv}\omega_{sv}\right)\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r} \\ &= \frac{1}{4}\, G^{ursv}\omega_{sv}\,\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r}\,, \end{aligned} \tag{4.24}$$

where we have used the symmetries of G^{ursv}, the fact that $\omega_{vs} = -\omega_{sv}$, and the skew-symmetry of the wedge product. For fixed indices k and ℓ, if $\omega = dx^k \wedge dx^\ell$ then

$$\begin{aligned} \omega_{sv} &= (dx^k\wedge dx^\ell)\Big(\frac{\partial}{\partial x^s},\frac{\partial}{\partial x^v}\Big) = dx^k\Big(\frac{\partial}{\partial x^s}\Big)\,dx^\ell\Big(\frac{\partial}{\partial x^v}\Big) - dx^k\Big(\frac{\partial}{\partial x^v}\Big)\,dx^\ell\Big(\frac{\partial}{\partial x^s}\Big) \\ &= \delta^k_s\delta^\ell_v - \delta^\ell_s\delta^k_v\,, \end{aligned}$$

by the definition of the wedge product. Substituting such expression into (4.24) yields:

$$\begin{aligned} \widetilde{G}(dx^k\wedge dx^\ell) &= \frac{1}{4}\, G^{ursv}(\delta^k_s\delta^\ell_v - \delta^\ell_s\delta^k_v)\,\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r} = \frac{1}{4}\left(G^{urk\ell} - G^{ur\ell k}\right)\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r} \\ &= \frac{1}{2}\, G^{urk\ell}\,\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r}\,, \end{aligned} \tag{4.25}$$

where we have used the symmetries of G^{ursv}. For arbitrary indices i and j it turns out that

$$\Big(\frac{\partial}{\partial x^u}\wedge\frac{\partial}{\partial x^r}\Big)(dx^i, dx^j) = \delta^i_u\delta^j_r - \delta^j_u\delta^i_r\,;$$

combining the above expression with (4.25) finally yields:

$$\big[\widetilde{G}(dx^k \wedge dx^\ell)\big](dx^i, dx^j) \;=\; \frac{1}{2}\, G^{urk\ell}\,(\delta^i_u \delta^j_r - \delta^j_u \delta^i_r) \;=\; \frac{1}{2}\,(G^{ijk\ell} - G^{jik\ell}) \;=\; G^{ijk\ell}.$$

An analogous computation holds for $\widetilde{R}$. $\square$

We shall denote by $\Sigma = \big\{\sigma_1, \ldots, \sigma_m\big\}$, $m = \binom{n}{2}$, the generalized spectrum (the set of generalized eigenvalues) of the pair $(-\widetilde{R}, \widetilde{G})$, i.e. of the pair of matrices $\big(-\mathbf{R}, \mathbf{G}\big)$. An upper bound and a lower bound for sectional curvature are given by, respectively, $\sigma_{\max} \triangleq \max \Sigma$ and $\sigma_{\min} \triangleq \min \Sigma$.

Remark. Once the set of generalized eigenvalues and eigenvectors is known, one can consider, for example, the eigenvector $\omega \in \Lambda^2(T_p\mathcal{M})$ that corresponds to the maximum eigenvalue $\sigma_{\max}$. Unfortunately since the set of *simple* (or *decomposable*[5]) 2-forms $\big\{\eta \wedge \xi \,:\, \eta, \xi \in T^*_p\mathcal{M}\big\}$ is a *proper* subset $\Lambda^2(T_p\mathcal{M})$ it is not necessarily the case that $\omega = X^\flat \wedge Y^\flat$ for some $X, Y \in T_p\mathcal{M}$. Therefore it may happen that $K(X,Y)$ is *not* equal to $\sigma_{\max}$ for any pair of tangent vectors, and $\sigma_{\max}$ and $\sigma_{\min}$ only provide, respectively, an upper and a lower bound for sectional curvature. However in the case of three-dimensional manifolds (such as the manifold of three landmarks in one dimension) the maximum and minimum eigenvalues of the spectrum Σ *are* achieved by sectional curvature $K(X,Y)$ by appropriate choices of the tangent vectors, since it *is* the case that every 2-form in $\Lambda^2(\mathbb{R}^3)$ is decomposable. In fact, if

$$\omega \;=\; \omega_{12}\, dx^1 \wedge dx^2 + \omega_{23}\, dx^2 \wedge dx^3 + \omega_{31}\, dx^3 \wedge dx^1,$$

- when $\omega_{12} \neq 0$ one can pick, for example, $\eta = dx^1 - \frac{\omega_{23}}{\omega_{12}}\, dx^3$ and $\xi = \omega_{12}\, dx^2 - \omega_{31}\, dx^3$. It turns out that $\omega = \eta \wedge \xi$.
- when $\omega_{12} = 0$ one may instead choose $\eta = -\omega_{31}\, dx^1 + \omega_{23}\, dx^2$ and $\xi = dx^3$. Again, it is the case that $\omega = \eta \wedge \xi$.

Therefore in the case of a three-dimensional manifold once the generalized eigenvector $\omega_{\max}$ that corresponds to the maximum eigenvalue $\sigma_{\max}$ is known, the above formulas allow to compute the cotangent vectors $X^\flat, Y^\flat \in T^*_p\mathcal{M}$ such that $\omega_{\max} = X^\flat \wedge Y^\flat$.

[5] A k-form $\omega \in \Lambda^k V$ is decomposable if and only if it satisfies the Plücker relations [**12**].

The corresponding pair of tangent vectors X, Y are those for which sectional curvature is maximal, i.e. $K(X,Y) = \sigma_{\max}$. The same holds for the minimum eigenvalue $\sigma_{\min}$ in the generalized spectrum.

We will conclude this section by showing how the generalized spectrum of the pair $(-\widetilde{R}, \widetilde{G})$ is related to the *scalar curvature* of the manifold. We recall from Differential Geometry [**11, 23, 27**] that the *Ricci tensor*

$$\mathrm{Ric} = R_{jk}\, dx^j \otimes dx^k$$

is the covariant 2-tensor field defined as the trace of the Riemannian curvature tensor on its first and last indices; in other words its components are defined as:

$$R_{jk} \triangleq g^{i\ell} R_{ijk\ell}.$$

The *scalar curvature* is the function S defined as the trace of the Ricci tensor, i.e.:

$$S \triangleq g^{jk} R_{jk}\,,$$

so that we can express it in terms of the dual Riemannian curvature tensor as follows:

$$\begin{aligned} S = g^{i\ell} g^{jk} R_{ijk\ell} \;&=\; g^{i\ell} g^{jk} \big(g_{iu}\, g_{jr}\, g_{ks}\, g_{\ell v}\, R^{ursv} \big) \\ = \delta^{\ell}_{u}\, \delta^{k}_{r}\, g_{ks}\, g_{\ell v}\, R^{ursv} \;&=\; g_{uv}\, g_{rs}\, R^{ursv}. \end{aligned}$$

The following proposition holds.

Proposition 4.10. *Let* $\{\sigma_1, \ldots, \sigma_m\}$, $m = \binom{n}{2}$, *be the generalized spectrum of the pair of linear operators* $(-\widetilde{R}, \widetilde{G})$, *defined in* (4.22) *and* (4.23). *Then it is the case that*

$$\sum_{i=1}^{m} \sigma_i = \frac{1}{2}\, S.$$

Proof. The generalized eigenvalues of the pair $(-\widetilde{R}, \widetilde{G})$ are the ordinary eigenvalues of the linear transformation $-\widetilde{G}^{-1}\widetilde{R} : \Lambda(T_p\mathcal{M}) \to \Lambda(T_p\mathcal{M})$, since by Proposition 4.8 linear operator $\widetilde{G}$ is invertible; the components of inverse $\widetilde{G}^{-1}$ are given by $H_{svur} = g_{su}\, g_{vr} - g_{sr}\, g_{vu}$. An arbitrary two-form $\omega \in \Lambda^2(T_p\mathcal{M})$, which we can write as $\omega = \sum_{s<v} \omega_{sv}\, dx^s \wedge dx^v$, gets mapped by $\widetilde{R}$ to $\widetilde{R}\omega = \sum_{u<r} w^{ur} \frac{\partial}{\partial x^u} \wedge \frac{\partial}{\partial x^r} \in \Lambda^2(T_p^*\mathcal{M})$, with $w^{ur} = \sum_{s<v} R^{ursv}\omega_{sv}$. Therefore if we define two-form $\eta \in \Lambda^2(T_p\mathcal{M})$ as

$$\eta = \widetilde{G}^{-1}\widetilde{R}\,\omega = \sum_{y<z} \eta_{yz}\, dx^y \wedge dx^z,$$

its components can be written as

$$\eta_{yz} = \sum_{u<r} H_{yzur}\, w^{ur} = \sum_{u<r}\sum_{s<v} H_{yzur}\, R^{ursv}\, \omega_{sv} = \sum_{s<v} L_{yz}{}^{sv}\, \omega_{sv}\,,$$

with:

$$\begin{aligned} L_{yz}{}^{sv} &\triangleq \sum_{u<r} H_{yzur}\, R^{ursv} \;=\; \frac{1}{2}\, H_{yzur}\, R^{ursv} \;=\; \frac{1}{2}\big(g_{yu}g_{zr} - g_{yr}g_{zu}\big)R^{ursv} \\ &= \frac{1}{2}\big(g_{yu}g_{zr}\, R^{ursv} - g_{yr}g_{zu}\, R^{ursv}\big) \;=\; g_{yu}g_{zr}\, R^{ursv}. \end{aligned}$$

The summation of the ordinary eigenvalues of the linear transformation $-\widetilde{G}^{-1}\widetilde{R}$ is given by the ordinary trace of $-L_{yz}{}^{sv}$ (intended as the summation of its diagonal elements):

$$\begin{aligned} \sum_{i=1}^{m} \sigma_i &= \sum_{y<z} \delta^{yz}_{sv}\big(-L_{yz}{}^{sv}\big) \;=\; -\sum_{y<z} \delta^y_s\, \delta^z_v\, L_{yz}{}^{sv} \;=\; -\frac{1}{2}\, \delta^y_s\, \delta^z_v\, L_{yz}{}^{sv} \\ &= -\frac{1}{2}\, \delta^y_s\, \delta^z_v\, g_{yu}\, g_{zr}\, R^{ursv} \;=\; -\frac{1}{2}\, g_{su}\, g_{vr}\, R^{ursv} \;=\; -\frac{1}{2}\, g_{su}\, g_{vr}\, R^{urvs} \;=\; \frac{1}{2}\, S\,, \end{aligned}$$

which completes the proof. $\square$

CHAPTER 5

Curvature of the Landmarks Manifold

In the present chapter, which is central in this thesis, we apply the formulas that we developed in Chapter 4 to the Riemannian manifold of landmarks $\mathcal{I}$ introduced in Chapter 2. From now on we make the simplifying assumption that the smoothing parameter λ is equal to infinity, i.e. that we are dealing with the exact matching problem. We start with analyzing the simplest but nonetheless very informative case of two landmarks in one dimension, and then move on to the case of three or more landmarks in one dimension. Finally we provide a general formula for sectional curvature for N landmarks in D dimensions. As we did in the previous chapter we adopt Einstein's summation convention. In some computations the rules of such convention are broken, e.g. the summation index may appear three times in the same factor; in such cases we write the summation symbol explicitly. In any case in most of the present chapter we denote summation symbols with Greek letters $(\varphi, \psi, \xi, \eta, \ldots)$. In the next chapter we will study the effect of curvature on the qualitative dynamics of landmarks, for which we derived the differential equations in Chapter 3.

We start by computing the dual Riemannian curvature tensor for N one-dimensional landmarks, and then use the result for calculating sectional curvature for two one-dimensional and three one-dimensional landmarks. This is followed by the computation of the general expression for sectional curvature of N landmarks in D dimensions by means of the formula provided by Theorem 4.4.

1. The dual curvature tensor for one-dimensional landmarks

We shall use Proposition 4.3 on the general form of the dual Riemannian curvature tensor from the previous chapter, which we repeat here for our convenience.

PROPOSITION 5.1. *The Riemannian curvature tensor with raised indices R^{ursv} may be written in function of the cometric tensor and its derivatives as follows:*

$$
\begin{aligned}
&(\mathrm{T}_1) && 2R^{ursv} = - g^{rv,su} + g^{rs,uv} - g^{us,rv} + g^{uv,rs} \\
&(\mathrm{T}_2) && \quad - \frac{1}{2}\Big\{ g^{rs,}_{\ \ \eta}\, g^{\eta\sigma} g^{uv,}_{\ \ \sigma} - g^{rs,}_{\ \ \psi}\, \big(g^{\psi u,v} + g^{\psi v,u}\big) - g^{uv,}_{\ \ \varphi}\, \big(g^{\varphi r,s} + g^{\varphi s,r}\big) \Big\} \\
&(\mathrm{T}_3) && \quad + \frac{1}{2}\Big\{ g^{us,}_{\ \ \eta}\, g^{\eta\sigma} g^{vr,}_{\ \ \sigma} - g^{us,}_{\ \ \psi}\, \big(g^{\psi v,r} + g^{\psi r,v}\big) - g^{vr,}_{\ \ \varphi}\, \big(g^{\varphi u,s} + g^{\varphi s,u}\big) \Big\} \\
&(\mathrm{T}_4) && \quad - \frac{1}{2}\big(g^{\varphi r,s} - g^{\varphi s,r}\big) g_{\varphi\psi} \big(g^{\varphi u,v} - g^{\varphi v,u}\big) \\
&(\mathrm{T}_5) && \quad + \frac{1}{2}\big(g^{\varphi u,s} - g^{\varphi s,u}\big) g_{\varphi\psi} \big(g^{\varphi r,v} - g^{\varphi v,r}\big) \\
&(\mathrm{T}_6) && \quad + \big(g^{\varphi u,r} - g^{\varphi r,u}\big) g_{\varphi\psi} \big(g^{\varphi s,v} - g^{\varphi v,s}\big),
\end{aligned}
$$

where, as usual, $g^{ij,k} \triangleq g^{ij,}_{\ \ \xi}\, g^{\xi k}$ *and* $g^{ij,k\ell} \triangleq g^{ij,}_{\ \ \xi\eta}\, g^{\xi k} g^{\eta \ell}$.

When $D = 1$, expression (2.15) and differential equation (2.16) provided in Chapter 2 respectively take the forms:

$$\gamma(\varrho) = \frac{1}{2^{k-\frac{1}{2}} \pi^{\frac{1}{2}} (k-1)!} \, \frac{1}{a} \Big(\frac{\varrho}{a}\Big)^{k-\frac{1}{2}} K_{k-\frac{1}{2}}\Big(\frac{\varrho}{a}\Big), \tag{5.2}$$

and

$$\gamma'' = \frac{2k-2}{\varrho}\, \gamma' + \frac{1}{a^2}\, \gamma\,, \tag{5.3}$$

where k and a^2 are the parameters of differential operator $L = (\mathrm{id} - a^2\Delta)^k$. We should also note that the most remarkable difference between one-dimesional and D-dimensional landmark manifolds is topological in nature, in that in the former case the manifold is *not connected* since the ordering of landmarks cannot change.

1.1. Generic elements of the dual curvature tensor. In the case of one-dimensional landmarks the cometric tensor $g^{-1}(q)$ is a $N \times N$ matrix whose generic element $g^{ij}(q) = G(q^i, q^j)$ only depends on two of the N variables; the diagonal elements are actually constant. Therefore, for a fixed index k, the matrix of partial derivatives $\big[g^{ij,}_{\ \ k}\big]_{i,j=1}^{N}$ will have only one nonzero row and one nonzero column, namely the k-th ones. Similarly, for fixed k and m, with $k \neq m$, the the matrix of second partial derivatives $\left[g^{ij,}{}_{km}\right]_{i,j=1}^{N}$ will have only two nonzero elements, namely those in positions (k,m) and (m,k). We shall return on this after proving of the following result.

LEMMA 5.2. *Let $\gamma : [0,+\infty) \to \mathbb{R}$ be the function such that $G(x,y) = \gamma(|x-y|)$ is the kernel of admissible space V. Then the first and second partial derivatives of the cometric tensor for N one-dimensional landmarks are respectively given by:*

$$g^{ij,}{}_{k}\,(q) = (\delta^i_k - \delta^j_k)\,\gamma'(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j) \tag{5.4}$$

and

$$g^{ij,}{}_{km}\,(q) = (\delta^i_k - \delta^j_k)\,(\delta^i_m - \delta^j_m)\,\gamma''(\varrho^{ij})\,, \tag{5.5}$$

where $\varrho^{ij} \triangleq |q^i - q^j|$ and sgn *is the sign function.*

PROOF. We have that $g^{ij}(q) = \gamma(\varrho^{ij})$, where $\varrho^{ij} = |q^i - q^j|$. Therefore

$$g^{ij,}{}_{k}\,(q) = \frac{\partial}{\partial q^k}\,g^{ij}(q) = \begin{cases} 0 & \text{for } k \neq i, k \neq j \\ \frac{\partial}{\partial q^i}\gamma(\varrho^{ij}) & \text{for } k = i \\ \frac{\partial}{\partial q^j}\gamma(\varrho^{ij}) & \text{for } k = j \end{cases}.$$

Now choose $k = i$ without loss of generality. By the chain rule we have that

$$\frac{\partial}{\partial q^i}\gamma(\varrho^{ij}) \;=\; \gamma'(\varrho^{ij})\frac{\partial}{\partial q^i}|q^i - q^j| \;=\; \gamma'(\varrho^{ij})\frac{q^i - q^j}{|q^i - q^j|} \;=\; \gamma'(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j)\,.$$

Analogously, $\dfrac{\partial}{\partial q^j}\gamma(\varrho^{ij}) = \gamma'(\varrho^{ij})\,\mathrm{sgn}(q^j - q^i) = -\gamma'(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j)$. Combining the above formulas we finally have that:

$$g^{ij,}{}_{k}\,(q) = \delta^i_k\gamma'(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j) - \delta^j_k\gamma'(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j) = (\delta^i_k - \delta^j_k)\gamma'(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j).$$

Let us now compute the second derivatives:

$$\begin{aligned} g^{ij,}{}_{km}\,(q) &= \frac{\partial}{\partial q^m}\,g^{ij,}{}_{k}\,(q) \;=\; \frac{\partial}{\partial q^m}\,(\delta^i_k - \delta^j_k)\,\gamma'(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j) \\ &= (\delta^i_k - \delta^j_k)\Big\{\Big[\frac{\partial}{\partial q^m}\,\gamma'(\varrho^{ij})\Big]\mathrm{sgn}(q^i - q^j) + \gamma'(\varrho^{ij})\,\frac{\partial}{\partial q^m}\,\mathrm{sgn}(q^i - q^j)\Big\} \\ &= (\delta^i_k - \delta^j_k)\Big[\frac{\partial}{\partial q^m}\,\gamma'(\varrho^{ij})\Big]\mathrm{sgn}(q^i - q^j)\,. \end{aligned} \tag{5.6}$$

The derivative in square brackets above may be rewritten as:

$$\frac{\partial}{\partial q^m}\,\gamma'(\varrho^{ij}) = \begin{cases} 0 & \text{for } m \neq i, m \neq j \\ \frac{\partial}{\partial q^i}\gamma'(\varrho^{ij}) & \text{for } m = i \\ \frac{\partial}{\partial q^j}\gamma'(\varrho^{ij}) & \text{for } m = j \end{cases} . \tag{5.7}$$

Again, without loss of generality choose $m = i$. By the chain rule,

$$\begin{aligned} \frac{\partial}{\partial q^i}\gamma'(\varrho^{ij}) &= \gamma''(\varrho^{ij})\frac{\partial}{\partial q^i}\varrho^{ij} = \gamma''(\varrho^{ij})\frac{\partial}{\partial q^i}|q^i - q^j| \\ &= \gamma''(\varrho^{ij})\frac{q^i - q^j}{|q^i - q^j|} = \gamma''(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j) \end{aligned}$$

and analogously

$$\frac{\partial}{\partial q^j}\gamma'(\varrho^{ij}) = \gamma''(\varrho^{ij})\,\mathrm{sgn}(q^j - q^i) = -\gamma''(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j)\,.$$

Therefore we may write (5.7) in the following compact form:

$$\frac{\partial}{\partial q^m}\,\gamma'(\varrho^{ij}) = (\delta^i_m - \delta^j_m)\,\gamma''(\varrho^{ij})\,\mathrm{sgn}(q^i - q^j)\,.$$

Inserting such expression into (5.6) finally yields:

$$g^{ij,}{}_{km}\,(q) = (\delta^i_k - \delta^j_k)\,(\delta^i_m - \delta^j_m)\,\gamma''(\varrho^{ij})\,\big[\mathrm{sgn}(q^i - q^j)\big]^2,$$

that is:

$$g^{ij,}{}_{km}\,(q) = (\delta^i_k - \delta^j_k)\,(\delta^i_m - \delta^j_m)\,\gamma''(\varrho^{ij})\,,$$

which is precisely what we wanted to prove. $\square$

Remark. Some comments about the expressions provided by the above lemma are in order. Consider matrix $S(q)$, whose generic element is given by g^{ij}. We have that g^{ij} depends only on two out of the N variables $(q^1, \ldots, q^N)$, namely q^i and q^j. Therefore in matrix $\frac{\partial}{\partial q^k}S(q)$ only the k-th row and the k-th column are nonzero (with the notable exception of the element that lies on the diagonal, which must be zero). This fact is reflected in the right-hand side of expression (5.4), where the presence of factor $(\delta^i_k - \delta^j_k)$ implies that $g^{ij,}{}_k$ is nonzero only when $i = k$ or $j = k$ (again, note that when $i = j$ then $\delta^i_k - \delta^j_k = 0$ so that all the elements on the diagonal of matrix $\frac{\partial}{\partial q^k}S(q)$ are zero, for all k). For a fixed value of index k, matrix $\frac{\partial}{\partial q^k}S(q)$ turns

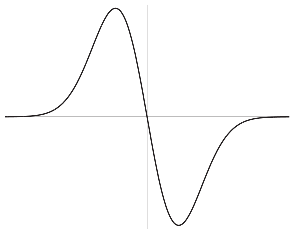


FIGURE 5.1. Typical shape of function $\widetilde{f} : \mathbb{R} \to \mathbb{R}$.

out to be *skew-symmetric*. As far as the second partial derivatives are concerned, since each of the "surviving" elements $g^{ij,}_{\ \ k}$ of matrix $\frac{\partial}{\partial q^k} S(q)$ depends only on the two variables q^i and q^j, the following holds:

- if $m \neq k$ then matrix $\frac{\partial^2}{\partial q^m \partial q^k} S(q)$ will have only *two* nonzero elements, namely those in positions (k, m) and (m, k). This is reflected by the right-hand side of (5.5), which is nonzero only when either $i = k$ and $j = m$, or $i = m$ and $j = k$. If $i = j$ (diagonal elements) the right-hand side of (5.5) is zero.
- if $m = k$ then only the k-th row and the k-th column of matrix $\frac{\partial^2}{\partial (q^k)^2} S(q)$ are nonzero (with the usual exception of the diagonal element). Note that when $m = k$ expression (5.5) takes the form: $g^{ij,}_{\ \ kk}\,(q) = (\delta^i_k - \delta^j_k)^2\, \gamma''(\varrho^{ij})$, which is nonzero when either $i = k$ or $j = k$, once again with the exception $i = j$.

In any case, matrix $\frac{\partial^2}{\partial q^m \partial q^k} S(q)$ is *symmetric* for any choice of indices k and m.

DEFINITION 5.3. Let $\widetilde{f} : \mathbb{R} \to \mathbb{R}$ be the map $x \mapsto \gamma'(|x|)\,\mathrm{sgn}(x)$. Define:

$$f(x, y) \triangleq \widetilde{f}(x - y) = \gamma'(|x - y|)\,\mathrm{sgn}(x - y), \qquad \text{for } \ x, y \in \mathbb{R}.$$

The above function is such that $f(x, y) = -f(y, x)$ and $f(x, x) = 0$. Also, $f^2(x, y) = \left[\gamma'(|x - y|)\right]^2$. The typical shape of function $\widetilde{f}(x)$ is shown in Figure 5.1. With the above definition we may write

$$g^{ij,}_{\ \ k}\,(q) = (\delta^i_k - \delta^j_k)\, f(q^i, q^j)$$

by the first part of Lemma 5.2. The proposition that follows provides the expression for the generic element of the dual Riemannian curvature tensor for one-dimensional landmarks, as a function of γ and its derivatives.

Proposition 5.4. *In the case of one-dimensional landmarks, the six terms of the dual Riemannian curvature tensor listed in Proposition 5.1 take the following form:*

$$
\begin{aligned}
\mathrm{T}_1 = & -\big[\gamma(\varrho^{ru})-\gamma(\varrho^{vu})\big]\big[\gamma(\varrho^{rs})-\gamma(\varrho^{vs})\big]\gamma''(\varrho^{rv}) \\
& +\big[\gamma(\varrho^{ru})-\gamma(\varrho^{su})\big]\big[\gamma(\varrho^{rv})-\gamma(\varrho^{sv})\big]\gamma''(\varrho^{rs}) \\
& -\big[\gamma(\varrho^{ur})-\gamma(\varrho^{sr})\big]\big[\gamma(\varrho^{uv})-\gamma(\varrho^{sv})\big]\gamma''(\varrho^{us}) \\
& +\big[\gamma(\varrho^{ur})-\gamma(\varrho^{vr})\big]\big[\gamma(\varrho^{us})-\gamma(\varrho^{vs})\big]\gamma''(\varrho^{uv})\,, \\
\mathrm{T}_2+\mathrm{T}_3 = & \frac{1}{2}\Bigg(\big[\gamma(\varrho^{ru})-\gamma(\varrho^{uv})-\gamma(\varrho^{sr})+\gamma(\varrho^{sv})\big]f(q^u,q^s)\,f(q^r,q^v) \\
& -\big[\gamma(\varrho^{ru})-\gamma(\varrho^{rv})-\gamma(\varrho^{su})+\gamma(\varrho^{sv})\big]f(q^r,q^s)\,f(q^u,q^v) \\
& +\big[\gamma(\varrho^{uv})-\gamma(\varrho^{rv})\big]\Big\{f(q^s,q^u)\,f(q^u,q^r)-f(q^s,q^r)\big[f(q^s,q^u)+f(q^u,q^r)\big]\Big\} \\
& +\big[\gamma(\varrho^{rs})-\gamma(\varrho^{us})\big]\Big\{f(q^v,q^r)\,f(q^r,q^u)-f(q^v,q^u)\big[f(q^v,q^r)+f(q^r,q^u)\big]\Big\} \\
& -\big[\gamma(\varrho^{su})-\gamma(\varrho^{vu})\big]\Big\{f(q^r,q^s)\,f(q^s,q^v)-f(q^r,q^v)\big[f(q^r,q^s)+f(q^s,q^v)\big]\Big\} \\
& -\big[\gamma(\varrho^{vr})-\gamma(\varrho^{sr})\big]\Big\{f(q^u,q^v)\,f(q^v,q^s)-f(q^u,q^s)\big[f(q^u,q^v)+f(q^v,q^s)\big]\Big\}\Bigg), \\
\mathrm{T}_4 = & -\frac{1}{2}\sum_{\varphi\psi}\Big\{\big[\gamma(\varrho^{rs})-\gamma(\varrho^{s\varphi})\big]f(q^r,q^\varphi)-\big[\gamma(\varrho^{rs})-\gamma(\varrho^{r\varphi})\big]f(q^s,q^\varphi)\Big\} \\
& \cdot\ g_{\varphi\psi}\Big\{f(q^\psi,q^u)\big[\gamma(\varrho^{\psi v})-\gamma(\varrho^{uv})\big]-f(q^\psi,q^v)\big[\gamma(\varrho^{\psi u})-\gamma(\varrho^{uv})\big]\Big\}, \\
\mathrm{T}_5 = & \frac{1}{2}\sum_{\varphi\psi}\Big\{\big[\gamma(\varrho^{us})-\gamma(\varrho^{s\varphi})\big]f(q^u,q^\varphi)-\big[\gamma(\varrho^{us})-\gamma(\varrho^{u\varphi})\big]f(q^s,q^\varphi)\Big\} \\
& \cdot\ g_{\varphi\psi}\Big\{f(q^\psi,q^r)\big[\gamma(\varrho^{\psi v})-\gamma(\varrho^{rv})\big]-f(q^\psi,q^v)\big[\gamma(\varrho^{\psi r})-\gamma(\varrho^{vr})\big]\Big\}, \\
\mathrm{T}_6 = & \sum_{\varphi\psi}\Big\{\big[\gamma(\varrho^{ur})-\gamma(\varrho^{r\varphi})\big]f(q^u,q^\varphi)-\big[\gamma(\varrho^{ru})-\gamma(\varrho^{u\varphi})\big]f(q^r,q^\varphi)\Big\} \\
& \cdot\ g_{\varphi\psi}\Big\{f(q^\psi,q^s)\big[\gamma(\varrho^{\psi v})-\gamma(\varrho^{sv})\big]-f(q^\psi,q^v)\big[\gamma(\varrho^{\psi s})-\gamma(\varrho^{vs})\big]\Big\}.
\end{aligned}
$$

Remark. As we anticipated at the beginning of the chapter we have written the *summation symbols* (such as $\sum_{\varphi\psi}$) explicitly where the rules of Einstein's summation conventions are broken, for example when the summation index appears more than twice in a product. In any case, in the above expressions and in the proof that follows Greek letters are used for summation indices.

Proof of Proposition 5.4. We will compute the six terms one by one. We have that

$$\begin{aligned} g^{rv,su} &= g^{rv,}{}_{\xi\eta}\, g^{\xi s} g^{\eta u} = (\delta^r_\xi - \delta^v_\xi)\,(\delta^r_\eta - \delta^v_\eta)\,\gamma''(\varrho^{rv})\,\gamma(\varrho^{\xi s})\,\gamma(\varrho^{\eta u}) \\ &= \big[\gamma(\varrho^{sr}) - \gamma(\varrho^{sv})\big]\,\big[\gamma(\varrho^{ur}) - \gamma(\varrho^{uv})\big]\,\gamma''(\varrho^{rv}); \end{aligned}$$

by appropriately rearranging the labels one can get analogous expressions for $g^{rs,uv}$, $g^{us,rv}$, and $g^{uv,rs}$; the expression for T_1 follows immediately. As far as terms T_2 and T_3 are concerned, it is the case that:

$$\begin{aligned} g^{rs,}{}_{\eta}\, g^{\eta\sigma} g^{uv,}{}_{\sigma} &= (\delta^r_\eta - \delta^s_\eta)\, f(q^r, q^s)\,\gamma(\varrho^{\eta\sigma})\,(\delta^u_\sigma - \delta^v_\sigma)\, f(q^u, q^v) \\ &= (\delta^r_\eta\delta^u_\sigma - \delta^r_\eta\delta^v_\sigma - \delta^s_\eta\delta^u_\sigma + \delta^s_\eta\delta^v_\sigma)\,\gamma(\varrho^{\eta\sigma})\, f(q^r, q^s)\, f(q^u, q^v) \\ &= \big[\gamma(\varrho^{ru}) - \gamma(\varrho^{rv}) - \gamma(\varrho^{su}) + \gamma(\varrho^{sv})\big]\, f(q^r, q^s)\, f(q^u, q^v). \end{aligned}$$

By appropriately relabeling the indices, we can use the above formula to express the summation of the first terms of T_1 and T_2:

$$\begin{aligned} (5.8)\qquad & -\frac{1}{2}\Big\{g^{rs,}{}_{\eta}\, g^{\eta\sigma} g^{uv,}{}_{\sigma} - g^{us,}{}_{\eta}\, g^{\eta\sigma} g^{vr,}{}_{\sigma}\Big\} \\ &= \frac{1}{2}\Big\{g^{us,}{}_{\eta}\, g^{\eta\sigma} g^{vr,}{}_{\sigma} - g^{rs,}{}_{\eta}\, g^{\eta\sigma} g^{uv,}{}_{\sigma}\Big\} \\ &= \frac{1}{2}\Big\{\big[\gamma(\varrho^{ru}) - \gamma(\varrho^{uv}) - \gamma(\varrho^{sr}) + \gamma(\varrho^{sv})\big]\, f(q^u, q^s)\, f(q^r, q^v) \\ &\qquad - \big[\gamma(\varrho^{ru}) - \gamma(\varrho^{rv}) - \gamma(\varrho^{su}) + \gamma(\varrho^{sv})\big]\, f(q^r, q^s)\, f(q^u, q^v)\Big\}. \end{aligned}$$

On the other hand,

$$\begin{aligned} & g^{rs,}{}_{\psi}\,\big(g^{\psi u,v} + g^{\psi v,u}\big) \\ &= g^{rs,}{}_{\psi}\,\big(g^{\psi u,}{}_{\sigma}\, g^{\sigma v} + g^{\psi v,}{}_{\sigma}\, g^{\sigma u}\big) \\ &= \sum_\psi (\delta^r_\psi - \delta^s_\psi)\, f(q^r, q^s)\,\big[(\delta^\psi_\sigma - \delta^u_\sigma)\, f(q^\psi, q^u)\,\gamma(\varrho^{\sigma v}) + (\delta^\psi_\sigma - \delta^v_\sigma)\, f(q^\psi, q^v)\,\gamma(\varrho^{\sigma u})\big] \\ &= \sum_\psi (\delta^r_\psi - \delta^s_\psi)\, f(q^r, q^s)\,\Big\{\big[\gamma(\varrho^{\psi v}) - \gamma(\varrho^{uv})\big]\, f(q^\psi, q^u) + \big[\gamma(\varrho^{\psi u}) - \gamma(\varrho^{vu})\big]\, f(q^\psi, q^v)\Big\}, \end{aligned}$$

so that we can finally write

$$
\begin{aligned}
&g^{rs,}{}_{\psi}\left(g^{\psi u,v}+g^{\psi v,u}\right)\\
&\quad = f(q^r,q^s)\Big\{\big[\gamma(\varrho^{rv})-\gamma(\varrho^{uv})\big]f(q^r,q^u)+\big[\gamma(\varrho^{ru})-\gamma(\varrho^{vu})\big]f(q^r,q^v)\\
&\qquad -\big[\gamma(\varrho^{sv})-\gamma(\varrho^{uv})\big]f(q^s,q^u)-\big[\gamma(\varrho^{su})-\gamma(\varrho^{vu})\big]f(q^s,q^v)\Big\}.
\end{aligned}
$$

We can use such expression to write the summation of the last two terms of both T_1 and T_2 as follows (modulo the multiplication by $\frac{1}{2}$):

$$
\begin{aligned}
(\star) \triangleq\ & g^{rs,}{}_{\psi}\left(g^{\psi u,v}+g^{\psi v,u}\right)+g^{uv,}{}_{\varphi}\left(g^{\varphi r,s}+g^{\varphi s,r}\right)\\
&\qquad - g^{us,}{}_{\psi}\left(g^{\psi v,r}+g^{\psi r,v}\right)-g^{vr,}{}_{\varphi}\left(g^{\varphi u,s}+g^{\varphi s,u}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\text{(5.9a)} && = + f(q^r,q^s)\Big\{\big[\gamma(\varrho^{rv})-\gamma(\varrho^{uv})\big]f(q^r,q^u)+\big[\underline{\gamma(\varrho^{ru})}-\gamma(\varrho^{vu})\big]f(q^r,q^v)\\
&\text{(5.9b)} && \qquad -\big[\underline{\gamma(\varrho^{sv})}-\gamma(\varrho^{uv})\big]f(q^s,q^u)-\big[\gamma(\varrho^{su})-\gamma(\varrho^{vu})\big]f(q^s,q^v)\Big\}\\
&\text{(5.9c)} && \quad + f(q^u,q^v)\Big\{\big[\gamma(\varrho^{us})-\gamma(\varrho^{rs})\big]f(q^u,q^r)+\big[\underline{\gamma(\varrho^{ru})}-\gamma(\varrho^{sr})\big]f(q^u,q^s)\\
&\text{(5.9d)} && \qquad -\big[\underline{\gamma(\varrho^{sv})}-\gamma(\varrho^{rs})\big]f(q^v,q^r)-\big[\gamma(\varrho^{vr})-\gamma(\varrho^{sr})\big]f(q^v,q^s)\Big\}\\
&\text{(5.9e)} && \quad - f(q^u,q^s)\Big\{\big[\gamma(\varrho^{uv})-\gamma(\varrho^{rv})\big]f(q^u,q^r)+\big[\underline{\gamma(\varrho^{ru})}-\gamma(\varrho^{vr})\big]f(q^u,q^v)\\
&\text{(5.9f)} && \qquad -\big[\underline{\gamma(\varrho^{sv})}-\gamma(\varrho^{rv})\big]f(q^s,q^r)-\big[\gamma(\varrho^{sr})-\gamma(\varrho^{vr})\big]f(q^s,q^v)\Big\}\\
&\text{(5.9g)} && \quad - f(q^r,q^v)\Big\{\big[\gamma(\varrho^{rs})-\gamma(\varrho^{us})\big]f(q^r,q^u)+\big[\underline{\gamma(\varrho^{ru})}-\gamma(\varrho^{su})\big]f(q^r,q^s)\\
&\text{(5.9h)} && \qquad -\big[\underline{\gamma(\varrho^{sv})}-\gamma(\varrho^{us})\big]f(q^v,q^u)-\big[\gamma(\varrho^{vu})-\gamma(\varrho^{su})\big]f(q^v,q^s)\Big\};
\end{aligned}
$$

the underlined terms in the above expression cancel two by two: precisely, (5.9a) cancels with (5.9g), (5.9b) with (5.9f), (5.9c) with (5.9e), and (5.9d) with (5.9h); in some cases the antisymmetry of function f is used. By recombining the surviving terms in the square brackets two by two yields:

$$(\star) \;= + f(q^r, q^s) \Big\{ \big[\gamma(\varrho^{rv}) - \gamma(\varrho^{uv})\big] f(q^r, q^u) - \big[\gamma(\varrho^{su}) - \gamma(\varrho^{vu})\big] f(q^s, q^v) \tag{5.10a}$$
$$- \underline{\big[\gamma(\varrho^{rv}) - \gamma(\varrho^{uv})\big] f(q^s, q^u)} \Big\}$$
$$+ f(q^u, q^v) \Big\{ \big[\gamma(\varrho^{us}) - \gamma(\varrho^{rs})\big] f(q^u, q^r) - \big[\gamma(\varrho^{vr}) - \gamma(\varrho^{sr})\big] f(q^v, q^s) \tag{5.10b}$$
$$- \underline{\big[\gamma(\varrho^{us}) - \gamma(\varrho^{rs})\big] f(q^v, q^r)} \Big\}$$
$$- f(q^u, q^s) \Big\{ \big[\gamma(\varrho^{uv}) - \gamma(\varrho^{rv})\big] f(q^u, q^r) - \big[\gamma(\varrho^{sr}) - \gamma(\varrho^{vr})\big] f(q^s, q^v) \tag{5.10c}$$
$$+ \underline{\big[\gamma(\varrho^{rs}) - \gamma(\varrho^{rv})\big] f(q^u, q^v)} \Big\}$$
$$- f(q^r, q^v) \Big\{ \big[\gamma(\varrho^{rs}) - \gamma(\varrho^{us})\big] f(q^r, q^u) - \big[\gamma(\varrho^{vu}) - \gamma(\varrho^{su})\big] f(q^v, q^s) \tag{5.10d}$$
$$+ \underline{\big[\gamma(\varrho^{uv}) - \gamma(\varrho^{us})\big] f(q^r, q^s)} \Big\},$$

where we have underlined the terms deriving from the recombinations: term (5.10a) derives from combining the surviving terms in the square brackets in (5.9b) and in (5.9f), term (5.10b) from (5.9d) and (5.9h), term (5.10c) from (5.9e) and (5.9c), and term (5.10d) from (5.9g) and (5.9a). Two terms in *each* pair of curly brackets can be factorized as:

$$(\star) \;=$$
$$+ f(q^r, q^s) \Big\{ \big[\gamma(\varrho^{uv}) - \gamma(\varrho^{rv})\big] \big[f(q^u, q^r) + f(q^s, q^u)\big] - \big[\gamma(\varrho^{su}) - \gamma(\varrho^{vu})\big] f(q^s, q^v) \Big\}$$
$$+ f(q^u, q^v) \Big\{ \big[\gamma(\varrho^{rs}) - \gamma(\varrho^{us})\big] \big[f(q^r, q^u) + f(q^v, q^r)\big] - \big[\gamma(\varrho^{vr}) - \gamma(\varrho^{sr})\big] f(q^v, q^s) \Big\}$$
$$- f(q^u, q^s) \Big\{ \big[\gamma(\varrho^{uv}) - \gamma(\varrho^{rv})\big] f(q^u, q^r) + \big[\gamma(\varrho^{rs}) - \gamma(\varrho^{rv})\big] \big[f(q^u, q^v) + f(q^v, q^s)\big] \Big\}$$
$$- f(q^r, q^v) \Big\{ \big[\gamma(\varrho^{rs}) - \gamma(\varrho^{us})\big] f(q^r, q^u) + \big[\gamma(\varrho^{uv}) - \gamma(\varrho^{us})\big] \big[f(q^r, q^s) + f(q^s, q^v)\big] \Big\},$$

so that we can finally write $(\star)$ as the summation of four terms:

$$\begin{aligned}(\star) \;=\; &+\big[\gamma(\varrho^{uv})-\gamma(\varrho^{rv})\big]\;\Big\{f(q^s,q^u)\,f(q^u,q^r)-f(q^s,q^r)\big[f(q^s,q^u)+f(q^u,q^r)\big]\Big\}\\ &+\big[\gamma(\varrho^{rs})-\gamma(\varrho^{us})\big]\;\Big\{f(q^v,q^r)\,f(q^r,q^u)-f(q^v,q^u)\big[f(q^v,q^r)+f(q^r,q^u)\big]\Big\}\\ &-\big[\gamma(\varrho^{us})-\gamma(\varrho^{uv})\big]\;\Big\{f(q^r,q^s)\,f(q^s,q^v)-f(q^r,q^v)\big[f(q^r,q^s)+f(q^s,q^v)\big]\Big\}\\ &-\big[\gamma(\varrho^{rv})-\gamma(\varrho^{rs})\big]\;\Big\{f(q^u,q^v)\,f(q^v,q^s)-f(q^u,q^s)\big[f(q^u,q^v)+f(q^v,q^s)\big]\Big\},\end{aligned}$$

which coincide precisely with the last four terms of $\mathrm{T}_2+\mathrm{T}_3$.

Moving on to terms T_4, T_5, and T_6 we have that, for example,

$$g^{\varphi u,s}=g^{\varphi u,}{}_{\eta}\,g^{\eta s}=(\delta^{\varphi}_{\eta}-\delta^{u}_{\eta})\,f(q^\varphi,q^u)\,\gamma(\varrho^{\eta s})=\big[\gamma(\varrho^{\varphi s})-\gamma(\varrho^{us})\big]f(q^\varphi,q^u), \tag{5.11}$$

so that

$$g^{\varphi u,s}-g^{\varphi s,u}=\Big\{\big[\gamma(\varrho^{\varphi s})-\gamma(\varrho^{us})\big]f(q^\varphi,q^u)-\big[\gamma(\varrho^{\varphi u})-\gamma(\varrho^{su})\big]f(q^\varphi,q^s)\Big\}, \tag{5.12}$$

while

$$g^{\psi r,v}-g^{\psi v,r}=\Big\{\big[\gamma(\varrho^{\psi v})-\gamma(\varrho^{rv})\big]f(q^\psi,q^r)-\big[\gamma(\varrho^{\psi r})-\gamma(\varrho^{vr})\big]f(q^\psi,q^v)\Big\}.$$

The expression for T_5 follows immediately from the antisymmetry of function f, whereas those for T_4 and T_6 are obtained in a completely similar manner. □

1.2. Diagonal elements of the tensor. One may be interested in the general form of the diagonal elements of the dual Riemannian curvature tensor; for example, in the simple case of two one-dimensional landmarks the Riemannian curvature tensor consists of only one element. The following result holds.

PROPOSITION 5.5. *The diagonal elements of the Riemannian curvature tensor are given by* $R^{ursv}\big|_{\substack{s=u\\v=r}} = R^{urur} = \frac{1}{2}\sum_{i=1}^{6} \mathrm{T}_i$, *with:*

$$
\begin{aligned}
\mathrm{T}_1 &= -2\big[\gamma(0)-\gamma(\varrho^{ur})\big]^2\,\gamma''(\varrho^{ur})\,,\\
\mathrm{T}_2+\mathrm{T}_3 &= \big[\gamma(0)-\gamma(\varrho^{ru})\big]\cdot\big[\gamma'(\varrho^{ru})\big]^2,\\
\mathrm{T}_5 &= 0\,,\\
\mathrm{T}_4+\mathrm{T}_6 &= \frac{3}{2}\sum_{\varphi\psi}\Big\{\big[\gamma(\varrho^{ur})-\gamma(\varrho^{r\varphi})\big]f(q^u,q^\varphi)-\big[\gamma(\varrho^{ur})-\gamma(\varrho^{u\varphi})\big]f(q^r,q^\varphi)\Big\}\\
&\quad\cdot g_{\varphi\psi}\Big\{\big[\gamma(\varrho^{ur})-\gamma(\varrho^{\psi r})\big]f(q^u,q^\psi)-\big[\gamma(\varrho^{ur})-\gamma(\varrho^{\psi u})\big]f(q^r,q^\psi)\Big\}.
\end{aligned}
$$

REMARK. Of the three nonzero terms in the expression for $2R^{urur}$ provided by Proposition 5.5 the last two, T_2 and $\mathrm{T}_4+\mathrm{T}_6$, are always positive for any landmark configuration, whereas the sign of T_1 is determined by the sign of function $-\gamma''(\varrho^{ur})$.

PROOF OF PROPOSITION 5.5. Setting $s=u$ and $v=r$ in term T_1 yields:

$$
\begin{aligned}
\mathrm{T}_1\big|_{\substack{s=u\\v=r}} &= -\,0+\big[\gamma(\varrho^{ru})-\gamma(0)\big]\big[\gamma(0)-\gamma(\varrho^{ur})\big]\gamma''(\varrho^{ur})\\
&\quad -\,0+\big[\gamma(\varrho^{ur})-\gamma(0)\big]\big[\gamma(0)-\gamma(\varrho^{ru})\big]\gamma''(\varrho^{ur})\\
&= -\,2\big[\gamma(0)-\gamma(\varrho^{ru})\big]^2\gamma''(\varrho^{ur}).
\end{aligned}
$$

We now note that the first line of the expression for term $\mathrm{T}_2+\mathrm{T}_3$ provided by Proposition 5.4 is equal to zero when $s=u$ and $v=r$; the remaining five lines of the same term become:

$$
\begin{aligned}
(\mathrm{T}_2+\mathrm{T}_3)\big|_{\substack{s=u\\v=r}} &= \frac{1}{2}\Bigg(-\,2\big[\gamma(0)-\gamma(\varrho^{ru})\big]\big[\gamma'(\varrho^{ru})\big]^2\\
&\quad -\big[\gamma(0)-\gamma(\varrho^{ru})\big]\Big\{-\big[\gamma'(\varrho^{ru})\big]^2\Big\}-\big[\gamma(0)-\gamma(\varrho^{ru})\big]\Big\{-\big[\gamma'(\varrho^{ru})\big]^2\Big\}\\
&\quad -\big[\gamma(0)-\gamma(\varrho^{ru})\big]\Big\{-\big[\gamma'(\varrho^{ru})\big]^2\Big\}-\big[\gamma(0)-\gamma(\varrho^{rv})\big]\Big\{-\big[\gamma'(\varrho^{ru})\big]^2\Big\}\Bigg)\\
&= \big[\gamma(0)-\gamma(\varrho^{ru})\big]\big[\gamma'(\varrho^{ru})\big]^2.
\end{aligned}
$$

As far as terms T_4, T_5 and T_6 are concerned it is convenient to turn directly to the expressions provided by Proposition 5.1; we have that:

$$\mathrm{T}_4\Big|_{\substack{s=u\\v=r}} = -\frac{1}{2}\big(g^{\varphi r,u} - g^{\varphi u,r}\big)g_{\varphi\psi}\big(g^{\psi u,r} - g^{\psi r,u}\big) = \frac{1}{2}\big(g^{\psi u,r} - g^{\psi r,u}\big)g_{\varphi\psi}\big(g^{\psi u,r} - g^{\psi r,u}\big),$$

which, with the notation introduced in subsection 3.3 of Chapter 4, we may also write as $\mathrm{T}_4\big|_{\substack{s=u\\v=r}} = \frac{1}{2}\big\|\big(B^{u,r} - B^{r,u}\big)^\sharp\big\|^2$ where the "sharp" operator $\sharp$ raises the indices of a cotangent vector, i.e. $\sharp : T_I^*\mathcal{I} \to T_I\mathcal{I} : X_i dx^i \mapsto (g^{ij}X_j)\partial_i$, and $\|\cdot\|$ is the norm induced by the metric. It is easy to verify that $\mathrm{T}_5\big|_{\substack{s=u\\v=r}} = 0$, while

$$\mathrm{T}_6\Big|_{\substack{s=u\\v=r}} = \big(g^{\varphi u,r} - g^{\varphi r,u}\big)g_{\varphi\psi}\big(g^{\varphi u,r} - g^{\varphi r,u}\big) = \big\|\big(B^{u,r} - B^{r,u}\big)^\sharp\big\|^2,$$

so that $(\mathrm{T}_4 + \mathrm{T}_6)\big|_{\substack{s=u\\v=r}} = \frac{3}{2}\big\|\big(B^{u,r} - B^{r,u}\big)^\sharp\big\|^2$; but by expression (5.12) we have

$$g^{\varphi u,r} - g^{\varphi r,u} = \Big\{\big[\gamma(\varrho^{\varphi r}) - \gamma(\varrho^{ur})\big]f(q^\varphi, q^u) - \big[\gamma(\varrho^{\varphi u}) - \gamma(\varrho^{ru})\big]f(q^\varphi, q^r)\Big\},$$

so that

$$\begin{aligned}(\mathrm{T}_4 + \mathrm{T}_6)\Big|_{\substack{s=u\\v=r}} &= \frac{3}{2}\big\|\big(B^{u,r} - B^{r,u}\big)^\sharp\big\|^2\\ &= \sum_{\varphi\psi}\Big\{\big[\gamma(\varrho^{\varphi r}) - \gamma(\varrho^{ur})\big]f(q^\varphi, q^u) - \big[\gamma(\varrho^{\varphi u}) - \gamma(\varrho^{ru})\big]f(q^\varphi, q^r)\Big\}\\ &\qquad\cdot g_{\varphi\psi}\Big\{\big[\gamma(\varrho^{\psi r}) - \gamma(\varrho^{ur})\big]f(q^\psi, q^u) - \big[\gamma(\varrho^{\psi u}) - \gamma(\varrho^{ru})\big]f(q^\psi, q^r)\Big\},\end{aligned}$$

which completes the proof. □

2. Sectional curvature for two one-dimensional landmarks

In the case of just two one-dimensional landmarks (q^1, q^2) matrices $\mathbf{G}$ and $-\mathbf{R}$ introduced in section 5 of Chapter 4 are just scalars, since $\Lambda^2(\mathbb{R}^2) \simeq \mathbb{R}$. For example, when the smoothing parameter λ is set to be equal to infinity (exact matching):

$$\begin{aligned}\mathbf{G} &= G^{1212} = \big(g^{us}g^{rv} - g^{uv}g^{rs}\big)\Big|_{\substack{u=s=1\\r=v=2}} = g^{11}g^{22} - g^{12}g^{21}\\ &= \gamma^2(0) - \gamma^2(\varrho^{12}) = \big[\gamma(0) + \gamma(\varrho^{12})\big]\big[\gamma(0) - \gamma(\varrho^{12})\big],\end{aligned} \tag{5.13}$$

which is *always positive* since $\gamma(0) > \gamma(\varrho^{12})$. Similarly it is the case that $-\mathbf{R} = -R^{1212} = R^{1221}$, whose expression can be computed by means of Proposition 5.5; more precisely the following result holds.

Proposition 5.6. *In the case of* two *one-dimensional landmarks term* $\mathrm{T}_4 + \mathrm{T}_6$ *of* $2R^{1212}$ *provided by Proposition 5.5 takes the following form, for a generic value of smoothing parameter* λ*:*

$$\mathrm{T}_4 + \mathrm{T}_6 = 3\,\frac{\big[\gamma(0) - \gamma(\varrho^{12})\big]^2}{\gamma(0) + \gamma(\varrho^{12})} \cdot \big[\gamma'(\varrho^{12})\big]^2,$$

so that the only element of matrix $-\mathbf{R} = R^{1221}$ *can be expressed as*

$$\begin{aligned} R^{1221} \quad &= \quad -R^{1212} \quad = \quad -\frac{1}{2}\big(\textstyle\sum_{i=1}^{6} \mathrm{T}_i\big)\big|_{\substack{s=u=1\\v=r=2}} \\ &= \quad \big[\gamma(0) - \gamma(\varrho^{12})\big]^2 \gamma''(\varrho^{12}) - \big[\gamma(0) - \gamma(\varrho^{12})\big] \frac{2\gamma(0) - \gamma(\varrho^{12})}{\gamma(0) + \gamma(\varrho^{12})} \big[\gamma'(\varrho^{12})\big]^2. \end{aligned}$$

Remark. Of the three terms in the expression for R^{1221} provided by the above proposition the last two are always negative, whereas the sign of the first one is determined by the sign of function $\gamma''(\varrho^{12})$.

Proof of Proposition 5.6. The first two terms of the expression for $R^{1221} = -R^{1212}$ derive directly form terms T_1 and $\mathrm{T}_2 + \mathrm{T}_3$ (multiplied by $\frac{1}{2}$) for R^{urur} provided by Proposition 5.5. As far as the last term is concerned, we note that in the case of two landmarks

$$g^{-1} = \begin{bmatrix} \gamma(0) & \gamma(\varrho^{12}) \\ \gamma(\varrho^{12}) & \gamma(0) \end{bmatrix}$$

$$\Longrightarrow \quad g = \frac{1}{\gamma^2(0) - \gamma^2(\varrho^{12})} \begin{bmatrix} \gamma(0) & -\gamma(\varrho^{12}) \\ -\gamma(\varrho^{12}) & \gamma(0) \end{bmatrix}. \tag{5.14}$$

By the proof of Proposition 5.5 we have that term $\mathrm{T}_4 + \mathrm{T}_6$ can be written as

$$\begin{aligned} (\mathrm{T}_4 + \mathrm{T}_6)\big|_{\substack{s=u=1\\v=r=2}} &= \frac{3}{2}\big\|\big(B^{1,2} - B^{2,1}\big)^\sharp\big\|^2 \\ &= \frac{3}{2}\big(g^{\varphi 1,2} - g^{\varphi 2,1}\big) g_{\varphi\psi} \big(g^{\psi 1,2} - g^{\psi 2,1}\big), \end{aligned}$$

which by the Einstein summation convention expands out to

$$\begin{aligned} (\mathrm{T}_4 + \mathrm{T}_6)\big|_{\substack{s=u=1\\v=r=2}} =& \frac{3}{2}\Big\{ \big(g^{11,2} - g^{12,1}\big) g_{11} \big(g^{11,2} - g^{12,1}\big) + \big(g^{11,2} - g^{12,1}\big) g_{12} \big(g^{21,2} - g^{22,1}\big) \\ & + \big(g^{21,2} - g^{22,1}\big) g_{21} \big(g^{11,2} - g^{12,1}\big) + \big(g^{21,2} - g^{22,1}\big) g_{22} \big(g^{21,2} - g^{22,1}\big) \Big\}; \end{aligned}$$

but $g^{11,2} = 0$, $g^{22,1} = 0$, $g_{11} = g_{22}$ and $g_{12} = g_{21}$, so the above expression reduces to:

$$(\mathrm{T}_4 + \mathrm{T}_6)\big|_{\substack{s=u=1\\v=r=2}} = \frac{3}{2}\Big\{ g_{11}\Big[\big(g^{12,1}\big)^2 + \big(g^{21,2}\big)^2\Big] - 2g_{12}\big(g^{12,1}g^{21,2}\big)\Big\}. \tag{5.15}$$

From equation (5.11) we have that

$$\begin{aligned} g^{12,1} &= \big[\gamma(0) - \gamma(\varrho^{12})\big] f(q^1, q^2) \\ \text{and} \quad g^{21,2} &= \big[\gamma(0) - \gamma(\varrho^{12})\big] f(q^2, q^1) = -g^{12,1}, \end{aligned}$$

so that:

$$\begin{aligned} \big(g^{12,1}\big)^2 &= \big(g^{12,1}\big)^2 = \big[\gamma(0) - \gamma(\varrho^{12})\big]^2 \big[\gamma'(\varrho^{12})\big]^2 \\ \text{and} \quad g^{12,1}g^{21,2} &= -\big[\gamma(0) - \gamma(\varrho^{12})\big]\big[\gamma'(\varrho^{12})\big]^2 = -\big(g^{12,1}\big)^2; \end{aligned}$$

on the other hand, by expression (5.14) we have

$$g_{11} = \frac{\gamma(0)}{\gamma^2(0) - \gamma^2(\varrho^{12})}, \qquad g_{12} = -\frac{\gamma(\varrho^{12})}{\gamma^2(0) - \gamma^2(\varrho^{12})}.$$

Whence we may rewrite (5.15) as:

$$\begin{aligned} (\mathrm{T}_4 + \mathrm{T}_6)\big|_{\substack{s=u=1\\v=r=2}} &= 3\,\big(g^{12,1}\big)^2\big\{g_{11} + g_{12}\big\} \\ &= 3\big[\gamma(0) - \gamma(\varrho^{12})\big]^2\big[\gamma'(\varrho^{12})\big]^2 \frac{\gamma(0) - \gamma(\varrho^{12})}{\gamma^2(0) - \gamma^2(\varrho^{12})} \\ &= 3\,\frac{\big[\gamma(0) - \gamma(\varrho^{12})\big]^2}{\gamma(0) + \gamma(\varrho^{12})}\,\big[\gamma'(\varrho^{12})\big]^2; \end{aligned}$$

Summing the above to the other terms (and multiplying by $-\frac{1}{2}$) yields

$$\begin{aligned} R^{1221} &= \big[\gamma(0) - \gamma(\varrho^{12})\big]^2\gamma''(\varrho^{12}) \\ &\quad - \frac{1}{2}\big[\gamma(0) - \gamma(\varrho^{12})\big]\big[\gamma'(\varrho^{12})\big]^2 - \frac{3}{2}\,\frac{\big[\gamma(0) - \gamma(\varrho^{12})\big]^2}{\gamma(0) + \gamma(\varrho^{12})}\big[\gamma'(\varrho^{12})\big]^2 \\ &= \big[\gamma(0) - \gamma(\varrho^{12})\big]^2\gamma''(\varrho^{12}) \\ &\quad - \frac{1}{2}\big[\gamma(0) - \gamma(\varrho^{12})\big]\frac{\big[\gamma(0) + \gamma(\varrho^{12})\big] + 3\big[\gamma(0) - \gamma(\varrho^{12})\big]}{\gamma(0) + \gamma(\varrho^{12})}\big[\gamma'(\varrho^{12})\big]^2 \\ &= \big[\gamma(\varrho^{12}) - \gamma(0)\big]^2\gamma''(\varrho^{12}) - \big[\gamma(0) - \gamma(\varrho^{12})\big]\frac{2\gamma(0) - \gamma(\varrho^{12})}{\gamma(0) + \gamma(\varrho^{12})}\big[\gamma'(\varrho^{12})\big]^2, \end{aligned}$$

which completes the proof. □

The problem of computing the generalized eigenvalues of the pair $(-\mathbf{R}, \mathbf{G})$ is at this point trivial: equation $-\widetilde{R}\omega = \sigma\,\widetilde{G}\omega$ (with $\omega = \omega_{12}\,dq^1 \wedge dx^2$) can simply be written as $-R^{1212}\omega_{12} = \sigma\,G^{1212}\omega_{12}$, whose only solution is obviously $\sigma = \frac{R^{1221}}{G^{1212}}$, which is the sectional curvature of the two-dimentional landmarks manifold $\mathcal{I}$ at the point of coordinates (q^1, q^2). In fact such result can also be obtained directly by applying formula (4.21) of the previous chapter:

$$K(X,Y) = \frac{-\sum_{u<r}\omega_{ur}\sum_{s<v}R^{ursv}\,\omega_{sv}}{\sum_{\bar{u}<\bar{r}}\omega_{\bar{u}\bar{r}}\sum_{\bar{s}<\bar{v}}G^{\bar{u}\bar{r}\bar{s}\bar{v}}\,\omega_{\bar{s}\bar{v}}} = \frac{-\omega_{12}\,R^{1212}\,\omega_{12}}{\omega_{12}\,G^{1212}\,\omega_{12}} = \frac{R^{1221}}{G^{1212}}\,, \tag{5.16}$$

which is *only* a function of $\varrho^{12} = |q^1 - q^2|$; we will denote it with $\kappa(\varrho^{12})$. By (5.13) and Proposition 5.5 we can finally compute the expression for sectional curvature for two one-dimensional landmarks.

Corollary 5.7. *Sectional curvature for two landmarks in one dimension* (5.16) *has the following expression for $\lambda = \infty$ (exact matching):*

$$\boxed{\kappa(\varrho^{12}) = \frac{\gamma(0)-\gamma(\varrho^{12})}{\gamma(0)+\gamma(\varrho^{12})}\,\gamma''(\varrho^{12}) - \frac{2\gamma(0)-\gamma(\varrho^{12})}{\big[\gamma(0)+\gamma(\varrho^{12})\big]^2}\,\big[\gamma'(\varrho^{12})\big]^2,} \tag{5.17}$$

where we have merged the two terms with factor $\big[\gamma'(\varrho^{12})\big]^2$ into one.

Proof. It suffices to divide R^{1221}, provided by Proposition 5.6, by G^{1212}, given by equation (5.13). $\square$

We implemented the above equations in the case of a *Gaussian* kernel,

$$\gamma(\varrho) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\exp\Big\{-\frac{1}{2}\frac{\varrho^2}{\sigma^2}\Big\}, \tag{5.18}$$

with $\sigma^2 = 1$. The shape of sectional curvature $\kappa(\varrho^{12})$, provided by equation (5.17), is plotted in Figure 5.2. It turns out that κ has a minimum, a zero (other than the one at $\varrho^{12} = 0$) and a maximum at points:

$$\varrho_{\mathrm{m}} = 0.953, \qquad \varrho_{\mathrm{z}} = 1.534, \qquad \varrho_{\mathrm{M}} = 2.198,$$

$$\kappa(\varrho_{\mathrm{m}}) = -0.1594, \qquad \kappa(\varrho_{\mathrm{z}}) = 0, \qquad \kappa(\varrho_{\mathrm{M}}) = 0.1786,$$

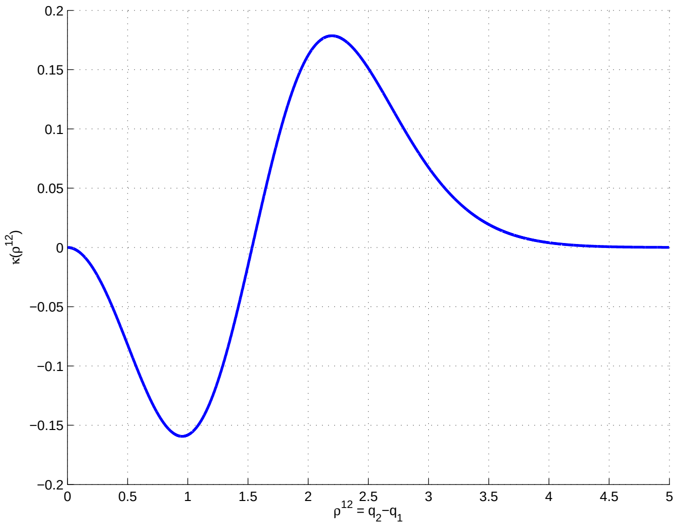


FIGURE 5.2. Sectional curvature $\kappa(\varrho^{12})$ for the Gaussian kernel.

respectively. So we see that when the two landmarks are close to each other curvature is negative, it is positive for higher values of the mutual distance of the landmarks, and finally converges to zero *from above* as $\varrho^{12} \to \infty$. Figure 5.2 refers to the Gaussian kernel (5.18) but if fact it turns out that graph of sectional curvature $\kappa(\varrho)$ is qualitatively similar to the one above for a wide class of Kernels. The following proposition is an adaptation of a result by François-Xavier Vialard [**44**].

PROPOSITION 5.8. *Consider function $\kappa(\varrho)$, $\varrho > 0$, given by (5.17), i.e. the sectional curvature for two one-dimensional landmarks in the case of exact matching. If $\gamma'(0) = 0$ (which is verified when the kernel is bell-shaped) then:*

$$\kappa(\varrho) = -\frac{1}{2}\frac{\left[\gamma''(0)\right]^2}{\gamma(0)}\varrho^2 + o(\varrho^2) \qquad \textit{as } \varrho \to 0, \tag{5.19}$$

so that $\kappa(\varrho)$ is negative in a neighborhood of zero. Under the following condition on the behavior of function γ at infinity:

$$\left[\gamma'(\varrho)\right]^2 = o\big(\gamma''(\varrho)\big) \qquad \textit{as } \varrho \to \infty, \tag{5.20}$$

it is the case that

$$\kappa(\varrho) = \gamma''(\varrho)\big(1 + o(1)\big) \qquad \textit{as } \varrho \to \infty,$$

therefore, since $\gamma''(\varrho) > 0$ for large ϱ, curvature κ is "convex at infinity", i.e. it converges to zero from above as $\varrho \to \infty$.

PROOF. First of all note that function (5.17) can be written as follows:

$$\kappa(\varrho) = \frac{\widetilde{\kappa}(\varrho)}{\gamma(0) + \gamma(\varrho)},$$

where

$$\widetilde{\kappa}(\varrho) \triangleq \big[\gamma(0) - \gamma(\varrho)\big]\gamma''(\varrho) - \frac{2\gamma(0) - \gamma(\varrho)}{\gamma(0) + \gamma(\varrho)}\big[\gamma'(\varrho^{12})\big]^2.$$

Since $\big[\gamma(0) + \gamma(\varrho)\big]^{-1}$ is strictly greater than zero for all $\varrho > 0$ and it converges to $\big[2\gamma(0)\big]^{-1}$ as $\varrho \to 0^+$, in order to prove (5.19) it is sufficient to study the properties of $\widetilde{\kappa}(\varrho)$ in a neighborhood of zero. First note that since $\gamma'(0) = 0$ the Taylor expansion of γ near the origin is given by:

$$\gamma(\varrho) = \gamma(0) + \frac{1}{2}\gamma''(0)\varrho^2 + o(\varrho^2);$$

therefore we may write the first term of $\widetilde{\kappa}(\varrho)$ as:

$$\begin{aligned} \big[\gamma(0) - \gamma(\varrho)\big]\gamma''(\varrho) =& \Big[-\frac{1}{2}\gamma''(0)\varrho^2 + o(\varrho^2)\Big]\big[\gamma''(0) + o(1)\big] \\ =& -\frac{1}{2}\big[\gamma''(0)\big]^2\varrho^2 + o(\varrho^2) \end{aligned} \tag{5.21}$$

in an neighborhood of zero. As far as the second term of $\kappa(\varrho)$ is concerned, we first note that $\gamma'(\varrho) = \gamma''(0)\varrho + o(\varrho)$, whence $\big[\gamma'(\varrho)\big]^2 = \big[\gamma''(0)\big]^2\varrho^2 + o(\varrho^2)$; consequently,

$$\begin{aligned} -\big[2\gamma(0) - \gamma(\varrho)\big]\big[\gamma'(\varrho)\big]^2 &= -\Big[\gamma(0) - \frac{1}{2}\gamma''(0)\varrho^2 + o(\varrho^2)\Big]\Big\{\big[\gamma''(0)\big]^2\varrho^2 + o(\varrho^2)\Big\} \\ &= -\gamma(0)\big[\gamma''(0)\big]^2\varrho^2 + o(\varrho^2), \end{aligned}$$

whence

$$\begin{aligned} -\frac{2\gamma(0) - \gamma(\varrho)}{\gamma(0) + \gamma(\varrho)}\big[\gamma'(\varrho)\big]^2 &= \Big[\frac{1}{2\gamma(0)} + o(1)\Big]\Big\{-\gamma(0)\big[\gamma''(0)\big]^2\varrho^2 + o(\varrho^2)\Big\} \\ &= -\frac{1}{2}\big[\gamma''(0)\big]^2\varrho^2 + o(\varrho^2), \end{aligned} \tag{5.22}$$

as $\varrho \to 0$. Summing (5.21) and (5.22) yields

$$\widetilde{\kappa}(\varrho) = -\big[\gamma''(0)\big]^2\varrho^2 + o(\varrho^2),$$

so that we finally have

$$\kappa(\varrho) = \frac{1}{\gamma(0)+\gamma(\varrho)}\,\widetilde{\kappa}(\varrho) = \Big[\frac{1}{2\gamma(0)} + o(1)\Big]\Big\{-\big[\gamma''(0)\big]^2\varrho^2 + o(\varrho^2)\Big\}$$
$$= -\frac{1}{2}\frac{\big[\gamma''(0)\big]^2}{\gamma(0)}\varrho^2 + o(\varrho^2).$$

which proves the first part of the proposition.

Now assume now that (5.20) holds. We may rewrite (5.17) in a neighborhood of infinity as:

$$\frac{\kappa(\varrho^{12})}{\gamma''(\varrho^{12})} = \underbrace{\frac{\gamma(0)-\gamma(\varrho^{12})}{\gamma(0)+\gamma(\varrho^{12})}}_{1+o(1)} - \underbrace{\frac{2\gamma(0)-\gamma(\varrho^{12})}{\big[\gamma(0)+\gamma(\varrho^{12})\big]^2}}_{\frac{2}{\gamma(0)}+o(1)}\underbrace{\frac{\big[\gamma'(\varrho^{12})\big]^2}{\gamma''(\varrho^{12})}}_{o(1)} = 1 + o(1)$$

as $\varrho \to \infty$, which concludes the proof since $\gamma''(\varrho) > 0$ for large values of ϱ. □

We will now verify the validity of condition (5.20) for three families of kernels.

Example 1 (Gaussian kernels). As we mentioned in Chapter (2) the Gaussian kernel $\gamma(\varrho) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\big\{-\frac{1}{2}\frac{\varrho^2}{\sigma^2}\big\}$ is such that equations (2.19) hold, i.e.:

$$\gamma'(\varrho) = -\frac{\varrho}{\sigma^2}\,\gamma(\varrho) \qquad \text{and} \qquad \gamma''(\varrho) = \frac{1}{\sigma^2}\Big(\frac{\varrho}{\sigma^2} - 1\Big),$$

therefore

$$\frac{[\gamma'(\varrho)]^2}{\gamma''(\varrho)} = \gamma(\varrho)\,\frac{\varrho^2}{\varrho^2 - \sigma^2}$$

which converges to zero as $\varrho \to \infty$; whence (5.21) is satisfied.

Example 2 (Sobolev kernels). With the notation introduced in Appendix B, i.e.:

$$\eta_{k,D} \triangleq \frac{1}{2^{k+\frac{D}{2}-1}\pi^{\frac{D}{2}}\,\Gamma(k)}\,\frac{1}{a^{k+\frac{D}{2}}}\,,$$

function (2.15) may be rewritten as:

$$\gamma(\varrho) = \eta_{k,D}\,\varrho^{k-\frac{D}{2}}K_{k-\frac{D}{2}}\Big(\frac{\varrho}{a}\Big),$$

while its derivative is provided by Corollary B.2:

$$\gamma'(\varrho) = -\eta_{k,D}\,\frac{1}{a}\,\varrho^{k-\frac{D}{2}}K_{k-\frac{D}{2}-1}\Big(\frac{\varrho}{a}\Big).$$

We will use the following property [**1**, §9.7.2] of modified Bessel functions in a neighborhood of infinity:

$$K_\nu(z) \sim \sqrt{\frac{\pi}{2z}}e^{-z}\Big\{1+\frac{\mu-1}{8z}+\frac{(\mu-1)(\mu-9)}{2!(8z)^2}+\frac{(\mu-1)(\mu-9)(\mu-25)}{3!(8z)^3}+\dots\Big\} \tag{5.23}$$

for $|\arg z| < \frac{3}{2}\pi$, where $\mu(\nu) = 4\nu^2$. In our case $D = 1$ and we shall also assume for simplicity $a = 1$; furthermore $\nu \triangleq k - \frac{D}{2} = k - \frac{1}{2}$. Function γ satisfies differential equation (5.3) with $a = 1$, which we rewrite for convenience:

$$\gamma'' = \frac{2k-2}{\varrho}\,\gamma' - \gamma;$$

hence the second derivative of γ can be written, in a neighborhood of infinity, as

$$\begin{aligned}
\gamma'' &= (2k-2)\varrho^{-1}\gamma' - \gamma \\
&= -\eta_{k,1}(2k-2)\varrho^{\nu-1}K_{\nu-1}(\varrho) + \eta_{k,1}\,\varrho^\nu K_\nu(\varrho) \\
&= -\eta_{k,1}(2k-2)\varrho^{\nu-1}\sqrt{\frac{\pi}{2\varrho}}\,e^{-\varrho}\Big\{1+\frac{(2k-3)^2-1}{8\varrho}+o\Big(\frac{1}{\varrho}\Big)\Big\} \\
&\qquad + \eta_{k,1}\,\varrho^\nu\sqrt{\frac{\pi}{2\varrho}}\,e^{-\varrho}\Big\{1+\frac{(2k-1)^2-1}{8\varrho}+o\Big(\frac{1}{\varrho}\Big)\Big\} \\
&= \eta_{k,1}\,\varrho^{\nu-\frac{1}{2}}\sqrt{\frac{\pi}{2}}\,e^{-\varrho}\Big\{1+\Big[\frac{(2k-3)^2-1}{8}-(2k-2)\Big]\frac{1}{\varrho}+o\Big(\frac{1}{\varrho}\Big)\Big\},
\end{aligned} \tag{5.24}$$

since $\mu(\nu) = 4\big(k-\frac{1}{2}\big)^2 = (2k-1)^2$ and $\mu(\nu-1) = 4(\nu-1)^2 = 4\big(k-\frac{3}{2}\big)^2 = (2k-3)^2$. On the other hand,

$$\begin{aligned}
\big[\gamma'(\varrho)\big]^2 &= \eta_{k,1}^2\,\varrho^{2\nu}\,K_{\nu-1}^2(\varrho) \\
&\sim \eta_{k,1}^2\,\varrho^{2\nu-1}\,\frac{\pi}{2}\,e^{-2\varrho}\Big\{1+2\,\frac{(2k-3)^2-1}{8}\,\frac{1}{\varrho}+o\Big(\frac{1}{\varrho}\Big)\Big\}.
\end{aligned}$$

Comparing (5.24) with the above expression one concludes that

$$\frac{\big[\gamma'(\varrho)\big]^2}{\gamma''(\varrho)} \to 0$$

as $\varrho \to \infty$; whence $\big[\gamma'(\varrho)\big]^2 = o\big(\gamma''(\varrho)\big)$, which is exactly (5.20).

Example 3 (Cauchy kernels). Again, as we mentioned in Chapter (2) the Cauchy-type kernel $\gamma(\varrho) = \frac{1}{1+a^2\varrho^2}$ is such that equations (2.20) hold, i.e.:

$$\gamma'(\varrho) = -2a^2\varrho\gamma^2(\varrho) \qquad \text{and} \qquad \gamma''(\varrho) = 8a^4\varrho^2\gamma^3(\varrho) - 2a^2\gamma^2(\varrho),$$

therefore $[\gamma'(\varrho)]^2 = 4a^4\varrho^2\gamma^4(\varrho)$ and

$$\frac{[\gamma'(\varrho)]^2}{\gamma''(\varrho)} = \frac{4a^2\big[a^2\varrho^2\gamma(\varrho)\big]\gamma^3(\varrho)}{8a^2\big[a^2\varrho^2\gamma(\varrho)\big]\gamma^2(\varrho) - 2a^2\gamma^2(\varrho)} = \gamma(\varrho)\,\frac{2\big[a^2\varrho^2\gamma(\varrho)\big]}{4\big[a^2\varrho^2\gamma(\varrho)\big] - 1}\,,$$

which converges to zero as $\varrho \to \infty$ since $a^2\varrho^2\gamma(\varrho) \to 1$; whence (5.20) is satisfied.

3. Sectional curvature for three one-dimensional landmarks

In this section we analyze sectional curvature for different configurations of three landmarks on the real line. We will use the techniques developed in section 5 of the previous chapter, motivated by the fact that in this case the shape manifold $\mathcal{I}$ has dimension 3 and that $\Lambda^2(T_I\mathcal{I}) \simeq \mathbb{R}^3$ is indeed decomposable. That is, for any $\omega \in \Lambda^2(T_I\mathcal{I})$ there exists a pair $(X, Y) \in T_I\mathcal{I} \times T_I\mathcal{I}$ such that $\omega = X^\flat \wedge Y^\flat$; so the maximum and minimum generalized eigenvalues of the pair of linear maps $(-\widetilde{R}, \widetilde{G})$ indeed provide the maximum and minimum sectional curvatures over possible choice of tangent 2-planes at the point I (landmark shape) under consideration.

Assuming the three one-dimensional landmarks are such that $q^1 < q^2 < q^3$ at all times, let $\varrho^{12} = |q^1 - q^2|$ and $\varrho^{23} = |q^2 - q^3|$. We have computed numerically the dual Riemannian curvature tensor R^{ursv} and tensor G^{ursv} using the formulas provided by Proposition 5.4 and equation (5.13), respectively, for $N = 3$ using the *Gaussian* kernel (5.18) with unit variance: maximum, median and minimum generalized eigenvalues were also computed numerically for $(\varrho^{12}, \varrho^{23}) \in (0, 5) \times (0, 5)$.

Figure 5.3 represents the maximum generalized eigenvalue: it is *positive* for any choice of ϱ^{12} and ϱ^{23}; for $(\varrho^{12}, \varrho^{23}) \to 0$ the maximum generalized eigenvalue converges to some finite, positive limit. Figure 5.4 represents the *minimum* generalized eigenvalue, for different values of ϱ^{12} and ϱ^{23}: it is *negative* for any choice of such parameters; note that for $(\varrho^{12}, \varrho^{23}) \to 0$ (i.e. when the three landmarks are close to each other) the minimum generalized eigenvalue converges to some finite, negative

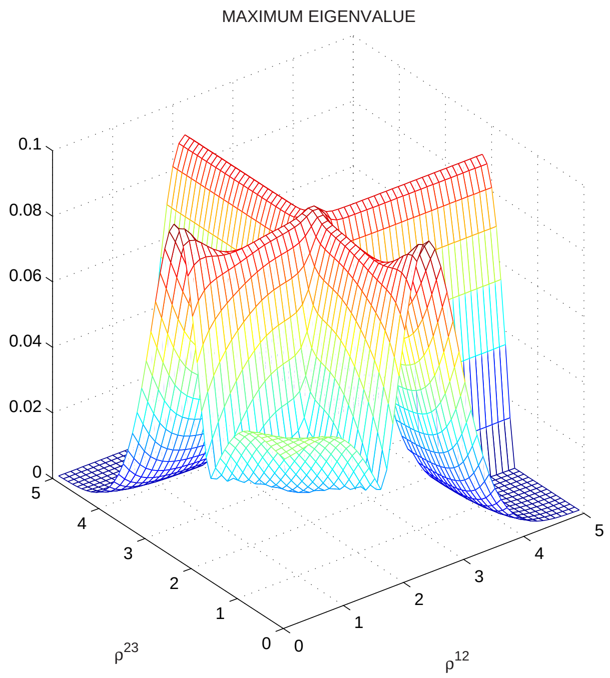


Figure 5.3. Maximum generalized eigenvalue (i.e. maximum sectional curvature) for three landmarks in one dimension, as a function of distances ϱ^{12} and ϱ^{23}.

limit. Figure 5.5 represents the median generalized eigenvalue (out of three): it has positive and negative values for different choices of ϱ^{12} and ϱ^{23}; for $(\varrho^{12}, \varrho^{23}) \to 0$ the median generalized eigenvalue converges to *the same* finite, positive limit to which the maximum eigenvalue converges.

Finally, Figure 5.6 shows the *generalized trace*, i.e. the summation of the three generalized generalized eigenvalues: we proved in Proposition 4.10 that this number is actually equal to ($\frac{1}{2}$ times) the *scalar curvature* of the manifold at a point; again, it has positive and negative values for different choices of ϱ^{12} and ϱ^{23}. Note, in particular, that for small values of distances ϱ^{12} and ϱ^{23} it is negative and it has a positive maximum when $\varrho^{12} = \varrho^{23} \simeq 2.2$. Also, note that when one of the two landmarks is "far away"—say, when ϱ^{23} is very large, the profile of the graph *as a function of* ϱ^{12} *only* is the same one that we got in the case of two one-dimensional landmarks (see Figure 5.2): curvature converges to zero as $\varrho^{12} \to 0$, then it has a

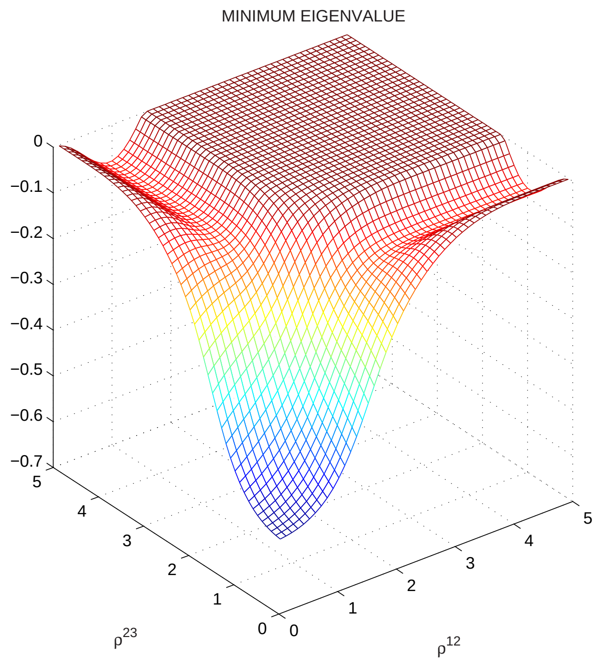


FIGURE 5.4. Minimum generalized eigenvalue (i.e. minimum sectional curvature) for three landmarks in one dimension, as a function of distances ϱ^{12} and ϱ^{23}.

negative bump, followed by a positive bump, and finally flattens again to zero (from above) as ϱ^{12} diverges.

Figure 5.7 compares the generalized eigenvalues along the line $\varrho^{12} = \varrho^{23}$ (symmetric landmark configurations); as anticipated above, for $\varrho \to 0$ along such line the maximum and median eigenvalues converge to the same positive limit. Also, we may note that the maximum and minimum eigenvalues are always, respectively, positive and negative, the median one has both positive and negative values. Also, the trace converges to zero from above as $\varrho \to \infty$ as it was the case for two one-dimensional landmarks—see Figure 5.2.

Table 5.1 shows possible choices of tangent vectors X and Y that achieve maximum and minimum curvature along the line $\varrho^{12} = \varrho^{23}$; they were computed using the techniques described at the end of the previous chapter. It is interesting to note that along the line $\varrho^{12} = \varrho^{23}$ the tangent plane of maximum curvature is the same i.e. it is spanned by the same vectors. On the other hand, the tangent

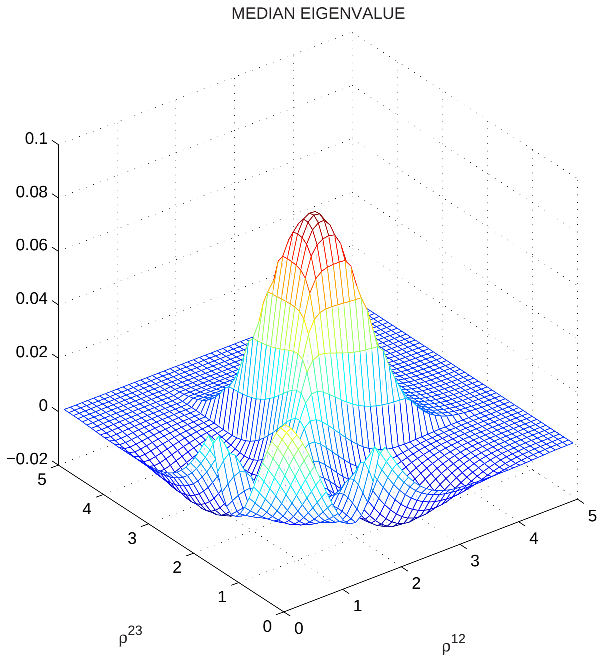


FIGURE 5.5. Median generalized eigenvalue for three landmarks in one dimension, as a function of distances ϱ^{12} and ϱ^{23}.

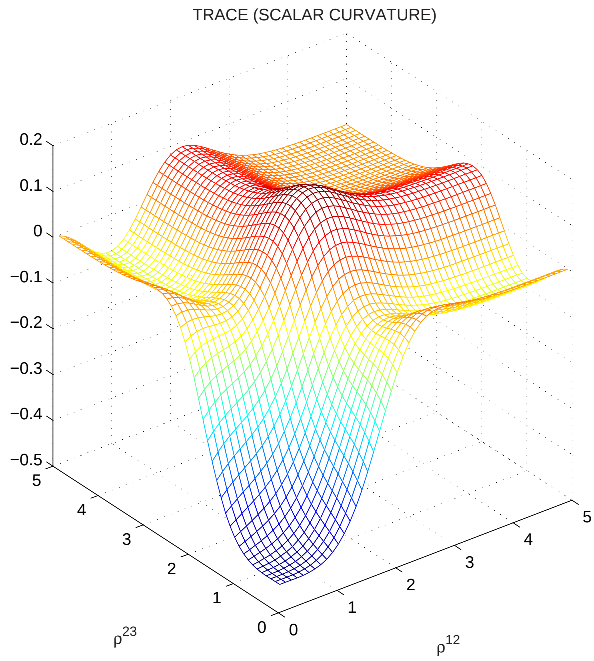


FIGURE 5.6. Trace (i.e. $\frac{1}{2}$ times scalar curvature) for three landmarks in one dimension, as a function of distances ϱ^{12} and ϱ^{23}.

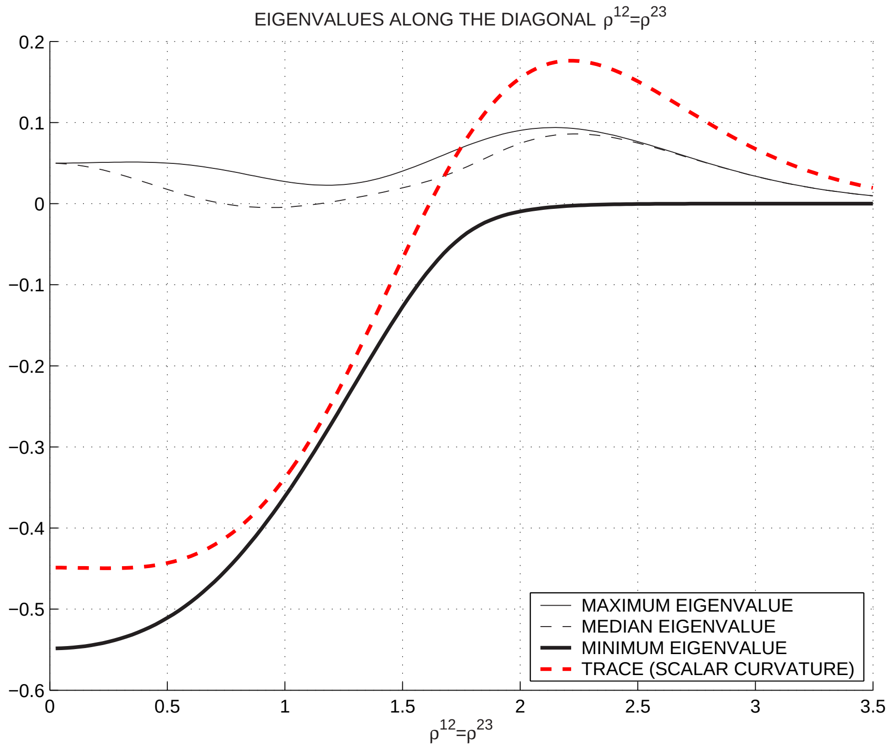


FIGURE 5.7. Generalized eigenvalues of $(-\widetilde{R}, \widetilde{G})$ for three landmarks in one dimension of the quadratic form along the line $\varrho^{12} = \varrho^{23}$.

plane of minimum curvature changes; one of the two tangent vectors can actually be fixed, $Y = (-1, 0, 1)$ while the other tangent vector can always be chosen to be of the form $X = \big(X_1(\varrho), 1, 0\big)$, where X_1 is a decreasing function of $\varrho \triangleq \varrho^{12} = \varrho^{23}$.

Finally, Table 5.2 shows possible choices of tangent vectors that achieve maximum and minimum curvature along a "ridge" of the graph for scalar curvature (the summation of the generalized eigenvalues, times $\frac{1}{2}$): the locations are shown in Figure 5.8 with different symbols, that also appear in Table 5.2. It is certainly interesting to follow the "evolution" along the ridge of the tangent plane that corresponds, for example, to the maximum eigenvalue (sectional curvature). In the position denoted by a square, that corresponds to the central local maximum of scalar curvature ($\varrho^{12} = \varrho^{23} = 2.225$), the two tangent vectors that maximize sectional curvature are $X = (0, 1, 0)$ and $Y = (1, 0, 1)$: moving along these directions makes the central

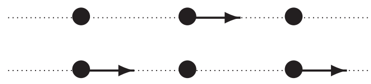
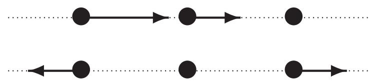
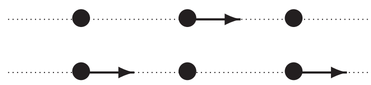
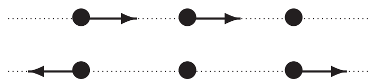
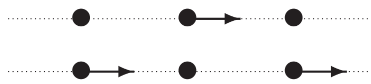
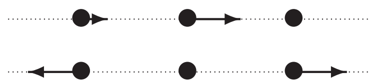
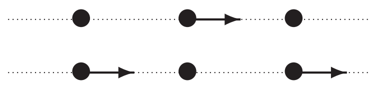
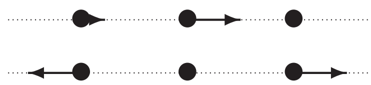
Position	Eigenvalues	Eigenvectors are $\omega = X^\flat \wedge Y^\flat$, with:	
$\varrho^{12} = 0.025$ $\varrho^{23} = 0.025$	$\sigma_{\max} = 4.99 \cdot 10^{-2}$		$X = (0, 1, 0)$ $Y = (1, 0, 1)$
	$\sigma_{\min} = -5.48 \cdot 10^{-1}$		$X \simeq (2, 1, 0)$ $Y = (-1, 0, 1)$
$\varrho^{12} \simeq 1.15$ $\varrho^{23} \simeq 1.15$	$\sigma_{\max} = 2.28 \cdot 10^{-2}$		$X = (0, 1, 0)$ $Y = (1, 0, 1)$
	$\sigma_{\min} = -2.94 \cdot 10^{-1}$		$X = (1.03, 1, 0)$ $Y = (-1, 0, 1)$
$\varrho^{12} = 2.15$ $\varrho^{23} = 2.15$	$\sigma_{\max} = 9.38 \cdot 10^{-2}$		$X = (0, 1, 0)$ $Y = (1, 0, 1)$
maximum of $\sigma_{\max}$ along $\varrho^{12} = \varrho^{23}$	$\sigma_{\min} = -3.95 \cdot 10^{-3}$		$X = (0.198, 1, 0)$ $Y = (-1, 0, 1)$
$\varrho^{12} = 3.5$ $\varrho^{23} = 3.5$	$\sigma_{\max} = 9.73 \cdot 10^{-3}$		$X = (0, 1, 0)$ $Y = (1, 0, 1)$
large distances	$\sigma_{\min} = -1.02 \cdot 10^{-7}$		$X \simeq (\varepsilon, 0, -1)$ $Y = (-1, 0, 1)$

TABLE 5.1. Eigenvalues and pairs (X, Y) that achieve them, along $\varrho^{12} = \varrho^{23}$; ε denotes a very small (but nonzero) number.

landmark closer to the other two, which on the other hand do not change their relative positions. On the other end of the path, in the position denoted by a circle (we are close to the boundary of the manifold, when two landmarks almost coincide), the two tangent vectors that maximize sectional curvature are close to $X = (1, 1, 0)$ and $Y = (0, 0, 1)$, "as if" the two close landmarks were actually clustered into one.

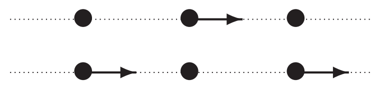
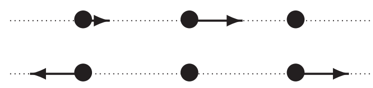
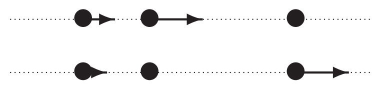
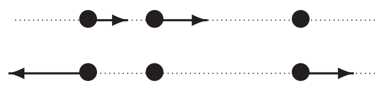
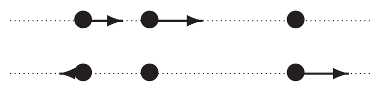
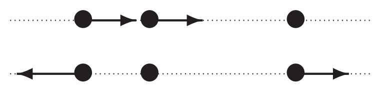
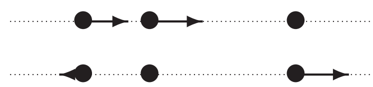
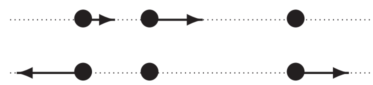
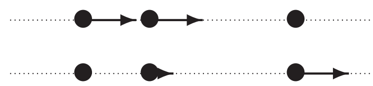
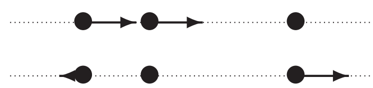
Position	Eigenvalues	Eigenvectors are $\omega = X^\flat \wedge Y^\flat$, with:	
$\varrho^{12} = 2.225$ $\varrho^{23} = 2.225$	$\sigma_{\max} = 9.27 \cdot 10^{-2}$		$X = (0, 1, 0)$ $Y = (1, 0, 1)$
maximum of trace along $\varrho^{12} = \varrho^{23}$ $\square$	$\sigma_{\min} = -2.51 \cdot 10^{-3}$		$X = (0.168, 1, 0)$ $Y = (-1, 0, 1)$
$\varrho^{12} = 1.525$ $\varrho^{23} = 2.30$	$\sigma_{\max} = 8.79 \cdot 10^{-2}$		$X = (0.316, 1, 0)$ $Y = (0.039, 0, 1)$
$\triangleright$	$\sigma_{\min} = -2.51 \cdot 10^{-2}$		$X = (0.445, 1, 0)$ $Y = (-2.12, 0, 1)$
$\varrho^{12} = 0.825$ $\varrho^{23} = 2.45$	$\sigma_{\max} = 8.18 \cdot 10^{-2}$		$X = (0.620, 1, 0)$ $Y = (-0.084, 0, 1)$
$\diamond$	$\sigma_{\min} = -9.40 \cdot 10^{-2}$		$X = (1.040, 1, 0)$ $Y = (-1.519, 0, 1)$
$\varrho^{12} = 0.425$ $\varrho^{23} = 2.65$	$\sigma_{\max} = 7.84 \cdot 10^{-2}$		$X = (0.807, 1, 0)$ $Y = (-0.122, 0, 1)$
$\triangleleft$	$\sigma_{\min} = -7.00 \cdot 10^{-2}$		$X = (0.318, 1, 0)$ $Y = (-1.532, 0, 1)$
$\varrho^{12} = 0.025$ $\varrho^{23} = 2.875$	$\sigma_{\max} = 7.89 \cdot 10^{-2}$		$X \simeq (1 - \varepsilon, 1, 0)$ $Y \simeq (0, \varepsilon, 1)$
$\circ$	$\sigma_{\min} = -5.75 \cdot 10^{-2}$		$X \simeq (1 + \varepsilon, 1, 0)$ $Y \simeq (-\varepsilon, 0, 1)$

TABLE 5.2. Eigenvalues and pairs (X, Y) that achieve them, in selected locations of along the ridge of scalar curvature, shown in Figure 5.8; ε denotes a very small (but nonzero) number.

4. Sectional curvature for N one-dimensional landmarks

In the previous chapter we computed the following formula for the numerator of sectional curvature of an n-dimensional Riemannian manifold in terms of the cometric

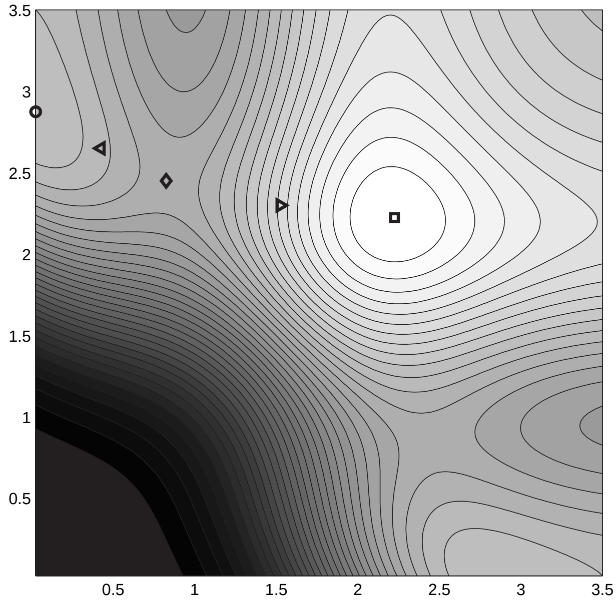


FIGURE 5.8. Locations along the "ridge" of scalar curvature for Table 5.2.

tensor g^{ij} and its partial derivatives $g^{ij,}_{\ \ k}$ and $g^{ij,}_{\ \ k\ell}$:

(5.25)

$$
\begin{aligned}
&2R^{ursv}X_uY_rY_sX_v \\
&= \big(X_uY_r - Y_uX_r\big)\Big(g^{su,rv} - \frac{1}{4}\,g^{us,}_{\ \ \ \varphi}\,g^{rv,\varphi} + g^{us,}_{\ \ \ \varphi}\,g^{\varphi r,v} - \frac{3}{2}\,g^{u\psi,r}g_{\psi\xi}\,g^{\xi s,v}\Big) \\
&\quad\cdot \big(X_sY_v - Y_sX_v\big) \\
&= \big(X_uY_r - Y_uX_r\big) \\
&\quad\cdot \Big(g^{r\varphi}g^{su,}_{\ \ \ \varphi\psi}\,g^{\psi v} - \frac{1}{4}\,g^{us,}_{\ \ \ \varphi}\,g^{\varphi\psi}g^{rv,}_{\ \ \ \psi} + g^{us,}_{\ \ \ \varphi}\,g^{\varphi r,}_{\ \ \ \psi}\,g^{\psi v} - \frac{3}{2}\,g^{r\varphi}g^{u\psi,}_{\ \ \ \varphi}\,g_{\psi\xi}\,g^{\xi s,}_{\ \ \ \eta}\,g^{\eta v}\Big) \\
&\quad\cdot \big(X_sY_v - Y_sX_v\big).
\end{aligned}
$$

Our notation is such that the sign of the above quantity the same as the sign of sectional curvature $K(X,Y)$. Our objective in this section is to provide a formula for sectional curvature in the case of N landmarks in one dimension; we will achieve this be inserting the expression for the partial derivative of the cometric, which are provided by Lemma 5.2, into the right-hand-side of expression (5.25). We will first "process" separately the four terms in the middle factor of (5.25). We should note that the dimension of the shape manifold $\mathcal{I}$ is in this case $n = N$; since $\Lambda^2(T_I\mathcal{I}) \simeq \mathbb{R}^{\binom{N}{2}}$ is not decomposable for $N \geq 3$ the method illustrated in section 5 of the previous chapter would only provide upper and lower bounds for sectional curvature at a point, not necessarily achievable at a pair of tangent vectors.

4.1. First term. An appropriate relabeling of the indices in (5.5) yields

$$g^{su,}{}_{\varphi\psi} = (\delta^s_\varphi - \delta^u_\varphi)\,(\delta^s_\psi - \delta^u_\psi)\,\gamma''(\varrho^{su})\,,$$

so that the first term in the middle factor of (5.25) can be expressed as

$$\begin{aligned} g^{r\varphi}g^{su,}{}_{\varphi\psi}\,g^{\psi v} &= (\delta^s_\varphi - \delta^u_\varphi)\,(\delta^s_\psi - \delta^u_\psi)\,\gamma''(\varrho^{su})\,\gamma(\varrho^{r\varphi})\,\gamma(\varrho^{\psi v}) \\ &= \big[\gamma(\varrho^{rs}) - \gamma(\varrho^{ru})\big]\,\big[\gamma(\varrho^{sv}) - \gamma(\varrho^{uv})\big]\,\gamma''(\varrho^{su}). \end{aligned}$$

Further simplifications can be derived after the multiplication by the components of the cotangent vectors. In fact we have that

$$\begin{aligned} g^{r\varphi}g^{su,}{}_{\varphi\psi}\,g^{\psi v} &= \big[\gamma(\varrho^{rs}) - \gamma(\varrho^{ru})\big]\,\big[\gamma(\varrho^{sv}) - \gamma(\varrho^{uv})\big]\,\gamma''(\varrho^{su}) \\ &= \big[\gamma(\varrho^{rs})\gamma(\varrho^{sv}) - \gamma(\varrho^{rs})\gamma(\varrho^{uv}) - \gamma(\varrho^{ru})\gamma(\varrho^{sv}) + \gamma(\varrho^{ru})\gamma(\varrho^{uv})\big]\,\gamma''(\varrho^{su}); \end{aligned} \tag{5.26}$$

but appropriately relabeling the indices ($u \to s$, $s \to u$, $r \to v$, $v \to r$) yields:

$$\begin{aligned} &\sum_u \big(X_uY_r - Y_uX_r\big)\big[\gamma(\varrho^{ru})\gamma(\varrho^{uv})\gamma''(\varrho^{su})\big]\big(X_sY_v - Y_sX_v\big) \\ &\quad = \sum_s \big(X_sY_v - Y_sX_v\big)\big[\gamma(\varrho^{vs})\gamma(\varrho^{sr})\gamma''(\varrho^{us})\big]\big(X_uY_r - Y_uX_r\big), \end{aligned}$$

so that the first and fourth terms within the square brackets of (5.26) can be combined. Note that we are summing over *all* four indices u, r, s and v; however, we write the summation symbols $\sum_u$ and $\sum_s$ because indices u and s break the rules of Einstein's summation convention, since they appear three times within the same term. In conclusion, we may write:

$$\begin{aligned} \big(X_uY_r - Y_uX_r\big)g^{r\varphi}g^{su,}{}_{\varphi\psi}\,g^{\psi v}\big(X_sY_v - Y_sX_v\big) &= \sum_{su} \big(X_uY_r - Y_uX_r\big)\cdot \\ \cdot\Big\{\big[2\gamma(\varrho^{rs})\gamma(\varrho^{sv}) - \gamma(\varrho^{rs})\gamma(\varrho^{uv}) - \gamma(\varrho^{ru})\gamma(\varrho^{sv})\big]\,&\gamma''(\varrho^{su})\Big\}\big(X_sY_v - Y_sX_v\big). \end{aligned}$$

4.2. Second term. First note that following auxiliary identity holds:

$$
\begin{aligned}
&\gamma(\varrho^{\varphi\psi})(\delta^u_\varphi - \delta^s_\varphi)(\delta^r_\psi - \delta^v_\psi) \;=\; \big[\gamma(\varrho^{\varphi\psi})\delta^u_\varphi - \gamma(\varrho^{\varphi\psi})\delta^s_\varphi\big](\delta^r_\psi - \delta^v_\psi)\\
&\quad= \big[\gamma(\varrho^{u\psi}) - \gamma(\varrho^{s\psi})\big](\delta^r_\psi - \delta^v_\psi) \;=\; \big[\gamma(\varrho^{u\psi}) - \gamma(\varrho^{s\psi})\big]\delta^r_\psi - \big[\gamma(\varrho^{u\psi}) - \gamma(\varrho^{s\psi})\big]\delta^v_\psi\\
&\quad= \gamma(\varrho^{ur}) - \gamma(\varrho^{sr}) - \gamma(\varrho^{uv}) + \gamma(\varrho^{sv}),
\end{aligned}
$$

so that the second term of the middle factor of (5.25) can be rewritten as:

$$
\begin{aligned}
-\frac{1}{4}\, g^{us,}{}_{\varphi}\, g^{\varphi\psi} g^{rv,}{}_{\psi} \;&=\; -\frac{1}{4}\,(\delta^u_\varphi - \delta^s_\varphi)\, f(q^u, q^s)\, \gamma(\varrho^{\varphi\psi})\,(\delta^r_\psi - \delta^v_\psi)\, f(q^r, q^v)\\
&=\; -\frac{1}{4}\big[\gamma(\varrho^{ur}) - \gamma(\varrho^{sr}) - \gamma(\varrho^{uv}) + \gamma(\varrho^{sv})\big] f(q^u, q^s)\, f(q^r, q^v).
\end{aligned}
$$

Simplifications occur after multiplying by the cotangent vector components. Relabeling the indices ($u \to s$, $s \to u$, $r \to v$, $v \to r$) yields:

$$
\begin{aligned}
&\sum_{sv} \big(X_u Y_r - Y_u X_r\big)\, \gamma(\varrho^{sv}) f(q^u, q^s) f(q^r, q^v)\, \big(X_s Y_v - Y_s X_v\big)\\
&\quad= \sum_{ur} \big(X_s Y_v - Y_s X_v\big)\, \gamma(\varrho^{ur}) f(q^s, q^u) f(q^v, q^r)\, \big(X_u Y_r - Y_u X_r\big)\\
&\quad= \sum_{ur} \big(X_s Y_v - Y_s X_v\big)\, \gamma(\varrho^{ur}) f(q^u, q^s) f(q^r, q^v)\, \big(X_u Y_r - Y_u X_r\big)
\end{aligned}
$$

(we have also used the antisymmetry of function f) and

$$
\begin{aligned}
&\sum_{uv} \big(X_u Y_r - Y_u X_r\big)\, \gamma(\varrho^{uv}) f(q^u, q^s) f(q^r, q^v)\, \big(X_s Y_v - Y_s X_v\big)\\
&\quad= \sum_{sr} \big(X_s Y_v - Y_s X_v\big)\, \gamma(\varrho^{sr}) f(q^s, q^u) f(q^v, q^r)\, \big(X_u Y_r - Y_u X_r\big)\\
&\quad= \sum_{sr} \big(X_s Y_v - Y_s X_v\big)\, \gamma(\varrho^{sr}) f(q^u, q^s) f(q^r, q^v)\, \big(X_u Y_r - Y_u X_r\big).
\end{aligned}
$$

In conclusion, we may write the second term as follows:

$$
\begin{aligned}
&\big(X_u Y_r - Y_u X_r\big)\Big(-\frac{1}{4}\, g^{us,}{}_{\varphi}\, g^{\varphi\psi} g^{rv,}{}_{\psi}\Big)\big(X_s Y_v - Y_s X_v\big)\\
&\quad= \sum_{urs} \big(X_u Y_r - Y_u X_r\big)\Big\{-\frac{1}{2}\,\big[\gamma(\varrho^{ur}) - \gamma(\varrho^{sr})\big] f(q^u, q^s)\, f(q^r, q^v)\Big\}\big(X_s Y_v - Y_s X_v\big).
\end{aligned}
$$

4.3. Third term. The following holds (we're summing over both φ and ψ):

$$g^{us,}{}_{\varphi}\, g^{\varphi r,}{}_{\psi}\, g^{\psi v} \;=\; \sum_{\varphi} (\delta^u_\varphi - \delta^s_\varphi)\, f(q^u, q^s)\, (\delta^\varphi_\psi - \delta^r_\psi)\, f(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) \;=\; A^{ursv}\, f(q^u, q^s),$$

with:

$$\begin{aligned}
A^{ursv} &\triangleq \sum_{\varphi} (\delta^u_\varphi - \delta^s_\varphi)\, (\delta^\varphi_\psi - \delta^r_\psi)\, f(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) \\
&= \sum_{\varphi}\; \delta^u_\varphi \delta^\varphi_\psi\, f(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) -\; \delta^u_\varphi \delta^r_\psi\, f(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) \\
&\quad - \sum_{\varphi}\; \delta^s_\varphi \delta^\varphi_\psi\, f(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) +\; \delta^s_\varphi \delta^r_\psi\, f(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) \\
&=\; f(q^u, q^r)\, \gamma(\varrho^{uv}) -\; f(q^u, q^r)\, \gamma(\varrho^{rv}) -\; f(q^s, q^r)\, \gamma(\varrho^{sv}) +\; f(q^s, q^r)\, \gamma(\varrho^{rv}) \\
&=\; f(q^u, q^r)\, \big[\gamma(\varrho^{uv}) - \gamma(\varrho^{rv})\big] + f(q^s, q^r)\, \big[\gamma(\varrho^{rv}) - \gamma(\varrho^{sv})\big].
\end{aligned}$$

Whence the third term takes the form:

$$g^{us,}{}_{\varphi}\, g^{\varphi r,}{}_{\psi}\, g^{\psi v} = \Big\{ f(q^u, q^r)\, \big[\gamma(\varrho^{uv}) - \gamma(\varrho^{rv})\big] + f(q^s, q^r)\, \big[\gamma(\varrho^{rv}) - \gamma(\varrho^{sv})\big] \Big\}\, f(q^u, q^s).$$

4.4. Fourth term. We need to compute: $-\dfrac{3}{2}\, g^{r\varphi} g^{u\psi,}{}_{\varphi}\, g_{\psi\xi}\, g^{\xi s,}{}_{\eta}\, g^{\eta v}$. Note that:

$$g^{r\varphi} g^{u\psi,}{}_{\varphi} \;= \gamma(\varrho^{r\varphi})\, (\delta^u_\varphi - \delta^\psi_\varphi)\, f(q^u, q^\psi) = \big[\gamma(\varrho^{ru}) - \gamma(\varrho^{r\psi})\big]\, f(q^u, q^\psi),$$

and analogously $g^{v\eta} g^{\xi s,}{}_{\eta} \;= \big[\gamma(\varrho^{v\xi}) - \gamma(\varrho^{vs})\big]\, f(q^\xi, q^s)$. Therefore:

$$\begin{aligned}
&-\frac{3}{2}\, g^{r\varphi} g^{u\psi,}{}_{\varphi}\, g_{\psi\xi}\, g^{\xi s,}{}_{\eta}\, g^{\eta v} \\
&\qquad = -\frac{3}{2} \sum_{\varphi\xi}\; \big[\gamma(\varrho^{ru}) - \gamma(\varrho^{r\psi})\big]\, f(q^u, q^\psi)\, g_{\psi\xi}\, f(x^\xi, q^s)\, \big[\gamma(\varrho^{v\xi}) - \gamma(\varrho^{vs})\big].
\end{aligned}$$

Remark. Let $Z^\psi \triangleq \big(X_u Y_r - Y_u X_r\big)\, g^{r\varphi} g^{u\psi,}{}_{\varphi}$. Then

$$\big(X_u Y_r - Y_u X_r\big) \Big(-\frac{3}{2}\, g^{r\varphi} g^{u\psi,}{}_{\varphi}\, g_{\psi\xi}\, g^{\xi s,}{}_{\eta}\, g^{\eta v} \Big) \big(X_s Y_v - Y_s X_v\big) = -\frac{3}{2}\, Z^\psi g_{\psi\xi}\, Z^\xi < 0,$$

since the metric tensor $g_{\psi\xi}$ is positive definite. Whence the fourth term provides a negative contribution to the numerator of sectional curvature, for any choice of tangent vectors $X = q^i \partial_i$ and $Y = Y^j \partial_j$ in $T_I \mathcal{I}$.

4.5. Summation of four terms. The above discussion boils down to the following way of rewriting formula (5.25):

$$
\begin{aligned}
2R^{ursv}X_uY_rY_sX_v \;=\; & \sum_{rsu} \big(X_uY_r - Y_uX_r\big)\cdot \\
\cdot \Bigg(& \big[2\gamma(\varrho^{rs})\gamma(\varrho^{sv}) - \gamma(\varrho^{rs})\gamma(\varrho^{uv}) - \gamma(\varrho^{ru})\gamma(\varrho^{sv})\big]\,\gamma''(\varrho^{su}) + f(q^u,q^s)\cdot \\
& \cdot \Big\{ f(q^u,q^r)\big[\gamma(\varrho^{uv}) - \gamma(\varrho^{rv})\big] + f(q^s,q^r)\big[\gamma(\varrho^{rv}) - \gamma(\varrho^{sv})\big] - \frac{1}{2} f(q^r,q^v)\big[\gamma(\varrho^{ur}) - \gamma(\varrho^{sr})\big]\Big\} \\
& - \frac{3}{2}\sum_{\varphi\xi}\,\big[\gamma(\varrho^{ru}) - \gamma(\varrho^{r\psi})\big]\, f(q^u,q^\psi)\, g_{\psi\xi}\, f(x^\xi,q^s)\,\big[\gamma(\varrho^{v\xi}) - \gamma(\varrho^{vs})\big]\Bigg)\big(X_sY_v - Y_sX_v\big)\,;
\end{aligned}
$$

in the next section we will extend the above formula to the D-dimensional case.

5. Sectional curvature for N D-dimensional landmarks

As we discussed in Chapter 3 in the case of landmarks in $D \geq 2$ dimensions it is convenient to introduce a metric tensor and a cometric tensor with "double indices", one that refers to the landmark index and the other that refers to the dimensional component of that specific landmark. Given the block-diagonal nature of the metric tensor we can write its generic element as follows:

$$
g^{iajb}(q) = h^{ij}(q)\,\delta^{ab}, \qquad i,j = 1,\ldots,N, \quad a,b = 1,\ldots,D \tag{5.27}
$$

where (in the case of exact matching, $\lambda = \infty$) $h^{ij}(q) = G(q^i,q^j) = \gamma\big(\|q^i - q^j\|_{\mathbb{R}^D}\big)$. The following result is the D-dimensional counterpart of Lemma 5.2.

LEMMA 5.9. *Let $\gamma : [0,+\infty) \to \mathbb{R}$ be such that $G(x,y) = \gamma\big(\|x-y\|_{\mathbb{R}^D}\big)$ is the kernel of admissible space V. Then the first and second partial derivatives of the cometric tensor for N one-dimensional landmarks are respectively given by:*

$$
g^{iajb,}{}_{kc} \;=\; \frac{\partial}{\partial q^{kc}}\, g^{iajb} \;=\; (\delta^i_k - \delta^j_k)\,\gamma'(\varrho^{ij})\,\frac{q^{ic} - q^{jc}}{\varrho^{ij}}\,\delta^{ab}, \tag{5.28}
$$

and

$$g^{iajb,}{}_{kc\ell d} = \frac{\partial}{\partial q^{\ell d}}\, g^{iajb,}{}_{kc} = (\delta^i_k - \delta^j_k)\,(\delta^i_\ell - \delta^j_\ell)\,\delta^{ab}\Big\{\gamma''(\varrho^{ij})\,\frac{\big(q^{ic}-q^{jc}\big)\big(q^{id}-q^{jd}\big)}{\big(\varrho^{ij}\big)^2}$$

$$+\,\frac{\gamma'(\varrho^{ij})}{\varrho^{ij}}\Big[\delta_{cd} - \frac{\big(q^{ic}-q^{jc}\big)\big(q^{id}-q^{jd}\big)}{\big(\varrho^{ij}\big)^2}\Big]\Big\}, \tag{5.29}$$

where $\varrho^{ij} \triangleq \|q^i - q^j\|_{\mathbb{R}^D}$.

PROOF. We have that $g^{iajb,}{}_{kc} = \frac{\partial}{\partial q^{kc}}\big(h^{ij}\big)\delta^{ab}$, where

$$\frac{\partial}{\partial q^{kc}}\, h^{ij}(q) = \frac{\partial}{\partial q^{kc}}\,\gamma(\varrho^{ij}) = \begin{cases} 0 & \text{for } k \neq i, k \neq j \\ \frac{\partial}{\partial q^{ic}}\gamma(\varrho^{ij}) & \text{for } k = i \\ \frac{\partial}{\partial q^{jc}}\gamma(\varrho^{ij}) & \text{for } k = j \end{cases}$$

By the chain rule

$$\frac{\partial}{\partial q^{ic}}\,\gamma(\varrho^{ij}) \;=\; \gamma'(\varrho^{ij})\,\frac{\partial}{\partial q^{ic}}\,\varrho^{ij} \;=\; \gamma'(\varrho^{ij})\,\frac{q^{ic}-q^{jc}}{\|q^i-q^j\|_{\mathbb{R}^D}}\,,$$

hence

$$g^{iajb,}{}_{kc} \;=\; \begin{cases} 0 & \text{for } k \neq i, k \neq j \\ \gamma'(\varrho^{ij})\,\dfrac{q^{ic}-q^{jc}}{\|q^i-q^j\|_{\mathbb{R}^D}}\,\delta^{ab} & \text{for } k = i \\ -\gamma'(\varrho^{ij})\,\dfrac{q^{ic}-q^{jc}}{\|q^i-q^j\|_{\mathbb{R}^D}}\,\delta^{ab} & \text{for } k = j \end{cases}$$

which we may write in the compact form

$$g^{iajb,}{}_{kc} \;=\; (\delta^i_k - \delta^j_k)\,\gamma'(\varrho^{ij})\,\frac{q^{ic}-q^{jc}}{\varrho^{ij}}\,\delta^{ab}, \tag{5.30}$$

which is precisely (5.28). We can use such expression to write:

$$\begin{aligned} g^{iajb,}{}_{kc\ell d} &= \frac{\partial}{\partial q^{\ell d}} g^{iajb,}{}_{kc} \;=\; (\delta^i_k - \delta^j_k)\,\frac{\partial}{\partial q^{\ell d}}\Big[\gamma'(\varrho^{ij})\,\frac{q^{ic}-q^{jc}}{\varrho^{ij}}\Big]\delta^{ab} \\ &= (\delta^i_k - \delta^j_k)\,\Big\{\frac{q^{ic}-q^{jc}}{\varrho^{ij}}\,\frac{\partial}{\partial q^{\ell d}}\,\gamma'(\varrho^{ij}) \\ &\qquad + \frac{\gamma'(\varrho^{ij})}{\varrho^{ij}}\,\frac{\partial}{\partial q^{\ell d}}\,(q^{ic}-q^{jc}) + \gamma'(\varrho^{ij})\big(q^{ic}-q^{jc}\big)\,\frac{\partial}{\partial q^{\ell d}}\,\frac{1}{\varrho^{ij}}\Big\}\delta^{ab}; \end{aligned} \tag{5.31}$$

we will now compute the three derivatives inside the curly brackets.

In a way that is analogous to how (5.30) was computed we can prove that:

$$\frac{\partial}{\partial q^{\ell d}}\,\gamma'(\varrho^{ij}) = (\delta^i_\ell - \delta^j_\ell)\,\gamma''(\varrho^{ij})\,\frac{q^{id}-q^{jd}}{\varrho^{ij}}. \tag{5.32}$$

Also,

$$\frac{\partial}{\partial q^{\ell d}}\left(q^{ic}-q^{jc}\right) = \begin{cases} 0 & \text{if } \ell \neq i \text{ and } \ell \neq j \\ 0 & \text{if } c \neq d \\ \frac{\partial}{\partial q^{ic}}\left(q^{ic}-q^{jc}\right) = 1 & \text{if } \ell = i \text{ and } d = c \\ \frac{\partial}{\partial q^{ic}}\left(q^{ic}-q^{jc}\right) = -1 & \text{if } \ell = j \text{ and } d = c \end{cases}$$

which may be expressed compactly as

$$\frac{\partial}{\partial q^{\ell d}}\left(q^{ic}-q^{jc}\right) \;=\; \delta^i_\ell\delta^c_d - \delta^j_\ell\delta^c_d \;=\; (\delta^i_\ell - \delta^j_\ell)\delta^c_d\,. \tag{5.33}$$

We should now note that

$$\frac{\partial}{\partial q^{\ell d}}\,\varrho^{ij} = \frac{\partial}{\partial q^{\ell d}}\,\|q^i - q^j\|_{\mathbb{R}^D} = \begin{cases} 0 & \text{if } \ell \neq i \text{ and } \ell \neq i \\ \dfrac{\partial}{\partial q^{id}}\,\varrho^{ij} = \dfrac{q^{id}-q^{jd}}{\varrho^{ij}} & \text{if } \ell = i \\ \dfrac{\partial}{\partial q^{jd}}\,\varrho^{ij} = \dfrac{q^{jd}-q^{id}}{\varrho^{ij}} & \text{if } \ell = j \end{cases}$$

that is,

$$\frac{\partial}{\partial q^{\ell d}}\,\varrho^{ij} \;=\; (\delta^i_\ell - \delta^j_\ell)\,\frac{q^{id}-q^{jd}}{\varrho^{ij}};$$

whence we can finally compute

$$\frac{\partial}{\partial q^{\ell d}}\,\frac{1}{\varrho^{ij}} \;=\frac{\partial}{\partial q^{\ell d}}\left(\varrho^{ij}\right)^{-1} \;=-\frac{1}{\left(\varrho^{ij}\right)^2}\,\frac{\partial}{\partial q^{\ell d}}\,\varrho^{ij} \;=\; -(\delta^i_\ell - \delta^j_\ell)\,\frac{q^{id}-q^{jd}}{\left(\varrho^{ij}\right)^3}. \tag{5.34}$$

Inserting the right-hand sides of equations (5.32), (5.33) and (5.34) into the right-hand side of (5.31) finally yields (5.29), thus completing the proof of the lemma. □

Remark. We should note that, as expected, in the case $D = 1$ equations (5.28) and (5.29) reduce, respectively, to equations (5.4) and (5.5) provided by Lemma 5.2. In fact in this case we have that $a = b = c = d = 1$, whence

$$\frac{q^{ic}-q^{jc}}{\varrho^{ij}} = \operatorname{sgn}(q^{i1}-q^{j1}),$$

$$\frac{\left(q^{ic}-q^{jc}\right)\left(q^{id}-q^{jd}\right)}{\left(\varrho^{ij}\right)^2} = \frac{\left(q^{i1}-q^{j1}\right)^2}{\left(\varrho^{ij}\right)^2} = 1$$

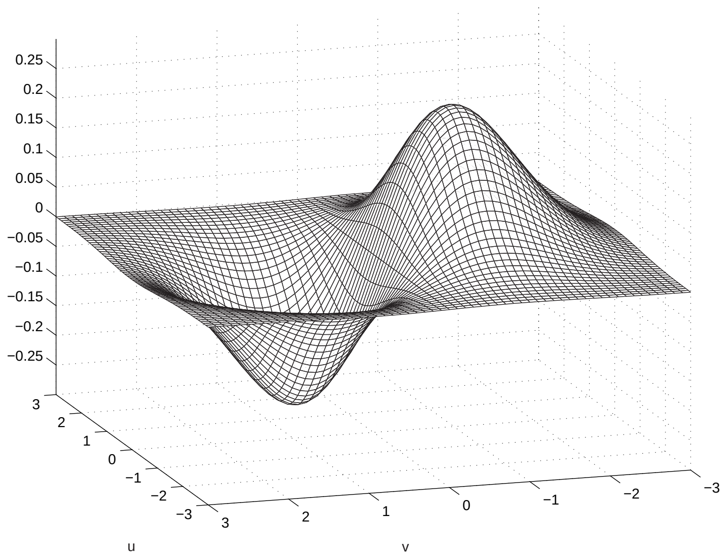


FIGURE 5.9. Typical shape of function $\widetilde{f}_1(x)$, $x = (v, u)$, for $D = 2$;

and obviously $\delta^{ab} = \delta^{cd} = 1$.

DEFINITION 5.10. For notational convenience, for a fixed index $a \in \{1, \ldots, D\}$ we will introduce function $\widetilde{f}_a : \mathbb{R}^D \to \mathbb{R} : x \mapsto \gamma'\big(\|x\|_{\mathbb{R}^D}\big)\,\frac{x^a}{\|x\|_{\mathbb{R}^D}}$ and define:

$$f_a(x,y) \triangleq \widetilde{f}_a(x-y) = \gamma'\big(\|x-y\|_{\mathbb{R}^D}\big)\,\frac{x^a - y^a}{\|x-y\|_{\mathbb{R}^D}},$$

where x^a and y^a are the a-th components of points $x, y \in \mathbb{R}^D$.

As we had for function f introduced in Definition 5.3, it is the case that $f_a(x,y) = -f_a(y,x)$ and $f_a(x,x) = 0$, for any choice of index $a \in \{1, \ldots, D\}$. Figure 5.9 shows the typical shape of function $\widetilde{f}_1 : \mathbb{R}^D \to \mathbb{R}$ in the case $D = 2$; compare it with the one-dimensional counterpart shown in Figure 5.1. With the above notation, equation (5.28) can be simply written as

$$g^{iajb,}{}_{kc} \;=\; (\delta^i_k - \delta^j_k)\, f_c(q^i, q^j)\, \delta^{ab}.$$

We are now going to apply the results of Lemma 5.9 to Theorem 4.4 so to compute the general formula for sectional curvature for the Riemannian manifold of N

landmarks in D dimensions. In order to do that, we first need to rewrite the formula provided by Theorem 4.4 in terms of the "double index" notation introduced in Chapter 3. We are going to use u, r, s, v for the indices concerning the landmark label, and a, b, c, d for those regarding the component of a specific landmark; as far as summation indices are concerned, we are going to use φ, ψ, ξ, η and $\alpha, \beta, \gamma, \varepsilon$ for the two types of indices, respectively. We have:

$$\begin{aligned}2R^{uarbscvd}X_{ua}Y_{rb}Y_{sc}X_{vd} \;=\; & \big(X_{ua}Y_{rb} - Y_{ua}X_{rb}\big)\Big(g^{scua,rbvd} - \frac{1}{4}\, g^{uasc,}{}_{\varphi\alpha}\, g^{rbvd,\varphi\alpha} \\ & + g^{uasc,}{}_{\varphi\alpha}\, g^{\varphi\alpha rb,vd} - \frac{3}{2}\, g^{ua\psi\beta,rb} g_{\psi\beta\xi\gamma}\, g^{\xi\gamma sc,vd}\Big)\big(X_{sc}Y_{vd} - Y_{sc}X_{vd}\big),\end{aligned}$$

with $u, r, s, v = 1, \ldots, N$ and $a, b, c, d = 1, \ldots, D$. As we did in section 4 we are going to analyze the four terms in the central factor of the right-hand side of the above equation one by one.

5.1. First term. We can use formula (5.29) from Lemma 5.9 to compute:

$$\begin{aligned}g^{scua,rbvd} &\triangleq g^{scua,}{}_{\varphi\alpha\psi\beta}\; g^{\varphi\alpha rb} g^{\psi\beta vd} \\ &= (\delta^s_\varphi - \delta^u_\varphi)\,(\delta^s_\psi - \delta^u_\psi)\,\delta^{ca}\Bigg\{\gamma''(\varrho^{su})\,\frac{\big(q^{s\alpha} - q^{u\alpha}\big)\big(q^{s\beta} - q^{u\beta}\big)}{\big(\varrho^{su}\big)^2} \\ &\quad + \frac{\gamma'(\varrho^{su})}{\varrho^{su}}\Bigg[\delta_{\alpha\beta} - \frac{\big(q^{s\alpha} - q^{u\alpha}\big)\big(q^{s\beta} - q^{u\beta}\big)}{\big(\varrho^{su}\big)^2}\Bigg]\Bigg\}\,\gamma(\varrho^{\varphi r})\,\delta^{\alpha b}\,\gamma(\varrho^{\psi v})\,\delta^{\beta d},\end{aligned}$$

that is, summing over indices α, β, φ and ψ,

$$\begin{aligned}g^{scua,rbvd} &= \big[\gamma(\varrho^{sr}) - \gamma(\varrho^{ur})\big]\big[\gamma(\varrho^{sv}) - \gamma(\varrho^{uv})\big]\delta^{ac}\Bigg\{\gamma''(\varrho^{su})\,\frac{\big(q^{sb} - q^{ub}\big)\big(q^{sd} - q^{ud}\big)}{\big(\varrho^{su}\big)^2} \\ &\quad + \frac{\gamma'(\varrho^{su})}{\varrho^{su}}\Bigg[\delta^{bd} - \frac{\big(q^{sb} - q^{ub}\big)\big(q^{sd} - q^{ud}\big)}{\big(\varrho^{su}\big)^2}\Bigg]\Bigg\},\end{aligned}$$

that is,

$$\begin{aligned}g^{scua,rbvd} \;=\; & \big[\underline{\gamma(\varrho^{sr})\gamma(\varrho^{sv})} - \gamma(\varrho^{sr})\gamma(\varrho^{uv}) - \gamma(\varrho^{ur})\gamma(\varrho^{sv}) + \underline{\gamma(\varrho^{ur})\gamma(\varrho^{uv})}\big]\delta^{ac} \\ & \cdot\Bigg\{\gamma''(\varrho^{su})\,\frac{\big(q^{sb} - q^{ub}\big)\big(q^{sd} - q^{ud}\big)}{\big(\varrho^{su}\big)^2} + \frac{\gamma'(\varrho^{su})}{\varrho^{su}}\Bigg[\delta^{bd} - \frac{\big(q^{sb} - q^{ub}\big)\big(q^{sd} - q^{ud}\big)}{\big(\varrho^{su}\big)^2}\Bigg]\Bigg\}.\end{aligned}$$

After multiplying by the components of the contangent vectors a simplification occurs, in the same way that they occurred in the one-dimensional case: that is, the two underlined terms in the above equation actually combine. The final result is:

$$\begin{aligned}
&\big(X_{ua}Y_{rb} - Y_{ua}X_{rb}\big)\, g^{scua,rbvd}\,\big(X_{sc}Y_{vd} - Y_{sc}X_{vd}\big) \\
&\quad = \sum_{su}\big(X_{ua}Y_{rb} - Y_{ua}X_{rb}\big)\Bigg(\big[2\gamma(\varrho^{sr})\gamma(\varrho^{sv}) - \gamma(\varrho^{sr})\gamma(\varrho^{uv}) - \gamma(\varrho^{ur})\gamma(\varrho^{sv})\big]\delta^{ac} \\
&\qquad\qquad \cdot\left\{\gamma''(\varrho^{su})\,\frac{\big(q^{sb}-q^{ub}\big)\big(q^{sd}-q^{ud}\big)}{\big(\varrho^{su}\big)^2} + \frac{\gamma'(\varrho^{su})}{\varrho^{su}}\Big[\delta^{bd} - \frac{\big(q^{sb}-q^{ub}\big)\big(q^{sd}-q^{ud}\big)}{\big(\varrho^{su}\big)^2}\Big]\right\}\Bigg) \\
&\qquad\qquad \cdot\big(X_{sc}Y_{vd} - Y_{sc}X_{vd}\big);
\end{aligned}$$

as usual, we have written the summation symbols explicitly only for those indices that break the rules of Einstein's summation convention.

5.2. Second term. This time we need formula (5.28) from Lemma 5.9:

$$\begin{aligned}
-\frac{1}{4}\, g^{uasc,}{}_{\varphi\alpha}\, g^{rbvd,\varphi\alpha} &= -\frac{1}{4} g^{uasc,}{}_{\varphi\alpha}\, g^{rbvd,}{}_{\psi\beta}\, g^{\varphi\alpha\psi\beta} \\
&= -\frac{1}{4}(\delta^u_\varphi - \delta^s_\varphi)\,\gamma'(\varrho^{us})\,\frac{q^{u\alpha}-q^{s\alpha}}{\varrho^{us}}\,\delta^{ac}\,(\delta^r_\psi - \delta^v_\psi)\,\gamma'(\varrho^{rv})\,\frac{q^{r\beta}-q^{v\beta}}{\varrho^{rv}}\,\delta^{bd}\,\gamma(\varrho^{\varphi\psi})\,\delta^{\alpha\beta} \\
&= -\frac{1}{4}\delta^{ac}\delta^{bd}\big[\gamma(\varrho^{ur}) - \gamma(\varrho^{uv}) - \gamma(\varrho^{sr}) + \gamma(\varrho^{sv})\big]\, f_\alpha(q^u,q^s)\delta^{\alpha\beta} f_\beta(q^r,q^v) \\
&= -\frac{1}{4}\delta^{ac}\delta^{bd}\big[\underline{\gamma(\varrho^{ur})} - \underline{\underline{\gamma(\varrho^{uv})}} - \underline{\underline{\gamma(\varrho^{sr})}} + \underline{\gamma(\varrho^{sv})}\big]\sum_{\alpha=1}^{D} f_\alpha(q^u,q^s)\, f_\alpha(q^r,q^v).
\end{aligned}$$

Once again, multiplying by the components of the cotangent vectors causes some simplifications: the two underlined and and the two doubly underlined terms combine. The final result is:

$$\begin{aligned}
&\big(X_{ua}Y_{rb} - Y_{ua}X_{rb}\big)\, g^{scua,rbvd}\,\big(X_{sc}Y_{vd} - Y_{sc}X_{vd}\big) \\
&\quad = \sum_{su}\big(X_{ua}Y_{rb} - Y_{ua}X_{rb}\big)\left(-\frac{1}{2}\,\delta^{ac}\,\delta^{bd}\big[\gamma(\varrho^{ur}) - \gamma(\varrho^{sr})\big]\sum_{\alpha=1}^{D} f_\alpha(q^u,q^s)\, f_\alpha(q^r,q^v)\right) \\
&\qquad\qquad \cdot\big(X_{sc}Y_{vd} - Y_{sc}X_{vd}\big).
\end{aligned}$$

5.3. Third term. Once again, by formula (5.28) from Lemma 5.9:

$$
\begin{aligned}
g^{uasc,}{}_{\varphi\alpha}\, g^{\varphi\alpha rb,vd} &= g^{uasc,}{}_{\varphi\alpha}\, g^{\varphi\alpha rb,}{}_{\psi\beta}\, g^{\psi\beta vd} \\
&= \sum_{\varphi} (\delta^u_\varphi - \delta^s_\varphi)\, f_\alpha(q^u, q^s)\, \delta^{ac}\, (\delta^\varphi_\psi - \delta^r_\psi)\, f_\beta(q^\varphi, q^r)\, \delta^{\alpha b}\, \gamma(\varrho^{\psi v})\, \delta^{\beta d} \\
&= \Big[\sum_{\varphi} \delta^u_\varphi \delta^\varphi_\psi\, f^d(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) - \delta^u_\varphi \delta^r_\psi\, f^d(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) \\
&\qquad - \sum_{\varphi} \delta^s_\varphi \delta^\varphi_\psi\, f^d(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) + \delta^s_\varphi \delta^r_\psi\, f^d(q^\varphi, q^r)\, \gamma(\varrho^{\psi v}) \Big]\, f^b(q^u, q^s)\, \delta^{ac} \\
&= \Big[\, f^d(q^u, q^r)\, \gamma(\varrho^{uv}) - f^d(q^u, q^r)\, \gamma(\varrho^{rv}) \\
&\qquad - \, f^d(q^s, q^r)\, \gamma(\varrho^{sv}) + f^d(q^s, q^r)\, \gamma(\varrho^{rv}) \Big]\, f^b(q^u, q^s)\, \delta^{ac} \\
&= \Big\{ \big[\gamma(\varrho^{uv}) - \gamma(\varrho^{rv})\big] f^d(q^u, q^r) + \big[\gamma(\varrho^{rv}) - \gamma(\varrho^{sv})\big] f^d(q^s, q^r) \Big\}\, f^b(q^u, q^s)\, \delta^{ac},
\end{aligned}
$$

where we have (only notationally) raised the indices in functions f^d and f^b to make them consistent with the left-hand side, e.g. defining $f^b \triangleq \delta^{\alpha b} f_\alpha$.

5.4. Fourth term. We have:

$$
\begin{aligned}
-\frac{3}{2}\, g^{ua\psi\beta,rb}\, g_{\psi\beta\xi\gamma}\, g^{\xi\gamma sc,vd} &= -\frac{3}{2}\, g^{ua\psi\beta,}{}_{\varphi\alpha}\, g^{\varphi\alpha rb}\, g_{\psi\beta\xi\gamma}\, g^{\xi\gamma sc,}{}_{\eta\varepsilon}\, g^{\eta\varepsilon vd} \\
= -\frac{3}{2} \sum_{\varphi\psi} (\delta^u_\varphi - \delta^\psi_\varphi)\, f_\alpha(q^u, q^\psi)\, \delta^{a\beta}\, &\gamma(\varrho^{\varphi r}) \delta^{\alpha b}\, g_{\psi\beta\xi\gamma}\, (\delta^\xi_\eta - \delta^s_\eta)\, f_\varepsilon(q^\xi, q^s)\, \delta^{\gamma c}\, \gamma(\varrho^{\eta v}) \delta^{\varepsilon d} \\
= -\frac{3}{2} \sum_{\varphi\psi} \big[\gamma(\varrho^{ur}) - \gamma(\varrho^{\psi r})\big]\, f^b(q^u, q^\psi)\, &h_{\psi\xi}\, f^d(q^\xi, q^s)\, \big[\gamma(\varrho^{\xi v}) - \gamma(\varrho^{sv})\big]\, \delta^{ac},
\end{aligned}
$$

where $h_{ij}(q)$ is the inverse tensor of $h^{ij}(q) = G(q^i, q^j)$, defined in (5.27): it is in fact the case that $g_{\psi\beta\xi\gamma}\, \delta^{a\eta}\, \delta^{\gamma c} = h_{\psi\xi}\, \delta_{\beta\gamma}\, \delta^{a\eta}\, \delta^{\gamma c} = h_{\psi\xi}\, \delta^{ac}$.

5.5. Summation of four terms. The discussion above finally leads to:

$$
\begin{aligned}
2R^{uarbscvd}X_{ua}Y_{rb}Y_{sc}X_{vd} \; = \; \sum_{rsu} \big(X_{ua}Y_{rb} - Y_{ua}X_{rb}\big)\big(X_{sc}Y_{vd} - Y_{sc}X_{vd}\big)\,\delta^{ac} \cdot \\
\cdot \Bigg(\big[2\gamma(\varrho^{sr})\gamma(\varrho^{sv}) - \gamma(\varrho^{sr})\gamma(\varrho^{uv}) - \gamma(\varrho^{ur})\gamma(\varrho^{sv})\big] \\
\cdot \left\{ \gamma''(\varrho^{su})\,\frac{\big(q^{sb}-q^{ub}\big)\big(q^{sd}-q^{ud}\big)}{\big(\varrho^{su}\big)^2} + \frac{\gamma'(\varrho^{su})}{\varrho^{su}}\Big[\delta^{bd} - \frac{\big(q^{sb}-q^{ub}\big)\big(q^{sd}-q^{ud}\big)}{\big(\varrho^{su}\big)^2}\Big]\right\} \\
- \frac{1}{2}\,\delta^{bd}\big[\gamma(\varrho^{ur}) - \gamma(\varrho^{sr})\big]\sum_{\alpha=1}^{D} f_\alpha(q^u,q^s)\,f_\alpha(q^r,q^v) \\
+ \Big\{\big[\gamma(\varrho^{uv}) - \gamma(\varrho^{rv})\big]f^d(q^u,q^r) + \big[\gamma(\varrho^{rv}) - \gamma(\varrho^{sv})\big]f^d(q^s,q^r)\Big\}f^b(q^u,q^s) \\
- \frac{3}{2}\sum_{\varphi\psi}\big[\gamma(\varrho^{ur}) - \gamma(\varrho^{\psi r})\big]\,f^b(q^u,q^\psi)\,h_{\psi\xi}\,f^d(q^\xi,q^s)\,\big[\gamma(\varrho^{\xi v}) - \gamma(\varrho^{sv})\big]\Bigg),
\end{aligned}
$$

where, as usual, we have written the summation symbols explicitly for the indices for which the rules of Einstein's summation conventions are broken, i.e. when the summation index appears more than twice in a product.

Remark. Once again, one can easily verify that the formula computed above for sectional curvature in the general case for N landmarks in D dimensions simplify to the one that we had computed for N one-dimensional landmarks reported at the end of section 4, simply by setting $a = b = c = d = 1$.

6. Conclusions

In this chapter we have provided explicit formulas for the dual Riemannian curvature tensor in the case of N one-dimensional landmarks and for sectional curvature in the case of N landmarks in one or $D \geq 2$ dimensions; in particular, we have have also analyzed in detail the graphs of sectional curvature for two and three landmarks on the real line. The formulas are expressed in terms of the function γ that defines the kernel $G(x,y) = \gamma\big(\|x-y\|_{\mathbb{R}^D}\big)$, and its first and second derivatives. For specific choices of γ its derivatives are related to it, e.g. by second order differential equation (2.15) in the case of Sobolev-type kernels or by equations (2.19) and (2.20) for the Gaussian and Cauchy kernels, respectively; using such expressions the formulas for sectional curvature could be made analytically more explicit, although we chose to leave them in their most general form—which is still numerically implementable.

In the next chapter we will use some of the formulas we worked out in the present one to investigate the effects of curvature on the qualitative dynamics of landmarks, i.e. the geodesic flow determined by the Hamiltonian system discussed in Chapter 3.

CHAPTER 6

The Qualitative Dynamics of Landmarks

In this chapter we explore the qualitative dynamics of geodesics for landmarks manifolds and analyze how such dynamics are influenced by curvature, which we studied in the previous chapter.

1. Introduction

Geodesics are determined by the Euler-Lagrange equations for the Riemannian energy that was introduced in Chapter 2. The geodesic flow on the tangent bundle can also be obtained from the *cogeodesic flow* determined by Hamilton's equations:

$$\left\{ \begin{array}{rcl} \dot{q}^i & = & \displaystyle\sum_{j=1}^{N} \Big(G(q^i, q^j) + \frac{\delta^{ij}}{\lambda} \Big)\, p_j \\ \dot{p}_i & = & -\displaystyle\sum_{j=1}^{N} \nabla_\xi G(q^i, q^j)\, \langle p_i, p_j \rangle_{\mathbb{R}^D} \end{array} \right. \qquad i = 1, \ldots, N. \tag{6.1}$$

which we derived in Chapter 3. Having studied curvature in the previous chapter we will analyze the effect that it has on the qualitative dynamics of landmarks, e.g. by verifying the existence of conjugate points (that is, points on the manifold that are connected by distinct geodesic curves) in regions of positive curvature, or by verifying the divergence of geodesics in regions of negative curvature.

Before proceeding we will express the conservation of the Hamiltonian for landmarks, whose expression is

$$\mathcal{H}(p, q) \;=\; \frac{1}{2} \sum_{i,j=1}^{N} \Big(G(q^i, q^j) + \frac{\delta^{ij}}{\lambda} \Big) \big\langle p_i, p_j \big\rangle_{\mathbb{R}^D}\,, \tag{6.2}$$

in a way that will be useful for our study. As usual, we assume that the kernel G of admissible Hilbert space V has the form $G(x, y) = \gamma\big(\|x - y\|_{\mathbb{R}^D}\big)$ for some function $\gamma : [0, \infty) \to \mathbb{R}$.

Proposition 6.1. *For any choice of the smoothing parameter λ, the following scalar quantity is conserved:*

$$M(p,q) \triangleq \sum_{1\le i<j\le N} \Big[\gamma(0) + \frac{1}{\lambda} - \gamma(\varrho^{ij})\Big] \langle p_i, p_j\rangle_{\mathbb{R}^D} \,,$$

where $\varrho^{ij} = \|q^i - q^j\|_{\mathbb{R}^D}$.

Proof. We can manipulate the expression for the Hamiltonian as follows:

$$\begin{aligned}
\mathcal{H}(p,q) &= \frac{1}{2}\Big\{ \sum_{i=1}^{N} \Big(\gamma(0) + \frac{1}{\lambda}\Big) \langle p_i, p_i\rangle_{\mathbb{R}^D} + 2\sum_{i<j} \gamma(\varrho^{ij}) \langle p_i, p_j\rangle_{\mathbb{R}^D} \Big\} \\
&= \frac{1}{2}\Big\{ \Big(\gamma(0) + \frac{1}{\lambda}\Big) \Big[\sum_{i=1}^{N} \langle p_i, p_i\rangle_{\mathbb{R}^D} + 2\sum_{i<j} \langle p_i, p_j\rangle_{\mathbb{R}^D} \Big] \\
&\qquad - 2\sum_{i<j} \Big(\gamma(0) + \frac{1}{\lambda}\Big) \langle p_i, p_j\rangle_{\mathbb{R}^D} + 2\sum_{i<j} \gamma(\varrho^{ij}) \langle p_i, p_j\rangle_{\mathbb{R}^D} \Big\} \\
&= \frac{1}{2}\Big(\gamma(0) + \frac{1}{\lambda}\Big) \Big\langle \sum_{i=1}^{N} p_i, \sum_{j=1}^{N} p_j \Big\rangle_{\mathbb{R}^D} - \sum_{i<j} \Big[\gamma(0) + \frac{1}{\lambda} - \gamma(\varrho^{ij})\Big] \langle p_i, p_j\rangle_{\mathbb{R}^D} \,;
\end{aligned}$$

but the first term on the right-hand side is conserved by the conservation of linear momentum. Since the Hamiltonian is conserved, so in the second term of the above expression and this completes the proof. □

Remark. We have not proven a “new” conservation law. The above proposition is a consequence of the conservation of the Hamiltonian function (6.2) and the conservation of linear momentum (Proposition 3.5).

For now on we shall assume, like we did in the previous chapter, that $\lambda = \infty$.

2. Dynamics of two one-dimensional landmarks

In the case of two landmarks on the real line Proposition 6.1 immediately implies the following result.

Corollary 6.2. *For $\lambda = \infty$ (exact matching), in the case of two one-dimensional landmarks the scalar quantity $\big[\gamma(0) - \gamma(\varrho^{12})\big] p_1 p_2$ is conserved.*

The conservation of linear momentum and the above result allow as to write the following expressions for the sum and product of momenta p_1 and p_2:

$$p_1(t) + p_2(t) = p_1(0) + p_2(0), \tag{6.3}$$

$$p_1(t)\, p_2(t) = \frac{\gamma(0) - \gamma\big(\varrho^{12}(0)\big)}{\gamma(0) - \gamma\big(\varrho^{12}(t)\big)}\, p_1(0)\, p_2(0)\,, \tag{6.4}$$

for all time t. Thanks to the above equations the evolution of momenta can be completely solved in function of the evolution of the mutual distance $\varrho^{12}(t)$ between the two landmarks and the initial distance $\varrho^{12}(0)$; in particular, when the asymptotic behavior of $\varrho^{12}(t)$ for $t \to \infty$ is known, equations (6.3) and (6.4) allow one to infer the asymptotic behavior of the momenta, as a function of the initial distance $\varrho^{12}(0)$.

Hamilton's equations (6.1) provided by Proposition 3.1 simplify to the following:

$$\dot{q}^1 = \gamma(0)\, p_1 + \gamma(\varrho^{12})\, p_2 \tag{6.5a}$$

$$\dot{q}^2 = \gamma(\varrho^{12})\, p_1 + \gamma(0)\, p_2 \tag{6.5b}$$

$$\dot{p}_1 = -f\big(q_1, q_2\big)\, p_1 p_2$$

$$\dot{p}_2 = -f\big(q_2, q_1\big)\, p_1 p_2 \;=\; -\dot{p}_1$$

with $\varrho^{12}(t) = |q_1(t) - q_2(t)|$ and $f(x, y) = \gamma'(|x - y|)\,\mathrm{sgn}(x - y)$ (see Definition 5.3). If we assume, as we do in the following, that $q_1 < q_2$ at all time, then $f(q_1, q_2) = -\gamma'(\varrho^{12})$ and the last two equations above can be rewritten as follows:

$$\dot{p}_1 = \gamma'\big(\varrho^{12}\big)\, p_1 p_2 \tag{6.5c}$$

$$\dot{p}_2 = -\gamma'\big(\varrho^{12}\big)\, p_1 p_2\,. \tag{6.5d}$$

Note also that the strong conservation law for momenta (Proposition 3.2) becomes:

$$\frac{\partial \varphi_t}{\partial \xi}\big(q_i(0)\big)\, p_i(t) = p_i(0), \qquad t \in [0, 1], \quad i = 1, 2. \tag{6.6}$$

Since $D = 1$ there is no conservation of angular momentum, whence from now on in this section by "weak conservation law" we will always mean the conservation of linear

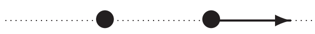
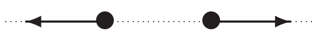
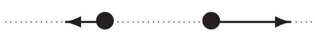
Case #	Initial momenta $p_1(0), p_2(0)$	
1.		$p_1 > 0,\ p_2 = 0$
2.		$0 < p_2 < p_1$
3.		$0 < p_1 = p_2$
4.		$0 < p_1 < p_2$
5.		$p_1 = 0,\ p_2 > 0$
6.		$p_1 > 0,\ p_2 < 0$, with $p_1 = \lvert p_2 \rvert$
7.		$p_1 > 0,\ p_2 < 0$, with $p_1 > \lvert p_2 \rvert$
8.		$p_1 < 0,\ p_2 > 0$, with $\lvert p_1 \rvert = p_2$
9.		$p_1 < 0,\ p_2 > 0$, with $\lvert p_1 \rvert < p_2$

TABLE 6.1. The nine cases of initial momenta $\big(p_1(0), p_2(0)\big)$.

momentum (6.3). We also note that since $q^2 > q^1$ for all t we have that $\varrho^{12} = q^2 - q^1$, whence the distance between landmarks satisfies the ordinary differential equation:

$$\frac{d}{dt}\,\varrho^{12} = \dot{q}^2 - \dot{q}^1 = \big[\gamma(0) - \gamma(\varrho^{12})\big](p_2 - p_1), \tag{6.7}$$

by (6.5a) and (6.5b).

As far as the dynamics of equations (6.5a)÷(6.5d) are concerned, we shall analyze the different cases listed in Table 6.1 (which are characterized by different *initial*

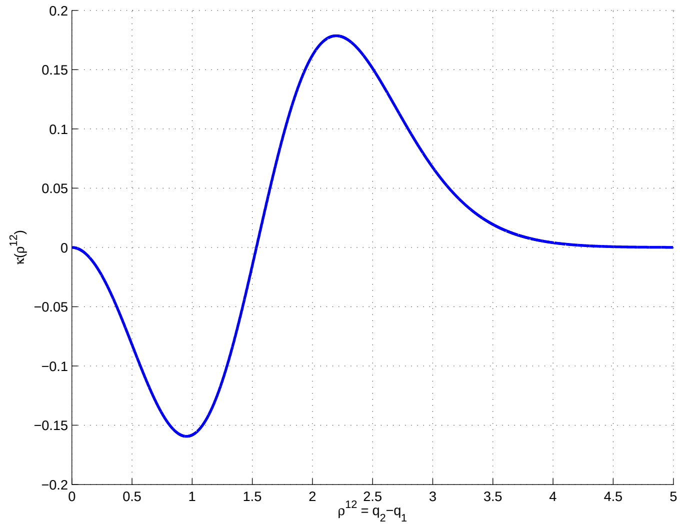


FIGURE 6.1. Sectional curvature $\kappa(\varrho^{12})$ for the Gaussian kernel.

values of the momenta p_1 and p_2) in various regions of the manifold, i.e. in areas with positive or negative curvature. From now on, we shall always assume that $q_1 < q_2$ (and since we assume that $\lambda = \infty$ such order cannot change in time). The qualitative behavior of the dynamical system in other cases not listed in Table 6.1 (e.g. $p_1 = 0$, $p_2 < 0$) can be inferred, by symmetry, from those listed.

We remind the reader that for the manifold of two one-dimensional landmarks sectional curvature κ, given by formula (5.17), only depends on ϱ^{12} and is plotted again for convenience in Figure 6.1 in the case of the Gaussian kernel (5.18) with unit variance. As we saw in the previous chapter function κ has a minimum, a zero (other than the one at $\varrho^{12} = 0$) and a maximum at points:

$$\varrho_{\mathrm{m}} = 0.953, \qquad \varrho_{\mathrm{z}} = 1.534, \qquad \varrho_{\mathrm{M}} = 2.198,$$

$$\kappa(\varrho_{\mathrm{m}}) = -0.1594, \qquad \kappa(\varrho_{\mathrm{z}}) = 0, \qquad \kappa(\varrho_{\mathrm{M}}) = 0.1786,$$

respectively. We have implemented numerically differential equations (6.5a)÷(6.5d) precisely in the case of the Gaussian kernel with unit variance.

In the several figures that follow, where $q_1(t)$ is plotted against $q_2(t)$, the thick diagonal line has equation $q_1 = q_2$ (or $\varrho^{12} = 0$). Since we assume that $q_2 > q_1$ the dynamics of the system take place *above* such line (see, for example, Figure 6.2: the details regarding such Figure will be discussed later; note also that the scale may change from figure to figure). Proceeding from bottom to top, the next (dashed) diagonal line represents the points where $q_2 - q_1 = \varrho_{\mathrm{m}}$, i.e. the points where curvature has its minimum; the next (continuous) diagonal line represents the points where $q_2 - q_1 = \varrho_{\mathrm{z}}$, i.e. the points where curvature has a zero; finally, the next (dash-dotted) diagonal line represents the points where $q_2 - q_1 = \varrho_{\mathrm{M}}$, i.e. the points where curvature has its maximum; above the latter line curvature is positive, and converges to zero from above as $q_2 - q_1 \to +\infty$. We shall now analyze the nine cases listed in Table 6.1 in some detail.

REMARK. Before proceeding, we should note some immediate consequences of the conservation laws that hold for *all* cases listed in Table 6.1. For example, when either one of the initial momenta is zero then by the strong conservation law (6.6) such momentum remains equal to zero for all time; if this is the case, then the other momentum is constant in time by the weak conservation law (6.3).

Also, note that for any t the map $\xi \mapsto \varphi_t(\xi)$ is a diffeomorphism, whence $\frac{\partial \varphi_t}{\partial \xi}(\xi) \neq 0$ for all pairs (t, ξ); but for any ξ the map $t \mapsto \frac{\partial \varphi_t}{\partial \xi}(\xi)$ is continuous and $\frac{\partial \varphi_t}{\partial \xi}\big|_{t=0} = 1$, therefore $\frac{\partial \varphi_t}{\partial \xi}(\xi) > 0$ for all pairs (t, ξ). Whence by the strong conservation law (6.6) if $p_i(0) \neq 0$ then $p_i(t)$ *never changes sign* in finite time and in particular *never becomes zero*. This implies, for example, that if the two landmarks collide into one another with opposite initial velocities then they will not “bounce” and escape to infinity in opposite directions, since this would imply a change of sign for both momenta; we shall provide the details of this instance later on (Case 6).

CASE 1 ($p_1 > 0$, $p_2 = 0$). By the strong conservation law (6.6) we have $p_2(t) \equiv 0$ for all $t \in [0, 1]$; a direct consequence of this, together with the weak conservation law (6.3), is that $p_1(t) \equiv p_1(0)$ for all $t \in [0, 1]$. By equations (6.5a) and (6.5b) the first landmark moves to the right with *constant* speed $\dot{q}^1(t) = \gamma(0)p_1(0)$ and “pushes” the second landmark, whose speed increases as ϱ^{12} decreases since $\dot{q}^2(t) = \gamma(\varrho^{12})p_1(0)$ and γ is monotone decreasing. The whole picture is provided by the following result.

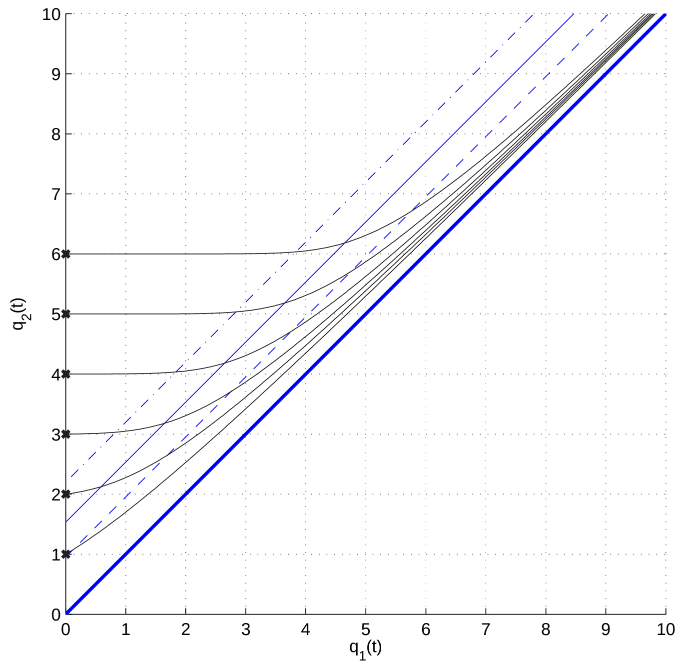


FIGURE 6.2. Six trajectories in Case 1: $\big(q_1(0), q_2(0)\big) = (0,1), (0,2), \ldots, (0,6)$ and, in all six cases, $\big(p_1(0), p_2(0)\big) = (10,0)$.

PROPOSITION 6.3. *Under the hypotheses of Case 1, i.e. $p_1(0) > 0$ and $p_2(0) = 0$, we have that $p_1(t) \equiv p_1(0)$, $p_2(t) \equiv 0$, $q^1(t) \to \infty$ and $q^2(t) \to \infty$ with $\varrho^{12}(t) \to 0$ as $t \to \infty$. Also, $\dot{q}^1(t) \equiv \gamma(0)p_1(0)$ for all t and $\dot{q}^2(t) \to \gamma(0)p_1(0)$ as $t \to \infty$.*

PROOF. Equation (6.7) becomes

$$\frac{d}{dt}\varrho^{12} = -\big[\gamma(0) - \gamma(\varrho^{12})\big]p_1(0) < 0,$$

so $\varrho^{12}(t)$ is a monotone decreasing and positive function; whence it converges. This implies that $\frac{d}{dt}\varrho^{12}(t) \to 0$ as $t \to \infty$, so that, by the above equation, $\gamma\big(\varrho^{12}(t)\big) \to \gamma(0)$; therefore $\varrho^{12}(t) \to 0$ by the monotonicity of function γ. The rest of the proof descends directly from the equations (6.5a) and (6.5b). $\square$

Typical trajectories in the (q_1, q_2) plane are shown in Figure 6.2.

CASE 2 $(0 < p_2 < p_1)$. In this case both landmarks initially move to the right, with the first one moving faster than the second one and therefore approaching it. It is not the case anymore that either $p_1(t)$ or $p_2(t)$ remain constant, but by the weak conservation law their summation $p_1(t)+p_2(t)$ does. The interesting phenomenon that happens is that the second landmark "bounces" off the first one, and *the two momenta are eventually "swapped" between the two*; we will later provide a rigorous justification of such a behavior. This is illustrated in Figure 6.3: the top graph shows the time evolution of $q_1(t)$ and $q_2(t)$, whereas the bottom portion illustrates the corresponding time evolution of momenta $p_1(t)$ and $p_2(t)$; note that $p_1(t)+p_2(t)$ is constant in time. Figure 6.4 shows the same trajectory in the (q_1, q_2) plane; since initial velocities are eventually swapped the final slope of the curve is the inverse of the initial slope. The following result holds.

PROPOSITION 6.4. *Under the hypotheses of Case 2, i.e.* $0 < p_2(0) < p_1(0)$, *we have that* $q^1(t) \to \infty$, $q^2(t) \to \infty$ *and* $\varrho^{12}(t) \to \infty$ *as* $t \to \infty$. *Also,*

$$\lim_{t\to\infty} p_1(t) = \lim_{t\to-\infty} p_2(t), \qquad \lim_{t\to\infty} p_2(t) = \lim_{t\to-\infty} p_1(t),$$

i.e. the two momenta are swapped.

PROOF. Differential equation (6.7) holds for the mutual distance ϱ^{12}:

$$\frac{d}{dt}\varrho^{12} = \big[\gamma(0)-\gamma\big(\varrho^{12}\big)\big](p_2-p_1).$$

Factor $[\gamma(0)-\gamma(\varrho^{12})]$ is always positive, so that the sign of the right-hand side of (6.7) depends only on the difference p_2-p_1, and is initially negative; whence $\varrho^{12}(t)$ initially decreases. By the strong conservation law momenta $p_1(t)$ and $p_2(t)$ are always positive in finite time therefore

$$\dot p_1 = \gamma'\big(\varrho^{12}\big)\, p_1 p_2 < 0,$$
$$\dot p_2 = -\gamma'\big(\varrho^{12}\big)\, p_1 p_2 > 0,$$

i.e. $p_1(t)$ and $p_2(t)$ are respectively monotone decreasing and increasing. But we have that $p_1(t) > 0$, so it will converge so some positive value $p_1(\infty)$; on the other hand $p_2(t) < p_1(0)+p_2(0)$ so it will also converge to some positive value $p_2(\infty)$.

We now claim that there exists a finite time t^* such that $p_1(t^*) = p_2(t^*)$. We will reason *by contradiction*, assuming that $p_1(t) > p_2(t)$ for all $t > 0$; in particular,

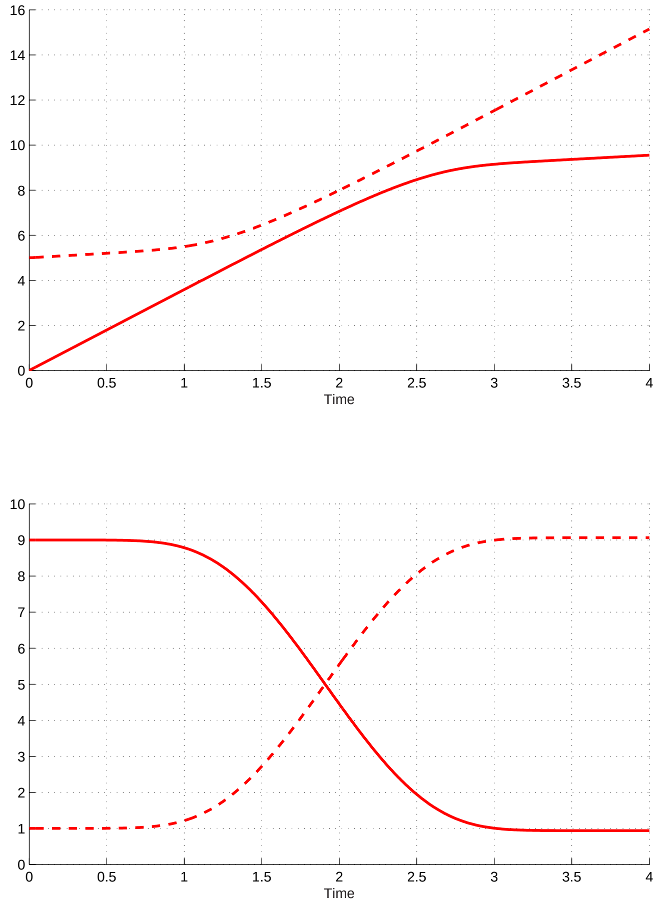


FIGURE 6.3. Trajectory in Case 2: $\big(q_1(0), q_2(0)\big) = (0, 1)$, $\big(p_1(0), p_2(0)\big) = (9, 1)$. The top and bottom graphs represent, respectively, the evolution of positions and momenta versus time. Note that the second landmark (dashed line) bounces off the first one (continuous line) and eventually the two momenta are swapped.

we assume that $p_1(\infty) \geq p_2(\infty) > 0$. If this is true $\varrho^{12}(t)$ is a monotone decreasing function for $t \geq 0$ by (6.7) and it must necessarily be the case that $\varrho^{12}(t) \to 0$ as $t \to \infty$; in fact if $\varrho^{12}(t) \to D$, for some $D > 0$, then $\dot{p}_1(t) \to \gamma'(D)\, p_1(\infty)\, p_2(\infty) < 0$ and $\dot{p}_2(t) \to -\gamma'(D)\, p_1(\infty)\, p_2(\infty) > 0$ as $t \to \infty$, which would imply that the graphs of p_1 and p_2 cross, immediately reaching a contradiction. So our hypothesis implies that $\varrho^{12}(t) \to 0$.

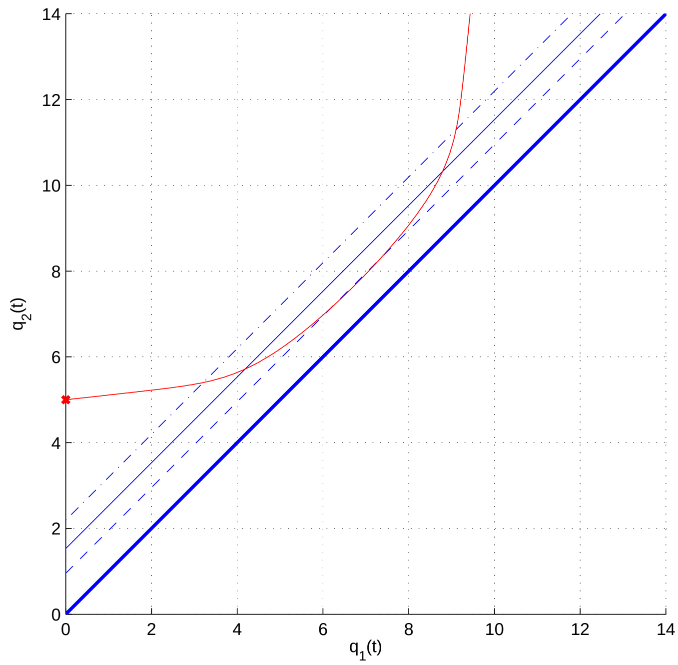


FIGURE 6.4. The same trajectory of Figure 6.3 drawn in the (q_1, q_2) plane.

Let us study the convexity of function $p_2(t)$. Differentiating equation (6.5d) and subsequently inserting equations (6.7), (6.5a) and (6.5b) yields:

$$
\begin{aligned}
\ddot{p}_2 &= -\gamma''\left(\varrho^{12}\right)\dot{\varrho}_{12}\,p_1p_2 - \gamma'\left(\varrho^{12}\right)\dot{p}_1p_2 - \gamma'\left(\varrho^{12}\right)p_1\dot{p}_2 \\
&= -\gamma''\left(\varrho^{12}\right)\left[\gamma(0) - \gamma\left(\varrho^{12}\right)\right](p_2 - p_1)p_1p_2 - \left[\gamma'\left(\varrho^{12}\right)\right]^2 p_1p_2^2 + \left[\gamma'\left(\varrho^{12}\right)\right]^2 p_1^2p_2 \\
&= \left\{\gamma''\left(\varrho^{12}\right)\left[\gamma(0) - \gamma\left(\varrho^{12}\right)\right] + \left[\gamma'\left(\varrho^{12}\right)\right]^2\right\} p_1p_2(p_1 - p_2). \qquad (6.8)
\end{aligned}
$$

By our hypothesis it must be the case that $p_1p_2(p_1 - p_2) > 0$, so that the sign of $\ddot{p}_2$ is determined by the factor in the curly braces; note that the first term in the braces is negative near zero and positive away from zero (this is determined by the sign of γ'') while the second term is always positive. However, since $\varrho^{12}(t) \to 0$ as $t \to \infty$, we are allowed to consider the following approximations:

$$
\begin{aligned}
\gamma(0) - \gamma(\varrho^{12}) &= -\frac{1}{2}\,\gamma''(0)\,(\varrho^{12})^2 + o\big((\varrho^{12})^2\big) && \text{since } \gamma'(0) = 0, \\
\gamma''(\varrho^{12}) &= \gamma''(0) + o(\varrho^{12}) && \text{since } \gamma'''(0) = 0, \\
\gamma'(\varrho^{12}) &= \gamma''(0)\,\varrho^{12} + o(\varrho^{12}) && \text{since } \gamma'(0) = 0, \\
\big[\gamma'(\varrho^{12})\big]^2 &= \big[\gamma''(0)\big]^2\,(\varrho^{12})^2 + o\big((\varrho^{12})^2\big).
\end{aligned}
$$

Inserting the above into the expression in curly braces in (6.8) yields:

$$
\begin{aligned}
&\gamma''(\varrho^{12})\big[\gamma(0) - \gamma(\varrho^{12})\big] + \big[\gamma'(\varrho^{12})\big]^2 \\
&= -\frac{1}{2}\big[\gamma''(0)\big]^2(\varrho^{12})^2 + o\big((\varrho^{12})^2\big) + \big[\gamma''(0)\big]^2(\varrho^{12})^2 + o\big((\varrho^{12})^2\big) \\
&= \frac{1}{2}\big[\gamma''(0)\big]^2(\varrho^{12})^2 + o\big((\varrho^{12})^2\big),
\end{aligned}
$$

so that we may conclude that $\ddot{p}_2 > 0$ (and $\ddot{p}_1 = -\ddot{p}_2 < 0$) once ϱ^{12} reaches a neighborhood of zero. This implies that from that time onwards $p_2(t)$ is convex (and $p_1(t)$ is concave); in particular it cannot be that $p_2(t)$ converges to a finite limit (nor can $p_1(t)$), which is a contradiction.

The time t^* defined above is the "crossing point" of Figure 6.3; when this occurs by equation (6.7) the time derivative $\dot{\varrho}^{12}$ changes sign, so that the distance between the landmarks starts increasing (in the (q_1, q_2) plane, the angle between the tangent vector to the trajectory and the q_1 axis reaches 45°, and keeps increasing). We should now note that the evolution of $p_1(t)$ and $p_2(t)$ depends on positions $q_1(t)$ and $q_2(t)$ only through *their difference* $\varrho^{12} = q_2 - q_1$, i.e. if one is interested only in the evolution of momenta it is sufficient to consider (6.5c) and (6.5d) in conjunction with ordinary differential equation (6.7), with the appropriate initial conditions. We will now *reverse* *time* by defining $\tilde{\varrho}^{12}(s) = \varrho^{12}(-s)$, $\tilde{p}_1(s) = p_1(-s)$, $\tilde{p}_2(s) = p_2(-s)$. We have that

$$\begin{aligned}
\frac{d}{ds}\,\tilde{p}_1(s) &= -\frac{dp_1}{dt}\Big|_{t=-s} = -\gamma'\big(\varrho^{12}(-s)\big)\,p_1(-s)\,p_2(-s) = -\gamma'\big(\tilde{\varrho}^{12}(s)\big)\,\tilde{p}_1(s)\,\tilde{p}_2(s),\\
\frac{d}{ds}\,\tilde{p}_2(s) &= -\frac{dp_2}{dt}\Big|_{t=-s} = +\gamma'\big(\varrho^{12}(-s)\big)\,p_1(-s)\,p_2(-s) = +\gamma'\big(\tilde{\varrho}^{12}(s)\big)\,\tilde{p}_1(s)\,\tilde{p}_2(s),\\
\frac{d}{ds}\,\tilde{\varrho}^{12}(s) &= -\frac{d\varrho^{12}}{dt}\Big|_{t=-s} = -\big[\gamma(0) - \gamma\big(\varrho^{12}(-s)\big)\big]\big(p_2(-s) - p_1(-s)\big)\\
&= \big[\gamma(0) - \gamma\big(\tilde{\varrho}^{12}(s)\big)\big]\big(\tilde{p}_1(s) - \tilde{p}_2(s)\big),
\end{aligned}$$

so that the evolution of $\tilde{\varrho}^{12}$, $\tilde{p}_1$ and $\tilde{p}_1$ is determined by the following system of ordinary differential equations:

$$\begin{aligned}
\dot{\tilde{\varrho}}^{12} &= \big[\gamma(0) - \gamma\big(\tilde{\varrho}^{12}\big)\big](\tilde{p}_1 - \tilde{p}_2),\\
\dot{\tilde{p}}_1 &= -\gamma'\big(\tilde{\varrho}^{12}\big)\,\tilde{p}_1\tilde{p}_2\,,\\
\dot{\tilde{p}}_2 &= \gamma'\big(\tilde{\varrho}^{12}\big)\,\tilde{p}_1\tilde{p}_2\,,
\end{aligned}$$

which are the same as (6.7), (6.5c) and (6.5d), except that the roles of p_1 and p_2 are *exchanged*. This implies that starting from the same initial conditions at the "crossing point" of Figure 6.3 the evolution of $\tilde{p}_1(s)$, $s > 0$ will be the same as that of $p_2(t)$, $t > 0$ and the evolution of $\tilde{p}_2(s)$, $s > 0$ will be the same as that of $p_1(t)$, $t > 0$; that is, the evolution *in the past* of p_1 will be the same as the evolution *in the future* of p_2, and vice versa. In conclusion, $p_1(-\infty) = p_2(\infty)$ and $p_2(-\infty) = p_1(\infty)$, so that momenta are eventually swapped between the two landmarks. $\square$

It is especially interesting in this case to study trajectories in different regions of the (q_1, q_2) plane; specifically, in regions of *positive* and *negative curvature*. It should first be noted that the behavior described above (the "bouncing" and swapping of momenta) is common to all regions, therefore, independently of where the trajectory originates, since $\varrho^{12} \to \infty$ the point $\big(q_1(t), q_2(t)\big)$ will eventually end up in the region of positive curvature (characterized by $q_2 - q_1 > \varrho_z$). However, it turns out that trajectories that originate at a common point in the region of positive curvature may exhibit *conjugate points*, as illustrated in Figure 6.5. Note that a trajectory that originates at a point in the region of positive curvature may enter the region of negative

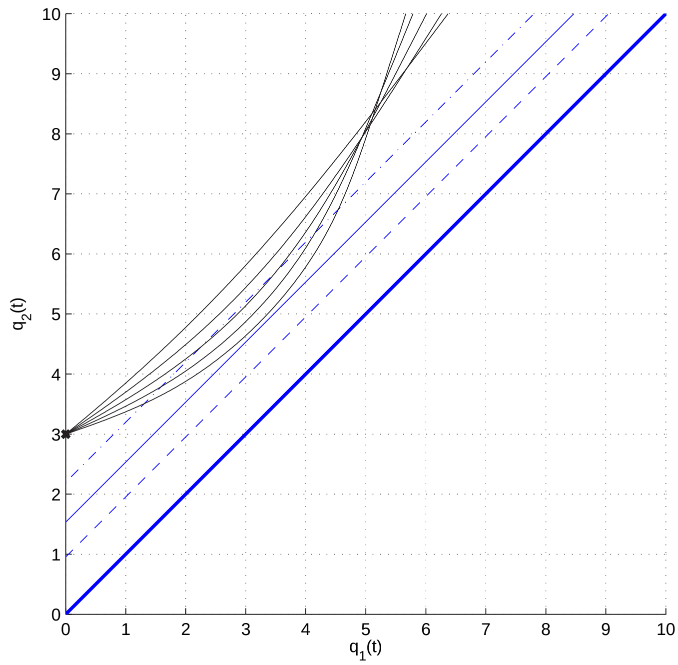


FIGURE 6.5. Conjugate points for trajectories that originate at a common point in the region of positive curvature (Case 2).

curvature (if, for example, $p_1 \gg p_2 > 0$) and then "bounce back" in the region of positive curvature, where it could have a conjugate point with another trajectory that originated at the same point. On the other hand, as illustrated in Figure 6.6, trajectories that originate at a common point in the region of *negative* curvature, never cross again within such region; however, they may meet again, i.e. have conjugate points, once they enter the region of positive curvature (Figure 6.7).

CASE 3 ($0 < p_1 = p_2$). The graphs and trajectories relative to this case can be inferred from those of Case 2. In fact, with reference to Figure 6.3, the trajectories start from the point where the momenta are the same (the "crossing point" of the bottom graph) and then are completely the same as in Case 2, with the two momenta settling to two final values $p_2 > p_1 > 0$ and both landmarks diverging (the second one faster than the first one).

For a fixed pair of (equal) initial momenta, the values of the final momenta *depend on the starting point* on the (q_1, q_2) plane, i.e. on the initial distance between the landmarks. More precisely, the following result holds.

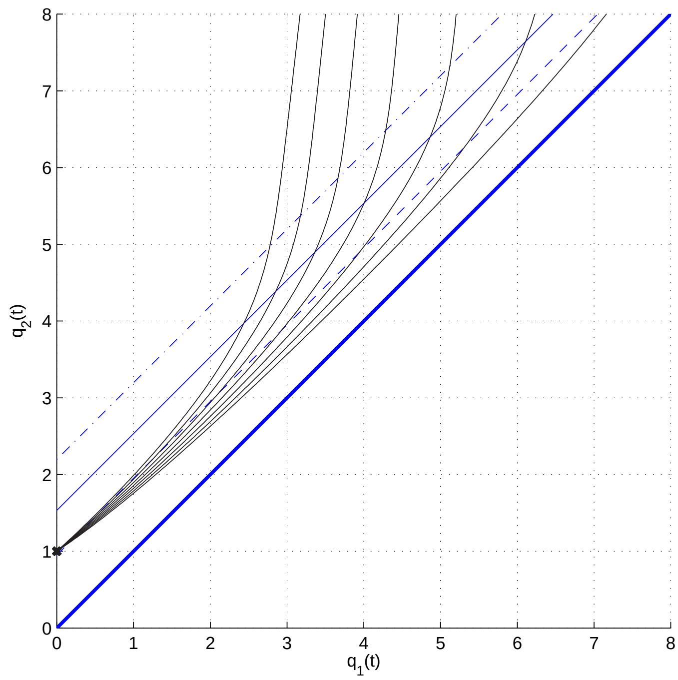


FIGURE 6.6. Trajectories that originate at the same point in the region of negative curvature (Case 2).

PROPOSITION 6.5. *Under the hypotheses of Case 3, i.e.* $0 < p_2(0) = p_1(0)$*, we have that* $q^1(t) \to \infty$*,* $q^2(t) \to \infty$ *with* $\varrho^{12}(t) \to \infty$ *as* $t \to \infty$*. Also,*

$$p_1(\infty) \triangleq \lim_{t\to\infty} p_1(t) = p_1(0)\left\{1 - \sqrt{\frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}}\right\} < p_1(0)$$

and

$$p_2(\infty) \triangleq \lim_{t\to\infty} p_2(t) = p_1(0)\left\{1 + \sqrt{\frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}}\right\} > p_2(0).$$

In particular if $\varrho^{12}(0) \simeq 0$ we have that $p_1(\infty) \simeq 0$ and $p_2(\infty) \simeq 2p_1(0)$, i.e. the transfer of momentum between the landmarks is almost complete.

PROOF. By the strong conservation law momenta must be positive for all time, whence by equations (6.5c) and (6.5d) we must have $\dot{p}_1 < 0$ and $\dot{p}_2 > 0$. So $p_1(t)$ increases and $p_2(t)$ decreases with time; consequently $p_2 - p_1 > 0$ for $t > 0$. The following equation holds:

$$\frac{d}{dt}\,\varrho^{12} = \big[\gamma(0) - \gamma(\varrho^{12})\big](p_2 - p_1) > 0;$$

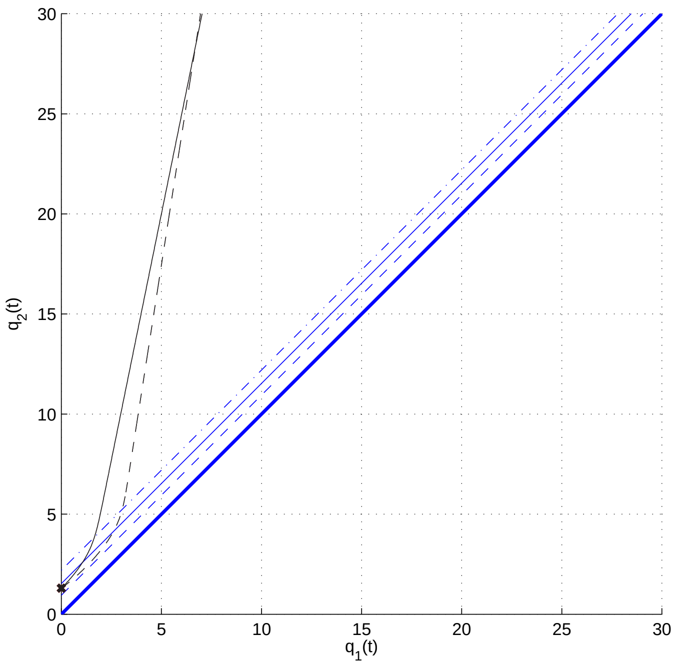


FIGURE 6.7. Trajectories that originate at the same point in the region of negative curvature and meet again in the region of positive curvature (Case 2). Dashed trajectory: $\big(q_1(0), q_2(0)\big) = (0, 1.3)$, $\big(p_1(0), p_2(0)\big) = (7, 3)$. Continuous trajectory: $\big(q_1(0), q_2(0)\big) = (0, 1.3)$, $\big(p_1(0), p_2(0)\big) = (5.01, 4.99)$.

since both factors on the right-hand side increase in time it must be the case that $\varrho^{12} \to \infty$ as $t \to \infty$. By (6.3) and (6.4) it is the case that

$$\begin{aligned} p_1(\infty) + p_2(\infty) &= 2p_1(0), \\ p_1(\infty)\, p_2(\infty) &= \left[1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}\right] p_1^2(0)\,, \end{aligned} \tag{6.9}$$

since $\lim_{\varrho\to\infty} \gamma(\varrho) = 0$. Combining the above equations implies

$$p_1^2(\infty) - 2p_1(0)p_1(\infty) + \left[1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}\right] p_1^2(0) = 0,$$

which in turn yields

$$p_1(\infty) = p_1(0)\left\{1 \pm \sqrt{\frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}}\right\};$$

the minus sign must be picked since $p_1(t)$ is a decreasing function. The result for $p_2(\infty)$ follows immediately from (6.9). □

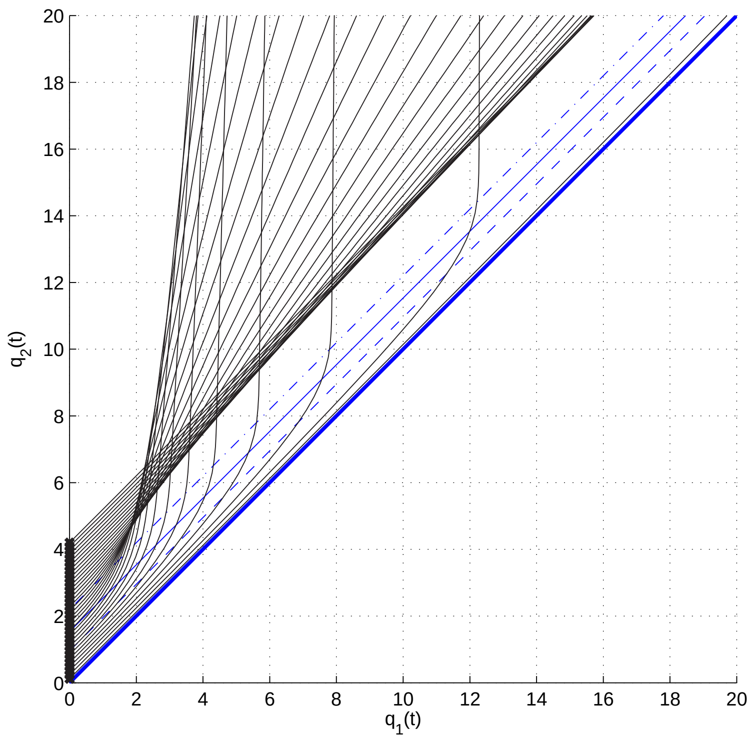


FIGURE 6.8. Trajectories with initial momenta $\big(p_1(0), p_2(0)\big) = (5, 5)$ but different initial positions (Case 3).

Figure 6.8 shows different landmark trajectories in the (q_1, q_2) plane which all have common initial momenta $\big(p_1(0), p_2(0)\big) = (5, 5)$ but different starting points $\big(q_1(0), q_2(0)\big) = (0, k \cdot 0.12)$, $k = 1, 2, \ldots, 35$.

CASE 4 $(0 < p_1 < p_2)$. Most of the relevant information on this case can be inferred from the previous graphs. For example, momenta evolve in time as in Figure 6.3, except that only the portion of the graph relative to some time after the "crossing point" of momenta should be considered. We should note that since $p_2 > p_1 > 0$ the initial tangent vectors to trajectories in the (q_1, q_2) plane form an angle with the q_1 axis that is greater than 45°.

PROPOSITION 6.6. *Under the hypotheses of Case 4, i.e.* $0 < p_1(0) < p_2(0)$, *we have that* $q^1(t) \to \infty$, $q^2(t) \to \infty$ *with* $\varrho^{12}(t) \to \infty$ *as* $t \to \infty$. *Also*

$$\lim_{t\to\infty} p_1(t) = \frac{p_1(0) + p_2(0)}{2}\left\{1 - \sqrt{1 - 4\Big[1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}\Big]\frac{p_1(0)p_2(0)}{[p_1(0) + p_2(0)]^2}}\right\}$$

and

$$\lim_{t\to\infty} p_2(t) \;=\; \frac{p_1(0)+p_2(0)}{2}\left\{1+\sqrt{1-4\Big[1-\frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}\Big]\frac{p_1(0)p_2(0)}{[p_1(0)+p_2(0)]^2}}\right\}.$$

Note that setting $p_1(0) = p_2(0)$ yields the result of Proposition 6.5.

PROOF. The fact that $\varrho^{12}(t) \to \infty$ as $t \to \infty$ is proven exactly as in Proposition 6.5. As far as momenta are concerned, we have that

$$p_1(\infty) + p_2(\infty) = p_1(0) + p_2(0),$$

$$p_1(\infty)\, p_2(\infty) = \left[1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}\right] p_1(0)p_2(0)\,,$$

since $\lim_{\varrho\to\infty} \gamma(\varrho) = 0$. Combining the above equations implies:

$$p_1^2(\infty) - \big[p_1(0) + p_2(0)\big]p_1(\infty) + \left[1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}\right] p_1(0)p_2(0) = 0.$$

Solving the above quadratic equation yields the result. □

Figure 6.9 shows two sets of eight trajectories. The two sets start from points $\big(q_1(0), q_2(0)\big) = (0, 1)$ and $(0, 5)$ respectively, and within each set the eight initial momenta are $\big(p_1(0), p_2(0)\big) = (4.5, 5.5)$, $(4, 6)$, $(3.5, 6.5)$, $(3, 7)$, $(2.5, 7.5)$, $(2, 8)$, $(1.5, 8.5)$, and $(1, 9)$, for both sets.

CASE 5 ($p_1 = 0$, $p_2 > 0$). As in Case 1, since one of the two momenta is initially equal to zero the conservation laws imply that both momenta remain constant at all time; whence $p_1(t) \equiv 0$ and $p_2(t) \equiv p_2(0)$ for all t. What happens is that if the two landmarks are initially close then the first one is "dragged" by the second one for a while, until the latter detaches itself and escapes to infinity, leaving the former virtually still. More precisely, the following holds.

PROPOSITION 6.7. *Under the hypotheses of Case 5, i.e.* $p_1(0) = 0$ *and* $p_2(0) > 0$, *we have that* $p_1(t) \equiv 0$, $p_2(t) \equiv p_2(0)$, $\dot q^2(t) \equiv \gamma(0)p_2(0)$ *and* $\dot q^1(t) \to 0$ *as* $t \to \infty$. *That is,* $q^1(t)$ *converges and* $q^2(t) \to \infty$ *as* $t \to \infty$.

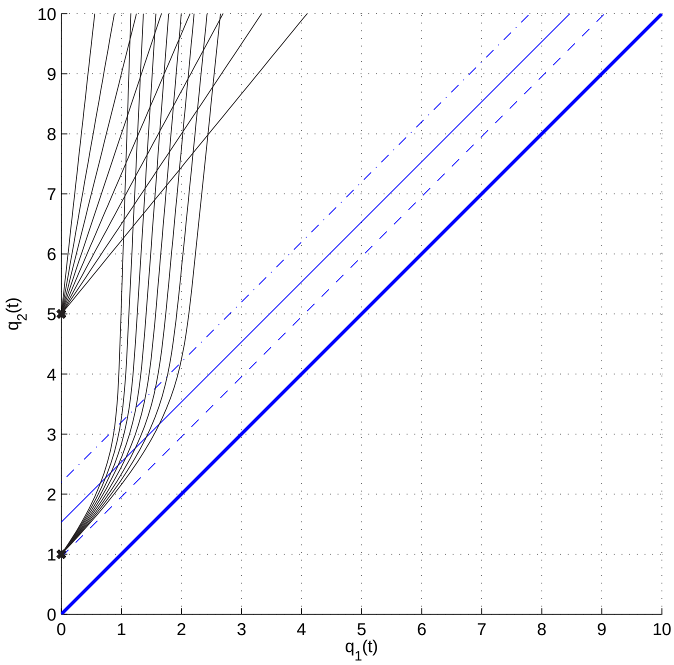


FIGURE 6.9. Two sets of trajectories for Case 4, which share common initial momenta but have two different starting positions.

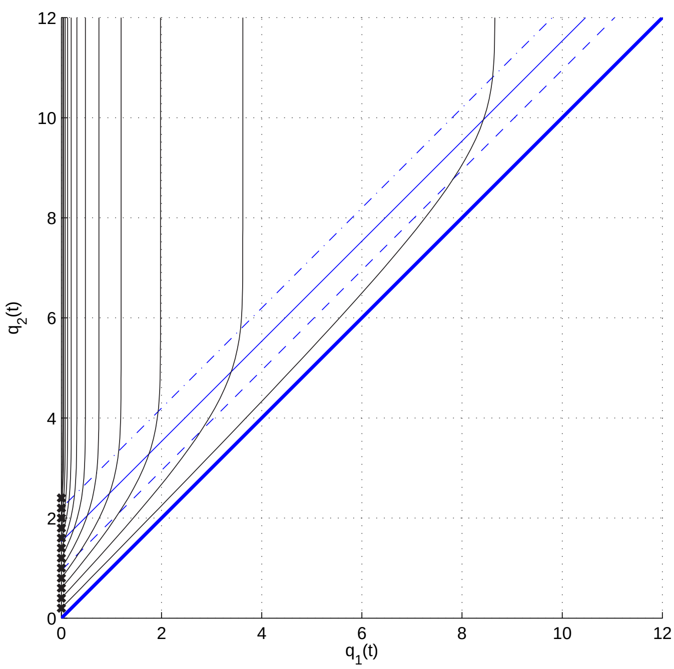


FIGURE 6.10. Trajectories for Case 5, all with initial momenta $\big(p_1(0), p_2(0)\big) = (0, 10)$ but different initial positions.

Proof. Equations (6.5a) and (6.5b) become, respectively, $\dot{q}^1 = \gamma(\varrho^{12})p_2(0)$ and $\dot{q}^2 = \gamma(0)p_2(0)$; the latter velocity is constant and positive, therefore $q_2(t) \to +\infty$ as t diverges. As far as the mutual distance ϱ^{12} is concerned, the ordinary differential equation

$$\frac{d}{dt}\,\varrho^{12} = \dot{q}^2 - \dot{q}^1 = \big[\gamma(0) - \gamma\big(\varrho^{12}\big)\big]p_2(0) > 0$$

holds; since the right-hand side is always positive function $\varrho^{12}(t)$ is monotone increasing. Whence $\dot{q}^1$ is monotone decreasing by (6.5a), i.e. the first landmark slows down in time; we will now argue that it comes to a virtual stop. In fact $\frac{d}{dt}\,\varrho^{12}(t) = \big[\gamma(0) - \gamma(\varrho^{12}(t))\big]p_2(0) > \big[\gamma(0) - \gamma\big(\varrho^{12}(0)\big)\big]p_2(0)$ for all t because $\varrho^{12}(t)$ is monotone increasing, which implies, by integration, that

$$\varrho^{12}(t) > \varrho^{12}(0) + \big[\gamma(0) - \gamma(\varrho^{12}(0))\big]p_2(0)\,t\,,$$

whence $\varrho^{12}(t)$ diverges as $t \to \infty$. When ϱ^{12} becomes large enough we have $\gamma(\varrho^{12}) \simeq 0$, therefore the dragging effect stops and the first landmark comes to a halt while the second diverges to infinity with constant speed. □

Figure 6.10 shows some typical trajectories for this case, with initial momenta all equal to $\big(p_1(0), p_2(0)\big) = (0, 10)$ and initial positions $\big(q_1(0), q_2(0)\big) = (0, k \cdot 0.2)$, $k = 1, \ldots, 12$; note that all trajectories eventually become virtually vertical.

Case 6 ($p_1 > 0$, $p_2 < 0$, with $p_1 = |p_2|$). What qualitatively happens in this case is that the two landmarks $q_1(t)$ and $q_2(t)$ converge towards each other *without bouncing back*, as we mentioned in a previous remark, while their momenta $p_1(t)$ and $p_2(t)$ *diverge* to plus and minus infinity, respectively.

Proposition 6.8. *Under the hypotheses of Case 6, i.e.* $0 < p_1(0) = -p_2(0)$, *it is the case that* $q^i(t) \to \frac{q^1(0)+q^2(0)}{2}$, *for* $i = 1, 2$, *and* $\varrho^{12}(t) \to 0$ *as* $t \to \infty$. *Also,* $p_1(t) \to \infty$ *and* $p_2(t) \to -\infty$ *as* $t \to \infty$.

Proof. Assuming $p_1(0) + p_2(0) = 0$, we have by the weak conservation law that $p_2(t) = -p_1(t)$ for all t. Therefore ordinary differential equations (6.5a)

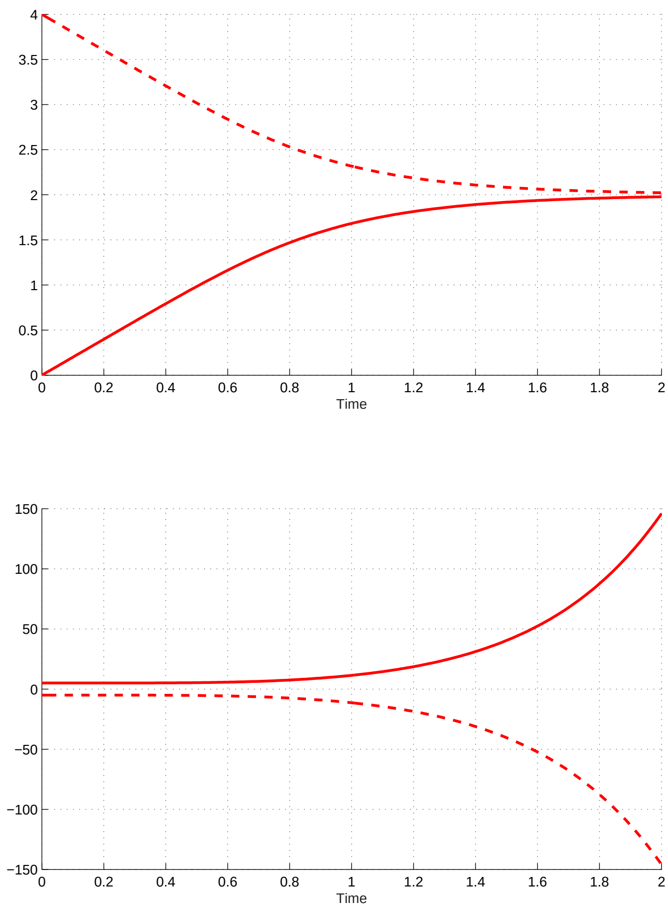


FIGURE 6.11. Evolution of positions (top) and momenta (bottom) versus time in Case 6. The continuous lines represent $q_1(t)$ and $p_1(t)$, while the dashed lines represent $q_2(t)$ and $p_2(t)$.

and (6.5b) may be written as:

$$\dot{q}^1 = \big[\gamma(0) - \gamma(\varrho^{12})\big] p_1(t), \tag{6.10a}$$

$$\dot{q}^2 = \big[\gamma(\varrho^{12}) - \gamma(0)\big] p_1(t); \tag{6.10b}$$

in particular, we note that $\dot{q}^2(t) = -\dot{q}^2(t)$ at all time. As far as the differential equations for momenta (6.5c) and (6.5d) are concerned, we have

$$\dot{p}_1 = -\gamma'(\varrho^{12})\, p_1^2(t) > 0, \tag{6.10c}$$

$$\dot{p}_2 = \gamma'(\varrho^{12})\, p_2^2(t) < 0; \tag{6.10d}$$

therefore $p_1(t)$ and $p_2(t)$ are monotone increasing and decreasing, respectively. Combining (6.10a) and (6.10b) yields the ordinary differential equation

$$\frac{d}{dt}\,\varrho^{12} = \dot{q}^2 - \dot{q}^1 = 2\big[\gamma(\varrho^{12}) - \gamma(0)\big]\,p_1(t); \tag{6.11}$$

the right-hand side is always *negative* since $\gamma(\varrho^{12}) < \gamma(0)$, $p_1(0) > 0$ and $p_1(t)$ is monotone increasing by (6.10c). This implies that the mutual distance $\varrho^{12}(t)$ is a monotone decreasing, continuously differentiable, lower bounded function; hence $\frac{d}{dt}\varrho^{12}(t) \to 0$ as $t \to \infty$. Since $p_1(t)$ is strictly positive and monotone increasing, by (6.11) we have that that $\gamma\big(\varrho^{12}(t)\big) \to \gamma(0)$, i.e., by the monotonicity of function γ, that $\varrho^{12}(t) \to 0$ as $t \to \infty$. In order to prove the divergence of the momenta we will use the formula on the product of momenta (6.3); since $p_1(t) = -p_2(t)$ at all time we have that

$$p_1(t) = \sqrt{\frac{\gamma(0) - \gamma\big(\varrho^{12}(0)\big)}{\gamma(0) - \gamma\big(\varrho^{12}(t)\big)}}\,p_1(0),$$

whence $p_1(t) \to +\infty$ (and $p_2(t) \to -\infty$) as $t \to \infty$. $\square$

Figure 6.11 shows exactly this behavior for initial positions $\big(q_1(0), q_2(0)\big) = (0, 4)$ and momenta $\big(p_1(0), p_2(0)\big) = (5, -5)$.

Case 7 ($p_1 > 0$, $p_2 < 0$, with $p_1 > |p_2|$). This case is similar to the previous one except that, since p_1 is initially larger (in absolute value) than p_2, when the collision occurs the "stronger" landmark prevails and the two eventually diverge together to $+\infty$, with their mutual distance $\varrho^{12}(t)$ quickly converging to zero (see Figure 6.13). Even though the second landmark eventually travels toward $+\infty$ its momentum never changes sign from negative to positive (by the strong conservation law) and in fact diverges to $-\infty$; $p_1(t)$ also diverges (to $+\infty$), but by the weak conservation law it is always the case that $p_1(t)$ maintains "the edge" against $p_2(t)$ (exactly by the difference $p_1(0) - |p_2(0)|$) so that the two landmarks finally travel together, at virtually constant positive speed.

Proposition 6.9. *Under the hypotheses of Case 7, i.e. $p_1(0) > 0$, $p_2(0) < 0$ with $p_1(0) > |p_2(0)|$, it is the case that $q^1(t) \to \infty$, $q^2(t) \to \infty$ with $\varrho^{12}(t) \to 0$*

as $t \to \infty$. Also, $p_1(t) \to \infty$, $p_2(t) \to -\infty$ and $\dot{q}^i(t) \to \gamma(0)[p_1(0) + p_2(0)]$, $i = 1, 2$, as $t \to \infty$.

Proof. Let $\Delta p \triangleq p_1(t) + p_2(t) = p_1(0) + p_2(0) > 0$; we can write

$$
\begin{aligned}
\frac{d}{dt}\,\varrho^{12} &= \bigl[\gamma(0) - \gamma\bigl(\varrho^{12}\bigr)\bigr](p_2 - p_1) \\
&= 2\bigl[\gamma(0) - \gamma\bigl(\varrho^{12}\bigr)\bigr]\,p_2 + \bigl[\gamma\bigl(\varrho^{12}\bigr) - \gamma(0)\bigr]\,\Delta p. \qquad (6.12)
\end{aligned}
$$

Since $\Delta p > 0$ and $p_2(t) < 0$ for all t both terms on the right-hand side of the above equation are always negative, therefore $\varrho^{12}(t)$ is a monotone decreasing, continuously differentiable function bounded from below, so that $\frac{d}{dt}\,\varrho^{12}(t) \to 0$ as $t \to \infty$; whence both terms on the right-hand side of (6.12) must converge to zero as $t \to \infty$. In particular $\gamma\bigl(\varrho^{12}(t)\bigr) \to \gamma(0)$ so that, by the monotonicity of γ, it must be the case that $\varrho^{12}(t) \to 0$. Now, combining equations (6.3) and (6.4) yields

$$
p_1^2(t) - \Delta p\, p_1(t) + \frac{\gamma(0) - \gamma\bigl(\varrho^{12}(0)\bigr)}{\gamma(0) - \gamma\bigl(\varrho^{12}(t)\bigr)}\, p_1(0)\, p_2(0) = 0,
$$

where we have used $p_2 = \Delta p - p_1$, which yields:

$$
p_1(t) = \frac{\Delta p}{2} \pm \sqrt{\Bigl(\frac{\Delta p}{2}\Bigr)^2 - \frac{\gamma(0) - \gamma\bigl(\varrho^{12}(0)\bigr)}{\gamma(0) - \gamma\bigl(\varrho^{12}(t)\bigr)}\, p_1(0)\, p_2(0)};
$$

the *plus* sign must be chosen since $p_1(t)$ is an increasing function. Analogously,

$$
p_2(t) = \frac{\Delta p}{2} - \sqrt{\Bigl(\frac{\Delta p}{2}\Bigr)^2 - \frac{\gamma(0) - \gamma\bigl(\varrho^{12}(0)\bigr)}{\gamma(0) - \gamma\bigl(\varrho^{12}(t)\bigr)}\, p_1(0)\, p_2(0)}.
$$

Taking limits for $t \to \infty$ proves the divergence of momenta. Finally, note that we can rewrite ordinary differential equation (6.5a) as follows:

$$
\dot{q}^1 = \gamma(0)\,\Delta p - \bigl[\gamma(0) - \gamma\bigl(\varrho^{12}\bigr)\bigr]\,p_2 = \gamma(0)\,\Delta p - \frac{\bigl[\gamma(0) - \gamma\bigl(\varrho^{12}(0)\bigr)\bigr]\,p_1(0)p_2(0)}{p_1},
$$

where we have used (6.4). Since $p_1(t) \to \infty$ as $t \to \infty$ we also have $\dot{q}^1(t) \to \gamma(0)\Delta p$. A similar reasoning holds for $\dot{q}^2(t)$. □

The situation described above is illustrated in Figure 6.12, where a the evolution of positions and momenta is shown for initial conditions $\bigl(q_1(0), q_2(0)\bigr) = (0, 4)$ and $\bigl(p_1(0), p_2(0)\bigr) = (8, -5)$. Note that in this particular instance the two landmarks are

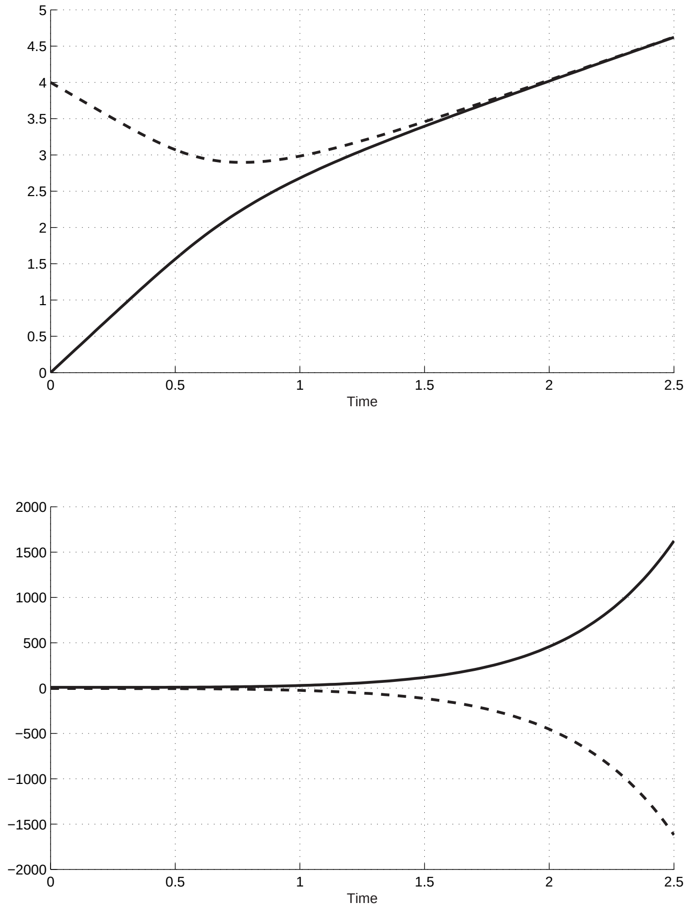


FIGURE 6.12. A typical trajectory for Case 7, plotted against time. Top: positions; bottom: momenta. Continuous line: first landmark; dashed line: second landmark.

initially relatively far, so that they initially travel at virtually constant speeds $\dot{q}^1 \simeq \gamma(0)p_1(0) > 0$ and $\dot{q}^2 \simeq \gamma(0)p_2(0) < 0$, with $|\dot{q}^2| < \dot{q}^1$; as they get closer they start feeling each other's influence, and eventually both move linearly together with positive speed. Figure 6.13 shows three sets of eight trajectories in the (q_1, q_2) plane. The three sets start at initial positions $\big(q_1(0), q_2(0)\big) = (0, 0.75)$, $(0, 2)$, and $(0, 4)$ respectively; within each set, the eight trajectories have initial momenta $\big(p_1(0), p_2(0)\big) = (5.5, -5)$, $(6, -5)$, $(6.5, -5)$, $(7, -5)$, $(7.5, -5)$, $(8, -5)$, $(8.5, -5)$, and $(9, -5)$. The farther away the two landmarks start from each other, the straighter are the trajectories initially; all trajectories eventually converge to the $q_1 = q_2$ diagonal line. It is perhaps useful

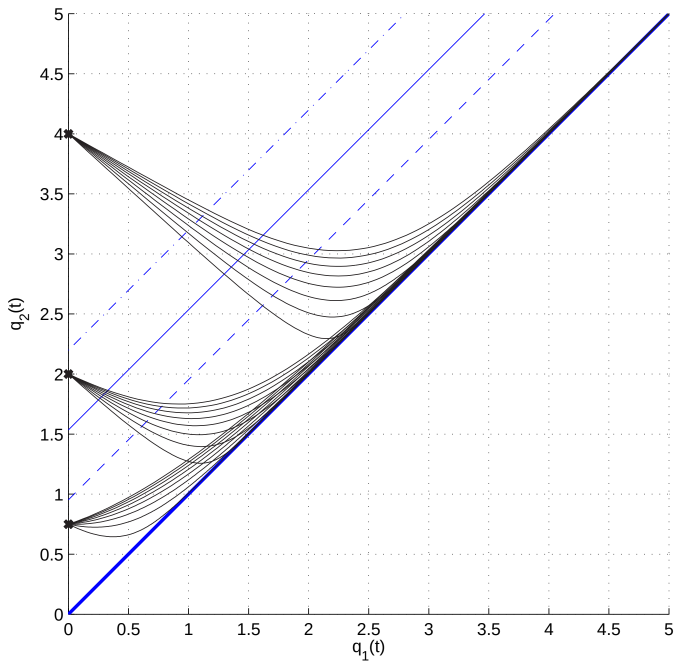


FIGURE 6.13. Three sets of trajectories in Case 7.

to remark that if it were the case that $p_2(0) = -p_1(0) < 0$ (Case 6) the trajectory in the (q_1, q_2) plane would be a straight line *perpendicular* to the $q_1 = q_2$ diagonal line.

CASE 8 ($p_1 < 0$, $p_2 > 0$, with $|p_1| = p_2$). In this situation the two landmarks are pulling away from each other, with equal (in absolute value) but opposite initial momenta. When the two landmarks are finally far away from each other, they break free and diverge with constant speed in opposite directions, having exchanged a certain amount of momentum prior to this.

PROPOSITION 6.10. *Under the hypotheses of Case 8, i.e.* $p_2(0) = -p_1(0) > 0$, *it is the case that* $q^1(t) \to -\infty$ *and* $q^2(t) \to +\infty$ *as* $t \to \infty$. *Also,*

$$\lim_{t\to\infty} p_i(t) = p_i(0)\sqrt{1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}}, \qquad i = 1, 2.$$

PROOF. Since $p_1(0) + p_2(0) = 0$, by the weak conservation law we will have that $p_1(t) + p_2(t) = 0$ for all t. By the strong law $p_1(t) < 0$ and $p_2(t) > 0$ for all time.

Equations (6.5a) and (6.5b) become, respectively:

$$\dot{q}^1 = \big[\gamma(\varrho^{12}) - \gamma(0)\big]\, p_2 < 0,$$

$$\dot{q}^2 = \big[\gamma(0) - \gamma(\varrho^{12})\big]\, p_2 = -\dot{q}^1 > 0,$$

so that $\dfrac{d}{dt}\,\varrho^{12} = \dot{q}^2 - \dot{q}^1 = 2\big[\gamma(0) - \gamma\big(\varrho^{12}\big)\big]\, p_2 > 0$. Therefore $\varrho^{12}(t)$ is an increasing function of time; we claim that $\varrho^{12}(t) \to \infty$.

Since $\varrho^{12}(t) > \varrho^{12}(0)$ and $\gamma(\cdot)$ is monotone decreasing, we have

$$\text{(6.13)} \qquad \frac{d}{dt}\,\varrho^{12}(t) = 2\big[\gamma(0) - \gamma\big(\varrho^{12}(t)\big)\big]p_2(t) > 2\big[\gamma(0) - \gamma\big(\varrho^{12}(0)\big)\big]p_2(t);$$

but by (6.4) and the fact that $p_1(t) = -p_2(t)$ we also have

$$p_2^2(t) = p_2^2(0)\frac{\gamma(0) - \gamma\big(\varrho^{12}(0)\big)}{\gamma(0) - \gamma\big(\varrho^{12}(t)\big)}\ ,$$

that is

$$\text{(6.14)} \qquad \begin{aligned} p_2(t) &= p_2(0)\sqrt{\frac{\gamma(0) - \gamma\big(\varrho^{12}(0)\big)}{\gamma(0) - \gamma\big(\varrho^{12}(t)\big)}} \\ &> p_2(0)\sqrt{1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}}. \end{aligned}$$

Therefore (6.13) becomes

$$\frac{d}{dt}\,\varrho^{12}(t) > 2\big[\gamma(0) - \gamma\big(\varrho^{12}(0)\big)\big]p_2(0)\sqrt{1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}};$$

the right-hand side is a positive number that does not depend on time, therefore $\varrho^{12}(t) \to \infty$ as $t \to \infty$. Finally, expression (6.14) implies

$$\lim_{t\to\infty} p_2(t) = p_2(0)\sqrt{1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}};$$

the result for $p_1(t)$ follows from $p_1(t) = -p_2(t)$. □

The situation is illustrated in Figure 6.14 for initial conditions $\big(q_1(0), q_2(0)\big) = (-0.15, 0.15)$, $\big(p_1(0), p_2(0)\big) = (-5, 5)$.

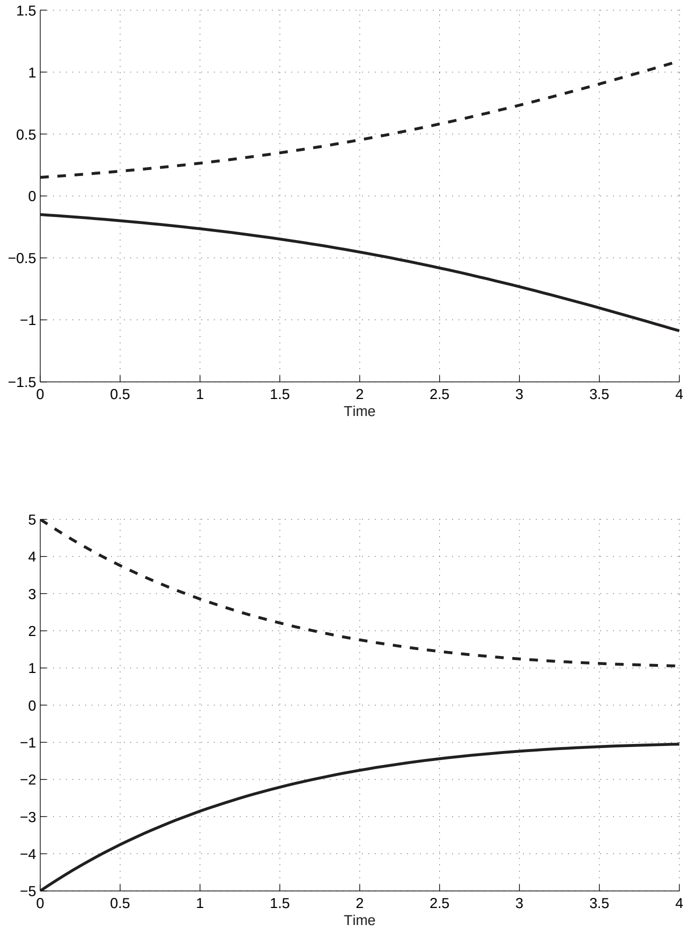


FIGURE 6.14. A typical trajectory for Case 8, plotted against time. Top: positions; bottom: momenta. Continuous line: first landmark; dashed line: second landmark. Note the exchange of momentum that occurs in the first part of the trajectory, when the two landmarks are relatively close.

CASE 9 ($p_1 < 0$, $p_2 > 0$, with $|p_1| < p_2$). This case is similar to the previous one, except that the second landmark "pulls to the right" more than the first one "pulls to the left", thus if the two start close enough to each other the former initially drags the latter to the right for a while, until the two eventually detach from each other. In the process some momentum is exchanged, as shown in Figure 6.15, which refers to the initial data $\big(q_1(0), q_2(0)\big) = (-0.1, 0.1)$ and $\big(p_1(0), p_2(0)\big) = (-5; 9)$.

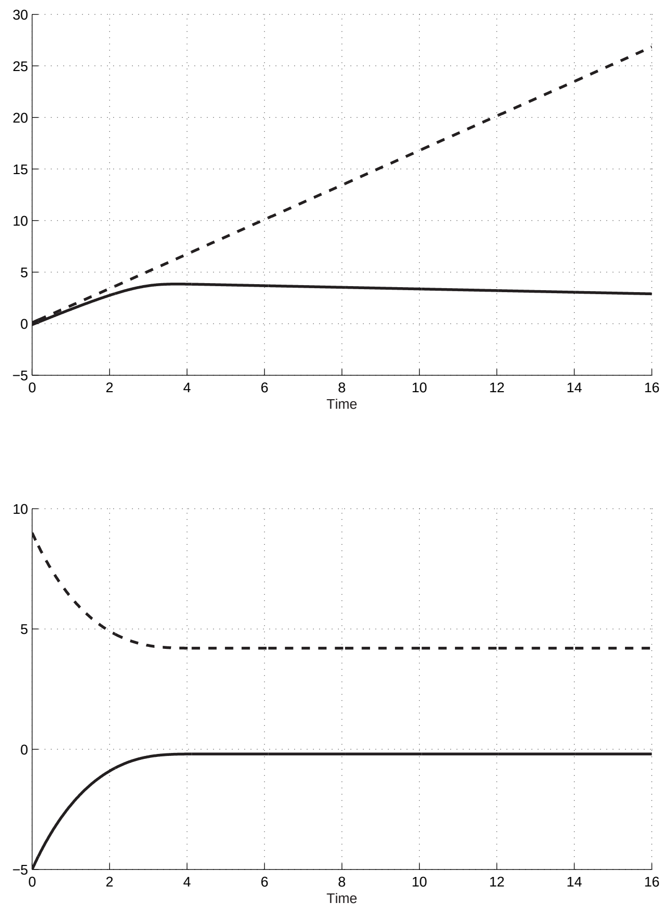


FIGURE 6.15. A typical trajectory for Case 9, plotted against time. Top: positions; bottom: momenta. Continuous line: first landmark; dashed line: second landmark.

PROPOSITION 6.11. *Under the hypotheses of Case 9, i.e.* $p_1(0) < 0$, $p_2(0) > 0$ *with* $|p_1(0)| < p_2(0)$, *it is the case that* $q^1(t) \to -\infty$ *and* $q^2(t) \to \infty$ *as* $t \to \infty$. *Also,*

$$\lim_{t\to\infty} p_1(t) = \frac{\Delta p}{2} + \sqrt{\left(\frac{\Delta p}{2}\right)^2 - \left[1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}\right] p_1(0)\, p_2(0)}$$

and

$$\lim_{t\to\infty} p_2(t) = \frac{\Delta p}{2} - \sqrt{\left(\frac{\Delta p}{2}\right)^2 - \left[1 - \frac{\gamma\big(\varrho^{12}(0)\big)}{\gamma(0)}\right] p_1(0)\, p_2(0)},$$

where $\Delta p \triangleq p_1(0) + p_2(0) = p_2(0) - |p_1(0)| > 0$.

PROOF. We have that $p_1(0) + p_2(0) = \Delta p > 0$. By the weak conservation law $p_2(t) = \Delta p - p_1(t)$ for all t. Equations (6.5a) and (6.5b) can be written as:

$$\dot{q}^1 = \gamma(0)\, p_1 + \gamma(\varrho^{12})\,(\Delta p - p_1) = \big[\gamma(0) - \gamma\big(\varrho^{12}\big)\big]\, p_1 + \gamma\big(\varrho^{12}\big)\,\Delta p$$

$$\dot{q}^2 = \gamma(\varrho^{12})\, p_1 + \gamma(0)\,(\Delta p - p_2) = \big[\gamma\big(\varrho^{12}\big) - \gamma(0)\big]\, p_1 + \gamma(0)\,\Delta p$$

so that $\frac{d}{dt}\,\varrho^{12} = 2\big[\gamma\big(\varrho^{12}\big) - \gamma(0)\big]\, p_1 + \big[\gamma(0) - \gamma\big(\varrho^{12}\big)\big]\,\Delta p > 0$; in fact, both terms on the right-hand side are positive; since $\varrho^{12}(t)$ is monotone increasing,

$$\frac{d}{dt}\,\varrho^{12} \geq \big[\gamma(0) - \gamma\big(\varrho^{12}(t)\big)\big]\,\Delta p \geq \big[\gamma(0) - \gamma\big(\varrho^{12}(0)\big)\big]\,\Delta p,$$

so that $\varrho^{12}(t) \geq \varrho^{12}(0) + \big[\gamma(0) - \gamma\big(\varrho^{12}(0)\big)\big]\,\Delta p \cdot t$, i.e. $\varrho^{12}(t) \to \infty$ as $t \to \infty$. Manipulating equations (6.3) and (6.4) yields

$$p_2^2(t) - \Delta p\, p_2(t) + \frac{\gamma(0) - \gamma\big(\varrho^{12}(0)\big)}{\gamma(0) - \gamma\big(\varrho^{12}(t)\big)}\, p_1(0)\, p_2(0) = 0,$$

whose solution is

$$p_2(t) = \frac{\Delta p}{2} + \sqrt{\Big(\frac{\Delta p}{2}\Big)^2 - \frac{\gamma(0) - \gamma\big(\varrho^{12}(0)\big)}{\gamma(0) - \gamma\big(\varrho^{12}(t)\big)}\, p_1(0)\, p_2(0)};$$

taking the limit for $t \to \infty$ completes the proof. $\square$

Figure 6.16 shows three sets of eight trajectories in the (q_1, q_2) plane. The three sets start at initial positions $\big(q_1(0), q_2(0)\big) = (-0.1, 0.1)$, $(-0.5, 0.5)$, $(-2.5, 2.5)$ respectively; within each set, the eight trajectories have initial momenta $\big(p_1(0), p_2(0)\big) =$ $(-5, 5.5)$, $(-5, 6)$, $(-5, 6.5)$, $(-5, 7)$, $(-5, 7.5)$, $(-5, 8)$, $(-5, 8.5)$, and $(-5, 9)$. Once again, the farther away the two landmarks start from each other, the straighter are the trajectories initially. In any case, all trajectories eventually straighten as the two landmarks get farther away from each other.

REMARK. After having analyzed in detail all nine cases listed in Table 6.1 we should note that we have implicitly shown that in the case of two landmarks in one dimension the dynamical system 6.1 admits *no closed orbits*. François-Xavier Vialard provides an alternative and elegant short proof of this fact [**44**]. However, it

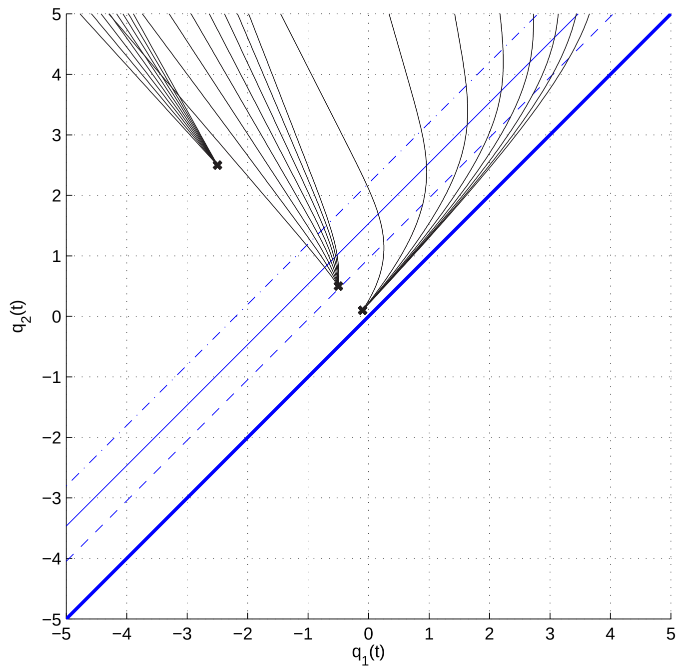


FIGURE 6.16. Three sets of trajectories for Case 9.

is currently not known weather such result can be extended to landmarks in $D \geq 2$ dimensions.

3. Dynamics of three one-dimensional landmarks

In this brief section we show the qualitative behavior of three one-dimensional landmarks for some sets of initial conditions and draw some comparison with the case of two landmarks in one dimension, which was thoroughly analyzed in the previous section. Hamilton's equations for three one dimensional landmarks are:

$$\dot{q}^1 = \gamma(0)\, p_1 + \gamma(\varrho^{12})\, p_2 + \gamma(\varrho^{13})\, p_3$$

$$\dot{q}^2 = \gamma(\varrho^{12})\, p_1 + \gamma(0)\, p_2 + \gamma(\varrho^{23})\, p_3$$

$$\dot{q}^3 = \gamma(\varrho^{13})\, p_1 + \gamma(\varrho^{23})\, p_2 + \gamma(0)\, p_3$$

$$\dot{p}_1 = +\gamma'\left(\varrho^{12}\right) p_1 p_2 + \gamma'\left(\varrho^{13}\right) p_1 p_3$$

$$\dot{p}_2 = -\gamma'\left(\varrho^{12}\right) p_1 p_2 + \gamma'\left(\varrho^{23}\right) p_2 p_3$$

$$\dot{p}_3 = -\gamma'\left(\varrho^{13}\right) p_1 p_3 - \gamma'\left(\varrho^{23}\right) p_2 p_3\,;$$

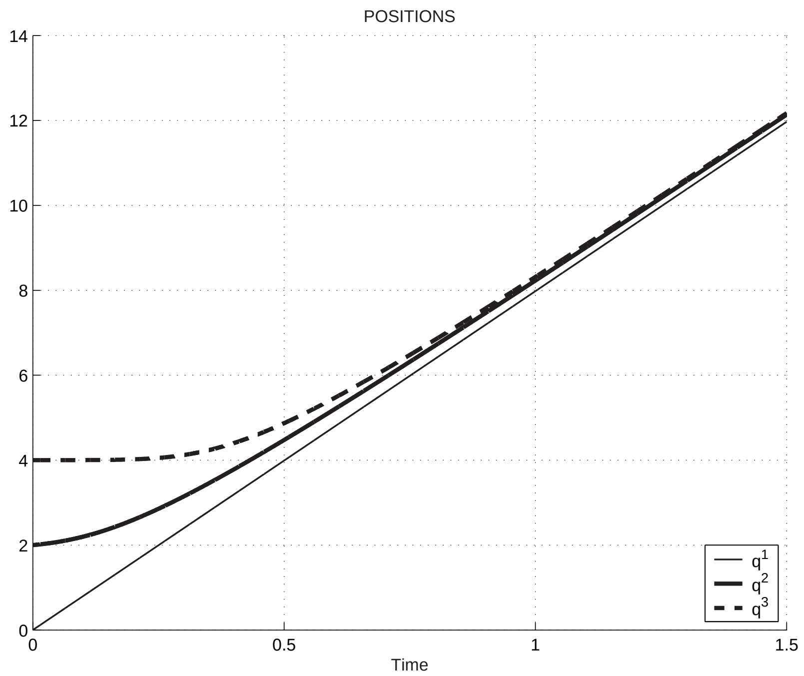
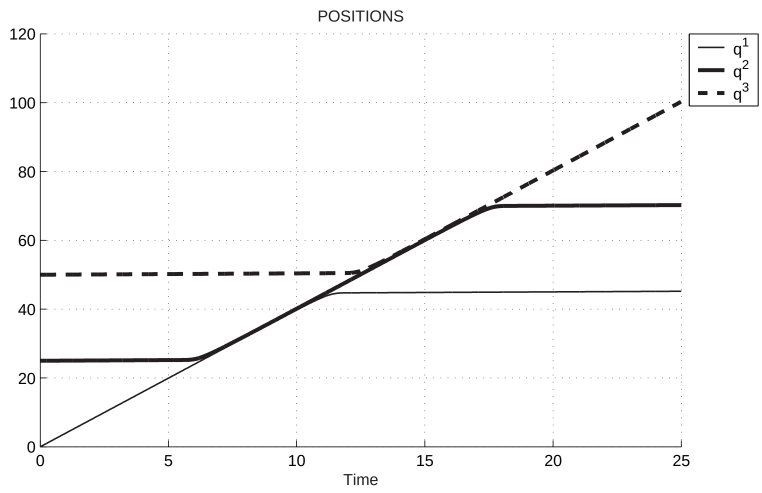


FIGURE 6.17. Positions and momenta versus time for Example 1 ($p_1(0) > 0$, $p_2(0) = p_3(0) = 0$). Landmark 1: thin line; Landmark 2: thick line; Landmark 3: thick dashed line.

we shall limit ourselves to illustrating three possible combinations of initial conditions for positions and momenta.

EXAMPLE 1. ($p_1(0) > 0$, $p_2(0) = p_3(0) = 0$) In this case, by the strong law, the momenta for the second and third landmarks are identically equal to zero in time and consequently $p_1(t) \equiv p_1(0)$ by the weak law. Therefore the velocity of the first landmark is constant it time, $\dot{q}^1 = \gamma(0)p_1(0)$, and the remaining two landmarks are "pushed" to $+\infty$ by the first one, without bouncing off; this type of behavior is completely analogous to the one of the two landmarks in one dimension of Case 1; see Figure 6.17.

EXAMPLE 2. ($p_1(0) \gg p_2(0) = p_3(0) > 0$) In this case all three landmarks have strictly positive momenta, so that "transfer" of momentum among them is possible, as it happened in Case 2 for two one-dimensional landmarks. The first landmark is initially much faster than the other two, and eventually bumps into the second one;



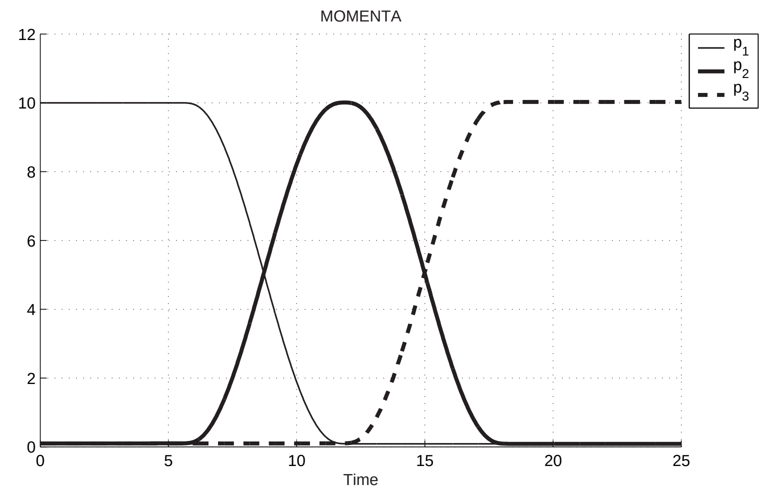


FIGURE 6.18. Positions and momenta versus time for Example 2 ($p_1(0) \gg p_2(0) = p_3(0) > 0$); large initial mutual distances. Landmark 1: thin line; Landmark 2: thick line; Landmark 3: dashed line.

momentum is swapped between the first and the second one, and later between the second one and the third one. Eventually the latter escapes to infinity at great speed, leaving the first and second one behind (at nonzero positive speed). Figures 6.18 and 6.19 refer, respectively, to the cases of large and small initial mutual distances; in the latter case the second landmark never gains “full” momentum, since this is transferred directly to the third one in an intermediate phase when all three landmarks travel closely to each other.

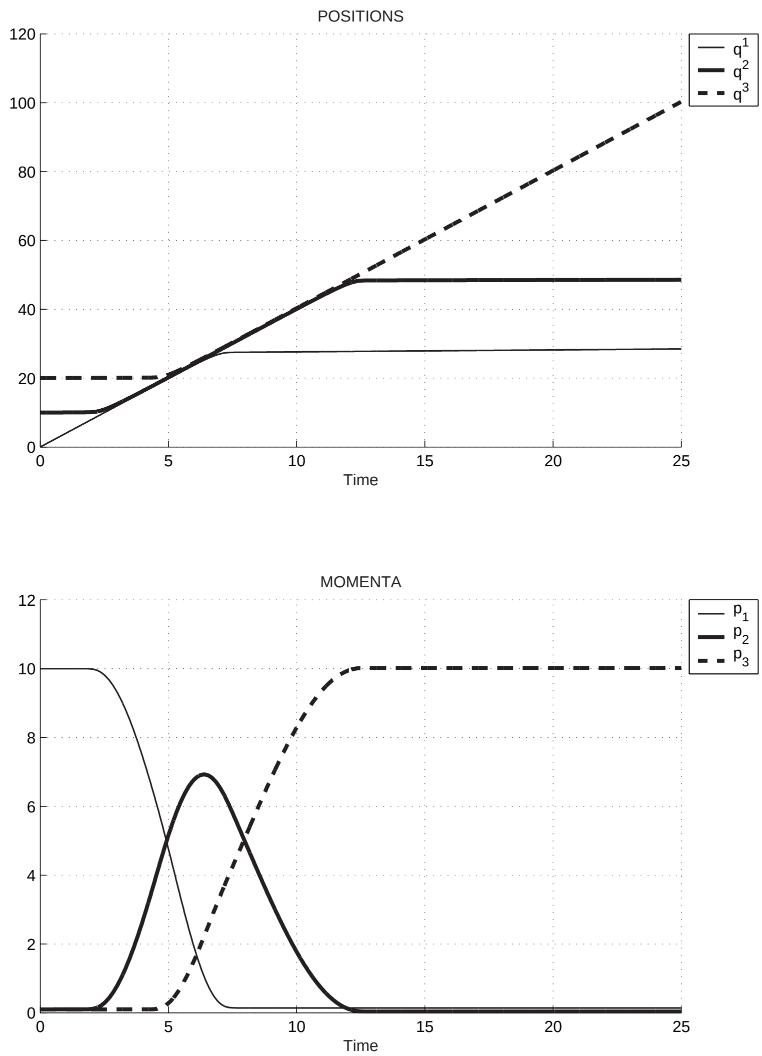


FIGURE 6.19. Positions and momenta versus time for Example 2 ($p_1(0) \gg p_2(0) = p_3(0) > 0$); small initial mutual distances. Landmark 1: thin line; Landmark 2: thick line; Landmark 3: dashed line.

EXAMPLE 3. ($p_1(0) > 0$, $p_2(0) < 0$, $p_3(0) < 0$, with $p_1(0) > |p_2(0) + p_3(0)|$) This case is analogous to Case 7 for two one-dimensional landmarks. The first landmark collides into the other two, which initially travel in the opposite direction. Since $p_1(0) > |p_2(0) + p_3(0)|$ the three eventually travel together towards $+\infty$; note that the three momenta diverge, to $+\infty$, $-\infty$ and $-\infty$ respectively. See Figure 6.20.

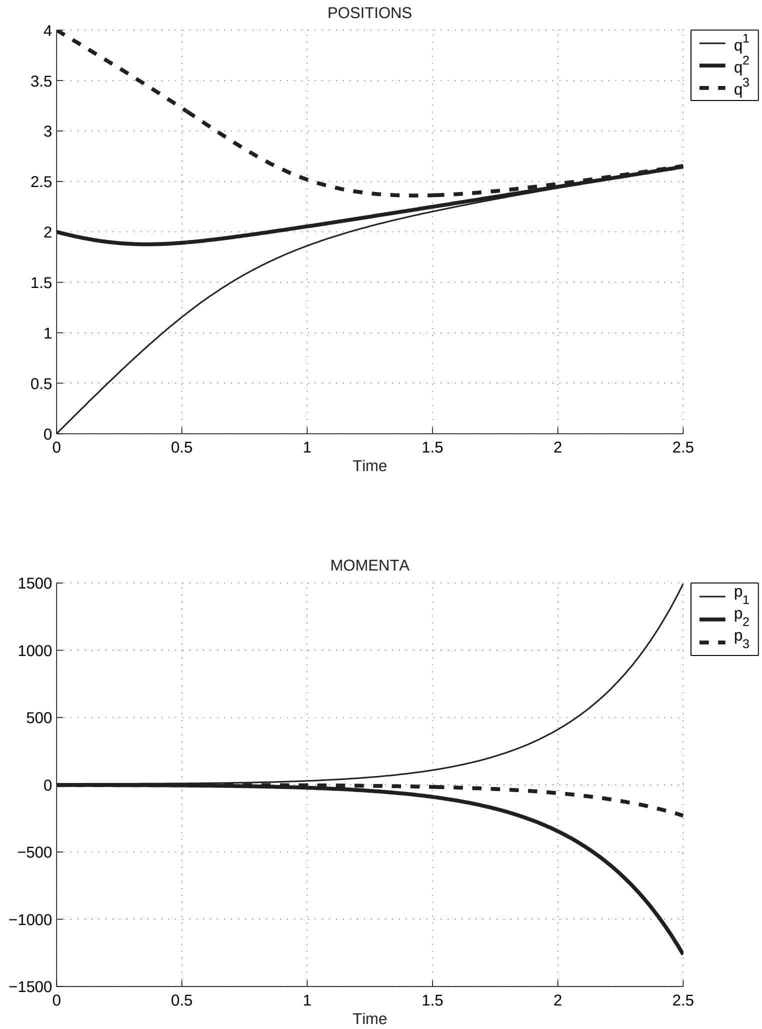


FIGURE 6.20. Positions and momenta versus time for Example 3 ($p_1(0) > 0$, $p_2(0) < 0$, $p_3(0) < 0$, with $p_1(0) > |p_2(0) + p_3(0)|$). Landmark 1: thin line; Landmark 2: thick line; Landmark 3: dashed line.

4. Dynamics of two two-dimensional landmarks

In this section we briefly analyze the qualitative behavior of two landmarks interacting with each other. We should say right away that while the dynamics of two landmarks are fully understood, it is not the case for two landmarks in two dimensions. Further research is needed, and here we shall limit ourselves to describe the observed qualitative phenomena.

If the two landmarks have initial momenta that lie on the same line, i.e. $p_1(0) = kp_2(0)$ for some $k \in \mathbb{R}$, then the problem is reduced to the one-dimensional case that was fully explored in section 2. For example, if

$$p_2(0) = -p_1(0) \tag{6.15}$$

and the momenta point "towards" each other, we are exactly in the situation described in Case 6 for two landmarks in one dimension. Since we are in two dimensions we are allowed to modify the *angle* of collision. For example Figure 6.21 shows the trajectories of two landmarks in two dimensions with initial positions

$$q^1(0) = (1, 0), \qquad q^2(0) = (-1, 0),$$

and initial momenta

$$p_1(0) = (-10, 8.6), \qquad p_2(0) = (10, -8.6)$$

(as usual, we are using a Gaussian kernel with unit variance); that is, situation (6.15) is perturbed by adding "opposite angles" to each of the initial momenta. As in case (6.15) the two landmarks eventually collide (in infinite time) by spinning around one another a countably infinite number of times. Figure 6.22 is a zoomed-in version of Figure 6.21 around the origin (note the different scale on the axes). As far as momenta are concerned, it is still the case that both $p_1(t)$ and $p_2(t)$ diverge as $t \to \infty$.

Remark. We should remind the reader that the landmark trajectories illustrated in Figures 6.21 and 6.22, despite being apparently complicated around the origin, are in fact generated by flows of diffeomorphisms of the plane. Such diffeomorphisms can be computed from the trajectories by implementing the techniques discussed in Chapter 2.

If we *increase* the value of the angle between the initial momenta beyond a certain threshold we have a *bifurcation* in the qualitative behavior of trajectories. This is illustrated in Figure 6.23 for initial positions

$$q^1(0) = (1, 0), \qquad q^2(0) = (-1, 0),$$

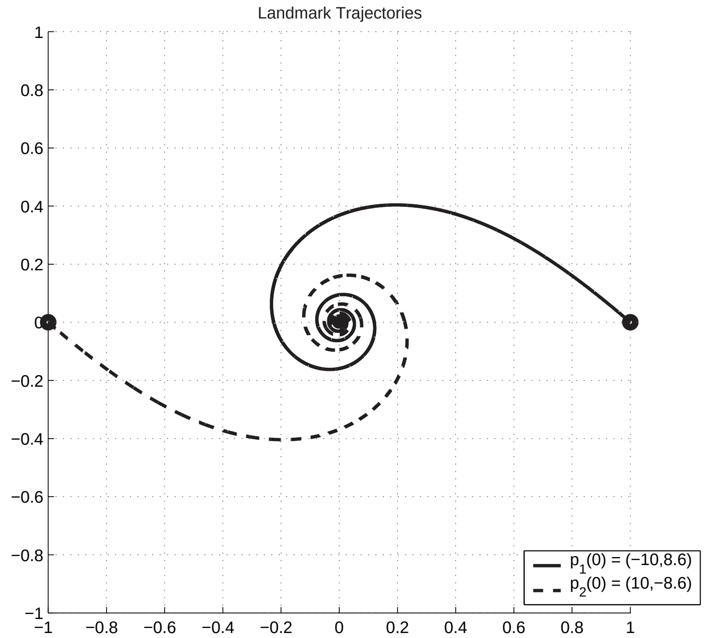


FIGURE 6.21. Converging trajectories for two landmarks in 2D.

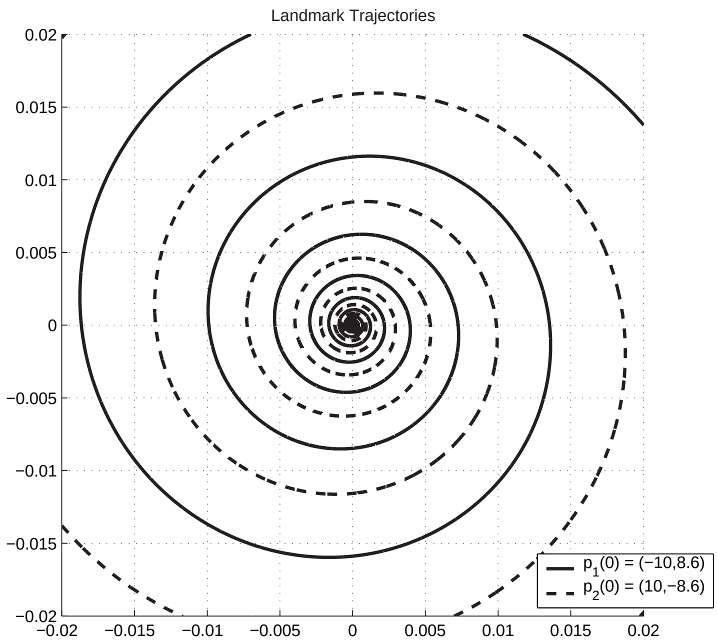


FIGURE 6.22. Converging trajectories for two landmarks in 2D—detail.

and initial momenta

$$p_1(0) = (-10, 9), \qquad\qquad p_2(0) = (10, -9).$$

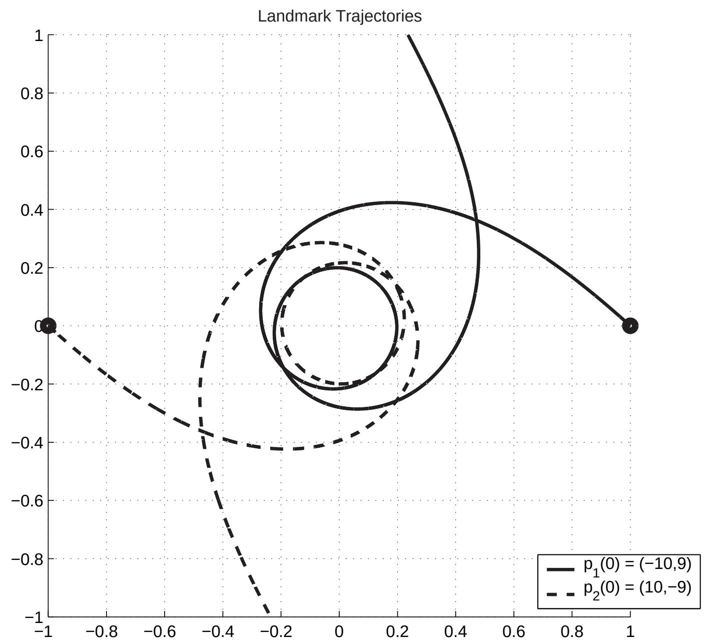


FIGURE 6.23. Diverging trajectories for two landmarks in 2D.

We observe that the two landmarks still spin around one another for a while, but eventually escape to infinity in opposite directions. The corresponding momenta (graphs not shown) eventually converge to finite values ($p_1(\infty) + p_2(\infty) = 0$ by the conservation of linear momentum) that determine the escape velocities of landmarks.

A similar behavioral pattern is observed when initial condition (6.15) is perturbed by adding an angle to each initial momentum "in the same direction"; this is depicted in Figure 6.24, which refers to initial positions

$$q^1(0) = (2, -4), \qquad\qquad q^2(0) = (-2, -4),$$

and initial momenta

$$p_1(0) = (-12, 10), \qquad\qquad p_2(0) = (12, 10);$$

in this case the two landmarks still converge and collide in infinite time. However, if the angle between the initial momenta is further increased beyond a certain threshold,

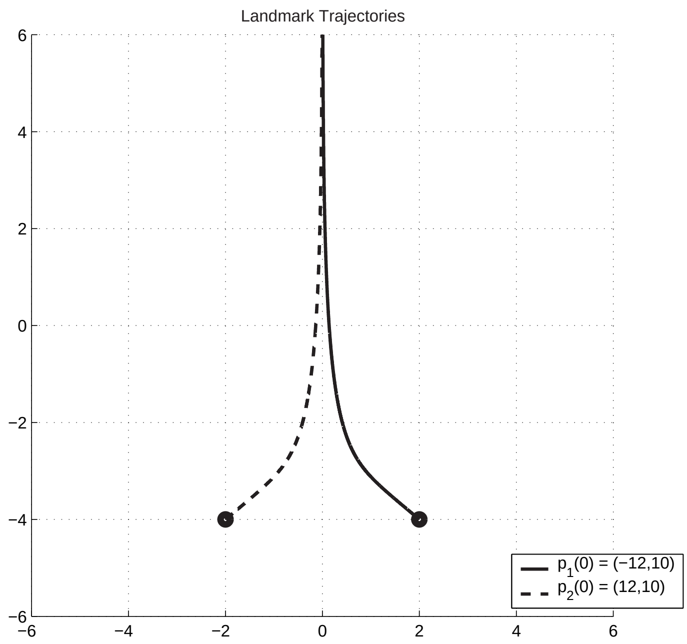


FIGURE 6.24. Converging trajectories for two landmarks in 2D.

e.g. if the initial momenta are chosen to be

$$p_1(0) = (-7, 10), \qquad\qquad p_2(0) = (7, 10)$$

with the same initial positions, then the qualitative behavior of Figure 6.25 is observed: the landmark trajectories eventually diverge.

As we mentioned at the beginning of this section we have not yet developed a rigorous explanation for the qualitative behavior of the dynamics of two landmarks in two dimensions. In particular, it would be of interest to find an analytical expression for the threshold "collision angle" beyond which the two landmarks eventually diverge, expressed in terms of the initial positions and the function γ that determines the kernel. Also, the question of whether periodic orbits exist in two dimensions is still an open one. The conservation of *angular* momentum is probably key to answering all these questions.

We conclude this section, and the chapter, by observing the effects of curvature on the qualitative dynamics of two landmarks on the plane: Figure 6.26 illustrates

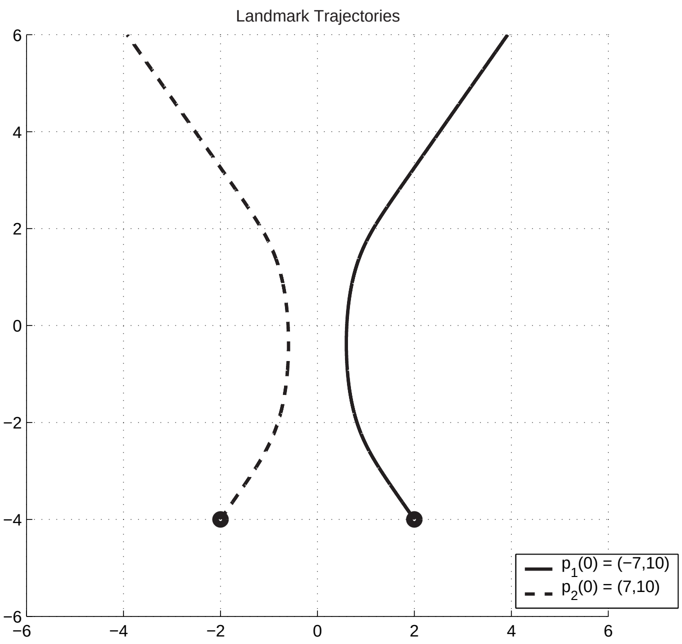


FIGURE 6.25. Diverging trajectories for two landmarks in 2D.

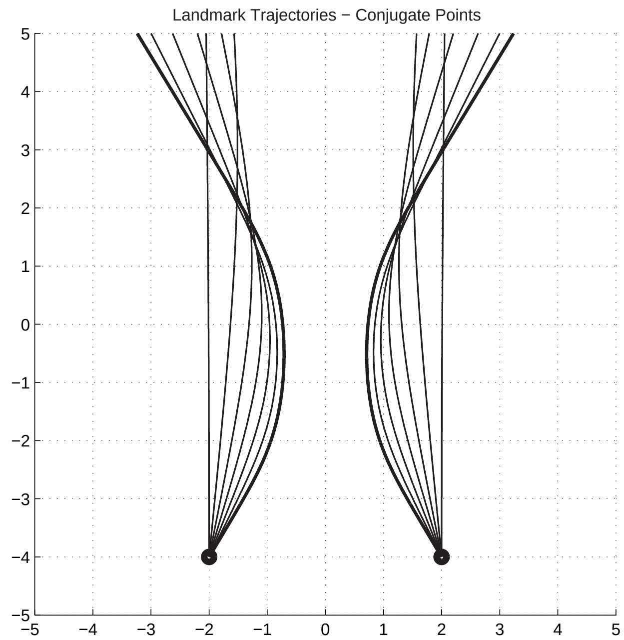


FIGURE 6.26. Existence of conjugate points for two landmarks in 2D.

the existence of *conjugate points* for trajectories originating at points

$$q^1(0) = (2, -4), \qquad\qquad q^2(0) = (-2, -4)$$

with different initial momenta, namely

$$p_1(0) = (-7 - k, 10), \qquad\qquad p_2(0) = (7 + k, 10),$$

for $k = 1, 2, \ldots, 7$ (the thicker line corresponds to the choice of $k = 1$). The above conditions correspond to *positive* sectional curvature at the point in question, and the existence of *conjugate points* is evident. The next chapter, which concludes this thesis, briefly illustrates the possible effects of the existence of conjugate points in the statistical analysis on shape spaces.

CHAPTER 7

Conclusions

1. Results

In this thesis we have endowed the set of N landmark points in D dimensions with the structure of Riemannian manifold: such structure derives directly from the notion that the distance between two shapes can be computed as the square root of the minimal energy that is necessary to match the first shape to the second one by means of a fluid flow. The metric tensor of the resulting DN-dimensional Riemannian manifold was made explicit, as well as the ordinary differential equations that determine the cogeodesic flow. Conservation laws, that follow from the translation-invariance and rotation-invariance of the metric tensor, were also explored.

Once the metric tensor of a generic Riemannian manifold is known, in principle one can compute the Riemannian curvature tensor by taking first and second partial derivatives of the elements of the metric and combining them. In our case following this procedure was unfeasible since the metric tensor is given by the inverse of a matrix; on the other hand, the structure of the matrix of partial derivatives of the cometric tensor happens to be very sparse, since each element of such tensor only depends on $2D$ out of the $n = DN$ coordinates. This suggested solving the problem of finding formulas for the Riemannian curvature tensor and sectional curvature that depend on the first and second derivatives of the elements of the cometric.

We then applied these formulas to computing the general expression of sectional curvature for landmark manifolds, both in one and D dimensions. In particular, we explored in great detail the simple but nonetheless very informative examples of two and three one dimensional landmarks: in the latter case we also identified the tangent 2-planes that correspond to maximum and minimum curvature in a number of different landmark configurations.

Last, but not least, we analyzed the effects that curvature has on the qualitative dynamics of landmarks. In particular, we verified the divergence of geodesics in regions of negative sectional curvature and the existence of conjugate points in regions of positive sectional curvature. These facts have important consequences in applications, as we will briefly discuss in the next section.

2. Applications and Future Work

Model (2.3) introduced in Chapter 2 and analogous ones are currently in use in the emerging discipline of computational anatomy [**10, 14, 15, 18, 22, 34, 35, 46**], specifically in analyzing Magnetic Resonance Imaging (MRI) data of the human brain. In such field one has the interest in building templates from data, i.e. models for a typically healthy brain and for a brain that is at a certain stage of a pathology, such as Alzheimer's disease, that modifies its geometry and structure in a characteristic manner. The ultimate goal of this exercise is the creation of diagnostic software.

Typically, statistical analysis on a data set $\mathcal{B} = \{J_1, \ldots, J_M\} \subset \mathcal{I}$ is performed as follows. First of all the so-called intrinsic (or Karcher's) mean [**38**] is computed:

$$m := \arg\min_{I \in \mathcal{I}} \sum_{J \in \mathcal{B}} d(I, J)^2, \tag{7.1}$$

where $d(\cdot, \cdot)$ is the geodesic distance. Then vectors $w_i \in T_m\mathcal{I}$, $i = 1, \ldots, M$ are found such that, for all i, $\exp_m(w_i) = J_i$ (the mean m is "shot" with initial velocity w_i "evolves" along a geodesic curve into datum J_i in unit time). At this point statistical analysis is done on the tangent space, for example by performing Principal Component Analysis (PCA, see [**5**]) on vectors $(w_1, \ldots, w_m)$ in the linear space $T_m\mathcal{I}$; equivalently, PCA can be performed on the cotangent vectors $(w_1^\flat, \ldots, w_m^\flat)$. in $T_m^*\mathcal{I}$. Note that statistical analysis should be ideally performed on the manifold *itself* using Principal Geodesic Analysis (PGA, see [**14, 15**]), which is however computationally unfeasible, whence such analysis is approximated with linear PCA on the tangent space at m.

Several remarks are in order. First of all, when the manifold is non-negatively curved and the data is not localized enough Karcher's mean (7.1) may not be unique (for example, the sphere has constant positive curvature and if the "data" happen to be precisely two opposite poles then any point on the corresponding equator solves the above minimization problem).

Also, *conjugate points* may exist in regions of positive curvature: in fact in the previous chapter we verified that this is precisely the case for landmark manifolds eve in low-dimensional settings. A consequence is that if the data points in $\mathcal{B}$ are not localized enough the vectors $w_i \in T_m\mathcal{I}$ such that $\exp_m(w_i) = J_i$ may not be well defined themselves (non-uniqueness).

Last, but not least, even when the Karcher mean is unique and the tangent vectors $w_i \in T_m\mathcal{I}$ are well defined, curvature causes *distortion* in the statistical analysis. For example geodesics that originate at the Karcher mean m with initial velocities $X, Y \in T_m\mathcal{I}$ locally *diverge* if $K(X, Y) < 0$, whereas they locally *converge* if $K(X, Y) > 0$. In any case, distortion is generated by the process of representing the data points on the tangent space at m: that is, in case of negative curvature two data points J_1 and J_2, whose corresponding vectors, respectively w_1 and w_2, appear close on $T_m\mathcal{I}$, may be far on the manifold $\mathcal{I}$ in terms of actual geodesic distance; on the other hand, in case of positive curvature the tangent vectors w_1 and w_2 could appear far on $T_m\mathcal{I}$ when in reality J_1 and J_2 are close on the manifold. Situations like these can potentially lead to inaccuracies in the statistical analysis.

Our current and future work aims precisely at estimating this distortion in the case of landmark shapes deriving from MRI databases of left and right hippocampi of three groups of patients: healthy patients, patients with Mild Cognitive Impairment (MCI, that corresponds to a Clinical Dementia Rating, or CDR, of approximately 0.5), and patients with Alzheimer's disease[1] (AD). For each of the three classes of patients the Karcher mean m is computed together with the tangent vectors $w_i \in T_m\mathcal{I}$ that correspond to the data, and PCA is performed on the tangent space. Now that a formula

[1]The three-dimensional landmark sets, which are hand-picked and hand-labeled from MRI images of the brain, are provided by the Center for Imaging Science at Johns Hopkins University.

for the sectional curvature of landmarks manifolds is known (see Chapter 5), we intend to compute sectional curvature for each pair of tangent vectors $w_i, w_j \in T_m\mathcal{I}$ and perform a study of the local distortion. Also, since the whole computation depends on the choice of the kernel (i.e. of the admissible Hilbert space V), quantifying how the choice of the kernel's parameters (e.g. the scaling factor) influence the distortion in the statistical analysis will certainly be of interest.

APPENDIX A

Admissible Hilbert Spaces and Reproducing Kernels

In this Appendix we will concisely summarize the properties of admissible Hilbert spaces and their reproducing kernels. We refer the reader to [**47**] for more details; we should note that Reproducing Kernel Hilbert Spaces (RKHS) were first introduced in [**3**], while [**45**] provides a modern and concise introduction.

We denote with $C_0^0(\mathbb{R}^D, \mathbb{R}^D)$ or simply with $C_0(\mathbb{R}^D, \mathbb{R}^D)$ the linear space of continuous functions $u : \mathbb{R}^D \to \mathbb{R}^D$ that vanish at infinity (that is, such that for every $\varepsilon > 0$ the set $\{x : \|u(x)\|_{\mathbb{R}^D} \geq \varepsilon\}$ is compact; see [**16**]) which is Banach with the norm:

$$\|u\|_\infty \triangleq \max_{x \in \mathbb{R}^D} \|u(x)\|_{\mathbb{R}^D}.$$

Also, we define

$$C_0^k(\mathbb{R}^D, \mathbb{R}^D) \triangleq \left\{u \in C^k(\mathbb{R}^D, \mathbb{R}^D) : \partial^\alpha u \in C_0(\mathbb{R}^D, \mathbb{R}^D) \text{ for } |\alpha| \leq k\right\},$$

where $C^k(\mathbb{R}^D, \mathbb{R}^D)$ is the linear space of functions $\mathbb{R}^D \to \mathbb{R}^D$ which are continuously differentiable k times; in the above definition we have used the multi-index notation for partial derivatives [**13, 16**]. C_0^k is a Banach space with the C^k (or $W^{k,\infty}$) norm

$$\|u\|_{k,\infty} \triangleq \sum_{|\alpha| \leq k} \max_{x \in \mathbb{R}^D} \left\|\partial^\alpha u(x)\right\|_{\mathbb{R}^D}.$$

In particular $C_0^1(\mathbb{R}^D, \mathbb{R}^D)$ is the linear space of continuously differentiable functions $u : \mathbb{R}^D \to \mathbb{R}^D$ that vanish at infinity with their first partial derivatives, which is Banach with the norm:

$$\|u\|_{1,\infty} = \max_{x \in \mathbb{R}^D} \|u(x)\|_{\mathbb{R}^D} + \sum_{i=1}^{D} \max_{x \in \mathbb{R}^D} \left\|\frac{\partial u}{\partial x_i}(x)\right\|_{\mathbb{R}^D}.$$

Admissible Hilbert spaces are defined as follows.

Definition A.1. A Hilbert space $(V, \langle \cdot, \cdot \rangle_V)$ of functions $\mathbb{R}^D \to \mathbb{R}^D$ is said to be *admissible* if the following conditions hold:

(1) V is continuously embedded in $C_0^1(\mathbb{R}^D, \mathbb{R}^D)$, i.e. there exists a positive constant C such that $\|u\|_{1,\infty} \leq C\|u\|_V$, for all $u \in V$;

(2) for any positive integer M, if $x_1, \ldots, x_N \in \mathbb{R}^D$ and $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^D$ are such that, for all $u \in V$, $\sum_{i=1}^N \left\langle \alpha_i, u(x_i) \right\rangle_{\mathbb{R}^D} = 0$, then $\alpha_1 = \ldots = \alpha_N = 0$.

We should note that property (2) establishes a certain "richness" of functions in space V. In fact it can be rewritten as follows: for fixed points $x_1, \ldots, x_N \in \mathbb{R}^D$, if there exist vectors $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^D$ of which at least one is non-zero, then there exists at least a function $v \in V$ such that $\sum_{i=1}^N \left\langle \alpha_i, u(x_i) \right\rangle_{\mathbb{R}^D} \neq 0$.

Example. As we do in Chapter 2, V can be chosen to be Sobolev space $H^k(\mathbb{R}^D, \mathbb{R}^D)$ with its norm:

$$||u||_V^2 \triangleq \int_{\mathbb{R}^D} \left\langle Lu(x), u(x) \right\rangle_{\mathbb{R}^D} dx;$$

in the above expression $L = (\mathrm{id} - a^2\Delta)^k$ is a self-adjoint spatial differential operator ($a \in \mathbb{R}$, $k \in \mathbb{N}$ and Δ is the Laplacian) that is applied to each of the D components of vector field u. By the Sobolev Embedding Theorem [**16**] we have in fact that if $k > \frac{D}{2} + 1$ then V is embedded in $C_0^1(\mathbb{R}^D, \mathbb{R}^D)$. Property (2) of Definition A.1 is also satisfied by $H^k(\mathbb{R}^D, \mathbb{R}^D)$ (in fact, for any value of k).

We will now prove the existence of so-called *reproducing kernels* for admissible Hilbert spaces, which is a consequence of property (1) of Definition A.1. For a fixed point $x \in \mathbb{R}^D$ and a fixed vector $\alpha \in \mathbb{R}^D$ consider the evaluation functional

$$\delta_x^\alpha : V \to \mathbb{R} : u \mapsto \left\langle \alpha, u(x) \right\rangle_{\mathbb{R}^D}. \tag{A.1}$$

Note that:

- functional δ_x^α is *linear*, since for all $u, v \in V$,

$$\begin{aligned} \delta_x^\alpha(u+v) &= \left\langle \alpha, (u+v)(x) \right\rangle_{\mathbb{R}^D} = \left\langle \alpha, u(x) + v(x) \right\rangle_{\mathbb{R}^D} \\ &= \left\langle \alpha, u(x) \right\rangle_{\mathbb{R}^D} + \left\langle \alpha, v(x) \right\rangle_{\mathbb{R}^D} = \delta_x^\alpha(u) + \delta_x^\alpha(v); \end{aligned}$$

- functional δ_x^α is also *bounded*, since for all $u \in V$,

$$\begin{aligned} |\delta_x^\alpha(u)| &= \left|\left\langle \alpha, u(x) \right\rangle_{\mathbb{R}^D}\right| \leq \|\alpha\|_{\mathbb{R}^D} \|u(x)\|_{\mathbb{R}^D} \\ &\leq \|\alpha\|_{\mathbb{R}^D} \|u\|_{1,\infty} \leq C\|\alpha\|_{\mathbb{R}^D} \|u\|_V , \end{aligned}$$

  where we have used the Schwartz inequality, the definition of $\|\cdot\|_\infty$ and property (1) of Definition A.1.

Therefore, by the Riesz Representation Theorem for Hilbert Spaces [**16**] there exists a unique function $K_x^\alpha(\cdot) \in V$ such that

$$\left\langle K_x^\alpha, u \right\rangle_V = \left\langle \alpha, u(x) \right\rangle_{\mathbb{R}^D} , \tag{A.2}$$

for all $u \in V$. Such function is called the *representer* of evaluation functional (A.1), and relation (A.2) is referred to the *reproducing property* of function $K_x^\alpha(\cdot)$.

Remark. In order for V to have a reproducing kernel we could loosen the requirements in the definition of admissible Hilbert space. In fact the existence of a constant C such that $\|u\|_\infty \leq C\|u\|_V$, for all $u \in V$ would imply the boundedness of functional δ_x^α; in other words, the continuous embedding of V in $C_0^0(\mathbb{R}^D, \mathbb{R}^D)$ would be sufficient. However we require that V is continuously embedded in $C_0^1(\mathbb{R}^D, \mathbb{R}^D)$ so that the flow generated by time dependent vector fields $v \in L^1\big([0,1], V\big)$, used in Chapter 2, are actually diffeomorphism from $\mathbb{R}^D$ to $\mathbb{R}^D$, i.e. continuously differentiable and invertible with continuously differentiable inverse.

Remark. The Hilbert space $L^2(\mathbb{R}, \mathbb{R})$ of square integrable functions, with inner product $\langle f, g \rangle = \int_{\mathbb{R}} fg \ dx$, does *not* have a reproducing kernel (indeed, it is not embedded in $C_0^0(\mathbb{R}, \mathbb{R})$); formally, only the Dirac delta function, which is not an element of the space, has the reproducing property.

Let V be an admissible Hilbert space. Note that the map $\mathbb{R}^D \to V : \alpha \mapsto K_x^\alpha(\cdot)$ is *linear* in α; in fact, for all $x \in \mathbb{R}^D$, $\alpha, \beta \in \mathbb{R}^D$, and $u \in V$,

$$\begin{aligned} \left\langle K_x^{\alpha+\beta}, u \right\rangle_V &= \left\langle \alpha + \beta, u(x) \right\rangle_{\mathbb{R}^D} = \left\langle \alpha, u(x) \right\rangle_{\mathbb{R}^D} + \left\langle \beta, u(x) \right\rangle_{\mathbb{R}^D} \\ &= \left\langle K_x^\alpha, u \right\rangle_V + \left\langle K_x^\beta, u \right\rangle_V = \left\langle K_x^\alpha + K_x^\beta, u \right\rangle_V ; \end{aligned}$$

whence $K_x^{\alpha+\beta} = K_x^\alpha + K_x^\beta$ by the uniqueness of the representer. In particular, for any point $y \in \mathbb{R}^D$ we have that $K_x^{\alpha+\beta}(y) = K_x^\alpha(y) + K_x^\beta(y)$; so for an arpitrary pair of points $x, y \in \mathbb{R}^D$ there exists a $D \times D$ matrix $K(y,x) \in \mathbb{R}^{D\times D}$ such that, for all $\alpha \in \mathbb{R}^D$, $K_x^\alpha(y) = K(y,x)\alpha$ (here we are treating α as a *column* vector). Such matrix-valued function $K : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D$ is called the *reproducing kernel* of Hilbert space V (whence the name of reproducing kernel Hilbert spaces).

By the reproducing property (A.2) we have that, for any pair of points $x, y \in \mathbb{R}^D$ and any pair of vectors $\alpha, \beta \in \mathbb{R}^D$,

$$\left\langle K_x^\alpha, K_y^\beta \right\rangle_V = \left\langle \alpha, K_y^\beta(x) \right\rangle_{\mathbb{R}^D} = \left\langle \alpha, K(x,y)\beta \right\rangle_{\mathbb{R}^D} = \alpha^T K(x,y)\beta,$$

but also

$$\left\langle K_y^\beta, K_x^\alpha \right\rangle_V = \left\langle \beta, K_x^\alpha(y) \right\rangle_{\mathbb{R}^D} = \left\langle \beta, K(y,x)\alpha \right\rangle_{\mathbb{R}^D} = \beta^T K(y,x)\alpha = \alpha^T K(y,x)^T \beta\,,$$

so that by the arbitrariness of α and β we have the symmetry: $K(x,y) = K(y,x)^T$. We will now use property (2) in Definition A.1 of admissible Hilbert spaces to prove positive-definiteness of the kernel.

Proposition A.2. *For any positive integer M, any points $x_1, \ldots, x_N \in \mathbb{R}^D$ and vectors $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^D$ the following holds:*

$$\sum_{i,j=1}^{M} \alpha_i^T K(x_i, x_j)\alpha_j \geq 0$$

with equality if and only if $\alpha_1 = \ldots = \alpha_M = 0$.

Proof. By the reproducing property of the representer function:

$$\begin{aligned}
\Big\| \sum_{i=1}^{M} K_{x_i}^{\alpha_i} \Big\|_V^2 &= \left\langle \sum_{i=1}^{M} K_{x_i}^{\alpha_i}, \sum_{j=1}^{M} K_{x_j}^{\alpha_j} \right\rangle_V = \sum_{i,j=1}^{M} \left\langle K_{x_i}^{\alpha_i}, K_{x_j}^{\alpha_j} \right\rangle_V \\
&= \sum_{i,j=1}^{M} \left\langle \alpha_i, K_{x_j}^{\alpha_j}(x_i) \right\rangle_{\mathbb{R}^D} = \sum_{i,j=1}^{M} \left\langle \alpha_i, K(x_i,x_j)\alpha_j \right\rangle_{\mathbb{R}^D} \\
&= \sum_{i,j=1}^{M} \alpha_i^T K(x_i,x_j)\alpha_j\,,
\end{aligned}$$

so that $\sum_{i,j=1}^{M} \alpha_i^T K(x_i, x_j)\alpha_j \geq 0$, with equality if and only if $\sum_{i=1}^{M} K_{x_i}^{\alpha_i} = 0$, that is when, for all $v \in V$,

$$\left\langle \sum_{i=1}^{M} K_{x_i}^{\alpha_i}, v \right\rangle_V = 0\,, \qquad \text{i.e.} \qquad \sum_{i=1}^{M} \left\langle \alpha_i, v(x_i) \right\rangle_{\mathbb{R}^D} = 0\,;$$

but by condition (2) in Definition A.1 the above relations holds if and only if $\alpha_i = 0$ for all $i = 1, \dots, N$. $\square$

PROPOSITION A.3. *When the admissible Hilbert space is of the Sobolev type, $V = H^k(\mathbb{R}^D, \mathbb{R}^D)$, with inner product:*

$$\left\langle u, v \right\rangle_V = \int_{\mathbb{R}^D} \left\langle Lu(x), v(x) \right\rangle_{\mathbb{R}^D} \, dx$$

where L is a self-adjoint differential operator that acts on each of the components of u, the reproducing kernel has the form:

$$K(x, y) = G(x, y) \cdot \mathrm{id} = \begin{bmatrix} G(x,y) & 0 & \cdots & 0 \\ 0 & G(x,y) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & G(x,y) \end{bmatrix},$$

where id *is the $D \times D$ identity matrix and $G(x, y)$ is the (scalar) Green's function*[1]*, or fundamental solution, of differential operator L.*

PROOF. For any point $x \in \mathbb{R}^D$, vector $\alpha \in \mathbb{R}^D$, and function $u \in V$, by the reproducing property (A.2) we have that

$$\begin{aligned} \left\langle \alpha, u(x) \right\rangle_{\mathbb{R}^D} &= \left\langle K_x^\alpha, u \right\rangle_V = \int_{\mathbb{R}^D} \left\langle K_x^\alpha(y), Lu(y) \right\rangle_{\mathbb{R}^D} \, dy \\ &= \int_{\mathbb{R}^D} \left\langle K(y,x)\alpha, Lu(y) \right\rangle_{\mathbb{R}^D} \, dy = \int_{\mathbb{R}^D} \alpha^T K(y,x)^T Lu(y) \, dy \\ &= \alpha^T \int_{\mathbb{R}^D} K(x,y) \, Lu(y) \, dy = \left\langle \alpha, \int_{\mathbb{R}^D} K(x,y) \, Lu(y) \, dy \right\rangle_{\mathbb{R}^D}, \end{aligned}$$

[1]By definition, the Green's function of a differential operator L is the solution of partial differential equation $Lu = \delta$, where δ is Dirac's delta function. The solution to equation $Lv = f$ is $v(x) = \int G(x, y) f(y) \, dy$, therefore the Green's function has the property that $v(x) = \int G(x, y) \, Lv(y) \, dy$ for any function v in the appropriate function space. See [**13**] for more details.

where we have treated $Lu(y)$ as a *column* vector. By the arbitrariness of $\alpha \in \mathbb{R}^D$ we have that

$$u(x) = \int_{\mathbb{R}^D} K(x,y)\, Lu(y)\; dy\,. \tag{A.3}$$

So if we indicate with u^i the i-th component of vector u and with K^{ij} the element of matrix (reproducing kernel) K in position (i,j), expression (A.3) can be written as

$$u^i(x) = \int_{\mathbb{R}^D} \sum_{j=1}^{D} K^{ij}(x,y)\, Lu^j(y)\; dy\,, \qquad i = 1, \ldots, D;$$

function $u \in V$ is also arbitrary, so if we choose a function whose only nonzero component is u^1 the above integral yields

$$u^1(x) = \int_{\mathbb{R}^D} K^{11}(x,y)\, Lu^1(y)\; dy\,,$$

so we may conclude, by the arbitrariness of u^1, that $K^{11}(x,y) = G(x,y)$, i.e. the Green's function, or fundamental solution, of differential operator L. In the same way we can prove that all the remaining diagonal elements $K^{ii}(x,y)$ are equal to $G(x,y)$, for $i = 2, \ldots, D$. On the other hand, if we choose again a function $u \in V$ whose only nonzero component is u^1 we also get

$$u^k(x) = 0 = \int_{\mathbb{R}^D} K^{k1}(x,y)\, Lu^1(y)\; dy\,, \qquad k = 2, \ldots, D,$$

so that, by the arbitrariness of u^1, it must be the case that $K^{k1}(x,y) \equiv 0$. In a completely similar way one can prove that all the remaining off-diagonal elements of the reproducing kernel K are identically equal to zero. □

In the case of admissible Hilbert spaces of the Sobolev type we sometimes say, with abuse of terminology, that scalar function G is the kernel of V.

COROLLARY A.4. *Under the hypotheses of Proposition A.3,*

$$\big\langle G(\cdot,x)\alpha, v\big\rangle_V = \big\langle \alpha, v(x)\big\rangle_{\mathbb{R}^D}$$

for any point $x \in \mathbb{R}^D$, any vector $x \in \mathbb{R}^D$, and any function $v \in V$.

Proof. In fact, by the reproducing property:

$$\left\langle \alpha, v(x) \right\rangle_{\mathbb{R}^D} = \left\langle K_x^\alpha, v \right\rangle_V = \left\langle K(\cdot, x)\alpha, v \right\rangle_V = \left\langle G(\cdot, x) \,\mathrm{id}\, \alpha, v \right\rangle_V = \left\langle G(\cdot, x)\alpha, v \right\rangle_V . \quad \square$$

Corollary A.5. *Under the hypotheses of Proposition A.3, for any choice of points $x_1, \ldots, x_M \in \mathbb{R}^D$ and for any vector $(a_1, \ldots, a_M) \in \mathbb{R}^M$ it is the case that*

$$\sum_{i,j=1}^{M} G(x_i, x_j)\, a_i a_j \geq 0\,,$$

with equality if and only if $a_1 = \ldots = a_M = 0$.

Proof. It is sufficient to apply Proposition A.2 to vectors $\alpha_i = \left(\frac{a_i}{\sqrt{D}}, \ldots, \frac{a_i}{\sqrt{D}}\right) \in \mathbb{R}^D$, for $i = 1, \ldots, M$. $\square$

APPENDIX B

Properties of Bessel Kernels

1. Introduction

When performing regularized landmark matching between the two labeled sets $I = (x^1, \ldots, x^N)$ and $I' = (y^1, \ldots, y^N)$, one's objective is to minimize the functional:

$$E[v, q] \triangleq \int_0^1 \int_{\mathbb{R}^D} \big\langle L v_t, v_t \big\rangle_{\mathbb{R}^D} \, dx \, dt + \lambda \int_0^1 \sum_{i=1}^N \left\| \frac{dq_i}{dt}(t) - v_t\big(q_i(t)\big) \right\|^2 dt$$

with respect to time-dependent velocity $v \in L^2([0, T], V)$, where V is an appropriate Hilbert space embedded in $C_0^1(\mathbb{R}^D, \mathbb{R}^D)$, and to landmark trajectories $q^i : [0, T] \to \mathbb{R}^D$ that satisfy the boundary conditions $q^i(0) = x^i$ and $q^i(1) = y^i$, for $i = 1, \ldots, N$. In the first term of the right-hand side of the above equation L is a *spatial* differential operator which typically has the form $L = (\mathrm{id} - a^2 \Delta)^k$, where Δ is the Laplacian operator and a^2 is just a scaling factor; exponent k is generally a positive integer, although the theory can be extended to the case of positive real values of k (pseudo-differential operators).

It turns out that the Green's function of operator L plays a fundamental role in both the solution of the above minimization problem and in the study of the Riemannian curvature tensor of landmark-based shape manifolds. When the differential operator is in fact $L = (\mathrm{id} - a^2 \Delta)^k$ such Green's function $G : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}$ takes the form $G(x, y) = \gamma\big(\|x - y\|_{\mathbb{R}^D}\big)$, with $\gamma : [0, \infty) \to \mathbb{R}$ given by:

$$\boxed{\gamma(\varrho) = \frac{1}{2^{k + \frac{D}{2} - 1} \pi^{\frac{D}{2}} \, \Gamma(k)} \, \frac{1}{a^D} \left(\frac{\varrho}{a}\right)^{k - \frac{D}{2}} K_{k - \frac{D}{2}}\left(\frac{\varrho}{a}\right),} \tag{B.1}$$

where K_ν is a modified Bessel function [**1**, Chap. 9] whereas Γ is the gamma function [**1**, Chap. 6]. Note that γ and its derivatives are defined at zero by continuity. Reference [**20**] provides a table of fundamental solutions for different differential operators, from which (B.1) can be computed. Function (B.1) depends on both D, the dimension of the ambient space for landmarks, and k, the exponent of differential operator L. Function γ is C^∞ in the open set $(0, \infty)$ while its regularity in the origin depends on $\nu = k - \frac{D}{2}$: its smoothness at zero in fact increases with parameter ν. We shall explore the asymptotic behavior of γ at zero later on in this appendix.

For notational convenience let us define, for a fixed value of scaling factor a:

$$\eta_{k,D} \triangleq \frac{1}{2^{k+\frac{D}{2}-1}\pi^{\frac{D}{2}}\,\Gamma(k)}\,\frac{1}{a^{k+\frac{D}{2}}}\,,$$

so that (B.1) can simply be written as

$$\gamma(\varrho) = \eta_{k,D}\,\varrho^{\nu} K_{\nu}\Big(\frac{\varrho}{a}\Big), \tag{B.2}$$

with $\nu = k - \frac{D}{2}$.

2. Differential equation

It is well known that modified Bessel function $K_\nu(z)$ is a solution to the following second order differential equation [**1**]:

$$z^2\frac{d^2w}{dz^2} + z\frac{dw}{dz} - (z^2 + \nu^2)w = 0\,; \tag{B.3}$$

note that $K_\nu(z)$ is a holomorphic function of z throughout the complex plane cut along the negative real axis, and for fixed z it is an entire function of ν. For our purposes, we are interested in $K_\nu(z)$ when $z \in \mathbb{R}^+$ and $\nu \in \mathbb{R}$.

Function $\gamma(\varrho)$, which is defined in terms of a modified Bessel function, also satisfies a second order differential equation which will differ from (B.3) due to factor $\varrho^{k-D/2}$ in (B.1). In fact the following result holds.

Proposition B.1. *Function γ defined in* (B.1) *satisfies the following second order differential equation:*

$$\boxed{\gamma'' = \frac{2\nu - 1}{\varrho}\,\gamma' + \frac{1}{a^2}\,\gamma\,,} \tag{B.4}$$

where $\nu = k - \frac{D}{2}$.

Proof. Perhaps the most straightforward way of proving the proposition is taking successive derivatives of the expression provided by (B.1), recombining the resulting terms and and then using differential equation (B.3). However we shall pursue a different path, which will also yield a useful expression for the first derivative of $\gamma(\varrho)$.

Differentiation of (B.2) yields the following expression:

$$\gamma'(\varrho) = \eta_{k,D}\Big\{\nu\varrho^{\nu-1}K_\nu\Big(\frac{\varrho}{a}\Big) + \frac{1}{a}\,\varrho^\nu K_\nu'\Big(\frac{\varrho}{a}\Big)\Big\}. \tag{B.5}$$

We will use an important property of the modified Bessel functions of our interest. Namely, the following formulas hold [**1**, §9.6.26] for any value of parameter μ:

$$\mathcal{Z}_\mu'(z) = \mathcal{Z}_{\mu-1}(z) - \frac{\mu}{z}\,\mathcal{Z}_{\mu+1}(z)\,, \tag{B.6}$$

$$\mathcal{Z}_\mu'(z) = \mathcal{Z}_{\mu+1}(z) + \frac{\mu}{z}\,\mathcal{Z}_{\mu}(z)\,, \tag{B.7}$$

where $\mathcal{Z}_\mu(z) = e^{\mu\pi i}K_\mu(z)$. Identity (B.6) implies that $K_\mu'(z) = -K_{\mu-1}(z) - \frac{\mu}{z}K_\mu(z)$ for all μ, so that, fixing $\mu = \nu$,

$$K_\nu'\Big(\frac{\varrho}{a}\Big) = -K_{\nu-1}\Big(\frac{\varrho}{a}\Big) - \frac{a\nu}{\varrho}K_\nu\Big(\frac{\varrho}{a}\Big); \tag{B.8}$$

the second one (B.7) implies that $K_\mu'(z) = -K_{\mu+1}(z) + \frac{\mu}{z}\,K_\mu(z)$, whence for $\mu = \nu - 1$

$$K_{\nu-1}'\Big(\frac{\varrho}{a}\Big) = -K_\nu\Big(\frac{\varrho}{a}\Big) + \frac{a}{\varrho}\,(\nu-1)K_{\nu-1}\Big(\frac{\varrho}{a}\Big). \tag{B.9}$$

Inserting (B.8) into the right-hand side of (B.5) yields:

$$\begin{aligned}\gamma'(\varrho) &= \eta_{k,D}\Big\{\nu\varrho^{\nu-1}K_\nu\Big(\frac{\varrho}{a}\Big) + \frac{1}{a}\,\varrho^\nu\Big[-K_{\nu-1}\Big(\frac{\varrho}{a}\Big) - \frac{a\nu}{\varrho}K_\nu\Big(\frac{\varrho}{a}\Big)\Big]\Big\}\\ &= -\eta_{k,D}\,\frac{1}{a}\,\varrho^\nu K_{\nu-1}\Big(\frac{\varrho}{a}\Big).\end{aligned} \tag{B.10}$$

Differentiating the above function gives:

$$
\begin{aligned}
\gamma''(\varrho) &= -\eta_{k,D}\Big\{\frac{1}{a}\,\nu\varrho^{\nu-1}K_{\nu-1}\Big(\frac{\varrho}{a}\Big)+\frac{1}{a^2}\,\varrho^{\nu}K'_{\nu-1}\Big(\frac{\varrho}{a}\Big)\Big\}\\
&\overset{(*)}{=} -\eta_{k,D}\Big\{\frac{1}{a}\,\nu\varrho^{\nu-1}K_{\nu-1}\Big(\frac{\varrho}{a}\Big)+\frac{1}{a^2}\,\varrho^{\nu}\Big[-K_{\nu}\Big(\frac{\varrho}{a}\Big)+\frac{a}{\varrho}\,(\nu-1)K_{\nu-1}\Big(\frac{\varrho}{a}\Big)\Big]\Big\}\\
&= -\eta_{k,D}\Big\{\frac{2\nu-1}{\varrho}\,\frac{1}{a}\,\nu\varrho^{\nu-1}K_{\nu-1}\Big(\frac{\varrho}{a}\Big)-\frac{1}{a^2}\,\varrho^{\nu}K_{\nu}\Big(\frac{\varrho}{a}\Big)\Big\}\\
&= \frac{2\nu-1}{\varrho}\Big\{-\eta_{k,D}\,\frac{1}{a}\,\varrho^{\nu}K_{\nu-1}\Big(\frac{\varrho}{a}\Big)\Big\}+\frac{1}{a^2}\,\eta_{k,D}\,\varrho^{\nu}K_{\nu}\Big(\frac{\varrho}{a}\Big)\\
&= \frac{2\nu-1}{\varrho}\,\gamma'(\varrho)+\frac{1}{a^2}\,\gamma(\varrho)\,,
\end{aligned}
$$

which is equation (B.4); note that we have used expression (B.9) in step $(*)$. □

A by-product of the above proof is the following result.

Corollary B.2. *The first derivative of function γ may be expressed ad follows:*

$$
\begin{aligned}
\gamma'(\varrho) &= -\frac{1}{2^{k+\frac{D}{2}-1}\pi^{\frac{D}{2}}\,\Gamma(k)}\,\frac{1}{a^{D+1}}\,\Big(\frac{\varrho}{a}\Big)^{k-\frac{D}{2}}K_{k-\frac{D}{2}-1}\Big(\frac{\varrho}{a}\Big)\\
&= -\eta_{k,D}\,\frac{1}{a}\,\varrho^{k-\frac{D}{2}}K_{k-\frac{D}{2}-1}\Big(\frac{\varrho}{a}\Big).
\end{aligned}
$$

Proof. The result follows immediately from expression (B.10). □

3. Asymptotic behavior at zero

We shall now study the behavior of function $\gamma(\varrho)$ in a neighborhood of zero. We will need the following properties of modified Bessel functions [**1**, §9.6.8, §9.6.9]: when $z \to 0$, $K_0 \sim -\ln z$, while for a fixed value of paramenter μ with $\Re\mu > 0$ we have that $K_\mu(z) \sim \frac{1}{2}\Gamma(\mu)\big(\frac{1}{2}z\big)^{-\mu}$. In other words,

$$\lim_{z\to 0}\frac{K_0(z)}{\ln z} = -1$$

and

$$\lim_{z\to 0} z^{\mu}K_\mu(z) = 2^{\mu-1}\Gamma(\mu) \tag{B.11}$$

for $\Re\mu > 0$.

Proposition B.3. *It is the case that:*

$$\lim_{\varrho\to 0}\gamma(\varrho) = \frac{1}{\left(2a\sqrt{\pi}\right)^D}\,\frac{\Gamma\left(k-\frac{D}{2}\right)}{\Gamma(k)},$$

provided that $k-\frac{D}{2}>0$. *Otherwise, when* $k-\frac{D}{2}\le 0$ *the above limit diverges to* $+\infty$.

Proof. The case $k>\frac{D}{2}$ follows directly from applying property (B.11) to expression (B.1). The other case follows form the facts that $K_\mu(x)\to+\infty$ as $x\to 0^+$ for any value of parameter μ, and that $x^\mu\to+\infty$ as $x\to 0^+$ for $\mu<0$. □

Proposition B.4. *The first derivative of function* γ *is such that:*

$$\lim_{\varrho\to 0}\frac{\gamma'(\varrho)}{\varrho} = -\frac{1}{2^{D+1}a^{D+2}\pi^{\frac{D}{2}}}\,\frac{\Gamma\left(k-\frac{D}{2}-1\right)}{\Gamma(k)},$$

provided that $k-\frac{D}{2}-1>0$.

Proof. It follows from Corollary B.2 that

$$\frac{\gamma'(\varrho)}{\varrho} = -\frac{1}{2^{k+\frac{D}{2}-1}\pi^{\frac{D}{2}}\,\Gamma(k)}\,\frac{1}{a^{D+2}}\left(\frac{\varrho}{a}\right)^{k-\frac{D}{2}-1}K_{k-\frac{D}{2}-1}\left(\frac{\varrho}{a}\right),$$

which, provided that $k-\frac{D}{2}-1>0$, by (B.11) converges to

$$\begin{aligned}\lim_{\varrho\to 0}\frac{\gamma'(\varrho)}{\varrho} &= -\frac{1}{2^{k+\frac{D}{2}-1}\pi^{\frac{D}{2}}\,\Gamma(k)}\,\frac{1}{a^{D+2}}\,2^{k-\frac{D}{2}-2}\,\Gamma\left(k-\frac{D}{2}-2\right)\\ &= -\frac{1}{2^{D+1}a^{D+2}\pi^{\frac{D}{2}}}\,\frac{\Gamma\left(k-\frac{D}{2}-1\right)}{\Gamma(k)},\end{aligned}$$

which is precisely what we wanted to prove. □

Proposition B.5. *The second derivative of function* γ *is such that:*

$$\lim_{\varrho\to 0}\gamma''(\varrho) = -\frac{1}{2^{D+1}a^{D+2}\pi^{\frac{D}{2}}}\,\frac{\Gamma\left(k-\frac{D}{2}-1\right)}{\Gamma(k)},$$

provided that $k-\frac{D}{2}-1>0$. *If* $k-\frac{D}{2}-1\le 0$, *the above limit diverges to* $+\infty$.

Proof. In order to compute the above limit we shall use differential equation (B.4) and the results provided by Propositions B.3 and B.4, which both hold since $k-\frac{D}{2}-1>0$. By Proposition B.3 we have that

$$\lim_{\varrho\to 0}\frac{1}{a^2}\,\gamma(\varrho) = \frac{k-\frac{D}{2}-1}{2^D a^{D+2}\pi^{\frac{D}{2}}}\,\frac{\Gamma\left(k-\frac{D}{2}-1\right)}{\Gamma(k)},$$

where we have used the property of the gamma function: $\Gamma(z+1) = z\Gamma(z)$. On the other hand by Proposition B.3 it is the case that

$$\lim_{\varrho\to 0} \frac{2\nu - 1}{\varrho}\,\gamma'(\varrho) = -\frac{2k - D - 1}{2^{D+1}a^{D+2}\pi^{\frac{D}{2}}}\,\frac{\Gamma\big(k - \frac{D}{2} - 1\big)}{\Gamma(k)}\,,$$

where we have fixed $\nu = k - \frac{D}{2}$. Combining the above limits with equation (B.4) finally yields

$$\begin{aligned}
\lim_{\varrho\to 0}\gamma''(\varrho) &= \lim_{\varrho\to 0}\Big[\frac{2\nu-1}{\varrho}\,\gamma'(\varrho) + \frac{1}{a^2}\,\gamma(\varrho)\Big] \\
&= -\frac{2k - D - 1}{2^{D+1}a^{D+2}\pi^{\frac{D}{2}}}\,\frac{\Gamma\big(k - \frac{D}{2} - 1\big)}{\Gamma(k)} + \frac{2k - D - 2}{2^{D+1}a^{D+2}\pi^{\frac{D}{2}}}\,\frac{\Gamma\big(k - \frac{D}{2} - 1\big)}{\Gamma(k)} \\
&= -\frac{1}{2^{D+1}a^{D+2}\pi^{\frac{D}{2}}}\,\frac{\Gamma\big(k - \frac{D}{2} - 1\big)}{\Gamma(k)}\,,
\end{aligned}$$

as we wanted to prove. □

Remark. Out of curiosity, note that: $\lim_{\varrho\to 0}\gamma''(\varrho) = \lim_{\varrho\to 0}\dfrac{\gamma'(\varrho)}{\varrho}$.

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, New York, 1964.

[2] V. I. Arnold. *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 1989.

[3] N. Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3):337–404, May 1950.

[4] G. Benettin, L. Galgani, and A. Giorgilli. *Appunti di Meccanica Razionale.* Edizioni Libreria Progetto, Padova, Italy, 1991.

[5] C. M. Bishop. *Pattern Recognition and Machine Learning.* Information Science and Statistics. Springer, 2006.

[6] K. W. Brodlie. Unconstrained minimization. In *The State of the Art in Numerical Analysis*, pages 229–268. Academic Press, London, Apr. 1977.

[7] C. Chicone. *Ordinary Differential Equations and Applications*, volume 34 of *Texts in Applied Mathematics*. Springer, 1999.

[8] G. E. Christensen, R. Rabbitt, and M. I. Miller. 3D brain mapping using a deformable neuroanatomy. *Physics in Medicine and Biology*, 39:609–618, 1994.

[9] G. E. Christensen, R. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, Oct. 1994.

[10] J. G. Csernansky, S. Joshi, L. Wang, J. W. Haller, M. Gado, J. P. Miller, U. Grenander, and M. I. Miller. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. *Proc. Natl. Acad. Sci. USA, Neurobiology*, 95:11406–11411, Sept. 1998.

[11] M. P. Do Carmo. *Riemannian Geometry.* Birkhäuser, Boston, 1992.

[12] M. G. Eastwood and P. W. Michor. Some remarks on the Plücker relations. *Rendiconti del Circolo Matematico di Palermo*, 63:85–88, 2000.

[13] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, Rhode Island, 1998.

[14] P. T. Fletcher and S. Joshi. Principal Geodesic Analysis on symmetric spaces: Statistics of diffusion tensors. In *Computer Vision - ECCV 2004*, volume 3024 of *Lecture Notes in Computer Science*. Springer-Verlag, Prague, Czech Republic, May 2004.

[15] P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi. Principal Geodesic Analysis for the Study of Nonlinear Statistics of Shape. *IEEE Transactions on Medical Imaging*, 23(8):995–1005, Aug. 2004.

[16] G. B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, second edition, 1999.

[17] J. Glaunès. *Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique*. PhD thesis, Université Paris 13, France, Sept. 2005.

[18] U. Grenander and M. I. Miller. Computational anatomy: an emerging discipline. *Quarterly of Applied Mathematics*, 56(3):617–694, Sept. 1998.

[19] B. K. P. Horn. *Robot vision*. The MIT Press, Cambridge, Massachusetts, 1986.

[20] K. Itô, editor. *Encyclopedic Dictionary of Mathematics*. MIT Press, Cambridge, Massachusetts, USA, second edition, 1993. Prepared by the Mathematical Society of Japan.

[21] K. Jänich. *Vector Analysis*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 2001.

[22] S. C. Joshi and M. I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370, Aug. 2000.

[23] J. Jost. *Riemannian Geometry and Geometric Analysis*. Springer-Verlag, New York, third edition, 2002.

[24] E. Kalssen, A. Srivastava, W. Mio, and S. Joshi. Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):372–383, Mar. 2004.

[25] P. R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Prentice Hall, Englewood Cliffs, New Jersey, 1986.

[26] A. J. Laub. *Matrix Analysis for Scientists and Engineers*. SIAM, 2004.

[27] J. M. Lee. *Riemannian Manifolds: an Introduction to Curvature*, volume 176 of *Graduate Texts in Mathematics*. Springer, New York, 1997.

[28] J. M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, 2003.

[29] P. W. Michor. Topics in Differential Geometry. Lecture Notes, University of Vienna, 2006.

[30] P. W. Michor. Curvature, covariantly expressed. Unpublished technical report, University of Vienna, July 2007.

[31] P. W. Michor and D. B. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Documenta Mathematica*, 10:217–245, 2005.

[32] P. W. Michor and D. B. Mumford. Riemannian geometries on spaces of plane curves. *Journal of the European Mathematical Society*, 8:1–48, 2006.

[33] P. W. Michor and D. B. Mumford. An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23:74–113, 2007.

[34] M. I. Miller, A. Trouvé, and L. Younes. On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4:375–405, 2002.

[35] M. I. Miller, A. Trouvé, and L. Younes. Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, 24(2):209–228, Mar. 2006.

[36] M. I. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41(1/2):61–84, 2001.

[37] C. Ovidiu and D.-C. Chang. *Geometric Mechanics on Riemannian Manifolds. Applications to Partial Differential Equations*. Birkhäuser, Boston, 2005.

[38] X. Pennec. *Probabilities and Statistics on Riemannian Manifolds: a Geometric Approach*. Research Report 5093, INRIA, January 2004.

[39] E. Polak. *Computational Methods in Optimization: a Unified Approach*. Academic Press, New York, 1971.

[40] F. Riesz and B. Sz.-Nagy. *Functional Analysis*. Dover, New York, 1990.

[41] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, third edition, 1976.

[42] E. Sharon and D. B. Mumford. Kalman filtering with intermittent observations. *International Journal of Computer Vision*, 70(1):55–75, Oct. 2006.

[43] A. Trouvé. *Infinite Dimensional Group Action and Pattern Recognition*. Technical report, DMI, École Nationale Supérieure, Paris, France, May 1995.

[44] F.-X. Vialard. Landmarks, geodesics and curvature in a simple case. Technical report, École Nationale Supérieure de Cachan, France, Aug. 2007.

[45] G. Wahba. *Spline Models for Observational Data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, Pennsylvania, 1990.

[46] L. Wang, F. Beg, T. Ratnanather, C. Ceritoglu, L. Younes, J. C. Morris, J. G. Csernansky, and M. I. Miller. Large deformation diffeomorphism and momentum based hippocampal shape

discrimination in dementia of the Alzheimer type. *IEEE Transactions on Medical Imaging*, 26(4):462–470, Apr. 2007.

[47] L. Younes. Deformation Analysis for Shape and Image Processing. Lecture notes, Johns Hopkins University, 2007.

[48] L. Younes, P. W. Michor, J. Shah, and D. B. Mumford. A metric on shape space with explicit geodesics. *Rendiconti Lincei – Matematica e Applicazioni*, 9:25–57, 2008.