

## Curve-Straightening in Closed Euclidean Submanifolds

Anders Linnér

Department of Mathematics, University of Missouri-St. Louis, St. Louis, Missouri 63121-4499, USA

Received March 10, 1989; in revised form November 5, 1990

**Abstract.** The flow in the negative direction of the gradient vector field associated with the functional total squared (geodesic) curvature  $\int k^2 ds$  is the so-called curve-straightening flow. This paper will consider spaces of closed curves in closed Euclidean submanifolds. It will define these spaces of curves as submanifolds of certain Hilbert manifolds representing all curves. The main result will then be to show the existence of a particular set of functionals defined on the entire Hilbert manifold which have the following four properties: 1. The directional derivatives of these functionals may be computed by solving an initial value problem for a system of ordinary differential equations. 2. By introducing a suitable Hilbert space basis for the Sobolev spaces used, the gradients may be effectively computed (but of course not explicitly computed, except in very special cases). 3. The gradients span the space normal to the tangent space of the space of closed curves. 4. Despite the fact that these gradients in general are not given explicitly it is nevertheless possible to compute the projection onto the tangent space to the space of closed curves. In particular we do this for the gradient of  $\int k^2 ds$ . When all details are worked out this gives us an algorithm (which we supply) for finding critical points in the space of closed curves. It is not known if the trajectories actually always converge to critical points. If the functional is modified to include a multiple of the length so the functional becomes  $\int k^2 + \lambda ds$  then the above convergence is known for  $\lambda > 0$ . The motivating application for the curve-straightening flow is the possibility of using it to find (non-trivial) closed (periodic) geodesics. Note that if  $\lambda = 0$  then a closed geodesic is a global minimum. For any  $\lambda$ , geodesics are critical but there are also other critical points, the so-called elastic curves. The paper concludes by deriving the second variation formula for  $\int k^2 + \lambda ds$  along closed geodesics. The quadratic functional associated with the second derivative is shown to be positive definite even for non-zero  $\lambda$  along some closed geodesics in some particular manifolds of interest.

## 0. Introduction

This paper will consider the problem of finding critical points of functionals restricted to spaces of closed curves in closed Euclidean submanifolds. In particular we are interested in functionals that are independent of parametrizations of the curve. An example of such a “geometric” quantity is the length of a curve  $\gamma$  and the corresponding length functional  $\int \gamma ds$ . Our main applications will involve functionals of the form  $\int \gamma f(k)ds$ , where  $f$  is some real valued function and  $k$  is the (geodesic) curvature along  $\gamma$ . One of the primary motivations for considering such functionals is the problem of finding geodesics, in particular closed geodesics. To avoid confusion we point out that we consider a geodesic closed only if it is periodic, i.e. not only should the endpoints be equal but the tangent directions should also be equal at the endpoints. We will also call a constant curve a point, rather than a closed geodesic.

Before we continue we remark that it is by no means obvious how to deal with the complications caused by the fact that we consider functionals independent of the parametrizations of the curves. The independence of parametrization will cause the Euler-Lagrange equations associated with the functional to be dependent (see [1]). There are several different approaches to this difficulty. For a discussion of the length functional see [2, pp. 243–252]. One possibility is to fix a parametrization and restrict the functional to curves with this parametrization. This fixed choice of parametrization is then considered as a pointwise constraint and requires a corresponding Lagrange multiplier. In this paper we will use a different approach than any of the above. The so-called (generalized) indicatrix will be introduced. We will suggest a way of uniquely associating to each regular curve parametrized proportional to arclength its indicatrix. The space of indicatrices is given a Hilbert manifold structure and the functional of interest will be defined on this new manifold. The indicatrices representing curves that satisfy some particular choice of boundary conditions (e.g. periodic) turn out to be submanifolds of the space of all indicatrices.

We will in this paper show how to split the tangent space of all indicatrices into a space tangent to the space representing closed curves and its orthogonal complement. We stress that for such a given Hilbert manifold this splitting depends only on the particular boundary conditions under consideration (and, to be sure, how the Euclidean submanifold and its orthonormal frame bundle are embedded). Given a functional defined on the space of all indicatrices its gradient may now be projected orthogonally onto the tangent space of the space of indicatrices representing closed curves. The steepest descent is now to flow along the trajectories in the negative direction of this projected gradient vector field.

A very important application of the above general method is when  $f(k) = k^2 + \lambda$ , here  $\lambda \geq 0$ . Note that if  $\lambda = 0$  then this functional is globally minimized at the closed geodesics ( $k=0$  at all points along a geodesic). The steepest descent flow as described above for this functional is known as the curve-straightening flow. When  $\lambda = 0$  it is not known if all trajectories will converge to critical points. On the other hand for  $\lambda > 0$  such convergence has been established (see [3]).

There are several fundamental differences between the methods used in [3] and the approach used here. In [3] variations of the indicatrix are related to variations of the curve via an integro-differential equation (see Proposition 1.2) involving among other things the curvature along the curve. This approach proved

successful as a means of showing the Palais-Smale condition. Our objective here is quite different. We want to develop an effective method that makes it possible to actually follow these trajectories inside this infinite dimensional manifold. Our main result (Theorem 2.4) will be used to show that the problem of computing the relevant gradients has been reduced to a problem of solving an initial value problem of a system of ordinary differential equations together with some linear algebraic equations. This is a definite improvement as compared to [3]. This is accomplished by allowing the geometry of the manifold containing the curves to enter through the second fundamental form rather than through the curvature. This is why the manifold containing the curves as well as its orthonormal frame bundle are embedded in Euclidean space. Another significant difference is our use of a Hilbert basis in the Sobolev space factor of the space of indicatrices to facilitate the computation of various gradients. The gradient computations in [3] and also in [4] are done with the help of the so-called duBois-Reymonds lemma. Note that apart from this, the present paper to a large extent generalizes the methods developed in [4].

To motivate why the total squared curvature functional is used we remark that the Palais-Smale conditions has been established for the energy functional  $\int ds^2$  as well (see for instance [2]). There are significant differences however. For starters, the energy functional is not independent of the parametrization of the curve. The spaces of curves on which it is defined include constant curves. As a consequence the steepest descent flow when applied to a curve in a trivial homotopy class will tend to shrink such a curve to a point. In particular on sphere-like surfaces with trivial fundamental group this is a serious drawback.

There have also been attempts to use the length functional and define some kind of “gradient” flow (see [5, 6]). These attempts have produced certain partial differential equations whose solutions yield the so-called curve-shortening flow. These flows have their roots in the Birkhoff curve-shortening process (see [7]) and all these flows suffer the same drawback as mentioned above.

Since the total squared curvature functional gets arbitrarily large on any sequence of closed curves approaching a point it is clear that the curve-straightening will not shrink curves down to points. Since geodesics are critical points of the length functional we see that for any  $\lambda \geq 0$  geodesics are critical points of  $\int k^2 + \lambda ds$ . There are also other critical points, the so-called elastic curves. In the case  $\lambda = 0$  examples of such critical points are given in [8] where two-dimensional spaces of constant curvature are considered. For  $\lambda > 0$  there are examples in [3] on the standard round sphere  $S^2$ .

To gain a better understanding of where an initial curve will flow under curve-straightening it is also important to study the stability of the critical points. For the standard round sphere we again refer to [3] and Theorem 3.1 which states that if  $0 \leq \lambda \leq \frac{8}{7}$  then all non-geodesic critical points are unstable. We will in this paper give additional results concerning the stability of the geodesics. In particular we will derive the second variation formula for normal variations along a geodesic in a two-dimensional manifold. We also consider some special cases related to sphere-like surfaces.

The paper is organized as follows:

0. Introduction
1. Indicatrices
2. Projections onto the tangent space of closed curves
3. The curve-straightening algorithm
4. The second variation along geodesics

## 1. Indicatrices

*1.1. M and OM.* We will restrict our attention to curves in closed Euclidean submanifolds. Let  $F: R^k \rightarrow R^l$  be smooth ( $C^\infty$ ). Assume that  $DF(p)$  is onto for all  $p \in F^{-1}(0)$ . We will consider functionals defined on curves in  $M = F^{-1}(0)$ . We let  $n = k - l$  and we assume that  $n \geq 2$ . A standard argument using the implicit function theorem shows that  $M$  is an  $n$ -dimensional submanifold of  $R^k$ . Note also that

$$T_p M = \ker DF(p) = \{v \in R^k \mid DF(p)v = 0\}$$

will be identified as the tangent space to  $M$  at  $p$ . Let the superscript  $T$  indicate transpose so that  $DF(p)^T: R^l \rightarrow R^k$ . The superscript  $\perp$  denotes orthogonal complements. We have

$$T_p M^\perp = \text{Im } DF(p)^T = \{v \in R^k \mid v = DF(p)^T w, w \in R^l\}$$

identified as the normal space to  $M$  at  $p$ .

Next we consider the orthonormal frame bundle  $OM$  associated with  $M$ . Put  $(p, v) = (p, v_1, \dots, v_n)$  and let

$$G: R^k \times R^{nk} \rightarrow R^l \times R^{nl} \times R^n \times R^m,$$

where  $m = \frac{n(n-1)}{2}$  be given by

$$G(p, v) = (F(p), DF(p)v_1, \dots, DF(p)v_n, \\ v_1 \cdot v_1 - 1, \dots, v_n \cdot v_n - 1, v_1 \cdot v_2, \dots, v_1 \cdot v_n, v_2 \cdot v_3, \dots, v_{n-1} \cdot v_n).$$

Now let  $OM = G^{-1}(0)$  and check that  $DG(p, v)$  is onto so another application of the implicit function theorem shows that  $OM$  is an  $\frac{n(n+1)}{2}$ -dimensional submanifold of  $R^k \times R^{nk}$ . Without getting into details we remark that analogous constructions give embeddings of the tangent bundles  $TM$  and  $TOM$  as well as the unit tangent bundle  $SM$  as closed submanifolds of higher dimensional Euclidean spaces.

*1.2. Curves in M and Their Indicatrices.* Let  $I = [0, 1]$  and consider a regular curve  $\gamma: I \rightarrow M$  so that  $|\gamma'(t)| = L$  for all  $t \in I$  (so  $L > 0$  is the length of  $\gamma$  and  $'$  will always mean derivative with respect to  $t$ ). Associated with such  $\gamma$  is its indicatrix, which comes about in the following way. Choose an orthonormal frame at  $\gamma(0)$  and parallel translate this frame along  $\gamma$ . It is now possible to write  $\gamma'(t) = L(y_1(t), \dots, y_n(t))$  with respect to this parallel frame. Note that  $y(t) = (y_1(t), \dots, y_n(t))$  is a map into  $S^{n-1}$  the standard  $(n-1)$ -dimensional sphere. In order to get a unique indicatrix associated to  $\gamma$  we suggest a particular choice of initial frame. Let  $\{b_1, \dots, b_k\}$  denote the standard basis for  $R^k$ . Let  $\tilde{b}_i$  denote the projection of  $b_i$  onto  $T_{\gamma(0)}M$ . Next we apply the Gram-Schmidt process to  $\{\gamma'(0), \tilde{b}_1, \dots, \tilde{b}_k\}$ . This will generate an orthonormal frame  $\{e_1, \dots, e_n\}$  in the tangent space of  $M$  at  $\gamma(0)$ . Note that with this choice of initial frame  $y(0) = e_1$  (i.e. the eastpole as in [4] rather than the northpole as in [3]).

In conjunction with these constructions it is appropriate to make explicit, in terms of  $DF(p)$ , the projections used.

**Proposition 1.1.** *Given  $v \in R^k$  if we let  $v_p = DF(p)^T(DF(p)DF(p)^T)^{-1}DF(p)v$  and  $v_t = v - v_p$  then  $v \in R^k$  is decomposed as  $(v_t, v_p) \in T_p M \oplus T_p M^\perp$ .*

*Proof.* By applying  $DF(p)$  to  $v_p$  we see that  $DF(p)v_p = DF(p)v$  it follows that  $v_t \in T_p M$ . Next we let  $u \in T_p M$  be arbitrary

$$\langle v_p, u \rangle = \langle (DF(p)DF(p)^T)^{-1}DF(p)v, DF(p)u \rangle = 0$$

so  $v_p \in T_p M^\perp$ .

A word of warning is appropriate here. In writing this paper we made every effort to keep the notation as consistent and standard as possible. However we did decide to blur the distinction between row and column vectors and in particular suppress transposes (except for matrices). Since the notation will get quite cumbersome (in Sect. 3.2 in particular) we hope this decision relieved some of the complicated expressions from at least one level of notational mushrooming. As a basis for  $T_p M^\perp$  we use the row vectors of  $DF(p)$ . This will of course in general not be an orthonormal basis. Let  $DF_i(p)$  denote the  $i^{\text{th}}$  such row vector. The next two results we include for completeness. The proofs are straightforward.

**Proposition 1.2.** *The covariant differentiation  $\frac{D}{dt}$  along  $\gamma$  (induced by  $R^k$  on  $M$ ) of tangent vector fields  $v$  along  $\gamma$  is given by*

$$\frac{D}{dt}v(t) = v'(t) - \sum_{i=1}^l \alpha_i(t)DF_i(\gamma(t)),$$

where  $\alpha_i(t)$  is given by the solution of the  $l \times l$  linear system

$$\sum_{i=1}^l \alpha_i(t) \langle DF_i(\gamma(t)), DF_j(\gamma(t)) \rangle = \langle v'(t), DF_j(\gamma(t)) \rangle$$

for  $j=1, \dots, l$ .

**Proposition 1.3.** *The parallel translation of  $v$  along  $\gamma$  with respect to the connection induced by  $R^k$  on  $M$  is given by*

$$v'(t) = \sum_{i=1}^l \alpha_i(t)DF_i(\gamma(t)),$$

where  $v(0)=v$  and  $\alpha_i(t)$  is given by the solution of the  $l \times l$  linear system

$$\sum_{i=1}^l \alpha_i(t) \langle DF_i(\gamma(t)), DF_j(\gamma(t)) \rangle = -\langle v(t), HF_j(\gamma(t))\gamma'(t) \rangle$$

for  $j=1, \dots, l$ .

Here  $HF_j(\gamma(t))$  is the Hessian (the symmetric matrix of second partial derivatives) of the  $j^{\text{th}}$  component function  $F_j$  of  $F$ . To each regular curve parametrized proportional to its length we now associate the triple  $\hat{y}=(y, L, \omega)$ , where  $y$  is the indicatrix,  $L$  is the length of the curve, and

$$\omega = (p, \{e_1, \dots, e_n\}) \in OM,$$

where  $p$  is the initial point of the curve and  $\{e_1, \dots, e_n\}$  is the orthonormal basis of  $T_p M$ . It is of course not necessary to use the basis suggested above but it is convenient to have an automatic way of picking such a frame. We strictly speaking should refer to the whole triple  $\hat{y}$  as the indicatrix. With the use of Proposition 1.3 it is now clear how to determine the indicatrix. Conversely we have:

**Proposition 1.4.** Given  $\hat{y}=(y, L, \omega)$  we can recover  $\gamma$  by solving

$$\begin{aligned}\gamma'(t) &= L \sum_{j=1}^n y_j(t) e_j(t), \\ e'_h(t) &= \sum_{i=1}^l \alpha_{hi}(t) DF_i(\gamma(t))\end{aligned}$$

$\gamma(0)=p$  and  $e_h(0)=e_h$  for  $h=1 \dots n$ .

We again omit the proof but point out that  $\alpha_{hi}(t)$  are solutions of the linear system in Proposition 1.3 with  $v$  replaced by  $e_h$ . Note that the solutions  $\alpha_{hi}$  will contain terms involving the  $e_h$ 's. Note also that the  $\gamma'(t)$  on the right-hand side of the linear system should be replaced by  $L \sum_{j=1}^n y_j(t) e_j(t)$ .

**1.3. The Hilbert Manifold Structure.** Now we supply the space of all indicatrices with a Hilbert manifold structure. There is of course a lot of freedom in defining this structure. The choices made are motivated by on the one hand the form of the functional and on the other hand a desire for computational simplicity. As seen in the introduction we will specialize to  $\int k^2 + \lambda ds$  the modified total squared curvature functional. A variational vector field  $Y$  along  $y$  will satisfy  $Y(0)=0$  since  $y(0)=(1, 0, \dots, 0)$  and it will also be tangent to  $S^{n-1}$ . The curves we consider are sufficiently smooth so that the functional is well defined. Let  $H^1$  denote the Sobolev space of curves from  $I$  into  $R^n$  whose first derivative is square integrable. Let

$$\mathbf{S} = \{y \in H^1 \mid y: I \rightarrow S^{n-1}, y(0)=(1, 0, \dots, 0)\}.$$

On this level the modified total squared curvature functional is:

$$J^\lambda: \mathbf{S} \times R^+ \times OM \rightarrow R$$

given by  $J^\lambda(\hat{y}) = \frac{1}{L} \left( \int_I \langle y'(t), y'(t) \rangle dt \right) + \lambda L$  because  $k^2(t) = \frac{1}{L^2} \langle y'(t), y'(t) \rangle$ . The tangent space to  $\mathbf{S}$  will be identified with  $H^1$  maps  $Y$  into  $R^n$  which satisfy  $Y(t)$  is tangent to  $S^{n-1}$  at  $y(t)$  and  $Y(0)=0$ . The only thing left to supply is the Riemannian structure. For the tangent spaces of  $R^+$  and  $OM$  we use the standard inner product induced from Euclidean space. On  $\mathbf{S}$  we use a particular Riemannian structure suitable for our purposes. Let the standard connection on  $S^{n-1}$  be  $\frac{D}{dt}$  then given  $Y$  and  $\hat{Y}$  in  $TS$  we define:

$$\langle Y, \hat{Y} \rangle = \int_I \left\langle \frac{D}{dt} Y, \frac{D}{dt} \hat{Y} \right\rangle.$$

We are in other words discarding the normal part of the derivatives. This will simplify the gradient computations that follow.

**1.4. Indicatrices Corresponding to Closed Curves.** We now turn our attention to curves satisfying some type of boundary conditions. It turns out that the set of indicatrices corresponding to the curves satisfying the particular boundary conditions under consideration is a Hilbert submanifold of  $\mathbf{S} \times R^+ \times OM$ . To see that this is indeed the case for closed curves we will look at the following map:

$$\Phi: \mathbf{S} \times R^+ \times OM \rightarrow SM \times SM,$$

where  $\Phi(\hat{y}) = \left( (p, e_1)(\hat{y}), \left( \gamma(1), \frac{1}{L} \gamma'(1) \right)(\hat{y}) \right)$  and  $SM$  is the unit tangent bundle of  $M$ . If we let

$$\Delta = \{(x, y) \in SM \times SM \mid x = y\}$$

be the diagonal then the space of closed curves is given to be  $\Omega = \Phi^{-1}(\Delta)$ . The appropriate transversality conditions may now be checked, see for instance [3] (and also [10] for general theory), which will show that  $\Omega$  is a submanifold of the manifold of all indicatrices.

*Remark.* The gradient of  $J^\lambda$  will of course in general not be tangent to  $\Omega$ . The main result of this paper is to show the existence of a set of functionals whose gradients will span the space normal to the tangent space of  $\Omega$ . As for the overall objective, to present an algorithm for the curve-straightening flow, a second important contribution is to show how these spanning gradients can be effectively computed. Moreover it is shown that despite the fact that these gradients are not explicitly known the projection onto the tangent space to the space of indicatrices representing closed curves can nevertheless be determined.

## 2. Projections onto the Tangent Space of Closed Curves

**2.1. A Basis for  $T_y S$ .** By introducing a parallel frame along  $y \in S^{n-1}$  it is possible to express tangent vector fields along  $y$  by its components with respect to this frame. Let  $\{b_1, \dots, b_n\}$  be the standard basis for  $R^n$ . We move this frame so that its origin is at  $y(0) = (1, 0, \dots, 0)$  and then parallel translate this frame along  $y$ . A vector field  $Y$  along  $y$  may now be written as

$$Y = (Y_2, \dots, Y_n) \quad \text{and} \quad \langle Y, Y \rangle = \sum_{r=2}^n \langle Y_r, Y_r \rangle,$$

[since  $Y_r(0) = 0$  we may use the following Sobolev  $H^1$  norm:

$$\langle Y_r, Y_r \rangle = \int_I \langle Y'_r(t), Y'_r(t) \rangle dt].$$

In order to proceed we need an orthonormal basis for

$$H = \{Y_r \in H^1 \mid Y_r(0) = 0\}.$$

Here is one possibility (it is of course not the only possibility but it has some advantages from a numerical standpoint). For  $p = 1, 2, \dots$  let  $P_p(t) = \sum_{q=1}^p c_q(p) t^q$  be given by:

$$c_q(p) = \frac{1}{q} d_{q-1}(p), \quad q = 1, \dots, p,$$

$$d_{-1}(p) = 0, \quad d_0(0) = 1, \quad d_0(1) = -1, \quad d_1(1) = 2,$$

$$d_q(p+1) = 4d_{q-1}(p) - 2d_q(p) - d_q(p-1) - \frac{1}{p+1} [2d_{q-1}(p) - d_q(p) - d_q(p-1)],$$

$$q = 0, \dots, p-1.$$

**Proposition 2.2.** *The collection  $\{P_p\}_{p=1}^\infty$  is a complete orthogonal set in  $H$  and  $\langle P_p, P_p \rangle = \frac{1}{2p-1}$ .*

*Proof.* Let  $\{\hat{Q}_p\}_{p=0}^\infty$  be the Legendre polynomials on  $[-1, 1]$  then  $\hat{Q}_0(t) = 1$ ,  $\hat{Q}_1(t) = t$ , and

$$\hat{Q}_{p+1}(t) = 2t\hat{Q}_p(t) - \hat{Q}_{p-1}(t) - \frac{1}{p+1}[t\hat{Q}_p(t) - \hat{Q}_{p-2}(t)]$$

so  $\langle \hat{Q}_p, \hat{Q}_p \rangle_{L^2} = \frac{2}{2p+1}$  in the standard  $L^2$  inner product on  $[-1, 1]$  (see [9]). We now shift these polynomials to  $I$  by  $Q(t) = \hat{Q}(2t-1)$  so now  $Q_0(t) = 1$ ,  $Q_1(t) = 2t-1$ , and

$$Q_{p+1}(t) = (4t-2)Q_p(t) - Q_{p-1}(t) - \frac{1}{p+1}[(2t-1)Q_p(t) - Q_{p-2}(t)]$$

and  $\langle Q_p, Q_p \rangle_{L^2} = \frac{1}{2p+1}$  in the standard  $L^2$  inner product on  $I$ . By defining  $d_q(p)t^q$

from  $Q_p(t) = \sum_{q=0}^p d_q(p)t^q$  and collecting coefficients of equal powers of  $t$  generates the

recursive relations for the  $d_q(p)$ 's. Now let  $P_p(t) = \int_0^t Q_{p-1}(s)ds$ . It is clear that  $P_p$  is in  $H$  and  $\langle P_p, P_q \rangle = \langle Q_{p-1}, Q_{q-1} \rangle_{L^2}$ . It follows that  $\{P_p\}_{p=1}^\infty$  is orthogonal and the norm relation follows. To show completeness we let  $Y_r \in H$  be such that  $\langle P_p, Y_r \rangle = \langle Q_{p-1}, Y_r' \rangle_{L^2} = 0$  for all  $p$ . By the completeness of the Legendre polynomials it must be that  $Y_r' = 0$  in  $L^2$  but then  $Y_r = 0$  in  $H$ .

*Remark.* Here are the first ten:

$$\begin{aligned} & t, t^2 - t, \quad 2t^3 - 3t^2 + t, \quad 5t^4 - 10t^3 + 6t^2 - t, \\ & 14t^5 - 35t^4 + 30t^3 - 10t^2 + t, \quad 42t^6 - 126t^5 + 140t^4 - 70t^3 + 15t^2 - t, \\ & 132t^7 - 462t^6 + 630t^5 - 420t^4 + 140t^3 - 21t^2 + t, \\ & 429t^8 - 1716t^7 + 2772t^6 - 2310t^5 + 1050t^4 - 252t^3 + 28t^2 - t, \\ & 1430t^9 - 6435t^8 + 12012t^7 - 12012t^6 + 6930t^5 - 2310t^4 + 420t^3 - 36t^2 + t, \\ & 4862t^{10} - 24310t^9 + 51480t^8 - 60060t^7 + 42042t^6 - 18018t^5 \\ & \quad + 4620t^4 - 660t^3 + 45t^2 - t. \end{aligned}$$

**2.2. A Basis for  $T_\omega OM$ .** We now turn our attention to  $OM$ . Again we need to find a basis for the tangent space. Recall that  $OM = G^{-1}(0)$ , where  $G$  was defined in Sect. 1.1. Our goal is to find a basis that may be expressed in terms of  $F$  and derivatives of  $F$ . As a first step we compute the derivative of  $G$ . It will be convenient to talk about block matrices building the  $(l + nl + n + m) \times (k + nk)$  matrix representing  $DG$ . We will let  $F_i$  denote the  $i^{\text{th}}$  component function of  $F$ , so there are  $l$  such functions. The hessian of  $F_i$ , i.e. the symmetric matrix of second partial derivatives of  $F_i$  is denoted by  $HF_i$ . Our block matrices will be as follows:  $DF(p)$  is  $l \times k$ ,  $e_j$  is  $k \times 1$ ,  $HF_i e_j$  is  $1 \times k$  and finally  $H_j$  is  $l \times k$  and its  $i^{\text{th}}$  row is  $HF_i e_j$ . We now describe  $DG$ . It splits into an upper  $(l + nl) \times (k + nk)$  matrix and a lower  $(n + m) \times (k + nk)$  matrix. The upper matrix has  $n + 1$  blocks  $DF(p)$  along the diagonal. Underneath



the leftmost  $DF(p)$  is  $H_1$  and then  $H_2$  so on and so forth until  $H_n$ . All remaining entries are zero. The lower matrix has all zeroes in the leftmost  $(n+m) \times k$  block. The remaining part splits into an upper  $n \times nk$  block and a lower  $m \times nk$  block. The upper block has  $2e_i^T$  along the diagonal starting with  $e_1^T$  in the upper left-hand corner and ending with  $e_n^T$  in the lower right-hand corner and zeroes elsewhere. Finally the lower matrix has  $e_2^T$  in the upper left-hand corner  $e_1^T$  as its right-hand neighbor. Underneath  $e_2^T$  is  $e_3^T$  and underneath  $e_1^T$  is a block of zeroes and then follows (if possible)  $e_1^T$ . This way of building  $e_i^T$ 's diagonally continues until there is no more room. The whole process (continuing going down) now starts all over shifted  $k$  entries over to the right and now  $e_3^T$  will have  $e_2^T$  as its right neighbor and so on and so forth. We illustrate the above with an example when  $n=3$ .

$$DG(p) = \begin{pmatrix} DF(p) & 0 & 0 & 0 \\ H_1 & DF(p) & 0 & 0 \\ H_2 & 0 & DF(p) & 0 \\ H_3 & 0 & 0 & DF(p) \\ 0 & 2e_1^T & 0 & 0 \\ 0 & 0 & 2e_2^T & 0 \\ 0 & 0 & 0 & 2e_3^T \\ 0 & e_2^T & e_1^T & 0 \\ 0 & e_3^T & 0 & e_1^T \\ 0 & 0 & e_3^T & e_2^T \end{pmatrix}$$

The tangent spaces of  $OM$  are identified with linear subspaces of  $R^k \times R^{nk}$ .

**Proposition 2.3.** *A basis for the tangent space of  $OM$  at  $\omega=(p, \{e_1, \dots, e_n\})$  is given by:*

$$\begin{aligned} &\{(e_1, v_{11}, \dots, v_{n1}), (e_2, v_{12}, \dots, v_{n2}), \dots, (e_n, v_{1n}, \dots, v_{nn}), \\ &(0, e_2, -e_1, 0, \dots, 0), (0, e_3, 0, -e_1, 0, \dots, 0), \dots, (0, e_n, 0, \dots, 0, -e_1), \\ &(0, 0, e_3, -e_2, 0, \dots, 0), (0, 0, e_4, 0, -e_2, 0, \dots, 0), \dots, (0, \dots, 0, e_n, -e_{n-1})\}, \end{aligned}$$

where the vectors  $v_{ij}=a_{1ij}DF_1(p)^T + \dots + a_{lij}DF_l(p)^T$  and  $x_{ij}=(a_{1ij}, \dots, a_{lij})$  is a solution of the following linear equation  $Ax_{ij}=-b_{ij}$ :

$$A = \begin{pmatrix} DF_1(p)DF_1(p)^T & DF_1(p)DF_2(p)^T & \dots & DF_1(p)DF_l(p)^T \\ DF_2(p)DF_1(p)^T & DF_2(p)DF_2(p)^T & \dots & DF_2(p)DF_l(p)^T \\ \dots & \dots & \dots & \dots \\ DF_l(p)DF_1(p)^T & DF_l(p)DF_2(p)^T & \dots & DF_l(p)DF_l(p)^T \end{pmatrix},$$

$$b_{ij} = \begin{pmatrix} e_i^T H F_1 e_j \\ e_i^T H F_2 e_j \\ \dots \\ e_i^T H F_l e_j \end{pmatrix}.$$

*Proof.* First note that since  $DF(p)$  is assumed onto it can be shown that  $A$  is invertible. Next counting the number vectors in the list we see that it equals  $\frac{n(n+1)}{2}$  (i.e. the dimension of  $OM$ ) so the proof is complete if we can show linear independence and tangency. To show linear independence we observe that the last  $m$  vectors are orthogonal among each other and also orthogonal to each of the first  $n$  vectors. Next note that the  $e_i$ 's are linearly independent it follows that the first  $n$  vectors are linearly independent and hence all  $n+m$  vectors. To show tangency we apply  $DG$  to each vector. We see that they are all annihilated so they are indeed in the tangent space of  $OM$ .

**2.3. The Gradient of  $J^\lambda$ .** Recall that given any smooth functional  $F: X \rightarrow \mathbb{R}$  on a Hilbert manifold  $X$ , we define the gradient  $\nabla F \in TX$  by  $DF(p)v = \langle \nabla F(p), v \rangle$  for all  $(p, v) \in TX$ . Let  $\nabla J^\lambda(\hat{y}) = (\alpha, \alpha_L, \alpha_{OM})$ . Here we write  $\alpha = (\alpha_2, \dots, \alpha_n)$  to indicate the components with respect to the parallel frame along  $y$ . The gradient of  $J^\lambda$  satisfies the following defining relation:

$$DJ^\lambda(\hat{y})v = \langle \nabla J^\lambda(\hat{y}), v \rangle = \langle \alpha, v_y \rangle + \alpha_L v_L + \alpha_{OM} \cdot v_{OM}$$

for all  $v = (v_y, v_L, v_{OM})$  in  $T(\mathbb{S} \times \mathbb{R}^+ \times OM)$ . The dot product used is the standard Euclidean and the first inner product is given by the sum of the Sobolev inner products in the factors  $\alpha_2, \dots, \alpha_n$ . Now suppose  $y' = (Y_2, \dots, Y_n)$  then the gradient

$\nabla J^\lambda(\hat{y})$  is given by  $\alpha_r = \frac{2}{L} Y_r$ ,  $r \in \{2, \dots, n\}$ ,  $\alpha_L = \lambda - \frac{1}{L^2} \sum_{r=2}^n \int_I Y_r^2(t) dt$  and  $\alpha_{OM} = 0$ .

When  $\alpha_r$  is expressed with respect to the orthogonal basis  $\{P_p\}_{p=1}^\infty$  we write  $\alpha_r^p$  for the ‘‘Fourier’’ coefficients. They may be computed by

$$\alpha_r^p = (2p-1) \int_I \alpha_r'(t) Q_{p-1}(t) dt.$$

**2.4. The Main Result.** We now take full advantage of the fact that our manifolds are embedded in Euclidean space with its linear structure. Consider:

$$A: \mathbb{S} \times \mathbb{R}^+ \times OM \rightarrow \mathbb{R}^k$$

given by  $A(\hat{y}) = (\gamma(1) - p)(\hat{y})$ . The dependence of  $\gamma(1)$  on  $\hat{y}$  is given by Proposition 1.4. Next we consider:

$$\hat{A}: \mathbb{S} \times \mathbb{R}^+ \times OM \rightarrow \mathbb{R}^k$$

given by  $\hat{A}(\hat{y}) = \left( \frac{1}{L} \gamma'(1) - e_1 \right)(\hat{y})$ . Note that  $\frac{1}{L} \gamma'(1) = \sum_{i=1}^n y_i(1) e_i(1)$  and the dependence of the  $e_i$ 's on  $\hat{y}$  is also given by Proposition 1.4. Now let  $\{A_{ij}\}_{j=1}^k$  and  $\{\hat{A}_{ij}\}_{j=1}^k$  denote the  $k$  component functions of  $A$  and  $\hat{A}$ . Using notation similar to the above we write

$$\nabla A_j(\hat{y}) = (\beta, \beta_L, \beta_{OM}), \quad \nabla \hat{A}_j(\hat{y}) = (\hat{\beta}, \hat{\beta}_L, \hat{\beta}_{OM}),$$

so we drop the  $j$ 's for the time being. We also write  $\beta = (\beta_2, \dots, \beta_n)$ ,  $\hat{\beta} = (\hat{\beta}_2, \dots, \hat{\beta}_n)$  for the components with respect to the parallel frame along  $y$ . Using the basis from Proposition 2.3 we write  $\beta_{OM} = (\beta_{OM}^1, \dots, \beta_{OM}^{n+m})$ ,  $\hat{\beta}_{OM} = (\hat{\beta}_{OM}^1, \dots, \hat{\beta}_{OM}^{n+m})$ . Since the basis is not orthonormal we also need the Euclidean dot products between the different vectors in the basis. If  $(w_1, \dots, w_{n+m})$  denotes the basis, let  $c_{ij} = w_i \cdot w_j$  for  $i, j$  in  $\{1, \dots, n+m\}$  denote these products. Note that most  $c_{ij}$ 's are in fact zero and  $c_{ij} = c_{ji}$ . We will use the basis for these Sobolev spaces to write  $\beta_r = \sum_{p=1}^\infty \beta_r^p P_p$ ,

$\hat{\beta}_r = \sum_{p=1}^{\infty} \hat{\beta}_r^p P_p$ . Let  $W_r^p \in TS$  with  $r \in \{2, \dots, n\}$  be given by  $W_r^p = (0, \dots, P_p, \dots, 0)$  with  $P_p$  in the  $(r-1)^{\text{th}}$  position out of the  $(n-1)$  possible.

**Theorem 2.4.** (i) *Given any  $X \in T(S \times R^+ \times OM)$  the derivatives  $DA_j(\hat{y})X$  and  $D\hat{A}_j(\hat{y})X$  can be found by solving an initial value problem for a system of ordinary differential equations.*

(ii) *The gradients can be computed from the following derivatives (the  $j$  subscripts are dropped on the right-hand sides of the equations):*

$$\begin{aligned} DA_j(\hat{y})(W_r^p, 0, 0) &= \beta_r^p, & D\hat{A}_j(\hat{y})(W_r^p, 0, 0) &= \hat{\beta}_r^p, \\ DA_j(\hat{y})(0, 1, 0) &= \beta_L, & D\hat{A}_j(\hat{y})(0, 1, 0) &= \hat{\beta}_L, \\ DA_j(\hat{y})(0, 0, w_s) &= \sum_{i=1}^{n+m} \beta_{OM}^i c_{is}, & D\hat{A}_j(\hat{y})(0, 0, w_s) &= \sum_{i=1}^{n+m} \hat{\beta}_{OM}^i c_{is}. \end{aligned}$$

(iii) *The collection  $\{\nabla A_j, \nabla \hat{A}_j\}_{j=1}^k$  span the normal space  $T\Omega^\perp$  to the tangent space  $T\Omega$  of the space of closed curves.*

(iv) *Projection of any  $X \in T_{\hat{y}}(S \times R^+ \times OM)$  onto  $T\Omega$  is done as follows. By (iii) there are some scalars  $\lambda_j$  and  $\hat{\lambda}_j$  so that  $X^{\tan} = X - \sum_{j=1}^k (\lambda_j \nabla A_j(\hat{y}) + \hat{\lambda}_j \nabla \hat{A}_j(\hat{y}))$  is tangent to  $\Omega$ .*

*Let  $X^{\tan} = (X_y, X_L, X_{OM})$  be the different factors of  $X^{\tan}$ . Use the embedding of  $OM$  to write  $X_{OM} = (X_p, X_{e_1}, \dots, X_{e_n})$ . Determine the  $\lambda_j, \hat{\lambda}_j$  by finding a solution of the under-determined system*

$$\begin{aligned} DA(\hat{y})X^{\tan} &= X_p, \\ D\hat{A}(\hat{y})X^{\tan} &= X_{e_1}. \end{aligned}$$

*Proof.* (i) Recall that  $A(\hat{y}) = (\gamma(1-p)(\hat{y}))$  and  $\hat{A}(\hat{y}) = \left(\frac{1}{L}\gamma'(1)-e_1\right)(\hat{y})$  and Proposition 1.4 gives the initial value problem relating  $\gamma$  and  $\{e_1, \dots, e_n\}$  to  $\hat{y}$ . Note that  $\hat{y}$  enters both into the equations as well as into the initial conditions. If we evaluate the solutions of this initial value problem at  $t=1$  then this creates a map

$$\Psi: S \times R^+ \times OM \rightarrow OM$$

that may be differentiated to get a map

$$D\Psi(\hat{y}): T(S \times R^+ \times OM) \rightarrow TOM.$$

The point now is that if we can compute this derivative in arbitrary directions then we can also compute the derivatives  $DA(\hat{y})$  and  $D\hat{A}(\hat{y})$ . We postpone the details of these computations to Sect. 3.2 (Step 11–12). But remark that it is possible to solve the algebraic linear systems involved once and for all and this would then leave us with an initial value problem.

(ii) By the above remarks and by using the bases in 2.1 and 2.2 we compute the corresponding directional derivatives. The defining relations for the gradients then gives the expressions in the theorem. The only remaining issue is to determine the  $\beta_{OM}^i$ 's for  $i \in \{1, \dots, n+m\}$ . We observe that the above gives  $n+m$  linear equations. This system has a unique solution because the matrix with entries consisting of the dot products of the  $w_i$ 's is indeed invertible.

(iii) Suppose there is some  $X \in T\Omega^\perp$  such that  $\langle \nabla A_j(\hat{y}), X \rangle = \langle \nabla \hat{A}_j(\hat{y}), X \rangle = 0$  for all  $j=1, \dots, k$ . We want to show that  $X=0$ . First observe that  $DA_j(\hat{y})X$

$= D\hat{A}_j(\hat{y})X = 0$  for  $j = 1, \dots, k$ . It follows that  $DA(\hat{y})X = D\hat{A}(\hat{y})X = 0$ . The transversality of  $\Phi$  (see the beginning of Sect. 1.4) implies (see [10]) that  $T\Omega^\perp$  and  $T\Delta^\perp$  have the same dimension. It also follows that

$$D\Phi(T\Omega^\perp) + T\Delta = T(SM \times SM)$$

so if  $v \in T\Omega^\perp$  satisfies  $D\Phi(\hat{y})v \in T\Delta$  then it must be that  $v = 0$ . Since

$$DA(\hat{y})X = D(\gamma(1) - p)(\hat{y})X = D\gamma(1)X - Dp(\hat{y})X = 0$$

and

$$D\hat{A}(\hat{y})X = D\left(\frac{1}{L}\gamma'(1) - e_1\right)(\hat{y})X = D\frac{1}{L}\gamma'(1)(\hat{y})X - De_1(\hat{y})X = 0$$

we see that  $D\Phi(\hat{y})X$  is of the form  $(x, x)$ . This means that  $D\Phi(\hat{y})X \in T\Delta$  and therefore  $X = 0$ .

(iv) Recall that the space of indicatrices corresponding to closed curves is given by  $\Omega = \Phi^{-1}(\Delta)$ , where  $\Delta$  is the diagonal in  $SM \times SM$ . If a vector is tangent to  $\Omega$  then its image under  $D\Phi$  will be tangent to  $\Delta$ . Since such vectors are of the form  $(x, x)$  the definition of  $\Phi$  when compared with the definitions of  $\Delta$  and  $\hat{A}$  yields the statement (iv) in the theorem.

*Remark.* In [4] the gradients are computed explicitly in the case of planar curves. For curves in arbitrary manifolds these gradients may not be computed explicitly (in fact already in the standard sphere  $S^2$  things are sufficiently complicated). The main reason for this is how the geometry enters in to the functionals, in our case via parallel translation along the curve. The point with Theorem 2.4 is that nevertheless it is possible to get expressions for the gradients provided certain derivatives are computed. In the next section we will show how the computation of these derivatives is reduced to solving initial value problems for certain ordinary differential equations.

### 3. The Curve-Straightening Algorithm

**3.1. Initializing.** We will list the main steps needed and we begin with:

1. Use Proposition 2.2 to generate the Legendre polynomials  $\{Q_p\}_{p=0}^\infty$  and also the complete set  $\{P_p\}_{p=1}^\infty$ .

This basis will be instrumental in computing the various gradients because we will in effect compute the “Fourier” coefficients with respect to the basis. This is a major difference from what was done in [3] and [4] where the duBois-Reymonds lemma was the principal tool in determining the gradients. Superscript  $p$  is used for the components with respect to  $P_p$ .

We will assume that  $F: R^k \rightarrow R^l$  together with its partial derivatives are given (there is no need for more than third partials in what follows). An initial curve  $\gamma: [a, b] \rightarrow M = F^{-1}(0)$  is also assumed given. Since the flow will take place on the space of indicatrices rather than on the space of curves the indicatrix corresponding to  $\gamma$  must be computed.

2. Compute the length  $L$  of  $\gamma$

$$L = \int_a^b |\gamma'(u)| du.$$

Next we need to parametrize proportional to arclength. Recall that the indicatrices are defined on  $I=[0,1]$  so we need a reparametrizing  $1-1$  map  $\sigma: I \rightarrow [a,b]$  so that  $|\gamma'(\sigma(t))|=L$  for all  $t \in I$ .

3. Find  $\sigma$  by solving the initial value problem:

$$\sigma'(t) = \frac{L}{|\gamma'(\sigma(t))|}, \quad \sigma(0) = a.$$

In order to find the indicatrix we proceed as in Sect. 1.2 to generate an orthonormal basis  $\{e_1, \dots, e_n\}$  of the tangent space of  $M$  at  $\gamma(a)$ .

4. Use projections as in Proposition 1.1 and Gram-Schmidt to find  $\{e_1, \dots, e_n\}$ .

Next we need the corresponding parallel frames along  $\gamma$ :

5. Use Proposition 1.3 to get the parallel frames  $\{e_1(t), \dots, e_n(t)\}$ , i.e.  $\{e_1(t), \dots, e_n(t)\}$  is an orthonormal basis for the tangent space of  $M$  at  $\gamma(\sigma(t))$ .

Let  $\hat{y}=(y, L, \omega)$  denote the initial indicatrix so that  $\omega=(\gamma(a), \{e_1, \dots, e_n\})$ . Let

$$y(t) = (y_1(t), \dots, y_n(t)).$$

6. The initial components are given by

$$y_i(t) = \frac{\gamma'(t)}{L} \cdot e_i(t), \quad i = 1, \dots, n.$$

**3.2. Steps Needed to Compute the Gradients.** The flow in the direction of steepest descent of  $J^\lambda$  will require repeated computations of the various gradients and therefore the subsequent steps starting with Step 7 will be part of an iteration. Recall that we use another set of parallel orthonormal frames along  $y$ , and with respect to this frame, tangent vectors to the space of indicatrices will have  $n-1$  components. Note that at this stage we need to know  $y$  at arbitrary points so  $\{P_p\}_{p=1}^\infty$  is used here.

7. Adapt Proposition 1.3 to the case  $S^{n-1}$  to compute the parallel frame  $\{f_2, \dots, f_n\}$  along  $y$  and we will use a subscript  $r$  to denote the components with respect to these frames.

To compute the gradient of  $J^\lambda$  we need:

8. Express  $y'$  with respect to  $\{f_2, \dots, f_n\}$  and we get  $y'(t) = (Y_2(t), \dots, Y_n(t))$  so

$$Y_r(t) = y'(t) \cdot f_r(t) \quad \text{for } r = 2, \dots, n.$$

9. Let  $Y_r^p = (2p-1) \int_0^1 Y_r(t) Q_{p-1}(t) dt$  then these are the coefficients of  $Y_r$  with respect to  $\{P_p\}_{p=1}^\infty$ .

With notation as in Sect. 2.3 we have:

10. The gradient  $\nabla J^\lambda(\hat{y})$  is given by  $\alpha_r^p = \frac{2}{L} Y_r^p$ ,  $\alpha_L = \lambda - \frac{1}{L^2} \int \sum_{r=2}^n Y_r^2(t) dt$ , and  $\alpha_{OM} = 0$ .

We now get to the heart of the algorithm. The following steps will show how the remaining gradients are computed.

11. Use Proposition 1.4 to compute  $\gamma(t)$  and  $\{e_1(t), \dots, e_n(t)\}$ . This step is of course redundant the first time around.

Note that Proposition 1.4 for each  $t \in I$  gives a map from  $S \times R^+ \times OM$  into  $R^k \times R^{nk}$ , but its values are always in  $OM$  as a submanifold of  $R^k \times R^{nk}$ . It follows that a variation of the indicatrix  $\hat{y}$  creates a variational vector field along a curve in

*OM*. The components of this vector field may be given either with respect to the basis of *OM* or with respect to the basis of  $R^k \times R^{nk}$ . In the latter case we denote the components by  $\frac{\partial \gamma}{\partial w}$  and  $\frac{\partial e_h}{\partial w}$  for  $h = 1, \dots, n$ . Write  $v_y = (v_{y_2}, \dots, v_{y_n})$  using the frames along  $y$  and

$$v_w = v_1 w_1 + \dots + v_{n+m} w_{n+m} = (v_p, v_{e_1}, \dots, v_{e_n})$$

for the factors in  $R^k \times R^{nk}$ . Let  $\frac{\partial y_h}{\partial w} = (v_{y_2} f_2 + \dots + v_{y_n} f_n) \cdot b_h$  for  $h = 1, \dots, n$ ; here  $\{b_1, \dots, b_n\}$  is the standard basis in  $R^n$ . The variational vector fields are solutions of the following initial value problem:

$$\begin{aligned} \frac{\partial \gamma'}{\partial w}(t) &= v_L \sum_{h=1}^n y_h(t) e_h(t) + L \sum_{h=1}^n \left( \frac{\partial y_h}{\partial w} e_h(t) + y_h(t) \frac{\partial e_h}{\partial w} \right), \\ \frac{\partial e'_h}{\partial w}(t) &= \sum_{i=1}^l \left( \frac{\partial \alpha_{hi}}{\partial w} DF_i(\gamma(t)) + \alpha_{hi}(t) HF_i(\gamma(t)) \frac{\partial \gamma}{\partial w} \right). \end{aligned}$$

The initial conditions are given by  $\frac{\partial \gamma}{\partial w}(0) = v_p$  and  $\frac{\partial e_h}{\partial w}(0) = v_{e_n}$  for all  $h = 1, \dots, n$ . The functions  $\frac{\partial \alpha_{hi}}{\partial w}$  are determined by differentiating the solutions of the linear system in Proposition 1.3 with respect to the variational parameter  $w$ , and then evaluate at zero. The equation looks as follows:

$$\begin{aligned} & \frac{\partial \alpha_{h1}}{\partial w} \langle DF_1(\gamma(t)), DF_f(\gamma(t)) \rangle + \dots + \frac{\partial \alpha_{hl}}{\partial w} \langle DF_l(\gamma(t)), DF_f(\gamma(t)) \rangle \\ &= - \left\langle \frac{\partial e_h}{\partial w} HF_f(\gamma(t)) \gamma'(t) \right\rangle \\ & \quad - \alpha_{h1} \left( \left\langle HF_1(\gamma(t)) \frac{\partial \gamma}{\partial w}, DF_f(\gamma(t)) \right\rangle + \left\langle DF_1(\gamma(t)), HF_f(\gamma(t)) \frac{\partial \gamma}{\partial w} \right\rangle \right) - \dots \\ & \quad - \alpha_{hl} \left( \left\langle HF_l(\gamma(t)) \frac{\partial \gamma}{\partial w}, DF_f(\gamma(t)) \right\rangle + \left\langle DF_l(\gamma(t)), HF_f(\gamma(t)) \frac{\partial \gamma}{\partial w} \right\rangle \right) \\ & \quad - \left\langle e_h, DHF_f(\gamma(t)) \left\{ \gamma'(t), \frac{\partial \gamma}{\partial w} \right\} + HF_f(\gamma(t)) \frac{\partial \gamma'}{\partial w}(t) \right\rangle. \end{aligned}$$

Note that at this stage there is a need to know the third partial derivatives of  $j^{\text{th}}$  component function  $F_j$ . They show up in the map  $DHF_j(\gamma(t))$  which appears as follows. The map  $HF_j$  is a map from  $R^k$  into  $L(R^k; R^k)$  the space of linear maps on  $R^k$ . Its derivative  $DHF_j$  is hence a map from  $R^k$  into  $L(R^k; L(R^k; R^k))$ , using the same notation. It follows that  $DHF_j(\gamma(t))$  is a linear map from  $R^k$  into  $L(R^k; R^k)$  and  $DHF_j(\gamma(t))\gamma'(t)$  is a linear map from  $R^k$  to  $R^k$ , and we write  $DHF_j(\gamma(t)) \left\{ \gamma'(t), \frac{\partial \gamma}{\partial w} \right\}$  for this map's value at  $\frac{\partial \gamma}{\partial w}$ . Note also that  $\frac{\partial \gamma'}{\partial w}(t)$  in the last term in the previous equation should be replaced by

$$v_L \sum_{h=1}^n y_h(t) e_h(t) + L \sum_{h=1}^n \left( \frac{\partial y_h}{\partial w} e_h(t) + y_h(t) \frac{\partial e_h}{\partial w} \right).$$

12. By evaluating solutions of the above system at  $t=1$  we compute all the derivatives needed in Theorem 2.4(ii).
13. Use the derivatives in Step 12 and Theorem 2.4(ii) to determine the gradients  $\nabla A_j$  and  $\nabla \hat{A}_j$ .
14. Project  $\nabla J^\lambda(\hat{y})$  onto the tangent space of indicatrices representing closed curves by finding the tangential part from solving the linear algebraic system in Theorem 2.4(iv).

We are now in a position to flow. As an example the simplest possibility would be to take a short step in the negative direction of the projected gradient. This would give us an approximate new indicatrix  $\hat{y}$  and we would then proceed with Step 7 as before. To be precise let  $\hat{y}_0$  be the previous indicatrix and let the approximate new indicatrix be denoted by  $\hat{y}_h$ . Suppose the projected gradient is  $(v_y, v_L, v_\omega)$ , then its components are given by  $v_y = (v_{y_2}, \dots, v_{y_n})$ , where each of these components are computed from the Hilbert space basis components in the standard way. An approximate flow in the negative direction is now given by:

$$\begin{aligned}\frac{y_h - y_0}{h} &= -(v_{y_2}f_2 + \dots + v_{y_n}f_n), \\ \frac{L_h - L_0}{h} &= -v_L, \\ \frac{\omega_h - \omega_0}{h} &= -(v_1w_1 + \dots + v_{n+m}w_{n+m}).\end{aligned}$$

This approximate flow will give a new  $y_h$  that is a function that is not quite in  $S^{n-1}$  and also an  $\omega$  that is not quite in  $OM$ . A possible alternative would be to flow a distance  $h$  along geodesics in  $S^{n-1}$  and  $OM$  instead of along lines in the tangent spaces. We leave problems such as these (they really belong to the field of numerical analysis) as sources for future investigations.

## 4. The Second Variation Along Geodesics

**4.1. Background.** This last section will initiate a study of the stability of the closed geodesics under the total squared curvature flow. We will give some partial results but there are many open questions still to be answered. Recall that the flow is known to satisfy the Palais-Smale condition only when the parameter  $\lambda$  is greater than zero. On the other hand if  $\lambda=0$  then all closed geodesics are global minima but in this case it is not known if all the trajectories of the negative gradient flow converges to critical points.

On  $S^2$ , the standard two-sphere of constant Gaussian curvature  $G$ , it is known (see [3]) that if  $0 \leq \lambda \leq \frac{8G}{7}$  then all non-geodesic critical points are unstable (i.e. they are not local minima). Moreover the second variation formula is positive semi-definite at the  $n$ -fold great circle if and only if  $\frac{2n-1}{n^2} \geq \frac{\lambda}{2G}$ . We will discuss the second variation formula evaluated at closed geodesics in two-dimensional manifolds of arbitrary Gaussian curvature. This formula remarkably turns out to be identical to the second variation formula in the constant Gaussian curvature case when it is evaluated along a closed geodesic.

#### 4.2. The Second Variation Formula along Geodesics in Two-Dimensional Manifolds.

We will not go into the lengthy details of the derivation of the second variation formula. The main reference is [8] Lemma 1.1 and Formula 1.4 and its proof. We work on the level of curves rather than on the level of indicatrices. Consider a variation

$$\gamma = \gamma_w = \gamma(w, t) : (-\varepsilon, \varepsilon) \times I \rightarrow M$$

of closed curves such that  $\gamma_0 = \gamma(0, t)$  is a closed geodesic and  $M$  is two-dimensional. With

$$\mathbf{F}^2(\gamma_w) = \int_{\gamma_w} k^2 + \lambda ds$$

and variational vector field  $W(w, t) = \frac{\partial \gamma_w}{\partial w}(w, t)$  along  $\gamma_w$ . The tangential part of the variation will not show up in the second variation formula so we write  $W = \phi N$ , where  $N$  is the unit normal to the curve. With  $\dot{\phi}$  denoting the derivative we have:

$$\frac{d^2}{dw^2} \mathbf{F}^2(\gamma_w)|_{w=0} = \int_{\gamma_0} 2\ddot{\phi}^2 - (4G - \lambda)\dot{\phi}^2 + G(2G - \lambda)\phi^2 ds;$$

here  $G$  is the Gaussian curvature along the curve. Note that surprisingly this is exactly the same as Formula 1.4 in [8] when the geodesic curvature  $k$  equals zero. To see why this is the case we list the differences in the proof,

$$W(k)_{ss} = \text{same as before} + 2\dot{\phi}\dot{G} + \phi\ddot{G}.$$

This difference propagates through the computation until the end where an integration by parts cancels the two new terms because for closed curves:

$$\int_{\gamma} \phi^2 \ddot{G} ds = - \int_{\gamma} 2\phi \dot{\phi} \dot{G} ds.$$

To see how the length  $L$  of the curve enters into the formula an alternative formula is:

$$\frac{d^2}{dw^2} \mathbf{F}^2(\gamma_w)|_{w=0} = \frac{1}{L} \int_I \frac{2\ddot{\phi}^2}{L^2} - (4G - \lambda)\phi^2 + L^2 G(2G - \lambda)\phi^2 dt.$$

**4.3. Some Consequences of the Second Variation Formula.** If we consider a surface of revolution then a circular closed geodesic will have constant Gaussian curvature along itself. In fact if  $(x(u), y(u))$  is revolved about the  $x$ -axis then if we again let  $\dot{y}, \dot{x}$  denote the derivatives it must be that  $\dot{y} = 0$  at such a geodesic. It follows that

$$G = - \frac{\ddot{y}}{x\dot{y}^2}$$

along such a geodesic. Suppose first of all that  $\lambda = 0$ , then after integration by parts we may write:

$$\frac{d^2}{dw^2} \mathbf{F}^0(\gamma_w)|_{w=0} = \frac{2}{L} \int_I \left( \frac{\ddot{\phi}}{L} + LG\phi \right)^2 dt.$$

Moreover this is zero only if the curvature and the length are related by  $L^2 G = (2\pi n)^2$  for some integer  $n$ . Since  $L = p2\pi y$  for some positive integer  $p$  it follows that

$$\frac{\ddot{y}}{\dot{x}^2} = - \frac{n^2}{p^2}.$$



Hence the second variation is positive definite generically for positive Gaussian curvature and always for negative Gaussian curvature at the circular geodesics on a surface of revolution. It follows for instance that on a non-spherical ellipsoid of revolution the second variation is positive definite along any  $p$ -fold circular geodesic provided the ratio of the two axes is not  $\frac{n}{p}$  for any integer  $n$ . Hence for sufficiently small positive  $\lambda$  it will remain positive definite. By perturbing the ellipsoid of revolution we may also find ellipsoids so that the second variation is positive definite along the shortest as well as the longest simple closed geodesic.

In conclusion we remark that it is exactly the facts sampled above that make curve-straightening a competitive alternative to other flows used to find closed geodesics. For instance both the length and the energy functional tend to flow curves away from the closed geodesics on ellipsoids.

## References

1. Akhiezer, N.I.: The calculus of variations, pp. 37–45. Blaisdell Publishing 1962
2. Palais, R.S., Terng, C.: Critical point theory and submanifold geometry. Lect. Notes in Math., vol. 1353, pp. 233–259. Berlin, Heidelberg, New York: Springer 1988
3. Langer, J., Singer, D.A.: Curve-straightening in Riemannian manifolds. *Ann. Global Anal. Geom.* **5**, 133–150 (1987)
4. Linnér, A.: Some properties of the curve straightening flow in the plane. *Trans. Am. Math. Soc.* **314**, 605–617 (1989)
5. Gage, M., Hamilton, R.: The heat equation shrinking convex plane curves. *J. Diff. Geom.* **23**, 69–96 (1986)
6. Grayson, M.: The heat equation shrinks embedded plane curves to round points. *J. Diff. Geom.* **26**, 285–314 (1987)
7. Birkhoff, G.D.: Dynamical systems with two degrees of freedom. *Trans. Am. Math. Soc.* **18**, 199–300 (1917)
8. Langer, J., Singer, D.A.: The total squared curvature of closed curves. *J. Diff. Geom.* **20**, 1–22 (1984)
9. Arfken, G.: Mathematical methods for physicists. 2<sup>nd</sup> ed., pp. 534–558. New York: Academic Press 1970
10. Abraham, R., Marsden, J.E., Ratiu, T.: Manifolds, tensor analysis and applications. 2<sup>nd</sup> ed., pp. 201–203. Berlin, Heidelberg, New York: Springer 1988

Communicated by S.-T. Yau

