1

First and second variational formulas for area

In this chapter, we will derive the first and second variational formulas for the area of a submanifold. This will be useful in our later discussion on the volume and Laplacian comparison theorems. It will also be used in our studies of the stability issues of minimal submanifolds.

Let M be a Riemannian manifold of dimension m with metric denoted by ds^2. In terms of local coordinates $\{x_1, \ldots, x_m\}$ the metric is written in the form

$$ds^2 = g_{ij}\, dx_i\, dx_j,$$

where we are adopting the convention that repeated indices are being summed over. If X and Y are tangent vectors at a point $p \in M$, we will also denote their inner product by

$$ds^2(X, Y) = \langle X, Y\rangle.$$

If we let $\mathcal{S}(TM)$ be the set of smooth vector fields on M, then the Riemannian connection $\nabla : \mathcal{S}(TM) \times \mathcal{S}(TM) \to \mathcal{S}(TM)$ satisfies the following properties:

(1) $\nabla_{(f_1 X_1 + f_2 X_2)} Y = f_1 \nabla_{X_1} Y + f_2 \nabla_{X_2} Y$;
(2) $\nabla_X (f_1 Y_1 + f_2 Y_2) = X(f_1) Y_1 + f_1 \nabla_X Y_1 + X(f_2) Y_2 + f_2 \nabla_X Y_2$;
(3) $X\langle Y_1, Y_2\rangle = \langle \nabla_X Y_1, Y_2\rangle + \langle Y_1, \nabla_X Y_2\rangle$; and
(4) $\nabla_X Y - \nabla_Y X = [X, Y]$, for all $X, Y \in \mathcal{S}(TM)$,

for all $X, X_1, X_2, Y, Y_1, Y_2 \in \mathcal{S}(TM)$ and for all $f_1, f_2 \in C^\infty(M)$. Property (3) says that ∇ is compatible with the Riemannian metric, while property (4) means that ∇ is torsion free. Moreover, the Riemannian connection is the only connection satisfying the above properties. The curvature tensor of the Riemannian metric is then given by

$$\mathcal{R}_{XY}Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z,$$

for $X, Y, Z \in \mathcal{S}(TM)$, and it satisfies the properties:

(1) $\mathcal{R}_{XY}Z = -\mathcal{R}_{YX}Z$;
(2) $\mathcal{R}_{XY}Z + \mathcal{R}_{YZ}X + \mathcal{R}_{ZX}Y = 0$; and
(3) $\langle \mathcal{R}_{XY}Z, W\rangle = \langle \mathcal{R}_{ZW}X, Y\rangle$,

for all $X, Y, Z, W \in \mathcal{S}(TM)$. The sectional curvature of the 2-plane section spanned by a pair of orthonormal vectors X and Y is defined by

$$K(X, Y) = \langle \mathcal{R}_{XY}Y, X\rangle.$$

If we take $\{e_1, \ldots, e_m\}$ to be an orthonormal basis of the tangent space of M, then the Ricci curvature is defined to be the symmetric 2-tensor given by

$$\mathcal{R}_{ij} = \sum_{k=1}^{m} \langle \mathcal{R}_{e_i, e_k} e_k, e_j\rangle.$$

Observe that the diagonal elements of the Ricci curvature are given by

$$\mathcal{R}_{ii} = \sum_{k \neq i} K(e_i, e_k).$$

Let N be an n-dimensional submanifold in M with $n < m$. The Riemannian metric ds_M^2 defined on M when restricted to N induces a Riemannian metric ds_N^2 on N. One can easily check that for vector fields $X, Y \in \mathcal{S}(TN)$, if we define

$$\nabla_X^t Y = (\nabla_X Y)^t$$

to be the tangential component of $\nabla_X Y$ to N, then ∇^t is the Riemannian connection of N with respect to ds_N^2. The normal component of ∇ yields the negative of the second fundamental form of N. In particular, one defines the second fundamental form by

$$\overrightarrow{II}(X, Y) = -(\nabla_X Y)^n,$$

and checks that it is tensorial with respect to $X, Y \in \mathcal{S}(TN)$. Taking the trace of the bilinear form $\overrightarrow{II}$ over the tangent space of N yields the mean curvature vector, given by

$$\mathrm{tr}(\overrightarrow{II}) = \overrightarrow{H}.$$

Let us now consider a one-parameter family of deformations of N given by $N_t = \phi(N, t)$ for $t \in (-\epsilon, \epsilon)$ with $N_0 = N$. Let $\{x_1, \ldots, x_n\}$ be a coordinate system around a point $p \in N$. We can consider $\{x_1, \ldots, x_n, t\}$ to be a coordinate system of $N \times (-\epsilon, \epsilon)$ near the point $(p, 0)$. Let us denote $e_i = d\phi\,(\partial/\partial x_i)$ for $i = 1, \ldots, n$ and $T = d\phi\,(\partial/\partial t)$. The induced metric on N_t from M is then given by $g_{ij} = \langle e_i, e_j\rangle$. We may further assume that $\{x_1, \ldots, x_n\}$ form a normal coordinate system at $p \in N$. Hence $g_{ij}(p, 0) = \delta_{ij}$ and $\nabla_{e_i} e_j(p, 0) = 0$. Let us define dA_t to be the area element of N_t with respect to the induced metric. For t sufficiently close to 0, we can write $dA_t = J(x, t)\, dA_0$. With respect to the normal coordinate system $\{x_1, \ldots, x_n\}$, the function $J(x, t)$ is given by

$$J(x, t) = \frac{\sqrt{g(x, t)}}{\sqrt{g(x, 0)}}$$

with $g(x, t) = \det(g_{ij}(x, t))$. To compute the first variation for the area of N, we compute $J'(p, t) = (\partial J/\partial t)(p, t)$. By the assumption that $g_{ij}(p, 0) = \delta_{ij}$, we have $J'(p, 0) = \frac{1}{2} g'(p, 0)$. However,

$$\begin{aligned} g &= \det(g_{ij}) \\ &= \sum_{j=1}^{n} g_{1j}\, c_{1j}, \end{aligned}$$

where c_{ij} are the cofactors of g_{ij}. Therefore

$$\begin{aligned} g'(p, 0) &= \sum_{j=1}^{n} g'_{1j}(p, 0)\, c_{1j}(p, 0) + \sum_{j=1}^{n} g_{1j}(p, 0)\, c'_{1j}(p, 0) \\ &= g'_{11}(p, 0) + c'_{11}(p, 0). \end{aligned}$$

By induction on the dimension, we conclude that $g'(p, 0) = \sum_{i=1}^{n} g'_{ii}$. On the other hand,

$$\begin{aligned} g'_{ii} &= T\langle e_i, e_i\rangle \\ &= 2\langle \nabla_T e_i, e_i\rangle \\ &= 2\langle \nabla_{e_i} T, e_i\rangle, \end{aligned}$$

because $\{x_1, \ldots, x_n, t\}$ form a coordinate system for $N \times (-\epsilon, \epsilon)$. Let us point out that the quantity

$$\sum_{i=1}^{n} \langle \nabla_{e_i} T, e_i\rangle$$

is now well defined under an orthonormal change of basis and hence is globally defined. If we write $T = T^t + T^n$, where T^t is the tangential component of T on N and T^n is its normal component, then

$$\begin{aligned}\sum_{i=1}^{n} \langle \nabla_{e_i} T, e_i \rangle &= \sum_{i=1}^{n} \langle \nabla_{e_i} T^t, e_i \rangle + \sum_{i=1}^{n} \langle \nabla_{e_i} T^n, e_i \rangle \\ &= \operatorname{div}(T^t) + \sum_{i=1}^{n} e_i \langle T^n, e_i \rangle - \sum_{i=1}^{n} \langle T^n, \nabla_{e_i} e_i \rangle \\ &= \operatorname{div}(T^t) + \langle T^n, \overrightarrow{H} \rangle,\end{aligned}$$

where $\overrightarrow{H}$ is the mean curvature vector of N. Hence the first variation for the volume form at the point $(p, 0)$ is given by

$$\frac{d}{dt} dA_t|_{(p,0)} = \left(\operatorname{div} T^t + \left\langle T^n, \overrightarrow{H} \right\rangle \right) dA_0 \Big|_{(p,0)} .$$

However, the right-hand side is intrinsically defined independent of the choice of coordinates and hence this formula is valid at any arbitrary point.

If T is a compactly supported variational vector field on N, then using the divergence theorem the first variation of the area of N is given by

$$\frac{d}{dt} A(N_t) \Big|_{t=0} = \int_N \left\langle T^n, \overrightarrow{H} \right\rangle .$$

This shows that the mean curvature of N is identically 0 if and only if N is a critical point of the area functional.

Definition 1.1 An immersed submanifold $N \hookrightarrow M$ is said to be minimal if its mean curvature vector vanishes identically, i.e., $\overrightarrow{H} \equiv 0$.

When N is a curve in M that is parametrized by arc-length with unit tangent vector e, then the first variational formula for length can be written as

$$\begin{aligned}\frac{d}{dt} L \Big|_{t=0} &= \langle T^t, e \rangle |_0^l - \int_0^l \langle T^n, \nabla_e e \rangle \\ &= \langle T, e \rangle |_0^l - \int_0^l \langle T, \nabla_e e \rangle .\end{aligned}$$

We will now proceed to derive the second variational formula for area. Let $\phi : N \times (-\epsilon, \epsilon) \times (-\epsilon, \epsilon) \longrightarrow M$ be a two-parameter family of variations of N. Using similar notation, we write $d\phi(\partial/\partial x_i) = e_i$ for $i = 1, \ldots, n$, and denote the variational vector fields by $d\phi(\partial/\partial t) = T$ and $d\phi(\partial/\partial s) = S$.

In terms of a general coordinate system, the first partial derivative of J can be written as

$$\frac{\partial J}{\partial t}(x,t,s) = \sum_{i,j=1}^{n} g^{ij}\langle \nabla_{e_i} T, e_j\rangle\, J(x,t,s),$$

where (g^{ij}) denotes the inverse matrix of (g_{ij}). Differentiating this with respect to s and evaluating at $(p,0,0)$ we have

$$\begin{aligned}
\frac{\partial^2 J}{\partial s \partial t} &= \sum_{i,j=1}^{n} S\left(g^{ij}\langle \nabla_{e_i} T, e_j\rangle J\right)\\
&= \sum_{i,j=1}^{n} (Sg^{ij})\langle \nabla_{e_i} T, e_j\rangle J + \sum_{i,j=1}^{n} g^{ij}\left(S\langle \nabla_{e_i} T, e_j\rangle\right) J\\
&\quad + \sum_{i,j=1}^{n} g^{ij}\langle \nabla_{e_i} T, e_j\rangle S(J)\\
&= \sum_{i,j=1}^{n} (Sg^{ij})\langle \nabla_{e_i} T, e_j\rangle + \sum_{i=1}^{n} S\langle \nabla_{e_i} T, e_i\rangle\\
&\quad + \left(\sum_{i=1}^{n}\langle \nabla_{e_i} T, e_i\rangle\right)\left(\sum_{j=1}^{n}\langle \nabla_{e_j} S, e_j\rangle\right). \qquad (1.1)
\end{aligned}$$

However, differentiating the formula $\sum_{k=1}^{n} g^{ik} g_{kj} = \delta_{ij}$, we obtain

$$\sum_{k=1}^{n}(Sg^{ik})g_{kj} = -\sum_{k=1}^{n} g^{ik}(Sg_{kj}),$$

hence

$$\begin{aligned}
Sg^{ij} &= -\sum_{k,l=1}^{n} g^{ik}(Sg_{kl})g^{lj}\\
&= -Sg_{ij}\\
&= -S\langle e_i, e_j\rangle\\
&= -\langle \nabla_S e_i, e_j\rangle - \langle \nabla_S e_j, e_i\rangle\\
&= -\langle \nabla_{e_i} S, e_j\rangle - \langle \nabla_{e_j} S, e_i\rangle.
\end{aligned}$$

The first term on the right-hand side of (1.1) now becomes

$$\sum_{i,j=1}^{n}(Sg^{ij})\langle\nabla_{e_i}T, e_j\rangle = -\sum_{i,j=1}^{n}\langle\nabla_{e_i}S, e_j\rangle\langle\nabla_{e_i}T, e_j\rangle - \sum_{i,j=1}^{n}\langle\nabla_{e_j}S, e_i\rangle\langle\nabla_{e_i}T, e_j\rangle.$$

The second term on the right-hand side of (1.1) can be written as

$$\begin{aligned}\sum_{i=1}^{n}S\langle\nabla_{e_i}T, e_i\rangle &= \sum_{i=1}^{n}\langle\nabla_S\nabla_{e_i}T, e_i\rangle + \sum_{i=1}^{n}\langle\nabla_{e_i}T, \nabla_S e_i\rangle \\ &= \sum_{i=1}^{n}\langle\mathcal{R}_{Se_i}T, e_i\rangle + \sum_{i=1}^{n}\langle\nabla_{e_i}\nabla_S T, e_i\rangle + \sum_{i=1}^{n}\langle\nabla_{e_i}T, \nabla_{e_i}S\rangle,\end{aligned}$$

where the term $\langle\mathcal{R}_{Se_i}T, e_i\rangle$ on the right-hand side denotes the curvature tensor of M. Therefore, we have

$$\begin{aligned}\frac{\partial^2 J}{\partial s\partial t} = &-\sum_{i,j=1}^{n}\langle\nabla_{e_i}S, e_j\rangle\langle\nabla_{e_i}T, e_j\rangle - \sum_{i,j=1}^{n}\langle\nabla_{e_j}S, e_i\rangle\langle\nabla_{e_i}T, e_j\rangle \\ &+ \sum_{i=1}^{n}\langle\mathcal{R}_{Se_i}T, e_i\rangle + \sum_{i=1}^{n}\langle\nabla_{e_i}\nabla_S T, e_i\rangle + \sum_{i=1}^{n}\langle\nabla_{e_i}T, \nabla_{e_i}S\rangle \\ &+ \left(\sum_{i=1}^{n}\langle\nabla_{e_i}T, e_i\rangle\right)\left(\sum_{j=1}^{n}\langle\nabla_{e_j}S, e_j\rangle\right).\end{aligned} \tag{1.2}$$

We will now consider some special cases that will simplify (1.2). Let us first assume that N is a curve parametrized by arc-length in M with unit tangent vector given by e, then the second variational formula for the length is given by

$$\begin{aligned}\left.\frac{\partial^2 L}{\partial s\partial t}\right|_{(s,t)=(0,0)} &= \int_0^l\{-\langle\nabla_e S, e\rangle\langle\nabla_e T, e\rangle + \langle\mathcal{R}_{Se}T, e\rangle\} \\ &\quad + \int_0^l\{\langle\nabla_e\nabla_S T, e\rangle + \langle\nabla_e T, \nabla_e S\rangle\}.\end{aligned}$$

If we further assumed that N is a geodesic satisfying the geodesic equation $\nabla_e e \equiv 0$, then we have

$$\begin{aligned}\left.\frac{\partial^2 L}{\partial s \partial t}\right|_{(s,t)=(0,0)} &= \int_0^l \{-(e\langle S, e\rangle)(e\langle T, e\rangle) + \langle \mathcal{R}_{Se} T, e\rangle\} \\ &\quad + \int_0^l \{e\langle \nabla_S T, e\rangle + \langle \nabla_e T, \nabla_e S\rangle\} \\ &= \int_0^l \left\{\langle \nabla_e T, \nabla_e S\rangle + \langle \mathcal{R}_{Se} T, e\rangle - (e\langle S, e\rangle)(e\langle T, e\rangle)\right\} \\ &\quad + \langle \nabla_S T, e\rangle|_0^l .\end{aligned}$$

The second special case is when N is a general n-dimensional manifold and then if the two variational vector fields are the same and are normal to N, (1.2) becomes

$$\begin{aligned}\left.\frac{\partial^2 J}{\partial t^2}\right|_{t=0} &= -\sum_{i,j=1}^n \langle \nabla_{e_i} T, e_j\rangle^2 - \sum_{i,j=1}^n \langle \nabla_{e_j} T, e_i\rangle\langle \nabla_{e_i} T, e_j\rangle + \sum_{i=1}^n \langle \mathcal{R}_{Te_i} T, e_i\rangle \\ &\quad + \sum_{i=1}^n \langle \nabla_{e_i} \nabla_T T, e_i\rangle + \sum_{i=1}^n |\nabla_{e_i} T|^2 + \left(\sum_{i=1}^n \langle \nabla_{e_i} T, e_i\rangle\right)^2 \\ &= -\sum_{i,j=1}^n \langle \nabla_{e_i} T, e_j\rangle^2 - \sum_{i,j=1}^n \langle \nabla_{e_j} T, e_i\rangle\langle \nabla_{e_i} T, e_j\rangle - \sum_{i=1}^n \langle \mathcal{R}_{e_i T} T, e_i\rangle \\ &\quad + \operatorname{div}(\nabla_T T)^t + \left\langle (\nabla_T T)^n, \vec{H}\right\rangle + \sum_{i=1}^n |\nabla_{e_i} T|^2 + \left\langle T, \vec{H}\right\rangle^2 . \qquad (1.3)\end{aligned}$$

On the other hand, if $\{e_{n+1}, \ldots, e_m\}$ denotes an orthonormal set of vectors normal to N in M, then

$$\sum_{i=1}^n \langle \nabla_{e_i} T, \nabla_{e_i} T\rangle = \sum_{i,j=1}^n \langle \nabla_{e_i} T, e_j\rangle^2 + \sum_{i=1}^n \sum_{\nu=n+1}^m \langle \nabla_{e_i} T, e_\nu\rangle^2 .$$

Also

$$\begin{aligned}\langle \nabla_{e_i} T, e_j\rangle &= \left\langle T, \vec{II}_{ij}\right\rangle \\ &= \langle \nabla_{e_j} T, e_i\rangle,\end{aligned}$$

where $\overrightarrow{II}_{ij}$ denotes the second fundamental form with value in the normal bundle of N. Hence, (1.3) becomes

$$\frac{\partial^2 J}{\partial t^2}\bigg|_{t=0} = -\sum_{i,j}\left\langle T, \overrightarrow{II}_{ij}\right\rangle^2 - \sum_{i=1}^{n}\langle \mathcal{R}_{e_i T}T, e_i\rangle + \operatorname{div}(\nabla_T T)^t + \left\langle (\nabla_T T)^n, \overrightarrow{H}\right\rangle + \sum_{i=1}^{n}\sum_{\nu=n+1}^{m}\langle \nabla_{e_i}T, e_\nu\rangle^2 + \left\langle T, \overrightarrow{H}\right\rangle^2.$$

Therefore, the second variational formula for area in terms of compactly supported normal variations is given by

$$\frac{d^2}{dt^2}A(N_t)\bigg|_{t=0} = \int_N \left\{ -\sum_{i,j}\left\langle T, \overrightarrow{II}_{ij}\right\rangle^2 - \sum_{i=1}^{n}\langle \mathcal{R}_{e_i T}T, e_i\rangle + \left\langle (\nabla_T T)^n, \overrightarrow{H}\right\rangle \right\} + \int_N \left\{ \sum_{i=1}^{n}\sum_{\nu=n+1}^{m}\langle \nabla_{e_i}T, e_\nu\rangle^2 + \left\langle T, \overrightarrow{H}\right\rangle^2 \right\}.$$

Definition 1.2 A minimally immersed submanifold $N \hookrightarrow M$ is said to be stable if the second variation for area with respect to all compactly supported normal variations is nonnegative. This means that the stability inequality

$$0 \leq -\int_N \sum_{i,j}\left\langle T, \overrightarrow{II}_{ij}\right\rangle^2 - \int_N \sum_{i=1}^{n}\langle \mathcal{R}_{e_i T}T, e_i\rangle + \int_N \sum_{i=1}^{n}\sum_{\nu=n+1}^{m}\langle \nabla_{e_i}T, e_\nu\rangle^2$$

is valid for any compactly supported normal vector field T.

If we further restrict N to be an orientable codimension-1 minimal submanifold of an orientable manifold M, we can write any normal variation in the form $T = \psi e_m$, where ψ is a differentiable function on N and e_m is a unit normal vector field to N. Then the second variational formula can be written as

$$\begin{aligned}\frac{d^2}{dt^2}A(N_t)\bigg|_{t=0} &= \int_N \left\{ -\sum_{i,j}\left\langle T, \overrightarrow{II}_{ij}\right\rangle^2 - \mathcal{R}(T,T) + \sum_{i=1}^{n}\langle \nabla_{e_i}T, e_m\rangle^2 \right\} \\ &= \int_N \left\{ -\psi^2 h_{ij}^2 - \psi^2\mathcal{R}(e_m, e_m) + |\nabla\psi|^2 \right\},\end{aligned}$$

where $\overrightarrow{II}_{ij} = h_{ij}\, e_m$ with h_{ij} being the component of the second fundamental form and $\mathcal{R}(T,T)$ denotes the Ricci curvature of M in the direction of T. Here we have also used the fact that

$$\begin{aligned}\langle \nabla_{e_i} T, e_m \rangle &= \psi \langle \nabla_{e_i} e_m, e_m \rangle + e_i(\psi) \langle e_m, e_m \rangle \\ &= e_i(\psi).\end{aligned}$$

In particular, the stability inequality in this case is given by

$$\int_N |\nabla \psi|^2 \geq \int_N \psi^2 h_{ij}^2 + \int_N \psi^2 \mathcal{R}(e_m, e_m). \tag{1.4}$$

The last special case is again to assume that N is an oriented hypersurface in an oriented manifold M and we restrict the variation to be given by hypersurfaces which are a constant distant from N. The variational vector field is then given by e_m with $\nabla_{e_m} e_m \equiv 0$. This situation is particularly useful for the purpose of controlling the growth of the volume of geodesic balls of radius r. In this case, if we write $\overrightarrow{H} = H\, e_m$, the first variational formula for the area element becomes

$$\frac{\partial J}{\partial t}(x, 0) = H(x)\, J(x, 0), \tag{1.5}$$

and the second variational formula can be written as

$$\begin{aligned}\frac{\partial^2 J}{\partial t^2}(x, 0) = &- \sum_{i,j=1}^{m-1} h_{ij}^2(x)\, J(x, 0) \\ &- \mathcal{R}(e_m, e_m)(x)\, J(x, 0) + H^2(x)\, J(x, 0).\end{aligned} \tag{1.6}$$

2

Volume comparison theorem

In this chapter, we will develop a volume comparison theorem originally proved by Bishop (see [BC]). Let $p \in M$ be a point in a complete Riemannian manifold of dimension m. In terms of polar normal coordinates at p, we can write the volume element as

$$J(\theta, r)dr \wedge d\theta,$$

where $d\theta$ is the area element of the unit $(m-1)$-sphere. The Gauss lemma asserts that the area element of submanifold $\partial B_p(r)$, which is the boundary of the geodesic ball of radius r, is given by $J(\theta, r)d\theta$. By the first and second variational formulas (1.5) and (1.6), if $x = (\theta, r)$ is not in the cut-locus of p, we have

$$\begin{aligned} J'(\theta, r) &= \frac{\partial J}{\partial r}(\theta, r) \\ &= H(\theta, r)\, J(\theta, r) \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} J''(\theta, r) &= \frac{\partial^2 J}{\partial r^2}(\theta, r) \\ &= -\sum_{i,j=1}^{m-1} h_{ij}^2(\theta, r)\, J(\theta, r) - \mathcal{R}_{rr}(\theta, r)\, J(\theta, r) + H^2(\theta, r)\, J(\theta, r), \end{aligned} \tag{2.2}$$

where $\mathcal{R}_{rr} = \mathcal{R}(\partial/\partial r, \partial/\partial r)$, $H(\theta, r)$, and $(h_{ij}(\theta, r))$ denote the Ricci curvature in the radial direction, the mean curvature and the second fundamental form of $\partial B_p(r)$ at the point $x = (\theta, r)$ with respect to the unit normal vector $\partial/\partial r$, respectively.