



In classical differential geometry for embedded two-dimensional surfaces  $M$  in  $\mathbb{R}^3$ , the Riemannian metric for  $M$  induced from the dot product of  $\mathbb{R}^3$  is called the *first fundamental form*.

There are various notations for the first fundamental form; a common notation is  $\mathcal{I}$ , for the roman letter one. Thus,

$$\mathcal{I}(v, w) = v \cdot w \quad (1)$$

for vectors  $v, w \in \mathbb{R}^3$ . (We consider the tangent planes of  $M$  to be two-dimensional subspaces of  $\mathbb{R}^3$ .)

## Quadratic form representation

Recall, in linear algebra, that a symmetric bilinear form  $T$  over  $\mathbb{R}$  can always be represented by its quadratic form  $Q$ :

$$T(u, v) = \frac{1}{2}(Q(u + v) - Q(u) - Q(v)), \quad Q(w) = T(w, w),$$

for any vectors  $u$  and  $v$ . This process may be applied to the first fundamental form, and classically, the first fundamental form is expressed as

$$ds^2 = E du^2 + 2F dudv + G dv^2. \quad (2)$$

In modern terminology, (??) is the quadratic form that represents the bilinear form  $\mathcal{I}$ . The use of the letters  $E, F, G$  for the coefficients of the quadratic form is traditional, and dates back to Gauss; in terms of the metric tensor  $g_{ij}$ , these coefficients are defined by  $E = g_{11}$ ,  $F = g_{12} = g_{21}$ ,  $G = g_{22}$ .

The letters  $u$  and  $v$  in (??) denote local coordinates on  $M$ . Classically,  $du$  and  $dv$  meant “infinitesimally small” changes in  $u$  and  $v$ , but in modern differential geometry,  $du$  and  $dv$  have been given a precise meaning using differential forms.

In tensor notation, (??) is written as

$$\mathcal{I} = E du \otimes du + F du \otimes dv + F dv \otimes du + G dv \otimes dv, \quad (3)$$

Although the tensor notation is more clumsy, it allows us to rigorously justify a change of variables, by the rule  $\alpha^*(du \otimes dv) = d(\alpha^*u) \otimes d(\alpha^*v)$ . See the example below.

The symbol  $ds$  in (??) alludes to

$$ds^2 = dx^2 + dy^2 + dz^2,$$

the infinitesimal length of a curve. Compare with the modern notation

$$\mathcal{I} = dx \otimes dx + dy \otimes dy + dz \otimes dz. \quad (4)$$

(This is just an alternate way of writing the definition of  $\mathcal{I}$ : the restriction of the dot product on  $\mathbb{R}^3$ .)

## Example: sphere

We illustrate an example: we compute the first fundamental form  $\mathcal{I}$  of the sphere  $S^2$  in spherical coordinates (latitude/longitude system). We set

$$x = \cos \phi \cos \theta$$

$$y = \cos \phi \sin \theta$$

$$z = \sin \phi,$$

and substitute these in (??):

$$\begin{aligned} \mathcal{I} &= d(\cos \phi \cos \theta) \otimes d(\cos \phi \cos \theta) + d(\cos \phi \sin \theta) \otimes d(\cos \phi \sin \theta) \\ &\quad + d(\sin \phi) \otimes d(\sin \phi) \\ &= (-\sin \phi \cos \theta d\phi - \cos \phi \sin \theta d\theta) \otimes (-\sin \phi \cos \theta d\phi - \cos \phi \sin \theta d\theta) \\ &\quad + (-\sin \phi \sin \theta d\phi + \cos \phi \cos \theta d\theta) \otimes (-\sin \phi \sin \theta d\phi + \cos \phi \cos \theta d\theta) \\ &\quad + (\cos \phi d\phi) \otimes (\cos \phi d\phi) \\ &= (\sin \phi \cos \theta)^2 d\phi \otimes d\phi + (\cos \phi \sin \theta)^2 d\theta \otimes d\theta \\ &\quad + (\sin \phi \sin \theta)^2 d\phi \otimes d\phi + (\cos \phi \cos \theta)^2 d\theta \otimes d\theta \\ &\quad + (\cos \phi)^2 d\phi \otimes d\phi \end{aligned}$$

(note that the cross terms with  $d\phi \otimes d\theta$  and  $d\theta \otimes d\phi$  cancel)

$$\begin{aligned} &= \sin^2 \phi d\phi \otimes d\phi + \cos^2 \phi d\theta \otimes d\theta + \cos^2 \phi d\phi \otimes d\phi \\ &= d\phi \otimes d\phi + \cos^2 \phi d\theta \otimes d\theta. \end{aligned}$$

Of course this was a very cumbersome calculation; the writing would be simplified if we had just dropped the  $\otimes$  signs and wrote  $d\phi^2$  for  $d\phi \otimes d\phi$ , etc.

And even then the calculation would be more organized if we computed the coefficients  $g_{ij}$  directly. We only show this kind calculation in order to justify what exactly is meant by the classical expression

$$d\phi^2 + \cos^2 \phi d\theta^2$$

for the first fundamental form of the sphere.

## Use of first fundamental form to compute lengths and areas

The first fundamental form is related to the area form. If

$$ds^2 = E du^2 + 2F dudv + G dv^2$$

then

$$dA = \sqrt{EG - F^2} du \wedge dv$$

is the area form. For the sphere, this is

$$dA = \sqrt{\cos^2 \phi - 0} d\phi \wedge d\theta = \cos \phi d\phi \wedge d\theta ,$$

which is just the formula given in calculus for evaluating surface integrals on the sphere using spherical coordinates.

The first fundamental form itself may be used to find the length  $s$  of a curve  $\gamma$  on a surface  $M$ , when  $\gamma$  is parameterized by local coordinates:

$$\begin{aligned} s &= \int_{\gamma} ds = \int_{\gamma} \sqrt{ds^2} \\ &= \int_{\gamma} \sqrt{E du^2 + 2F dudv + G dv^2} , \quad \gamma(t) = (u, v) , \\ &= \int_a^b \sqrt{E \left( \frac{du}{dt} \right)^2 + 2F \frac{du}{dt} \frac{dv}{dt} + G \left( \frac{dv}{dt} \right)^2} dt . \end{aligned}$$

Although in practice it is probably easier to directly use cartesian coordinates, rather than the above expressions, to compute the length of  $\gamma$ , the first fundamental form plays an essential role in the theoretical investigation of the lengths of curves on a surface.

## Example: plane and cylinder

For the plane  $\mathbb{R}^2 \subset \mathbb{R}^3$  with  $z = 0$ , the first fundamental form is just

$$ds^2 = dx^2 + dy^2. \quad (5)$$

For the cylinder with the coordinates

$$x = \cos u$$

$$y = \sin u$$

$$z = v$$

the first fundamental form is

$$ds^2 = du^2 + dv^2. \quad (6)$$

## Relation with isometric maps

Notice that (5) looks the same as (6) after renaming the variables. This is evidence that the plane and cylinder should be locally isometric: a flat sheet can be rolled into a cylinder. An isometry between two surfaces, by definition, preserves the metric on the two surfaces, so an isometry preserves the first fundamental form.

Of course, (5) and (6) are expressions of the first fundamental form in local coordinates of two different surfaces, so it makes no sense to say they are equal. But it is not hard to see that:

Suppose  $\Phi: M \rightarrow N$  is an isometry of two surfaces, and  $u, v$  are coordinates on  $M$ . If we use the coordinates  $u' = u \circ \Phi^{-1}$  and  $v' = v \circ \Phi^{-1}$  on  $N$ , then the first fundamental form of  $N$  is obtained by taking the first fundamental form of  $M$  and renaming  $u, v$  to  $u', v'$ .

## Relation with conformal, equiareal maps

There is also a notion of a conformal mapping: a diffeomorphism  $\Phi: M \rightarrow N$  is called *conformal* if  $\Phi$  preserves the first fundamental form up to a non-zero constant of proportionality. (The proportion may vary at each point of  $M$  and  $N$ .) It may be verified that this is the same as saying that  $\Phi$  preserves the angles of intersecting tangent vectors.

For example, the stereographic projection from the sphere to the plane is conformal.

Yet another notion is that of an equiareal mapping: a diffeomorphism  $\Phi: M \rightarrow N$  is called *equiareal* if  $\Phi$  preserves areas of all subregions of the surfaces. This amounts to saying that the quantity  $\sqrt{EG - F^2}$  is invariant under  $\Phi$  (provided we rename the variables as explained above).

For example, the projection of the sphere to the cylinder wrapping it is equiareal. This fact was used by Archimedes to show the sphere of radius  $r$  has area  $4\pi r^2$ , because the cylinder that wraps it also has area  $4\pi r^2$ .

If  $\Phi$  is both conformal and equiareal, then it is an isometry. As a well-known example, a sphere is not isometric to the plane, not even locally, so we cannot draw maps of the Earth that preserve both directions and relative proportion of lands. We must give up at least one of these properties: e.g. the Mercator projection preserves direction only; maps with Mercator look “strange” the first time one sees them, because such maps do not preserve area.

## Relation with Gaussian curvature

There is a formula for the Gaussian curvature  $K(p)$  at a point on a surface:

$$K = \frac{\begin{vmatrix} -\frac{1}{2}E_{vv} + F_{uv} - \frac{1}{2}G_{uu} & \frac{1}{2}E_u & F_u - \frac{1}{2}E_v \\ F_v - \frac{1}{2}G_u & E & F \\ \frac{1}{2}G_v & F & G \end{vmatrix} - \begin{vmatrix} 0 & \frac{1}{2}E_v & \frac{1}{2}G_u \\ \frac{1}{2}E_v & E & F \\ \frac{1}{2}G_u & F & G \end{vmatrix}}{(EG - F^2)^2}$$

where the bars denote the determinant, and the subscripts denote partial derivatives.

This formula is known as Brioschi’s formula; Brioschi had stated it without proof in 1854, and later it was calculated by Gauss.

The immediate corollary of this strange formula is:

**Theorem 1** (Theorema Egregium). *The Gaussian curvature of a surface is unchanged under isometries (because it only depends on the first fundamental form).*

This theorem is not obvious, since the usual definitions of the Gaussian curvature are not invariant (they depend on the particular embedding of the surface in  $\mathbb{R}^3$ ).

## References

- [1] Michael Spivak. *A Comprehensive Introduction to Differential Geometry*, volumes I and II. Publish or Perish, 1979.
- [2] Andrew Pressley. *Elementary Differential Geometry*. Springer-Verlag, 2003.