Theory of Surfaces—A Rapid Introduction

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

Abstract

In this article, we introduce the reader to the basic facts of the theory of surfaces in a geometric way. This is based on a series o ftwo lectures given in Summer School in Advanced Real Analysis and its Applications to PDE, IISc-TIFR Mathematics Programme, Bangalore, May 18–June 8, 1992.

1 Theory of Surfaces–Preliminary Notions

Our aim here is to introduce the reader to some basic concepts and results in the theory of surfaces in $\mathbb{R}^3$. We shall endeavour to instill geometric intuition and a feeling for the subject. In particular, we shall not stop to prove various equivalent definitions and concepts introduced.

Let $\mathbb{R}^n$ be the n-dimensional vector space over $\mathbb{R}$. We have a *natural basis* $\{e_i : 1 \le i \le n\}$ of $\mathbb{R}^n$ over $\mathbb{R}$, where $e_i := (0, \dots, 0, 1, 0, \dots, 0)$, 1 at the i-th place. We identify any $x \in \mathbb{R}^n$ with the n-tuple $(x_1, \dots, x_n)$ where $x := \sum x_i e_i$. We then have the Euclidean inner product $\langle x, y\rangle := \sum_{i=1}^n x_i y_i$, where $x = \sum x_i e_i$ and $y = \sum y_i e_i$. We let E^n stand for $\mathbb{R}^n$ with the Euclidean inner product. E^n is called the n-dimensional Euclidean space. We have the norm on E^n induced by the inner product $\|x\| := \sqrt{\langle x, x\rangle}$ for $x \in E^n$. We often denote E^n by E. We shall denote by $\nabla f(p)$ the gradient of f at p.

Intuitively, our notion of a surface in $\mathbb{R}^{n+1}$ is a nonempty subset $S \subset \mathbb{R}^{n+1}$ such that each $p \in S$ has a relative open neighbourhood in S which is homeomorphic (diffeomorphic?) to an open set in $\mathbb{R}^n$ and further (this is the most important) there exists a(n) (affine) tangent hyperplane that best approximates S at p.

We shall start with the definition of a level surface. Let $f : U \subset \mathbb{R}^{n+1} \to \mathbb{R}$ be a smooth function on an open set U. Let c be in the image of f. Let $S := f^{-1}(c)$. To get a tangent plane we impose a further condition on f: For each $p \in S$, $\nabla f(p) \neq 0$. Such an S is called a *level surface* of f at c. We look at some examples.

Example 1. The simplest function one can think of is an affine linear function of the form $f(x) := \langle x, N\rangle - c$ for some $c \in \mathbb{R}$ where $N := (a_1, \dots, a_n) \in \mathbb{R}^{n+1}$. Since ∇f is the constant vector N, we must assume that $N \neq 0$. In this case $S := f^{-1}(0)$ is a hyperplane.

Example 2. Let $g(x) := \langle x, x \rangle - R^2$, $x \in \mathbb{R}^{n+1}$, $R > 0$. Then $S = g^{-1}(0)$ is the sphere of radius R centered at the origin. Note that the gradient $\nabla g(x) = (x_1, \dots, x_{n+1}) \neq 0$, as $\|x\| = R$.

Example 3. A right circular cylinder of base radius R is a level surface in $\mathbb{R}^3$ if we consider the function $h(x) := x_1^2 + x_2^2 - R^2$.

Example 4. The saddle surface $S := \{(x, y, z) \in \mathbb{R}^3 : z = xy\}$ is a level surface. (Proof?)

Before we define the tangent plane to a surface let us recall the concept of tangent vectors to a curve in E.

A curve in U, an open subset of E, is a "smooth" map $c\colon (a, b) \to U$. By smooth we mean that it has sufficiently many continuous derivatives on $(-\varepsilon, \varepsilon)$. This is same as saying that the component functions c_i have sufficiently many continuous derivatives. Here, of course, $c(t) := (c_1(t), \dots, c_n(t))$. We say that it passes through the point $p := c(0)$. The *tangent* to this curve at $c(t)$ is by definition the vector

$$c'(t) := \lim_{h \to 0} \frac{c(t+h) - c(t)}{h} = (c_1'(t), \dots, c_n'(t)).$$

$c'(t)$ is also called the *velocity* vector of c at $c(t)$. The length $l(c)$ of the curve c is defined by setting $l(c) := \int_{-\varepsilon}^{\varepsilon} \|c'(t)\| \, dt$.

Ex. 5. Let $h : (a, b) \to (-\varepsilon, \varepsilon)$ be a smooth map such that $h'(s) \neq 0$ for any $a < s < b$. Then $c \circ h : (a, b) \to U$ is a curve tracing the same image as c. h is called a reparameterization of c. Show that $l(c \circ h) = l(c)$.

We can put this in a picturesque language: The train travels the same distance whether it is slow or fast!

A continuous map $c\colon (a, b) \to S \subset \mathbb{R}^{n+1}$ is said to be a smooth curve in S if $c\colon (a, b) \to \mathbb{R}^{n+1}$ is smooth. If $c\colon (-\varepsilon, \varepsilon) \to S$ and if $p := c(0)$ then we say that c passes through p. The tangent vector $v := c'(0)$ is said to be a tangent vector to S at p. We denote by $T_p(S)$ the set of all tangent vectors to S at p. That is,

$$T_p S := \{v \in \mathbb{R}^{n+1} : \exists \text{ a curve in } S \text{ through } p \text{ with } c'(0) = v\}.$$

This geometric definition does not reveal the fact that $T_p S$ is an n-dimensional vector subspaces of $\mathbb{R}^{n+1}$. So, we give an analytic definition of $T_p S$.

Let $S := f^{-1}(0)$. Let $v \in T_p S$. Let c be any curve in S passing through p with $c'(0) = v$. The map $f \circ c\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$ is then constant so that we have

$$0 = \frac{d}{dt}(f \circ c(t)) \,|_{t=0} = f'(c(0))(c'(0)) = \langle \nabla f(p), v \rangle \,.$$

Thus the $\nabla f(p)$ is orthogonal to any vector tangent to S at p. Hence $\nabla f(p)$ can be considered as normal to the surface at p. This suggests us the following definition:

$$T_p S := \{v \in \mathbb{R}^{n+1} : \langle \nabla f(p), v \rangle = 0\}.$$

Clearly, T_pS is an n-dimensional linear subspace of $\mathbb{R}^{n+1}$. The above two definitions can be seen to be equivalent. (See Remark 9 below for an idea towards a proof.)

We invite the reader to convince himself this definition coincides with the notion of the affine tangent plane if we translate tangent plane by p. That is, $p + T_pS$ is the affine tangent plane. See the pictures below (?!) in the case of a sphere and a cylinder.

An important property of the level surfaces is the existence of a smooth nowhere vanishing *normal field* on S, viz., the map

$$p \mapsto N_p := N(p) := \frac{\nabla f(p)}{\| \nabla f \| (p)}.$$

We now define a surface to be a nonempty subset in E that looks like a level surface around each point. More precisely,

Definition 6. A nonempty subset $S \in \mathbb{R}^{n+1}$ is said to be a surface in $\mathbb{R}^{n+1}$ if for each $p \in S$, there exists an open set $V_p \ni p$ and a smooth function $g_p : V_p \to \mathbb{R}$ such that i) $S \cap V_p = g_p^{-1}(0)$ and ii) $\nabla g_p(x) \neq 0$ for all$x \in S \cap V_p$.

Remark 7. (May be omitted on first reading.) Def. 6 is the most convenient one for us to work with. The more conventional definition is as follows:

Definition 8. A surface in E^{n+1} is a set $S \subset E^{n+1}$ such that every point $p \in S$ has an open neighbourhood V in E^{n+1} with the following property:
There exists an open set U in E^n and a smooth map $\varphi : U \to V$ such that i) $\varphi : U \to V \cap S$ is a homeomorphism and ii) φ is regular, i.e., the Frechet derivative, or what is the same, the Jacobian

$$d\varphi(u_1, \ldots, u_n) := \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial x_{n+1}}{\partial u_1} & \cdots & \frac{\partial x_{n+1}}{\partial u_n} \end{pmatrix} |_{(u_1,\ldots,u_n)} : E^n \to E^{n+1}$$

is of rank n at every point in U. Here we have written

$$\varphi(u_1, \ldots, u_n) = (x_1(u_1, \ldots, u_n), \ldots, x_{n+1}(u_1, \ldots, u_n)).$$

The pair (φ, U) is called a *parameterization* of $V \cap S$.

The tangent space in this case may be taken to be the image of the derivative of the parameterizing map: $T_pS := d\varphi(u)(\mathbb{R}^n)$ where $\varphi(u) = p$.

That Def. 6 and Def. 8 are equivalent follows from an instructive application of the inverse mapping theorem. (End of Remark.)

We shall however work with Def. 6. If S is a surface then T_pS is defined as $T_pS := \{v \in \mathbb{R}^{n+1} : \langle v, \nabla g_p(p) \rangle = 0\}$.

Remark 9. (May be omitted on first reading.) The equality of the all the tangent spaces defined so far can be seen if one realizes the following facts:
i) T_pU is the n-dimensional vector space $\mathbb{R}^n$.
ii) φ carries smooth curves in U into smooth curves in $V \cap S$.

Note that given any point $p \in S$, there exits a unit normal field in $S \cap V_p$, viz.,

$$x \mapsto \nabla g_p(x)/\, \| \nabla g_p(x) \| \,.$$

Our next goal is to attach a numerical quantity to any point of S which will tell us how curved S is at that point. Even though most of what we say below continues to be true in higher dimensions, we shall restrict ourselves to $n = 3$.

In any kind of measurement we need a standard object against which we compare other objects. Intuitively, we should like to think a plane in $\mathbb{R}^3$ is as "straight" and not curved at all. Hence one way of measuring the curvature of S at $p \in S$ is to see how much it deviates from being a plane. As the tangent plane T_pS is thought of as the plane best approximating the surface at p, our first tentative definition of a curvature of the surface is the rate of change of T_pS as p varies over a path. (Differentiation or rate of change of quantities are best done via curves!) But there are many directions or curves through any given point of the surface and hence the question arises which are to be considered. As there are only two linearly independent directions at each point p we may start with finding the rate of change of the tangent spaces along two curves c_i passing through p where $c_i'(0)$ form a basis of T_pS.

Since the tangent spaces are 2-dimensional objects, we wonder whether there is any 1-dimensional object whose rate of change will allows us to infer that of the tangent spaces. As you may have guessed, there is an obvious choice, viz., the map $p \mapsto N_p$ in the neighbourhood $S \cap V_p$. Hence our definition of curvature reads as follows:

The curvature at a point is the rate of change of a unit normal N along two linearly independent directions at that point.

More precisely, if $v \in T_pS$ and c is any curve through p with $c'(0) = v$ we then compute $D_vN := \frac{d}{dt}(N \circ c(t))\,|_{t=0}$. An easy application of the chain rule shows that D_vN is independent of the choice of the curve as long as $c'(0) = v$. Where does this vector D_vN belong to? Since $\langle N(c(t)), N(c(t)) \rangle = 1$, on differentiation we get $2\,\langle D_vN, N_p \rangle = 0$. That is, $D_vN \in T_pS$.

Thus, we get a map $L_p : T_pS \to T_pS$ given by $L_pv := D_vN$. It is easy to see that L_p is linear. It can be shown that L_p is symmetric with respect to the inner product T_pS inherits from $\mathbb{R}^3$. Thus curvature of S is to be got as a numerical quantity from this symmetric linear map L_p. Once we have a symmetric linear map of a finite dimensional inner product space we think of the natural numerical quantities associated with it, viz., the eigenvalues, which are real numbers. The eigen values $\lambda_i(p)$ of L_p are called the *principal curvatures* of S at p. The symmetric functions $H_p := (1/2)[\lambda_1(p) + \lambda_2(p)]$ and $K_p := \lambda_1(p)\lambda_2(p)$ are respectively called the *mean curvature* and the *Gaussian curvature* of S at p. Of these K_p remains invariant if we choose $\tilde{N} := -N$ as the unit normal to define $\tilde{L} = -L$.

Remark 10. (May be omitted on first reading) The above symmetric functions have the following property: They generate algebraically all the polynomial functions $f : M(2, \mathbb{R}) \to \mathbb{R}$ which are "invariant under conjugation": $f(AXA^{-1}) = f(X)$ for all $X \in M(2, \mathbb{R})$ and invertible A.

$\lambda_i(p)$'s have geometric meaning, if we recall how we got them out of L_p: They are the maximum and minimum of the function $v \mapsto \langle L_pv, v \rangle$ on the compact space $\{v \in T_pS : \|v\| =$

$1\}$. Thus, if v_i is an eigenvector of unit norm corresponding to the eigenvalue λ_i, then v_i is called a "principal direction." Thus v_i is a direction of "an extremum" for L_p.

Let us look at some of our earlier examples.

1) If S is the plane given by $\langle X, N\rangle + c = 0$ then $T_pS = \{v \in \mathbb{R}^3 : \langle v, N\rangle = 0\}$, i.e., the plane itself is translated paralleley to pass through the origin. Hence the unit normal field $p \mapsto N_p$ can be taken as $N_p := (N/\|N\|)$ for all $p \in S$. Hence $D_vN = 0$ for $v \in T_pS$. Hence the plane has the curvatures 0 as it should be.

2) Let S be a cylinder of base radius R. Let $p := (x_0, y_0, z_0) \in S$. There is an obvious choice of 2 linearly independent directions through p, viz., those corresponding to the curves c_1, the straight line through p parallel to the z-axis and the circle which is the intersection of the cylinder and the plane $z = z_0$. The unit normal can be taken as $N_p := (x_0, y_0)/R$. Now on c_1, $N_q = N_p$ for all $q \in c_1$ and hence $D_{c'(0)}N = 0$. (For, $q \in c_1$ will have the same x, y coordinates but different "height" z. Hence $N_q = (x_0, y_0, 0)/R$. Also, $c_1(t) := (x_0, y_0, z_0 + t) = p + te_3$ so that $c_1'(0) = e_3$. Hence $e_3 \in T_pS$ and $D_{e_3}N = 0$.) Thus, e_3 is a principal direction with principal curvature 0.

Since $x_0^2 + y_0^2 = R^2$, we can find $\theta_0 \in (0, 2\pi]$ such that $x_0 = R\cos\theta_0$, $y_0 = R\sin\theta_0$. Hence c_2 is given by

$$c_2 := (R\cos(\theta_0 + t), R\sin(\theta_0 + t), z_0)$$

so that $c_2'(0) = (-R\sin\theta_0), R\cos(\theta_0), 0)$. That is, $c_2'(0) = (-y_0, x_0, 0)$, the usual tangent to the circle. The unit normal field along c_2 is given by

$$N(c_2(t)) := (\cos(\theta_0 + t), \sin(\theta_0 + t), 0).$$

Hence $D_{c_2'(0)}N = (1/R)c_2'(0)$. Hence L_p is given by $L_p = \begin{pmatrix} 0 & 0 \\ 0 & 1/R \end{pmatrix}$ with respect to this basis. Hence $H_p = 1/2R$ and $K_p = 0$.

3) Let $S := S^2(R)$ be the sphere of radius R centered at the origin. Then the unit normal filed is given by $N_p = p/R$. Then,

$$D_vN := \frac{d}{dt}|_{t=0}\, N \circ c(t) = \frac{d}{dt}|_{t=0}\, \frac{c(t)}{R} = \frac{c'(0)}{R} = \frac{v}{R}.$$

Thus L_p is the scalar operator $(1/R)$ Identity. Hence every direction is principal and the principal curvatures are $\frac{1}{R}, \frac{1}{R}$ so that $H_p = (1/R)$ and $K_p = 1/R^2$. Note that this implies that larger the radius less curved is the sphere, which is intuitively appealing.

4) For the saddle surface $S = \{z = xy \text{ in } \mathbb{R}^3\}$ and $p = 0$, L_p is given by $L_pe_1 = -e_2$ and $L_{e_2} = -e_2$. Hence $L_p = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ with eigen values $1, -1$ with eigen directions $e_1 - e_2$ and $e_1 + e_2$. Details are left to the reader. We have $H_p = 0$ and $K_p = -1$.

2 Examples of Surfaces acc. to Definition 1.2

This section may be omitted on first reading.

Example 11. $S(r)$ has a parameterization $(u, v) \mapsto (r \cos u \cos v, r \cos u \sin v, r \sin u)$ for $0 < v < 2\pi$ and $0 < u < \pi$.

Example 12. A surface of revolution. Let $c\colon (a, b) \ni u \mapsto (x(u), 0, z(u))$ be a curve in the xz-plane. We revolve it around the z-axis to get a surface of revolution with the parameterization: $\varphi(u, v) := (x(u) \cos v, x(u) \sin v, z(u))$ for $(u, v) \in (a, b) \times (0, 2\pi)$. (Notice that Eg. 11 is a special case.)

Example 13. Let $f : U \to \mathbb{R}$ be a smooth function on an open set U of E^2. Then $S := \{(x, y, z) : z = f(x, y), x, y, z \in \mathbb{R}\}$ is a surface. The parameterization is given by $(x, y) \mapsto (x, y, f(x, y))$.

Example 14. Let $f : \Omega \subset E^3 \to \mathbb{R}$ be a smooth function. We say that $a \in f(\Omega)$ is a *regular value* if for any $p \in \Omega$ with $f(p) = a$, the gradient of f at p is nonzero: $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})(p) \neq 0$. If a is a regular value of f then $S := \{p \in \Omega : f(p) = a\}$ is a surface. I shall leave this as an instructive exercise in the use of inverse function theorem.

We define a (smooth) curve in a surface S as follows: $c\colon (-\varepsilon, \varepsilon) \to S$ is a curve if for every $s \in (-\varepsilon, \varepsilon)$, we have a neighbourhood V of $c(s)$ in E^3 and a $\delta > 0$ such that the restriction $c\colon (s - \delta, s + \delta) \to V$ is smooth. We can therefore speak of tangents and the lengths of curves as earlier.

Example 15. In $S(r)$, consider the equator c in the xy-plane: $c\colon s \mapsto r \cos s e_1 + r \sin s e_2 \equiv (r \cos s, r \sin s, 0)$ for $0 \leq s \leq 2\pi$. Then $c'(s) = -r \sin s e_1 + r \cos s e_2 \equiv (-r \sin s, r \cos s, 0)$ so that $\| c'(s) \| = r$ for all s and hence $l(c) = 2\pi r$, as is to be expected.

Ex. 16. 1) Let $\varphi : U \to V \cap S$ be a parameterization around $p \in S$. We let $X_1 := (\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u})$ and $X_2 := (\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v})$. X_i are tangent vectors at $\varphi(u, v)$ to S and they are linearly independent for all $(u, v) \in U$. (Can you think of some natural curves which have these as their tangent vectors?)

Ex. 17. We let $T_pS := \mathbb{R}X_1 \oplus \mathbb{R}X_2$ for all $p = \varphi(u, v)$ with $(u, v) \in U$. Then T_pS is a 2-dimensional subspace of E^3 independent of the parameterization. It is called the *tangent space* of S at p.

Ex. 18. The vector or the cross product $\mathbf{n}_p := X_1 \times X_2(u, v)$ is nonzero. It is called the *normal field* to the surface at p, since it is perpendicular to T_pS.

Ex. 19. Verify that for $S = S(r)$ the normal at $x \in S$ is given by $\mathbf{n}_x = x/r$. In the case of $S = f^{-1}(a)$ (Eg. 14), we have $\mathbf{n}_p = \nabla f(p)$, the gradient of f at p.

3 Gaussian Curvature

We now indicate the geometric meaning of the Gaussian curvature.

If S is a surface in E^3 and (φ, U) is a parameterization of a piece of S, the vector field $\mathbf{n}_p := X_1 \times X_2$ is a nonzero normal vector field on $\varphi(U) \subset S$. Hence we have a normal field $\mathbf{n}$ of unit norm on $\varphi(U)$, unique up to sign. Thus we have a map $\mathbf{n} : \varphi(U) \to S^2(1)$, called the *Gauss map*.

Example 20. 1) For $S = S(r)$, the Gauss map is $x \mapsto x/r$.
2) For $S = E^2$, the Gauss map is the constant $x \mapsto e_3$.
3) For the cylinder in Eg. 3, the Gauss map is $(x, y, z) \mapsto (1/r)(x, y, 0)$.

Definition 21. As x varies over a small *oriented* area around p in S, the Gauss map $\mathbf{n}_x$ sweeps out an *oriented* area in $S^2(1)$. The rate of change of this *oriented* area is called the *Gaussian curvature* of S at p.

A more precise definition is analytical. We define $K(p)$, the Gaussian curvature of S at p to be the determinant of the Jacobian $d\mathbf{n}_p$ of the Gauss map. This is what we should expect if we remember the change of variable formula from calculus of several variables. It is important to realize that $K(p)$ does not depend on the choice of the sign of $\mathbf{n}$.

Example 22. 1) For the sphere $S(r)$ the Gauss map can be thought of the restriction of the linear map $x \mapsto x/r$ from E^2 to itself and hence the curvature is given by $K(p) = 1/r^2$. Thus a sphere of smaller radius is more 'curved' than a sphere of larger radius. This is certainly clear if you look at the way the Gauss map sweeps out areas.

2) The plane has curvature 0. This is clear either from the geometric definition or from the analytical one.

3) The cylinder springs a surprise on almost everybody. Contrary to general expectations, it is not 'curved' at all, i.e., its Gaussian curvature is zero. For as x varies over a region of the surface, $\mathbf{n}_x$ varies over a part of the equator only, whose *area* is 0!

The theorem of Gauss says that the Gaussian curvature is intrinsic in the sense that it depends only on the inner product structure on the tangent spaces of S.

A more precise statement is given below.

A function $f : S \to \mathbb{R}$ is said to be smooth if for each $p \in S$ there exists an open set W_p in R^3 and a smooth function $F : W_p \to \mathbb{R}$ such that $F \mid_{W \cap S} = f$. Given a surface S, there exists two important classes of smooth functions:
i) $f_a : S \to \mathbb{R}$ given by $f(x) := \|x - a\|^2$, for a fixed $a \in \mathbb{R}^3$.
ii) $h_u : S \to \mathbb{R}$ given by $h_u(x) := \langle x, u\rangle$ for a fixed unit vector $u \in \mathbb{R}^3$. g_u is called the height function in the direction u.

For any smooth function $f : S \to \mathbb{R}$ we have a linear map $f'(p)$ or $df(p)$ on T_pS given by $f'(p)(v) := (d/dt)(f \circ c(t)) \mid_{t=0}$, for any curve c with $c(0) = p$ and $c'(0) = v$. By Riesz representation theorem there exists a unique vector, say, $\nabla f(p) \in T_pS$ such that $f'(p)(v) = \langle v, \nabla f(p)\rangle$ for all $v \in T_pS$. The map $p \mapsto \nabla f(p)$ is called the *gradient* of f.

Let $h := h_u$ be as above, Then

$$h'(p)(v) = \frac{d}{dt} g \circ c(t) \mid_{t=0} == \frac{d}{dt} \langle c(t), u\rangle \mid_{t=0} = \langle c'(0), u\rangle .$$

Hence the $\nabla h(p) = u^T$, the projection of u onto the T_pS part of the orthogonal decomposition $\mathbb{R}^3 = T_pS \oplus \mathbb{R}N_p$. Note that p is a critical point of h_u (i.e., $h_u'(p) = 0$) iff u is normal to S at p.

For $f = f_a$, proceeding as above, we find that $\nabla f(p) = (p - a)^T$.

If $\varphi : S_1 \to S_2$ is any continuous map between surfaces, the expression $\varphi = (\varphi_1, \varphi_2, \varphi_3)$ allows us to define smoothness of φ. The derivative $d\varphi(p)$ or $\varphi'(p)$ is the linear map from $T_pS_1 \to T_{\varphi(p)}S_2$ given by $\varphi'(p)(v) := (d/dt)\varphi \circ c(t)\,|_{t=0}$ if $v = c'(0) \in T_pS_1$. A smooth map $\varphi : S_1 \to S_2$ is said to be isometry if φ is bijective, φ^{-1} is smooth and $\varphi'(p)$ is linear isometry of T_pS_1 onto $T_{\varphi(p)})S_2$. Equivalently, if c is any curve in S_1 then $\ell(c) = \ell(\varphi \circ c)$.

Theorem 23 (Gauss)**.** *If $\varphi\colon S_1 \to S_2$ is an isometry of the surfaces, then $K_{S_1}(p) = K_{S_2}(\varphi(p))$.* □

An illustrative example of this result is the map φ from the piece of the plane $S_1 := (0, 2\pi) \times \mathbb{R} \times \{0\} \subset \mathbb{R}^3$ to the cylinder $S_2 : \{x^2 + y^2 = 1\}$ given by $\varphi(u, v, 0) = (\cos u, \sin u, v)$. Note however that the mean curvatures are different.

A most instructive way of proving Gauss' theorem is to define an intrinsic derivation of tangent fields on S. You may recall that the curvature of a plane curve c with unit speed is got as the length of the derivative of the tangent field $c'(s)$. Hence, on a surface, we wish to differentiate a tangent field Y with respect to a direction $v \in T_pS$ to get a tangent vector, say, D_vY. The natural thing to do is to define $D_vY := \frac{d}{dt}(Y \circ c(t))\,|_{t=0}$ with an obvious notation. The trouble here is that $\frac{d}{dt}(Y \circ c(t))\,|_{t=0}$ may not be a tangent vector. (See what happens if you consider D_XX where X is a tangent field on the unit sphere S.) But however there is no need to despair as we can easily repair this. We set

$$D_vY := (\nabla_vY)^T,$$

where ∇_vY is the usual gradient in $\mathbb{R}^3$. It is easy to see that D satisfies the following properties:

Let X_i denote the coordinate tangent fields in an open set of S. That is, $X_i := d\varphi(\frac{\partial}{\partial x_i})$. For X, Y, Z tangent fields on S and f, g smooth functions on S we have

1. $D_X(Y + Z) = D_XY + D_XZ$
2. $D_{(X+Y)}Z = D_XZ + D_YZ$
3. $D_{fX}Y = fD_XY$
4. $D_X(gY) = gD_XY + X(g)Y$
5. $D_{X_i}X_j - D_{X_j}X_i = 0$
6. $Z\langle X, Y\rangle = \langle D_ZX, Y\rangle + \langle X, D_ZY\rangle$

Now it is easy to show that there is a unique D satisfying the above properties and it is given by

$$\langle D_{X_i}X_j, X_k\rangle = \frac{1}{2}\left[\frac{\partial}{\partial x_i}\langle X_j, X_k\rangle + \frac{\partial}{\partial x_j}\langle X_k, X_i\rangle - \frac{\partial}{\partial x_k}\langle X_i, X_j\rangle\right]. \tag{1}$$

Thus D_XY is defined intrinsically, i.e., using only the concepts involving the objects defined on S and not anything extraneous as the normal. The way to prove Eq. 1 is to start with $X_i(\langle X_j, X_k\rangle$ and use the properties of D to expand and cyclically change i, j, k.

Use the definition $D_{X_i}X_j = (\nabla_{X_i}X_j)^T = \nabla_{X_i}X_j - \langle L_{X_i}, X_j\rangle N$ in the equation $D_{X_i}X_j - D_{X_j}X_i = 0$ and collect the tangent and normal terms (using the symmetry of L) to get

$$\left\langle D_{X_i}D_{X_j} - D_{X_j}D_{X_i}X_j, X_i\right\rangle = \langle Lx_i, X_i\rangle \langle LX_j, X_j\rangle - \langle LX_i, X_j\rangle^2 = \det(L) = K$$

This last equation proves Gauss theorem.

For further details the reader is referred to [1] or [2].

4 Geodesics

Definition 24. A *geodesic* on a surface S is a curve $c\colon(-\varepsilon,\varepsilon)\to S$ with zero acceleration as observed from the surface. By this we mean that the *acceleration* $c''(s) \perp T_{c(s)}S$ for all s. Thus, the tangential component $(c'')^\top = 0$.

A more intuitive definition runs as follows: The length of the segment between any two of its sufficiently nearby points is less than or equal to the length of any curve joining these points. That is, given any two points $c(t), c(t+\delta)$ sufficiently nearby, the length of the segment of the curve between these points, viz., $\int_t^{t+\delta}\|c'(s)\|\,ds$ is less than or equal to $l(\sigma)$, the length of any curve σ that joins these points. This definition is equivalent to the above one, but we shall not prove it in our lectures.

The analytical definition $(c'')^\top = 0$ can be translated into a system of second order ordinary differential equations via parameterization. Hence by the existence and uniqueness theorem in the theory of ordinary differential equations, it follows that there exists a unique geodesic with given initial data. That is, if $p \in S$ and $v \in T_pS$ are given there exists a "unique" geodesic $c_{p,v} : (-\varepsilon,\varepsilon) \to S$ such that $c(0) = p$ and $c'(0) = v$. (ε may depend upon v.)

Example 25. The great circles, i.e., the intersection of the planes through the origin and $S(r)$, are the geodesics on a sphere $S(r)$. We can describe these without messy notation. Let $x \in S$ and $v(\neq 0) \in T_pS$. Note that this means $\langle x, v\rangle = 0$. (See Ex. 16.) We then want the description of a circle centered at the origin of radius r in the plane spanned by x and v:

$$c(s) \equiv c_{x,v} := \cos sx + \frac{r}{\|v\|}\sin sv.$$

One easily sees that $c''(s) = -c(s) \perp T_pS$ and that $c_{x,av}(s) = c_{x,v}(as)$ for any $a \in \mathbb{R}$.

Example 26. The curves $c(s) := (r\cos s, r\sin s, rs)$ are geodesics on a cylinder $S := \{(x,y,z) : x^2 + y^2 = r^2\}$.

A more natural way of parameterizing a geodesic through the arc length, i.e., a parameterization with $\|c'\| = 1$. If $c\colon [0,T] \to S$ is a geodesic, we first of all note that, for any t, $c'(t) \neq 0$ since otherwise it has to be the constant curve due to uniqueness. (Do you understand this reasoning completely?) If we introduce $s(t) := \int_0^t \|c'(t)\|\,dt$, s is then a strictly increasing function and hence we can use it to reparametrize c. In the new s-parameterization, we have $l(c|_{[0,a]}) = a$ for any $0 \le a \le T$.

The geodesics of the sphere with respect to the arc length are given by $c(s) = \cos \frac{s}{r} x + r \sin \frac{s}{r} v$ for $v \in T_pS$ with $\| v \| = 1$. (This is enough; see the last line of Eg. 25.) The reader is urged to check the details of this.

References

[1] N.J. Hicks, *Notes on Differential Geometry*, D. Van Nostrand, 1966

[2] S. Kumaresan, *A Course in Riemannian Geometry*, (Preliminary Version-1990)

[3] do Carmo, *Differential Geometry of Curves and Surfaces*