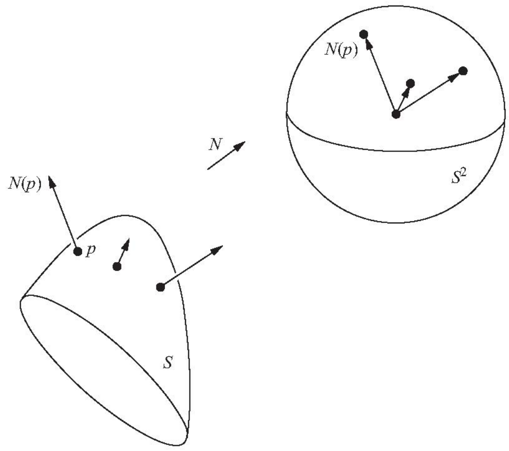
The Geometry of the Gauss Map

Throughout this chapter, S will denote a regular orientable surface in which an orientation (i.e., a differentiable field of unit normal vectors N) has been chosen; this will be simply called a surface S with an orientation N.

Definition Let $S \subset \mathbb{R}^3$ be a surface with an orientation N. The map $N : S \to \mathbb{R}^3$ takes its values in the unit sphere

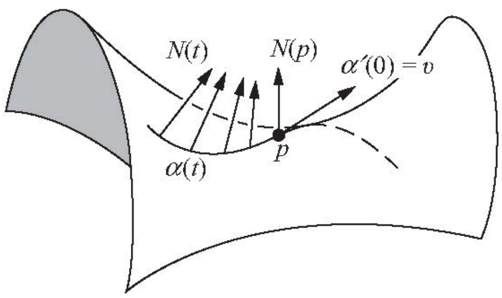
$$S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$$

The map $N : S \to S^2$, thus defined, is called the Gauss map of S.



It is straightforward to verify that the Gauss map is differentiable. Thus the differential dN_p is a linear map from T_pS to $T_{N(p)}S^2$. Since T_pS and $T_{N(p)}S^2$ are the same vector spaces, dN_p can be viewed as a linear map $dN_p : T_pS \to T_pS$ from T_pS to itself defined as follows.

For each parametrized curve $\alpha(t)$ in S with $\alpha(0) = p$, we consider the parametrized curve $N \circ \alpha(t) = N(t)$ in the sphere S^2; thhis amounts to restricting the normal vector N to the curve $\alpha(t)$. The tangent vector $N'(0) = dN_p(\alpha'(0))$ is a vector in T_pS. It measures the rate of change of the normal vector N, restricted to the curve $\alpha(t)$, at $t = 0$. Thus, dN_p measures how N pulls away from $N(p)$ in a neighborhood of p. In the case of curves, this measure is given by a number, the curvature. In the case of surfaces, this measure is characterized by a linear map.



Examples

1. Let $S = \{(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz + d = 0\}$ be a plane in $\mathbb{R}^3$. Then the unit normal vector $N = (a, b, c)/\sqrt{a^2 + b^2 + c^2}$ is a constant, and therefore $dN \equiv 0$, i.e. $dN_p(v) = 0v = 0 \in T_{N(p)}S = T_pS$ for all $p \in S$ and all $v \in T_pS$.

2. Consider the unit sphere

$$S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}.$$

If $\alpha(t) = (x(t), y(t), z(t))$ is a parametrized curve in S^2, then

$$2x(t)x'(t) + 2y(t)y'(t) + 2z(t)z'(t) = 0 \iff \langle \alpha(t), \alpha'(t) \rangle = 0,$$

which shows that the vector (x, y, z) is normal to the sphere at the point (x, y, z). Thus, $\bar{N} = (x, y, z)$ and $N = (-x, -y, -z)$ are fields of unit normal vectors in S^2. Restricted to the curve $\alpha(t)$, the normal vectors

$$\begin{aligned} N(t) &= (-x(t), -y(t), -z(t)) &\implies dN(x'(t), y'(t), z'(t)) = N'(t) = (-x'(t), -y'(t), -z'(t)), \\ \bar{N}(t) &= (x(t), y(t), z(t)) &\implies d\bar{N}(x'(t), y'(t), z'(t)) = \bar{N}'(t) = (x'(t), y'(t), z'(t)) \end{aligned}$$

that is, $dN_p(v) = -v$ and $d\bar{N}_p(v) = v$ for all $p \in S^2$ and all $v \in T_pS^2$.

3. Consider the cylinder $S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1\}$. By an argument similar to that of the previous example, we see that $\bar{N} = (x, y, 0)$ and $N = (-x, -y, 0)$ are unit normal vectors at (x, y, z). If $(x(t), y(t), z(t))$ is a parametrized curve in the cylinder, since

$$(x(t))^2 + (y(t))^2 = 1 \implies 2x(t)x'(t) + 2y(t)y'(t) = 0,$$

we are able to see that, along this curve,

$$N(t) = (-x(t), -y(t), 0) \implies dN(x'(t), y'(t), z'(t)) = N'(t) = (-x'(t), -y'(t), 0).$$

This implies that if v is a vector tangent to the cylinder and parallel to the z axis and if w is a vector tangent to the cylinder and parallel to the xy plane, then

$$dN(v) = 0 = 0v; \quad dN(w) = -w.$$

It follows that v and w are eigenvectors of dN with eigenvalues 0 and -1, respectively.

4. Let us analyze the point $p = (0, 0, 0)$ of the hyperbolic paraboloid $z = y^2 - x^2$. For this, we consider a parametrization $X(u, v)$ given by

$$X(u, v) = (u, v, v^2 - u^2),$$

and compute the normal vector$N(u, v)$. We obtain $X_u = (1, 0, -2u)$, $X_v = (0, 1, 2v)$ and

$$N = X_u \wedge X_v = \left(\frac{u}{\sqrt{u^2 + v^2 + \frac{1}{4}}}, \frac{-v}{\sqrt{u^2 + v^2 + \frac{1}{4}}}, \frac{1}{2\sqrt{u^2 + v^2 + \frac{1}{4}}} \right).$$

If $\alpha(t) = X(u(t), v(t))$ is a curve with $\alpha(0) = p$ then the tangent vector $\alpha'(0)$ has coordinates $(u'(0), v'(0), 0)$. Restricting $N(u, v)$ to this curve and computing $N'(0)$, we obtain

$$N'(0) = (2u'(0), -2v'(0), 0),$$

and therefore, at p,

$$dN_p(u'(0), v'(0), 0) = (2u'(0), -2v'(0), 0).$$

It follows that the vectors $(1, 0, 0)$ and $(0, 1, 0)$ are eigenvectors of dN_p with eigenvalues 2 and -2, respectively.

5. The method of the previous example, applied to the point $p = (0,0,0)$ of the paraboloid $x = x^2 + ky^2$, $k > 0$, shows that the unit vectors of the x axis and the y axis are eigenvectors of dN_p, with eigenvalues 2 and $2k$, respectively (assuming that N is pointing outwards from the region bounded by the paraboloid).

Proposition The differential $dN_p : T_pS \to T_pS$ of the Gauss map is a self-adjoint linear map.

Proof Since $dN_p : T_pS \to T_pS$ is linear, it suffices to verify that

$$\langle dN_p(w_1), w_2\rangle = \langle w_1, dN_p(w_2)\rangle \quad \text{for a basis } \{w_1, w_2\} \text{ of } T_pS.$$

Let $X(u,v)$ be a parametrization of S at p and $\{X_u, X_v\}$ the associated basis of T_pS. If $\alpha(t) = X(u(t), v(t))$ is a parametrized curve in S, with $\alpha(0) = p$, we have

$$\begin{aligned} dN_p(\alpha'(0)) &= dN_p(X_u u'(0) + X_v v'(0)) \\ &= \frac{d}{dt} N(u(t), v(t))|_{t=0} \\ &= N_u u'(0) + N_v v'(0); \end{aligned}$$

in particular, $dN_p(X_u) = N_u$ and $dN_p(X_v) = N_v$. Therefore, to prove that $dN_p : T_pS \to T_pS$ is self-adjoint, it suffices to show that

$$\langle N_u, X_v\rangle = \langle dN_p(X_u), X_v\rangle = \langle X_u, dN_p(X_v)\rangle = \langle X_u, N_v\rangle.$$

Differentiating the equations $\langle N, X_u\rangle = 0$ and $\langle N, X_v\rangle = 0$ with respect to v and u, respectively, we get

$$\begin{aligned} &\langle N_v, X_u\rangle + \langle N, X_{uv}\rangle = 0 \implies \langle N_v, X_u\rangle = -\langle N, X_{uv}\rangle \\ &\langle N_u, X_v\rangle + \langle N, X_{vu}\rangle = 0 \implies \langle N_u, X_v\rangle = -\langle N, X_{vu}\rangle \\ \implies \quad &\langle N_v, X_u\rangle = -\langle N, X_{uv}\rangle = \langle N_u, X_v\rangle. \end{aligned}$$

This proves that $dN_p : T_pS \to T_pS$ is a self-adjoint linear map.

Remark Let V be a vector space of dimension 2, endowed with an inner product $\langle\ ,\ \rangle$. We say that a linear map $A : V \to V$ is self-adjoint if $\langle Av, w\rangle = \langle v, Aw\rangle$ for all $v, w \in V$.

If $\{e_1, e_2\}$ is an orthonormal basis for V and (α_{ij}), $i, j = 1, 2$, is the matrix of A relative to that basis, then

$$\langle Ae_j, e_i\rangle = \alpha_{ij} = \langle e_j, Ae_i\rangle = \alpha_{ji},$$

that is, the matrix (α_{ij}) is symmetric.

To each self-adjoint linear map we associate a map $B : V \times V \to \mathbb{R}$ defined by

$$B(v, w) = \langle Av, w\rangle \quad \forall\, v, w \in V.$$

B is clearly bilinear; that is, it is linear in both v and w. Moreover, the fact that A is self-adjoint implies that $B(v, w) = B(w, v)$; that is, B is a bilinear symmetric form in V.

Conversely, if B is a bilinear symmetric form in V, we can define a linear map $A : V \to V$ by $\langle Av, w\rangle = B(v, w)$ and the symmetry of B implies that A is self-adjoint.

On the other hand, to each symmetric, bilinear form B in V, there corresponds a quadratic form Q in V given by

$$Q(v) = B(v, v), \quad v \in V,$$

and the knowledge of Q determines B completely, since

$$B(v,w) = \frac{1}{2}\left[Q(v+w) - Q(v) - Q(w)\right].$$

Thus, a one-to-one correspondence is established between quadratic forms in V and self-adjoint linear maps of V.

The fact that $dN_p : T_pS \to T_pS$ is a self-adjoint linear map allows us to associate to dN_p a quadratic form in T_pS defined as follows.

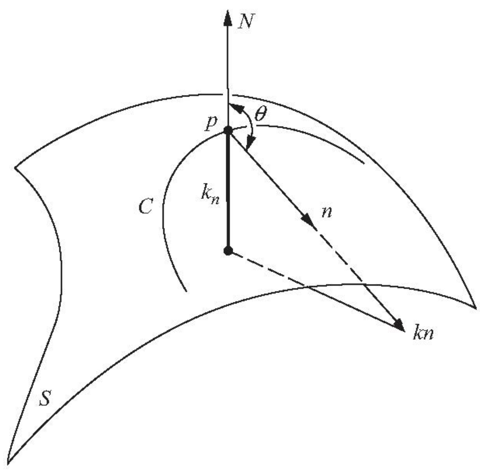
Definition The quadratic form $II_p : T_pS \to \mathbb{R}$, defined by

$$II_p(v) = -\langle dN_p(v), v\rangle, \quad v \in T_pS,$$

is called the second fundamental form of S at p.

Definition Let C be a regular curve in S passing through $p \in S$, k the curvature of C at p, and $\cos\theta = \langle n, N\rangle$, where n is the normal vector to C and N is the normal vector to S at p. The number $k_n = k\cos\theta$ is then called the normal curvature of $C \subset S$ at p.

In other words, k_n is the length of the projection of the vector kn over the normal to the surface at p, with a sign given by the orientation N of S at p.



Remarks

- The normal curvature of C does not depend on the orientation of C but changes sign with a change of orientation for the surface.
- Let $v \in T_pS$ be a unit tangent vector to S at p, and let $C \subset S$ be a regular curve parametrized by $\alpha(s) : I \to S$, where s is the arc length of C, and with $\alpha(0) = p$, $\alpha'(0) = v$. Let $N(s) = N \circ \alpha(s)$ be the restriction of the normal vector N to the curve $\alpha(s)$. For all $s \in I$, since

$$\langle N(s), \alpha'(s)\rangle = 0 \implies \langle N'(s), \alpha'(s)\rangle + \langle N(s), \alpha''(s)\rangle = 0 \implies -\langle N'(s), \alpha'(s)\rangle = \langle N(s), \alpha''(s)\rangle,$$

we have

$$\begin{aligned} II_p(\alpha'(0)) &= -\langle dN_p(\alpha'(0)), \alpha'(0)\rangle = -\langle dN_p(-\alpha'(0)), -\alpha'(0)\rangle = II_p(-\alpha'(0)) \\ &= -\langle N'(0), \alpha'(0)\rangle = \langle N(0), \alpha''(0)\rangle \\ &= \langle N(p), kn(p)\rangle \overset{(*)}{=} k_n(p) \implies k_n(p) \text{ depends only on } v = \alpha'(0) \text{ and } II_p \end{aligned}$$

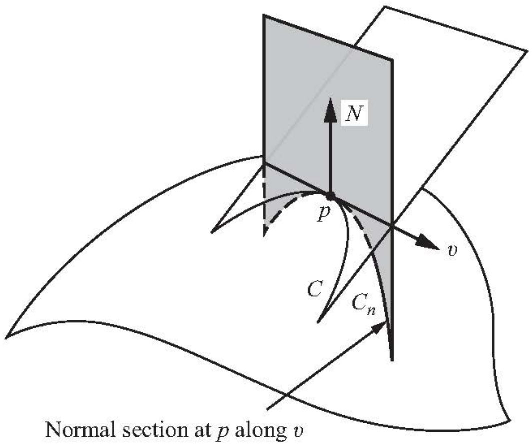
that is, the value of the second fundamental form II_p for a unit vector $v \in T_pS$ is equal to the normal curvature of a regular curve passing through p and tangent to v. In particular, we obtained the following result.

Proposition (Meusnier) All curves lying on a surface S and having at a given point $p \in S$ the same tangent line have at this point the same normal curvatures.

Remark Given a unit vector $v \in T_pS$, the intersection of S with the plane containing v and $N(p)$ is called the normal section of S at p along v. In a neighborhood of p, a normal section of S at p is a regular plane curve on S whose normal vector

$$n(p) = \pm N(p) \text{ or } 0 \text{ (a zero vector when } k = 0);$$

and, by $(*)$, $k(p) = |k_n(p)|$. With this terminology, the above proposition says that the absolute value of the normal curvature at p of a curve $\alpha(s)$ is equal to the curvature of the normal section of S at p along $\alpha'(0)$.



C and C_n have the same normal curvature at p along v.

Examples

1. Let S be the surface of revolution obtained by rotating the curve $z = y^4$, parametrized by $\alpha(t) = (0, t, t^4)$, $t \in \mathbb{R}$, about the z axis and let $p = \alpha(0) = (0,0,0) \in S$. Since
   - $k(p) = \dfrac{|\alpha'(0) \times \alpha''(0)|}{|\alpha'(0)|^3} = 0,$
   - $T_pS = \{(x,y,0) \mid (x,y) \in \mathbb{R}^2\}$, the xy plane,

   $\Longrightarrow$ $N(p) \mathbin{/\!/} (0,0,1)$, i.e. the normal vector $N(p)$ is parallel to the z axis,

   and any normal section at p is obtained from the curve $z = y^4$ by rotation; hence, it has curvature zero. It follows that all normal curvatures are zero at p, and thus $dN_p = 0$.

2.
   - Let $S = \{(x,y,z) \in \mathbb{R}^3 \mid ax+by+cz+d = 0\}$ be a plane in $\mathbb{R}^3$. Since all normal sections are straight lines, all normal curvatures are zero. Thus, the second fundamental form is identically zero at all points. This agrees with the fact that $dN_p = 0$ for all $p \in S$.
   - Let $S^2 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 = 1\}$ with N as orientation, the normal sections through a point $p \in S^2$ are circles with radius 1. Thus, all normal curvatures are equal to 1, and the second fundamental form is $II_p(v) = 1$ for all $p \in S^2$ and all $v \in T_pS^2$ with $|v| = 1$.

- Let $S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1\}$ be a cylinder in $\mathbb{R}^3$. Since the normal sections at a point p vary from a circle perpendicular to the axis of the cylinder to a straight line parallel to the axis of the cylinder, passing through a family of ellipses, the normal curvatures varies from 1 to 0. It is not hard to see geometrically that 1 is the maximum and 0 is the minimum of the normal curvature at p.

Lemma If the function $Q(x, y) = ax^2 + 2bxy + cy^2$, restricted to the unit circle $x^2 + y^2 = 1$, has a maximum at the point $(1, 0)$, then $b = 0$.

Proof Parametrize the circl $x^2 + y^2 = 1$ by $x = \cos t$, $y = \sin t$, $t \in (-\varepsilon, 2\pi + \varepsilon)$. Thus, $t = 0$ is an interior point of $(-\varepsilon, 2\pi + \varepsilon)$, and Q, restricted to that circle, becomes a function of t:

$$Q(t) = a\cos^2 t + 2b\cos t\sin t + c\sin^2 t, \quad t \in (-\varepsilon, 2\pi + \varepsilon).$$

Since Q has a maximum at the interior point $(1, 0)$ we have

$$\left(\frac{dQ}{dt}\right)_{t=0} = 2b = 0.$$

Hence, $b = 0$ as we wished.

Proposition Given a quadratic form Q in V, there exists an orthonormal basis $\{e_1, e_2\}$ of V such that if $v \in V$ is given by $v = xe_1 + ye_2$, then

$$Q(v) = \lambda_1 x^2 + \lambda_2 y^2,$$

where λ_1 and λ_2 are the maximum and minimum, repectively, of Q on the unit circle $|v| = 1$.

Proof Let λ_1 be the maximum of Q on the unit circle $|v| = 1$, and let e_1 be a unit vector with $\max_{|v|=1} Q(v) = Q(e_1) = \lambda_1$. Such an e_1 exists by continuity of Q on the compact set $|v| = 1$. Let e_2 be a unit vector that is orthogonal to e_1, and set $\lambda_2 = Q(e_2)$. We shall show that the basis $\{e_1, e_2\}$ satisfies the conditions of the proposition.

Let B be the symmetric bilinear form that is associated to Q and set $v = xe_1 + ye_2$. Then

$$Q(v) = B(v, v) = B(xe_1 + ye_2, xe_1 + ye_2) = \lambda_1 x^2 + 2bxy + \lambda_2 y^2, \quad \text{where } b = B(e_1, e_2).$$

By the lemma, $b = B(e_1, e_2) = 0$, and thus $Q(v) = \lambda_1 x^2 + \lambda_2 y^2$, for $v = xe_1 + ye_2 \in V$.

Furthermore, for any $v = xe_1 + ye_2$ with $x^2 + y^2 = 1$, since $\lambda_1 = \max_{|v|=1} Q(v) \geq Q(e_2) = \lambda_2$,

$$Q(v) = \lambda_1 x^2 + \lambda_2 y^2 \geq \lambda_2(x^2 + y^2) = \lambda_2 = Q(e_2) \implies \lambda_2 = \min_{|v|=1} Q(v).$$

Since

Theorem Let $A : V \to V$ be a self-adjoint linear map. Then there exists an orthonormal basis $\{e_1, e_2\}$ of V such that $A(e_1) = \lambda_1 e_1$, $A(e_2) = \lambda_2 e_2$ (that is, e_1 and e_2 are eigenvectors and λ_1, λ_2 are eigenvalues of A). In the basis $\{e_1, e_2\}$, the matrix of A is clearly diagonal and the elements λ_1, λ_2, $\lambda_1 \geq \lambda_2$, on the diagonal are the maximum and the minimum, respectively, of the quadratic form $Q(v) = \langle Av, v\rangle$, on the unit circle of V.

Proof Consider the quadratic form $Q(v) = \langle Av, v\rangle$. By proposition above, there exists an orthonormal basis $\{e_1, e_2\}$ of V such that

$$Q(e_1) = \lambda_1 = \max_{|v|=1|} Q(v) \geq \min_{|v|=1|} Q(v) = \lambda_2 = Q(e_2).$$

By setting $Ae_1 = \alpha_{11}e_1 + \alpha_{21}e_2$ and $Ae_2 = \alpha_{12}e_1 + \alpha_{22}e_2$, since

$$\alpha_{11} = \langle Ae_1, e_1\rangle = Q(e_1) = \lambda_1 \quad \text{and} \quad \alpha_{21} = \langle Ae_1, e_2\rangle = B(e_1, e_2) = 0 \implies Ae_1 = \lambda_1 e_1,$$

and since

$$\alpha_{12} = \langle Ae_2, e_1\rangle = B(e_2, e_1) = 0 \quad \text{and} \quad \alpha_{22} = \langle Ae_2, e_2\rangle = Q(e_2) = \lambda_2 \implies Ae_2 = \lambda_2 e_2,$$

and in the basis $\{e_1, e_2\}$, the matrix of A is $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

Remark For each $p \in S$ there exists an orthonormal basis $\{e_1, e_2\}$ of T_pS such that

$$dN_p(e_1) = -k_1 e_1 \text{ and } dN_p(e_2) = -k_2 e_2, \quad \text{where } k_1 = \max_{v\in T_pS,\, |v|=1} II_p(v) \text{ and } k_2 = \min_{v\in T_pS,\, |v|=1} II_p(v).$$

Definition The maximum normal curvature k_1 and the minimum normal curvature k_2 are called the principal curvatures at p; the corresponding directions, that is, the directions given by the eigenvectors e_1, e_2, are called principal directions at p.

Definition If a regular connected curve C on S is such that for all $p \in C$ the tangent line of C is a principal direction at p, then C is called a line of curvature of S.

Proposition (Olinde Rodrigues) A necessary and sufficient condition for a connected regular curve C on S to be a line of curvature of S is that

$$N'(t) = \lambda(t)\alpha'(t),$$

for any parametrization $\alpha(t)$ of C, where $N(t) = N \circ \alpha(t)$ and $\lambda(t)$ is a differentiable function of t. In this case, $-\lambda(t)$ is the (principal) curvature along $\alpha'(t)$.

Proof It suffices to observe that if $\alpha'(t)$ is contained in a principal direction, then $\alpha'(t)$ is an eigenvector of dN and

$$N'(t) = dN(\alpha'(t)) = \lambda(t)\alpha'(t).$$

The converse is immediate (since $dN(\alpha'(t)) = N'(t) = \lambda(t)\alpha'(t)$).

Remark For $p \in S$, let $\{e_1, e_2\}$ be the principal directions at p such that

$$dN_p(e_1) = -k_1 e_1 \text{ and } dN_p(e_2) = -k_2 e_2,$$

where $k_1 = \max_{v\in T_pS,\, |v|=1} II_p(v)$, $k_2 = \min_{v\in T_pS,\, |v|=1} II_p(v)$ are the principal curvatures at p. For each unit vector $v \in T_pS$, since $\{e_1, e_2\}$ forms an orthonormal basis of T_pS, we have

$$v = e_1 \cos\theta + e_2 \sin\theta,$$

where θ is the angle from e_1 to v in the orientation of T_pS. The normal curvature k_n along v is given by

$$\begin{aligned} k_n &= II_p(v) = -\langle dN_p(v), v\rangle \\ &= -\langle dN_p(e_1\cos\theta + e_2\sin\theta), e_1\cos\theta + e_2\sin\theta\rangle \\ &= \langle e_1k_1\cos\theta + e_2k_2\sin\theta, e_1\cos\theta + e_2\sin\theta\rangle \\ &= k_1\cos^2\theta + k_2\sin^2\theta. \end{aligned}$$

The last expression is known classically as the Euler formula; actually, it is just the expression of the second fundamental form in the basis $\{e_1, e_2\}$.

Given a linear map $A : V \to V$ of a vector space of dimension 2 and given a basis $\{v_1, v_2\}$ of V, we recall that

$$\text{determinant of } A = a_{11}a_{22} - a_{12}a_{22}, \quad \text{trace of } A = \frac{a_{11} + a_{22}}{2},$$

where (a_{ij}) is the matrix of A in the basis $\{v_1, v_2\}$. It is known that these numbers do not depend on the choice of the basis $\{v_1, v_2\}$ and are, therefore, attached to the linear map A.

Definition Let $p \in S$ and let $dN_p : T_pS \to T_pS$ be the differential of the Gauss map. Then the determinant of dN_p is called the Gaussian curvature K of S at p, and the negative of half of the trace of dN_p is called the mean curvature H of S at p.

In terms of the principal curvatures we can write

$$K = \det dN_p = \begin{vmatrix} -k_1 & 0 \\ 0 & -k_2 \end{vmatrix} = \det(-dN_p) = k_1k_2, \quad H = \operatorname{tr}(-dN_p) = \operatorname{tr}\begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} = \frac{k_1 + k_2}{2}.$$

Definition A point p of a surface S is called

1. elliptic if $\det dN_p > 0$, e.g. points of a sphere $(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2$ and the point $p = (0,0,0)$ of the paraboloid $z = x^2 + ky^2$, $k > 0$.
2. hyperbolic if $\det dN_p < 0$, e.g. the point $p = (0,0,0)$ of the hyperbolic paraboloid $z = y^2 - x^2$.
3. parabolic if $\det dN_p = 0$, with $dN_p \neq 0$, e.g. points of a cylinder $(x-a)^2 + (y-b)^2 = r^2$.
4. planar if $dN_p = 0$, e.g. points of a plane $ax + by + cz + d = 0$.

It is clear that this classification does not depend on the choice of the orientation.

Definition A point $p \in S$ is called an umbilical point of S if $k_1 = k_2$.

Observe that all points of a plane ($k_1 = k_2 = 0$) are (planar) umbilical points, and all points of a sphere of radius r ($k_1 = k_2 = \frac{1}{r}$) or the point $p = (0,0,0)$ of the paraboloid $z = x^2 + y^2$ ($k_1 = k_2 = 2$) are (nonplanar) umbilical points.

It is an interesting fact that the only surfaces made up entirely of umbilical points are essentially spheres and planes.

Proposition If all points of a connected surface S are umbilical points, then S is either contained in a sphere or in a plane.

Proof Let $p \in S$ and let $X(u,v)$ be a parametrization of S at p such that the coordinate $V = X(U) \subset S$ is connected. Since each $q \in V$ is an umbilical point, there exists a differentiable function $\lambda : V \to \mathbb{R}$ such that

$$\begin{aligned} & dN_q(w) = \lambda(q)w \quad \forall\, q \in V, \forall\, w = a_1X_u + a_2X_v \in T_qS \\ \iff\ & N_ua_1 + N_va_2 = \lambda(q)(a_1X_u + a_2X_v) \quad \forall\, a_1,\, a_2 \in \mathbb{R} \\ \implies\ & N_u = \lambda(q)X_u \quad \text{and} \quad N_v = \lambda(q)X_v. \end{aligned}$$

Differentiating the first equation in v and the second one in u and subtracting the resulting equations, we obtain

$$\lambda_v(q)X_u - \lambda_u(q)X_v = 0.$$

Since X_u and X_v are linearly independent, we have

$$\lambda_u(q) = \lambda_v(q) = 0 \quad \forall\, q \in V.$$

Also since V is connected, $\lambda(q) = \lambda$ (a constant) for all $q \in V = X(U) \subset S$.

If $\lambda = 0$, then $N_u(q) = N_v(q) = 0$ for all $q \in V$ and therefore $N(q) = N_0$ (a constant vector) for all $q \in V$. Thus

$$\langle X(u,v), N_0\rangle_u = \langle X(u,v), N_0\rangle_v = 0 \quad \forall\,(u,v) \in U.$$

Since $U \subset \mathbb{R}^2$ is connected, we have

$$\langle X(u,v), N_0\rangle = d\,(\text{a constant}) \quad \forall\,(u,v) \in U$$

and all points $X(u,v)$ of V belong to a plane.

If $\lambda \neq 0$, since $N_u = \lambda X_u$, $N_v = \lambda X_v$, we have

$$\left(X(u,v) - \frac{1}{\lambda}N(u,v)\right)_u = \left(X(u,v) - \frac{1}{\lambda}N(u,v)\right)_v = 0 \quad \forall\,(u,v) \in U.$$

Thus there exists a fixed point $Y \in \mathbb{R}^3$ such that

$$X(u,v) - \frac{1}{\lambda}N(u,v) = Y \quad \forall\,(u,v) \in U \implies |(X(u,v) - Y|^2 = \frac{1}{\lambda^2}|N|^2 = \frac{1}{\lambda^2} \quad \forall\,(u,v) \in U.$$

Hence all points of $V = X(U)$ are contained in a sphere of center Y and radius $\dfrac{1}{|\lambda|}$.

Furthermore, observe that if $V = X(U)$ and $W = \bar{X}(\bar{U})$ are connected coordinate neighborhoods of $p = X(u_0, v_0) = \bar{X}(\bar{u}_0, \bar{v}_0) \in S$, then $V = X(U)$ and $W = \bar{X}(\bar{U})$ are contained in the same plane or in the same sphere by the continuity. This proves that if all points of a connected surface S are umbilical points, then S is either contained in a sphere or in a plane.

The Gauss Map in Local Coordinates

Let $X(u,v)$ be a parametrization at a point $p \in S$ of a surface S, and let $\alpha(t) = X(u(t), v(t))$ be a parametrized curve on S, with $\alpha(0) = p$. To simplify the notation, we shall make the convention that all functions to appear below denote their values at the point p.

The tangent vector to $\alpha(t)$ at p $\alpha' = X_u u' + X_v v'$ and

$$dN(\alpha') = \frac{d}{dt}N(u(t), v(t)) = N_u u' + N_v v'.$$

Since $N = \dfrac{X_u \wedge X_v}{|X_u \wedge X_v|}$, N_u, $N_v \in T_pS$, in basis $\{X_u, X_v\}$ we may write

$$(*) \quad \begin{cases} N_u = a_{11}X_u + a_{21}X_v, \\ N_v = a_{12}X_u + a_{22}X_v, \end{cases}$$

and therefore,

$$dN(\alpha') = (a_{11}u' + a_{12}v')X_u + (a_{21}u' + a_{22}v')X_v;$$

hence,

$$dN\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} u' \\ v' \end{pmatrix}.$$

This shows that in the basis $\{X_u, X_v\}$, dN is given by the matrix (a_{ij}), $i,j = 1,2$. Notice that this matrix is not necessarily symmetric, unless $\{X_u, X_v\}$ is an orthonormal basis.

On the other hand, the expression of the second fundamental form in the basis $\{X_u, X_v\}$is given by

$$\begin{aligned} II_p(\alpha') &= -\langle dN(\alpha'), \alpha' \rangle = -\langle N_u u' + N_v v', X_u u' + X_v v' \rangle \\ &= e(u')^2 + 2fu'v' + g(v')^2, \end{aligned}$$

where, since $\langle N, X_u \rangle = \langle N, X_v \rangle = 0$,

$$\begin{aligned} e &= -\langle N_u, X_u \rangle = \langle N, X_{uu} \rangle, \\ f &= -\langle N_v, X_u \rangle = \langle N, X_{uv} \rangle = \langle N, X_{vu} \rangle = -\langle N_u, X_v \rangle, \\ g &= -\langle N_v, X_v \rangle = \langle N, X_{vv} \rangle. \end{aligned}$$

We shall now obtain the values of a_{ij} in terms of the coefficients e, f, g. We shall now obtain the values of aij in terms of the coefficients e, f, g. From equations $(*)$ for N_u, N_v, we have

$$\begin{aligned} -f &= -\langle N_u, X_v \rangle = a_{11}F + a_{21}G, \\ -f &= -\langle N_v, X_u \rangle = a_{12}E + a_{22}G, \\ -e &= -\langle N_u, X_u \rangle = a_{11}E + a_{21}F, \\ -g &= -\langle N_v, X_v \rangle = a_{12}F + a_{22}G. \end{aligned}$$

where E, F and G are the coefficients of the first fundamental form in the basis $\{X_u, X_v\}$. In matrix form, we have

$$-\begin{pmatrix} e & f \\ f & g \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \iff \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} = -\begin{pmatrix} e & f \\ f & g \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \quad (\dagger),$$

and thus

$$\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} = \frac{-1}{EG - F^2} \begin{pmatrix} e & f \\ f & g \end{pmatrix} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix} = \frac{-1}{EG - F^2} \begin{pmatrix} eG - fF & -eF + fE \\ fG - gF & -fF + gE \end{pmatrix}.$$

Note that the Equations $(*)$, with (a_{ij}) defined in $(\dagger)$, are nonlinear partial differential equations of 2^{nd} order for $X = X(u, v)$, called the equations of Weingarten.

From Eq. $(\dagger)$, we immediately obtain the Gaussian curvature

$$K(p) = \det(-dN_p) = \det(a_{ij}) = \frac{eg - f^2}{EG - F^2}.$$

To compute the mean curvature, we recall that $-k_1$, $-k_2$ are the eigenvalues of dN Therefore, k_1 and k_2 satisfy the equation

$$dN(v) = -kv = -kIv \quad \text{for some } v \in T_pS,\ v \neq 0,$$

where I is the identity map. It follows that the linear map $dN + kI$ is not invertible; hence, it has zero determinant. Thus,

$$\det \begin{pmatrix} a_{11} + k & a_{12} \\ a_{21} & a_{22} + k \end{pmatrix} = 0 \iff k^2 + (a_{11} + a_{22})k + (a_{11}a_{22} - a_{12}a_{21}) = 0.$$

Since k_1 and k_2 are the roots of the above quadratic equation, we conclude that

$$H = \frac{1}{2}(k_1 + k_2) = -\frac{1}{2}(a_{11} + a_{22}) = \frac{1}{2}\frac{eG - 2fF + gE}{EG - F^2};$$

hence,

$$k^2 - 2Hk + K = 0 \iff k = H \pm \sqrt{H^2 - K}.$$

From this relation, it follows that if we choose $k_1(q) \geq k_2(q)$, $q \in S$, then the functions k_1 and k_2 are continuous in S. Moreover, k_1 and k_2 are differentiable in S, except perhaps at the umbilical points $(H^2 = K)$ of S.

Examples

1. Let U be an open subset of $\mathbb{R}^2$ and let S be the graph of a differentiable function $z = h(x, y)$, $(x, y) \in U$. Then S is parametrized by

$$X(x, y) = (x, y, h(x, y)), \quad (x, y) \in U.$$

A simple computation shows that

$$K = \frac{h_{xx}h_{yy} - h_{xy}^2}{(1 + h_x^2 + h_y^2)^2}, \quad 2H = \frac{(1 + h_x^2)h_{yy} - 2h_x h_y h_{xy} + (1 + h_y^2)h_{xx}}{(1 + h_x^2 + h_y^2)^{3/2}}.$$

2. Consider a surface of revolution parametrized by

$$X(u, v) = (\varphi(v)\cos u, \varphi(v)\sin u, \psi(v)) \quad 0 < u < 2\pi,\ a < v < b,\ \varphi(v) \neq 0.$$

The coefficients of the first fundamental form are given by

$$E = \varphi^2, \quad F = 0, \quad G = (\varphi')^2 + (\psi')^2.$$

It is convenient to assume that the rotating curve is parametrized by arc length, that is, that

$$(\varphi')^2 + (\psi')^2 = G = 1.$$

The computation of the coefficients of the second fundamental form is straightforward and yields

$$\begin{aligned}
e &= \frac{(X_u, X_v, X_{uu})}{\sqrt{EG - F^2}} = \frac{1}{\sqrt{EG - F^2}} \begin{vmatrix} -\varphi\sin u & \varphi\cos u & 0 \\ \varphi'\cos u & \varphi'\sin u & \psi' \\ -\varphi\cos u & -\varphi\sin u & 0 \end{vmatrix} = -\varphi\psi' \\
f &= \frac{(X_u, X_v, X_{uv})}{\sqrt{EG - F^2}} = \frac{1}{\sqrt{EG - F^2}} \begin{vmatrix} -\varphi\sin u & \varphi\cos u & 0 \\ \varphi'\cos u & \varphi'\sin u & \psi' \\ -\varphi'\sin u & \varphi'\cos u & 0 \end{vmatrix} = 0 \\
g &= \frac{(X_u, X_v, X_{vv})}{\sqrt{EG - F^2}} = \frac{1}{\sqrt{EG - F^2}} \begin{vmatrix} -\varphi\sin u & \varphi\cos u & 0 \\ \varphi'\cos u & \varphi'\sin u & \psi' \\ \varphi''\cos u & \varphi''\sin u & \psi'' \end{vmatrix} = \psi'\varphi'' - \psi''\varphi'
\end{aligned}$$

Since $F = f = 0$, we conclude that the parallels ($v =$const.) and the meridians ($u =$const.) of a surface of revolution are lines of curvature of such a surface.

Because

$$K = \frac{eg - f^2}{EG - F^2} = -\frac{\psi'(\psi'\varphi'' - \psi''\varphi')}{\varphi}$$

and φ is always positive, it follows that the parabolic points are given by either $\psi' = 0$ (the tangent line to the generator curve is perpendicular to the axis of rotation) or $\psi'\varphi'' - \psi''\varphi' = 0$

(the curvature of the generator curve is zero). A point which satisfies both conditions is a planar point, since these conditions imply that $e = f = g = 0$.

It is convenient to put the Gaussian curvature in still another form. By differentiating $(\varphi')^2 + (\psi')^2 = 1$ we obtain $\varphi'\varphi'' = -\psi'\psi''$. Thus

$$K = -\frac{\psi'(\psi'\varphi'' - \psi''\varphi')}{\varphi} = -\frac{(\psi')^2\varphi'' + (\varphi')^2\varphi''}{\varphi} = -\frac{\varphi''}{\varphi}.$$

3. Let $a > r > 0$, and consider the parametrization

$$X(u,v) = ((r\cos u + a)\cos v, (r\cos u + a)\sin v, r\sin u), \quad 0 < u < 2\pi,\ 0 < v < 2\pi$$

of the torus generated by rotating $S^1 = \{(y,z) \mid (y-a)^2 + z^2 = r^2\}$ about z-axis. Since

$$\begin{aligned} X_u &= (-r\sin u\cos v, -r\sin u\sin v, r\cos u), \\ X_v &= (-(r\cos u + a)\sin v, (r\cos u + a)\cos v, 0), \\ X_{uu} &= (-r\cos u\cos v, -r\cos u\sin v, -r\sin u), \\ X_{uv} &= (r\sin u\sin v, -r\sin u\cos v, 0), \\ X_{vv} &= (-(r\cos u + a)\cos v, -(r\cos u + a)\sin v, 0), \end{aligned}$$

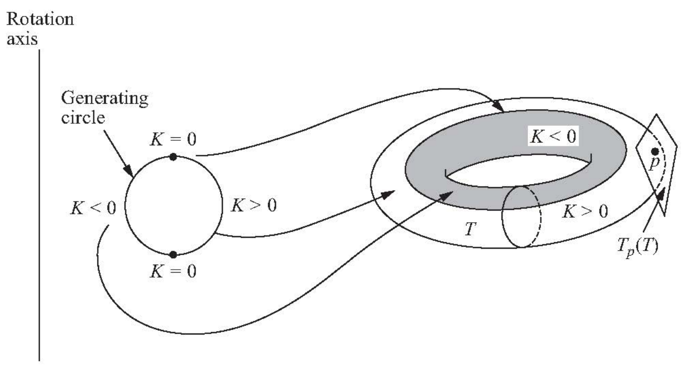
we obtain $X_u \wedge X_v = (-r\cos u\,(r\cos u + a)\cos v, -r\cos u\,(r\cos u + a)\sin v, -r\sin u\,(r\cos u + a))$,

$$E = \langle X_u, X_u\rangle = r^2, \quad F = \langle X_u, X_v\rangle = 0, \quad G = \langle X_v, X_v\rangle = (r\cos u + a)^2,$$

$|X_u \wedge X_v| = \sqrt{EG - F^2} = r(r\cos u + a)$, and

$$\begin{aligned} e &= \langle N, X_{uu}\rangle = \langle\frac{X_u \wedge X_v}{|x_u \wedge X_v|}, X_{uu}\rangle = \frac{\langle X_u \wedge X_v, X_{uu}\rangle}{\sqrt{EG - F^2}} = \frac{r^2(r\cos u + a)}{r(r\cos u + a)} = r, \\ f &= \langle N, X_{uv}\rangle = \langle\frac{X_u \wedge X_v}{|x_u \wedge X_v|}, X_{uv}\rangle = \frac{\langle X_u \wedge X_v, X_{uv}\rangle}{\sqrt{EG - F^2}} = 0, \\ g &= \langle N, X_{vv}\rangle = \langle\frac{X_u \wedge X_v}{|x_u \wedge X_v|}, X_{vv}\rangle = \frac{\langle X_u \wedge X_v, X_{vv}\rangle}{\sqrt{EG - F^2}} = \frac{r\cos u\,(r\cos u + a)^2}{r(r\cos u + a)} = \cos u(r\cos u + a), \end{aligned}$$

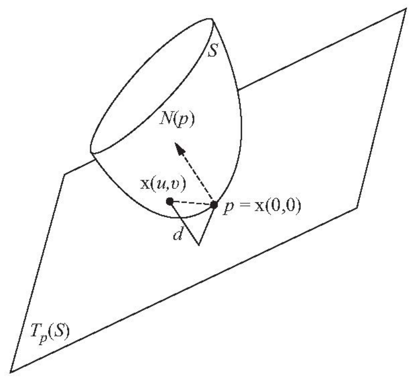
and $K = \dfrac{eg - f^2}{EG - F^2} = \dfrac{\cos u}{r(r\cos u + a)}$. Note that $K = 0$ when $u = \pi/2$ or $u = 3\pi/2$, the points of such parallels are parabolic points; $K < 0$ when $\pi/2 < u < 3\pi/2$, the points in this region are hyperbolic points; and $K > 0$ when $0 < u < \pi/2$, or $3\pi/2 < u < 2\pi$, the points in this region are elliptic points.



Proposition Let $p \in S$ be an elliptic point of a surface S. Then there exists a neighborhood V of p in S such that all points in V belong to the same side of the tangent plane T_pS. Let $p \in S$ be a hyperbolic point. Then in each neighborhood of p there exist points of S in both sides of T_pS.

Proof Let $X(u,v)$ be a parametrization of S at p, with $X(0,0) = p$, and let $d : X(U) \to \mathbb{R}$ be a function defined by

$$d(q) = \langle X(u,v) - X(0,0), N(p) \rangle, \quad \text{for } q = X(u,v) \in X(U).$$



Since $X(u,v)$ is differentiable and by the Taylor's formula, we have

$$X(u,v) = X(0,0) + X_u(0,0)u + X_v(0,0)v + \frac{1}{2}(X_{uu}(0,0)u^2 + 2X_{uv}(0,0)uv + X_{vv}(0,0)v^2) + \bar{R},$$

where the remainder $\bar{R}$ satisfies that

$$\lim_{(u,v)\to(0,0)} \frac{\bar{R}}{u^2+v^2} = 0.$$

It follows that $\langle X_u(0,0), N(p) \rangle = \langle X_v(0,0), N(p) \rangle = 0$ and

$$\begin{aligned} d(q) &= \langle X(u,v) - X(0,0), N(p) \rangle \\ &= \frac{1}{2}\left\{ \langle X_{uu}(0,0), N(p) \rangle u^2 + 2\langle X_{uv}(0,0), N(p) \rangle uv + \langle X_{vv}(0,0), N(p) \rangle v^2 \right\} + \langle \bar{R}, N(p) \rangle \\ &= \frac{1}{2}\left\{ eu^2 + 2fuv + gv^2 \right\} + \langle \bar{R}, N(p) \rangle \\ &= \frac{1}{2} II_p(w) + \langle \bar{R}, N(p) \rangle, \end{aligned}$$

where $w = X_u(0,0)u + X_v(0,0)v \in T_pS$ and $\displaystyle\lim_{(u,v)\to(0,0)} \frac{\langle \bar{R}, N(p) \rangle}{u^2+v^2} = 0$.

For an elliptic point p, $K(p) > 0$, so the principal curvatures k_1, k_2 have the same sign and thus $II_p(w) = k_n$ has a fixed sign for all $w \in T_pS$ satisfying $|w| = 1$ (by the Euler formula). Therefore, for all (u,v) sufficiently near p, d has the same sign as $II_p(w)$; that is, all such (u,v) belong to the same side of T_pS.

For a hyperbolic point p, $K(p) < 0$, so the principal curvatures k_1, k_2 have the opposite signs, and in each neighborhood of p there exist points (u,v) and $(\bar{u},\bar{v})$ such that $II_p(w/|w|) = k_1$

and $II_p(\bar{w}/|\bar{w}|) = k_2$ have opposite signs (here $w = X_u(0,0)u + X_v(0,0)v$ and $\bar{w} = X_u(0,0)\bar{u} + X_v(0,0)\bar{v}$ are principal directions); such points belong therefore to distinct sides of T_pS.

Proposition Let p be a point of a surface S such that the Gaussian curvature $K(p) \neq 0$, and let V be a connected neighborhood of p where K does not change sign. Then

$$K(p) = \lim_{A \to 0} \frac{A'}{A},$$

where

- A is the area of a a region $B \subset V$ containing p,
- A' is the area of the region $N(B)$ in $\mathbb{S}^2$,

and the limit is taken through a sequence of regions B_n that converges to p, in the sense that any sphere around p conatins all B_n, for n sufficiently large.

Proof Suppose $K > 0$ in V. Let $X : U \to S$ be a parametrization of S at p such that $V \subset X(U)$ and let $B = X(R)$. Since

$$A = \iint_R |X_u \wedge X_v|\, du\, dv, \quad \text{and} \quad A' = \iint_R |N_u \wedge N_v|\, du\, dv = \iint_R K|X_u \wedge X_v|\, du\, dv,$$

we have

$$\lim_{A \to 0} \frac{A'}{A} = \lim_{A \to 0} \frac{A'/A(R)}{A/A(R)} = \frac{\displaystyle\lim_{A(R)\to 0} \frac{1}{A(R)} \iint_R K|X_u \wedge X_v|\, du\, dv}{\displaystyle\lim_{A(R)\to 0} \frac{1}{A(R)} \iint_R |X_u \wedge X_v|\, du\, dv} = \frac{K|X_u \wedge X_v|}{|x_u \wedge X_v|} = K(p).$$

Remark In the proof, we have used the following Theorems from Advanced Calculus.

- **Change of Variables Theorem** Let $F : U \to V$ be a diffeomorphism between open subsets of $U, V \subset \mathbb{R}^n$, let $D^* \subset U$ and $D = F(D^*) \subset V$ be bounded subsets, and let $f : D \to \mathbb{R}$ be a bounded function. Then

$$\begin{aligned} \int_D f(y_1, \ldots, y_n)\, dy_1 \cdots dy_n &= \int_{D^*} f(F(x_1, \ldots, x_n))\, |\det DF(x_1, \ldots, x_n)|\, dx_1 \cdots dx_n \\ &= \int_{D^*} f(F(x_1, \ldots, x_n)) \left| \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)} \right| dx_1 \cdots dx_n. \end{aligned}$$

- **Theorem** Let $f : B_r(p) \to \mathbb{R}$ be a function defined on the ball $B_r(p) \subset \mathbb{R}^n$ of radius r and center p. If f is continuous at p, then

$$\lim_{\rho \to 0} \frac{1}{V(B_\rho(p))} \int_{B_\rho(p)} f(x)\, dx = f(p), \quad \text{where } V(B_\rho(p)) = \int_{B_\rho(p)} dx = \text{the volume of } B_\rho(p).$$

  Proof Since

$$f(p) = f(p) \cdot \frac{1}{V(B_\rho(p))} \int_{B_\rho(p)} dx = \frac{1}{V(B_\rho(p))} \int_{B_\rho(p)} f(p)\, dx \quad \text{and} \quad \lim_{x \to p} f(x) = f(p),$$

we have for any $\varepsilon > 0$, there is a $\delta > 0$ such that if $x \in B_\delta(p)$ then $|f(x) - f(p)| < \varepsilon$, so for all $0 < \rho < \delta$, we have

$$\begin{aligned}
\left|\lim_{\rho\to 0}\frac{1}{V(B_\rho(p))}\int_{B_\rho(p)} f(x)\,dx - f(p)\right| &= \left|\lim_{\rho\to 0}\frac{1}{V(B_\rho(p))}\int_{B_\rho(p)} [f(x) - f(p)]\,dx\right| \\
&\leq \lim_{\rho\to 0}\frac{1}{V(B_\rho(p))}\int_{B_\rho(p)} |f(x) - f(p)|\,dx \\
&< \lim_{\rho\to 0}\frac{1}{V(B_\rho(p))}\int_{B_\rho(p)} \varepsilon\,dx \\
&= \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we have

$$\lim_{\rho\to 0}\frac{1}{V(B_\rho(p))}\int_{B_\rho(p)} f(x)\,dx = f(p).$$