Differential Geometry of Implicit Surfaces in 3-Space - A Primer

John F. Hughes

Department of Computer Science
Brown University
Providence, Rhode Island 02912

CS-03-05
February 2003

Differential Geometry of Implicit Surfaces in 3-Space – A Primer (Tech-Report CS-03-05)

John F. Hughes, Brown University

February 3, 2003

Abstract

Differential geometry is typically presented for parametric surfaces of the form $(u, v) \mapsto (X(u, v), Y(u, v), Z(u, v))$. But in computer graphics, we often encounter smooth surfaces defined *implicitly*, by some equation of the form $G(x, y, z) = 0$; for example, the unit sphere in 3-space can be defined by $x^2 + y^2 + z^2 - 1 = 0$. If (x, y, z) is a point of such an implicit surface, how can we compute the differential geometry of the surface near that point – its surface normal, its curvature in any tangent direction, its principle curvatures, its mean and gaussian curvatures? Fortunately, all of these questions are answered by a somewhat obscure paper by Dombrowski [1]. Unfortunately for many readers, it is in German, *and* it treats the subject in the full generality of n dimensions. In this note, we state some of the results of that paper, and give their proofs. The reader should have some familiarity with differential geometry already; we use the notation of Millman and Parker [2].

1 Introduction

In this paper, we'll discuss the differential geometry of an *implicit surface* in 3-space, i.e., a surface M defined by an equation

$$G(x, y, z) = 0$$

where G is a C^2 function, and $\nabla G \neq 0$ at all points in some open set[1] containing M. The results are quite old; they were developed by Dombrowski [1] in the n-dimensional case almost 40 years ago. We re-present them here because the 3D case is far easier to understand, and the proofs are not hard to follow. Having worked through some of the German in Dombrowski's original paper, we're happy to share our new-found knowledge with others who know no German or who have no need for the n-dimensional results.

Our approach is to determine the *Weingarten map* L at each point p of the surface M. Recall [2] that L is a map from the tangent space to M at p, denoted T_pM, to itself.

[1] If M is compact, then we need only assume that $\nabla G \neq 0$ on the surface M; the continuity of ∇G then ensures the existence of the necessary open set.

For a given vector **v** in this tangent space the value $\mathsf{L}(\mathbf{v})$ tells how the unit surface normal changes as you move through the point p with velocity **v**.

Let's make this more precise with the introduction of some notation. First, let **n** be the vector field defined on M which gives the normal vector at each point of M. Now suppose that

$$\gamma : [-1, 1] \to M$$

is a path on the surface M, and that $\gamma(0) = p$, and that $\gamma'(0) = \mathbf{v}$. Then $\mathsf{L}(\mathbf{v})$ can be defined as

$$\frac{d(\mathbf{n}(\gamma(t)))}{dt}(0),$$

the derivative of $\mathbf{n}(\gamma(t))$ at time $t = 0$ (i.e., as $\gamma(t)$ passes through p (see Figure 1).

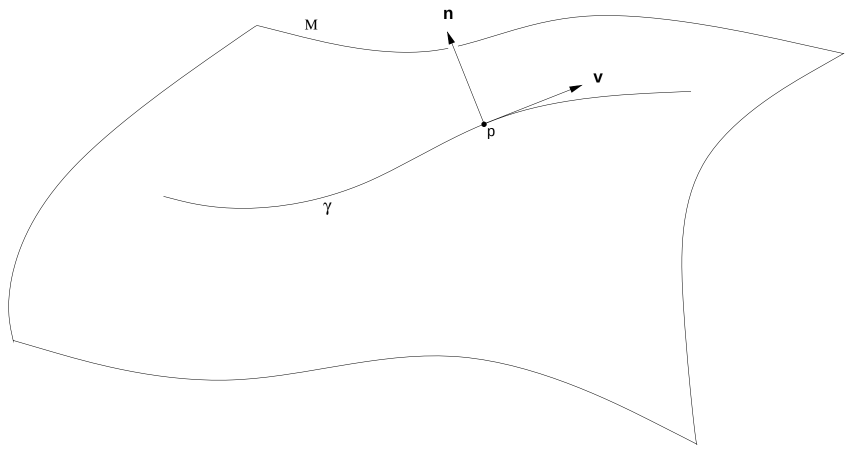

Figure 1: The surface M, and the path γ on M.

There remains a natural question "Is the value of L well-defined? It seems to depend on the choice of γ." We'll show, presently, that the value is independent of γ; alternatively, one can mimic the proof of Proposition 7.1 of Millman and Parker ([2], p. 123) to prove the same thing.

The Weingarten map is a linear map from T_pM to itself. By writing down its matrix with respect to some basis for T_pM, we can compute its eigenvalues (which are the principle curvatures of M at p), its determinant—the product of those eigenvalues—which is the Gaussian curvature at p, and half its trace—the average of those eigenvalues—which is the mean curvature of M at p. In short, the Weingarten map captures the local differential geometry at the point p.

Here's the plan of action, in broad strokes.

- Establish notation, especially for derivatives.
- Since we're working with a function G that's defined on all of R^3, we'll extend the domain of **n** from just M to all of R^3, i.e., we'll find a function **N** defined on all of R^3 which has the property that for all $q \in M$, $\mathbf{N}(q) = \mathbf{n}(q)$. We can then work with this function **N**, which is a map from R^3 to R^3, easily.

- Use the map $\mathbf{N}$ to show the independence of the definition above from the particular choice of a path γ.
- Differentiate the formula for $\mathbf{N}$ in terms of ∇G to get a formula for the transformation L.
- Write down a matrix ℓ for the transformation L with respect to a basis for $T_p M$, so that we can compute curvatures, etc.

1.1 Step-by-step details

Establish notation, especially for derivatives.

Points in R^3, when it's considered merely as the coordinate space in which our surface lives, will be written as ordinary (row) triples, so that we can write $G(x, y, z)$ instead of $G([x, y, z]^t)$.

By contrast, elements of R^3, considered as a *vector space*, will be written as column-vectors in general. Boldface will denote vector-valued functions, so that for example,

$$\mathbf{n} : M \to R^3$$

denotes the function that takes a point of the surface M (i.e., a point (x, y, z) of R^3 such that $G(x, y, z) = 0$) and returns the unit normal vector to the surface M at that point.

Following Millman and Parker, we'll use superscripts to denote indices, so that a typical point of R^k is $(x^1, \ldots, x^k)$. A function from R^k to R^n has component functions, so we'll write

$$f(x^1, \ldots, x^k) = (f^1(x^1, \ldots, x^k), \ldots, f^n(x^1, \ldots, x^k)).$$

Thus a curve in 3-space, being a function from R to R^3, might be written

$$\gamma(t) = (\gamma^1(t), \gamma^2(t), \gamma^3(t)).$$

while a function from R^3 to R could be written

$$f(x^1, x^2, x^3) = f^1(x^1, x^2, x^3).$$

1.1.1 Derivatives

We'll need to talk about the derivative of a function from R^k to R^n for various values of k and n, so let's get that out of the way directly. If

$$f : R^k \to R^n$$

is such a function, then its derivative is a function from R^k to the space of linear transformations from R^k to R^n. In particular, if p is a point of R^k, then $df(p)$ is a linear transformation from R^k to R^n. As such, it can be applied to a vector $\mathbf{v}$; the value is written $df(p)(\mathbf{v})$, and intuitively represents the following: "If I moved a little way from p in the direction $\mathbf{v}$, and applied f to each point along the way, the corresponding values of f would move from $f(p)$ to some nearby point. The net change in that value is $df(p)(\mathbf{v})$." Fortunately for us, the *matrix* of this linear transformation with respect to the standard basis is easy to write down. It's a $n \times k$ matrix, denoted $Df(p)$ and its (i, j) coordinate is simply

$$\frac{\partial f^i(x^1, \ldots, x^k)}{\partial x^j}(p),$$

the j-th partial derivative of the i-th component of f, evaluated at the point p.

It's common to denote $df(p)$ by df_p, so that when you apply the transformation to a vector $\mathbf{v}$, it's written $df_p(\mathbf{v})$, and there's only a single set of parentheses.

There are some notational conveniences. In particular, for a curve

$$\gamma(t) = (\gamma^1(t), \gamma^2(t), \gamma^3(t)),$$

there is a shorthand for $D\gamma(t)$, namely $\gamma'(t)$. And for a function

$$G : R^3 \to R : (x^1, x^2, x^3) \mapsto G(x^1, x^2, x^3),$$

there is also a common name for $DG(x^1, x^2, x^3)$, namely

$$\nabla G(x^1, x^2, x^3) = \left[\frac{\partial G}{\partial x^1}(x^1, x^2, x^3), \frac{\partial G}{\partial x^2}(x^1, x^2, x^3), \frac{\partial G}{\partial x^3}(x^1, x^2, x^3)\right],$$

which is called the *gradient* of G. Note that the gradient is a 1×3 vector (i.e., a *row* vector).

A less-common (but important for us) common name is that of the derivative of the gradient. The gradient is a function from R^3 to R^3, so its derivative is a 3×3 matrix, called the *Hessian*. The Hessian of a function G (i.e., the derivative of the gradient of G) is often denoted $H(G)$ or simply HG. Its matrix form is

$$HG(x^1, x^2, x^3) = \begin{bmatrix} \frac{\partial^2 G}{\partial x^1 \partial x^1}(x^1, x^2, x^3) & \frac{\partial^2 G}{\partial x^1 \partial x^2}(x^1, x^2, x^3) & \frac{\partial^2 G}{\partial x^1 \partial x^3}(x^1, x^2, x^3) \\ \frac{\partial^2 G}{\partial x^2 \partial x^1}(x^1, x^2, x^3) & \frac{\partial^2 G}{\partial x^2 \partial x^2}(x^1, x^2, x^3) & \frac{\partial^2 G}{\partial x^2 \partial x^3}(x^1, x^2, x^3) \\ \frac{\partial^2 G}{\partial x^3 \partial x^1}(x^1, x^2, x^3) & \frac{\partial^2 G}{\partial x^3 \partial x^2}(x^1, x^2, x^3) & \frac{\partial^2 G}{\partial x^3 \partial x^3}(x^1, x^2, x^3) \end{bmatrix}.$$

The *chain rule* tells us that derivatives of composite functions can be computed in terms of the derivatives of the functions themselves. In particular, if $f : R^n \to R^p$, $g : R^k \to R^n$, and $h : R^k \to R^p$ are functions related by

$$h = f \circ g,$$

i.e., for $x \in R^k$, we have

$$h(x) = f(g(x)),$$

then the derivatives of h, f, and g are related. The chain rule (version 1) says (informally) that

$$dh = df \circ dg,$$

which is taken to mean that

$$dh_p(\mathbf{v}) = df_{g(p)}(dg_p(\mathbf{v})).$$

In matrix form, the chain rule is somewhat simpler: since the composition of linear maps amounts to multiplication of their matrices, the chain rule, version 2, says that

$$Dh(p) = Df(g(p)) \cdot Dg(p),$$

where the dot indicates matrix multiplication.

1.2 Extending the normal field to all of R^3

Since we're working with a function G that's defined on all of R^3, we'll extend the domain of $\mathbf{n}$ *from just M to all of R^3, i.e., we'll find a function* $\mathbf{N}$ *defined on all of R^3 which has the property that for all $q \in M$,* $\mathbf{N}(q) = \mathbf{n}(q)$*. We can then work with this function* $\mathbf{N}$*, which is a map from R^3 to R^3, easily.*

It's a standard result from calculus that at a "regular point" of a function $G : R^3 \to R$, i.e., a point p where $\nabla G(p) \neq 0$, the gradient vector is perpendicular to the level surface of G that passes through p. Since our gradient vector is a *row* vector, this means that if $\mathbf{v}$ is a tangent vector to the level surface at p, then $\nabla G(p)\mathbf{v} = 0$, where this represents a matrix multiplication between a 1×3 vector and a 3×1 vector.

Unfortunately, it's rarely the case that the gradient is a *unit* vector. We'll therefore let

$$\alpha : R^3 \to R : p \mapsto \frac{1}{\|\nabla G(p)\|}$$

denote the inverse-length of the gradient, and define

$$\mathbf{N} : R^3 \to R^3 : p \mapsto \alpha(p)\nabla G(p).$$

The function $\mathbf{N}$ is a unit-vector-valued function on all of R^3, and it agrees with the function $\mathbf{n}$ defined on M (the surface defined by $G = 0$). We'll therefore say that $\mathbf{N}$ is an *extension* of $\mathbf{n}$ to R^3. There's one slight problem: the function $\mathbf{N}$ might not be well-defined if there is someplace where the gradient of G is zero, at which point α would be undefined as well. But for all the remainder of this primer, it will only be necessary that α and $\mathbf{N}$ be defined on some open set containing M, and we have already assumed that about the function G. For notational simplicity, we'll continue to treat them as being defined on all of R^3.

1.3 Well-definedness of the Weingarten map

Use the map $\mathbf{N}$ *to show the independence of the definition above from the particular choice of a path γ.*

Suppose that γ is a curve on M, i.e. that

$$\gamma : [-1, 1] \to R^3 : t \mapsto \gamma(t),$$

and

$$G(\gamma(t)) = 0 \text{ for all } t.$$

and that $\gamma(0) = p$, and that $\gamma'(0) = \mathbf{v}$. Then recall that $\mathsf{L}(\mathbf{v})$ is defined as

$$\frac{d(\mathbf{n}(\gamma(t)))}{dt}(0)$$

the derivative of $\mathbf{n}(\gamma(t))$ at time $t = 0$. Now because $\mathbf{n}(q) = \mathbf{N}(q)$ for every point q on our surface, i.e., every point q for which $G(q) = 0$, we rewrite the expression defining $\mathsf{L}(\mathbf{v})$ as

$$\frac{d(\mathbf{N}(\gamma(t)))}{dt}(0),$$

i.e., replace **n** with **N**. But now we can apply the chain rule:

$$\begin{aligned}\frac{d(\mathbf{N}(\gamma(t)))}{dt}(0) &= D\mathbf{N}(\gamma(0)) \cdot D\gamma(0) \\ &= D\mathbf{N}(\gamma(0)) \cdot \gamma'(0) \\ &= D\mathbf{N}(p) \cdot \mathbf{v}.\end{aligned}$$

Because this last expression is independent of the choice of γ, the definition of $\mathsf{L}(\mathbf{v})$ is as well. (Note that again, the dot denotes matrix multiplication; in the last line, this is the product of a 3×3 matrix with a 3×1 vector.)

1.4 The relation between N and ∇G

Differentiate the formula for **N** *in terms of* ∇G *to get a formula for the transformation* L.

Recall that **N** and ∇G are related — one is just a scalar multiple of the other:

$$\mathbf{N}(x^1, x^2, x^3) = \alpha(x^1, x^2, x^3)\nabla G(x^1, x^2, x^3).$$

Using the product rule from multivariable calculus, we can compute the derivative of each side of this equality:

$$\begin{aligned}D\mathbf{N}(x^1, x^2, x^3) &= D\alpha(x^1, x^2, x^3)\nabla G(x^1, x^2, x^3) + \alpha(x^1, x^2, x^3)D(\nabla G)(x^1, x^2, x^3) \\ &= D\alpha(x^1, x^2, x^3)\nabla G(x^1, x^2, x^3) + \alpha(x^1, x^2, x^3)HG(x^1, x^2, x^3)\end{aligned}$$

Now recalling from above that for a tangent vector **v** at the point p of the surface M, we have

$$\mathsf{L}(\mathbf{v}) = D\mathbf{N}(p) \cdot \mathbf{v},$$

we can compute this by substitution:

$$\begin{aligned}\mathsf{L}(\mathbf{v}) &= D\mathbf{N}(p) \cdot \mathbf{v} \\ &= D\alpha(p)\nabla G(p) \cdot \mathbf{v} + \alpha(p)HG(p) \cdot \mathbf{v}\end{aligned}$$

Fortunately for us, this simplifies a great deal: the defining characteristic of a *tangent* vector is that it's orthogonal to the normal vector of the surface; but the gradient of G is just a multiple of this normal, and hence $\nabla G(p) \cdot \mathbf{v}$ must be zero. We therefore conclude that

$$\begin{aligned}\mathsf{L}(\mathbf{v}) &= \alpha(p)HG(p) \cdot \mathbf{v} \\ \mathsf{L}(\mathbf{v}) &= \frac{1}{||\nabla G(p)||}HG(p) \cdot \mathbf{v}.\end{aligned}$$

1.5 A matrix form for the Weingarten map

Write down a matrix ℓ *for the transformation* L *with respect to a basis for* T_pM*, so that we can compute curvatures, etc.*

Now that we have this explicit and simple expression for $\mathsf{L}(\mathbf{v})$, let's see what it looks like in matrix form. Suppose that $\{\mathbf{b}_1, \mathbf{b}_2\}$ is an orthonormal basis for the tangent plane to M at p. The the matrix ℓ of L at p, with respect to this basis, has entries

$$\ell_{ij} = \langle \mathbf{b}_i, \mathsf{L}(\mathbf{b}_j) \rangle := \mathbf{b}_i^t \mathsf{L}(\mathbf{b}_j)$$

Using the formula above, this becomes

$$\begin{aligned} \ell_{ij} &= \mathbf{b}_i^t \mathsf{L}(\mathbf{b}_j) \\ &= \mathbf{b}_i^t \frac{1}{\|\nabla G(p)\|} HG(p)\mathbf{b}_j \\ &= \frac{1}{\|\nabla G(p)\|} \mathbf{b}_i^t HG(p)\mathbf{b}_j \end{aligned}$$

so that the matrix ℓ is simply

$$\ell = \frac{1}{\|\nabla G(p)\|} \begin{bmatrix} \mathbf{b}_1^t HG(p)\mathbf{b}_1 & \mathbf{b}_1^t HG(p)\mathbf{b}_2 \\ \mathbf{b}_2^t HG(p)\mathbf{b}_1 & \mathbf{b}_2^t HG(p)\mathbf{b}_2 \end{bmatrix}.$$

Thus the principle curvatures are the eigenvalues of this matrix; the Gaussian curvature is its determinant, and the mean curvature is half its trace. Not surprisingly, these values are independent of the choice of the orthonormal basis.

Fortunately, the expression above works even if $\mathbf{b}_1$ and $\mathbf{b}_2$ are not unit vectors, except that the result is multiplied by the product of their squared lengths. Indeed, a little linear algebra shows that it works even if $\mathbf{b}_1$ and $\mathbf{b}_2$ are not a orthogonal – the computation carried out with $\mathbf{b}_1$ and $\mathbf{b}_2 + \alpha\mathbf{b}_1$ gives the same result as it does for $\mathbf{b}_1$ and $\mathbf{b}_2$. This allows us to find a form that doesn't depend on any choice of tangent-space basis at all:

Now we can define $\mathbf{c}_i = \nabla G(p) \times \mathbf{e}_i$ for $i = 1, 2, 3$. Carrying out the computation above for $\mathbf{c}_1$ and $\mathbf{c}_2$ yields an expression for $\|\mathbf{c}_1\|^2\|\mathbf{c}_2\|^2\kappa$ (where κ is the Gaussian curvature). Writing $\nabla G(p) = (G_x, G_y, G_z)$, and using double-subscripts to denote second partial derivatives, the expression is:

$$\begin{aligned} \|\mathbf{c}_1\|^2\|\mathbf{c}_2\|^2\kappa \quad = \quad & G_z^2 \left[\begin{vmatrix} G_{yy} & G_{yz} \\ G_{yz} & G_{zz} \end{vmatrix} G_x^2 + \begin{vmatrix} G_{zz} & G_{zx} \\ G_{zx} & G_{xx} \end{vmatrix} G_y^2 + \begin{vmatrix} G_{xx} & G_{xy} \\ G_{xy} & G_{yy} \end{vmatrix} G_z^2 + \right. \\ & \left. -2 \begin{vmatrix} G_{xy} & G_{yz} \\ G_{xz} & G_{zz} \end{vmatrix} G_x G_y - 2 \begin{vmatrix} G_{yz} & G_{zx} \\ G_{yx} & G_{xx} \end{vmatrix} G_y G_z - 2 \begin{vmatrix} G_{xz} & G_{yx} \\ G_{zy} & G_{yy} \end{vmatrix} G_x G_z \right]. \end{aligned}$$

Similar expressions result from the computation applied to $\mathbf{c}_2$ and $\mathbf{c}_3$, and to $\mathbf{c}_3$ and $\mathbf{c}_1$. Letting Q denote the expression in brackets above, and summing the three computations, we get

$$\kappa(\|\mathbf{c}_1\|^2\|\mathbf{c}_2\|^2 + \|\mathbf{c}_2\|^2\|\mathbf{c}_3\|^2\|\mathbf{c}_3\|^2\|\mathbf{c}_1\|^2) = (G_x^2 + G_y^2 + Gz^2)Q.$$

Since $\|\mathbf{c}_1\|^2 = G_y^2 + G_z^2$, and similar formulas hold for $\mathbf{c}_2$ and $\mathbf{c}_3$, it's a short step to a formula given by Dombrowski in a footnote; he says that it "is well known, but absent from most textbooks.” The formula is

$$
\begin{aligned}
\kappa &= (G_x^2+G_y^2+G_z^2)^{-2}\left[\begin{vmatrix} G_{yy} & G_{yz} \\ G_{yz} & G_{zz} \end{vmatrix} G_x^2 + \begin{vmatrix} G_{zz} & G_{zx} \\ G_{zx} & G_{xx} \end{vmatrix} G_y^2 + \begin{vmatrix} G_{xx} & G_{xy} \\ G_{xy} & G_{yy} \end{vmatrix} G_z^2 + \right. \\
&\quad \left. -2\begin{vmatrix} G_{xy} & G_{yz} \\ G_{xz} & G_{zz} \end{vmatrix} G_x G_y - 2\begin{vmatrix} G_{yz} & G_{zx} \\ G_{yx} & G_{xx} \end{vmatrix} G_y G_z - 2\begin{vmatrix} G_{xz} & G_{yx} \\ G_{zy} & G_{yy} \end{vmatrix} G_x G_z\right].
\end{aligned}
$$

This differs slightly from Dombrowski’s formula; his has a typographical error in the fifth term.

References

[1] Peter Dombrowski. Krümmungsgrößen gleichungsdefinierter untermannigfaltigkeiten riemmannscher mannigfaltigkeiten. *Mathematische Nachrichten*, 38(3/4):133–180, 1968.

[2] Richard S. Millman and George D. Parker. *Elements of Differential Geometry*. Prentice-Hall, Englewood Cliffs, NJ, 1977.