DIFFERENTIAL GEOMETRY

IAN MCINTOSH

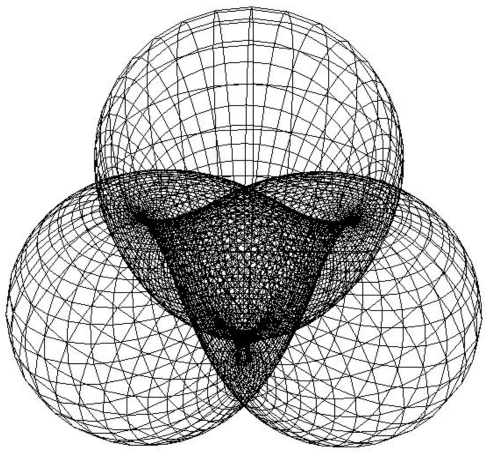

FIGURE 1. The Wente torus: an immersed torus with constant mean curvature.

CONTENTS

Date: January 15, 2019.

1. The Geometry of Space Curves.

Although the main purpose of the module is to understand surfaces using vector calculus, the study of curves in space is a useful warm-up exercise, and curves also play an important role in surface theory. So we will start with the study of curves, and show that we can find some quantities (the **curvature** and **torsion**) which completely describe each curve up to rotations, reflections and translations in Euclidean space.

1.1. Smooth Paths and Regular Paths.

Definition 1.1. *A **smooth path** in space is a smooth vector-valued function $p\colon I \to \mathbb{R}^3$, where $I \subset \mathbb{R}$ is an interval of positive length. We will call its image in space its **track**. We say the function p **parameterises** the track. We can write this parameterisation in standard coordinates as $p(t) = (x(t), y(t), z(t))$.*

(i) By a *smooth* function we mean one we can differentiate as often as we like. To be precise, one whose r-th derivative exists for all $r \in \mathbb{N}$; synonyms are **infinitely differentiable,** or **C^∞-smooth.**

(ii) The interval I need not be open. If not, then the understanding is that $p(t)$ is defined and smooth on some *open* interval containing I. The interval I need not be finite.

For $p(t) = (x(t), y(t), z(t))$ the derivative will be written as

$$\frac{dp}{dt} = \Big(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\Big), \qquad \text{or} \qquad p'(t) = \big(x'(t),\, y'(t),\, z'(t)\big).$$

This is usually visualised, and referred to, as the **tangent vector.** It is also known as the **velocity vector,** thinking of $p(t)$ as the position of a moving particle at time t.

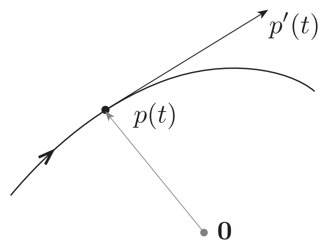

Figure 2. Position vector and tangent/velocity vector.

Example 1.1 (Helix). Define

$$p(t) = (a\cos t,\, a\sin t,\, bt), \quad \text{where } a, b > 0. \tag{1.1}$$

This path describes a **right-handed helix** on the cylinder $x^2 + y^2 = a^2$. The number $2\pi b$ is called the **pitch** of the helix. The tangent vector is:

$$p'(t) = (-a\sin t,\, a\cos t,\, b),$$

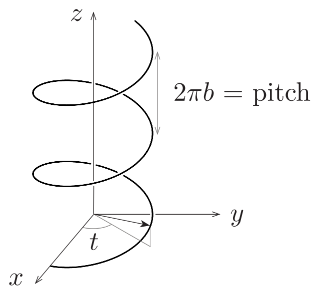

FIGURE 3. The right-handed helix.

which has constant length $\sqrt{a^2+b^2}$. Notice that the tangents describe a circle lying in a horizontal plane.[1]

It is very important to realise that smoothness is a property of the parametrisation, and not necessarily a property of its track. In particular, the tangent line to the track need not be well-defined for a smooth path: this can only happen when $p'(t)$ vanishes at the corresponding point. Also, the track of a smooth path can be self-intersecting. The following examples illustrate these points.

Example 1.2. For any $k \in \mathbb{R}$ define a smooth path $p_k(t)$ as follows[2]:

$$p_k(t) = (t^3 + kt, t^2 + k, 0), \quad t \in \mathbb{R}. \tag{1.2}$$

Since $x(t) = ty(t)$ the track of $p_k(t)$ can be described as

$$\{(x, y, 0) : x^2 = y^2(y - k)\}.$$

Each $p_k(t)$ is clearly a smooth planar path (lying in the plane $z = 0$), with tangent vectors

$$p_k'(t) = (3t^2 + k, 2t, 0).$$

Since this is a planar curve we can assign to each point $p_k(t)$ the slope of the tangent line, which is $2t/(3t^2 + k)$.

This example demonstrates three different types of behaviour, dependent on the sign of k. We can examine these by considering $k = -1, 0, 1$.

For $k = 1$, $p_1(t)$ parameterises an **embedded cubic**. Its path is one-to-one (injective) and the slope of its tangent line at $p_1(t)$ is a smooth function.

For $k = 0$ the path $p_0(t)$ is one-to-one but the slope of its tangent line is $2/3t$, which has a discontinuity at $p_0(0) = (0, 0)$: this gives the curve a **cusp**, and it is known as a **cuspidal cubic**.

For $k = -1$ there is a point of self-intersection, since $p_{-1}(1) = p_{-1}(-1) = (0, 0)$. At this point the tangent line is not well-defined, since the two tangent vectors $(2, \pm 2, 0)$ are not

[1]In general the path $p'(t)$ is called the **hodograph** of $p(t)$.

[2]These are historically known, perhaps confusingly for our modern terminology, as **Newton's diverging parabolas**. Newton classified cubic equations of the form $y^2 = ax^3 + bx^2 + cx + d$ by distinguishing them into four classes, allowing linear transformations of x: the diverging parabolas form the third class.

parallel. This family of curves are known as **nodal cubics**: each has a **node** or **ordinary double point** at the origin. But the tangent vector of the path is well-defined, since these tangents occur at different times $t = \pm 1$.

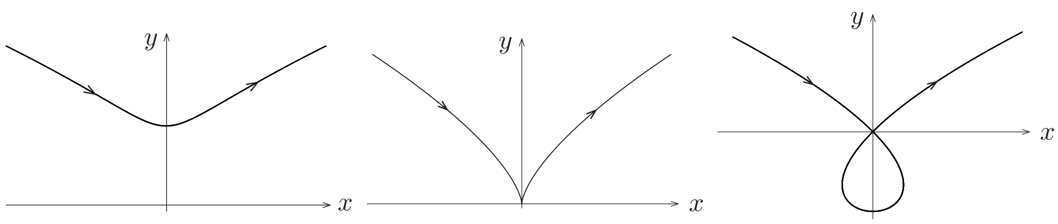

FIGURE 4. From left to right, embedded, cuspidal and nodal cubics.

These examples make clear the need to distinguish those smooth paths whose tangent vectors do not vanish at any point.

Definition 1.2. *Let $p : I \to \mathbb{R}^3$ be a smooth path. We say a parameter value $t_0 \in I$ is:*

(i) ***singular*** *or* ***critical*** *if $p'(t_0) = \mathbf{0}$, in which case $\mathbf{p} = p(t_0)$ is a* ***singular*** *or* ***critical point,*** *and;*
(ii) ***regular*** *if it is not singular, in which case $\mathbf{p}$ is a* ***regular point****.*

A smooth path is ***regular*** *if all its points are regular.*

In particular, by this definition whenever $p(t)$ has a point $\mathbf{p}$ of self-intersection, this is only a regular point if **each of** the parameter values t_0 for which $p(t_0) = \mathbf{p}$ is regular. It is still possible for regular curves to have multiple tangent lines at one point, but regularity only allows this to happen at point of self-intersection. For example, it can be shown that the cuspidal cubic in Example 1.2 does not admit any regular parametrisation about its cusp, but the nodal cubic is a regular path.

1.2. **Smooth Curves and Arc Length Parametrisation.** Any smooth path $p(t)$ has many different reparametrisations, defined as follows.

Definition 1.3. *If $p : I \to \mathbb{R}^3$ and $q : J \to \mathbb{R}^3$ are smooth paths for which $q(u) = p(t(u))$, $u \in J$, for some smooth change of variable function $t(u) : J \to I$ with $t'(u) \neq 0$ then we say that $q(u)$ is a* ***reparametrisation*** *of $p(t)$. When $t'(u) > 0$ (resp. $t'(u) < 0$) we say u is an* ***orientation preserving*** *(resp.* ***orientation reversing****) reparametrisation.*

Remark 1.1. We want the reparametrisation function $u(t)$ to be a smooth bijection $u : I \to J$, with smooth inverse. This will follow if its derivative is not allowed to vanish, for then it is either a strictly increasing function (when $u'(t) > 0$) or strictly decreasing function (when $u'(t) < 0$). In particular, a reparametrisation cannot "double back" on itself.

Both smooth paths p, q will have the same track. By the chain rule

$$\frac{dq}{du} = \frac{dt}{du}\frac{dp}{dt},$$

so that the condition $t'(u) \neq 0$ is necessary and sufficient to allow the parametrisation to be inverted, i.e., we can write $p(t) = q(u(t))$. We can see from this that an orientation preserving reparametrisation moves us in the same direction along the curve, whereas an orientation reversing reparametrisation reverses the direction of travel. Notice that $p(t)$ will be a regular path if and only if $q(u)$ is a regular path, since $|q'(u)| = |t'(u)||p'(t)|$.

Our definition for a smooth curve intends to get at the geometric object which remains unchanged by reparametrisation: this is slightly more subtle than than the idea of the track.

Definition 1.4. *A* ***smooth curve*** *(resp.* ***oriented smooth curve****) C is an equivalence class of smooth paths any two of which are equal after reparametrisation (resp. orientation preserving reparametrisation). Paths in the equivalence class are said to* ***parameterise*** *C. A smooth curve is* ***regular*** *if one, and hence all, its parametrisations are regular. If C is an oriented curve with parametrisation $p(t)$ then its* ***opposite*** *is the curve $-C$ with parametrisation $p(-t)$.*

Remark 1.2. What is the difference between a curve and its track? Clearly every curve determines a single track, but there are many examples of tracks which come from two different curves. One of the simplest examples is as follows. Let

$$p : [0, 2\pi) \to \mathbb{R}^3, \qquad p(u) = (\cos(u), \sin(u), 0),$$
$$q : [0, 2\pi) \to \mathbb{R}^3, \qquad q(t) = (\cos(2t), \sin(2t), 0).$$

Both curves have the unit circle in the plane $z = 0$ as their track. The curve p winds once around this circle, whereas the curve q winds twice around. There is no reparametrisation $u : [0, 2\pi) \to [0, 2\pi)$ for which $q(t) = p(u(t))$, because u is invertible (by the previous remark) and p is invertible, but q is not injective. [You might think that $u(t) = 2t$ should work, but it doesn't map $[0, 2\pi)$ to itself.]

There are a particular collection of (re)parametrisations which are geometrically meaningful, corresponding to the arc length along the curve from a chosen initial point and in a chosen direction along the path. Recall from Vector Calculus the definition of arc length.

Definition 1.5. *Let $p : I \to \mathbb{R}^3$ be a smooth path. The* ***(oriented) arc length*** *from $p(t_0)$ to $p(t)$ is defined to be*

$$s(t, t_0) = \int_{t_0}^{t} |p'(u)| \, du, \qquad t_0, t \in I. \tag{1.3}$$

Note that $s(t, t_0) < 0$ if $t < t_0$. By the Fundamental Theorem of Calculus, this is a differentiable function of t, with derivative

$$\frac{ds}{dt} = |p'(t)|. \tag{1.4}$$

Definition 1.6. *We say $p(t)$ is* ***arc length parameterised*** *(and t is an* ***arc length parametrisation****) whenever $s(t, t_0) = t - t_0$, or equivalently, when $|p'(t)| = 1$. For the latter reason, this is also referred to as a* ***unit speed parametrisation****, and $T(t) = p'(t)$ is called the* ***unit tangent vector field*** *along $p(t)$.*

It seems geometrically obvious that the track of any smooth curve should admit an arc length parametrisation. However, we can only ensure that this is smooth when the arc length function is smooth, and therefore we must avoid paths which have $|p'(t)| = 0$ at some point (because $f(x) = |x|$ is smooth everywhere except at $x = 0$, where it is not even differentiable).

Lemma 1.7. *Let $p(t)$ be a regular path. Then p can be reparameterised by arc length $s(t)$ to obtain a regular path $\tilde{p}(s)$ for which*

$$\frac{d\tilde{p}}{ds} = \frac{p'(t)}{|p'(t)|}.$$

Proof. The expression above is a consequence of the chain rule mentioned above, since

$$\frac{dt}{ds} = 1 \bigg/ \frac{ds}{dt} = \frac{1}{|p'(t)|},$$

and we have just seen that this is smooth whenever $|p'(t)|$ does not vanish, i.e., at all regular points. □

Example 1.3. Consider the helix $p(t) = (a \cos t, a \sin t, bt)$ with $a, b \in \mathbb{R}$. Then

$$p'(t) = (-a \sin t, a \cos t, b), \qquad |p'(t)|^2 = a^2 + b^2.$$

The arc length from $p(0)$ to $p(t)$ is therefore

$$s(t, 0) = \int_0^t \sqrt{a^2 + b^2}\, du = ct, \qquad \text{where } c = \sqrt{a^2 + b^2}.$$

It follows that $t(s) = s/c$ and an arc length reparametrisation is therefore

$$\tilde{p}(s) = \big(a \cos(s/c),\, a \sin(s/c),\, bs/c\big).$$

We will reserve the symbol s to denote arc length parametrisations. Such a parametrisation is not unique for a given regular curve C, but can only be changed by translation of s or a change of orientation.

Theorem 1.8. *Suppose $p(s)$ and $\tilde{p}(\sigma)$ are both arc length parametrisations of the same curve C. Then*

$$\sigma = \pm s + c, \quad \textit{from some } c \in \mathbb{R}.$$

Proof. Since $\tilde{p}(\sigma)$ and $p(s)$ are unit speed parameterisations,

$$|\tilde{p}'(\sigma)| = 1, \quad |p'(s)| = 1.$$

Since these parameterise the same curve, we can write $p(s) = \tilde{p}(\sigma(s))$. Using the chain rule gives

$$1 = |p'(s)| = |\tilde{p}'(\sigma)\sigma'(s)| = |\tilde{p}'(\sigma)||\sigma'(s)| = |\sigma'(s)|.$$

Thus $\sigma'(s) = \pm 1$, whence $\sigma(s) = \pm s + c$ for some constant of integration c. □

It follows that a smooth regular oriented curve has precisely one unit speed tangent vector field along it.

1.3. **Curvature and torsion.** Intuitively, the curvature of a space curve C is its rate of change of direction. Since rate of change is measured by the derivative, and direction is measure by the unit tangent vector, we measure curvature by taking the derivative of the unit tangent vector, i.e., the second derivative of the unit speed parameterised path.

Definition 1.9. *Let C be a regular smooth curve, and $p(s)$ an arc length parametrisation with unit speed tangent vector field $T(s) = p'(s)$. We call*

$$\boldsymbol{k}(s) = T'(s) = p''(s)$$

*the **curvature vector** of $p(s)$ and*

$$\kappa(s) = |\boldsymbol{k}(s)| = |p''(s)| \geq 0$$

*the **curvature.***

Note that

$$\boldsymbol{k} \bullet T = p'' \bullet p' = \tfrac{1}{2}(p' \bullet p')' = 0.$$

Since the length of T is constant, as s varies only the direction of T changes. Thus the

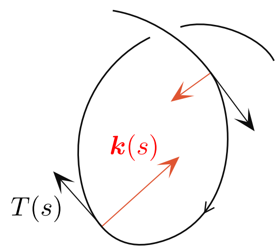

FIGURE 5. The curvature and unit tangent vectors.

magnitude of $\boldsymbol{k} = T'$ does indeed measure the rate of change of the curve's direction.

Example 1.4. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, with $\mathbf{b}$ of unit length. Then the path $p(s) = \mathbf{a} + s\mathbf{b}$ is an arc length parametrisation of the straight line through $\mathbf{a}$ in direction $\mathbf{b}$. Clearly all straight lines may be parameterised in this way. Then:

$$p'(s) = \mathbf{b}, \qquad p''(s) = \mathbf{0},$$

so $\kappa(s) = 0$ for all s, as we would expect. Conversely, if the curvature is identically zero then $p''(s) = \mathbf{0}$ for all s, and two integrations yield $p(s) = \mathbf{a} + s\mathbf{b}$. Hence a curve has zero curvature if and only if it is a straight line.

Provided $\kappa(s) \neq 0$ the two vectors $T(s), \boldsymbol{k}(s)$ span a plane, called the **osculating plane at the point** $p(s)$ on the curve. The **torsion** measures the extent to which the curve twists out of this plane. To define torsion we proceed as follows.

Definition 1.10. *Let $p(s)$ be an arc length parameterised curve. We say a point $p(s_0)$ is a **point of inflection** when $\kappa(s_0) = 0$. When $p(s)$ has no points of infection we define:*

- *the* ***principal normal vector***
$$N(s) = \frac{\boldsymbol{k}(s)}{|\boldsymbol{k}(s)|} = \frac{\boldsymbol{k}(s)}{\kappa(s)} = \frac{T'(s)}{\kappa(s)} = \frac{p''(s)}{|p''(s)|},$$
- *the* ***binormal vector***
$$B(s) = T(s) \times N(s). \tag{1.5}$$

At each point $p(s)$ on the curve the vectors $T(s), N(s), B(s)$ form a positively oriented orthonormal basis for $\mathbb{R}^3$, called the ***Frenet frame.***

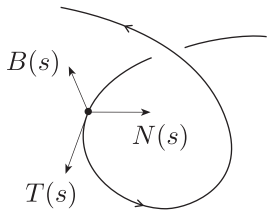

FIGURE 6. The Frenet frame.

Differentiating $B \bullet B = 1$ and (1.5) gives two identities:
$$B' \bullet B = 0, \quad B' = (T' \times N) + (T \times N') = T \times N', \tag{1.6}$$
using the fact that T' and N are parallel. Hence B' is orthogonal to T and B, which means it is parallel to N.

Definition 1.11. *The* ***torsion*** *$\tau(s)$ is the function for which $B'(s) = \tau(s)N(s)$. Equally, $\tau = B' \bullet N$.*

Recall that $\kappa(s) = |p''(s)|$. Combining the definition of τ with (1.6), we have
$$\begin{aligned} \tau &= B' \bullet N = (T \times N') \bullet N = [T, N', N], \quad \text{the triple scalar product,} \\ &= -[T, N, N'], \end{aligned}$$
using the antisymmetric property of the triple scalar product. Now
$$T = p', \qquad N = \frac{1}{\kappa}T' = \frac{1}{\kappa}p'', \qquad N' = \frac{1}{\kappa^2}(\kappa p''' - p''\kappa').$$
Therefore in arc length parametrisation
$$\tau(s) = -\frac{1}{\kappa(s)^2}\,[p'(s), p''(s), p'''(s)] = -\frac{[p'(s), p''(s), p'''(s)]}{|p''(s)|^2}. \tag{1.7}$$

Remark 1.3. We will be using the triple scalar product quite a lot. Notice our notation: for any three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ we set
$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = (\mathbf{a} \times \mathbf{b}) \bullet \mathbf{c} = \det(\mathbf{a}\ \mathbf{b}\ \mathbf{c}),$$
where the last expression is the determinant of the 3×3 matrix with columns $\mathbf{a}, \mathbf{b}, \mathbf{c}$. Recall from Linear Algebra that this makes $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$ linear in each slot, and it is totally

skew symmetric (swapping any two vectors changes the sign). Two important facts follow from this: (i) it is zero when any two entries are equal, (ii) its value is unchanged by any cyclic permutation of entries. Its geometric meaning is linked to this. When $\mathbf{a}, \mathbf{b}, \mathbf{c}$ form a right-handed basis (not necessarily orthogonal) $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$ is the volume of the parallelopiped generated by these three vectors.

Example 1.5 (Curvature and torsion of helices). From our previous work with helices, a unit speed parametrisation is

$$p(s) = (a\cos(s/c), a\sin(s/c), bs/c), \quad \text{where } c = \sqrt{a^2 + b^2}$$

Hence

$$\begin{aligned} p'(s) &= (-\frac{a}{c}\sin(s/c), \frac{a}{c}\cos(s/c), \frac{b}{c}), \\ p''(s) &= -\frac{a}{c^2}(\cos(s/c), \sin(a/c), 0), \\ p'''(s) &= \frac{a}{c^3}(\sin(s/c), -\cos(a/c), 0), \end{aligned}$$

and

$$[p', p'', p'''](s) = \frac{ba^2}{c^6}(\cos^2(s/c) + \sin^2(s/c)) = \frac{ba^2}{c^6}.$$

Hence

$$\begin{aligned} \boldsymbol{k}(s) &= p''(s) = \left(-\frac{a}{c^2}\cos\left(\frac{s}{c}\right), -\frac{a}{c^2}\sin\left(\frac{s}{c}\right), 0\right), \\ \kappa(s) &= \frac{|a|}{c^2} = \frac{|a|}{a^2 + b^2}, \\ \tau(s) &= -\frac{c^4}{a^2}\frac{a^2 b}{c^6} = -\frac{b}{c^2}. \end{aligned}$$

So a helix has constant curvature and constant torsion (in fact we will see later that a curve with such properties must be a helix). Notice that the torsion is negative since this helix is right-handed. Taking $b = 0$ gives a circle of radius $a > 0$, with $\kappa(s) = 1/a$ and torsion $\tau = 0$. Thus the curvature of a circle is the reciprocal of its radius.

1.3.1. *Planar curves.* We say a curve $p(s)$ is **planar** if it lies in a fixed plane in $\mathbb{R}^3$. Thus $p(s)$ is planar when there is a point $\mathbf{p} \in \mathbb{R}^3$ a non-zero vector $\mathbf{n} \in \mathbb{R}^3$ for which

$$(p(s) - \mathbf{p}) \bullet \mathbf{n} = 0,$$

i.e., $p(s)$ lies in the plane through $\mathbf{p}$ with normal $\mathbf{n}$. By differentiating this equation twice we see that for a planar curve without points of inflection

$$p'(s) \bullet \mathbf{n} = 0 = p''(s) \bullet \mathbf{n}, \quad \text{i.e., } T \bullet \mathbf{n} = 0 = N \bullet \mathbf{n}.$$

Thus $p(s)$ is planar if and only if the osculating plane spanned by $T(s)$ and $N(s)$ is constant. This is equivalent to saying that the binormal B is constant. Conversely, suppose the binormal B is constant, then for any choice of point $\mathbf{p} = p(s_0)$ on the curve

$$((p(s) - \mathbf{p}) \bullet B)' = p'(s) \bullet B = T \bullet B = 0,$$

thus $(p(s)-\mathbf{p})\bullet B)$ is a constant function of s. But it vanishes at $s=s_0$, so $(p(s)-\mathbf{p})\bullet B=0$, i.e., $p(s)$ is a planar curve. We have therefore proved:

Theorem 1.12. *A curve without inflections is planar if and only if it has constant binormal vector, that is, if and only if its torsion τ is identically zero all along the curve.*

A plane $\Pi\subset\mathbb{R}^3$ is **oriented** by making a choice of (constant) unit normal vector ξ. For a smooth curve in the oriented plane (Π,ξ) we can talk about the **signed** curvature. Given $T(s)=p'(s)$ and ξ, the vector $\nu(s)=\xi\times T(s)$ is a unit vector normal to both $T(s)$ and ξ, chosen so that (T,ν,ξ) is positively oriented. It follows that $N=\pm\nu(s)$. The **signed curvature** is defined to be

$$\kappa_s(s)=\boldsymbol{k}(s)\bullet\nu(s)=\boldsymbol{k}(s)\bullet(\xi\times T(s))=[\boldsymbol{k}(s),\xi,T(s)]. \tag{1.8}$$

Thus $\kappa_s=\pm\kappa$, with the sign determined by $N\bullet\nu$. Notice that $(T,N,B)=(T,\nu,\xi)$ precisely when $N=\nu$.

1.3.2. *Curvature and torsion in any regular parametrisation.* Since it is not usually practical to find the arc length parametrisation of a curve, we now derive the expressions for curvature and torsion in any regular parametrisation.

Theorem 1.13. *If $p(t)$ is a regular parametrisation of a space curve then the curvature vector and curvature are given by:*

$$\boxed{\boldsymbol{k}=\frac{1}{|p'|^2}\left(p''-\frac{p''\bullet p'}{|p'|^2}p'\right)}\qquad\qquad \boxed{\kappa=\frac{|p'\times p''|}{|p'|^3}} \tag{1.9}$$

Further, when $p(t)$ has no inflection points it has torsion

$$\boxed{\tau=-\frac{1}{\kappa^2}\frac{[p',p'',p''']}{|p'|^6}=-\frac{[p',p'',p''']}{|p'\times p''|^2}} \tag{1.10}$$

Proof. Define $\tilde{p}(s)=p(t(s))$, so that $d\tilde{p}/ds=p'(t(s))/|p'(t(s))|$. By the chain rule we have

$$\frac{d}{ds}=\frac{dt}{ds}\frac{d}{dt}=\frac{1}{ds/dt}\frac{d}{dt}=\frac{1}{|p'(t)|}\frac{d}{dt}.$$

Combining this with the definition of $\boldsymbol{k}$, we get

$$\boldsymbol{k}=\frac{d^2\tilde{p}}{ds^2}=\frac{1}{|p'(t)|}\frac{d}{dt}\left(\frac{p'(t)}{|p'(t)|}\right).$$

Now note that

$$\frac{d}{dt}|p'(t)|=\frac{d}{dt}\big(p'(t)\bullet p'(t)\big)^{1/2}=\frac{2p'(t)\bullet p''(t)}{2(p'(t)\bullet p'(t))^{1/2}}=\frac{p'(t)\bullet p''(t)}{|p'(t)|}.$$

Hence, by the quotient rule,

$$\begin{aligned}\boldsymbol{k} &= \frac{1}{|p'|}\left(\frac{1}{|p'|}p'' - \frac{p'\bullet p''}{|p'|^3}p'\right)\\ &= \frac{1}{|p'|^2}\left(p'' - \frac{p'\bullet p''}{|p'|^2}p'\right)\\ &= \frac{|p'|^2p'' - (p'\bullet p'')p'}{|p'|^4}.\end{aligned}$$

Then

$$\begin{aligned}\kappa^2 = |\boldsymbol{k}|^2 &= \left(\frac{|p'|^2p'' - (p'\bullet p'')p'}{|p'|^4}\right)\bullet\left(\frac{|p'|^2p'' - (p'\bullet p'')p'}{|p'|^4}\right)\\ &= \frac{|p'|^4|p''|^2 - 2|p'|^2(p'\bullet p'')^2 + (p'\bullet p'')^2|p'|^2}{|p'|^8}\\ &= \frac{|p'|^2\,|p''|^2 - (p'\bullet p'')^2}{|p'|^6} = \frac{|p'\times p''|^2}{|p'|^6},\end{aligned}$$

by Lagrange's formula (1.11) below. Now from (1.7)

$$\begin{aligned}\tau(t(s)) &= \frac{-1}{\kappa(s)^2}\left[\tilde{p}'(s), \tilde{p}''(s), \tilde{p}'''(s)\right]\\ &= \frac{-1}{\kappa^2}\left[\frac{p'}{|p'|}, \frac{p''}{|p'|^2} - \frac{p'\bullet p''}{|p'|^4}p', \tilde{p}'''(s)\right]\\ &= \left[\frac{p'}{|p'|}, \frac{p''}{|p'|^2}, \tilde{p}'''(s)\right],\end{aligned}$$

where we have used the expression for $\boldsymbol{k} = \tilde{p}''(s)$ above and the fact that triple scalar product is linear in each term and vanishes when two vectors are parallel. Now $\tilde{p}'''(s)$ is of the form $ap'(t) + bp''(t) + cp'''(t)$, and so for the same reason only the term $cp'''(t)$ is required. Now it is easy to see that

$$cp''' = \frac{1}{|p'|^2}\frac{d}{ds}p'' = \frac{1}{|p'|^2}\left(p'''\frac{dt}{ds}\right) = \frac{p'''}{|p'|^3}.$$

□

Remark 1.4. For any $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ we have:

$$|\mathbf{a}\times\mathbf{b}|^2 = |\mathbf{a}|^2|\mathbf{b}|^2\sin^2\theta = |\mathbf{a}|^2|\mathbf{b}|^2(1-\cos^2\theta) = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a}\bullet\mathbf{b})^2. \tag{1.11}$$

This is sometimes called **Lagrange's formula,** and can be used to evaluate κ without having to compute the vector product.

Example 1.6. Consider the path $p(t) = (t, \frac{1}{\sqrt{2}}t^2, \frac{1}{3}t^3)$, $t \in \mathbb{R}$, which is a type of **Veronese curve**. We have

$$p'(t) = (1, \sqrt{2}t, t^2), \quad |p'(t)| = \sqrt{1 + 2t^2 + t^4} = 1 + t^2.$$

Therefore this is not in arc length parameterisation, but it is regular since $|p'(t)| \neq 0$ for all t. Therefore to calculate its curvature and torsion we should use Theorem 1.13. For this we need

$$p''(t) = (0, \sqrt{2}, 2t), \quad p'''(t) = (0, 0, 2).$$

To get $\kappa(t)$ we first calculate, using Lagrange's formula,

$$\begin{aligned}|p' \times p''|^2 &= |p'|^2|p''|^2 - (p' \bullet p'')^2 \\ &= (1+t^2)^2(2+4t^2) - (2t+2t^3)^2 \\ &= 2(1+t^2)^2[1+2t^2-2t^2] = 2(1+t^2)^2\end{aligned}$$

Therefore

$$\kappa(t) = \frac{\sqrt{2}}{(1+t^2)^2}.$$

For the torsion, we compute first

$$[p', p'', p'''] = \begin{vmatrix} 1 & \sqrt{2}t & t^2 \\ 0 & \sqrt{2} & 2t \\ 0 & 0 & 2 \end{vmatrix} = 2\sqrt{2}.$$

Therefore

$$\tau(t) = -\frac{\sqrt{2}}{(1+t^2)^2}.$$

So this curve has $\tau = -\kappa$.

Remark 1.5. It is interesting to note that there is a **vector** triple product expression for the curvature vector:

$$\boldsymbol{k} = \frac{1}{|p'|^4}(p' \times p'') \times p'.$$

This comes from the general expression

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{a} = |\mathbf{a}|^2\mathbf{b} - (\mathbf{a} \bullet \mathbf{b})\mathbf{a}, \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^3.$$

This is probably not so useful for calculation as the formula above, but it is more memorable, and the expression for κ follows very easily from it. Since $p' \times p''$ and p' are orthogonal, $|(p' \times p'') \times p'|$ is the area of a rectangle with sides $p' \times p''$ and p', i.e., its area is $|p' \times p''||p'|$. Thus $|\boldsymbol{k}| = |p' \times p''|/|p'|^3$.

1.4. **Congruence and Frenet formulas.** Our aim now is to show that the curvature and torsion of a curve determine it completely up to rotations and translations. These transformations generate the group of **orientation preserving Euclidean motions**. We start with a short discussion of Euclidean motions.

Definition 1.14. *A transformation $E : \mathbb{R}^3 \to \mathbb{R}^3$ is called a* ***Euclidean motion*** *(or* ***isometry****) when it preserves distances and angles, i.e.,*

$$(E\mathbf{p}_2 - E\mathbf{p}_1) \bullet (E\mathbf{q}_2 - E\mathbf{q}_1) = (\mathbf{p}_2 - \mathbf{p}_1) \bullet (\mathbf{q}_2 - \mathbf{q}_1), \quad \forall\ \mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2 \in \mathbb{R}^3. \tag{1.12}$$

It can be shown that all Euclidean motions must have the form $E\mathbf{v} = L\mathbf{v} + \mathbf{c}$, where $L : \mathbb{R}^3 \to \mathbb{R}^3$ is an invertible linear map, and $\mathbf{c}$ is a constant vector. It follows that the set of Euclidean motions is a group (under composition of transformations). The condition (1.12) amounts to the condition

$$L\mathbf{v} \bullet L\mathbf{w} = \mathbf{v} \bullet \mathbf{w}, \quad \forall\ \mathbf{v}, \mathbf{w} \in \mathbb{R}^3.$$

This implies that L must represented by an **orthogonal matrix**, i.e, $LL^t = I_3$, the 3×3 identity matrix. It follows from this that $\det(L)^2 = 1$, since $\det(L^t) = \det(L)$, and therefore $\det(L) = \pm 1$. L is a rotation about the origin when $\det(L) = 1$. When $\det(L) = -1$ it is a reflection through some plane containing the origin combined with a rotation about the axis normal to that plane. Thus every Euclidean motion is the composition of a rotation or reflection with a translation.

Definition 1.15. *We say the Euclidean motion $E\mathbf{v} = L\mathbf{v} + \mathbf{c}$ is* ***orientation preserving*** *when $\det(L) = 1$, equally, when*

$$[L\mathbf{e}_1, L\mathbf{e}_2, L\mathbf{e}_3] = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$$

where $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is the standard orthonormal basis for $\mathbb{R}^3$. Otherwise it is called ***orientation reversing****.*

Definition 1.16. *We will say two space curves $p, q : [a, b] \to \mathbb{R}^3$ are* ***congruent*** *when there is a Euclidean motion E for which $(E \circ p)(t) = q(t)$ for all $t \in [a, b]$, and we will say they are* ***properly congruent*** *when E is also orientation preserving.*

Remark 1.6. The set of all 3×3 orthogonal matrices is denoted by $O(3)$. This forms a group under matrix multiplication, which is called the **orthogonal group**. The subset of those matrices which also have determinant 1 is denoted by $SO(3)$. It is a subgroup, called the **special orthogonal** (sub)group. Thus $SO(3)$ represents the group of all rotations of $\mathbb{R}^3$. Notice that the set $\{L \in O(3) : \det(L) = -1\}$ does not form a subgroup of $O(3)$ (since it doesn't contain the identity transformation), but that the product of any two of these transformations is a rotation.

These orthogonal transformations $L \in O(3)$ can be understood by considering their eigenvalues and eigenvectors. We need two facts. First, the eigenvalues of L are roots of the characteristic polynomial $\det(L - \lambda I)$, which is a cubic polynomial with real coefficients when $L \in O(3)$. Therefore there is always one real eigenvalue. The other two eigenvalues either have to be both real, or complex conjugates. Second, the eigenvalues satisfy $|\lambda| = 1$. Therefore each real eigenvalue is ± 1. So we can can always write the eigenvalues of L as $\pm 1, e^{i\theta}, e^{-i\theta}$ for some $0 \le \theta \le 2\pi$. The case $+1$ gives a rotation through angle θ about the axis which is the eigenline for eigenvalue $+1$. The case -1 gives a reflection through the plane normal to the eigenline for eigenvalue -1, combined with a rotation through angle θ about that axis.

The **Frenet** (sometimes called **Frenet-Serret**) formulas tell us how to rebuild a curve from its curvature and torsion, up to rotations and translations in space. They show how the Frenet frame can be obtained from a system of linear differential equations whose data is the curvature and torsion of the curve. These are derived as follows.

Let $p(s)$ be a unit speed path. By definition of the principal normal and binormal

$$T' = \kappa N, \qquad B' = \tau N.$$

Since $|N| = 1$ we have $N' \bullet N = 0$, so

$$N' = \alpha T + \beta B,$$

for some scalar functions α, β. Since $N \bullet T = 0 = N \bullet B$ it follows that

$$\alpha = N' \bullet T = -N \bullet T' = -\kappa, \qquad \beta = N' \bullet B = -N \bullet B' = -\tau,$$

and we thereby obtain the Frenet formulas.

$$T' = \kappa N, \qquad N' = -\kappa T - \tau B \qquad B' = \tau N. \tag{1.13}$$

Remark 1.7. Another way to derive the expression for N' is to use the fact that $N = B \times T$ (since T, N, B is a right handed orthonormal frame). Differentiating this gives

$$\begin{aligned} N' =& B' \times T + B \times T' \\ =& \tau N \times T + \kappa B \times N \\ =& -\tau B - \kappa T. \end{aligned}$$

It is useful to write (1.13) in matrix form

$$\begin{pmatrix} T' & N' & B' \end{pmatrix} = \begin{pmatrix} T & N & B \end{pmatrix} \begin{pmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix}. \tag{1.14}$$

We can simplify this notationally into $F' = FA$ where F is the matrix representing the Frenet frame and A is the skew-symmetric matrix (i.e., $A^t = -A$) containing the curvature and torsion. Notice that F is a special orthogonal matrix, since its columns come from an orthonormal basis with right-handed orientation. We should think of (1.13), equally (1.14), as a linear system of ordinary differential equations for the vector-valued functions $T(s), N(s), B(s)$ given the functions $\kappa(s)$ and $\tau(s)$. Using this perspective we can prove the following theorem.

Fundamental Theorem of Space Curves. *(i) Let $p, \hat{p} : I \to \mathbb{R}^3$ be unit speed paths with curvatures $\hat{\kappa}(s) = \kappa(s) \neq 0$ and torsions $\hat{\tau}(s) = \tau(s)$ for all $s \in I$. Then p and $\hat{p}$ are properly congruent. Moreover, if $\hat{\kappa}(s) = \kappa(s) \neq 0$ and $\hat{\tau}(s) = -\tau(s)$ for all $s \in I$ then p and $\hat{p}$ are congruent but not properly congruent.*

(ii) Let $\kappa, \tau : I \to \mathbb{R}$ be smooth functions, with $\kappa > 0$. There these exists a smooth unit speed path $p : I \to \mathbb{R}^3$ with curvature κ and torsion τ, and p is determined uniquely up to proper congruence.

We will not study the proof in detail: the interested Reader can find it in the appendix. Nevertheless, we can understand why the pair (κ, τ) only determines the curve up to proper congruence. The proof is based on the existence and uniqueness of solutions to the Frenet equations (1.13). In the construction of a path p from κ, τ, there are "constants of integration" chosen at two steps: first, the solution to $F' = FA$ requires an initial condition $F(t_0) \in SO(3)$, and second, $F(t)$ provides the tangent $T = p'$ from which $p(t)$ is obtained by integration. This requires another initial condition $p(t_0)$. These choices will, respectively, rotate and translate the path.

Example 1.7. Let us now show that a curve is a helix if and only if it has constant curvature $\kappa > 0$ and constant torsion τ. Example 1.5 gave an arc length parameterised helix with curvature and torsion

$$\kappa(s) = \frac{|a|}{a^2 + b^2}, \qquad \tau(s) = -\frac{b}{a^2 + b^2},$$

for arbitrary constants $a, b \in \mathbb{R}$. Conversely, given constants $\kappa > 0$ and τ we can find $a, b \in \mathbb{R}$ satisfying these equations. To be precise,

$$|a| = \frac{\kappa}{\kappa^2 + \tau^2}, \quad b = -\frac{\tau}{\kappa^2 + \tau^2}.$$

Thus every choice of constants $\kappa > 0$ and τ has a helix with those values for curvature and torsion. By the Fundamental Theorem any path with that curvature and torsion is a helix (congruent to the standard form given in Example 1.5).

2. Smooth Surfaces and their Calculus.

Our aim is to adapt vector calculus to work on smooth surfaces, but before we can do that we need to define the notion of a **smooth surface** itself. From a vector calculus course we expect that this idea should include two common types of construction.

Level surfaces. For a differentiable function $f : \mathbb{R}^3 \to \mathbb{R}$ for each $k \in \mathbb{R}$ the set

$$S_k = \{(x, y, z) : f(x, y, z) = k\}$$

is usually called the level set at level k. For example, the sphere of radius $r > 0$ is the level surface

$$S^2(r) = \{(x, y, z) : x^2 + y^2 + z^2 = r^2\}.$$

For simplicity the unit sphere $S^2(1)$ is usually denoted by S^2. This definition also gives a non-empty set when $r = 0$, but that set is $\{\mathbf{0}\}$, the single point set containing the origin. It would be absurd to call this a surface, so not every level set S_k gives a surface. Notice that the problem with $S^2(0)$ is that it has the wrong dimension: a surface ought to be two-dimensional.

A second family of examples turns up another reason for caution. Consider the level sets

$$S_k = \{(x, y, z) : x^2 + y^2 - z^2 = k\}.$$

For $k < 0$ these are the **hyperboloids of one sheet**, while for $k > 0$ we get the **hyperboloids of two sheets**. In between these cases lies the **cone** S_0. This looks two dimensional at almost every point except the origin, where its tangent plane fails to be well-defined. Our definition will exclude S_0 from being a **smooth** surface for this reason.

Later, the Regular Value Theorem will give us sufficient conditions for a level set to be a surface.

Parametric Surfaces. Another common construction of surfaces is to describe them using the two dimensional version of the parametrisation of a curve, in the form

$$S = \{p(u, v) = (x(u, v), y(u, v), z(u, v)) : (u, v) \in U \subset \mathbb{R}^2\}$$

For example, the map $p : \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$p(u, v) = (\cos u \cos v,\ \sin u \cos v,\ \sin v),$$

has image S^2, i.e., this is a parametrisation of S^2. It is clearly not injective, since it is made from periodic functions. By restricting the domain it can be made injective, but to make it simultaneously surjective we have to carefully choose the domain. For example, the domain

$$D = \{(0, v) : -\pi/2 \le v \le \pi/2\} \cup \{(u, v) : 0 < u < 2\pi,\ -\pi/2 < v < \pi/2\},$$

will make $p : D \to S^2$ bijective.

A good parametrisation ought not the have redundant variables. For example, the map $p(u, v) = (u, 0, 0)$ doesn't really depend upon v, and its image is clearly not a surface. We will see later that what makes for a good parametrisation of a surface is the independence of the vectors of partial derivatives $p_u = \partial p/\partial u$ and $p_v = \partial p/\partial v$ at each point. This also indicates the two dimensional nature of a proper surface: these two vectors should span the tangent plane.

2.1. Fundamental Concepts.

Definition 2.1. *If* $\mathbf{a} \in \mathbb{R}^n$*,* $r > 0$*, then:*

$$B_r(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{a}| < r\},$$

is the ***open ball*** *of radius* r *and centre* $\mathbf{a}$*. Of course:*

$$|\mathbf{x} - \mathbf{a}| = \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2}.$$

A subset $U \subset \mathbb{R}^n$ *is* ***open*** *if for each* $\mathbf{a} \in U$ *there exists* $r > 0$ *such that* $B_r(\mathbf{a}) \subset U$*. A subset* $E \subset \mathbb{R}^n$ *is* ***closed*** *if its complement* E' *is open.*

An open set containing $\mathbf{a} \in \mathbb{R}^n$ *is called a* ***neighbourhood*** *of* $\mathbf{a}$*.*

Remark 2.1. These ideas come from topology (the so-called standard topology of $\mathbb{R}^n$), but we do not need to study topology to understand them in our context. An important observation is that when we use standard inclusions like $\mathbb{R} \subset \mathbb{R}^2$, $x \to (x, y)$, open sets are never mapped to open sets, since every open interval $(a, b) \subset \mathbb{R}$ is mapped to the subset

$$\{(x, y) : a < x < b,\ y = 0\} \subset \mathbb{R}^2,$$

which is neither open nor closed[3]. Thus the notion of "open" is dimension dependent, as the definition indicates.

A function $f : U \to \mathbb{R}^m$ on an open set $U \subset \mathbb{R}^n$ will be called **smooth at** $\mathbf{a} \in U$ if for every $k \in \mathbb{N}$ all k-th order partial derivatives of f exist in a neighbourhood of $\mathbf{a}$ and are continuous at $\mathbf{a}$, and f is **smooth on** U if f is smooth at all $\mathbf{a} \in U$.

Definition 2.2. *Suppose* $U_1 \subset \mathbb{R}^n$ *and* $U_2 \subset \mathbb{R}^m$ *are open subsets. A smooth function* $f : U_1 \to U_2$ *is said to be a* ***diffeomorphism*** *if* f *is invertible and the inverse function* $f^{-1} : U_2 \to U_1$ *is smooth. In this case* U_1 *and* U_2 *are said to be* ***diffeomorphic*** *open sets.*

One of our main tasks will be to adapt the idea of a smooth map to non-open subsets of $\mathbb{R}^3$ (because surfaces, being two dimensional, are never open subsets of $\mathbb{R}^3$). We achieve this using the following idea.

Definition 2.3. *Suppose* $f : D \to \mathbb{R}^m$ *with* $D \subset \mathbb{R}^n$ *and* $\mathbf{a} \in D$*. If* V *is a neighbourhood of* $\mathbf{a}$ *and* $\tilde{f} : V \to \mathbb{R}^m$ *satisfies:*

$$\tilde{f}(\mathbf{x}) = f(\mathbf{x}), \quad \textit{for all } \mathbf{x} \in V \cap D,$$

then $\tilde{f}$ *is called a* ***local extension*** *of* f *at* $\mathbf{a}$*.*

The idea now is that a function on a not-necessarily-open subset $D \subset \mathbb{R}^n$ will be smooth if it admits about each point a local extension which is smooth. From this we can extend the notion of diffeomorphism.

[3]Use the definition to convince yourself that this set is not open and its complement is also not open.

Definition 2.4. *Function $f : D \to \mathbb{R}^m$ is* ***smooth*** *at $\mathbf{a} \in D$ if there exists a local extension of f at $\mathbf{a}$ which is smooth at $\mathbf{a}$. If f is smooth at all $\mathbf{a} \in D$ then f is* ***smooth on*** *D.*

Let $D_1 \subset \mathbb{R}^n$ and $D_2 \subset \mathbb{R}^m$ be subsets which are not necessarily open. A function $f : D_1 \to D_2$ is smooth if it is smooth as a function into $\mathbb{R}^m$. Given this, a smooth function $f : D_1 \to D_2$ is a ***diffeomorphism*** *if f is invertible and $f^{-1} : D_2 \to D_1$ is smooth. Then D_1 is said to be* ***diffeomorphic*** *to D_2, written $D_1 \cong D_2$.*

It is easy to see that diffeomorphism is an equivalence relation. To proceed further with **differential** geometry, and develop our understanding of when two sets are diffeomorphic, we have to recall the notion of the total derivative from vector calculus.

Definition 2.5. *For a smooth map $f : U \to \mathbb{R}^m$ defined on an open set $U \subset \mathbb{R}^n$ its* ***differential*** *or* ***total derivative*** *at $\mathbf{a} \in U$ is the linear map*

$$df(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^m; \quad df(\mathbf{a})[\mathbf{h}] = \frac{d}{dt} f(\mathbf{a} + t\mathbf{h})\Big|_{t=0}.$$

The matrix which represents this linear map (using the canonical bases of $\mathbb{R}^n$ and $\mathbb{R}^m$) is the ***Jacobian matrix*** *$J_f(\mathbf{a})$, the $m \times n$ matrix whose ij-th entry is the partial derivative:*

$$D_j f_i(\mathbf{a}) = \frac{\partial f_i}{\partial x_j}(\mathbf{a}).$$

Remark 2.2. Recall that the vector $df(\mathbf{a})[\mathbf{h}]$ gives the directional derivative of f in the direction $\mathbf{h}$. When $f : \mathbb{R}^n \to \mathbb{R}$ is a scalar function then the 1×3 matrix J_f is also denoted by ∇f, the **gradient vector** of f:

$$\nabla f(\mathbf{p}) = \Big(\frac{\partial f}{\partial x}\Big|_{\mathbf{p}}, \frac{\partial f}{\partial y}\Big|_{\mathbf{p}}, \frac{\partial f}{\partial z}\Big|_{\mathbf{p}} \Big) = (f_x(\mathbf{p}), f_y(\mathbf{p}), f_z(\mathbf{p})).$$

Then, if $\mathbf{h} = (h_1, h_2, h_3)$ we have:

$$df(\mathbf{p})[\mathbf{h}] = f_x(\mathbf{p})h_1 + f_y(\mathbf{p})h_2 + f_z(\mathbf{p})h_3 = \nabla f(\mathbf{p}) \bullet \mathbf{h}.$$

We recall also the vector calculus version of the chain rule.

Theorem 2.6 (Chain Rule)**.** *Suppose we have a "chain" of maps:*

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m \xrightarrow{g} \mathbb{R}^p,$$

satisfying the following conditions:

- *f is smooth at $\mathbf{a} \in \mathbb{R}^n$ (thus $df(\mathbf{a})\colon \mathbb{R}^n \to \mathbb{R}^m$ exists).*
- *g is smooth at $\mathbf{b} = f(\mathbf{a}) \in \mathbb{R}^m$ (thus $dg(\mathbf{b})\colon \mathbb{R}^m \to \mathbb{R}^p$ exists).*

Then $g \circ f\colon \mathbb{R}^n \to \mathbb{R}^p$ is smooth at $\mathbf{a}$, and:

$$\boxed{d(g \circ f)(\mathbf{a}) = dg(\mathbf{b}) \circ df(\mathbf{a})}$$

From the chain rule we can immediately derive a necessary condition for a smooth map $f : U \to V$, between open subset $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$, to be a diffeomorphism. For this necessarily requires

$$f^{-1} \circ f = \mathrm{id}_U, \qquad f \circ f^{-1} = \mathrm{id}_V,$$

the identity maps on U and V respectively. Since f and f^{-1} are both smooth the chain rule implies that

$$df^{-1}(\mathbf{b}) \circ df(\mathbf{a}) = \mathrm{id}_{\mathbb{R}^n}, \qquad df(\mathbf{a}) \circ df^{-1}(\mathbf{b}) = \mathrm{id}_{\mathbb{R}^m}.$$

This means the linear map $df(\mathbf{a})$ has both a left and right inverse. Thus $n = m$ and $df(\mathbf{a})$ must be invertible; equally, a linear isomorphism (in particular, diffeomorphic sets must have the same dimension). It turns out that this condition is also sufficient locally. That is the content of the Inverse Function Theorem, whose proof belongs in a course on calculus of several variables (see, for example, [Bartle]). So we will simply state it and make use of it without further comment.

Inverse Function Theorem. *Let $U \subset \mathbb{R}^n$ be open, and let $f\colon U \to \mathbb{R}^n$ be smooth at $\mathbf{a} \in U$. Suppose that $df(\mathbf{a})\colon \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism. Then there exist neighbourhoods A of $\mathbf{a}$ and B of $\mathbf{b} = f(\mathbf{a})$ such that $f\colon A \to B$ is a smooth diffeomorphism. Furthermore:*

$$df^{-1}(\mathbf{b}) = (df(\mathbf{a}))^{-1}$$

2.2. **Charts, Atlases and Surfaces.** Our definition of a surface will essentially say that about each point the surface has a patch which is diffeomorphic to an open subset of $\mathbb{R}^2$, together with a requirement that this patch sits well in $\mathbb{R}^3$. The next two definitions make this precise.

Definition 2.7. *A **(smooth) surface patch** is a subset $D \subset \mathbb{R}^3$ together with a diffeomorphism $\varphi : D \to U$ to an open subset $U \subset \mathbb{R}^2$.*

Definition 2.8 (Smooth surface)**.** *Let $S \subset \mathbb{R}^3$. A **chart** on S is a surface patch (D, φ) where $D = S \cap V$ for some open subset $V \subset \mathbb{R}^3$. The chart map $\varphi : D \to U \subset \mathbb{R}^2$ is also referred to as **local coordinates** on D, and its inverse*

$$p = \varphi^{-1} : U \to D \subset S,$$

*is called a **local parametrisation**.*

*If each point of S lies in a chart then S is called a **smooth surface**. A collection of charts whose union is S is called an **atlas** for S.*

As the example of the sphere above shows, it would be too restrictive to insist that the whole of a surface can be covered by one chart, since many surfaces are not diffeomorphic to an open subset of $\mathbb{R}^2$.

Example 2.1 (Graphs). Let $f\colon U \to \mathbb{R}$ be a smooth function, where $U \subset \mathbb{R}^2$ is open, and let $S_f \subset \mathbb{R}^3$ to be its **graph:**

$$S_f = \{(u, v, f(u, v)) : (u, v) \in U\}.$$

Figure 7 gives a visualisation of this. We claim that S_f is a surface patch, and therefore a surface. For, if we define $\varphi(u, v, f(u, v)) = (u, v)$ then $\varphi : S_f \to \mathbb{R}^2$ is smooth: it has smooth extension $\tilde{\varphi} : \mathbb{R}^3 \to \mathbb{R}^2$ given by the projection onto the plane $\tilde{\varphi}(x, y, z) = (x, y)$. Its inverse is $p : \mathbb{R}^2 \to S_f$, $p(u, v) = (u, v, f(u, v))$, which is also smooth since each component is a smooth function of u, v. Thus (S_f, φ) is a global chart for S_f.

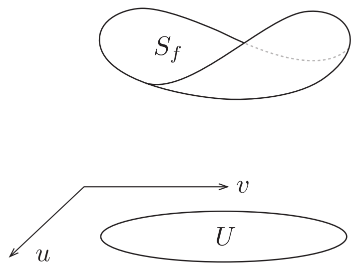

FIGURE 7. Graph of a function of two variables.

Example 2.2 (Sphere: hemispherical charts). We can use graphs to put an atlas on the unit sphere S^2. Take $D = \{(x, y, z) \in S^2 : z > 0\}$. Then $D = S_f$, where $f(u, v) =$

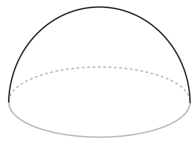

FIGURE 8. Northern hemisphere of the unit sphere.

$\sqrt{1-(u^2+v^2)}$, defined on the open disc $U = \{(u, v) \in \mathbb{R}^2 : u^2+v^2 < 1\}$. So D is a surface patch by the previous example. Furthermore $D = S^2 \cap V$ where $V = \{(x, y, z) : z > 0\}$ is the open upper half space; so D is a chart of S. We can use 6 hemispherical charts like this (or 4, if we want to be economical) to make an atlas for S^2. Thus S^2 is a smooth surface.

Example 2.3 (Sphere: spherical polar charts). Every point on S^2 has **equatorial spherical polar** coordinates, which are obtained from the inverse of the parametrisation seen earlier:

$$p(u, v) = (\cos u \cos v,\ \sin u \cos v,\ \sin v).$$

If we restrict this to the open subset

$$U = \{(u, v) : 0 < u < 2\pi,\ -\pi/2 < v < \pi/2\} = (0, 2\pi) \times (-\pi/2, \pi/2) \subset \mathbb{R}^2,$$

we obtain a smooth bijection $p : U \to S^2 \setminus C$, where $C \subset S^2$

$$C = \{(x, 0, z) \in S^2 : x \geq 0\},$$

is the "Greenwich meridian". See Figure 9. Let $D = S^2 \setminus C$. The inverse of p is the map $\varphi : D \to U$ which gives coordinates (u, v) to $p(u, v)$. These can be interpreted geometrically as the longitude and latitude relative to C. To show that (D, φ) is a chart it suffices to show that φ is smooth (since we know it has smooth inverse p). Its smooth local extension is given by

$$\tilde{\varphi}\colon V \to \mathbb{R}^3; (x, y, z) \mapsto \big(\theta(x, y), \arcsin(z)\big),$$

where $V = \{(x, y, z) : |z| < 1\} \setminus \{(x, 0, z) : x \geq 0$ and $\theta(x, y)$ is the polar angle of the point $(x, y) \in \mathbb{R}^2$. Therefore D is a surface patch. Finally, $D = S^2 \cap V$, so D is a chart. It may be combined with one other similar chart to obtain an atlas.

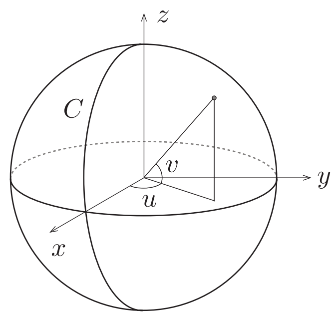

FIGURE 9. Longitude and latitude on the sphere.

2.3. Tangent Spaces.

Definition 2.9. *For a smooth surface S, $\mathbf{p} \in S$, the* ***tangent space to*** S ***at*** $\mathbf{p}$ *is the set*

$$T_{\mathbf{p}}S = \{c'(0) \in \mathbb{R}^3 : c(t) \text{ is a smooth curve on } S \text{ with } c(0) = \mathbf{p}\}.$$

We will call the elements of $T_{\mathbf{p}}S$ tangent vectors, although this definition puts them at the origin $\mathbf{0} \in \mathbb{R}^3$. We will use the phrase **tangent plane** to refer to the parallel plane parallel to $T_{\mathbf{p}}S$ through $\mathbf{p}$, i.e., $\mathbf{p} + T_{\mathbf{p}}S$.

Theorem 2.10. *$T_{\mathbf{p}}S \subset \mathbb{R}^3$ is a two dimensional vector space.*

Proof. Choose a chart (D, φ) about $\mathbf{p}$, with local parametrisation $p : U \to D$ about $\mathbf{a} = \varphi(\mathbf{p}) \in U$. We claim $dp(\mathbf{a}) : \mathbb{R}^2 \to \mathbb{R}^3$ is injective with image $T_{\mathbf{p}}S$. That it is injective follows at once from the existence of a local smooth extension $\tilde{\varphi}$ of φ about $\mathbf{p}$, since $\tilde{\varphi} \circ p = \mathrm{id}_U$, the identity map on U, therefore by the chain rule

$$d\tilde{\varphi}(\mathbf{p}) \circ dp(\mathbf{a}) = id_{\mathbb{R}^2},$$

so that $dp(\mathbf{a})$ must have trivial kernel. To see that the image is $T_{\mathbf{p}}S$ we argue as follows.

(i) First, $\operatorname{im} dp(\mathbf{a}) \subset T_{\mathbf{p}}S$. For if $\mathbf{h} \in \mathbb{R}^2$ then $\gamma(t) = \mathbf{a} + t\mathbf{h}$ lies in U for t small, so $p \circ \gamma$ lies on S. Thus

$$dp(\mathbf{a})[\mathbf{h}] = (p \circ \gamma)'(0) \in T_{\mathbf{p}}S.$$

(ii) Next, $T_{\mathbf{p}}S \subset \operatorname{im} dp(\mathbf{a})$, for if $X \in T_{\mathbf{p}}S$ then $X = c'(0)$ for some smooth curve $c(t)$ in S with $c(0) = \mathbf{p}$. Now $(\varphi \circ c)(t)$ is a smooth curve in $\mathbb{R}^2$. Define $\mathbf{h} = (\varphi \circ c)'(0)$, and observe that

$$X = (p \circ \varphi \circ c)'(0) = dp(\mathbf{a})[(\varphi \circ c)'(0)] = dp(\mathbf{a})[\mathbf{h}] \in \operatorname{im} dp(\mathbf{a}).$$

□

Remark 2.3. When S has local parametrisation $p : U \to S$, for which $p(\mathbf{a}) = \mathbf{p}$ for some $\mathbf{a} \in U$, we can write $p(u, v) = (x(u, v), y(u, v), z(u, v))$. The coordinate lines through $\mathbf{a}$ in U are given by $\gamma(t) = \mathbf{a} + t\mathbf{e}_1$ and $\delta(t) = \mathbf{a} + t\mathbf{e}_2$, for $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$. Since $\gamma'(0) = \mathbf{e}_1$ and $\delta'(0) = \mathbf{e}_2$ we have

$$\frac{\partial p}{\partial u}(\mathbf{a}) = dp(\mathbf{a})[\mathbf{e}_1] = (p \circ \gamma)'(0),$$

$$\frac{\partial p}{\partial v}(\mathbf{a}) = dp(\mathbf{a})[\mathbf{e}_2] = (p \circ \delta)'(0),$$

Since $dp(\mathbf{a})$ is injective these span $T_{\mathbf{p}}S$. Thus

$$T_{\mathbf{p}}S = \mathrm{Sp}\{p_u(\mathbf{a}), p_v(\mathbf{a})\}.$$

2.4. **The Regular Value Theorem.** Recall that for any smooth function $f : \mathbb{R}^3 \to \mathbb{R}$ its level sets are the subsets of $\mathbb{R}^3$ of the form

$$S_k = \{(x, y, z) : f(x, y, z) = k\}, \quad k \in \mathbb{R}.$$

Definition 2.11. *Let $V \subset \mathbb{R}^3$ be an open set and $f : V \to \mathbb{R}$ a smooth function. We say that:*

(i) *$\mathbf{p} \in V$ is a **critical point** of f if $df(\mathbf{p}) = 0$, the zero linear map $\mathbb{R}^3 \to \mathbb{R}$,*
(ii) *$\mathbf{p}$ is a **regular point** of f if $\mathbf{p}$ is not a critical point,*
(iii) *k is a **critical value** of f if $k = f(\mathbf{p})$ for some critical point $\mathbf{p} \in V$,*
(iv) *k is a **regular value** of f if k is not a critical value.*

Remark 2.4.

(i) Since $df(\mathbf{p})[\mathbf{h}] = \nabla f(\mathbf{p}) \bullet \mathbf{h}$, $\mathbf{p}$ is a critical point if and only if $\nabla f(\mathbf{p}) = \mathbf{0}$, and a regular point if and only if $\nabla f(\mathbf{p}) \neq \mathbf{0}$.
(ii) If $S_k = \emptyset$ then k is a regular value of f. For, if not then S_k contains a critical point. Thus, for example, -1 is a regular value of $f(x, y, z) = x^2 + y^2 + z^2$.
(iii) If k is a regular value of f then all points of S_k are regular points; however if k is a critical value then **at least one** point of S_k is a critical point.

Example 2.4. Suppose $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^2 + y^2 - z^2$. Then $\nabla f = (2x, 2y, -2z)$, so the only critical point is $(0, 0, 0)$ and the corresponding critical value is $f(0, 0, 0) = 0$. The level set S_0 is the cone $x^2 + y^2 - z^2 = 0$, but notice that even though it corresponds to a critical value most of the points on S_0 are regular points (in fact, every point except the origin $\mathbf{0}$ - the 'cone point').

But suppose we take the same function restricted to $V = \mathbb{R}^3 \setminus \{\mathbf{0}\}$. Then $\nabla f = \mathbf{0}$ has no solution in V, so every point in V is a regular point and every value is a regular value for this domain - including 0! In this case S_0 is a disjoint union of two punctured half cones.

With the definitions above in hand, we can now state the Regular Value Theorem. The proof would take us more time than we have, and so it has been placed in Appendix B.

Regular Value Theorem. *Suppose $V \subset \mathbb{R}^3$ is open, and $f : V \to \mathbb{R}$ is a smooth function. If $k \in \mathbb{R}$ is a regular value of f, and the level set $S_k = \{\mathbf{p} \in V : f(\mathbf{p}) = k\}$ is non-empty, then S_k is a smooth surface. Furthermore $T_{\mathbf{p}}S_k = \ker df(\mathbf{p})$ for all $\mathbf{p} \in S_k$; equivalently*

$$T_{\mathbf{p}}S_k = \nabla f(\mathbf{p})^{\perp} := \{\mathbf{v} \in \mathbb{R}^3 : \nabla f(\mathbf{p}) \bullet \mathbf{v} = 0\}.$$

If $\nabla f(\mathbf{p}) = (a, b, c)$ then the last part says $T_{\mathbf{p}}S$ is the plane with equation $ax+by+cz = 0$.

Example 2.5. From the previous example, for $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^2 + y^2 - z^2$, by the Regular Value Theorem every level set S_k for $k \neq 0$ is a smooth surface: it is easy to see these are non-empty, since $(\sqrt{k}, 0, 0) \in S_k$ for $k > 0$, while $(0, 0, \sqrt{-k}) \in S_k$ for $k < 0$. We get the hyperboloids of one sheet for $k > 0$ and the hyperboloids of two sheets for $k < 0$. The cone S_0 is not a smooth surface, but $S_0 \setminus \{\mathbf{0}\}$ (the cone minus its cone point) is a smooth surface: it is the zero level set for $f : V \to \mathbb{R}$ with $V = \mathbb{R}^3 \setminus \{\mathbf{0}\}$.

As an example of finding the tangent space, consider $\mathbf{p} = (1, 0, 0) \in S_1$. Then $\nabla f(\mathbf{p}) = (2, 0, 0)$ so $T_{\mathbf{p}}S$ is the plane $x = 0$.

2.5. **Maps between Surfaces.** Let $f : S_1 \to S_2$ be a smooth map between two smooth surfaces. For any $\mathbf{p} \in S_1$ and $X \in T_{\mathbf{p}}S_1$ we have $X = c'(0)$ for some smooth curve $c(t)$ on S_1, and $(f \circ c)(t)$ is a smooth curve on S_2. Therefore

$$df(\mathbf{p})[X] = (f \circ c)'(0) \in T_{f(\mathbf{p})}S_2.$$

Therefore the following definition makes sense.

Definition 2.12. *For a smooth map $f : S_1 \to S_2$ between two smooth surfaces the **differential of f at $\mathbf{p}$** is the linear map*

$$df(\mathbf{p}) : T_{\mathbf{p}}S_1 \to T_{f(\mathbf{p})}S_2; \quad df(\mathbf{p})[X] = (f \circ c)'(0),$$

where $c(t)$ is any smooth curve on S_1 with $c(0) = \mathbf{p}$ and $c'(0) = X$.

Remark 2.5.

(i) Since every smooth map $f : S_1 \to S_2$ is, by definition, smooth as a map $f : S_1 \to \mathbb{R}^3$, it follows immediately that this differential is a linear map.

(ii) Recall that if $p(u, v)$, $p : U \to S_1$, is a local parametrisation about $\mathbf{p}$ ($p(\mathbf{a}) = \mathbf{p}$ for some $\mathbf{a} \in U \subset \mathbb{R}^2$) then $T_{\mathbf{p}}S_1 = \mathrm{Sp}\{p_u(\mathbf{a}), p_{\mathbf{v}}(\mathbf{a})\}$. Now for a smooth map $f : S_1 \to S_2$ between two smooth surfaces, and with $\gamma(t) = \mathbf{a} + t\mathbf{e}_1$, $\delta(t) = \mathbf{a} + t\mathbf{e}_2$ on U, we have

$$\begin{aligned} df(\mathbf{p})[p_u(\mathbf{a})] &= (f \circ p \circ \gamma)'(0) = (f \circ p)_u(\mathbf{a}), \\ df(\mathbf{p})[p_v(\mathbf{a})] &= (f \circ p \circ \delta)'(0) = (f \circ p)_v(\mathbf{a}). \end{aligned}$$

This gives us a way of calculating df in the local coordinate expression $f(p(u, v))$. For if $X \in T_{\mathbf{p}}S_1$ then $X = X^u p_u(\mathbf{a}) + X^v p_v(\mathbf{a})$ for coefficients $X^u, X^v \in \mathbb{R}$, and then by linearity

$$df(\mathbf{p})[X] = X^u df(\mathbf{p})[p_u(\mathbf{a})] + X^v df(\mathbf{p})[p_v(\mathbf{a})] = X^u (f \circ p)_u(\mathbf{a}) + X^v (f \circ p)_u(\mathbf{a}). \quad (2.1)$$

Removing the clutter of notation labelling points, we can simplify this:

$$df[X] = X^u (f \circ p)_u + X^v (f \circ p)_v. \quad (2.2)$$

(iii) When $\tilde{f}$ is a local smooth extension for f about $\mathbf{p}$ and $X \in T_{\mathbf{p}}S_1$ is tangent to the smooth curve $c(t)$ on S_1, we have

$$d\tilde{f}(\mathbf{p})[X] = (\tilde{f} \circ c)'(0) = (f \circ c)'(0) = df(\mathbf{p})[X].$$

This gives us a way of calculating df given a local smooth extension for f.

Lemma 2.13 (Chain Rule for Surface Maps)**.** *Let $f : S_1 \to S_2$ and $g : S_2 \to S_3$ be two smooth maps between smooth surfaces, and let $\mathbf{p} \in S_1$ with $\mathbf{q} = f(\mathbf{p})$. Then $g \circ f : S_1 \to S_3$ is smooth with differential at $\mathbf{p}$ given by*

$$d(g \circ f)(\mathbf{p}) = dg(\mathbf{q}) \circ df(\mathbf{p}) : T_{\mathbf{p}}S_1 \to T_{g(\mathbf{q})}S_3.$$

The proof is a trivial application of the usual chain rule for maps in $\mathbb{R}^3$, given that f and g both have local smooth extensions into $\mathbb{R}^3$.

Corollary 2.14. *Suppose $f : S_1 \to S_2$ is a diffeomorphism of surfaces. Then at every $\mathbf{p} \in S_1$ the differential $df(\mathbf{p}) : T_{\mathbf{p}}S_1 \to T_{f(\mathbf{p})}S_2$ is a linear isomorphism, with inverse $df^{-1}(f(\mathbf{p}))$.*

This follows immediately from the chain rule applied to $f^{-1} \circ f = \mathrm{id}_{S_1}$. The more interesting fact is that the Inverse Function Theorem can also be adapted to maps between surfaces. For the statement of this we need to adapt the notion of a local diffeomorphism.

Definition 2.15. *A smooth map $f : S_1 \to S_2$ is a* ***local diffeomorphism at*** $\mathbf{p} \in S_1$ *if the differential $df(\mathbf{p}) : T_{\mathbf{p}}S_1 \to T_{f(\mathbf{p})}S_2$ is a linear isomorphism. We say $f : S_1 \to S_2$ is a* ***local diffeomorphism*** *when it is a local diffeomorphism about each point.*

Theorem 2.16 (Inverse Function Theorem for Surface Maps.)**.** *A smooth map $f : S_1 \to S_2$ is a local diffeomorphism at $\mathbf{p}$ if and only if there exist chart domains $D_1 \subset S_1$ containing $\mathbf{p}$ and $D_2 \subset S_2$ containing $f(\mathbf{p})$ such that $f : D_1 \to D_2$ is a diffeomorphism.*

Proof. This is simply a consequence of the standard Inverse Function Theorem 2.1, applied to a **local representative** F of f at $\mathbf{p}$. This means the following. Let (φ_1, D_1) be a chart on S_1 about $\mathbf{p}$, and (φ_2, D_2) be a chart about $f(\mathbf{p})$ on S_2. We will denote the corresponding inverses (local parametrisations) by $p_j : U_j \to D_j$. Now define $F : U_1 \to U_2$ to be the composition $F = \varphi_2 \circ f \circ p_1$;

$$U_1 \xrightarrow{p_1} D_1 \xrightarrow{f} D_2 \xrightarrow{\varphi_2} U_2.$$

Then if $\varphi_1(\mathbf{p}) = \mathbf{a} \in U_1$ and $\mathbf{q} = f(\mathbf{p})$ it follows from the Chain Rule (for Surfaces) that:

$$dF(\mathbf{a}) = d\varphi_2(\mathbf{q}) \circ df(\mathbf{p}) \circ dp_1(\mathbf{a}).$$

Now all three linear maps on the right hand side are isomorphisms, so $dF(\mathbf{a}) : \mathbb{R}^2 \to \mathbb{R}^2$ must also be one. Therefore by the Inverse Function Theorem there exists a neighbourhood $W_1 \subset U_1$ of $\mathbf{a}$ such that the restriction of F to W_1 is a diffeomorphism. By reducing the size of the chart domains D_1, D_2 if necessary (so that $D_1 = p_1(W_1)$, $D_2 = f \circ p_1(W_1)$, we have a diffeomorphism $F : W_1 \to W_2$. It follows that the restriction of f to D_1 is

$$f = p_2 \circ F \circ \varphi_1 : D_1 \to D_2,$$

which is a composite of diffeomorphisms, and therefore a diffeomorphism. □

Notice what this theorem says: if the differential $df(\mathbf{p})$ is invertible at each $\mathbf{p}$ then **locally** f possesses an inverse, but this need not mean f itself is invertible. The following example demonstrates this point.

Example 2.6. Let $S_1 \subset \mathbb{R}^3$ be the plane $z = 0$ and S_2 be the cylinder $\{(x, y, z) : x^2+z^2 = 1\}$. The map

$$f : S_1 \to S_2; \quad f(x, y, 0) = (\cos(x), y, \sin(x)),$$

wraps the plane S_1 around the cylinder S_2 infinitely many times, so it is not injective and cannot have an inverse. It is plainly smooth and at each $\mathbf{p} = (x, y, 0)$ clearly $T_{\mathbf{p}}S_1$ is the subspace of $\mathbb{R}^3$ with equation $z = 0$. S_1 can be globally parameterised by $p(u, v) = (u, v, 0)$, for which

$$(f \circ p)_u = (-\sin(u), 0, \cos(u)), \quad (f \circ p)_v = (0, 1, 0).$$

These are linearly independent at every point $p(u, v)$, therefore

$$df(\mathbf{p})[X] = X^u(-\sin(u), 0, \cos(u)) + X^v(0, 1, 0) = (-X^u \sin(u), X^v, X^u \cos(u)),$$

has trivial kernel. So by Theorem 2.16 f is locally invertible.

3. The Geometry of Smooth Surfaces.

Throughout this chapter S will be a smooth surface in $\mathbb{R}^3$.

3.1. **The Riemannian Metric.** Euclidean space $\mathbb{R}^3$ comes equipped with a canonical inner product: the dot product. Since each tangent space $T_{\mathbf{p}}S \subset \mathbb{R}^3$ is a vector subspace each inherits this inner product:

$$\langle X, Y\rangle_{\mathbf{p}} = X \bullet Y,\ \ X, Y \in T_{\mathbf{p}}S \subset \mathbb{R}^3. \tag{3.1}$$

Recall that the defining properties of an inner product are

$$\begin{aligned} \langle X, Y\rangle &= \langle Y, X\rangle \\ \langle aX + bY, Z\rangle &= a\langle X, Z\rangle + b\langle Y, Z\rangle,\ \ a, b, \in \mathbb{R} \\ \langle X, X\rangle &> 0,\ \forall\ X \neq 0. \end{aligned}$$

The first two say that $\langle\ ,\ \rangle$ is a **symmetric bilinear form**, and the third one says it is **positive definite**.

Definition 3.1. *The **first fundamental form** of a smooth surface $S \subset \mathbb{R}^3$ is the inner product* (3.1) *induced on each tangent space $T_{\mathbf{p}}S$ by the dot product in $\mathbb{R}^3$. This is also called the **induced Riemannian metric**.*

We should keep in mind that the first fundamental form is a **family** of inner products, one for each tangent space. In modern language it is an example of a tensor on S. When we want to emphasise that we are making a calculation at a particular point $\mathbf{p}$ we can write $\langle\ ,\ \rangle_{\mathbf{p}}$.

We can express the information carried by the Riemannian metric in coordinates. Let $p : U \to S$ be a local parametrisation corresponding to a coordinate chart $\varphi : D \to U$ on S. Recall that at each $\mathbf{p} = p(\mathbf{a})$ we have $T_{\mathbf{p}}S = \mathrm{Sp}\{p_u(\mathbf{a}), p_v(\mathbf{a})\}$. Thus $X, Y \in T_{\mathbf{p}}S$ can be expressed in this basis as

$$X = X^u p_u + X^v p_v, \quad Y = Y^u p_u + Y^v p_v,$$

where $X^u, X^v, Y^u, Y^v \in \mathbb{R}$ are the **components** of X, Y in these coordinates. Then

$$\begin{aligned} \langle X, Y\rangle &= \langle X^u p_u + X^v p_v,\ Y^u p_u + Y^v p_v\rangle \\ &= X^u Y^u\, |p_u|^2 + (X^u Y^v + X^v Y^u)\langle\, p_u, p_v\rangle + X^v Y^v\, |p_v|^2 \\ &= \begin{pmatrix} X^u & X^v \end{pmatrix} \begin{pmatrix} |p_u|^2 & \langle\, p_u, p_v\rangle \\ \langle\, p_v, p_u\rangle & |p_v|^2 \end{pmatrix} \begin{pmatrix} Y^u \\ Y^v \end{pmatrix}. \end{aligned}$$

Notice that this is nothing other than the usual process for expressing an inner product as a matrix using a basis for the vector space. The matrix will, of course, be symmetric and positive definite.

Following the terminology due to Gauss himself we define the components of this matrix to be

$$E = |p_u|^2 = p_u \bullet p_u, \quad F = \langle\, p_u, p_v\rangle = p_u \bullet p_v, \quad G = |p_v|^2 = p_v \bullet p_v.$$

These are called the **components** of the Riemannian metric in the given chart. With this notation

$$\begin{pmatrix} X^u & X^v \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} Y^u \\ Y^v \end{pmatrix}. \tag{3.2}$$

N.B. Notice that as we move though the chart domain $D = p(U)$, E, F, G will be functions of (u, v). If we choose another parametrisation then we should expect all the components E, F, G (and X^u, X^v, Y^u, Y^v) to be different, but the quantity $\langle X, Y\rangle$ is geometric and will not change.

Since p_u, p_v are the tangent vectors along the coordinate curves through the chart domain $D = p(U)$ we make the following definitions.

Definition 3.2. *We say a coordinate parametrisation $p(u, v)$ is **orthogonal** when $F(u, v) = \langle p_u, p_v\rangle = 0$ throughout U. We say these coordinates are **isothermal** (or **conformal**) when additionally $E(u, v) = G(u, v)$ throughout U, and **orthonormal** when we also have $E = G = 1$.*

These are properties of the coordinates, not the Riemannian metric.

Example 3.1 (Planes). Let S be an arbitrary plane in $\mathbb{R}^3$, containing the point $\mathbf{b} \in \mathbb{R}^3$ and with normal $\mathbf{n} \in \mathbb{R}^3$. For any two linearly independent $X, Y \in \mathbb{R}^3$ with $X \bullet \mathbf{n} = 0 = Y \bullet \mathbf{n}$ we can parameterise S by

$$p(u, v) = \mathbf{b} + uX + vY, \quad u, v \in \mathbb{R}^2.$$

Clearly $p_u = X$ and $p_v = Y$ so in this parametrisation the Riemannian metric has components

$$E = |X|^2, \quad F = X \bullet Y, \quad G = |Y|^2.$$

The coordinate system (u, v) will therefore be orthogonal precisely when X, Y are orthogonal, isothermal when X, Y are orthogonal and have the same length, and orthonormal when X, Y are orthonormal.

Example 3.2 (Cylinder). We can parameterise part of the cylinder $S = \{(x, y, z) : x^2 + z^2 = 1\}$ by

$$p(u, v) = (\cos(u), v, \sin(u)); \quad -\pi < u < \pi, \ v \in \mathbb{R}.$$

In this case

$$p_u = (-\sin(u), 0, \cos(u)), \quad p_v = (0, 1, 0),$$

so that the Riemannian metric components are

$$E = |p_u|^2 = \sin^2(u) + \cos^2(u) = 1, \quad F = p_u \bullet p_v = 0, \quad G = |p_v|^2 = 1.$$

So these coordinates on the cylinder are also orthonormal.

Example 3.3 (Helicoid). The helicoid may be defined parametrically as follows:

$$S = \{p(u, v) = (u \cos v, \ u \sin v, \ av) : u, v \in \mathbb{R}\},$$

where $a \neq 0$ is constant. In this parametrisation

$$p_u = (\cos v, \sin v, 0), \quad p_v = (-u \sin v, u \cos v, a),$$

so that

$$E = \cos^2 v + \sin^2 v = 1, \quad F = 0, \quad G = u^2(\sin^2 v + \cos^2 v) + a^2 = u^2 + a^2.$$

So this is an orthogonal coordinate system on the helicoid which is not isothermal: see Figure 10.

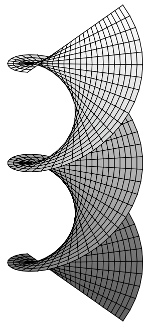

FIGURE 10. Orthogonal coordinate curves on a helicoid.

3.2. **Lengths and Areas.** The Riemannian metric provides all the information required to calculate the arc length of curves on the surface S, or the area of bounded regions on S.

Lengths of curves. Let $c : [a, b] \to S$ be a smooth curve, and suppose it lies entirely in a chart domain (D, φ). In this chart $c(t)$ has coordinates $(u(t), v(t)) = \varphi(c(t))$, equally, $c(t) = p(u(t), v(t))$. Therefore

$$\frac{dc}{dt} = \frac{\partial p}{\partial u}\frac{du}{dt} + \frac{\partial p}{\partial v}\frac{dv}{dt} = u'p_u + v'p_v.$$

Therefore

$$|c'|^2 = E.(u')^2 + 2Fu'v' + G.(v')^2 = \langle c', c' \rangle.$$

Of course, this is just the expression (3.2) for the squared length of the tangent vector c' and only involves the Riemannian metric. It follows that the arc length of c is

$$s(a, b) = \int_a^b |c'(t)|dt = \int_a^b \sqrt{E.(u')^2 + 2Fu'v' + G.(v')^2}dt. \tag{3.3}$$

In case $c(t)$ does not lie inside one coordinate domain, we can break c into a finite number of segments each of which lies in some coordinate domain, and add the arc lengths of each segment.

Remark 3.1. This is the origin of the classical (and still quite commonly used) expression for the Riemannian metric:

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2,$$

sometimes referred to as "the element of arc length". This has the virtue that it makes clear the role of E, F, G as coordinate dependent components of a symmetric bilinear form.

Areas. We can compute the area of any region $R \subset S$ which is the image of a region $Q \subset U \subset \mathbb{R}^2$, by some local parametrisation $p : U \to S$, over which integration in the plane is well-defined[4]. For in that case standard vector calculus tells us that

$$\text{Area}(R) = \iint_Q |p_u \times p_v| du\, dv = \iint_Q \sqrt{EG - F^2}\, du\, dv. \tag{3.4}$$

using Lagrange's formula $|p_u \times p_v|^2 = |p_u|^2|p_v|^2 - (p_u \bullet p_v)^2$. As with arc length, the "element of surface area" $\sqrt{EG - F^2}\, du\, dv$ is a coordinate invariant quantity even though the function $\sqrt{EG - F^2}$ may change.

3.3. **Isometries and Local Isometries.** Recall that a smooth map $f\colon S \to \bar{S}$ between smooth surfaces S and $\bar{S}$ is said to be a **local diffeomorphism** at $\mathbf{p} \in S$ if $df(\mathbf{p})\colon T_{\mathbf{p}}S \to T_{f(\mathbf{p})}\bar{S}$ is a linear isomorphism.

Definition 3.3. *A smooth map $f\colon S \to \bar{S}$ is said to be a* ***local isometry*** *at $\mathbf{p} \in S$ if $df(\mathbf{p})\colon T_{\mathbf{p}}S \to T_{f(\mathbf{p})}\bar{S}$ is a linear isometry of vector spaces:*

$$\big\langle df(\mathbf{p})[X], df(\mathbf{p})[Y]\big\rangle = \langle X, Y\rangle, \quad \forall X, Y \in T_{\mathbf{p}}S. \tag{3.5}$$

Alternatively, we say that f ***preserves the first fundamental forms*** *of S and $\bar{S}$.*

Remark 3.2. It suffices to check equation (3.5) on all pairs of vectors in a basis of $T_{\mathbf{p}}S$. In particular, if $df(\mathbf{p})$ maps an orthonormal basis of $T_{\mathbf{p}}S$ to an orthonormal basis of $T_{f(\mathbf{p})}\bar{S}$ then $df(\mathbf{p})$ is a linear isometry, and conversely. It follows that a linear isometry is a linear isomorphism. Hence if f is a local isometry at $\mathbf{p}$ then f is a local diffeomorphism at $\mathbf{p}$.

Definition 3.4. *We say that f is a* ***local isometry*** *if f is a local isometry at each point. If in addition f is a diffeomorphism then f is called an* ***isometry*** *and S, $\bar{S}$ are said to be* ***isometric.***

The relation of being isometric is an equivalence relation on the set of all smooth surfaces. The relationship between these different types of map may be summarised by the following diagram:

$$\begin{array}{ccc} \text{Isometry} & \Rightarrow & \text{Diffeomorphism} \\ \Downarrow & & \Downarrow \\ \text{Local isometry} & \Rightarrow & \text{Local diffeomorphism} \end{array}$$

Example 3.4. In example 2.6 we showed that the plane S with equation $z = 0$ and the cylinder $\bar{S}$ with equation $x^2 + z^2 = 1$ are locally diffeomorphic, but not diffeomorphic. Recall we used the map

$$f : S \to \bar{S}; \quad f(x, y, 0) = (\cos x, y, \sin x).$$

We will show that this is a local isometry. In terms of the global parametrisation $p(u, v) = (u, v, 0)$ on S we have:

$$p_u = (1, 0, 0), \quad p_v = (0, 1, 0),$$

[4]For technical reasons, we stick to regions which are the closure of bounded open sets whose boundary is a finite union of piecewise continuous curves.

So this is an orthonormal basis for $T_{p(u,v)}S$. Now

$$df[p_u] = \frac{\partial}{\partial u} f(p(u,v)) = (-\sin u, 0, \cos u), \qquad df[p_v] = \frac{\partial}{\partial v} f(p(u,v)) = (0,1,0),$$

which is an orthonormal basis of $T_{f(p(u,v))}\bar{S}$. Thus df maps an orthonormal basis to an orthonormal basis, hence f is a local isometry. However, f is cannot be an isometry because it is not a diffeomorphism.

We can check whether a map is a local isometry by looking at how it relates the components E, F, G on S with the components $\bar{E}, \bar{F}, \bar{G}$ on $\bar{S}$, but since these components are calculated in charts we need to first ensure that we are making the calculation in compatible charts. The correct notion for compatibility is given by the next definition.

Definition 3.5. *Suppose $f\colon S \to \bar{S}$ is smooth. Parametrisations $p : U \to S$ and $\bar{p} : U \to \bar{S}$ are said to be **f-adapted** if $\bar{p} = f \circ p$. Equally, local charts (φ, D) and $(\bar{\varphi}, \bar{D})$ are f-adapted when $\varphi(D) = \bar{\varphi}(\bar{D})$ and $\bar{\varphi} \circ f = \varphi$ on D.*

Notice that since both φ and $\bar{\varphi}$ are diffeomorphisms the only way they can be f-adapted is if $f : D \to \bar{D}$ is a diffeomorphism.

With this notion of f-adapted parametrisations we have the following useful result.

Lemma 3.6 (E, F, G Lemma)**.** *Suppose $f\colon S \to \bar{S}$ is smooth. Then f is a local isometry at $\mathbf{p} \in S$ if and only if there exist f-adapted parametrisations (p, U) about $\mathbf{p} = p(\mathbf{a})$ and $(\bar{p}, U)$ about $f(\mathbf{p}) = \bar{p}(\mathbf{a})$ such that*

$$(\bar{E}, \bar{F}, \bar{G}) = (E, F, G), \quad \textit{at } \mathbf{a}.$$

Proof. For f-adapted parametrisations we have

$$\bar{p}_u = (f \circ p)_u = df[p_u], \quad \bar{p}_v = (f \circ p)_v = df[p_v].$$

Hence

$$\bar{E} = |\bar{p}_u|^2 = |df[p_u]|^2, \quad \bar{F} = \langle \bar{p}_u, \bar{p}_v \rangle = \langle df[p_u], df[p_v] \rangle, \quad \bar{G} = |\bar{p}_v|^2 = |df[p_v]|^2.$$

Write $X_1 = p_u(\mathbf{a})$ and $X_2 = p_v(\mathbf{a})$, where $\mathbf{a} = \varphi(\mathbf{p})$. Then

$$\begin{aligned}
& (\bar{E}, \bar{F}, \bar{G}) = (E, F, G) \quad \text{at } \varphi(\mathbf{p}) \\
\iff\ & \big\langle df(\mathbf{p})[X_i],\, df(\mathbf{p})[X_j] \big\rangle = \langle X_i, X_j \rangle, \quad i, j = 1, 2 \\
\iff\ & df(\mathbf{p}) \text{ is a linear isometry} \\
\iff\ & f \text{ is a local isometry at } \mathbf{p}.
\end{aligned}$$

It remains to show that if f is a local isometry at $\mathbf{p}$ then f-adapted charts exist. Since f is a local diffeomorphism at $\mathbf{p}$ we can find charts D about $\mathbf{p}$ and $\bar{D}$ about $f(\mathbf{p})$ such that $f\colon D \to \bar{D}$ is a diffeomorphism, by the Inverse Function Theorem for surfaces. If the chart map for D is $\varphi\colon D \to U$, redefine the chart map for $\bar{D}$ by $\bar{\varphi}\colon \bar{D} \to U$; $\bar{\varphi} = \varphi \circ f^{-1}$. Then $\bar{\varphi}$ is smooth (by the Chain Rule for surfaces), and invertible with inverse $\bar{p} = f \circ p$ also smooth; hence $\bar{\varphi}$ is indeed a chart map for $\bar{D}$. ☐

3.4. **The Shape Operator.** At each point on a smooth surface S we have a choice of two unit normal vectors. In any coordinate chart it is always possible to smoothly assign a unit normal to each point, using the local parametrisation $p : U \to S$ as follows:

$$\xi(p(\mathbf{a})) = \frac{p_u(\mathbf{a}) \times p_v(\mathbf{a})}{|p_u(\mathbf{a}) \times p_v(\mathbf{a})|}, \ \mathbf{a} \in U.$$

We say that the coordinates give an **orientation** to S in $D = p(U)$. For some surfaces it is not possible to extend such an orientation smoothly to the whole surface: the Möbius band is one example.

Definition 3.7. *A smooth surface S is said to be **orientable** if there exists a **smooth unit normal field** on S, i.e., a smooth function $\xi\colon S \to \mathbb{R}^3$ satisfying for all $\mathbf{p} \in S$:*

$$\xi(\mathbf{p}) \perp T_{\mathbf{p}}S \quad \textit{and} \quad |\xi(\mathbf{p})| = 1.$$

*If S is orientable then there are precisely two smooth unit normal fields, choice of either of which constitutes an **orientation** of S. An **oriented surface** is an orientable surface together with an orientation.*

Example 3.5. Every smooth level surface $S_k = \{(x, y, z) : F(x, y, z) = k\}$, for a regular value k, is orientable by the unit normal obtained from ∇F:

$$\xi(\mathbf{p}) = \frac{\nabla F(\mathbf{p})}{|\nabla F(\mathbf{p})|}.$$

Definition 3.8. *For a smooth oriented surface S the chosen unit normal field, thought of as a surface map $\xi : S \to S^2$, is called the **Gauss map**.*

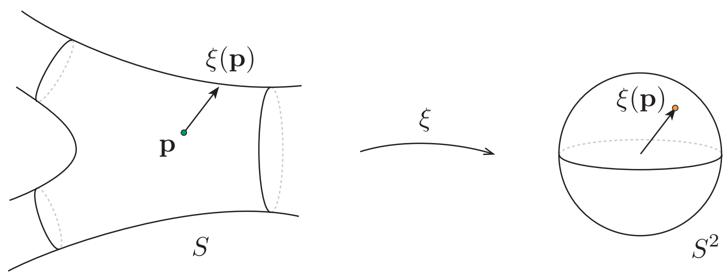

FIGURE 11. The Gauss map

The differential $d\xi(\mathbf{p})$ of the unit normal field measures the rate of turning of the tangent plane $T_{\mathbf{p}}S$ as $\mathbf{p}$ moves along S. It is a linear map $d\xi(\mathbf{p}) : T_{\mathbf{p}}S \to T_{\xi(\mathbf{p})}S^2$, but we know that

$$T_{\xi(\mathbf{p})}S^2 = \xi(\mathbf{p})^{\perp} = T_{\mathbf{p}}S. \tag{3.6}$$

Using this identification we regard this differential as a linear map from $T_{\mathbf{p}}S$ to itself:

$$d\xi(\mathbf{p}) : T_{\mathbf{p}}S \to T_{\mathbf{p}}S.$$

Definition 3.9. *The **shape operator** (or **Weingarten map**) of S at $\mathbf{p}$ is the linear map/operator:*

$$A_{\mathbf{p}}\colon T_{\mathbf{p}}S \to T_{\mathbf{p}}S;\;\; A_{\mathbf{p}} = -d\xi(\mathbf{p}).$$

We often abbreviate $A_{\mathbf{p}} = A$ when the point $\mathbf{p} \in S$ is understood, or to indicate the family of maps, one at each point.

Notice that choosing the opposite orientation, $-\xi$, changes the sign of the shape operator: $A \mapsto -A$.

Example 3.6 (Plane). Whichever orientation is selected, ξ is constant and hence $d\xi(\mathbf{p}) = 0$ (the zero map $T_{\mathbf{p}}S \to \mathbb{R}^3$) for all $\mathbf{p} \in S$. Thus $A(X) = \mathbf{0}$ for all $X \in T_{\mathbf{p}}S$.

Example 3.7 (Sphere). Let S be the sphere of radius R centred at the origin, and choose $\xi(\mathbf{p}) = \mathbf{p}/R$ for all $\mathbf{p} \in S$, which is the outward-pointing unit normal. Then:

$$d\xi(\mathbf{p})[X] = \xi(X) = X/R, \quad \text{for all } X \in T_{\mathbf{p}}S,$$

because ξ is the restriction of the linear map $\mathbf{v} \mapsto \mathbf{v}/R$ on $\mathbb{R}^3$. Thus $A(X) = (-1/R)X$ for all tangent vectors X.

Recall from Linear Algebra that a linear map $A : V \to V$ on an inner product space $(V, \langle\ ,\ \rangle)$ is **self-adjoint** when

$$\langle AX, Y\rangle = \langle X, AY\rangle, \quad \forall\ X, Y \in V.$$

Recall also that this means A is represented by a symmetric matrix when any basis of V is chosen. In particular, V has an orthonormal basis of eigenvectors of A which diagonalises A, and the eigenvalues are all real.

Lemma 3.10 (Shape Lemma)**.** *For any smooth surface S, the shape operator $A_{\mathbf{p}}$ is a self-adjoint operator on $T_{\mathbf{p}}S$, for all $\mathbf{p} \in S$.*

Proof. Since A is a linear operator, it suffices to establish its symmetry on a basis (X_1, X_2) of $T_{\mathbf{p}}S$. Clearly $\langle A(X_i), X_i\rangle = \langle X_i, A(X_i)\rangle$, so it suffices to show:

$$\langle A(X_1), X_2\rangle = \langle X_1, A(X_2)\rangle.$$

For this, we choose a chart about $\mathbf{p}$ and take the basis $X_1 = p_u$ and $X_2 = p_v$. Then

$$A(X_1) = -d\xi(\mathbf{p})[\,p_u] = -(\xi \circ p)_u, \quad A(X_2) = -d\xi(\mathbf{p})[\,p_v] = -(\xi \circ p)_v.$$

Thus

$$\langle A(X_1), X_2\rangle = -(\xi \circ p)_u \bullet p_v = (\xi \circ p) \bullet p_{vu},$$

where $p_{vu} = \partial^2 p/\partial u \partial v$. The last equality follows from $(\xi \circ p) \bullet p_v = 0$, by differentiation with respect to u. Now swapping u with v in the previous calculation gives

$$\langle A(X_2), X_1\rangle = -(\xi \circ p)_v \bullet p_u = (\xi \circ p) \bullet p_{uv}.$$

Since $p_{uv} = p_{vu}$ we have shown that $\langle A(X_1), X_2\rangle = \langle A(X_2), X_1\rangle$. ∎

Example 3.8 (Paraboloids). For each $r \in \mathbb{R}$ we define a smooth function $f_r : \mathbb{R}^2 \to \mathbb{R}$ by:

$$f_r(x, y) = x^2 + ry^2.$$

Let S_r be the graph of f_r. Then S_r is a smooth surface, with a global chart. The intersection of S_r with any vertical plane is a parabola. If $r > 0$ (resp. $r < 0$) then the level curves of f_r are ellipses (resp. hyperbolas). Consequently S_r is called an **elliptic** (resp. **hyperbolic**) **paraboloid**. The hyperbolic paraboloid is the archetypal saddle surface. When $r = 0$ the surface is a **parabolic cylinder:** The paraboloid S_r may also be viewed as the level set

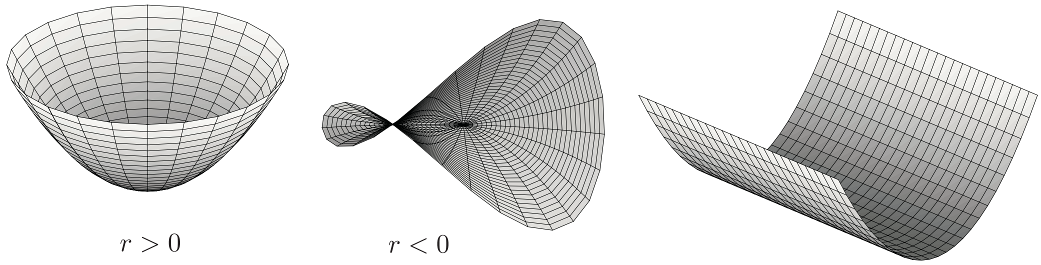

FIGURE 12. Elliptic and hyperbolic paraboloids; parabolic cylinder.

of the smooth function:

$$F_r(x, y, z) = z - x^2 - ry^2,$$

corresponding to the regular value $0 \in \mathbb{R}$. We give S_r the induced orientation:

$$\xi(x, y, z) = \frac{\nabla F_r(x, y, z)}{|\nabla F_r(x, y, z)|} = \frac{(-2x, -2ry, 1)}{\sqrt{4x^2 + 4r^2y^2 + 1}}.$$

Let $\mathbf{p} = (0, 0, 0)$, the unique point in common to all the S_r. Since $\nabla F_r(\mathbf{p}) = (0, 0, 1)$ the tangent space $T_{\mathbf{p}}S_r$ is the plane $z = 0$. We will compute the matrix which represents the shape operator of S_r at $\mathbf{p}$ by choosing the (orthonormal) basis $X_1 = (1, 0, 0)$, $X_2 = (0, 1, 0)$ for $T_{\mathbf{p}}S_r$.

For this we must find curves $c_1(t)$, $c_2(t)$ on S_r with $c_j(0) = \mathbf{p}$ and $c_j'(0) = X_j$. We can take

$$c_1(t) = (t, 0, t^2), \quad c_2(t) = (0, t, rt^2).$$

Now

$$A(X_j) = -d\xi(\mathbf{p})[X_j] = -(\xi \circ c_j)'(0),$$

so

$$A(X_1) = \frac{d}{dt}|_{t=0} \frac{(2t, 0, -1)}{\sqrt{4t^2 + 1}} = (2, 0, 0),$$

and

$$A(X_2) = \frac{d}{dt}|_{t=0} \frac{(0, 2rt, -1)}{\sqrt{4r^2t^2 + 1}} = (0, 2r, 0).$$

Then

$$\langle A(X_1), X_1 \rangle = 2, \quad \langle A(X_1), X_2 \rangle = 0, \quad \langle A(X_2), X_2 \rangle = 2r.$$

Thus $A_{\mathbf{p}}$ is represented by the diagonal matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & 2r \end{pmatrix}$$

from which it follows that X_1 and X_2 are eigenvectors of $A_{\mathbf{p}}$, with corresponding eigenvalues 2 and $2r$ respectively.

3.5. **Normal Curvature and Principal Curvatures.** For a smooth surface S with unit normal $\xi(\mathbf{p})$ at $\mathbf{p} \in S$, we say a plane $\Pi_{\mathbf{p}} \subset \mathbb{R}^3$ is a **normal plane** at $\mathbf{p}$ if Π_p contains the **normal line** $\mathbf{p} + t\xi(\mathbf{p})$. Each normal plane intersects S along a **normal section** $\Pi_{\mathbf{p}} \cap S$, which is an unparameterised smooth curve on S. It is not hard to see that for every unit tangent vector $T \in T_{\mathbf{p}}S$ there is a normal plane $\Pi_{\mathbf{p}}$ whose normal section is tangent to T: simply take

$$\Pi_{\mathbf{p}} = \{\mathbf{p} + uT + v\xi(\mathbf{p}) : u, v \in \mathbb{R}\}.$$

As an oriented plane (with orientation $T \times \xi(\mathbf{p})$), $\Pi_{\mathbf{p}}$ is uniquely determined by the unit tangent vector T.

It is important to realise that although $\Pi_{\mathbf{p}}$ is a normal plane to S at $\mathbf{p}$, $\Pi_{\mathbf{p}}$ is not necessarily normal to S at other points of $\Pi_{\mathbf{p}} \cap S$. So, in general, normal sections at $\mathbf{p}$ will not be normal sections at other points. For a unit tangent vector $T \in T_{\mathbf{p}}S$ its normal

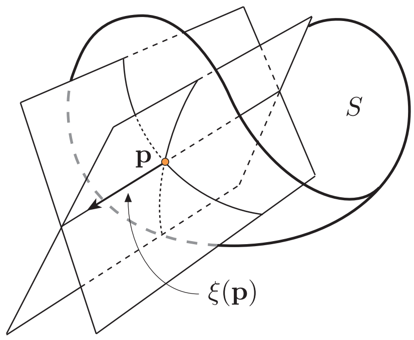

FIGURE 13. Normal sections

section $\Pi_{\mathbf{p}} \cap S$ has a unique unit speed parametrisation $c(t)$ satisfying:

$$c(0) = \mathbf{p}, \quad c'(0) = T, \quad c(t) \in \Pi_{\mathbf{p}} \cap S, \quad |c'(t)| = 1.$$

Definition 3.11. *We define the* ***normal curvature*** *of this (oriented) normal section to be*

$$\kappa_n(T) = c''(0) \bullet \xi(\mathbf{p}) = \boldsymbol{k} \bullet \xi(\mathbf{p}). \tag{3.7}$$

This depends only on T, since T completely determines c. It can be easily shown to equal the signed curvature of the planar curve $c(t)$ (i.e., as a curve in the oriented plane $\Pi_{\mathbf{p}}$) since the choice of orientation makes $\xi(\mathbf{p})$ the intrinsic normal for $c(t)$.

Lemma 3.12. *Given a unit vector $T \in T_{\mathbf{p}}S$, the normal curvature of the normal section through $\mathbf{p}$ which it determines satisfies*

$$\kappa_n(T) = \langle T, A_{\mathbf{p}}(T)\rangle, \tag{3.8}$$

where $A_{\mathbf{p}}$ is the shape operator on $T_{\mathbf{p}}S$.

Proof. First, $c'(t) \bullet \xi(c(t)) = 0$ and therefore (by differentiation)

$$c''(t) \bullet \xi(c(t)) + c'(t) \bullet (\xi \circ c)'(t) = 0.$$

So, from the definition above, along $T = c'(t)$,

$$\begin{aligned}\kappa_n(T) &= c''(t) \bullet \xi(c(t)) \\ &= -c'(t) \bullet (\xi \circ c)'(t) \\ &= -c'(t) \bullet d\xi[c'(t)], \\ &= \langle c'(t), A(c'(t))\rangle.\end{aligned}$$

□

The equation (3.8) gives a geometric interpretation of the shape operator as a quadratic form. Further, since $A_{\mathbf{p}}$ is a symmetric operator it is diagonalisable, with real eigenvalues κ_1, κ_2: there exists an orthonormal basis (Z_1, Z_2) of $T_{\mathbf{p}}S$ for which Z_1 and Z_2 are eigenvectors of $A_{\mathbf{p}}$:

$$A_{\mathbf{p}}(Z_1) = \kappa_1(\mathbf{p})Z_1, \qquad A_{\mathbf{p}}(Z_2) = \kappa_2(\mathbf{p})Z_2.$$

Then $\kappa_i = \langle Z_i, A(Z_i)\rangle = \kappa_n(Z_i)$.

Definition 3.13. *The eigenvalues $\kappa_i(\mathbf{p})$ are called the **principal curvatures** (of S at $\mathbf{p}$), and the unit eigenvectors $\pm Z_i$ are called the **principal directions** (of S at $\mathbf{p}$).*

Remark 3.3. If the orientation of S is reversed (i.e. we choose $-\xi$ instead of ξ) then the shape operator changes sign and hence so do κ_1 and κ_2.

Theorem 3.14 (Euler)**.** *The principal curvatures (of S at $\mathbf{p}$) are the maximum and minimum normal curvatures (of S at $\mathbf{p}$).*

We should keep in mind that it is possible for $\kappa_1 = \kappa_2$, in which case $\kappa_n(T) = \kappa_1$ for all T and every direction is a principal direction.

Proof. Any unit vector $T \in T_{\mathbf{p}}S$ may be written

$$T = (\cos\theta)Z_1 + (\sin\theta)Z_2,$$

for some $\theta \in [0, 2\pi)$. Then

$$A(T) = (\cos\theta)\kappa_1 Z_1 + (\sin\theta)\kappa_2 Z_2,$$

so

$$\kappa_n(T) = \langle T, A(T)\rangle = (\cos^2\theta)\kappa_1 + (\sin^2\theta)\kappa_2,$$

from which it follows that κ_n always lies between κ_1 and κ_2. □

Remark 3.4. Euler's Theorem is nothing other than an application of the more general statement that any real quadratic form $q(x, y) = ax^2 + 2bxy + cy^2$, when restricted to the unit circle $x^2 + y^2 = 1$, takes its maximum and minimum values on the eigenvectors of the symmetric matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

which corresponds to it. It is also a general fact that when A has distinct eigenvalues the eigenvectors are orthogonal, hence for $\kappa_1 \neq \kappa_2$ the principal directions are orthogonal.

3.6. **Gaussian and Mean Curvatures.** Given any basis E_1, E_2 for the tangent space $T_{\mathbf{p}}S$ the shape operator $A_{\mathbf{p}}$ is represented by a matrix $\mathbf{A}$ defined by the property that, for any $X = X^1E_1 + X^2E_2 \in T_{\mathbf{p}}S$,

$$\mathbf{A} \begin{pmatrix} X^1 \\ X^2 \end{pmatrix} = \begin{pmatrix} U^1 \\ U^2 \end{pmatrix}$$

for $AX = U^1E_1 + U^2E_2$.

By standard linear algebra, a change of this basis E_1, E_2 transforms the matrix $\mathbf{A}$ into $P^{-1}\mathbf{A}P$, where P is an invertible 2×2 matrix determined by the expressions for the new basis in terms of the old basis. The **invariants** of this tranformation will be geometric information, and not basis dependent. Recall that

$$\operatorname{tr}(P^{-1}\mathbf{A}P) = \operatorname{tr}(\mathbf{A}), \quad \text{and} \quad \det(P^{-1}\mathbf{A}P) = \det(\mathbf{A}),$$

so these are geometric information, and depend only on the shape operator itself. Therefore we may write these as $\operatorname{tr}(A)$ and $\det(A)$. Of course, we may assume $\mathbf{A}$ is the diagonal form, whose diagonal entries are the principal curvatures (in any order).

Definition 3.15. *The **Gaussian curvature** of S at $\mathbf{p}$ is defined to be*

$$K(\mathbf{p}) = \det(A_{\mathbf{p}}) = \kappa_1(\mathbf{p})\kappa_2(\mathbf{p}),$$

*and the **mean curvature** of S at $\mathbf{p}$ is defined to be*

$$H(\mathbf{p}) = \tfrac{1}{2}\operatorname{tr}(A_{\mathbf{p}}) = \tfrac{1}{2}(\kappa_1(\mathbf{p}) + \kappa_2(\mathbf{p})).$$

The Gaussian and mean curvatures together carry all the information carried by the principal curvature. This is because the characteristic polynomial of A (whose roots are κ_1, κ_2) is

$$\chi_A(\lambda) = \lambda^2 - 2H\lambda + K,$$

and therefore

$$\kappa_1, \kappa_2 = H \pm \sqrt{H^2 - K}. \tag{3.9}$$

Notice that in particular $H^2 - K \geq 0$.

3.7. **Geometric classification of points on a surface.** The points of S are classified according to the relative signs of the principal curvatures. There are four mutually exclusive classes.

Definition 3.16. *We say that* $\mathbf{p} \in S$ *is:*

- *an* ***elliptic point*** *if* κ_1, κ_2 *are non-zero and have the same sign at* $\mathbf{p}$*.*
- *a* ***hyperbolic point*** *if* κ_1, κ_2 *are non-zero and have opposite signs at* $\mathbf{p}$*.*
- *a* ***parabolic point*** *if precisely one of* κ_1, κ_2 *vanishes at* $\mathbf{p}$*.*
- *a* ***planar point*** *if* $\kappa_1 = \kappa_2 = 0$ *at* $\mathbf{p}$*.*

This point classification is independent of the orientation of S.

The terminology arises from the special case when S is a paraboloid. From Example 3.8, the principal curvatures of the paraboloid $S = S_r$ at $\mathbf{p} = (0, 0, 0)$ are 2 and $2r$; so $\mathbf{p}$ is an elliptic (resp. hyperbolic) point precisely when S is an elliptic (resp. hyperbolic) paraboloid, and $\mathbf{p}$ is a parabolic point if and only if S is a parabolic cylinder. The paradigm surface for locating planar points is of course any plane, all of whose points are planar.

The geometric point classification of S may be achieved by inspecting the Gaussian and mean curvatures:

$$\begin{aligned}
\mathbf{p} \text{ elliptic} &\iff K(\mathbf{p}) > 0; \\
\mathbf{p} \text{ hyperbolic} &\iff K(\mathbf{p}) < 0; \\
\mathbf{p} \text{ parabolic} &\iff K(\mathbf{p}) = 0,\ H(\mathbf{p}) \neq 0; \\
\mathbf{p} \text{ planar} &\iff K(\mathbf{p}) = 0 = H(\mathbf{p}).
\end{aligned}$$

There are some additional geometric conditions which may apply at some (or all) points $\mathbf{p}$ of S.

Definition 3.17. *We say that* $\mathbf{p} \in S$ *is:*

- *an* ***umbilic point*** *if* $\kappa_1(\mathbf{p}) = \kappa_2(\mathbf{p})$*; equivalently, if* $H(\mathbf{p})^2 - K(\mathbf{p}) = 0$*;*
- *a* ***minimal point*** *if* $\kappa_1(\mathbf{p}) = -\kappa_2(\mathbf{p})$*; equivalently, if* $H(\mathbf{p}) = 0$*;*
- *a* ***flat point*** *if* $\kappa_1(\mathbf{p}) = 0$ *or* $\kappa_2(\mathbf{p}) = 0$*; equivalently, if* $K(\mathbf{p}) = 0$*.*

We say that S *is an* ***umbilic*** *(resp.* ***minimal****, resp.* ***flat****)* ***surface*** *if all points of* S *are umbilic (resp. minimal, resp. flat).*

Remark 3.5. A minimal surface gets its name from the fact that on any such surface a small enough bounded patch of the surface minimises area amongst all surfaces with the same boundary.

This geometric point classification can be summarised in the following table.

Point type	**Principal curvatures**	**Gaussian & mean curvatures**
elliptic	$\kappa_1, \kappa_2 \neq 0$ have same sign	$K > 0$
hyperbolic	$\kappa_1, \kappa_2 \neq 0$ have opposite signs	$K < 0$
parabolic	precisely one κ_i vanishes	$K = 0,\ H \neq 0$
planar	$\kappa_1 = 0 = \kappa_2$	$K = 0 = H$
flat	at least one κ_i vanishes	$K = 0$
umbilic	$\kappa_1 = \kappa_2$	$H^2 - K = 0$
minimal	$\kappa_1 = -\kappa_2$	$H = 0$

Example 3.9 (Circular Cylinder). Let S be the circular cylinder with equation $x^2+y^2 = R^2$, where $R > 0$. Orient this by outward normal

$$\xi(x, y, z) = \frac{1}{R}\,(x, y, 0).$$

This is the restriction of a linear map on $\mathbb{R}^3$, so

$$A_{\mathbf{p}}(X) = -d\xi(\mathbf{p})[X] = -\frac{1}{R}\,(X_1, X_2, 0), \text{ for } X = (X_1, X_2, X_3).$$

Notice that $Z_1 = (0, 0, 1) \in T_{\mathbf{p}}S$ for all $\mathbf{p} \in S$, and $A(Z_1) = 0$, so 0 is an eigenvalue. Thus Z_1 is a principal direction, with principal curvature $\kappa_1 = 0$. Since $Z_2 \perp Z_1$ it follows that $Z_2 = (X_1, X_2, 0)$ where $\xi \bullet Z_2 = xX_1 + yX_2 = 0$. Then:

$$A(Z_2) = -\frac{1}{R}\,Z_2.$$

Hence $\kappa_2 = -1/R$, so:

$$K = 0 \quad \text{and} \quad H = -\frac{1}{2R}.$$

Thus S is flat, but not planar.

3.8. **The Second Fundamental Form.** Given an oriented surface S with shape operator A we have seen, in the Shape Lemma 3.10, that $\langle X, A(Y)\rangle$ is a symmetric bilinear form on each tangent space.

Definition 3.18. *The* ***second fundamental form*** *of S at $\mathbf{p}$ is the symmetric bilinear form*

$$\alpha_{\mathbf{p}} : T_{\mathbf{p}}S \times T_{\mathbf{p}}S \to \mathbb{R}; \quad \alpha_{\mathbf{p}}(X, Y) = \langle X, A_{\mathbf{p}}(Y)\rangle.$$

Just like the first fundamental form we can represent this by a symmetric matrix in any coordinate chart (D, φ). We choose this so that its local parametrisation $p = \varphi^{-1}$ is compatible with the orientation, i.e.

$$\xi(p(u, v)) = \frac{p_u \times p_v}{|p_u \times p_v|}.$$

Define:

$$e = \alpha(p_u, p_u), \qquad f = \alpha(p_u, p_v) = \alpha(p_v, p_u), \qquad g = \alpha(p_v, p_v),$$

which are called the **components of the second fundamental form** in the given chart. They are analogous to the components E, F, G of the Riemannian metric. For example, if $X = X^u p_u + X^v p_v$ and $Y = Y^u p_u + Y^v p_v$ then

$$\alpha(X, Y) = \begin{pmatrix} X^u & X^v \end{pmatrix} \begin{pmatrix} e & f \\ f & g \end{pmatrix} \begin{pmatrix} Y^u \\ Y^v \end{pmatrix}$$

Furthermore, e, f, g may be readily computed.

Lemma 3.19. *In an orientation preserving parametrisation $p : U \to S$ the components of the second fundamental form are given by*

$$e = \frac{[p_u, p_v, p_{uu}]}{\sqrt{EG - F^2}}, \quad f = \frac{[p_u, p_v, p_{uv}]}{\sqrt{EG - F^2}}, \quad g = \frac{[p_u, p_v, p_{vv}]}{\sqrt{EG - F^2}}. \tag{3.10}$$

[Recall that $[\mathbf{u}, \mathbf{v}, \mathbf{w}] = (\mathbf{u} \times \mathbf{v}) \bullet \mathbf{w}$, the scalar triple product.]

Proof. The proof is by straightforward computation, exploiting the identities

$$(\xi \circ p) \bullet p_u = 0 = (\xi \circ p) \bullet p_v.$$

By taking partial derivatives of these we see that

$$\begin{aligned} e &= \langle p_u, A(p_u) \rangle = -p_u \bullet (\xi \circ p)_u = (\xi \circ p) \bullet p_{uu}, \\ f &= \langle p_u, A(p_v) \rangle = -p_u \bullet (\xi \circ p)_v = (\xi \circ p) \bullet p_{uv}, \\ g &= \langle p_v, A(p_v) \rangle = -p_v \bullet (\xi \circ p)_v = (\xi \circ p) \bullet p_{vv}. \end{aligned}$$

Now (3.10) follow from the identity

$$\xi \circ p = \frac{p_u \times p_v}{|p_u \times p_v|} = \frac{p_u \times p_v}{\sqrt{EG - F^2}},$$

in which we have used

$$|p_u \times p_v|^2 = |p_u|^2 \, |p_v|^2 - (p_u \bullet p_v)^2 = EG - F^2.$$

□

The Gaussian and mean curvatures can now be directly computed from the components of the first and second fundamental forms.

Theorem 3.20. *Let S be an oriented surface whose first and second fundamental forms have components (E, F, G) and (e, f, g) respectively in some orientation preserving coordinate chart. Then the Gaussian and mean curvatures in that chart are given by*

$$K = \frac{eg - f^2}{EG - F^2}, \quad H = \frac{eG - 2fF + gE}{2(EG - F^2)}. \tag{3.11}$$

Proof. To compute K and H, we need the determinant and trace of A. Suppose the matrix of A with respect to the basis (p_u, p_v) is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \text{where} \qquad \begin{aligned} A(p_u) &= ap_u + cp_v, \\ A(p_v) &= bp_u + dp_v. \end{aligned} \tag{3.12}$$

Then $\det A = ad - bc$ and $\operatorname{tr} A = a + d$. In order to express a, b, c, d in terms of e, f, g we note that

$$\begin{aligned}
\begin{pmatrix} e & f \\ f & g \end{pmatrix} &= \begin{pmatrix} \langle p_u, A(p_u)\rangle & \langle p_u, A(p_v)\rangle \\ \langle p_v, A(p_u)\rangle & \langle p_v, A(p_v)\rangle \end{pmatrix} \\
&= \begin{pmatrix} a|\,p_u|^2 + c\langle\, p_v, p_u\rangle & b|\,p_u|^2 + d\langle\, p_v, p_u\rangle \\ a\langle\, p_u, p_v\rangle + c|\,p_v|^2 & b\langle\, p_u, p_v\rangle + d|\,p_v|^2 \end{pmatrix} \qquad \text{by (3.12)} \\
&= \begin{pmatrix} aE + cF & bE + dF \\ aF + cG & bF + dG \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}
\end{aligned}$$

Therefore:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} e & f \\ f & g \end{pmatrix} \tag{3.13}$$

Note that

$$\det \begin{pmatrix} E & F \\ F & G \end{pmatrix} = EG - F^2 = |p_u \times p_v|^2 \neq 0,$$

so the inverse matrix exists. Hence

$$K = \det A = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det \begin{pmatrix} e & f \\ f & g \end{pmatrix} \Big/ \det \begin{pmatrix} E & F \\ F & G \end{pmatrix} = \frac{eg - f^2}{EG - F^2}$$

Notice that the multiplicative property of determinants allowed us to achieve this without having to explicitly perform the matrix inversion in (3.13). However the trace is **not** multiplicative, so we now need to develop (3.13)

$$\begin{aligned}
\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \frac{1}{EG - F^2} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix} \begin{pmatrix} e & f \\ f & g \end{pmatrix} \\
&= \frac{1}{EG - F^2} \begin{pmatrix} eG - fF & \star \\ \star & gE - fF \end{pmatrix}
\end{aligned}$$

Hence

$$H = \tfrac{1}{2} \operatorname{tr} A = \frac{a + d}{2} = \frac{eG - 2fF + gE}{2(EG - F^2)}$$

□

3.9. **Gauss's Theorema Egregium.** The classical approach to the definition of the Gaussian curvature, given above, is via the shape operator, which measures normal curvatures. The normal curvatures measure **extrinisic** geometry of the surface: they rely on the way the surface sits in the ambient Euclidean space. Gauss's Theorema Egregium ("Remarkable Theorem") states that despite this the Gaussian curvature $K = \det(A) = \kappa_1 \kappa_2$ relies only on the intrinsic geometry determined by the surface's metric.

Theorema Egregium. *Let S, $\bar{S}$ be smooth surfaces with Gaussian curvatures K, $\bar{K}$ respectively. If $f\colon S \to \bar{S}$ is a local isometry then $\bar{K}(f(\mathbf{p})) = K(\mathbf{p})$ for all $\mathbf{p} \in S$.*

Proof. Our aim is to show that K can be expressed entirely in terms of E, F, G, the components of the metric for S, hence $\bar{K}$ can be expressed entirely in terms of $\bar{E}, \bar{F}, \bar{G}$, the components of the metric for $\bar{S}$. In that case, since by the (E, F, G) Lemma 3.6 a local isometry $f : S \to \bar{S}$ has the property that $(E, F, G) = (\bar{E}, \bar{F}, \bar{G})$ in f-adapted charts, it follows that $K \circ f = \bar{K}$.

We start from the expression (3.11) for K, which we rewrite as

$$W^4 K = [\, p_u, p_v, p_{uu}]\,[\, p_u, p_v, p_{vv}] - [p_u, p_v, p_{uv}]^2,$$

where $W = \sqrt{EG - F^2}$ and we have used (3.10) for e, f, g. For any vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ we write $(\mathbf{a}\ \mathbf{b}\ \mathbf{c})$ for the 3×3 matrix with columns $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and note that the matrix

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{pmatrix},$$

with rows $\mathbf{a}, \mathbf{b}, \mathbf{c}$, is its transpose and therefore has the same determinant. Then we can write:

$$\begin{aligned}
W^4 K &= \det\begin{pmatrix} p_u & p_v & p_{uu} \end{pmatrix} \det\begin{pmatrix} p_u & p_v & p_{vv} \end{pmatrix} - \det\begin{pmatrix} p_u & p_v & p_{uv} \end{pmatrix}^2 \\
&= \det\begin{pmatrix} p_u \\ p_v \\ p_{uu} \end{pmatrix} \det\begin{pmatrix} p_u & p_v & p_{vv} \end{pmatrix} - \det\begin{pmatrix} p_u \\ p_v \\ p_{uv} \end{pmatrix} \det\begin{pmatrix} p_u & p_v & p_{uv} \end{pmatrix} \\
&= \det\left(\begin{pmatrix} p_u \\ p_v \\ p_{uu} \end{pmatrix} \begin{pmatrix} p_u & p_v & p_{vv} \end{pmatrix}\right) - \det\left(\begin{pmatrix} p_u \\ p_v \\ p_{uv} \end{pmatrix} \begin{pmatrix} p_u & p_v & p_{uv} \end{pmatrix}\right) \\
&= \begin{vmatrix} E & F & p_u \bullet p_{vv} \\ F & G & p_v \bullet p_{vv} \\ p_u \bullet p_{uu} & p_v \bullet p_{uu} & p_{uu} \bullet p_{vv} \end{vmatrix} - \begin{vmatrix} E & F & p_u \bullet p_{uv} \\ F & G & p_v \bullet p_{uv} \\ p_u \bullet p_{uv} & p_v \bullet p_{uv} & |p_{uv}|^2 \end{vmatrix} \qquad (3.14) \\
&= \begin{vmatrix} E & F & p_u \bullet p_{vv} \\ F & G & p_v \bullet p_{vv} \\ p_u \bullet p_{uu} & p_v \bullet p_{uu} & p_{uu} \bullet p_{vv} - |p_{uv}|^2 \end{vmatrix} - \begin{vmatrix} E & F & p_u \bullet p_{uv} \\ F & G & p_v \bullet p_{uv} \\ p_u \bullet p_{uv} & p_v \bullet p_{uv} & 0 \end{vmatrix},
\end{aligned}$$

where to obtain the last line we have used the observation that both determinants in (3.14) have the same top left hand 2×2 block and therefore it is possible to move the bottom right hand corner element (viz. $|p_{uv}|^2$) of the second determinant across to the first determinant. (To see this, simply expand the determinants along their final column, or row.)

Now we claim that all the entries involving second order derivatives of $p(u,v)$ can be expressed in terms of E, F, G are their derivatives. The computations are:

$$\begin{aligned}
p_{uu} \bullet p_u &= \tfrac{1}{2}(p_u \bullet p_u)_u = \tfrac{1}{2}E_u,\\
p_{uu} \bullet p_v &= (p_u \bullet p_v)_u - p_u \bullet p_{vu} = F_u - \tfrac{1}{2}(p_u \bullet p_u)_v = F_u - \tfrac{1}{2}E_v,\\
p_{uv} \bullet p_u = p_{vu} \bullet p_u &= \tfrac{1}{2}(p_u \bullet p_u)_v = \tfrac{1}{2}E_v,\\
p_{uv} \bullet p_v = p_{vu} \bullet p_v &= \tfrac{1}{2}(p_v \bullet p_v)_u = \tfrac{1}{2}G_u,\\
p_{vv} \bullet p_u &= (p_v \bullet p_u)_v - p_v \bullet p_{uv} = F_v - \tfrac{1}{2}(p_v \bullet p_v)_u = F_v - \tfrac{1}{2}G_u,\\
p_{vv} \bullet p_v &= \tfrac{1}{2}(p_v \bullet p_v)_v = \tfrac{1}{2}G_v.
\end{aligned}$$

Finally, from these it follows that

$$\begin{aligned}
p_{uu} \bullet p_{vv} &= (p_u \bullet p_{vv})_u - p_u \bullet p_{vvu} = (F_v - \tfrac{1}{2}G_u)_u - p_u \bullet p_{vvu},\\
|p_{uv}|^2 &= (p_u \bullet p_{uv})_v - p_u \bullet p_{uvv} = \tfrac{1}{2}E_{vv} - p_u \bullet p_{vvu}.
\end{aligned}$$

The equality of the third order partial derivatives gives

$$p_{uu} \bullet p_{vv} - |p_{uv}|^2 = F_{uv} - \tfrac{1}{2}G_{uu} - \tfrac{1}{2}E_{vv}.$$

□

Remark 3.6.

(i) This proof shows that one can write down an explicit expression for the Gaussian curvature purely in terms of the metric components E, F, G. Such expressions were first published by Brioschi, in 1852, 24 years after Gauss gave the first proof of his Remarkable Theorem. Our proof follows Brioschi's: Gauss's is not so easy to follow.

(ii) By contrast, the mean curvature H is not an intrinsic quantity. This follows at once from the fact that the plane and the cylinder in Example 3.4 are local isometric but have different mean curvatures.

3.10. **Example: curvatures for the Circular Torus.** We will find the Gaussian and mean curvature of the circular torus of revolution $S = T^2(a,b)$ where $0 < b < a$. This surface is obtained by rotating around the z-axis a circle in the y,z-plane, of radius b and centre $y = a$, $z = 0$. The whole surface can be described using cylindrical polar coordinates (r, θ, z), where $r^2 = x^2 + y^2$, as

$$T(a,b) = \{(r,\theta,z) : (r-a)^2 + z^2 = b^2\}.$$

We will make our calculations using the following local parametrisation:

$$p(u,v) = \big(\cos u(a + b\cos v),\, \sin u(a + b\cos v),\, b\sin v\big), \quad 0 < u, v < 2\pi.$$

From this we will determine the geometric point classification of S, locate special points, and compute the principal curvatures.

We have:

$$p_u = (a + b\cos v)(-\sin u,\, \cos u,\, 0), \qquad p_v = b(-\cos u \sin v,\, -\sin u \sin v,\, \cos v),$$

and

$$\begin{aligned} p_{uu} &= -(a+b\cos v)(\cos u,\, \sin u,\, 0), \\ p_{uv} &= b\sin v(\sin u,\, -\cos u,\, 0) \;=\; p_{vu}, \\ p_{vv} &= -b(\cos u\cos v,\, \sin u\cos v,\, \sin v). \end{aligned}$$

Therefore

$$E = |p_u|^2 = (a+b\cos v)^2, \qquad F = p_u \bullet p_v = 0, \qquad G = |p_v|^2 = b^2,$$

so that $\sqrt{EG-F^2} = b(a+b\cos v)$. Hence by (3.10)

$$\begin{aligned} e &= \frac{[p_u, p_v, p_{uu}]}{\sqrt{EG-F^2}} = \frac{-b(a+b\cos v)^2}{b(a+b\cos v)} \begin{vmatrix} -\sin u & \star & \cos u \\ \cos u & \star & \sin u \\ 0 & \cos v & 0 \end{vmatrix} \\ &= -\cos v(a+b\cos v), \end{aligned}$$

while

$$f = \frac{[p_u, p_v, p_{uv}]}{\sqrt{EG-F^2}} = 0,$$

since p_{uv} and p_u are linearly dependent, and

$$\begin{aligned} g &= \frac{[p_u, p_v, p_{vv}]}{\sqrt{EG-F^2}} = \frac{-b^2(a+b\cos v)}{b(a+b\cos v)} \begin{vmatrix} -\sin u & -\cos u\sin v & \cos u\cos v \\ \cos u & -\sin u\sin v & \sin u\cos v \\ 0 & \cos v & \sin v \end{vmatrix} \\ &= -b\left(-\cos^2 v \begin{vmatrix} -\sin u & \cos u \\ \cos u & \sin u \end{vmatrix} + \sin^2 v \begin{vmatrix} -\sin u & -\cos u \\ \cos u & -\sin u \end{vmatrix} \right) \\ &= -b. \end{aligned}$$

So

$$eg - f^2 = b\cos v(a+b\cos v),$$

Hence by (3.11)

$$K = \frac{eg-f^2}{EG-F^2} = \frac{\cos v}{b(a+b\cos v)}$$

and

$$H = \frac{eG - 2fF + gE}{2(EG-F^2)} = \frac{-(a+2b\cos v)}{2b(a+b\cos v)}$$

Notice that K and H only depend on the latitude (v), as might be expected on a surface of revolution. Notice also that $a + b\cos v > 0$. Since K and H are smooth functions on the whole of S, we can obtain their values at the edges of the parametrisation $p(u,v)$ by continuity (taking limits).

The geometric point classification of S is primarily determined by the sign of K

$$\begin{aligned} &K > 0, && \text{if } 0 \le v < \pi/2, \text{ or } 3\pi/2 < v \le 2\pi; \\ &K < 0, && \text{if } \pi/2 < v < 3\pi/2. \end{aligned}$$

Thus the "outside" of S is elliptic, whereas the "inside" is hyperbolic. Furthermore S has flat points ($K = 0$) on the "top" and "bottom" circles of latitude $v = \pi/2$ and $v = 3\pi/2$. Since $H = -1/2b \neq 0$ at these latitudes, points on these circles are parabolic. Of the

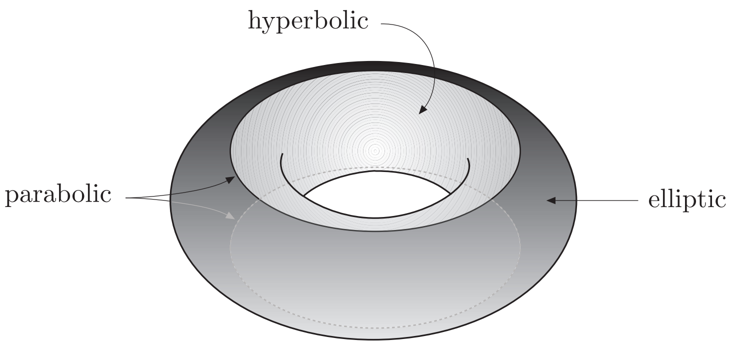

FIGURE 14. Geometric point classification of the torus.

possible special geometric point types, minimal points occur when $\cos v = -a/2b$, which is only possible if $b \geq a/2$. If $b = a/2$ the minimal points constitute the "inner equator" ($v = \pi$) of S; otherwise they constitute the pair of latitudes $v = \pi \pm \arccos(a/2b)$, which lie in the hyperbolic region of S. To investigate the possibility of umbilic points, we compute:

$$H^2 - K = \frac{a^2}{4b^2(a + b\cos v)^2} > 0.$$

Thus there are no umbilic points. Finally, the principal curvatures may be determined using (3.9)

$$\kappa_1, \kappa_2 = H \pm \sqrt{H^2 - K} = -\frac{\cos v}{a + b\cos v}, \ -\frac{1}{b}.$$

3.11. **The Geometry of Curves on a Surface.** Let $c(t)$ be a regular smooth path in a smooth surface S, with unit tangent vector $T(t) = c'(t)/|c'(t)| \in T_{c(t)}S$.

Definition 3.21 (Darboux Frame)**.** *Let ξ be the unit normal field (Gauss map) on S. Along the curve $c(t)$ in S we define*

$$V(t) = \xi(c(t)), \qquad U(t) = V(t) \times T(t). \tag{3.15}$$

*Notice that $U(t) \in T_{c(t)}S$. It is called the **intrinsic normal** to $c(t)$. The triple of vectors $(T(t), U(t), V(t))$ is a positively oriented (i.e. right-handed) orthonormal basis of $\mathbb{R}^3$, called the **Darboux frame** along $c(t)$.*

The intrinsic normal may or may not agree with the principal normal $N(t)$ of $c(t)$, and so one does not expect the Darboux frame to agree with the Frenet frame. In fact the Darboux frame is defined at every point on $c(t)$, including points of inflection (where $N(t)$ is not defined).

Recall that $c(t)$ has curvature vector $\boldsymbol{k}(t)$ orthogonal to $T(t)$. Therefore

$$\boldsymbol{k} = (\boldsymbol{k} \bullet U)U + (\boldsymbol{k} \bullet V)V. \tag{3.16}$$

The second term here generalises the normal curvature of a normal section.

Definition 3.22. *In the previous decomposition, the component* $(\boldsymbol{k} \bullet U)U$ *is called the* ***geodesic curvature vector,*** *and* $\kappa_g = \boldsymbol{k} \bullet U$ *is called the* ***geodesic curvature****. If* κ_g *vanishes identically then* $c(t)$ *is called a* ***geodesic (curve)*** *of* S*.*

The component $(\boldsymbol{k} \bullet V)V$ *is called the* ***normal curvature vector,*** *and* $\kappa_n = \boldsymbol{k} \bullet V$ *is called the* ***normal curvature****. If* κ_n *vanishes identically then* $c(t)$ *is said to be an* ***asymptotic curve*** *of* S*.*

If the orientation of S is reversed then both U and V, and hence κ_g and κ_n, change sign.

Remark 3.7. Geodesics are a very important family of curves on a surface, but sadly we don't have time in this course to study them properly. The condition $\kappa_g = 0$ says that the geodesic is doing the least amount of turning possible to stay on the surface. It turns out that this makes geodesics locally distance minimising, i.e., given two points on a geodesic which are close enough together, that segment of the geodesic has shortest length over all curves on the surface with the same end points. This is the reason why they are of such interest: they generalise the idea of a straight line to "curved" spaces.

The geodesic and normal curvatures can be explicitly calculated from the expression for the curvature vector.

Lemma 3.23. *For a curve* $c(t)$ *on a surface* S*, not necessarily unit speed, the geodesic curvature and normal curvature are given by*

$$\kappa_g = \frac{[c', c'', V]}{|c'|^3}, \qquad \kappa_n = \frac{c'' \bullet V}{|c'|^2}. \tag{3.17}$$

Further, the curvature $\kappa = |\boldsymbol{k}|$ *of* $c(t)$ *satisfies* $\kappa^2 = \kappa_g^2 + \kappa_n^2$*.*

Proof. The equation $\kappa^2 = \kappa_g^2 + \kappa_n^2$ follows immediately from (3.16) since U, V are unit vectors. Now we recall from earlier the expression

$$\boldsymbol{k} = \frac{1}{|c'|^2} c'' - \frac{c' \bullet c''}{|c'|^4} c'.$$

Since $U = V \times T$ and $c' \times T = 0$ we notice that

$$c' \bullet U = [c', V, T] = [V, T, c'] = 0,$$

Thus

$$\kappa_g = \boldsymbol{k} \bullet U = \frac{1}{|c'|^2}[c'', V, T] = \frac{1}{|c'|^3}[c', c'', V],$$

since $T = c'/|c'|$. Similarly, $c' \bullet V = 0$ so that

$$\kappa_n = \boldsymbol{k} \bullet V = \frac{1}{|c'|^2} c'' \bullet V.$$

□

The geodesic curvature is the amount of curvature "perceivable" in S, and geodesics are therefore the analogues in S of straight lines in the Euclidean plane. They play an essential rôle in any deeper investigation of the Riemannian geometry of S.

Example 3.10. Let S be the helicoid from Example 3.3, with the standard parametrisation $p(u,v) = (u\cos v, u\sin v, av)$. The u coordinate lines are the straight lines with equation, for each $v \in \mathbb{R}$,

$$c_v(t) = (t\cos(v), t\sin(v), av).$$

Every point of the helicoid lies on one of these lines, so it is called a **ruled surface** (and these lines are called the **rulings**). Since every straight line has $\kappa = 0$ it follows immediately from $\kappa^2 = \kappa_g^2 + \kappa_n^2$ that $\kappa_g = 0 = \kappa_n$ for these rulings. Therefore they are simultaneously geodesics and asymptotic curves.

Now fix $u \in \mathbb{R}$ and $c(t) = (u\cos t, u\sin t, at)$ be the corresponding v-coordinate line. This is a helix. We choose the orientation of S induced by the parametrisation, that is

$$\xi = \frac{p_u \times p_v}{|\, p_u \times p_v|}.$$

Since

$$p_u = (\cos v, \sin v, 0), \qquad p_v = (-u\sin v, u\cos v, a),$$

we have

$$p_u \times p_v = (a\sin v, -a\cos v, u),$$

and hence:

$$\xi = \Big(\frac{a\sin v}{\sqrt{a^2+u^2}}, \frac{-a\cos v}{\sqrt{a^2+u^2}}, \frac{u}{\sqrt{a^2+u^2}}\Big). \tag{3.18}$$

Therefore

$$V(t) = \xi(c(t)) = \Big(\frac{a\sin t}{\sqrt{a^2+u^2}}, \frac{-a\cos t}{\sqrt{a^2+u^2}}, \frac{u}{\sqrt{a^2+u^2}}\Big).$$

Since also

$$c'(t) = (-u\sin t, u\cos t, a), \qquad c''(t) = (-u\cos t, -u\sin t, 0),$$

we finally compute

$$\kappa_n = \frac{c'' \bullet V}{|c'|^2} = 0,$$

and

$$\kappa_g = \frac{[c', c'', V]}{|c'|^3} = \frac{-u}{(a^2+u^2)^2}\begin{vmatrix} -u\sin t & \cos t & a\sin t \\ u\cos t & \sin t & -a\cos t \\ a & 0 & u \end{vmatrix} = \frac{u}{a^2+u^2}.$$

Thus, all the v-coordinate lines are asymptotic curves. Notice also that each one has constant geodesic curvature, and precisely one (the axis $u = 0$ of the helicoid) is a geodesic of the helicoid (of course, since it is a Euclidean straight line).

Appendix A. Proof of the Fundamental Theorem of Space Curves.

The proof of Theorem 1.4 rests on the standard existence and uniqueness theorem for linear ordinary differential equations, which we state here without proof.

Theorem A.1. *Let $A(t)$ be a smooth 3×3 matrix-valued function on some interval $I \subset \mathbb{R}$, and $t_0 \in I$. Then for each $\mathbf{x}_0 \in \mathbb{R}^3$ the linear o.d.e. $\mathbf{x}' = \mathbf{x}A$ for the vector-valued function $\mathbf{x}(t) = (x(t), y(t), z(t))$ has a unique solution $\mathbf{x}(t)$ over $t \in I$ satisfying the initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$.*

We will make use of this in the form of the following corollary. Let $M_3^3(\mathbb{R})$ denote the set (indeed, vector space) of all 3×3 matrices with real entries.

Corollary A.2. *Let $A : I \to M_3^3(\mathbb{R})$ be a smooth matrix-valued function. For each invertible matrix $F_0 \in M_3^3(\mathbb{R})$ there exists a unique matrix-valued solution $F : I \to M_3^3(\mathbb{R})$ to the o.d.e. $F' = FA$ satisfying the initial condition $F(t_0) = F_0$. This solution $F(t)$ is invertible for all $t \in I$. If $\hat{F}(t)$ is the unique solution to the same equation with initial condition $\hat{F}(t_0) = \hat{F}_0$ then $\hat{F} = LF$ where $L = \hat{F}_0 F_0^{-1}$ is constant.*

Further, when $A^t = -A$ and $F_0 \in O(3)$ then $F(t)$ is an orthogonal matrix for all $t \in I$ (and $F(t)$ is special orthogonal if $F_0 \in SO(3)$).

The relationship between the Theorem and its Corollary is that the rows of F_0 give three linearly independent initial conditions $\mathbf{x}_1(t_0), \mathbf{x}_2(t_0), \mathbf{x}_3(t_0)$ which span $\mathbb{R}^3$, and the corresponding solutions $\mathbf{x}_1(t), \mathbf{x}_2(t), \mathbf{x}_3(t)$ must therefore span $\mathbb{R}^3$ at each $t \in I$. Since $\mathbf{x}_j' = \mathbf{x}_j A$ for each of these, the matrix F with rows $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ satisfies $F' = FA$ and is invertible for all time. It is called **a fundamental matrix solution**, since every solution to the o.d.e. can be obtained from its rows by linear combination. Now if $\hat{F}$ satisfies the same equation with a different initial condition $\hat{F}(t_0) = \hat{F}_0$, then both $F_0^{-1}F$ and $\hat{F}_0^{-1}\hat{F}$ satisfy the equation with intial condition I_3, so

$$\hat{F}_0^{-1}\hat{F} = F_0^{-1}F, \text{ i.e., } \hat{F} = \hat{F}_0 F_0^{-1} F.$$

To see the last part of the corollary, when $A^t = -A$ it follows that

$$\begin{aligned} (FF^t)' &= F'F^t + F(F')^t, \quad \text{since } (F^t)' = (F')^t, \\ &= FAF^t + FA^tF^t, \quad \text{since } (FA)^t = A^tF^t, \\ &= F(A + A^t)F^t = 0. \end{aligned}$$

So FF^t is constant. Thus $F(t_0)F(t_0)^t = I_3$ implies $FF^t = I_3$ for all time. Hence F is an orthogonal matrix. Now $F(t)$ is a continuous function of t, therefore so is $\det(F(t))$. But $\det(F(t)) = \pm 1$, so either $\det(F) = 1$ or $\det(F) = -1$ for all time. The sign is therefore determined by the initial condition $\det(F_0)$.

Proof of the Fundamental Theorem.
(i) Suppose we have two arc length parameterised space curves $p, \hat{p} : I \to \mathbb{R}^3$ with the same curvature and torsion, κ and τ. Let T, N, B and $\hat{T}, \hat{N}, \hat{B}$ be their respective Frenet frames,

with corresponding matrices F and $\hat{F}$. These are both solutions to the o.d.e. $F' = FA$ where

$$A = \begin{pmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix}. \tag{A.1}$$

By Corollary A.2 there is a constant matrix L for which $\hat{F} = LF$. Since $F, \hat{F}$ are special orthogonal, so is L. This means in particular that $\hat{T} = LT$. Now $T = p'$ and $\hat{T} = \hat{p}'$, so

$$\hat{p}(t) - \hat{p}(t_0) = \int_{t_0}^{t} \hat{T}(\tau) d\tau = \int_{t_0}^{t} LT(\tau) d\tau = L \int_{t_0}^{t} T(\tau) d\tau = L(p(t) - p(t_0)),$$

since L is constant. Therefore

$$\hat{p} = Lp + \mathbf{c}, \quad \mathbf{c} = \hat{p}(t_0) - Lp(t_0).$$

i.e., the two paths are properly congruent.
(ii) Given smooth functions $\kappa, \tau : I \to \mathbb{R}$, with $\kappa > 0$, we solve $F' = FA$ for the matrix A in (A.1). Use the columns of F to define T, N, B, then these satisfy the Frenet equations, by construction. Now define

$$p(t) = \int_{t_0}^{t} T(\tau) d\tau.$$

Then $p : I \to \mathbb{R}^3$ is a smooth path with $p' = T$, parameterised by arc length since $|T| = 1$, and with normal N and binormal B, and therefore with curvature κ and torsion τ. By part (i), any other curve with the same curvature and torsion is properly congruent to it.

Finally, let us consider the effect of an orientation **reversing** isometry $L \in O(3)$, $\det(L) = -1$, on a path p. Let $\hat{p} = Lp$. Then $\hat{p}$ has Frenet frame $\hat{T} = LT$, $\hat{N} = LN$, but $\hat{B} = -LB$, since

$$[LT, LN, LB] = \det(L)[T, N, B] = \det(L) = -1.$$

Using the Frenet formulas this means $\hat{\kappa} = \kappa$ but $\hat{\tau} = -\tau$. We deduce that two paths with the same curvature but opposite torsion are still congruent, but not **properly** congruent. □

Appendix B. Proof of the Regular Value Theorem.

The proof of the Regular Value Theorem rests largely on the following lemma.

Lemma B.1 (Regular Point Lemma)**.** *Suppose $V \subset \mathbb{R}^3$ is open and $f : V \to \mathbb{R}$ is smooth at $\mathbf{p} \in V$. If $\mathbf{p}$ is a regular point of f then there exist:*

- *a neighbourhood $A \subset V$ of $\mathbf{p}$,*
- *an open subset $B \subset \mathbb{R}^3$,*
- *a smooth diffeomorphism $\psi : B \to A$,*

such that $f(\psi(u, v, w)) = w$ for all $(u, v, w) \in B$.

Proof. For convenience define:

$$\pi_1(x,y,z) = x, \qquad \pi_2(x,y,z) = y, \qquad \pi_3(x,y,z) = z,$$

and maps $F_i : \mathbb{R}^3 \to \mathbb{R}^3$, $i = 1,2,3$:

$$\begin{aligned} F_1(x,y,z) &= \big(f(x,y,z), y, z\big), \\ F_2(x,y,z) &= \big(x, f(x,y,z), z\big), \\ F_3(x,y,z) &= \big(x, y, f(x,y,z)\big). \end{aligned}$$

Thus $\pi_i \circ F_i = f$. We claim that at least one F_i is invertible on a neighbourhood of $\mathbf{p}$. First note that:

$$J_f(\mathbf{p}) = \begin{pmatrix} f_x(\mathbf{p}) & f_y(\mathbf{p}) & f_z(\mathbf{p}) \end{pmatrix},$$

a 1×3 matrix. So, since $\mathbf{p}$ is a regular point, at least one of these partial derivatives is non-zero; say $f_x(\mathbf{p})$. Now:

$$J_{F_1}(\mathbf{p}) = \begin{pmatrix} f_x(\mathbf{p}) & f_y(\mathbf{p}) & f_z(\mathbf{p}) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

hence:

$$\det J_{F_1}(\mathbf{p}) = f_x(\mathbf{p}) \neq 0.$$

So $dF_1(\mathbf{p})$ is a linear isomorphism, and by the Inverse Function Theorem F_1 is locally smoothly invertible about $\mathbf{p}$, with inverse $F_1^{-1} : \tilde{B} \to A$ for some open sets $A \subset V$, $\tilde{B} \subset \mathbb{R}^3$. Now let $\sigma : \mathbb{R}^3 \to \mathbb{R}^3$ be the diffeomorphism which permutes coordinates in the order $\sigma(u,v,w) = (w,u,v)$, and let $B = \sigma^{-1}(\tilde{B})$. Define $\psi : B \to A$ by $\psi = F_1^{-1} \circ \sigma$, then

$$(f \circ \psi)(u,v,w) = (\pi_1 \circ F_1 \circ F_1^{-1} \circ \sigma)(u,v,w) = \pi_1(w,u,v) = w.$$

It is easy to see how to adapt this argument to the case where only $f_y(\mathbf{p})$ or $f_z(\mathbf{p})$ is non-zero, by replacing F_1 by F_2 or F_3, and σ by σ^2 or $\sigma^3 = \text{id}$. □

Proof of the Regular Value Theorem. Given $f : V \to \mathbb{R}$ with regular value k, every $\mathbf{p} \in S_k$ is regular. By the Regular Point Lemma each $\mathbf{p}$ has an open neighbourhood $A \subset V$ in which $f \circ \psi = w$ in the language of that lemma. Thus

$$D = S_k \cap A = \{\psi(u,v,w) : w = k\}.$$

is the domain for a chart $\varphi : D \to \mathbb{R}^2$, $\varphi(\psi(u,v,w)) = (u,v)$. Since every point has a chart of this type, S_k is a smooth surface.

It remains to show that $T_{\mathbf{p}}S = \ker df(\mathbf{p})$. If $X \in T_{\mathbf{p}}S$ then $X = c'(0)$ for some smooth path $c(t)$ in S with $c(0) = \mathbf{p}$. Now

$$df(\mathbf{p})[X] = (f \circ c)'(0) = 0,$$

since $(f \circ c)(t) = k$ for all t. Thus $T_{\mathbf{p}}S \subset \ker df(\mathbf{p})$. Since $\mathbf{p}$ is a regular point, the linear map $df(\mathbf{p}) : \mathbb{R}^3 \to \mathbb{R}$ has rank 1, and its kernel is therefore 2-dimensional. But $T_{\mathbf{p}}S$ is also 2-dimensional. So $T_{\mathbf{p}}S = \ker df(\mathbf{p})$. By Remark 2.4(i) this equals $\nabla f(\mathbf{p})^{\perp}$. □

Appendix C. Brioschi's intrinsic formulae for the Gaussian curvature.

For completeness we give Brioschi's general formulae for the Gaussian curvature.

Theorem C.1 (Brioschi)**.** *The Gauss curvature of a smooth surface may be calculated locally using any one of the following three intrinsic ways:*

$$2WK = \frac{\partial}{\partial u}\Big(\frac{2GF_v - FG_v - GG_u}{GW}\Big) + \frac{\partial}{\partial v}\Big(\frac{FG_u - GE_v}{GW}\Big) \tag{C.1}$$

$$2WK = \frac{\partial}{\partial u}\Big(\frac{FE_v - EG_u}{EW}\Big) + \frac{\partial}{\partial v}\Big(\frac{2EF_u - EE_v - FE_u}{EW}\Big) \tag{C.2}$$

$$\begin{aligned} 2WK &= \frac{\partial}{\partial u}\Big(\frac{F_v - G_u + F\big(\ln\sqrt{E/G}\,\big)_v}{W}\Big) \\ &\qquad + \frac{\partial}{\partial v}\Big(\frac{F_u - E_v + F\big(\ln\sqrt{G/E}\,\big)_u}{W}\Big) \end{aligned} \tag{C.3}$$

where $W = \sqrt{EG - F^2}$. If the coordinates are orthogonal ($F = 0$) then these reduce to the following formula:

$$K = -\frac{1}{2\sqrt{EG}}\left(\frac{\partial}{\partial u}\Big(\frac{G_u}{\sqrt{EG}}\Big) + \frac{\partial}{\partial v}\Big(\frac{E_v}{\sqrt{EG}}\Big)\right). \tag{C.4}$$

The aim here is to derive the formulae for the Gaussian curvature given in Theorem C.1. Expanding the right hand side and making some initial cursory groupings of terms produces the following:

$$\begin{aligned} W^4K = &-\frac{1}{2}\Big((E_{vv} + G_{uu} - 2F_{uv})W^2 \\ &+E_uF_vG + E_vFF_v + EF_uG_v + FF_uG_u - 2FF_uF_v \\ &-\frac{1}{2}\Big(E_uFG_v + E_uGG_u + EE_vG_v + E_vFG_u - E_vFG_u \\ &-E_vFG_u + (E_v)^2G + E(G_u)^2\Big)\Big) \end{aligned}$$

We now note:

$$2WW_u = (W^2)_u = E_uG + EG_u - 2FF_u, \tag{C.5}$$

and a similar formula for WW_v, allowing us to regroup as follows:

$$\begin{aligned} -2W^4K &= (E_{vv} + G_{uu} - 2F_{uv})W^2 \\ &\quad -\tfrac{1}{2}(E_uG + EG_u - 2FF_u)G_u - \tfrac{1}{2}(E_vG + EG_v - 2FF_v)E_v \\ &\quad +E_uF_vG + EF_uG_v - 2FF_uF_v - \tfrac{1}{2}E_uFG_v + \tfrac{1}{2}E_vFG_u \\ &= (E_{vv} + G_{uu} - 2F_{uv})W^2 - WW_uG_u - WW_vE_v + \mathbf{R}, \end{aligned}$$

where $\mathbf{R}$ denotes the following residual terms:

$$\mathbf{R} = E_uF_vG + EF_uG_v - 2FF_uF_v - \tfrac{1}{2}(E_uG_v - E_vG_u)F.$$

These may be written in two ways. First

$$\begin{aligned}\mathbf{R} &= (E_uG + EG_u - 2FF_u)F_v - EF_vG_u + EF_uG_v - \tfrac{1}{2}(E_uG_v - E_vG_u)F \\ &= 2WW_uF_v + E(F_uG_v - F_vG_u) - \tfrac{1}{2}F(E_uG_v - E_vG_u),\end{aligned}$$

yielding

$$\begin{aligned}-2W^4K &= W^3\left(\frac{\partial}{\partial v}\left(\frac{E_v}{W}\right) + \frac{\partial}{\partial u}\left(\frac{G_u}{W}\right) - 2\frac{\partial}{\partial u}\left(\frac{F_v}{W}\right)\right) \\ &\quad + E(F_uG_v - F_vG_u) - \tfrac{1}{2}F(E_uG_v - E_vG_u). \qquad \text{(C.6)}\end{aligned}$$

On the other hand

$$\begin{aligned}\mathbf{R} &= (E_vG + EG_v - 2FF_v)F_u + E_uF_vG - E_vF_uG - \tfrac{1}{2}(E_uG_v - E_vG_u)F \\ &= 2WW_vF_u + G(E_uF_v - E_vF_u) - \tfrac{1}{2}F(E_uG_v - E_vG_u),\end{aligned}$$

yielding

$$\begin{aligned}-2W^4K &= W^3\left(\frac{\partial}{\partial v}\left(\frac{E_v}{W}\right) + \frac{\partial}{\partial u}\left(\frac{G_u}{W}\right) - 2\,\frac{\partial}{\partial v}\left(\frac{F_u}{W}\right)\right) \\ &\quad + G(E_uF_v - E_vF_u) - \tfrac{1}{2}F(E_uG_v - E_vG_u). \qquad \text{(C.7)}\end{aligned}$$

We now require a slightly more complicated calculation.

Lemma C.2.

$$\frac{\partial}{\partial v}\Big(\frac{FE_u}{EW}\Big) - \frac{\partial}{\partial u}\Big(\frac{FE_v}{EW}\Big) = \frac{1}{W^3}\Big(G(E_uF_v - E_vF_u) - \tfrac{1}{2}F(E_uG_v - E_vG_u)\Big). \qquad \text{(C.8)}$$

$$\frac{\partial}{\partial v}\Big(\frac{FG_u}{GW}\Big) - \frac{\partial}{\partial u}\Big(\frac{FG_v}{GW}\Big) = \frac{1}{W^3}\Big(\tfrac{1}{2}F(E_uG_v - E_vG_u) - E(F_uG_v - F_vG_u)\Big). \qquad \text{(C.9)}$$

Proof. It suffices to prove (C.8). We have:

$$\begin{aligned}& E^2W^2\left(\frac{\partial}{\partial v}\Big(\frac{FE_u}{EW}\Big) - \frac{\partial}{\partial u}\Big(\frac{FE_v}{EW}\Big)\right) \\ =\ & EW(E_{uv}F + E_uF_v) - E_uF(E_vW + EW_v) \\ & -EW(E_{vu}F + E_vF_u) + E_vF(E_uW + EW_u) \\ =\ & EW(E_uF_v - E_vF_u) - EF(E_uW_v - E_vW_u).\end{aligned}$$

Bearing in mind (C.5) we write:

$$\begin{aligned}& 2EW^3\left(\frac{\partial}{\partial v}\Big(\frac{FE_u}{EW}\Big) - \frac{\partial}{\partial u}\Big(\frac{FE_v}{EW}\Big)\right) \\ =\ & 2W^2(E_uF_v - E_vF_u) - F\big(E_u(2WW_v) - E_v(2WW_u)\big) \\ =\ & 2W^2(E_uF_v - E_vF_u) - E_uF(E_vG + EG_v - 2FF_v) \\ & +E_vF(E_uG + EG_u - 2FF_u) \\ =\ & (2W^2 - 2F^2)(E_uF_v - E_vF_u) - EF(E_uG_v - E_vG_u) \\ =\ & 2EG(E_uF_v - E_vF_u) - EF(E_uG_v - E_vG_u),\end{aligned}$$

and the result follows on division by $2EW^3$. □

Now Brioschi's formula (C.1) follows from (C.6) and (C.9), whereas (C.2) follows from (C.7) and (C.8). Formula (C.3) is simply the average of (C.1) and (C.2).

References

[Bartle] R Bartle, *The elements of real analysis,* 2nd ed. Wiley, 1976.
[Spivak] M Spivak, *A comprehensive introduction to differential geometry*, Vol. 2, Publish or Perish, 1999.

Dept. of Mathematics, University of York