SURFACES IN $\mathbb{R}^n$.

KEVIN MCGERTY

1. SURFACES: DEFINITION AND BASIC PROPERTIES

These notes contain an account of the material we have covered on surfaces in $\mathbb{R}^n$.

Definition 1.1. A subset $S \subset \mathbb{R}^n$ is said to be a k-dimensional surface if for each point $s \in S$ there is a neighborhood U of s in S and a homeomorphism $\varphi : \mathbb{R}^k \to U$.

Here S is given the topology inherited from $\mathbb{R}^n$, and so an open set in S is the intersection of S with an open set in $\mathbb{R}^n$, thus a neighborhood of $x \in S$ is of the form $N \cap S$ where N is a neighborhood of x in $\mathbb{R}^n$. Recall that a *homeomorphism* is a continuous bijection whose inverse is also continuous.

Informally then, surfaces in $\mathbb{R}^n$ are subsets which "locally look like $\mathbb{R}^k$". The maps φ occuring in the definition are known as *charts*. A collection of charts $\{\varphi_i \colon \mathbb{R}^k \to S\}_{i \in I}$ such that $S = \cup_{i \in I} \varphi_i(\mathbb{R}^k)$ is called an *atlas*. Thus in order to demonstrate that a subset $S \subset \mathbb{R}^n$ is a k-surface, it is enough to give an atlas for S. There is a related notion which is that of a *k-submanifold* of $\mathbb{R}^n$:

Definition 1.2. A subset $S \subset \mathbb{R}^n$ is a submanifold of $\mathbb{R}^n$ if for each $s \in S$ there is an open set V of $\mathbb{R}^n$ which contains s and a homeomorphism $\psi : \mathbb{R}^n \to V$ such that (if we think of $\mathbb{R}^k$ lying inside $\mathbb{R}^n$ as the span of the first k standard basis vectors) $\psi_{|\mathbb{R}^k}$ maps $\mathbb{R}^k$ homeomorphically to S.

In other words we require that if S is a submanifold, then locally at a point $s \in S$ the pair $S \subset \mathbb{R}^n$ looks like $\mathbb{R}^k \subset \mathbb{R}^n$. Notice that a k-submanifold is certainly a k-surface, since we can take $\psi_{|\mathbb{R}^k}$ as a chart for the point s. On the other hand, if our charts are simply continuous, it is not necessarily that case that a surface will be a submanifold ([Z] chapter 12 briefly mentions the *Alexander horned sphere* which is an example of this, for more details see [H]). In analysis however, we as interested in the case where the chart maps are differentiable. In this case the the notions of surface and submanifold above are very close: in fact if we insist that the charts φ in the definition of a surface are $\mathcal{C}^{(1)}$ then the notions coincide wherever the derivative has maximal rank. (Since the derivative is a matrix with n rows and k columns, and $k \leq n$, this means where the derivative has rank k.) Before proving this, we first show that the set of points in the domain of a $\mathcal{C}^{(1)}$ chart φ where the derivative has maximal rank is an open set (which, however could be empty), and so the maximal rank condition should be thought of as "stable"– see also the statement of Sard's theorem at the end of this section for a much deeper result which shows that it is also in some sense "generic".

Date: May, 2005.

Lemma 1.3. *Let $k \leq n$ and $\varphi : \mathbb{R}^k \to \mathbb{R}^n$ be a $\mathcal{C}^{(1)}$ map. If $D\varphi(0)$ has rank k, then there is a neighborhood U of 0 in $\mathbb{R}^k$ such that $d\varphi$ has maximal rank at every point in U.*

Proof. Since the matrix $D\varphi(0)$ has maximal rank we can find k coordinates (which for notional convenience we assume are the first k) such that the matrix $\Delta_k = (\frac{\partial \varphi_i}{\partial t_j})_{1 \leq i,j \leq k}$ is invertible at 0. But now recall that a matrix is invertible precisely when its determinant is nonzero, and so this means that $\det(\Delta_k(0) \neq 0$. But now since $D\varphi$ is continuous so is $\det(\Delta_k(t))$ and hence the preimage of $\mathbb{R} - \{0\}$ contains a neighborhood U of $0 \in \mathbb{R}^k$. Thus on U the matrix $(\frac{\partial \varphi_i}{\partial t_j})_{1 \leq i,j \leq k}$ is invertible and hence $d\varphi$ has maximal rank. ☐

We now show that there is no difference between a k-surface and a k-submanifold when the charts of the surface are $\mathcal{C}^{(1)}$ and have maximal rank. Since open balls centered at 0, open intervals centered at 0 and $\mathbb{R}^k$ itself are all diffeomorphic (convince yourself of this!) we are free to take any of them as the domain of our charts. We will prefer to choose intervals. For any k we set $I^k_\varepsilon = \{x \in \mathbb{R}^k : |x_i| < \varepsilon, 1 \leq i \leq k\}$ and $I^k = I^k_1$.

Proposition 1.4. *Suppose S is a k-surface in $\mathbb{R}^n$ such that each point $s \in S$ has a $\mathcal{C}^{(1)}$ chart $\varphi : I^k \to S$ with $\varphi(0) = s$ and $D\varphi(0)$ of maximal rank. Then S is a submanifold of $\mathbb{R}^n$. In fact more precisely there is an $\varepsilon > 0$ and a diffeomorphism $\psi : I^n_\varepsilon \to \mathbb{R}^n$ such that $\psi_{|I^k \cap I^n_\varepsilon} = \varphi_{|I^k \cap I^n_\varepsilon}$.*

Proof. Since we are assuming that such charts exist for each point $s \in S$ it is clear that it is enough to show the existence of the diffeomorphism ψ. Now since $D\varphi(0)$ has rank k we can assume that the submatrix $(\frac{\partial \varphi_i}{\partial t_j})_{1 \leq i,j \leq k}$ is invertible. Let U be a neighbourhood of $0 \in \mathbb{R}^k$ on which that submatrix of $D\varphi$ is invertible (see the previous lemma). Pick $\varepsilon_1 > 0$ such that $I^n_{\varepsilon_1} \subset U \times \mathbb{R}^{n-k}$. Define $\psi : I^n_{\varepsilon_1} \to \mathbb{R}^n$ by

$$\psi(t_1, t_2, \ldots, t_n) = \varphi(t) + (0, 0, \ldots, 0, t_{k+1}, t_{k+2}, \ldots, t_n).$$

It is then easy to check that the derivative of ψ is invertible on $I^n_{\varepsilon_1}$ and so by the inverse function theorem, ψ is a diffeomorphism on some (perhaps smaller) interval I^n_ε, and we are done. ☐

This leads us to the following definition.

Definition 1.5. A $\mathcal{C}^{(m)}$ (sometimes *smooth*) k-surface in $\mathbb{R}^n$ is a k-surface S such that at each point $s \in S$ there is neighborhood U and a chart $\varphi : I^k \to U$ which is $\mathcal{C}^{(m)}$ and for which the derivative at $\varphi^{-1}(s)$ has rank k.

Thus the proposition shows that a $\mathcal{C}^{(1)}$ surface is a submanifold of $\mathbb{R}^n$. We also want to have a notion of differentiability for a function $f\colon S \to \mathbb{R}$. The easiest way to do this is to insist that f is the restriction of a differentiable function on $\mathbb{R}^n$, at least locally, and indeed this makes sense for any subset of $\mathbb{R}^n$.

Definition 1.6. If A is a subset of $\mathbb{R}^n$ we say that a function $f\colon A \to \mathbb{R}$ is $\mathcal{C}^{(m)}$ at a point $x \in A$ if there exists a neighborhood U of x in $\mathbb{R}^n$ and a differentiable function $g\colon U \to \mathbb{R}$ such that $g_{|A \cap U} = f_{|A \cap U}$.

A consequence of the previous proposition is that we can use charts to test if a function is differentiable in the above sense.

Lemma 1.7. *Let $S \subset \mathbb{R}^n$ be a smooth surface in $\mathbb{R}^n$.*

1. *If* $\varphi\colon I^k \to U \subset S$ *is a smooth* $\mathcal{C}^{(m)}$ *chart of* S *and* $x \in U$ *then a function* $f\colon U \to \mathbb{R}$ *is in* $\mathcal{C}^{(p)}(U,\mathbb{R})$ *for* $p \leq m$ *if and only if* $f \circ \varphi : I^k \to \mathbb{R}$ *is in* $\mathcal{C}^{(p)}(I^k,\mathbb{R})$.
2. *Suppose that* $\varphi\colon I^k \to U \subset S$ *and* $\psi\colon I^k \to V \subset S$ *are smooth* $\mathcal{C}^{(m)}$ *charts such that* $U \cap V \neq \emptyset$. *Then*
$$\theta = \psi^{-1} \circ \varphi\colon \varphi^{-1}(U \cap V) \to \psi^{-1}(U \cap V)$$
is a $\mathcal{C}^{(m)}$ *map between open subsets of* $\mathbb{R}^k$.

Proof. For the first part, clearly if f is in $\mathcal{C}^{(p)}(S,\mathbb{R})$ then the composite $f \circ \varphi$ is in $\mathcal{C}^{(p)}(I^k,\mathbb{R})$ whenever $p \leq m$. Conversely, Proposition 1.4 shows that if $x \in S$ is the image of $t \in I$ we may extend φ to a $\mathcal{C}^{(m)}$ diffeomorphism Φ from an open neighborhood of $t \in I^n \supset I^k$ to an open neighborhood of $x \in \mathbb{R}^n \supset S$ (this is a slightly more general statement than what is given in Proposition 1.4 but it has exactly the same proof). Thus if we know that $f \circ \varphi$ is in $\mathcal{C}^{(p)}(I^k,\mathbb{R})$ then we may set
$$\tilde{g}(t_1,t_2,\ldots,t_n) = f \circ \varphi(t_1,t_2,\ldots,t_k),$$
and then $g = \tilde{g} \circ \Phi^{-1}$ is a $\mathcal{C}^{(p)}$ function on the image of Φ which extends f as required.

For the second part, suppose that $x \in U \cap V$, and $t = \varphi^{-1}(x) \in \varphi^{-1}(U \cap V)$. Then $\theta = \psi \circ \Phi^{-1}$ with Φ as above, and so it is clear that θ is $\mathcal{C}^{(m)}$ as required. □

Remark 1.8. The second part of the previous lemma shows that if $f \circ \varphi$ is differentiable for one chart it is differentiable for any chart. (Of course this also follows from the first part, since our definition of differentiability did not depend on any chart). We could therefore *define* a function $f\colon S \to \mathbb{R}$ on a smooth $\mathcal{C}^{(m)}$ surface S to be in $\mathcal{C}^{(p)}(S,\mathbb{R})$ for $p \leq m$ if for some (and hence any) atlas $\{\varphi_i\colon I^k \to U_i\}_{i \in I}$ the functions $f_i = f \circ \varphi_i : I^k \to \mathbb{R}$ are all in $\mathcal{C}^{(p)}(I^k,\mathbb{R})$.

Example 1.9. It is relatively easy to check directly that things like the sphere
$$S^2 = \{(x_1,x_2,x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$$
are smooth surfaces. One way to do this is to use spherical polar coordinates, $P\colon \mathbb{R}^2 \to \mathbb{R}^3$ where
$$P(\phi,\theta) = (\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi)).$$
A direct check shows that the image of P is exactly S^2, however since the functions $\sin$ and $\cos$ are periodic, P is not quite a chart – it is not a homeomorphism onto its image (indeed since the sphere S^2 is compact it cannot be covered by any single chart). This can easily be remedied by using P to define an atlas with two charts. Let $R : \mathbb{R}^3 \to \mathbb{R}^3$ be the map given by $R(x_1,x_2,x_3) = (x_2,x_3,x_1)$. Then R maps S^2 to itself, and if we put $P_1 = P_{|(0,2\pi)\times(0,\pi)}$ and $P_2 = R \circ P_{|(0,2\pi)\times(0,\pi)}$ then $\{P_1,P_2\}$ is an atlas, since P_1 omits only the points $(0,0\pm 1)$ from it's image, and P_2 omits only the points $(0,\pm 1,0)$.

In the previous example we were lucky to have a simple enough geometric picture that it was not too difficult to produce an explicit parametrization, but if we are to work with more general surfaces it is clear that we will need a method of showing that something is a surface which is less cumbersome than searching for explicit charts each time. Thankfully we can essentially turn the proof of the Proposition 1.4 upside-down to give us a general method for producing surfaces.

Definition 1.10. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a differentiable function, we say that $x \in \mathbb{R}^n$ is a *regular point* if $Df(x)$ has maximal rank at x. If $y \in \mathbb{R}^m$ we say that y is a *regular value* of f if every x in

$$f^{-1}(y) = \{x \in \mathbb{R}^n : f(x) = y\}$$

is a regular point of f, otherwise y is a *critical value*. Sets of the form $f^{-1}(y)$ are known as the *fibers* of f.

Proposition 1.11. *Suppose that $f : \mathbb{R}^n \to \mathbb{R}^m$ is $\mathcal{C}^{(1)}$ and that $y \in \mathbb{R}^m$ is a regular value of f. Then $S = f^{-1}(y)$ is either empty or a $\mathcal{C}^{(1)}$ surface in $\mathbb{R}^n$ of dimension $n - m$.*

Proof. We assume that S is nonempty, so that for each $s \in S$ we must produce a $\mathcal{C}^{(1)}$ chart. Now since y is a regular value, we know that $Df(s)$ has rank m, thus reordering the coordinates if necessary we may assume that the submatrix $(\frac{\partial f_i}{\partial x_j})_{1 \leq i,j \leq m}$ is invertible. Consider now the map $g : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$g(x) = (f_1(x) - y_1, f_2(x) - y_2, \ldots, f_m(x) - y_m, x_{m+1} - s_{m+1}, \ldots, x_n - s_n).$$

(where $y = (y_1, y_2, \ldots, y_m)$ and $s = (s_1, s_2, \ldots, s_n)$ so that $g(s) = 0$.) It is immediate that $g \in \mathcal{C}^{(1)}(\mathbb{R}^n, \mathbb{R}^n)$ and that $Dg(s)$ is invertible. Hence by the inverse function theorem there is a neighborhood U of s on which g is a diffeomorphism onto a neighborhood of $g(0) = 0$. Thus if $\psi : I_\varepsilon^n \to U$ is the inverse of g defined in some sufficiently small interval around 0, we see that the restriction of ψ to the last $n - m$ coordinates provides a $\mathcal{C}^{(1)}$ chart for S. ☐

Example 1.12. It is now easy to see, for example, that the n-sphere $S^n = \{x \in \mathbb{R}^{n+1} : x_1^2 + x_2^2 + \ldots + x_{n+1}^2 = 1\}$ is a smooth surface of dimension n. With a little more thought you can see that, for example the set of $n \times n$ orthogonal matrices

$$O_n(\mathbb{R}) = \{A \in \mathrm{Mat}_{n,n}(\mathbb{R}) : A.A^t = I_n\}$$

is also smooth surface in $\mathrm{Mat}_{n,n}(\mathbb{R})$ (which is just $\mathbb{R}^{n^2}$) of dimension $\frac{1}{2}n(n-1)$. (The point is to notice that the map $f : \mathrm{Mat}_{n,n}(\mathbb{R}) \to \mathrm{Mat}_{n,n}(\mathbb{R})$ which sends A to A^tA lands in the linear subspace of *symmetric* matrices, *i.e.* matrices X for which $X = X^t$. Since this space has dimension $\frac{1}{2}n(n+1)$ we can think of f as a map from $\mathbb{R}^{n^2}$ to $\mathbb{R}^{\frac{1}{2}n(n+1)}$, and so if one shows that I_n is a regular value of f we get the desired result. But since $Df(A)(H) = A^tH + H^tA$ this is easy.)

Remark 1.13. Lemma 1.7 and the remark following it suggest that all these results might be extended to maps between surfaces. Indeed given surfaces S and T we can define a function $f \colon S \to T$ to be m-times differentiable at a point $x \in S$ if for some chart $\varphi : I^k \to S$ containing x in its image and some chart $\psi : I^l \to T$ containing $f(x)$ in it's image, the function $g = \psi^{-1} \circ f \circ \varphi$ is m-times differentiable at $\varphi^{-1}(x)$. It is then obvious how to extend the notions of regular value, critical point *etc.* and the proof of the corresponding version of Proposition 1.11 is almost the same as the one we give for $\mathbb{R}^n$.

This last proposition shows that fibers of differentiable maps are smooth surfaces at a regular value. We mention here without proof a fundamental result that shows that this will usually give a large supply of smooth surfaces.

Theorem 1.14. *(Sard's theorem) If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable then the set of critical values of f has measure zero.*

Remark 1.15. Note that a point which is not in the image of f is a regular value, vacuously.

2. TANGENT SPACES

We now want to make sense of the notion of a tangent vector to a smooth surface. We start by considering the case of a curve C inside the plane $\mathbb{R}^2$. Then we give two ways of thinking about the tangent space at a point $(x_0, y_0) \in C$. First, appealing to physical intuition, consider a particle p moving along the curve, say with position $p(t) \in C$ at time t. Now since the point p always lies on the curve C, it's velocity at point should be tangent to the curve, that is, if p is at (x_0, y_0) at time $t = 0$ then the vector $\dot{p}(0) = \frac{dp}{dt}(0)$ should be a tangent vector to C at (x_0, y_0). The set of all possible velocities of a particle as it passes through the point (x_0, y_0) is then the tangent line T_0C to C at (x_0, y_0).

We can give a second desciption in the case when we can write the curve as the graph of a function $f : \mathbb{R} \to \mathbb{R}$, *i.e.* $C = \{(x, y) \in \mathbb{R}^2 : y = f(x)\}$. One of the standard descriptions of the derivative at a point $x \in \mathbb{R}$ is as the slope of the tangent line at the corresponding point $(x, f(x))$ of the graph of f. Thus elementary calculus tells us that the tangent line T_0C is the $\{(x, y) : y = f(x_0) + f'(x)(x - x_0)\}$, which I would like to write in the form:

$$\{(x_0, y_0) + t(1, f'(x_0)) : t \in \mathbb{R}\}$$

Thus if we think of f giving a chart $\phi : \mathbb{R} \to C$ given by $t \mapsto (t, f(t))$ then the tangent space is just the set $\{(x_0, y_0) + t\phi'(x_0) : t \in \mathbb{R}\}$.

Of course in this simple case it is immediate that these viewpoints are extremely close to each other, and it is not hard to see that they give the same answer (this will also follow from what we show later). However even in the case of curves inside $\mathbb{R}^2$ the first notion makes sense more generally than the second.

Example 2.1. Consider the curve

$$C = \{(x, y) \in \mathbb{R}^2 : y^2 = x^3 + x^2\},$$

then there is a map $\psi : \mathbb{R} \to C$ given by $\psi(t) = (t^2 - 1, t(t^2 - 1))$, and moreover $\psi'(t) = (2t, 3t^2 - 1)$, which never vanishes, thus it looks tantalizingly like a smooth chart. However $\psi(1) = \psi(-1) = (0, 0)$ and so ψ is not a bijection, and indeed at $(0, 0)$ the curve C is not smooth. On the other hand if we define the tangent space at $(0, 0)$ to be the set of all possible velocity vectors at $(0, 0)$ of particles moving on C it is not hard to convince yourself that the tangent space at $(0, 0)$ is the union of two lines – the lines $\{(x, y) \in \mathbb{R}^2 : x - y = 0\}$ and the line $\{(x, y) \in \mathbb{R}^2 : x + y = 0\}$. The second definition of tangent space does not even make sense, because you cannot write C as the graph of a function of x near $(0, 0)$ (in fact you cannot find a map $\phi : I^1 \to C$ which is a $\mathcal{C}^{(1)}$ homeomorphism onto a neighborhood of $(0, 0)$ in C).

Note that the example above is not a smooth curve, and the fact that the tangent space at the origin is a pair of lines, which may feel surprising, is not a phenomenon that can happen for smooth curves, as we will shortly see.

We now give our definition for the tangent space to a surface in $\mathbb{R}^n$.

Definition 2.2. Let $S \subset \mathbb{R}^n$ be a subset of $\mathbb{R}^n$ and let $s \in S$ be a point in S. Then let

$$\mathcal{G} = \{\alpha : (-\varepsilon, \varepsilon) \to S : \alpha \in \mathcal{C}^{(1)} \text{ and } \alpha(0) = s\}$$

be the set of all differentiable curves on S defined in some neighborhood of $0 \in \mathbb{R}$ which pass through s at time 0. The tangent space at s is then

$$T_sS = \{\frac{d\alpha}{dt}(0) : \alpha \in \mathcal{G}\}.$$

Remark 2.3. You should of course notice that there is in some sense a lot of redundancy in this definition – many different curves α will have the same velocity at s. This redundancy also makes it look rather difficult to calculate, so we want another description which is more computable. For smooth surfaces, the charts will provide us with exactly that. This description will also show that for smooth surfaces the tangent space is a linear space.

Remark 2.4. The interval $(-\varepsilon, \varepsilon)$ which is the domain of the curve α is allowed to be arbitrarily small – we are only interested in the derivative of α at 0 and this can be calculated once we know α in some neighborhood of 0.

Suppose now that S is a smooth surface in $\mathbb{R}^n$.

Proposition 2.5. *Let $\phi : I^k \to S$ be a smooth chart with range $U \subset S$, and set $s = \phi(0)$. The tangent space T_sS is a linear subspace of $\mathbb{R}^n$. In fact we have*

$$T_sS = \{D\phi(0)(v) : v \in \mathbb{R}^k\}.$$

Proof. Since ϕ is a smooth chart, the derivative $D\phi(0)$ has full rank, so we can apply Proposition 1.4 to view ϕ as the restriction of a diffeomorphism $\psi : I^n_{\varepsilon_0} \to \mathbb{R}^n$, where $\varepsilon_0 > 0$. Now suppose that $\alpha : (-\varepsilon, \varepsilon) \to S$ is $\mathcal{C}^{(1)}$ and $\alpha(0) = s$. Notice that by Remark 2.4 we may assume that the image of α is contained in the image of ψ. Then

$$\beta = \psi^{-1} \circ \alpha : (-\varepsilon, \varepsilon) \to I^n_{\varepsilon_0}$$

is a $\mathcal{C}^{(1)}$ curve. Moreover because $\alpha = \psi \circ \beta$ the chain rule shows that $\alpha'(0) = D\psi(0)(\beta'(0))$, and hence the vector $\alpha'(0)$ lies in the image of $D\psi(0)$. Since the image of β lies in $\mathbb{R}^k$ we see that $\alpha'(0)$ lies in fact in the image of $D\phi(0)$. Thus we see that $T_sS \subset \text{im}(D\phi(0))$. To obtain the reverse inclusion consider for each $v \in \mathbb{R}^k$ the path $\alpha(t) = \psi(tv)$. $\square$

Remark 2.6. I have defined the tangent space at a point of a smooth surface to be a linear subspace of $\mathbb{R}^n$. This is probably not quite what one's standard picture of a tangent vector is in the case of a curve – we tend to draw the tangent vector starting at the point on the curve. It is of course fine to think of the tangent space as I defined it simply translated to the point in the same way, however the more modern idea is to consider the set of *all* tangent vectors to *all* points on the smooth surface S as a new surface TS in $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ the points of which consist of pairs (s, v) where $s \in S$ and $v \in T_sS$. The object TS is called the *tangent bundle* of S. (If you have done Lagrangian mechanics at some point, then the tangent bundle is essentially the phase space – that is the space of possible positions and velocities of a point moving on the surface S).

Remark 2.7. The fact that TS is a surface is perhaps the first time we need to pay attention to the degree of differentiability of our surfaces. If S has charts ϕ which are infinitely differentiable (*i.e.* if S is a $\mathcal{C}^{(\infty)}$ surface) then so is TS. However if S is a $\mathcal{C}^{(m)}$ surface, then TS is only a $\mathcal{C}^{(m-1)}$ surface). Indeed suppose that $\varphi : I^k \to U \subset S$ is a $\mathcal{C}^{(m)}$ chart for S. Then $\mathcal{D}\varphi : I^k \times \mathbb{R}^k \to TS$ given by

$$\mathcal{D}\varphi(x, v) = (\varphi(x), D\varphi(x)(v))$$

is a $\mathcal{C}^{(m-1)}$ chart of TS, with range $\bigcup_{s\in U} T_sS$.

The last notion we want to introduce is that of a *vector field* or *section* of the tangent bundle.

Definition 2.8. A vector field or section of TS over a subset $U \subset S$ is a map $t\colon U \to TS$ such that for every $x \in U$ we have $t(x) = (x, v)$ where $v \in T_xS$.

In other words, if we let $\pi\colon TS \to S$ be obvious map sending a pair $(x, v) \in TS$ to $x \in S$, then a map $t : U \to TS$ is a section if $\pi \circ t$ is the identity map on U. Since we have shown that TS is a surface, it makes sense to say that a section is continuous, differentiable *etc*.

3. ORIENTATION

For surfaces in $\mathbb{R}^3$ we familiar with the idea that there are "two sides" – for example if we think of the sphere of points at distance 1 from the origin, we can imagine ourselves standing on the "outside" or "inside" of the sphere. Slightly less intuitively, it is in fact that case that some surfaces have only one side, the most famous example being the "Möbius band" which you can make by gluing together the ends of a strip of paper with a half-twist. (Formally if we set

$$\begin{aligned}\mathbf{e}(\varphi) &= (cos(\varphi), sin(\varphi), 0),\\ \mathbf{f}(\varphi) &= (cos(\varphi)sin(\varphi/2), sin(\varphi)sin(\varphi/2), cos(\varphi/2)),\end{aligned}$$

we can define a surface M in $\mathbb{R}^3$ given by

$$\{\mathbf{e}(\varphi) + t\mathbf{f}(\varphi) : \varphi \in [0, 2\pi], t \in (-\frac{1}{2}, \frac{1}{2})\}$$

which is a Möbius band). We want to make sense of this notion of "sides" for a general surface. It is our first example of a global phenomenon: Locally any smooth 2-surface in $\mathbb{R}^3$ looks like $\mathbb{R}^2 \subset \mathbb{R}^3$, and so clearly it has two sides. The M´obius band shows that it is not always possible to label the sides we see locally, in a way which is globally consistent. What we actually do to formalize our discussion is to rephrase the question of in terms of an *orientation* for the surface, thus we will first define what the notion of an orientation for a surface is, and then return to see that it captures the idea of "sides" of a surface.

An orientation is a generalization of the notion of "clockwise" and "anticlockwise" for $\mathbb{R}^2$ and (if you have done physics) the "right-hand" and "left-hand" rule for $\mathbb{R}^3$. Thus for $\mathbb{R}^2$ we say two nonzero vectors v_1, v_2 are oriented "clockwise" if one rotates clockwise when turning from the direction of v_1 to the direction of v_2, and oriented "anti-clockwise" in the opposite case (notice that this only makes sense if the two vector are not collinear). The key to generalizing this is to observe that if A is the two-by-two matrix whose first row is v_1 and whose second row is v_2 then the vectors are oriented anticlockwise if $\det(A) > 0$ and they are oriented clockwise if $\det(A) < 0$ (make sure you convince yourself of this!)

For $\mathbb{R}^n$ we generalize this as follows: A *frame* in $\mathbb{R}^n$ is a sequence of n linearly independent vectors $(v_1, v_2, \dots, v_n)$. If we have two frames $(v_1, v_2, \dots v_n)$ and $(w_1, w_2, \dots, w_n)$ then since they both give bases of $\mathbb{R}^n$ we may write

$$w_i = \sum_{j=1}^{n} a_{ij} v_j, \qquad 1 \leq i, j \leq n,$$

for some constants a_{ij}. The matrix $A = (a_{ij})$ is called the *transition matrix* from $(v_1, v_2, \dots, v_n)$ to $(w_1, w_2, \dots, w_n)$. We put a relation on the set of frames in $\mathbb{R}^n$ by setting $(v_1, v_2, \dots, v_n) \sim (w_1, w_2, \dots, w_n)$, if $\det(A) > 0$ where A is the transition matrix above. It is easy to check, using the standard properties of the determinant, that this gives an equivalence relation, and an *orientation* of $\mathbb{R}^n$ is an equivalence class of this relation. We denote the set of orientations by $\mathcal{O}(\mathbb{R}^n)$. In fact it is easy to see that there are exactly two equivalence classes: the equivalence class containing $(e^1, e^2, \dots, e^n)$ the standard frame, and the equivalence class containing the frame $(-e^1, e^2, \dots, e^n)$. Indeed a more concrete way of defining an orientation would have been to declare set of frames in $\mathbb{R}^n$ to be partitioned according to the sign of $\det(C)$ where C is the matrix whose i-th row is i-th vector of the frame (this is essentially what we did for $\mathbb{R}^2$. It's worth convincing yourself that this notion of orientation coincides with the one you are familiar with for $\mathbb{R}^3$.

Lemma 3.1. *Let $\Lambda^n(\mathbb{R}^n)$ be the space of alternating multilinear functions on n vectors in $\mathbb{R}^n$, a one dimensional vector space. Then an orientation of $\mathbb{R}^n$ corresponds to a direction in $\Lambda^n(\mathbb{R}^n)$.*

Proof. If $\omega = dx_1 \wedge \dots \wedge dx_n$ denotes the alternating function corresponding to taking the determinant of the matrix whose rows are the n vectors in order (see the appendix), then $\Lambda^n(\mathbb{R}^n)$ is spanned by the function ω. The above "concrete" description of the equivalence classes shows that one class corresponds to the positive multiples of ω and the other to the negative multiples of ω. □

We now want to define an orientation for a k-surface in $\mathbb{R}^n$.

Definition 3.2. An orientation $\mathcal{O}$ of a smooth surface S is a choice of orientation $\mathcal{O}(T_xS)$ for the tangent space T_xS of each point $x \in S$ which varies continuously with $x \in S$.

Of course in order for this defintion to make sense we need to say what we mean by "varying continously" here. A family of frames F over $U \subset S$ is just a k-tuple $F = (t_1, t_2, \dots, t_k)$ of sections over U such that if $t_i(s) = (s, v_i(s))$ (so $v_i(s) \in T_sS$) then for each $s \in U$ the vectors $(v_1(s), v_2(s), \dots, v_k(s))$ are a frame in T_sS. It is then natural to say a family of frames is continuous if the sections t_i are continuous. We then define a family of orientations to be continuous at a point $s \in S$ if there is a neighborhood U of s, and a continuous family of frames F on U such that the orientation at s is the equivalence class of the frame $F(s)$.

We now want to know when it is that a surface can be given an orientation.

Definition 3.3. An smooth k-surface $S \subset \mathbb{R}^n$ is said to be *orientable* if it has a smooth atlas

$$\{\varphi_i \colon I^k \to U_i\}_{i \in I}$$

(*i.e.* the φ_i are charts such that $S = \cup_{i\in I} U_i$) such that for each $i, j \in I$ the charts φ_i and φ_j are *compatible* in the sense that either $U_i \cap U_j = \emptyset$ or

$$\theta_{ij} = \varphi_i^{-1} \circ \varphi_j \colon \varphi_j^{-1}(U_i \cap U_j) \to \varphi_i^{-1}(U_i \cap U_j)$$

has $\det(D\theta_{ij}(t)) > 0$ for all $t \in \varphi_j^{-1}(U_i \cap U_j)$.

Notice that $\theta_{ij} \in \mathcal{C}^{(m)}(\varphi_j^{-1}(U_i \cap U_j), \varphi_i^{-1}(U_i \cap U_j))$ by Lemma 1.7. The following lemma justifies the terminology "orientable".

Lemma 3.4. *An orientable surface $S \subset \mathbb{R}^n$ has an orientation.*

Proof. We first define an orientation for the open sets U_i ($i \in I$). For this we take the standard frame $(e^1, e^2, \dots e^k)$ in $\mathbb{R}^k$ and take the associated family of frames over U_i, that is $F : U_i \to (TS)^k$ given by

$$F(\varphi(t)) = \big((\varphi(t), D\varphi_i(t)(e^1)), (x, D\varphi_i(t)(e^2)), \dots, (x, D\varphi_i(t)(e^k)\big),$$

It is clear that this is a continuous family of frames, and hence we get, by taking orientation classes, a continuous family of orientations on each U_i. In order to obtain a family of orientation on all of S we must only check that if $x \in U_i \cap U_j$ then the orientations of T_xS given by the frames

$$(D\varphi_i(t)(e^1), D\varphi_i(t)(e^2), \dots D\varphi_i(t)(e^k)),$$

and

$$(D\varphi_j(t)(e^1), D\varphi_j(t)(e^2), \dots D\varphi_j(t)(e^k)),$$

are the same. But by the chain rule, the transition matrix between these two frames is exactly given by $\theta(t)$, which by assumption has $\det(D\theta(t)) > 0$. ☐

Remark 3.5. We say that the atlas in the previous lemma is compatible with the orientation. Indeed it is clear from the definitions that given any chart $\varphi \colon I^k \to S$ either it is compatible with the orientation, or the chart $\tilde{\varphi} : I^k \to S$ where $\tilde{\varphi}(t_1, t_2, \dots, t_k) = \varphi(-t_1, t_2, \dots, t_k)$ is compatible. Using this observation it is easy to see that any orientation has a compatible atlas. (See also the proof of Proposition 7.5).

Example 3.6. The charts P_1, P_2 for the two sphere S^2 described in Example 1.9 given a compatible atlas for S^2, and hence it is an orientable surface.

The following lemma is not used in subsequent sections, but it clarifies the notion of an orienting atlas:

Lemma 3.7. ($*$) *Let $\mathcal{A} = \{\varphi_i : I^k \to U_i\}_{i\in I}$ be an atlas of compatible charts. Then if $\psi \colon I^k \to S$ is a chart, either ψ is compatible with all the φ_i or it is compatible with none.*

Proof. It is sufficient to show that if $\psi \colon I^k \to V \subset S$ is compatible with some φ_{i_0} with $V \cap U_{i_0} \neq \emptyset$, it is compatible with all other φ_j. For $i, j \in I$, let $\theta_{ij} = \varphi_i^{-1} \circ \varphi_j$, (thus by assumption $\det(D\theta_{ij}) > 0$), and let $\eta_j = \psi^{-1} \circ \varphi_j$. We want to show that $\det(D\eta_j) > 0$ for all j, given that $\det(D\eta_{i_0}) > 0$.

Fix $x \in V \cap U_{i_0}$, and suppose that $y \in V \cap U_j$ for some $j \in I$. Take a smooth path $c \colon [0, 1] \to V$ with $c(0) = x$ and $c(1) = y$. By compactness we can find a finite set of U_i, say $U_i = U_{i_0}, U_{i_1}, \dots U_{i_p} = U_j$ which covers $c([0, 1])$. Moreover we can arrange that $U_{i_r} \cap U_{i_{r+1}}$ is nonempty. We show by induction on r that U_{i_r} is compatible with ψ. For $r = 0$ this is true by assumption. For $r > 0$ pick $z \in U_{i_{r-1}} \cap U_{i_r}$. Then since ψ since $\eta_{i_r} = \eta_{i_{r-1}} \circ \theta_{i_{r-1} i_r}$, by induction and the chain rule we see that $\det(D\eta_{i_r}(z)) > 0$. But since $\det(D\eta_{i_r})$ is always nonzero, it follows that it is positive on $c([0,1]) \cap V \cap U_{i_r}$. Then since $U_{i_p} = U_j$ we see that ψ is compatible with φ_j as required. □

Remark 3.8. An alternative definition of an orientation can be given by taking an orientation to simply *be* an atlas $\mathcal{A} = \{\varphi_i : I^k \to U_i\}$ of compatible charts as in Definition 3.3. In this case it is normal to take the atlas to be *maximal*, that is, if $\psi : I^k \to V \subset S$ is a chart, and ψ is compatible with each φ_i then in fact $\psi \in \mathcal{A}$. The previous lemma shows that it is sufficient to check ψ is compatible with a single chart in the atlas. It also has the following corollary:

Corollary 3.9. $(*)$ *A connected surface S which is orientable has exactly two orientations. Moreover if S has a sequence of charts $\varphi_i : I^k \to U_i$ for $i = 1, 2, \dots, n$ such that for each i we have $U_i \cap U_{i+1} \neq \emptyset$ and φ_i and φ_{i+1} are compatible but $U_1 \cap U_n \neq \emptyset$ and φ_1, φ_n are incompatible then S is not orientable.*

Proof. The previous lemma shows that a given orienting atlas $\mathcal{A}$ splits the set of charts for a connected surface up into two nonempty equivalence classes – those charts which are compatible with $\mathcal{A}$, and those which are not. From this is it easy to deduce there is at most two possible orientations.

For the second part suppose that an orienting $\mathcal{A}$ exists. Considering which equivalence class of $\mathcal{A}$ the charts in the sequence lie in, we immediately obtain a contradiction. □

Remark 3.10. The sequence of charts in the second part of the Corollary is known as an "incompatible chain". It is easy to use the explicit description of the Möbius strip given at the start of the section to produce a three term incompatible chain for it, thereby showing it is not orientable.

Finally, we prove a result which relates orientation to the idea of the "sides" of a surface, at least for surfaces of dimension $(n-1)$ (or *hypersurfaces*) in $\mathbb{R}^n$. To do this we need the notion of a normal vector field for a k-surface S in $\mathbb{R}^n$. This is a continuous function $\mathbf{n}\colon S \to \mathbb{R}^n$ such that at each $x \in S$ we have $(\mathbf{n}(x), v) = 0$ for each $v \in T_xS$.

Lemma 3.11. *Let S be a smooth $(n-1)$ surface in $\mathbb{R}^n$. Then an orientation of S corresponds to a continuous normal vector field $\mathbf{n}$ such that $||\mathbf{n}(x)|| = 1$ for all $x \in S$.*

Proof. Fix an orientation for $\mathbb{R}^n$, say the standard one given by the frame. Let $\mathcal{O}$ be an orientation of S and suppose that $\varphi : I^{n-1} \to U \subset S$ is a chart of S such that the standard frame F given by φ over U induces $\mathcal{O}$. Since S is a hypersurface, there is a unique normal vector $\mathbf{n}(x)$ of length 1 such that

$$(\mathbf{n}(x), D\varphi(\varphi^{-1}(x))(e^1), D\varphi(\varphi^{-1}(x))(e^2), \dots, D\varphi(\varphi^{-1}(x))(e^{n-1}))$$

is in the chosen orientation of $\mathbb{R}^n$. Hence given a choice of orientation we get a unit normal vector field (check that it is continuous).

Conversely, suppose that we are given a unit normal vector field $\mathbf{n} : S \to \mathbb{R}^n$. Then we get an induced orientation on each tangent space taking the frames $(v_1, v_2, \dots v_{n-1})$ such that $(\mathbf{n}(x), v_1, v_2, \dots, v_{n-1})$ is in the standard orientation of $\mathbb{R}^n$. It is easy to check that this gives a continuous orientation if $\mathbf{n}$ is continuous. □

This gives us another way to show that the sphere S^2 is orientable, since the vector field $\mathbf{n}(x) = x$ gives a nowhere vanishing normal vector field (and indeed this method also shows the n-sphere is orientable).

Since a choice of a unit normal vector field is just a formal way choosing a side of the hypersurface, this explains the relation between orientations and sides mentioned heuristically at the start of this section. The fact that a nonorientable hypersurface has no unit normal vector field is just a formal way of expressing the fact that, even though a hypersurface always locally appears to have two sides, you cannot always make this local fact globally consistent, (as the Möbius band shows).

4. SURFACES WITH BOUNDARY

We now want to extend our notion of surfaces to one which allows a boundary – for example we certainly want to be able to say that the closed unit disc in the plane is a surface with a boundary consisting of the unit circle. Our original definition of a surface insisted that each point had a neighborhood which looked like $\mathbb{R}^k$, but this makes each point "interior" to the surface. In order to allow boundaries, we need another local picture to pemit boundary points. We do this in the simplest way: Let $H^k = \{(x_1, x_2, \ldots, x_k) \in I^k : x_1 \leq 0\}$ be the half-cube of points in I^k with nonpositive first coordinate. Clearly I^{k-1} thought of as the points in I^k with $x_1 = 0$ is the boundary of H^k (in the sense of point set topology).

Definition 4.1. A subset S of $\mathbb{R}^n$ is a surface with boundary if for each point $x \in S$ there is a neighborhood U of x in S and either a homeomorphism $\varphi : I^k \to U$ or a homeomorphism $\psi : H^k \to U$.

We obtain the notion of a $\mathcal{C}^{(m)}$-smooth k-surface by insisting that the charts are $\mathcal{C}^{(m)}$ and their derivatives have maximal rank (the notion of differentiablility for functions on H^k is that of Definition 1.6). For a surface with boundary we have a natural notion of a boundary.

Definition 4.2. A point $x \in S$ is said to be a *boundary* point if for some chart $\varphi : H^k \to S$ we have $\varphi^{-1}(x) \in \partial H^k = I^{k-1}$. The set of boundary points is denoted ∂S.

Notice that if the preimage of x lies in ∂H^k for some chart whose image contains x then this will be true of any chart whose image contains x, since the inverse function theorem implies that diffeomorphisms send boundary points to boundary points. We say that a surface has no boundary if $\partial S = \emptyset$.

Lemma 4.3. *The boundary ∂S is a smooth $k-1$ surface (without boundary).*

Proof. Let $x \in \partial S$, and suppose that $\varphi\colon I^k \to S$ is a chart such that $\varphi^{-1}(x) \in I^{k-1}$. Then $\varphi_{|I^{k-1}}$ is a chart for ∂S near x. It is now also clear that ∂S is a surface without boundary. □

Lemma 4.4. *If S is an oriented surface, then so is ∂S.*

Proof. Suppose that we have an atlas of compatible charts for S. Let $\varphi\colon H^k \to U$ and $\psi\colon H^k \to V$ be two of these charts such that $U \cap V \neq \emptyset$. Let $\theta\colon \psi^{-1}(U \cap V) \to \varphi^{-1}(U \cap V)$ be the composite $\varphi^{-1} \circ \psi$. Since θ takes I^{k-1} to itself, it follows that on I^{k-1} we have $\frac{\partial \theta_1}{\partial t_i} = 0$ for $i > 1$. Moroever since it maps H^k to itself,

$t_1 \leq 0 \iff \theta_1(t) \leq 0$ and so $\frac{\partial\theta_1}{\partial t_1} > 0$ for $t \in I^{k-1}$. Thus the derivative $D\theta$ at points on I^{k-1} looks like

$$\begin{pmatrix} \frac{\partial\theta_1}{\partial t_1} & 0 & \dots & 0 \\ \frac{\partial\theta_2}{\partial t_1} & \dots & \frac{\partial\theta_2}{\partial t_{k-1}} & \frac{\partial\theta_2}{\partial t_k} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial\theta_k}{\partial t_1} & \dots & \frac{\partial\theta_k}{\partial t_{k-1}} & \frac{\partial\theta_k}{\partial t_k} \end{pmatrix}.$$

Hence it is clear that the restrictions $\varphi_{|I^{k-1}}$ and $\psi_{|I^{k-1}}$ are compatible. □

Remark 4.5. Later we will need to be precise about the actual orientations used. Let $\{\varphi_i\colon D_i^k \to U_i \subset S\}_{i\in I}$ be an atlas for, where D_i^k is either I^k or H^k. For $x \in \partial S$, pick φ_i such that $x \in U_i$. Then we say that $(v_1, v_2, \dots, v_{k-1})$ is a positively oriented frame in $T_x\partial S$ if $(e^1, D\varphi_i(x)^{-1}(v_1), D\varphi_i(x)^{-1}(v_2), \dots, D\varphi_i(x)^{-1}(v_{k-1}))$ is in the orientation class of $(e^1, e^2, \dots, e^k)$.

Example 4.6. In the previous section the Möbius band M was mentioned as an example of a nonorientable surface. There it is described as an open surface, but if we instead take it's closure

$$\bar{M} = \{\mathbf{e}(\varphi) + t\mathbf{f}(\varphi) : \varphi \in [0, 2\pi], t \in [-\frac{1}{2}, \frac{1}{2}]\}$$

we get an example of a 2-surface with boundary. It's instructive to check this and to see that $\partial\bar{M}$ is a circle.

5. DIFFERENTIAL FORMS ON $\mathbb{R}^n$

In this section we use the exterior algebra $\Lambda^*(\mathbb{R}^n)$ (see the appendix for a detailed discussion, or [E] chapter V sections 3 and 5). A differential k-form on an open subset U of $\mathbb{R}^n$ assigns to each point $x \in U$ an alternating k-multilinear function in smoothly varying way, that is, a differential k-form is a smooth function $\omega\colon U \to \Lambda^k(\mathbb{R}^n)$. In the notation of the appendix, ω may be written in the form

$$\omega(x) = \sum_I f_I(x) dx_I,$$

where I runs over all subsets of $\{1, 2, \dots, n\}$ of size k, and the requirement that ω be smooth is simply that each $f_I\colon \mathbb{R}^n \to \mathbb{R}$ is a smooth function. We write $\Omega^k(U)$ for the set (in fact vector space) of differential k-forms on U. Since $\Lambda^*(\mathbb{R}^n)$ is an algebra under the $\wedge$ product, the space of all differential forms $\Omega^*(U) = \bigoplus_k \Omega^k(U)$ is also: given α and β we set

$$(\alpha \wedge \beta)(x) = \alpha(x) \wedge \beta(x).$$

$\Omega^0(U)$ is just the vector space of smooth functions on U. Since the derivative of a smooth function $f\colon U \to \mathbb{R}$ is a smooth map which assigns to each point $x \in U$ a linear map $\mathbb{R}^n \to \mathbb{R}$, (and any such linear map is alternating), we see that the derivative can be thought of as a map $d\colon \Omega^0(U) \to \Omega^1(U)$. Explicitly, given $f \in \Omega^0(U)$

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n \in \Omega^1(U).$$

We want to extend d to a map which sends k-forms to $(k+1)$-forms.

Definition 5.1. Let $d^k\colon \Omega^k(U) \to \Omega^{k+1}(U)$ be the linear map defined by

$$d^k(f_I dx_I) = df_I \wedge dx_I.$$

for each k-subset $I \subset \{1, 2, \ldots, n\}$.

(When there is no possibility for confusion, we usually drop the superscript k. We denote the derivative by d when dealing with real-valued functions, where it coincides with d^0, while we will write D for the derivative of functions taking values in $\mathbb{R}^n$.) The map d is known as the *exterior derivative*. Its basic properties of the map d are as follows:

Lemma 5.2. *If $\alpha \in \Omega^k(U)$ and $\beta \in \Omega^l(U)$ then*

(1)
$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta.$$
(2) $d \circ d(\alpha) = 0$.

Proof. For the first statement, it suffices to check the case where $\alpha = f_I dx_I$ and $\beta = g_J dx_J$. Then

$$\begin{aligned}
d(\alpha \wedge \beta) &= d(f_I g_J dx_I \wedge dx_J) \\
&= (f_I dg_J + g_J df_I) \wedge dx_i \wedge dx_J \\
&= (d(f_I)dx_I) \wedge (g_J dx_J) + f_I(dg_J) \wedge dx_I \wedge dx_J \\
&= (d(f_I)dx_I) \wedge (g_J dx_J + (-1)^k f_I dx_I \wedge (dg_J) \wedge dx_J \\
&= d\alpha \wedge (\beta) + (-1)^k \alpha \wedge d\beta.
\end{aligned}$$

For the second part, we again need only check on the forms $f_I dx_I$. But then

$$\begin{aligned}
d \circ d(f_I dx_I) &= d(df_I \wedge dx_I) \\
&= (d^2(f_I) \wedge dx_I) - (df_I) \wedge d(dx_I) \\
&= (d^2(f_I)) \wedge dx_I,
\end{aligned}$$

since $d(dx_I) = 0$ by definition. Thus we are reduced to checking that $d(df_I) = 0$. But now

$$\begin{aligned}
d(df_I) &= d(\sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i) \\
&= \sum_{i=1}^n (\sum_{j=1}^n \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i}) dx_j) \wedge dx_i \\
&= \sum_{1 \leq i,j \leq n} (\frac{\partial^2 f}{\partial x_i \partial x_j} - \frac{\partial^2 f}{\partial x_j \partial x_i}) dx_i \wedge dx_j.
\end{aligned}$$

But this last expression vanishes by the symmetry of mixed partial derivatives (thus we needed the f_I to be at least $\mathcal{C}^{(2)}$). □

Remark 5.3. The operator $d\colon \Omega^k(U) \to \Omega^{k+1}(U)$ is determined by the conditions:

1. d is linear, that is $d(\alpha + \beta) = d\alpha + d\beta$,
2. d is an antiderivation: $d(\alpha \wedge \beta) = d(\alpha) \wedge \beta + (-1)^k \alpha \wedge d\beta$ for $\alpha \in \Omega^k(U)$.
3. $d \circ d\colon \Omega^k(U) \to \Omega^{k+2}(U)$ is zero.
4. On $\Omega^0(U)$, the space of smooth functions, d is the derivative.

In order to extend the notion of differential forms to surfaces, we need to first understand how they behave under smooth maps from $\mathbb{R}^k$ to $\mathbb{R}^n$. Let $\psi : \mathbb{R}^k \to \mathbb{R}^n$ be a smooth map. Recall that $\Omega^0(U)$ is simply the set of smooth functions on U. Now given $f \in \Omega^0(U)$ the map ψ allows us to define an element of $\Omega^0(\psi^{-1}(U))$ by setting $\psi^*(f)(x) = f(\psi(x))$. We want to generalize this to obtain a map

$$\psi^* \colon \Omega^p(U) \to \Omega^p(\psi^{-1}(U)).$$

First observe that any linear map from $\mathbb{R}^n$ to itself induces a pullback on multilinear functions on $\mathbb{R}^n$. That is, if $A : \mathbb{R}^n \to \mathbb{R}^n$ is linear, there is a map A^* on the space of multilinear functions given by setting, for $M \colon (\mathbb{R}^n)^k \to \mathbb{R}$ multilinear,

$$A^*(M)(v_1, v_2, \dots, v_k) = M(A(v_1), A(v_2), \dots, A(v_k)),$$

for $v_1, v_2, \dots, v_k \in \mathbb{R}^n$. Since $A^*(M)$ is clearly alternating if M is, A^* also restricts to give a pullback on Ω^k. But now we are requiring that ψ is smooth, and so at each $x \in \psi^{-1}(U)$ we have a linear map $D\psi(x) : \mathbb{R}^k \to \mathbb{R}^n$. Thus we obtain our desired pullback of forms by combining the two pullbacks we have just described:

Definition 5.4. Given $\psi \colon \mathbb{R}^k \to \mathbb{R}^n$ and $\alpha \in \Omega^p(U)$ let $\psi^*(\alpha)$ be given at $t \in \psi^{-1}(U)$ by setting, $\psi^*(\alpha)(t) = (D\psi(t))^*(\alpha(\psi(t))$, that is,

$$\psi^*(\alpha)(t)(v_1, v_2, \dots, v_p) = \alpha(\psi(t))(D\psi(t)(v_1), D\psi(t)(v_2), \dots, D\psi(t)(v_p)),$$

$v_1, v_2, \dots v_p \in \mathbb{R}^k$. Notice that we need ψ to be at least differentiable in order to be able to pull back at all. Since ψ is infinitely differentiable, it is straightforward to check (though elaborate to write out explicitly) that $\psi^*(\alpha)$ is a smooth form if α is (see the example below).

We now show that ψ^* is has all the compatibilities we could want.

Lemma 5.5. *Let $\alpha, \beta \in \Omega^*(U)$, and $\psi : \mathbb{R}^k \to \mathbb{R}^n$ be a smooth map. Then we have*

1. $\psi^*(\alpha \wedge \beta) = \psi^*(\alpha) \wedge \psi^*(\beta)$
2. $d(\psi^*(\alpha)) = \psi^*(d\alpha)$.

Proof. For the first part notice that the wedge product is defined pointwise, and so we only need to show that if $L \colon \mathbb{R}^n \to \mathbb{R}^k$ is a linear map and M, N are alternating multilinear functions, then

$$L^*(M \wedge N) = L^*(M) \wedge L^*(N).$$

We show this in the appendix.

For the second part, we first prove it for functions, *i.e.* for $f \in \Omega^0(U)$. In this case for $x \in \psi^{-1}(U)$ and $v \in \mathbb{R}^k$, we have

$$d(\psi^*(f))(x)(v) = d(f \circ \psi)(x)(v) = df(\psi(x)) \circ d\psi(x)(v) = \psi^*(df)(x)(v).$$

by the chain rule. Now given a form $\alpha = f dx_I$ where $I = (i_1 < i_2 < \dots < i_k)$, we see that if $\psi = (\psi_1, \psi_2, \dots, \psi_n)$ we have

$$\begin{aligned} d\psi^*(f dx_I) &= d(\psi^*(f)\psi^*(dx_{i_1}) \wedge \dots \wedge \psi^*(dx_{i_k})) \\ &= d(\psi^*(f) d\psi_{i_1} \wedge \dots \wedge d\psi_{i_k}), \end{aligned}$$

using the result for 0-forms, since $\psi^*(x_i) = \psi_i$. But then using Lemma 5.2 and the result for the 0-form f we see that

$$\begin{aligned} d(\psi^*(f) d\psi_{i_1} \wedge \ldots \wedge d\psi_{i_k}) &= d(\psi^*(f)) \wedge d\psi_{i_1} \wedge \ldots \wedge d\psi_{i_k} \\ &= \psi^*(df) \wedge d\psi_{i_1} \wedge \ldots \wedge d\psi_{i_k} \\ &= \psi^*(df \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_k}), \end{aligned}$$

where in the last equality we once again used the result for 0-forms. □

Example 5.6. Consider the special case where $\psi\colon \mathbb{R}^n \to \mathbb{R}^n$. Then if U is an open subset of $\mathbb{R}^n$ and $\alpha \in \Omega^n(U)$, we want to compute what $\psi^*(\alpha)$ is. We may write $\alpha = f dx_1 \wedge dx_2 \wedge \ldots dx_n$ where f is a smooth function on U. but then

$$\begin{aligned} \psi^*(\alpha)(x) &= f(\psi(x)) d\psi_1 \wedge d\psi_2 \wedge \ldots \wedge d\psi_n \\ &= f(\psi(x)) \det(D\psi(x)) dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n. \end{aligned}$$

The close relation of this expression to the change of variables formula is what will allow us to define the integral of a k-form on an orientable k-surface.

Example 5.7. Suppose that $\psi\colon \mathbb{R}^2 \to \mathbb{R}^3$ so that

$$\psi(t_1, t_2) = (\psi_1(t_1, t_2), \psi_2(t_1, t_2), \psi_3(t_1, t_2)),$$

and α be the 2-form $dx_1 \wedge dx_3$. Then using the above properties of the pullback ψ^* we see that

$$\begin{aligned} \psi^*(dx_1 \wedge dx_3) &= \psi^*(dx_1) \wedge \psi^*(dx_3) \\ &= d(\psi^*(x_1)) \wedge d(\psi^*(x_3)) \\ &= d\psi_1 \wedge d\psi_3 \\ &= (\frac{\partial \psi_1}{\partial t_1} dt_1 + \frac{\partial \psi_1}{\partial t_2} dt_2) \wedge (\frac{\partial \psi_3}{\partial t_1} dt_1 + \frac{\partial \psi_3}{\partial t_2} dt_2) \\ &= (\frac{\partial \psi_1}{\partial t_1} \frac{\partial \psi_3}{\partial t_2} - \frac{\partial \psi_3}{\partial t_1} \frac{\partial \psi_1}{\partial t_2}) dt_1 \wedge dt_2. \end{aligned}$$

Example 5.8. We define a 2-form ω on $\mathbb{R}^3$ as follows: given $x \in \mathbb{R}^3$ let

$$\omega(x)(v, w) = \det \begin{pmatrix} x_1 & v_1 & w_1 \\ x_2 & v_2 & w_2 \\ x_3 & v_3 & w_3 \end{pmatrix}$$

Expanding this by the first column we see that with respect to the forms $\{dx_i \wedge dx_j : 1 \leq i, j \leq 3\}$ the form is

$$\omega(x) = x_1 dx_2 \wedge dx_3 - x_2 dx_1 \wedge dx_3 + x_3 dx_1 \wedge dx_2,$$

then the same kind of calculation as the previous example shows that $\psi^*(\omega)$ is

$$\begin{aligned} \big(\psi_1(\frac{\partial \psi_2}{\partial t_1} \frac{\partial \psi_3}{\partial t_2} - \frac{\partial \psi_3}{\partial t_1} \frac{\partial \psi_2}{\partial t_2}) - \psi_2(\frac{\partial \psi_1}{\partial t_1} \frac{\partial \psi_3}{\partial t_2} - \frac{\partial \psi_3}{\partial t_1} \frac{\partial \psi_1}{\partial t_2}) \\ + \psi_3(\frac{\partial \psi_1}{\partial t_1} \frac{\partial \psi_2}{\partial t_2} - \frac{\partial \psi_2}{\partial t_1} \frac{\partial \psi_1}{\partial t_2})\big) dt_1 \wedge dt_2. \end{aligned}$$

Hence if, say, $\psi(t_1, t_2) = (\cos(t_2)\sin(t_1), \sin(t_2)\sin(t_1), \cos(t_1))$ then we find

$$\psi^*(\alpha) = \sin(t_1) dt_1 \wedge dt_2.$$

6. PARTITIONS OF UNITY

This section proves a fundamental result which allows us to move from local considerations on a surface to global ones. The key fact is that smooth functions are local objects. To make this idea precise we make the following definitions.

Definition 6.1. Given a topological space X and a continuous function $\rho : X \to \mathbb{R}$ we say the *support* of ρ to be

$$\overline{\{x \in X : \rho(x) \neq 0\}}.$$

An *open covering*, or simply *cover* of X is a collection of open sets $\{U_i\}_{i \in I}$ in X such that $X = \cup_{i \in I} U_i$. We say that an cover $\{V_j\}_{j \in J}$ *refines* the cover $\{U_i\}_{i \in I}$ if for each open set V_j there is some U_i such that $V_j \subset U_i$. We say the covering $\{V_j\}_{j \in J}$ is *locally finite* if each $x \in X$ has a neighborhood which intersects only finitely many of the V_j.

We can now define the key concept of this section.

Definition 6.2. Let X be a topological space with a locally finite open cover $\mathcal{U} = \{U_i\}_{i \in I}$. A partition of unity subordinate to $\mathcal{U}$ is a collection of continuous functions $\{\rho_i \colon X \to [0,1]\}_{i \in I}$ such that $\mathrm{supp}(\rho_i) \subset U_i$ and

$$\sum_{i \in I} \rho_i(x) = 1,$$

where this sum is finite for each $x \in X$ since the cover is locally finite.

We are interested in the case where X is a smooth surface (from now on, we mean $\mathcal{C}^\infty$ when we say smooth), and when the covering is given by the images of the charts of an atlas. A smooth partition of unity for X is then a partition of unity where the functions ρ_i are $\mathcal{C}^\infty$. The main result of this section is that such partitions always exist. We begin with an easy lemma which produces a supply of compactly supported smooth functions.

Lemma 6.3. *Let* $x_0 = (a_1, a_2, \ldots, a_k) \in \mathbb{R}^k$*, and let* $I^k_r(x_0)$ *be the open interval* $\{x \in \mathbb{R}^k : |a_i - x_i| < r, \forall i \in \{1, 2, \ldots, k\}\}$. *Then there is a smooth function* $\rho : \mathbb{R}^k \to [0,1]$ *such that* $\rho(x) > 0$ *on* $I^k_r(x_0)$ *and* $\rho(x) = 0$ *outside of* $I^k_r(x_0)$.

Proof. Let $\psi \colon \mathbb{R} \to \mathbb{R}$ be given by

$$\psi(t) = \begin{cases} \exp(-1/x), & x > 0 \\ 0, & x \leq 0 \end{cases}$$

It is easy to check that this is a smooth function on all of $\mathbb{R}$. Thus $\chi(t) = \psi(1+t)\psi(1-t)$ is a smooth function vanishing outside $(-1,1)$ and positive on $(-1,1)$. Setting

$$\rho(x) = \prod_{i=1}^{k} \chi(\frac{a_i - x_i}{r}),$$

it is then clear that ρ has the required properties. $\square$

The function ρ in the lemma is known as a *bump function*. We now prove that partitions of unity exist. Here we will state the most general result, but only demonstrate it only for the case of a compact surface, relegating the proof of the general case to an appendix.

Theorem 6.4. *Let S be a smooth $\mathcal{C}^\infty$ k-surface in $\mathbb{R}^n$, and let $\{U_i\}_{i\in I}$ be a cover of S. Then there is a locally finite refinement $V_{j\,j\in J}$ of the cover $\{U_i\}$ such that $\bar{V_j} \subset U_{i(j)}$ if $V_j \subset U_{i(j)}$, and a partition of unity $\{\rho_j\}_{j\in J}$ subordinate to $\{V_j\}_{j\in J}$. Moreover one can assume that J is countable.*

Proof. We give a proof here only for the case where S is compact. The general case is relegated to an appendix. Fix a smooth function $\rho\colon I^k \to \mathbb{R}$ as in the previous lemma, which vanishes outside $I^k_{1/2}$ and is positive in $I^k_{1/2}$. For each point $x \in S$, pick a coordinate neighborhood $\varphi_x\colon I^k \to S$ with image V_x in some U_i, and let $\psi_x\colon V_x \to \mathbb{R}$ be $\varphi_x \circ \rho \circ \varphi_x^{-1}$ so that we can extend ψ_x to a smooth function on all of S by letting it equal zero outside V_x. If $W_x = \psi_x(I^k_{1/2})$, the collection $\{W_x : x \in S\}$ is a covering of S, so that by compactness we may take a finite subcover of the $\{W_x : x \in S\}$, say $\{W_{x_1}, W_{x_2}, \dots, W_{x_m}\}$ (this cover is then clearly also locally finite). Then set

$$\rho_j = \psi_{x_j}/(\sum_{k=1}^{m} \psi_{x_k}).$$

(since each $x \in S$ lies in some W_{x_k} the denominator of the expression is never zero). It is clear that the functions ρ_j are a partition of unity as required (the refinement in the statement of the theorem is the cover $\{V_{x_1}, V_{x_2}, \dots V_{x_m}\}$, and the indexing set is $J = \{1, 2, \dots, m\}$ is finite and so countable.) □

The results of this section imply that smooth functions are completely local objects – if you know the values of a smooth function on some closed set in $\mathbb{R}^k$, you know *absolutely nothing* about that function outside of that set. We use the following corollary to show this.

Corollary 6.5. *Let $U \subset \mathbb{R}^k$ be an open set, and let A be a closed subset of U. Then there is a smooth function $\psi\colon \mathbb{R}^k \to \mathbb{R}$ such that $\psi(x) = 1$ for $x \in A$, and $\psi(x) = 0$ outside of U.*

Proof. Consider the cover $\{U, \mathbb{R}^k - A\}$ and apply the previous theorem to obtain a locally finite refinement $\{V_j\}_{j\in\mathbb{N}}$, and smooth functions $\{\eta_j\}_{j\in\mathbb{N}}$ with $\operatorname{supp}(\eta_j) \subset V_j$. Then set ψ to be the sum of the η_j with $V_j \subset U$. □

Of course this corollary used the full version of the theorem, which we prove in an appendix. The function ψ is known as a *cut-off function*. If we are given a smooth function f and a closed set A of $\mathbb{R}^k$ on which it is defined and any point $y \notin A$, we can find a cut-off function ψ such that ψf vanishes at x but is equal to f on all of A. Thus the existence of smooth functions is a purely local question. This will be essential for us in, for example, establishing Stokes' theorem, or even more basically in defining what it means to integrate a differential form.

7. DIFFERENTIAL FORMS ON SURFACES

We now extend our definition of differential forms to a smooth k-surface S. A p-form on S assigns to each point $x \in S$ an alternating multilinear function of degree p on the space T_xS. It is smooth, said differently, is a differential p-form, if the multilinear functions vary smoothly with $x \in S$. Now the tangent space T_xS is a k-dimensional subspace of $\mathbb{R}^n$. If you are comfortable with alternating functions on abstract vector spaces, then all you need to make sense of an alternating function on T_xS is the fact that T_xS is a vector space. If you do not want to think this way, it is also legitimate to assume that the alternating function is defined on all of $\mathbb{R}^n$ and we are just restricting it to tuples of vectors in T_xS (see Remark 10.23 in the appendix for more details).

There are two ways of making the notion of "smoothly varying" precise, just as in section 1 we had two ways of describing differentiability (or smoothness) for functions on a surface. Indeed since 0-forms are just smooth functions, we are simply generalizing that discussion. Our first definition has the advantage of being explicit.

Definition 7.1. A p-form ω on a surface is smooth if for each $x \in S$ there is a neighborhood N of x in $\mathbb{R}^n$ and $\alpha \in \Omega^p(N)$ such that ω is the restriction of α to $N \cap S$.

Notice here that "restriction" is being used in two senses: The first being that we restrict α from U to the points of S, and the second being that for each $x \in S$, we restrict the alternating function $\alpha(x)$ to the linear subspace $T_xS \subset \mathbb{R}^n$. The disadvantage of this definition, although it is entirely rigorous, is that it is hard to test if a given p-form actually *is* a differential p-form. In practice, we compute most things using charts, and so our second definition is in terms of charts for the surface.

Definition 7.2. A differential p-form ω on S is an assignment for each $x \in S$ of an alternating p-form $\omega(x)$ on T_xS such that if $\psi\colon I^k \to U \subset S$ is a chart then $\psi^*(\omega)$ is a differential p-form on I^k.

It follows from our discussion of pullbacks in section 5 that if we have any atlas $\{\psi_i : I^k \to S\}_{i \in I}$ for which $\psi_i^*(\omega)$ is a differential p-form for each ψ_i, then ω is a differential form on S. Moreover it is clear that a p-form which is smooth in the sense of our first definition is smooth in the sense of our second definition. In fact the two definitions are equivalent.

Proposition 7.3. $(*)$ *If ω is a p-form in the sense of Definition 7.2 then for each $x \in S$ there exists an open set N in $\mathbb{R}^n$ and $\alpha \in \Omega^p(N)$ such that $\omega_{|N\cap S} = \alpha_{|N\cap S}$.*

Proof. As noted above, it is clear that a p-form which is smooth in the sense of Definition 7.1 is smooth in the sense of Definition 7.2. This is proved in the same manner as Lemma 1.7. Suppose that $x \in S$ and $\varphi\colon I^k \to S$ is a coordinate chart such that $\varphi(0) = x$. Informally, if $(t_1, t_2, \dots t_k)$ are the coordinates on I^k then Proposition 1.4 says we can think of the t_i as smooth function on some neighborhood of x in $\mathbb{R}^n$. But then since in terms of the t_is we have $\omega = \sum_I f_I dt_I$ it is clear we can view this expression as giving a smooth form on that neighborhood as required. More formally, Proposition 1.4 shows that we can find a smooth diffeomorphism $\Phi\colon I^n_\varepsilon \to N \subset \mathbb{R}^n$ such that $\Phi_{|I^k \cap I^n_\varepsilon} = \varphi_{|I^k \cap I^n_\varepsilon}$ for some $\varepsilon > 0$ (and hence N is a neighborhood of x). Let $\pi\colon \mathbb{R}^n \to \mathbb{R}^k$ be the obvious projection map, sending $(t_1, t_2, \dots, t_n) \mapsto (t_1, t_2, \dots t_k)$. We can then define $\alpha \in \Omega^p(U)$

$$\alpha = (\Phi^{-1})^* \circ \pi^* \circ \psi^*(\omega).$$

It is clear that α is a differential form on N such that $\alpha_{|S\cap N} = \omega$ as required. □

Denote the space of differential p-forms on S by $\Omega^p(S)$. It is clear that there is a wedge product on $\Omega^*(S)$. Somewhat more subtly, there is an exterior derivative d

Lemma 7.4. *Given a differential form $\omega \in \Omega^p(S)$ there is a unique differential form $d\omega \in \Omega^{k+1}(S)$ such that for each chart $\psi : I^k \to S$*

$$\psi^*(d\omega) = d\psi^*(\omega).$$

Proof. Proof of the lemma in terms of the second definition: Because the differential $D\psi(t)$ of a chart $\psi\colon I^k \to S$ gives an isomorphism between $\mathbb{R}^k$ and $T_{\psi(t)}(S)$, it is clear that such a differential form is unique if it exist. Let $\{\varphi_i\colon I^k \to U_i \subset S\}_{i\in I}$ be an atlas for S. Then on each U_i we can consider $\varphi_i^*(\omega) \in \Omega^p(I^k)$. Let $\beta = d(\varphi_i^*(\omega)) \in \Omega^{p+1}(I^k)$. Then define

$$d\omega(x) = (D\varphi_i(\varphi_i^{-1}(x))^{-1})^*(\beta(\varphi^{-1}(x)).$$

Lemma 5.5 shows that this gives a well defined $(k+1)$-form on S as required. □

Because of the compatibility of d with pullbacks, it is easy to see that if ω is given in a neighborhood as $\alpha_{|S}$ then $d\omega = (d\alpha)_{|S}$ in that neighborhood. We can also extend the notion of pullback to the case of smooth maps $\psi\colon S_1 \to S_2$ between surfaces S_1 and S_2. Again there is an issue concerning what we mean by a smooth map between surfaces – there is a definition in terms of charts, and another in terms of restrictions, and again they coincide. The definition in terms of charts is given in Remark 1.13. The important point is that a smooth map ψ induces, for $x \in S_1$ a linear map $D\psi(x)\colon T_xS_1 \to T_{\psi(x)}S_2$, and hence given $\alpha \in \Omega^p(S_2)$, in the same fashion as we did for maps from $\mathbb{R}^k \to \mathbb{R}^n$, we can define $\psi^*(\alpha) \in \Omega^p(S_1)$ by setting

$$\psi^*(\alpha)(t) = (D\psi(t))^*(\alpha(\psi(t))),$$

for $t \in \mathbb{R}^k$, as we did in the case of a smooth map $\mathbb{R}^k \to \mathbb{R}^n$. Then we check on charts that this does indeed give a smooth form. It is then also routine (though tedious) to check via charts that,

$$d\psi^*(\alpha) = \psi^* d\alpha.$$

In most cases that we will deal with the map ψ will the the restriction to S of a map from $\mathbb{R}^n$, and so we can work directly with an extension of our form (again in most cases the form we work with will be the restriction of a form on some open subset of $\mathbb{R}^n$ also). However it is important in future generalizations of the material we are discussing to realize that all these concepts can be expressed just in terms of the chart maps.

We also want to mention the relation between orientations and k-forms on a k-surface, as an example of using Definition 7.2.

Proposition 7.5. *A k-surface S in orientable if and only if there exists nowhere vanishing differential k-form on S.*

Proof. This is just a globalization of Lemma 3.1. Suppose that S is orientable, and let $\{\varphi_i\colon I^k \to U_i \subset S\}_{i\in I}$ be an atlas of compatible charts. Then on each U_i we have the form ω_i defined by the condition that $\varphi_i^*(\omega_i) = dt_1 \wedge dt_2 \wedge \ldots \wedge dt_k \in \Omega^k(I^k)$. Let $\{\rho_m\}_{m\in\mathbb{N}}$ be a partition of unity subordinate to $\{U_i\}_{i\in I}$ and set $\omega = \sum_{m\in\mathbb{N}} \rho_m \omega_{i(m)}$ where $i(m)$ is such that $\operatorname{supp}(\rho_m) \subset U_{i(m)}$. It is easy to check using Example 5.6 that the compatibility of the charts ensures that ω never vanishes: indeed if $x \in S$ is such that $\rho_n(x) \neq 0$ for some $n \in \mathbb{N}$ then if $j = i(n)$ we can pullback to the domain of the chart φ_j to see that, if $\varphi_j(t) = x$ then

$$\varphi_j^*(\omega)(t) = \sum_m \varphi_j^*(\rho_m)(t) \det(D\theta_{j,i(m)}(t)) dt_1 \wedge dt_2 \wedge \ldots \wedge dt_k,$$

where $\theta_{j,i(m)} = \varphi_i^{-1} \circ \varphi_{i(m)}$. But the compatibility of the charts says exactly that $\det(D\theta_{j,i(m)}(t)) > 0$, and so it is clear that ω is nonzero at $x = \varphi_j(t)$.

Conversely, suppose that ω is a nowhere vanishing differential k-form on S, and let $\{\varphi_i I^k \to U_i \subset S\}_{i \in I}$ be an atlas. Then on each U_i we may write $\varphi_i^*(\omega) = f(t) dt_1 \wedge dt_2 \wedge \ldots dt_k$, where $f(t) \neq 0$ for all $t \in I^k$. But then clearly either $f(t) > 0$ for all t or $f(t) < 0$ for all t. In the first case set $\psi_i = \varphi_i$ while in the latter set $\psi_i(t_1, t_2, \ldots, t_k) = \varphi_i(-t_1, t_2, \ldots, t_k)$. It is then easy to see that $\{\psi_i\}_{i \in I}$ is an atlas of compatible charts for S. □

We will call such a form an *orientation form* for S. We wish to make a specific choice of orientation form, known as the *volume form*: Given an oriented surface S in $\mathbb{R}^n$ then there is an unique k-form ν_S such that at each $x \in S$ we have

$$\nu_S(x)(v_1, v_2, \ldots, v_k) = 1,$$

for all orthonormal frames $(v_1, v_2, \ldots, v_k)$ in the specified orientation on $T_x S$ or equivalently, by the linear change of variables formula, ν_S assigns to any frame in the orientation the volume of the parallelepiped described by that frame. It is then easy to check that if $\varphi \colon I^k \to S$ is a chart,

$$\varphi^*(\nu_S) = \sqrt{\det(D\varphi(t)^t D\varphi(t))} dt_1 \wedge dt_2 \wedge \ldots \wedge dt_k.$$

You can check that the volume form on the 2-sphere S^2 is the 2-form in Example 5.8. We define $\mathrm{vol}(S) = \int_S \nu_S$ to be the volume of S

8. INTEGRATION OF FORMS

If U is an open subset of $\mathbb{R}^n$ and α is a differential n-form on U with compact support, then there is a smooth function $f \colon \mathbb{R}^n \to \mathbb{R}$ such that $\alpha = f dx_1 \wedge \ldots \wedge dx_n$. We define

$$\int_U \alpha = \int_U f dx,$$

i.e. the integral of f over U (apologies for the fact that here "dx" on the right hand side of the equation isn't a differential form, just the standard notation for the Riemann integral). Now let $\psi \colon V \to U$ be a diffeomorphism between open subsets of $\mathbb{R}^n$ then

$$\begin{aligned} \int_V \psi^*(\alpha) &= \int_V f(\psi(x)) \det(D\psi(x)) dx_1 \wedge \ldots \wedge dx_n \\ &= \int_V f(\psi(x)) \det(D\psi(x)) d\mu = \pm \int_U f d\mu \\ &= \pm \int_U \alpha, \end{aligned}$$

using the change of variables formula in the second line. Thus we see that

$$\int_V \psi^*(\alpha) = \pm \int_U \alpha \tag{8.1}$$

with the sign ± according as ψ preserves or reserves the orientation of $\mathbb{R}^n$. This suggests that we should be able to define a consistent notion of the integral of a differential k-form over a k-surface when that surface is orientable. In order to ensure that our integrals exist, we need to suppose that our differential forms have compact support, that is, $\omega \in \Omega^k(S)$ has compact support if

$$\{x \in S : \omega(x) \neq 0\}$$

has compact closure. We write $\Omega^p_c(S)$ for the space of p-forms with compact support. For such forms we can make the following definition.

Proposition/Definition 8.1. Let S be a smooth oriented k-dimensional surface. There is a unique linear map $\int_S \colon \Omega^k_c(S) \to \mathbb{R}$ such that if $\varphi : D^k \to U \subset S$ is any chart compatible with the orientation, and α is a k-form such that $\text{supp}(\alpha) \subset U$ then

$$\int_S \alpha = \int_{D^k} \varphi^*(\alpha).$$

Proof. Let $\{\varphi_i \colon D^k_i \to U_i \subset S\}_{i \in I}$ be an atlas of charts compatible with the orientation of S (here the notation D^k_i stands for I^k or H^k depending on what kind of chart φ_i is), and let $\{\rho_m \colon S \to \mathbb{R}\}_{m \in \mathbb{N}}$ a partition of unity subordinate to $\{U_i\}_{i \in I}$ such that $\text{supp}(\rho_m) \subset U_{i(m)}$. Given any $\alpha \in \Omega^k(S)$ we have $\alpha = \sum_{j \in \mathbb{N}} \rho_m \alpha$. Each of the $\rho_m \alpha$ are supported in the image of a chart of the atlas, and as α has compact support there are only finitely many nonzero terms in the sum. Thus since the integral is supposed to be linear, we are led to define

$$\int_S \alpha = \sum_{m \in \mathbb{N}} \int_{D^k_i} \varphi^*_{i(m)}(\rho_m \alpha).$$

It is then easy to see using equation 8.1 and the compatibility of the charts in the atlas that $\int_S$ has the required property. □

Remark 8.2. Notice that we have used a partition of unity to do exactly what it was advertised to do – to let us go from a local situation to a global one. In this case we knew how to integrate a differential form locally (in the image of a single chart), and the partition of unity lets us use this to define the integral globally – *i.e.* on the whole of S.

We are now ready to state the main theorem of this course: Stokes' theorem. This is a vast generalization of the fundamental theorem of calculus, using all of the objects that we have constructed over the course. It is an example of a result where the true difficultly lies in finding what class of objects it is really a statement about, and this is the work we have been doing so far. Thus the actual proof at this point will seem almost a triviality, but this should not be allowed to be mislead one into thinking that the result is without content.

Theorem 8.3. *(Stokes' Theorem): Let S be an oriented k-surface in $\mathbb{R}^n$ with boundary and let $\alpha \in \Omega^{k-1}_c(S)$ be a $(k-1)$-form with compact support, then*

$$\int_S d\alpha = \int_{\partial S} \alpha.$$

Proof. We first reduce to the case where α is supported in the image of a single chart. As in the definition of the integral, take an atlas of charts $\{\varphi_i\}_{i \in I}$ compatible with the orientation, and take a partition of unity $\{\rho_s\}_{s \in \mathbb{N}}$ subordinate to a refinement of $\{U_i\}$.

$$\alpha = \sum_s \rho_s \alpha.$$

Applying the exterior derivative we see that

$$d\alpha = \sum_s d(\rho_s \alpha)$$

But each of the forms $\rho_s \alpha$ is supported in a compact subset of some U_i and so it is indeed enough to prove the theorem for α supported in the image of a single chart, say $\varphi : D^k \to U$, where D^k is either I^k or H^k. By definition we have

$$\int_S d\alpha = \int_{D^k} \varphi^*(d\alpha) = \int_{D^k} d\beta,$$

where $\beta = \varphi^*(\alpha) \in \Omega^{k-1}(D^k)$. Similarly we have

$$\int_{\partial S} \alpha = \int_{D^k \cap \varphi^{-1}(\partial S)} \beta,$$

and thus it is enough to show that

$$\int_{D^k} d\beta = \int_{D^k \cap \varphi^{-1}(\partial S)} \beta.$$

On D^k we can write

$$\beta = \sum_{i=1}^{k} (-1)^{i-1} f_i(t) dt_1 \wedge dt_2 \ldots \wedge \widehat{dt_i} \wedge \ldots \wedge dt_k,$$

where each of the f_i are smooth functions, and the notation

$$dt_1 \wedge dt_2 \ldots \wedge \widehat{dt_i} \wedge \ldots \wedge dt_k$$

denotes the exterior product of all the dt_p except dt_i. Hence applying the exterior derivative we see that

$$\int_{D^k} d\beta = \int_{D^k} (\sum_{i=1}^{k} \frac{\partial f_i}{\partial t_i}) dt_1 \wedge \ldots \wedge dt_k.$$

There are two cases: when D^k is I^k and when it is H^k. In the first case the functions f_i vanish outside of some compact subset of I^k by assumption and so we have

$$\begin{aligned} \int_{I^k} (\sum_{i=1}^{k} \frac{\partial f_i}{\partial t_i}) dt_1 \wedge \ldots \wedge dt_k &= \sum_{i=1}^{k} \int_{I^k} \frac{\partial f_i}{\partial t_i} dt \\ &= \sum_{i=1}^{k} \int_{s \in I^{k-1}} (\int_{-1}^{1} \frac{\partial f_i}{\partial t_i} dt_i) ds \\ &= \sum_{i=1}^{k} \int_{s \in I^{k-1}} [f_i(t)]_{t_i=-1}^{t_i=1} ds \\ &= 0, \end{aligned}$$

where in the second line Fubini's theorem allows us to integrate in any order we like. Since in this case the support of α does not intersect ∂S we are done.

In the case where $D_j^k = H^k$ the calculation is similar: the functions f_i now have to vanish outside of some compact subset of H^k. Moreover on $\partial H^k = I^{k-1}$ we have $dx_1 = 0$, thus

$$\beta_{|\partial H^k} = f_1(t) dt_2 \wedge \ldots \wedge dt_{k-1} \wedge dt_k.$$

Now we have

$$\begin{aligned}\int_{H^k} (\sum_{i=1}^{k} \frac{\partial f_i}{\partial t_i}) dt_1 \wedge \ldots \wedge dt_k &= \sum_{i=1}^{k} \int_{H^k} \frac{\partial f_i}{\partial t_i} dt \\ &= \int_{s \in I^{k-1}} (\int_{-1}^{0} \frac{\partial f_1}{\partial t_1} dt_k) ds + \sum_{i=2}^{k} \int_{s \in H^{k-1}} (\int_{-1}^{1} \frac{\partial f_i}{\partial t_i} dt_i) ds \\ &= \int_{s \in I^{k-1}} [f_k(t)]_{t_k=-1}^{t_k=0} ds + 0 \\ &= \int_{s \in I^{k-1}} f_k(s_1, s_2, \ldots, s_{k-1}, 0) ds.\end{aligned}$$

and so our choice of orientation for ∂S exactly guarantees that the last expression is just

$$\int_{H^k \cap \varphi^{-1}(U)} \beta.$$

□

9. APPLICATIONS OF STOKES' THEOREM

We start with an immediate corollary of Stokes' theorem: Recall that the unit ball B^n is the set $\{x \in \mathbb{R}^n : ||x|| \geq 1\}$, and the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : ||x|| = 1\}$, which is the boundary of B^n.

Theorem 9.1. *(Brower fixed point theorem): Let $F\colon B^n \to B^n$ be a smooth map from B^n to itself. Then F has a fixed point.*

Proof. Suppose for a contradiction that there was an F with $F(x) \neq x$ for all $x \in B^n$. Then we can define a map $G\colon B^n \to \partial B^n$ by taking the line segment from x to $F(x)$ and extending it until we meet S^{n-1}. Since F is smooth, the function G is also smooth. Now let ω be an $(n-1)$-form on S^{n-1} such that $\int_{S^{n-1}} \omega = 1$. Now the function G is the identity on S^{n-1} by construction, hence

$$\int_{\partial B^n} G^*(\omega) = \int_{S^{n-1}} \omega = 1.$$

However, applying Stokes' theorem we see that

$$\int_{\partial B^n} G^*(\omega) = \int_{B^n} dG^*(\omega) = \int_{B^n} G^*(d\omega) = 0,$$

since $d\omega = 0$ (there are no nonzero n-forms on an $(n-1)$-surface). Thus we have a contradiction. □

Remark 9.2. In fact, this theorem holds if we just assume that F is continuous. Surprising as it may seem, what we have proved is close to this – one can bootstrap up to an arbitrary continuous function once you know that smooth functions are dense (in a suitable sense) in the set of continuous functions.

Definition 9.3. Let S be a surface. We say that a differential form $\alpha \in \Omega^p(S)$ is *closed* if $d\alpha = 0$. We say that α is *exact* if $\alpha = d\beta$ for some $\beta \in \Omega^{p-1}(S)$. Since $d^2 = 0$ any exact form is automatically closed.

If S is a k-surface, then $\Omega^{k+1}(S) = 0$ and so *every* k-form is closed. Stokes theorem shows that if S is a compact surface without boundary then

$$H_S^k \colon \Omega^k(S) \to \mathbb{R}, \qquad \alpha \mapsto \int_S \alpha$$

vanishes on exact k-forms. We want to use this to attach to a smooth map between compact surfaces a numerical invariant called the *degree*. Before we can do this, we need to check that H_S^k is a nonzero map, that is we need to find a form whose integral is nonzero.

Lemma 9.4. *Let S be a compact oriented surface, and let α be an orientation form compatible with the orientation of S. The*

$$\int_S \alpha > 0.$$

In particular the map H_S^k is surjective, moreover since the volume form ν_S is an orientation form, we see that the volume of S is a positive number.

Proof. Let $\{\varphi_i \colon I^k \to U_i \subset S\}_{i \in I}$ be an atlas compatible with the orientation, and $\{\rho_m\}_{m \in \mathbb{N}}$ a partition of unity subordinate to $\{U_i\}_{i \in I}$ as in Definition 8.1. Then we have

$$\int_S \alpha = \sum_{m \in \mathbb{N}} \int_{I^k} \varphi_{i(m)}^*(\rho_m \alpha).$$

But now each of the terms of the sum is positive because the $\{\rho_m\}$ are nonnegative and $\varphi_i^*(\alpha) = g_i(t) dt_1 \wedge \ldots \wedge dt_k$ for some smooth function $g_i \colon I^k \to \mathbb{R}_{>0}$ for each $i \in I$. ☐

Definition 9.5. Let $f \colon S \to T$ be a smooth map between compact surfaces of dimension k. Then given $\alpha \in \Omega^k(T)$ the pullback $f^*(\alpha) \in \Omega^k(S)$ is exact if and only if α is, because the exterior derivative commutes with pullback. Let $\nu_T \in \Omega^k(T)$ be an orientation form on T normalized so that $\int_T \nu_T = 1$. We define the *degree* of f to be

$$\deg(f) = \frac{\int_S f^*(\nu_T)}{\int_T \nu_T} = \frac{1}{\mathrm{vol}(T)} H_S^k(f^*(\nu_T)).$$

Example 9.6. Let $S^2 = \{(x_1, x_2, x_2) : x_1^2 + x_2^2 + x_3^2 = 1\}$ be the 2-sphere, and $a \colon S^2 \to S^2$ the *antipodal* map, $a(x_1, x_2, x_3) = -(x_1, x_2, x_3)$. We calculate the degree of a. It is easy to see as in example 5.8 that if

$$\alpha(x) = x_1 dx_2 \wedge dx_3 - x_2 dx_1 \wedge dx_3 + x_3 dx_1 \wedge dx_2.$$

then α restricts to a nowhere vanishing 2-form ω on S^2 which is in fact the volume form ν_{S^2}. Moreover since a extends to $\mathbb{R}^3$, and $a^*(\alpha) = (-1)^3 \alpha = -\alpha$ then

$$\deg(a) = \frac{\int_{S^2} a^*(\omega)}{\int_{S^2} \omega} = \frac{\int_{S^2} (-\omega)}{\int_{S^2} \omega} = -1$$

We wish to show that the degree of a map f is constant when f varies smoothly – so once again we are forced to make precise the notion of a smoothly varying family.

Definition 9.7. Let S and T be smooth compact k-surfaces, and f_0, f_1 a pair of smooth maps from S to T. A *smooth homotopy* from S to T is a smooth function $F\colon S\times\mathbb{R}\to T$ (where $S\times\mathbb{R}$ is naturally a $(k+1)$-surface in $\mathbb{R}^{n+1}$ if S is a k-surface in $\mathbb{R}^n$) such that $F_{|S\times\{0\}} = f_0$ and $F_{|S\times\{1\}} = f_1$. We write $f_t\colon S\to T$ for the map $F_{|S\times\{t\}}$.

We need a preparatory lemma:

Lemma 9.8. *Let S be a smooth k-surface and let $i_t\colon S\to S\times\mathbb{R}$ denote the inclusions $i_t(x) = (x,t)$. Then there are maps $H_p\colon \Omega^p(S\times\mathbb{R})\to\Omega^{p-1}(S)$ such that*

$$dH_p(\omega) + H_{p+1}d(\omega) = i_1^*(\omega) - i_0^*(\omega).$$

Proof. As for the definition of the integral of a k-form, we can reduce to the case of where S is an open subset U of $\mathbb{R}^k$. If we give $U\times\mathbb{R}$ coordinates $(x,t) = (x_1,x_2,\dots,x_k,t)$, then any $\omega\in\Omega^p(U\times\mathbb{R})$ can be written as

$$\omega = \sum_I f_I(x,t)dx_I + \sum_J g_J(x,t)dt\wedge dx_J,$$

where I runs over subsets $\{i_1 < i_2 < \ldots < i_p\}$ and J runs over subsets $\{j_1 < j_2 < \ldots < j_{p-1}\}$, and f_I and g_J are smooth functions on $U\times\mathbb{R}$. Define

$$H_p(\omega) = \sum_J (\int_0^1 g_J(x,t)dt)dx_J$$

Then we see that

$$\begin{aligned} dH_p(\omega) + H_{p+1}d(\omega) &= \sum_{J,i}\big(\int_0^1 \frac{\partial g_J}{\partial x_i}(x,t)dt\big)dx_i\wedge dx_J \\ &\quad + \sum_I (\int_0^1 \frac{\partial f_I}{\partial t}dt)dx_I - \sum_{J,i}\big(\int_0^1 \frac{\partial g_J}{\partial x_i}(x,t)dt\big)dx_i\wedge dx_J \\ &= \sum_I (\int_0^1 \frac{\partial f_I}{\partial t}dt)dx_I \\ &= \sum_I f_I(x,1)dx_I - \sum_I f_I(x,0)dx_I \\ &= i_1^*(\omega) - i_0^*(\omega). \end{aligned}$$

□

We are now ready to prove the invariance of degree under smooth homotopy.

Proposition 9.9. *Let $F\colon S\times\mathbb{R}\to T$ be a smooth homotopy between compact surfaces. Then*

$$deg(f_0) = deg(f_1).$$

Proof. Let ν_T be the volume form for T. As before, for $t\in\mathbb{R}$ let $i_t\colon S\to S\times I$ be the inclusion given by $i_t(x) = (x,t)$. Then $f_0^*(\nu_T) = i_0^*\circ F^*(\nu_T)$ and $f_1^*(\nu_T) = i_1^*\circ F^*(\nu_T)$, hence if we set $\alpha = F^*(\nu_T)$ we are reduced to showing that

$$\int_S i_0^*(\alpha) = \int_S i_1^*(\alpha).$$

Since S is compact and without boundary, it therefore follows from Stokes' theorem that it is enough to show that $i_1^*(\alpha) - i_0^*(\alpha) = d\beta$ for some $\beta \in \Omega^{k-1}(S)$. To do this we simply apply Lemma 9.8: Set $\beta = H_k(\alpha)$. Then we know that

$$i_1^*(\alpha) - i_0^*(\alpha) = dH_k(\alpha) + H_{k+1}(d\alpha) = d\beta,$$

since $d\alpha = dF^*(\nu_T) = F^*(d\nu_T) = 0$. □

Finally we use this invariance of the degree to show that the sphere S^2 does not have a nowhere vanishing vector field. In fact the same proof shows that S^{2n} has no such vector field.

Theorem 9.10. *Let S^2 the the unit sphere in $\mathbb{R}^3$. Then there is no nowhere vanishing vector field $v : S^2 \to TS^2$*

Proof. Suppose for contradiction that there is such a v. By rescaling if necessary we may assume that $||v(x)|| = 1$ for all $x \in S^2$. Consider the map $F_t \colon S^2 \times \mathbb{R} \to S^2$ given by

$$F_t(x) = \cos(\pi t)x + \sin(\pi t)v(x).$$

It is easy to check that $F_t(x) \in S^2$ for all t and that F_0 is the identity, whereas $F_1 = a$ the antipodal map, $a(x) = -x$. It is easy to check that $\deg(a) = -1$ whereas $\deg(Id) = 1$. Thus we have a contradiction as required. □

Remark 9.11. The above proof easily generalizes to show that there is no nowhere vanishing smooth vector field on an even-dimensional sphere. It is easy to see that, on the other hand there are nowhere vanishing vector fields on odd-dimensional spheres – see if you can find one, starting first with the unit circle.

Remark 9.12. The applications in this section are just a glimpse at the many connections between analysis and topology on surfaces. The connection has been the subject of a huge amount of mathematics in the latter part of the twentieth century continuing through to today. [MT] is a natural continuation of what we have been talking about, while [BT] and [M] are, if it is allowable to say such a thing, canonical.

10. APPENDIX: MULTILINEAR ALGEBRA

We begin by discussing the content (or n-volume) of an n-dimensional version of a parallelogram: Given a set of n vectors, $v_1, v_2, \dots, v_n$, the n-dimensional parallelepiped $P(v_1, v_2, \dots, v_n)$ is the set:

$$\{t_1 v_1 + t_2 v_2 + \dots t_n v_n : t_i \in [0,1], \forall i, 1 \leq i \leq n\}.$$

We want to consider the function on n-tuples of vectors $c\colon (\mathbb{R}^n)^n \to \mathbb{R}$ given by $c(v_1, v_2, \dots, v_n) = \nu(P(v_1, v_2, \dots, v_n))$, the n-volume of $P(v_1, v_2, \dots, v_n)$.

Example 10.1. We consider the smallest cases, when $n = 1$ or 2. For $n = 1$ there is almost nothing to do: any vector $v \in \mathbb{R}$ is of the form λe_1, and $V(v) = |\lambda|$. Now let $n = 2$, and suppose we have two vectors $v = (a, b)$ and $w = (c, d)$. Then the area of the parallelogram that they span is $|ad - bc|$. (You can see this by cutting up the parallelogram they yield into two triangles say). Notice that in each of the cases we have examined, the content function (length or area) was the absolute value of a polynomial function, which gives a "signed volume". It turns out that this function is in fact easier to study. For the $n = 2$, let S be this signed area. Then $S(v, w)$ has the following properties:

- $S(v, v) = 0$ for all $v \in \mathbb{R}^2$;
- $S(e_1, e_2) = 1$ where $e_1 = (1, 0), e_2 = (0, 1)$;
- $S(v_1 + v_2, w) = S(v_1, w) + S(v_2, w)$, for all $v_1, v_2, w \in \mathbb{R}^2$;
- $S(v, w_1 + w_2) = S(v, w_1) + S(v, w_2)$, for all $v.w_1, w_2 \in \mathbb{R}^2$.

Of course all of these can be checked directly from the formula, but they can also be explained as follows: the first property simply asserts that the degenerate parallelogram consisting of the line segment from 0 to v does not have any area. The second is simply a normalization of area (a choice of units, if you like). Finally, the last two properties follow from the fact that the area is the product of the length of one of the vectors, w say, times the (signed) length of the component of the other vector in the direction perpendicular to w. Since taking the signed length of this component is a linear function, we see that S is linear in the vector v.

Similar reasoning suggests that a signed volume in $\mathbb{R}^3$ should have similar properties. In fact it turns out that the generalization of these properties uniquely determine a function on n-tuples of vectors in $\mathbb{R}^n$. We formalize this with the following definition:

Definition 10.2. A function $A\colon (\mathbb{R}^n)^k \to \mathbb{R}$ is an *alternating k-multilinear map* if

- (*alternating*): $A(v_1, v_2, \dots, v_n) = 0$ whenever at least two of the v_i are equal;
- (*multilinear*): A is linear in each factor: given $v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots v_n$, and w_1, w_2 we have

$$\begin{aligned} A(v_1, v_2, \dots v_{i-1}, w_1 + w_2, v_{i+1}, \dots, v_n) =& A(v_1, v_2, \dots, v_{i-1}, w_1, v_{i+1}, \dots, v_n) \\ &+ A(v_1, v_2, \dots, v_{i-1}, w_2, v_{i+1}, \dots, v_n). \end{aligned}$$

Notice that the set $\Lambda^k(\mathbb{R}^n)$ of such functions for a fixed integer k is a vector space. For convenience, we also define $\Lambda^0(\mathbb{R}^n) = \mathbb{R}$. The functions on k-vectors which only satisfy the second property are called *k-multilinear functions*. They also form a vector space, denoted $T^k(\mathbb{R}^n)$. These spaces are finite dimensional, as we will soon see.

Lemma 10.3. *Let* $v_1, v_2, \dots v_k \in \mathbb{R}^n$ *and suppose that* $v_i = \sum_{j=1}^n a^i_j e_j$ $(1 \leq i \leq k)$*. Then if* $A \in T^k(\mathbb{R}^n)$ *we have*

$$A(v_1, v_2, \dots, v_k) = \sum_{j_1, j_2, \dots, j_k} (a^1_{j_1} a^2_{j_2} \dots a^k_{j_k}) A(e_{j_1}, e_{j_2}, \dots, e_{j_k}).$$

Proof. We show this by induction on k: for $k = 0$ there is nothing to prove. Suppose now that the result is known for $k-1$. Then the function $(v_1, v_2, \dots, v_{k-1}) \mapsto A(v_1, v_2, \dots, v_{k-1}, v_k)$ is an element of $\Lambda^{k-1}(\mathbb{R}^n)$, and so we obtain

$$A(v_1, v_2, \dots, v_k) = \sum_{j_1, j_2, \dots, j_k} (a^1_{j_1} a^2_{j_2} \dots a^k_{j_k}) A(e_{j_1}, e_{j_2}, \dots, e_{j_{k-1}} v_k).$$

But now, since A is linear in v_k we have

$$A(e_{j_1}, e_{j_2}, \dots, e_{j_{k-1}}, v_k) = \sum_{j_k=1}^n a^k_{j_k} A(e_{j_1}, e_{j_2}, \dots, e_{j_{k-1}}, e_{j_k}).$$

Substituting this into the above expression gives the required result. ☐

Using the alternating property, we see that each of the terms in the formula given by the lemma vanishes for $A \in \Lambda^k(\mathbb{R}^n)$ unless all the e_{j_k} are distinct. But in fact more is true:

Lemma 10.4. *Let* $v_1, v_2, \dots, v_k$ *vectors in* $\mathbb{R}^n$*, and* $A \in \Lambda^k(\mathbb{R}^n)$ *we have*

$$\begin{aligned} &A(v_1, v_2, \dots, v_{i-1}, v_i, v_{i+1} \dots, v_{j-1}, v_j, v_{j+1}, \dots v_k) = \\ &-A(v_1, v_2, \dots, v_{i-1}, v_j, v_{i+1} \dots, v_{j-1}, v_i, v_{j+1}, \dots v_k), \end{aligned}$$

for every $1 \leq i < j \leq k$.

Proof. Given vectors $v_1, v_2, \dots, v_k$ consider a pair of indices $i < j$. Then

$$0 = A(v_1, v_2, \dots, v_{i-1}, v_i + v_j, v_{i+1} \dots, v_{j-1}, v_i + v_j, v_{j+1}, \dots v_k)$$

But then using the fact that A is linear in each of the ith and jth entries we find that

$$\begin{aligned} 0 =& A(v_1, v_2, \dots, v_{i-1}, v_i, v_{i+1} \dots, v_{j-1}, v_i, v_{j+1}, \dots v_k) \\ &+ A(v_1, v_2, \dots, v_{i-1}, v_i, v_{i+1} \dots, v_{j-1}, v_j, v_{j+1}, \dots v_k) \\ &+ A(v_1, v_2, \dots, v_{i-1}, v_j, v_{i+1} \dots, v_{j-1}, v_i, v_{j+1}, \dots v_k) \\ &+ A(v_1, v_2, \dots, v_{i-1}, v_j, v_{i+1} \dots, v_{j-1}, v_j, v_{j+1}, \dots v_k) \end{aligned}$$

Using the alternating property for the right-hand side, we see that the first and last of terms are zero, and so the lemma follows. ☐

A multilinear map satisfying the conclusion of the lemma is said to be *skew-symmetric*. It is easy to check that this property implies the alternating property (because the only number c with $c = -c$ is zero), so the space of alternating multilinear maps is the same as the space of skew-symmetric multilinear maps. Using Lemma 10.4, the formula in the preceeding lemma can be made more precise: each of the terms $A(e_{j_1}, e_{j_2}, \dots, e_{j_k})$ is zero unless all the indices $j_1, j_2, \dots, j_k$ are distinct, and then given a set of distinct indices $\{j_1, j_2, \dots, j_k\}$ if we swap any pair of them, the value of A simply changes sign. In particular, if $k = n$ then the indices $\{j_1, j_2, \dots, j_n\}$ must just be a reordering of the integers $\{1, 2, \dots, n\}$, and if $k > n$ then there are *no* nonzero alternating k-multilinear functions. For $n = 2$ we have shown that the signed area function is a skew-symmetric bilinear function ("bilinear" = 2-multilinear), but we do not as yet know if there are *any* such functions (apart from the zero function) for $n > 2$.

Definition 10.5. A *permutation* of $\{1, 2, \dots, n\}$ is a bijection

$$\sigma \colon \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}.$$

We denote the set of all permutations of the set $\{1, 2, \dots, k\}$ by S_k, the *symmetric group*. Note that the composition of two permutations is clearly a permutation. A permutation which interchanges two elements of the set $\{1, 2, \dots, k\}$ and leaves the remaining $k - 2$ unchanged is called a *transposition*.

Suppose that we have a subset $\{j_1, j_2, \dots, j_r\}$ of $\{1, 2, \dots, k\}$. We may define a permutation σ by setting

$$\sigma(i) = \begin{cases} j_{s+1}, & \text{if } i = j_s, s < r; \\ j_1 & i = j_r; \\ i & 0 \end{cases}$$

We say that σ is a *cycle* of length r, and write $\sigma = (j_1 j_2 \dots j_r)$. Thus a transposition is cycle of length 2. We say two cycles $\sigma_1 = (i_1 i_2 \dots i_t)$ and $\sigma_2 = (j_1 j_2, \dots j_r)$ are *disjoint* if the sets $\{i_1, i_2, \dots, i_t\}$ and $\{j_1, j_2, \dots, j_s\}$ are disjoint. Notice that if σ_1 and σ_2 are disjoint cycles, then they commute, *i.e.* $\sigma_1\sigma_2 = \sigma_2\sigma_1$. The next lemma shows that there is a kind of unique factorization of a permutation as a product of cycles.

Lemma 10.6. *Let $\sigma \in S_k$ be a permutation. Then there are disjoint cycles $\gamma_1, \gamma_2, \gamma_l$ such that*

$$\sigma = \gamma_1\gamma_2 \dots \gamma_l$$

Proof. To find the cycles we introduce a relation on $\{1, 2, \dots, k\}$ as follows: say that $i \sim j$ if for some integer $r \in \mathbb{Z}$ we have $\sigma^r(i) = j$ (where σ^0 is the identity permutation by definition, and if $n > 0$ we set $\sigma^{-n} = (\sigma^{-1})^n$). Now $i = \sigma^0(i)$, so that $\sim$ is reflexive. Next, if $\sigma^r(i) = j$, then $\sigma^{-r}(j) = i$ so that $\sim$ is symmetric. Finally, if $i \sim j$ and $j \sim p$, then there are integers r, s such that $\sigma^r(i) = j$ and $\sigma^s(j) = p$. But then $\sigma^{r+s}(i) = \sigma^s(\sigma^r(i)) = \sigma^s(j) = p$, and so $i \sim p$, that is, $\sim$ is transitive. Thus $\sim$ is an equivalence relation.

It follows that $\{1, 2, \dots, k\}$ is the disjoint union of the equivalence classes of $\sim$. Now let $\mathcal{O}$ be an equivalence class. Then if $j \in \mathcal{O}$, it follows $\sigma^s(j) \in \mathcal{O}$ for all $s \in \mathbb{Z}$. Since $\{1, 2, \dots, k\}$ is finite, $\mathcal{O}$ is also finite, and so at some point we must have $\sigma^s(j) \in \{j, \sigma(j), \sigma^2(j), \dots, \sigma^{s-1}(j)\}$. Suppose that r is the first s for which this is true. Then we have $\sigma^r(j) = \sigma^p(j)$ for some $p < r$. But then $\sigma^{r-p}(j) = j$, and so if r is minimal, we must have $p = 0$, and hence $\sigma^r(j) = j$. It follows that $\mathcal{O} = \{j, \sigma(j), \dots, \sigma^r(j)\}$. Let $\{\mathcal{O}_i\}_{1 \leq i \leq l}$ be the equivalence classes of $\sim$ and $j_t \in \mathcal{O}_t$ be an element of $\mathcal{O}_t$, $(1 \leq t \leq l)$, so that $\mathcal{O}_t = \{j_t, \sigma(j_t), \dots, \sigma^{r_t - 1}(j)\}$ where $r_t = |\mathcal{O}_t|$. Then if γ_i is the cycle $(j_t, \sigma(j_t), \dots, \sigma^{r_t-1}(j_t))$ we see immediately that $\sigma = \gamma_1\gamma_2 \dots \gamma_l$ as required. □

Example 10.7. The lemma is easiest to understand in an example: Suppose that $\sigma \in S_6$ is the permutation which sends $1 \mapsto 4$, $2 \mapsto 1$, $3 \mapsto 5$, $4 \mapsto 2$, $5 \mapsto 6$ and $6 \mapsto 3$. Then if we start with 1 we find that $1 \mapsto 4 \mapsto 2 \mapsto 1$, so that (142) is one of the cycles for σ. Picking 3 as the smallest term not in the cycle we already have, we find that $3 \mapsto 5 \mapsto 6 \mapsto 3$, and so σ is the product $(142)(356) = (356)(142)$.

The number of cycles in the expression for $\sigma \in S_k$ as a product of disjoint cycles (more precisely, the number of equivalence classes for $\sim$ in the previous lemma) is called the *cycle length* of σ, written $c(\sigma)$. We define a map $\varepsilon \colon S_k \to \{\pm 1\}$ by setting

$$\varepsilon(\sigma) = (-1)^{k-c(\sigma)}.$$

We call $\varepsilon(\sigma)$ the *sign* of σ. The existence of non-zero skew-symmetric multilinear functions follows from the fact that the function ε we have just defined is compatible with composition of permutations, in that if $\sigma_1, \sigma_2 \in S_k$ then

$$\varepsilon(\sigma_1 \circ \sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2). \tag{10.1}$$

The following lemmas gives a proof of this basic fact.

Lemma 10.8. *Let $\sigma \in S_k$ be a permutation. Then there are transpostions $\tau_1, \tau_2, \dots, \tau_m$ such that*

$$\sigma = \tau_1 \tau_2 \dots \tau_m.$$

(where this product is taken to be the identity map from $\{1, 2, \dots, n\}$ to itself if $m = 0$).

Proof. Since any permutation can written as a product of cycles, it is enough to show that any cycle can be written as a product of transpositions. Let $(j_1, j_2, \dots, j_s)$ be a cycle. It is easy to check that

$$(j_1, j_2, \dots, j_s) = (j_1 j_s)(j_1 j_{s-1}) \dots (j_1, j_2)$$

□

Example 10.9. Let $\sigma = (1234) \in S_4$, the we have

$$(1234) = (14)(13)(12)$$

The general case works in exactly the same way.

Notice that a representation of a permutation as a product of transpositions, in contrast to the representation as a product of disjoint cycles, is far from unique (as a trivial example, if τ is a transposition, then $\tau^2 = 1$, so given any product of transpositions which equals σ, we can add on 2 copies of τ at the end to get a new product of transpositions equaling σ).

Lemma 10.10. *Let $\sigma \in S_k$ and let τ be a transposition. Then $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$.*

Proof. It is enough to show that $c(\sigma\tau) = c(\sigma) \pm 1$. Let $\tau = (ij)$ be the presentation of τ as a product of cycles, and let $\sigma = \gamma_1\gamma_2, \dots \gamma_l$ be the presentation of σ as a product of cycles. . To see this, observe that there are two cases. Either the elements (ij) are in the same cycle γ_r, or they are in two different cycles γ_r, γ_s. In the first case one checks that $\gamma_r\tau$ splits into two disjoint cycles: suppose that $\gamma_r = (i, n_1, n_2, \dots, n_p, j, m_1, m_2, \dots, m_q)$, then $\gamma_r\tau = (i, n_1, n_2, \dots, n_p)(j, m_1, m_2, \dots, m_q)$. In the second case we have $\gamma_r = (i, n_1, n_2, \dots, n_p)$, $\gamma_s = (j, m_1, m_2, \dots, m_q)$ say, and so $\gamma_r\gamma_s\tau = (i, n_1, n_2, \dots, n_p, j, m_1, m_2, \dots, m_q)$, is a single cycle. The result follows. □

Combining the two lemmas, we see that ε is compatible with composition of permutations as claimed.

Lemma 10.11. *Let $A \in \Lambda^k(\mathbb{R}^n)$, and let $v_1, v_2, \dots, v_k \in \mathbb{R}^n$. Then for any $\sigma \in S_k$ we have*

$$A(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(k)}) = \varepsilon(\sigma)A(v_1, v_2, \dots, v_k).$$

Proof. For σ a transposition, this follows from skew-symmetry. For an arbitrary permutation σ, write is as a product of transpositions, and use induction. $\square$

We now consider again the expression we obtained in Lemma 10.3 for $A \in \Lambda^k(\mathbb{R}^n)$, an alternating multilinear function: for vectors $v_1, v_2, \dots, v_k$ with $v_i = \sum_{j=1}^n a^i_j e_j$, ($a^i_j \in \mathbb{R}$) we have

$$A(v_1, v_2, \dots, v_k) = \sum_{j_1, j_2, \dots, j_k} a^1_{j_1} a^2_{j_2} \dots a^k_{j_k} A(e_{j_1}, e_{j_2}, \dots, e_{j_k}).$$

Each of the terms $A(e_{j_1}, e_{j_2}, \dots e_{j_k})$ vanishes unless all the $\{j_1, j_2, \dots, j_k\}$ are distinct, and moreover, in this case we may find $\sigma \in S_k$ such that the sequence $(\{j_{\sigma(1)}, j_{\sigma(2)}, \dots, j_{\sigma(k)})$ is increasing. It follows using Lemma 10.11 that if we group together the nonzero terms in the above summation according to the k-element sequences $J = (j_1 < j_2 < \dots < j_k)$, for $j_s \in \{1, 2, \dots, n\}$, $(1 \leq s \leq k)$ we have

$$A(v_1, v_2, \dots, v_k) = \sum_{J \subset \{1,2,\dots,n\}} A(e_{j_1}, e_{j_2}, \dots, e_{j_k}) \big(\sum_{\sigma \in S_k} \varepsilon(\sigma) a^1_{j_{\sigma(1)}} a^2_{j_{\sigma(2)}} \dots a^k_{j_{\sigma(k)}} \big)$$

Now let, for each J as above, let $D_J \colon (\mathbb{R}^n)^k \to \mathbb{R}$ be the function given by

$$D_J(v_1, v_2, \dots, v_k) = \sum_{\sigma \in S_k} \varepsilon(\sigma) a^1_{j_{\sigma(1)}} a^2_{j_{\sigma(2)}} \dots a^k_{j_{\sigma(k)}},$$

so the previous equation shows that for an arbitrary $A \in \Lambda^k(\mathbb{R}^n)$,

$$A = \sum_{J \subset \{1,2,\dots,n\}} A(e_{j_1}, e_{j_2}, \dots, e_{j_k}) D_J. \tag{10.2}$$

Lemma 10.12. *For any J a k-element subset of $\{1, 2, \dots, n\}$ (where $k < n$), the function D_J is a nonzero element of $\Lambda^k(\mathbb{R}^n)$. Moreover the functions D_J as J runs over the sequences $(j_1 < j_2 < \dots < j_k)$, for $j_s \in \{1, 2, \dots, n\}, (1 \leq s \leq k)$ form a basis of $\Lambda^k(\mathbb{R}^n)$.*

Proof. Let $J = (j_1 < j_2 < \dots < j_k)$. It is clear that D_J is multilinear, so we need only check that it is alternating and nonzero. To see that it is alternating, let $\tau = (ij)$, and $v_i = \sum_{j=1}^n a^i_j e_j \in \mathbb{R}^n$ $(1 \leq i \leq k)$, then we have

$$\begin{aligned} D_J(v_{\tau(1)}, v_{\tau(2)}, \dots, v_{\tau(k)}) &= \sum_{\sigma \in S_k} \varepsilon(\sigma) a^1_{j_{\sigma(\tau(1))}} \dots a^k_{j_{\sigma(\tau(k))}} \\ &= \sum_{\rho \in S_k} \varepsilon(\rho\tau) a^1_{j_{\rho(1)}} \dots a^k_{j_{\rho(k)}} \\ &= \varepsilon(\tau) \sum_{\rho \in S_k} \varepsilon(\rho) a^1_{j_{\rho(1)}} \dots a^k_{j_{\rho(k)}} \\ &= -D_J(v_1, v_2, \dots, v_k). \end{aligned}$$

where in the second equality we set $\rho = \sigma\tau$. To see that the functions D_J form a basis of $\Lambda^k(\mathbb{R}^n)$, notice that if $(v_1, v_2, \dots, v_k) = (e_{i_1}, e_{i_2}, \dots, e_{i_k})$ where $I = (i_1, i_2, \dots, i_k)$ is an increasing sequence of elements of $\{1, 2, \dots, k\}$. then we have

$$D_J(e_{i_1}, \dots, e_{i_k}) = \begin{cases} 1 & \text{if } J = I \\ 0 & \text{if } I \neq J \end{cases}$$

This shows that the D_J are linearly independent, but Equation (10.2) already shows that they span $\Lambda^k(\mathbb{R}^n)$, so we are done. $\square$

Suppose now that $k = n$, then we have shown that $\Lambda^n(\mathbb{R}^n)$ is one dimensional with basis vector $D = D_{(1,2,\dots,n)}$ (as $(1, 2, \dots, n)$ is the only n-term increasing sequence in $\{1, 2, \dots, n\}$). This allows us to define the determinant of a linear map.

If $\alpha\colon \mathbb{R}^n \to \mathbb{R}^n$ is linear, then the function $\alpha^*(D)$ given by

$$\alpha^*(D)(v_1, v_2, \dots, v_n) = D(\alpha(v_1), \alpha(v_2), \dots, \alpha(v_n))$$

is clearly an alternating multilinear map, *i.e.* $\alpha^*(D) \in \Lambda^n(\mathbb{R}^n)$. But since this vector space is one-dimensional, with basis $\{D\}$, we see that

$$\alpha^*(D) = \lambda D,$$

and using Equation (10.2) above we see that $\lambda = D(\alpha(e_1), \alpha(e_2), \dots, \alpha(e_n))$. Thus we define

$$\det\colon \text{End}(\mathbb{R}^n) \to \mathbb{R}$$

to be the map which sends $\alpha \mapsto D(\alpha(e_1), \alpha(e_2), \dots, \alpha(e_n))$, the *determinant* of α.

Lemma 10.13. *Let $\alpha \in End(\mathbb{R}^n)$, and let $A = (a_{ij})_{1\leq i,j\leq n}$ be the matrix of α, then*

$$\det(\alpha) = \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}.$$

Moreover if $\alpha, \beta \in End(\mathbb{R}^n)$ then

$$\det(\alpha \circ \beta) = \det(\alpha). \det(\beta).$$

Proof. The first part is immediate from the columns of A are the vectors $\alpha(e_i)$, as i runs from 1 to n. The second part follows from the fact that $(\alpha \circ \beta)^*(D) = \beta^*(\alpha^*(D))$ so that

$$\det(\alpha \circ \beta) = \det(\alpha). \det(\beta).$$

$\square$

Corollary 10.14. *A linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^n$ is invertible if and only if $\det(\alpha) \neq 0$.*

Proof. Suppose that $\alpha \circ \beta = Id$. Then since it is easy to see that $\det(Id) = 1$, the previous lemma shows that

$$\det(\alpha)\det(\beta) = 1,$$

and hence $\det(\alpha) \neq 0$. For the converse, suppose that α is not invertible. Then the vectors $\alpha(e_1), \alpha(e_2), \dots, \alpha(e_n)$ are linearly independent. Thus we may find an i such that

$$\alpha(e_i) = \sum_{j\neq i} \lambda_j \alpha(e_j).$$

But then

$$\begin{aligned}
\det(\alpha) &= D(\alpha(e_1), \alpha(e_2), \dots, \alpha(e_n)) \\
&= D(\alpha(e_1), \dots, \sum_{j\neq i} \lambda_j \alpha(e_j), \dots, \alpha(e_n)) \\
&= \sum_{j\neq i} \lambda_j D(\alpha(e_1), \dots \alpha(e_j), \dots \alpha(e_n)) \\
&= 0.
\end{aligned}$$

using the fact that D is alternating and multilinear (in the $j-th$ term in the last sum, $\alpha(e_j)$ occurs twice, and so D vanishes). $\square$

Definition 10.15. If we have a collection of vector spaces V_k where k runs over the nonnegative integers $\{0,1,2,\ldots\}$ (which we will denote $\mathbb{N}$), then we define

$$\bigoplus_{i\geq 0} V_i = \{(v_i)_{i\in\mathbb{N}} \colon v_i \in V_i \text{ for all } i\in\mathbb{N}, v_i = 0 \text{ for all but finitely many } i\in\mathbb{N}\},$$

where we make the left-hand side into a vector space by adding and scaling componentwise.

Definition 10.16. An algebra is a vector space A together with a map $A\times A\to A$ where $(v,w)\mapsto v*w$ such that

$$(\lambda_1 v_1 + \lambda_2 v_2) * w = \lambda(v_1 * w) + \lambda_2(v_2 * w)$$

($v_1, v_2, w \in A$, $\lambda_1, \lambda_2 \in \mathbb{R}$), and

$$v\cdot(\lambda_1 w_1 + \lambda_2 w_2) * w = \lambda(v * w_1) + \lambda_2(v * w_2)$$

($v, w_1, w_2 \in A, \lambda_1, \lambda_2 \in \mathbb{R}$. We say A is *associative* if $u*(v*w) = (u*w)*w$. It has a *unit* if there is an $e\in A$ such that $e*v = v*e = v$ for all $v\in A$. Every algebra we will come across will be associative and have a unit. If A is a graded vector space, say $A = \bigoplus_{i\in\mathbb{N}} A_i$ then we say that A is a *graded algebra* if the multiplication is compatible with the grading in the sense that $*$ restricts to a map $*\colon A_i\times A_j$ to A_{i+j}.

Remark 10.17. The symbol "$*$" used to denote the multiplication is often changed to "$\cdot$" or something else depending on the particular algebra we are talking about (*e.g.* $\wedge$ in the case of the soon-to-be-defined exterior algebra). In fact, as with multiplication of real numbers, sometimes the symbol is suppressed altogether.

Example 10.18. Let P be the set of polynomials in two variables x, y. Then it is easy to see that P is an algebra under the standard multiplication of polynomials. Moreover if P_k is the linear span of the functions $\{x^k, x^{k-1}y, \ldots, xy^{k-1}, y^k\}$ we see that $P = \bigoplus P_i$ and that P is a graded algebra. Notice that P is commutative, that is, if p, q are polynomials, then $p*q = q*p$. Most of the algebras we will be dealing with will *not* be commutative. One of the simplest examples of a noncommutative algebra is the algebra $M_n(\mathbb{R})$ of n-by-n matrices, though it is not a graded algebra.

We want to make $\bigoplus_{k\in\mathbb{N}} T^k(\mathbb{R}^n)$ and $\bigoplus_{k\geq 0}\Lambda^k(\mathbb{R}^n)$ into algebras. It is easy to make $T^*(\mathbb{R}^n)$ into an algebra, in fact a graded algebra: Given $M\in T^k(\mathbb{R}^n)$ and $N\in T^l(\mathbb{R}^n)$ we define $M.N\in T^{k+l}(\mathbb{R}^n)$ by

$$M.N(v^1, v^2, \ldots v^{k+l}) = M(v^1, v^2, \ldots, v^k)\cdot N(v^{k+1}, \ldots, v^{k+l}).$$

Extending this linearly we get a multiplication on all of $T^*(\mathbb{R}^n)$. Notice that this multiplication is *not* commutative. (We will sometimes suppress the "$\cdot$" to make some expressions more compact).

On the other hand, it is not immediately clear how to make $\Lambda^*(\mathbb{R}^n)$ into an algebra. Although we have defined a product on $T^*(\mathbb{R}^n)$, it is *not* the case that if $\alpha\in\Lambda^k(\mathbb{R}^n)$ and $\beta\in\Lambda^l(\mathbb{R}^n)$ then their product $\alpha\cdot\beta$ lies in $\Lambda^{k+l}(\mathbb{R}^n)$. To fix this we need a way to take the "alternating part" of this product:

Lemma 10.11 shows that a multilinear function is alternating exactly when

$$M(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) = \varepsilon(\sigma) M(v_1, v_2, \ldots, v_k). \tag{10.3}$$

for every $\sigma \in S_k$. If for an arbitrary $N \in T^k(\mathbb{R}^n)$ we write ${}^{\sigma}N$ for the multilinear function given by

$$ {}^{\sigma}N(v_1, v_2, \dots, v_k) = N(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(k)}), $$

then the condition that M is alternating can be more concisely written as ${}^{\sigma}M = \varepsilon(\sigma)M$. Notice that the map $M \mapsto^{\sigma} M$ is linear for any $\sigma \in S_k$. We use this characterization of alternating functions to produce a projection map from the space $T^k(\mathbb{R}^n)$ to the subspace $\Lambda^k(\mathbb{R}^n)$ as follows:

Lemma 10.19. *Let* $A\colon T^k(\mathbb{R}^n) \to T^k(\mathbb{R}^n)$ *be the linear map given by*

$$ \begin{aligned} A(M)(v_1, v_2, \dots v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \varepsilon(\sigma) M(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(k)}) \\ &= \frac{1}{k!} \big(\sum_{\sigma \in S_k} \varepsilon(\sigma)^{\sigma} M \big)(v_1, v_2, \dots, v_k). \end{aligned} $$

Then $A(M) \in \Lambda^k(\mathbb{R}^n)$ *for all* $M \in T^k(\mathbb{R}^n)$*, and moreover* $A({}^{\tau}M) = \varepsilon(\tau)A(M)$ *for any* $\tau \in S_k$*, and hence* $A(M) = M$ *if* $M \in \Lambda^k(\mathbb{R}^n)$*, hence* $A^2 = A$ *and the image of* A *is exactly* $\Lambda^k(\mathbb{R}^n)$*.*

Proof. It is clear that A is a linear map. Suppose that $\tau \in S_k$, and $M \in T^k(\mathbb{R}^n)$. Then we have

$$ \begin{aligned} {}^{\tau}(A(M))(v_1, v_2, \dots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \varepsilon(\sigma)^{\tau} M(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(k)}) \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} \varepsilon(\sigma) M(v_{\tau\circ\sigma(1)}, v_{\tau\circ\sigma(2)}, \dots, v_{\tau\circ\sigma(k)}) \\ &= \frac{1}{k!} \sum_{\sigma'} \varepsilon(\tau^{-1} \circ \sigma') M(v_{\sigma'(1)}, v_{\sigma'(2)}, \dots, v_{\sigma'(k)}), \end{aligned} $$

where in the second line we have set $\sigma' = \tau \circ \sigma$. But as σ runs over S_k so does σ' and since $\varepsilon(\sigma' \circ \tau^{-1}) = \varepsilon(\sigma')\varepsilon(\tau^{-1}) = \varepsilon(\sigma')\varepsilon(\tau)$ we see immediately that this is just $\varepsilon(\tau)A(M)$. Thus $A(M) \in \Lambda^k(\mathbb{R}^n)$ as claimed. The proof of the "moreover" part is almost identical:

$$ \begin{aligned} A({}^{\tau}M)(v_1, v_2, \dots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \varepsilon(\sigma) M(v_{\sigma\circ\tau(1)}, v_{\sigma\circ\tau 2)}, \dots, v_{\sigma\circ\tau(k)}) \\ &= \frac{1}{k!} \sum_{\sigma'} \varepsilon(\sigma' \circ \tau^{-1}) M(v_{\sigma'(1)}, v_{\sigma'(2)}, \dots, v_{\sigma'(k)}) \\ &= \varepsilon(\tau^{-1}) \frac{1}{k!} \sum_{\sigma'} \varepsilon(\sigma') M(v_{\sigma'(1)}, v_{\sigma'(2)}, \dots, v_{\sigma'(k)}) \\ &= \varepsilon(\tau) A(M) \end{aligned} $$

where here $\sigma' = \sigma \circ \tau$, and we have again used the fact that $\varepsilon(\tau) = \varepsilon(\tau^{-1})$. This immediately implies the remaining statements in the lemma, since $|S_k| = k!$. ☐

Definition 10.20. If $\alpha \in \Lambda^k(\mathbb{R}^n)$ and $\beta \in \Lambda^k(\mathbb{R}^n)$ then set $(\alpha \wedge \beta)$ to be

$$ \frac{(k+l)!}{k!l!} A(\alpha \cdot \beta). $$

It is now of course clear that $\alpha \wedge \beta$ is alternating. Moreover, this product is *associative*, that is, if $\alpha, \beta, \gamma \in \Lambda^*(\mathbb{R}^n)$ then

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma),$$

indeed if α, β, and γ have degree k, l, m respectively, then

$$(\alpha \wedge \beta) \wedge \gamma = \alpha(\beta \wedge \gamma) = \frac{(k+l+m)!}{k!l!m!} A(\alpha \cdot \beta \cdot \gamma),$$

and so we have made $\Lambda^*(\mathbb{R}^n)$ into an algebra, which we call the *exterior algebra*. Although, like $T^*(\mathbb{R}^n)$, the alternating algebra $\Lambda^*(\mathbb{R}^n)$ is not commutative, it is *anticommutative*, that is, if $\alpha \in \Lambda^k(\mathbb{R}^n)$ and $\beta \in \Lambda^l(\mathbb{R}^n)$ then

$$\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha.$$

Given a linear map between $L\colon \mathbb{R}^k \to \mathbb{R}^n$, we obtain a map $L^*\colon T^*(\mathbb{R}^n) \to T^*(\mathbb{R}^k)$ known as the *pullback* map associated to L as follows: If $\alpha \in \Lambda^p(\mathbb{R}^n)$ then we set

$$L^*(\alpha)(v_1, v_2, \dots, v_p) = \alpha(L(v_1), L(v_2), \dots, L(v_p)),$$

where $v_1, v_2, \dots, v_p \in \mathbb{R}^k$. It is easy to see that L^* is a linear map $T^p(\mathbb{R}^n) \to T^p(\mathbb{R}^k)$, and moreover it is compatible with the product on the two algebras $T^*(\mathbb{R}^n)$ and $T^*(\mathbb{R}^k)$, in other words, given $M, N \in T^*(\mathbb{R}^n)$

$$L^*(M \cdot N) = L^*(M) \cdot L^*(N).$$

Now it is clear from the definitions that the maps A and L^* commute, and so L^* restricts to give a linear map $L^*\colon \Lambda^p(\mathbb{R}^n) \to \Lambda^p(\mathbb{R}^k)$, which moreover is compatible with the $\wedge$ product, *i.e.* if $\alpha, \beta \in \Lambda^*(\mathbb{R}^k)$ then

$$L^*(\alpha \wedge \beta) = L^*(\alpha) \wedge L^*(\beta).$$

Example 10.21. Let $\alpha_0 = dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n \in \Lambda^n(\mathbb{R}^n)$, then if $v^1, v_2, \dots, v^n \in \mathbb{R}^n$

$$\alpha_0(v^1, v^2, \dots, v^n) = \det(V)$$

where V is the n-by-n matrix whose columns are the vectors $v_1, v_2, \dots, v_n$. Indeed if we use the definition of $\wedge$ we obtain the formula

$$\det(V) = \sum_{\sigma \in S_n} \varepsilon(\sigma) v^1_{\sigma(1)} v^2_{\sigma(2)} \dots v^n_{\sigma(n)}.$$

Now if $L\colon \mathbb{R}^n \to \mathbb{R}^n$ is a linear map, then $L^*(\alpha_0)$ is also an n-form. But $\Lambda^n(\mathbb{R}^n)$ is one dimensional, and so there is a constant $c(L)$ such that $L^*(\alpha_0) = c(L)\alpha_0$. By applying $L^*(\alpha_0)$ to the n-tuple $(e^1, e^2, \dots, e^n)$ we see that $c(L) = \det(L)$, and so since $\Lambda^n(\mathbb{R}^n)$ is one dimensional we see that

$$L^*(\alpha) = \det(L)\alpha, \qquad \alpha \in \Lambda^n(\mathbb{R}^n).$$

Remark 10.22. We have given this whole discussion in the context of $\mathbb{R}^n$, where we could instead have talked about an abstract n-dimensional vector space V. In order to deal with this slightly more general case the only thing that changes is that one makes an arbitrary choice of a basis $w_1, w_2, \dots, w_n$, and then in place of the linear functions dx_i we use the *dual basis* of $\Lambda^1(V)$ consisting of the functions $\{dw_i\}_{1 \leq i \leq n}$ where

$$dw_i(\sum_{j=1}^n a_j w_j) = a_j.$$

Remark 10.23. Given a linear subspace $V \subset \mathbb{R}^n$, multilinear functions on $\mathbb{R}^n$ clearly restrict to give multilinear functions on V. In fact *every* multilinear function on V arises this way: if we pick a basis $\{v_1, v_2, \dots, v_k\}$ of V we can extend to a basis $\{v_1, v_2, \dots, v_k, v_{k+1}, \dots, v_n\}$ of $\mathbb{R}^n$. Let W be the linear span of the vectors $\{v_{k+1}, v_{k+2}, \dots, v_n\}$. Since any vector w in $\mathbb{R}^n$ can be written uniquely as $\sum_{i=1}^n \lambda_i v_i$, setting $w_1 = \sum_{i=1}^k \lambda v_i$ and $w_2 = \sum_{i=k+1}^n \lambda_i v_i$ we see that w can be written uniquely as $w_1 + w_2$ with $w_1 \in V$ and $w_2 \in W$. It is easy to check that the map $\pi\colon \mathbb{R}^n \to V$ such that $\pi(w) = w_1$ is then linear, and if $w \in V$ then $\pi(w) = w$. The map π is called a projection map. Now given any degree p multilinear function M on V and any linear projection $\pi\colon \mathbb{R}^n \to V$, it follows that $\pi^*(M)$ is a multilinear function on $\mathbb{R}^n$ which extends M.

Remark 10.24. There are many linear projections from $\mathbb{R}^n$ to V, as the choices above make clear, but if we are willing to use the inner product on $\mathbb{R}^n$ there is a natural projection known as the *orthogonal projection*. This takes $w \in \mathbb{R}^n$ to $w_1 \in V$ such that $w - w_1$ is perpendicular to every element of W. Explicitly, if $v_1, v_2, \dots v_k$ is an orthonormal basis for V (that is, $(v_i, v_j) = 1$ if $i = j$ and is 0 otherwise) then

$$w_1 = (w, v_1)v_1 + (w, v_2)v_2 + \dots (w, v_k)v_k,$$

where $(\cdot, \cdot)$ denote the inner product.

11. APPENDIX: PARTITIONS OF UNITY.

We show here the proof of existence of partitions of unity for a general surface. We begin with a lemma on the topology of our surfaces. A *basis* for the topology of a space X, is a collection of open sets $\{U_\alpha\}_{\alpha \in I}$ such that for any open set $U \subset X$ there is an $\alpha \in I$ with $U_\alpha \subset U$.

Lemma 11.1. *A subset of $\mathbb{R}^n$ has a countable basis for its topology. In particular any surface has a countable collection $\{U_i\}_{i \in \mathbb{N}}$ of open sets which form a basis for its topology. Moreover in the case of a surface S these sets can be chosen such that their closures $\{\bar{U}_j\}$ in S are compact.*

Proof. Observe that the collection $\mathcal{B}$ of open balls in $\mathbb{R}^n$ which are centered at point with rational coordinates and which have rational radii form a countable set. They form a *basis* of the set of open subsets of $\mathbb{R}^n$ in the sense that for any open set U there is a ball $B \in \mathcal{B}$ such that $B \subset U$. Intersecting these with our surface S we get a countable collection $\{U_j\}_{j \in \mathbb{N}}$ of open subsets of S which are a basis for the open sets in S.

For each point $x \in S$ we may take a chart $\varphi\colon I^k \to S$ such that x lies in the image of φ. Take an open ball in B in I^k such that $\bar{B} \subset I^k$ and $x \in \varphi(B)$. Then the image $U_x = \varphi(B)$ is an neighborhood of x in S with compact closure. By the previous paragraph, S has a countable basis for its open sets $\{U_j\}_{j \in \mathbb{N}}$ and so we can find for each x an open set $U_j \subset U_x$, and hence $\bar{U}_j \subset \bar{U}$ has compact closure, so we can assume that the $\{U_j\}_{j \in \mathbb{N}}$ have compact closure. $\square$

Recall the statement of the theorem:

Theorem 11.2. *Let S be a smooth $\mathcal{C}^\infty$ k-surface in $\mathbb{R}^n$, and let $\{V_\alpha\}_{\alpha \in A}$ be a cover of S. Then there is a locally finite refinement $\{W_k\}_{k \in \mathbb{N}}$ of the cover $\{V_\alpha\}$, and a partition of unity $\{\rho_k\}_{k \in \mathbb{N}}$ subordinate to $\{W_k\}_{k \in \mathbb{N}}$.*

Proof. Pick a basis $\{U_j\}_{j\in\mathbb{N}}$ for the topology of S as in the previous lemma, such that $\bar{U}_j$ is compact for each $j \in \mathbb{N}$. We use these sets to *exhaust* the surface S. Set $G_1 = U_1$. Then notice that since

$$\bar{G}_1 \subset S = \bigcup_{j\geq 1} U_j,$$

and $\bar{G}_1$ is compact, there is a $k \geq 0$ such that $\bar{G}_1 \subset \bigcup_{j=1}^{k} U_j = G_2$. Continuing in this way we find open sets G_k such that $\bar{G}_j$ is compact, and $\bar{G}_j \subset G_{j+1}$

By our construction, we have

$$\bar{G}_j \backslash G_{j-1} \subset G_{j+1} \backslash \bar{G}_{j-2}$$

where the set on the left is compact, and the one on the right is compact. Take the cover $\{V_\alpha\}$. The sets $V_\alpha \cap (G_{j+1}\backslash \bar{G}_{j-2})$ cover $\bar{G}_j \backslash G_{j-1}$. Since this last set is compact, we may find a finite subcover, and continuing with $j+1$ *etc.* we obtain a countable refinement $\{W_j\}_{j\in\mathbb{N}}$ of the cover $\{V_\alpha\}$ which is locally finite because the sets $G_{j+1}\backslash \bar{G}_{j-2}$ and $G_{j+4}\backslash \bar{G}_{j+1}$ are disjoint.

For each $x \in S$ let j be the largest integer such that $x \in S\backslash G_j$. Take a chart φ_x whose image lies within some $W_k \subset G_{j+1}\backslash \bar{G}_{j-1}$, and pick a bump function which is strictly positive in a neighborhood Q_x of x. The neighborhoods Q_x for each such x give a covering of $\bar{G}_{j+1}\backslash G_j$, a compact set, hence we can find a finite subcover. Doing this for all $x \in S$ we obtain a countable collection of functions $\{\psi_i\}_{i\in\mathbb{N}}$ subordinate to the cover $\{W_k\}_{k\in\mathbb{N}}$ such that $\psi_i \geq 0$. Now since the supports of the ψ_i are locally finite, the function

$$\psi = \sum_i \psi_i,$$

is well defined on S. Then the functions $\rho_i = \psi_i/\psi$ obviously satisfy our requirements. ☐

REFERENCES

[BT] R. Bott; L. Tu, *Differential forms in algebraic topology*, Graduate Texts in Mathematics, 82. Springer Verlag.

[E] C. H. Edwards, *Advanced Calculus of several variables*, Dover, 1994.

[H] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.

[MT] I. Madsen; J. Tornehave, *From Calculus to Cohomology*, Cambridge University Press, 1997.

[M] J. Milnor, *Topology from the differential viewpoint*, Princeton Landmarks in Mathematics, Princeton University Press, 1997.

[Z] V. Zorich, *Mathematical Analysis*, Vol. 1 & 2. Universitext, Springer-Verlag, 2004.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO.