# Frobenius Theorem Two Ways

Robert Lyons

March 4, 2016

Every smooth vector field $X$ on a manifold has integral curves which are embedded $1-$dimensional sub-spaces of $M$ that have tangent spaces spanned by the vector field $X$. Let $M$ be a manifold of dimension $m$ and assume that for every $x \in M$ there is a $p-$dimensional subspace $E_x$ of the tangent space of $M$ at $x$, $T_x(M)$. Is it possible to find a submanifold $N \subset M$ so that at every $y \in N$ the tangent space at $y \in N$ of $N$ is just $E_y = T_y(M)$. Is it enough that the planes $E_x$ vary smoothly with $x$? It turns out one can construct $p-$dimensional sub-spaces of the tangent spaces of $M$ that vary in a smooth way that are not the tangent spaces of any sub-manifold. Frobenius theorem gives necessary and sufficient conditions required of these $p-$planes so that they are the tangent planes of a $p-$dimensional sub-manifold.

In this paper we first present Frobenius theorem using the language of vector fields. The condition that the $p-$planes are the tangent spaces of a $p-$submanifold is that the distribution is closed under the Lie bracket. Following this we shall give an equivalent dual presentation using the language of differential forms. The $p-$planes are now the kernels of the differential forms. This second approach is in the spirit of the Cartan-Kahler theorem where the extension problem is turned into a problem of finding the zeros of an ideal of functions generated by the differential ideal.

Paper source:
This paper downloaded from: `www.pdx.edu/~rlyons`.
Send comments to: `rlyons@pdx.edu`.

## Contents

# 1 Differentiable Manifolds

We start with a short introduction to the definitions of differential geometry. There are many good books on the subject. An excellent modern treatment is [5]. Fora more elementary book that traces differential geometry back to it's roots in non-Euclidean geometry see [3].

## 1.1 Manifolds

To make the definition of a manifold precise the local Euclidean structure is modeled as a homeomorphism (continuous map with a continuous inverse) from a neighborhood to a subset of $\mathbb{R}^n$. This is called a coordinate map and is fundamental to the subject.

**Definition 1.1.1.** *A smooth differentiable manifold $M$ of dimension $m$ is a second countable Hausdorff space (for example some subset of $\mathbb{R}^K$ for some $K$) together with a collection of $C^\infty$ coordinate maps. In particular this means that there is a collection of open sets $U_\alpha \subset M$ and a collection of homeomorphisms $\varphi_\alpha : U_\alpha \to \mathbb{R}^N$ that satisfy,*

- $\bigcup_\alpha U_\alpha = M$,
- *When $U_\alpha \cap U_\beta \neq \emptyset$ then the mapping $\varphi_\alpha \circ \varphi_\beta^{-1} : \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$ is $C^\infty$.*

*We could replace $C^\infty$ in this definition with "real analytic" and we would get a real analytic differentiable manifold.*

The open sets $U_\alpha$ are called coordinate neighborhoods. The second countable axiom means that $M$, as a topological space, has a countable base. This axiom is essential so that $M$ has a partition of unity which allows us to define an integral in a natural way and show every manifold has a Riemannian metric. This axiom does not enter our discussion directly.

We denote the $k$-th coordinate projection by $r^k : \mathbb{R}^N \to \mathbb{R}$ which is defined as

$$r^k(a^1, \cdots, a^N) = a^k. \tag{1}$$

Using this we define coordinates on $M$ by,

$$x^k = r^k \circ \varphi_\alpha : U_\alpha \to \mathbb{R}. \tag{2}$$

If $x \in U_\alpha$ then $x$ is uniquely determined by its coordinates $\big(x^1(x), \cdots x^m(x)\big)$.

We use these coordinate neighborhoods to define many of the features found in Euclidean vector calculus. For example, a smooth map $f : M \to \mathbb{R}$ is defined to be smooth if the maps

$$f \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha) \to \mathbb{R}$$

are smooth for every coordinate neighborhood $(U_\alpha, \varphi_\alpha)$. Notice that $\varphi_\alpha$ is a homeomorphism so we can take the inverse map and it is continuous. We can take the derivative of a smooth function in the direction of coordinates $x^k$ by taking,

$$\frac{\partial(f \circ \varphi_\alpha^{-1})}{\partial x^k} = \frac{\partial f}{\partial r^k}.$$

We extend this idea and define tangent vectors and vector fields using the coordinate neighborhoods we defined above.

## 1.2 Tangent Vectors

In $\mathbb{R}^m$ a smooth curve has a derivative which is the tangent to the curve. In an abstract manifold we will define a tangent vector $X$ to be a directional derivative that takes a smooth function $f$ and associates a derivative $X(f)$ which we think of as the directional derivative of $f$ in the direction $X$.

For example, given a smooth differentiable manifold $M$ and given a curve $\gamma : (-a, a) \to M$ the vector $X$ tangent to $\gamma$ maps a smooth function $f : M \to \mathbb{R}$ to its directional derivative,

$$X(f) = \frac{d\gamma}{dt}(f) = \frac{d(f \circ \gamma)}{dt}. \tag{3}$$

We shall define a tangent vector, at a point $x \in M$, to be a operator that takes smooth functions $f : M \to \mathbb{R}$ to a real number $X(f)_x \in \mathbb{R}$. This operator is a derivation, which means that if $f, g$ are smooth functions at the point $x \in M$ then

$$X(fg)_x = X(f)_x g(x) + f(x)X(g)_x.$$

The set of vectors at a point $x \in M$ is denoted by $T_x(M)$ and is called the tangent space at $x$. We can add derivations and scale by real scalars $a, b \in \mathbb{R}$,

$$(aX + bY)(f) = aX(f) + bY(f).$$

This makes the set of derivations at $x$ into a real vector space. One can prove that this vector space has dimension $m$ (see [6] ). This makes intuitive sense since we think of $m-$dimensional surfaces in $\mathbb{R}^K$ have $m-$dimensional tangent spaces, just as $2-$dimensional surfaces in $\mathbb{R}^3$ has $2-$dimensional tangent planes.

A vector field is an assignment of a vector $X \in T_x(M)$ for every $x \in M$. A vector field is smooth if for every smooth $f : M \to \mathbb{R}$ the function $X(f) : M \to \mathbb{R}$ is a smooth mapping. Given a coordinate system $\varphi : U \to \mathbb{R}^N$ with coordinates $x^j = r^j \circ \varphi$ we define vector fields $\frac{\partial}{\partial x^k}$ by,

$$\frac{\partial}{\partial x^k}(f)_x = \left.\frac{\partial(f \circ \varphi^{-1})}{\partial r^k}\right|_{\varphi(x)}. \tag{4}$$

These $m$ tangent vectors $\frac{\partial}{\partial x^1}, \cdots, \frac{\partial}{\partial x^m}$ form a basis of the tangent space $T_x(M)$. For example, if $X \in T_x(M)$ we write,

$$X = \sum_{k=1}^{m} X^k \frac{\partial}{\partial x^k}.$$

To find the coefficients $X^k$ we apply $X$ to the function $x^j : M \to \mathbb{R}$. We note that,

$$\frac{\partial}{\partial x^k}\left(x^j\right) = \left.\frac{\partial(x^j \circ \varphi^{-1})}{\partial r^k}\right|_{\varphi(x)} = \left.\frac{\partial(r^j \circ \varphi \circ \varphi^{-1})}{\partial r^k}\right|_{\varphi(x)} = \left.\frac{\partial(r^j)}{\partial r^k}\right|_{\varphi(x)} = \delta^j_k.$$

Using this we see that we have,

$$X(x^j) = \sum_{k=1}^{m} X^k \frac{\partial}{\partial x^k}(x^j) = \sum_{k=1}^{m} X^k \delta^j_k = X^j.$$

This means that,

$$X = \sum_{j=1}^{m} X(x^j) \frac{\partial}{\partial x^j}.$$

Let $M$ and $N$ be differentiable manifolds of dimension $m$ and $n$ respectively. Let $\phi : M \to N$ be a smooth mapping. If $\gamma : I \to M$ is a smooth curve then $(\phi \circ \gamma) : I \to N$ is a smooth curve on $N$ The differential maps $T_x(M)$ in a way that corresponds to this mapping of curves. Using the association in equation 3 we write,

$$d\phi(\frac{d\gamma}{dt}) = \frac{d(\phi \circ \gamma)}{dt}$$

If we apply this to a function $f$ we get,

$$d\phi\left(\frac{d\gamma}{dt}\right)(f) = \frac{d(\phi \circ \gamma)}{dt}(f) = \frac{d(f \circ \phi \circ \gamma)}{dt} = \frac{d\gamma}{dt}\left(f \circ \phi\right).$$

So the differential $d\phi$ maps vectors in $T_x(M)$ to vectors in $T_{\phi(x)}(N)$. We formalize this with the following definition.

**Definition 1.2.1.** *The differential $d\phi : T_x(M) \to T_{\phi(x)}(N)$ is a mapping that maps $X$ to the derivation $d\phi(X)$ which takes a function $f : N \to \mathbb{R}$ to,*

$$d\phi(X)(f) = X(f \circ \phi).$$

**Example 1.2.2.** *Let $M = S^2 = \left\{x \in \mathbb{R}^3 | \|x\|^2 = R^2\right\}$. We define a parameterization for a fixed $R$*

$$\varphi^{-1} : (0, \pi) \times (0, 2\pi) \to S^2$$

*given by,*

$$\varphi^{-1}(\theta,\phi) = \begin{bmatrix} Rsin(\theta)cos(\varphi) \\ Rsin(\theta)sin(\varphi) \\ Rcos(\theta) \end{bmatrix}$$

*The coordinates* $(x^1, x^2)$ *have,*

$$x^1 : S^2 \to \mathbb{R} \text{ defined by } x^1\left(\varphi^{-1}(\theta,\phi)\right) = \theta,$$
$$x^2 : S^2 \to \mathbb{R} \text{ defined by } x^2\left(\varphi^{-1}(\theta,\phi)\right) = \phi,$$

## 1.3 Integral Curves

Given a smooth vector field $X$ and an $x \in M$ is there a curve $\gamma : (-a,a) \to M$ that has $X$ as its tangent vector? We need to solve the equation,

$$X(x^k) = X^k = \frac{d(x^k \circ \gamma)}{dt}$$

This ODE system has a unique solution which is called the integral curve of $X$.

Let $M$ be a $C^\infty$ manifold and let $X$ be a smooth vector field on $M$. If $x_0 \in M$ has $X(x_0) \neq 0$ then there is a smooth curve $\gamma : I \to M$ with $I = (a,b)$ with $a < 0 < b$ that satisfies,

$$\gamma(0) = x_0$$
$$\left.\frac{d\gamma}{dt}\right|_x = X(x)$$

For a more complete discussion of these issues in this section see [5].

This can be done in a smooth way so that for every $x_0 \in M$ with $X(x_0) \neq 0$ there is a neighborhood $U \subset M$ and an $\epsilon > 0$ and a smooth mapping $\varphi : (-\epsilon,\epsilon) \times M \to M$ that satisfies the following properties.

1. $\varphi(0,x) = x$
2. $\gamma(t) = \varphi_t(x)$ satisfies $\left.\frac{d\gamma}{dt}\right|_t = X(\varphi_t(x))$.
3. If $|t|, |s|, |t+s| < \epsilon$ and if $\varphi_t(x) \in U$ then $\varphi_{s+t}(x) = \varphi(s) \circ \varphi_t(x)$.

We denote $\varphi(t,x) = \varphi_t(x)$. This is called a local 1−parameter sub-group of diffeomorphisms.

## 1.4 Lie Derivative

Let $X$ and $Y$ be smooth vector fields. We define the Lie Bracket as,

$$[X,Y](f) = X\left(Y(f)\right) - Y\left(X(f)\right)$$

One can show that the Lie Bracket is a derivation and so $[X,Y]$ is a vector field on $M$.

The bracket is easily computed in coordinates. Let $(U,\varphi)$ be a coordinate neighborhood with coordinates $x^1,\cdots,x^m$. We write $X$ and $Y$ in coordinates,

$$X = \sum_{j=1}^{m} X^j \frac{\partial}{\partial x^j} \text{ and } Y = \sum_{j=1}^{m} Y^j \frac{\partial}{\partial x^j}.$$

Now compute the coordinates of $[X,Y]$ by applying to the coordinate function $x^k$,

$$\begin{aligned}[X,Y](x^k) &= X(Y(x^k)) - Y(X(x^k)) = X(Y^k) - Y(X^k) \\ &= \sum_{j=1}^{m}\left(X^j\frac{\partial Y^k}{\partial x^j} - Y^j\frac{\partial X^k}{\partial x^j}\right)\end{aligned}$$

Using the $1-$parameter group of diffeomorphisms we can define the Lie Derivative.

**Definition 1.4.1.** *Let $X$ be a vector field and let $Z$ be a tensor field.*

$$L_X(Z) = \lim_{t\to 0}\left(\tfrac{1}{t}\left(d\varphi_{-t}(Z_{\varphi_t(x)})_p - Z_p\right)\right)$$

The bracket and Lie derivative are, in fact, related.

**Proposition 1.4.2.** *Let $X$ and $Y$ be smooth vector fields then,*

$$L_X(Y) = [X,Y].$$

### 1.4.1 Exterior Algebras

Let $V$ be a vector space of dimension $n$. A $k-$tensor $T \in T(V)$ is a multi-linear mapping,

$$T : V \times V \times \cdots \times V = V^k \to \mathbb{R}.$$

A $k-$tensor is alternating if,

$$T(v_1,\cdots v_i,\cdots v_j,\cdots v_k) = -T(v_1,\cdots v_j,\cdots v_i,\cdots v_k).$$

This means permuting any two argument elements results in a negative value. We denote the alternating $k-$tensors by $\Lambda^k(V)$. From a $k-$tensor we can construct an alternating tensor with the definition,

$$Alt(T)(v_1, \cdots v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} T(v_{\sigma 1}, v_{\sigma 2}, \cdots, v_{\sigma k}).$$

If $dim(V) = n$ and $T \in \Lambda^n(V)$. If $e_1, \cdots e_n$ is a basis for $V$ and $v_i = \sum_j a_i^j e_j$ then

$$\begin{aligned} T(v_1, \cdots, v_n) &= \sum_{j_1, \cdots j_n = 1}^{n} a_1^{j_1} a_2^{j_2} \cdots a_n^{j_n} T(e_{j_1}, e_{j_2}, \cdots, e_{j_n})) \\ &= \sum_{\sigma \in S_n} a_1^{\sigma 1} a_2^{\sigma 2} \cdots a_n^{\sigma n} T(e_{\sigma 1}, e_{\sigma 2}, \cdots, e_{\sigma n})) \\ &= \sum_{\sigma \in S_n} (-1)^{sgn(\sigma)} a_1^{\sigma 1} a_2^{\sigma 2} \cdots a_n^{\sigma n} T(e_1, e_2, \cdots, e_n)) \\ &= det(a_i^j) T(e_1, \cdots e_n)). \end{aligned}$$

**Definition 1.4.3.** *Let $\omega \in \Lambda^k(V)$ and $\eta \in \Lambda^m(V)$ the the wedge product is defined by*

$$\omega \wedge \eta = \frac{(deg(\omega) + deg(\eta))!}{deg(\omega)! deg(\eta)!} Alt(\omega \otimes \eta).$$

Although this abstract formula might seem rather oblique, in simple cases is comes out to simple expressions. For example, if $\omega_1, \omega_2, \omega_3 \in \Lambda^1(V)$ then,

$$\begin{aligned} \omega^1 \wedge \omega^2 \wedge \omega^3 &= \omega^1 \wedge \left( \frac{2!}{1!1!} Alt(\omega^2 \otimes \omega^3) \right) \\ &= \frac{2!}{1!1!} \frac{3!}{2!1!} Alt\left(\omega^1 \otimes Alt(\omega^2 \otimes \omega^3)\right) \\ &= 3! Alt\left(\omega^1 \otimes \omega^2 \otimes \omega^3\right) \end{aligned}$$

This means that for $v_1, v_2, v_3 \in V$ we have,

$$\begin{aligned} \left(\omega^1 \wedge \omega^2 \wedge \omega^3\right)(v_1, v_2, v_3) &= 3! Alt\left(\omega^1 \otimes \omega^2 \otimes \omega^3\right)(v_1, v_2, v_3) \\ &= 3! \frac{1}{3!} \sum_{\sigma \in S_3} \left(\omega^1 \otimes \omega^2 \otimes \omega^3\right)(v_{\sigma 1}, v_{\sigma 2}, v_{\sigma 3}) \\ &= \sum_{\sigma \in S_3} \omega^1(v_{\sigma 1}) \omega^2(v_{\sigma 2}) \omega^3(v_{\sigma 2}) = det\left(\omega^i(v_j)\right) \end{aligned}$$

More generally we have the following,

**Proposition 1.4.4.** *Let $\theta^1, \cdots \theta^k \in \Lambda^1(V)$ and $v_1, \cdots, v_k \in V$ then*

$$\left(\theta^1 \wedge \theta^2 \cdots \wedge \theta^k\right)(v_1, \cdots, v_k) = \sum_{\sigma \in S_k} \theta^1(v_{\sigma 1})\theta^2(v_{\sigma 2}) \cdots \theta^k(v_{\sigma k}) = det\left(\theta^i(v_j)\right)$$

Given a basis $e_1, \cdots e_n$ of $V$ with dual basis $\omega^1, \cdots \omega^n$ then every $k-$form $\eta$ can be written as,

$$\eta = \sum_{i_1, i_2, \cdots i_k = 1}^{n} \omega_{i_1 i_2 \cdots i_k} \omega^{i_1} \wedge \omega^{i_2} \wedge \cdots \wedge \omega^{i_k}.$$

There is redundancy in the coordinates $\omega_{i_1 i_2 \cdots i_k}$. They are alternating, in the sense that

$$\omega_{i_1 \cdots i_a \cdots i_b \cdots i_k} = -\omega_{i_1 \cdots i_b \cdots i_a \cdots i_k}.$$

## 1.5 Differential Forms

At $x \in M$ the dual space to the tangent space is denoted by $T^*(M)$ and it consists of $1-$forms $\omega : T_x(M) \to \mathbb{R}$. A differential form is an element of the dual space at every point $x \in M$. A differential form is smooth if $\omega(X) : M \to \mathbb{R}$ for all smooth vector fields $X$. For example, for $f \in C^\infty(M)$ we can construct the $1-$form we call the gradient and for every vector field $X$ we define

$$df(X) = X(f). \tag{5}$$

We can define the set of multi-linear functionals on the tangent space

$$\omega : T_x(M) \times \cdots \times T_x(M) \to \mathbb{R}$$

We define an operator $\wedge$ defined for $\omega \in \Lambda^k(T_x(M))$ and $\eta \in \Lambda^m(T_x(M))$ and results in a form $\omega \wedge \eta \in \Lambda^{k+m}(T_x(M))$. For example, if $k = m = 1$ and for $v, w \in T_x(M)$, we have,

$$(\omega \wedge \eta)(v, w) = \omega(v)\eta(w) + (-1)^k \omega(w)\eta(v) = \omega(v)\eta(w) - \omega(w)\eta(v).$$

A $k-$form $\omega$ on $M$ is a $k-$form at every point $x \in M$. The form $\omega$ is smooth if for every smooth vector fields $X_1, \cdots X_k$ the function,

$$x \to \omega_x(X_1, \cdots X_k),$$

is a smooth function on $M$. We denote the smooth $k-$forms on $M$ by $\Lambda^k(M)$.

Given a coordinate neighborhood $U$, the vector fields $\frac{\partial}{\partial x^k}$ form a basis of $T_x(M)$ for every $x \in U$. The dual forms to these vector fields are denoted $dx^j$ and satisfy

$$dx^j \left(\frac{\partial}{\partial x^k}\right) = \frac{\partial}{\partial x^k}\left(x^j\right) = \left.\frac{\partial(r^j \circ \varphi \circ \varphi^{-1})}{\partial r^k}\right|_{\varphi(x)} = \delta^j_k.$$

If $f \in C^\infty(M)$ then $df$ is a differential defined by equation 5. We can write this $1-$form in terms of the basis $dx^j$,

$$df = \sum_{j=1}^{m} \lambda_j dx^j.$$

To find the coefficients we apply this $1-$form to the vector $\frac{\partial}{\partial x^k}$,

$$df\left(\frac{\partial}{\partial x^k}\right) = \frac{\partial}{\partial x^k}(f) = \frac{\partial f}{\partial x^k} = \lambda^k.$$

This means we can write $df$ as,

$$df = \sum_{j=1}^{m} \frac{\partial f}{\partial x^j} dx^j.$$

We define an operator $d : \Lambda^k(M) \to \Lambda^{k+1}(M)$ called the differential. Given a form with coordinates

$$\omega = \sum_{i_1,\cdots k_p} \omega_{i_1,\cdots,i_p} dx^{i_1} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}$$
$$d\omega = \sum_{i_1,\cdots k_p} \frac{\partial \omega_{i_1,\cdots i_p}}{\partial x^j} dx^{i_1} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}$$

**Example 1.5.1.** *The sphere is the constant surfaces of the function,*

$$f(x, y, z) = x^2 + y^2 + z^2 - R^2.$$

*The differential of this is given by,*

$$\omega = 2xdx + 2ydy + 2zdz.$$

*The tangent to the sphere is the kernel of this 1-form.*

One can define this without coordinates. In two dimensions this formula is,

$$d\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]). \tag{6}$$

## 1.6 Coordinate Transforms

In this section we describe how vectors and forms transform under coordinate transforms. We use the formal description of vectors and forms to deduce the transform properties but try to establish a link with older notation systems.

Let $M$ be a differentiable manifold of dimension $m$ and $(U, \varphi)$ be a coordinate neighborhood with coordinates $(x^1, x^2, \cdots x^m)$ where $x^k = r^k \circ \varphi : U \to \mathbb{R}$. Recall that $r^k : \mathbb{R}^m \to \mathbb{R}$ is the projection onto the $k$th coordinate so that,

$$r^k \left( \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^m \end{bmatrix} \right) = a^k.$$

If we transform the open set $\varphi(U)$ with a diffeomorphism $T : \mathbb{R}^m \to \mathbb{R}^m$ then we get a new coordinate mapping,

$$U \xrightarrow{\varphi} \varphi(U) \xrightarrow{T} \mathbb{R}^m, \tag{7}$$

and we define $\psi = T \circ \varphi$. The mapping $T$ is a diffeomorphism so $\psi$ is a homeomorphism and it has coordinates,

$$y^k = r^k \circ \psi = r^k \circ T \circ \varphi.$$

The vector fields $\frac{\partial}{\partial x^k}$ form a basis of $T_x(M)$ for all $x \in U$ are defined by,

$$\begin{aligned} \frac{\partial}{\partial x^k}(f) &= \frac{\partial (f \circ \varphi^{-1})}{\partial r^k} \\ &= \frac{\partial (f \circ \varphi^{-1} \circ T^{-1} \circ T)}{\partial r^k} = \frac{\partial (f \circ \psi^{-1} \circ T)}{\partial r^k} \\ &= \frac{\partial (f \circ \psi^{-1})}{\partial r^j} \frac{\partial (r^j \circ T)}{\partial r^k} \\ &= \frac{\partial (r^j \circ T)}{\partial r^k} \frac{\partial}{\partial y^j}(f). \end{aligned}$$

Since this is true for all $f$ we write this as,

$$\frac{\partial}{\partial x^k} = \frac{\partial (r^j \circ T)}{\partial r^k} \frac{\partial}{\partial y^j} \tag{8}$$

In a similar way, we can see that,

$$\begin{aligned} \frac{\partial}{\partial y^k}(f) &= \frac{\partial (f \circ \varphi^{-1} \circ T^{-1})}{\partial r^k} \\ &= \frac{\partial (f \circ \varphi^{-1})}{\partial r^i} \frac{\partial (r^i \circ T^{-1})}{\partial r^k} \end{aligned}$$

This leads to the inverse transform,

$$\frac{\partial}{\partial y^k} = \frac{\partial (r^j \circ T^{-1})}{\partial r^k} \frac{\partial}{\partial x^j} \tag{9}$$

We can write this in a different way by performing the following computation,

$$\frac{\partial x^j}{\partial y^k} = \frac{\partial}{\partial x^j}\left(y^k\right) = \frac{\partial(y^k \circ \varphi^{-1})}{\partial r^j} = \frac{\partial(r^k \circ T \circ \varphi \circ \varphi^{-1})}{\partial r^j} = \frac{\partial(r^k \circ T)}{\partial r^j}$$

This expression is the same as is found in equation 8. Similarly, we see that,

$$\frac{\partial y^j}{\partial x^k} = \frac{\partial}{\partial y^j}\left(x^k\right) = \frac{\partial(x^k \circ \varphi^{-1} \circ T^{-1})}{\partial r^j} = \frac{\partial(r^k \circ T^{-1})}{\partial r^j}$$

Using these we re-write equations 8 and 9 as following,

$$\frac{\partial}{\partial x^k} = \frac{\partial(r^j \circ T)}{\partial r^k}\frac{\partial}{\partial y^j} = \frac{\partial y^j}{\partial x^k}\frac{\partial}{\partial y^j} \tag{10}$$

$$\frac{\partial}{\partial y^k} = \frac{\partial(r^j \circ T^{-1})}{\partial r^k}\frac{\partial}{\partial x^j} = \frac{\partial x^j}{\partial y^k}\frac{\partial}{\partial x^j}. \tag{11}$$

We say that vectors transform contravariantly.

Using this formula we can compute the change of basis for differential forms,

$$\begin{aligned} dy^k\left(\frac{\partial}{\partial x^i}\right) &= dy^k\left(\frac{\partial(r^j \circ T)}{\partial r^i}\frac{\partial}{\partial x^j}\right) \\ &= \frac{\partial(r^j \circ T)}{\partial r^i} dy^k\left(\frac{\partial}{\partial x^j}\right) = \frac{\partial(r^j \circ T)}{\partial r^i}\delta^k_j \\ &= \frac{\partial(r^k \circ T)}{\partial r^i} \end{aligned}$$

This means that,

$$dy^k = \frac{\partial y^k}{\partial x^j} dx^j. \tag{12}$$

The converse is computed in a similar fashion,

$$dx^k = \frac{\partial x^k}{\partial y^j} dy^j. \tag{13}$$

We say that vectors transform covariantly.

### 1.6.1 Linear Coordinate Transforms

If we assume that $T : \mathbb{R}^m \to \mathbb{R}^m$ is linear then the coordinates transform described in equation 7. Let $(U, \varphi)$ be a coordinate neighborhood with coordinates $(x^1, \cdots, x^m)$ and let $e_1, \cdots e_m$ be the standard basis for $\mathbb{R}^m$. If $T : \mathbb{R}^m \to \mathbb{R}^m$ is a linear transform then

$$y^j = r^j \circ T \circ \varphi = r^j \circ T(e_i)(r^i \circ \varphi) = T^j_i x^i. \tag{14}$$

Now we compute the coordinate transform matrix,

$$\frac{\partial y^k}{\partial x^j} = \frac{\partial (r^k \circ T^{-1})}{\partial r^j} = \left(T^{-1}\right)^k_j, \tag{15}$$

$$\frac{\partial x^k}{\partial y^j} = \frac{\partial (r^k \circ T)}{\partial r^j} = T^k_j. \tag{16}$$

### 1.7 Implicit Function Theorem

The implicit function theorem is needed in several proofs in the following sections. Wew state this theorem here without proof.

**Theorem 1.7.1.** *The Inverse Function Theorem. Let $F : \mathbb{R}^m \to \mathbb{R}^m$ be a smooth mapping on an open set $V \subset \mathbb{R}^m$. If det $\left(\frac{\partial F^i}{\partial r^j}\right) \neq 0$ on $V$. Then for all $x_0 \in V$ there is a neighborhood $W$, $x_0 \in W \subset V$ such that $F : W \to F(W)$ is a diffeomorphism (e.g. $F^{-1}$ exists and is a diffeomorphism).*

For a proof see [4].

**Theorem 1.7.2.** *The Implicit Function Theorem. Let $F : \mathbb{R}^{n+k} \to \mathbb{R}^k$ be a smooth function and assume that $F(a) = 0$. If the Jacobian $Df(a)$ has rank $k$ then there is an open neighborhood $U \subset \mathbb{R}^{n+k}$ and a smooth diffeomorphism $h : \mathbb{R}^{m+n} \to \mathbb{R}^{m+n}$ so that $F \circ h : U \to \mathbb{R}^k$ so that,*

$$(F \circ h)\left(x^1, x^2, \cdots x^{m+n}\right) = \left(x^{n+1}, \cdots x^{n+m}\right).$$

This follows readily from Theorem 1.7.1. For a proof see [4].

An immediate consequence of this is that if $F$ has maximal rank at the point $a$ and $F(a) = 0$ then

$$(F \circ h)^{-1}(0) \cap U = \left\{ (x^1, \cdots, x^n) \middle| (x^1, \cdots, x^n, x^{n+1}, \cdots x^{n+k}) \in U \right\}$$

## 2 Frobenius Theorem And Vector Fields

In this section we shall discuss Frobenius thoerem using the language of vector fields. All the details of this section are explained nicely in [6].

Every smooth vector field generates a family of smooth curves that have tangent vector $X$. We would like to examine whether this has an analogue in higher dimensions. For example, given two smooth vector fields $X, Y$ on a manifold $M$, is there a $2-$dimensional submanifold $N$ that has $span\{X, Y\}$ as its tangent space. Of course we must have $X$ and $Y$ linear independent at every point so that the span is actually $2-$dimensional. Are there any other conditions? In Frobenius theorem we shall see that there is one condition, along with linearly independence, that guarantees the existence of this submanifold. In this section we discuss the theorem using the language of vector fields and in section 3 we present Frobenius' theorem using the language of differential forms and exterior differential analysis.

But we state Frobenius theorem we start with a few simple examples. Let $M = \mathbb{R}^3$ and define two vector fields on the manifold $M$, which we write write in terms of the coordinate basis and as standard $\mathbb{R}^3$ vectors.

$$X(x, y, z) = \frac{\partial}{\partial x} + x\frac{\partial}{\partial z} = \begin{bmatrix} 1 \\ 0 \\ x \end{bmatrix} \tag{17}$$

$$Y(x, y, z) = \frac{\partial}{\partial y} + y\frac{\partial}{\partial z} = \begin{bmatrix} 0 \\ 1 \\ y \end{bmatrix} \tag{18}$$

Fix $(x, y, z) \in \mathbb{R}^3$. Can we find a 2−dimensional surface $N$ that has tangent spaces formed by the basis $X, Y$. If $\gamma$ is a curve with tangent vector $X$ and it has $\varphi(\gamma(0)) = (x, y, z)$ then it must satisfy,

$$\begin{aligned} \frac{d(x \circ \gamma)}{dt} &= 1, \\ \frac{d(y \circ \gamma)}{dt} &= 0, \\ \frac{d(z \circ \gamma)}{dt} &= x. \end{aligned}$$

Let $(x, y, z) \in \mathbb{R}^3$ be any point in $M$ and let $\gamma(0) = (x, y, z)$ then we can solve for $\gamma$ uniquely. We get the following solution as a 1−parameter group of diffeomorphisms,

$$\varphi_t(x, y, z) = \begin{bmatrix} x + t \\ y \\ z + xt + \frac{1}{2}t^2 \end{bmatrix} \tag{19}$$

Similarly we can solve $Y$ globally and get the 1−parameter group given by $\eta$,

$$\eta_s(x, y, z) = \begin{bmatrix} x \\ s + y \\ z + ys + \frac{1}{2}s^2 \end{bmatrix} \tag{20}$$

So, around the point $(x, y, z)$ we can construct a 2−dimensional manifold given by

$$\Sigma(t, s) = \begin{bmatrix} t + x \\ s + y \\ z + xt + ys + \frac{1}{2}t^2 + \frac{1}{2}s^2 \end{bmatrix}$$

This parameterization is the inverse of a coordinate neighborhood. In this case we can weave the two 1-dimensional integral curves for $X$ and $Y$ together to form a smooth 2-dimensional manifold.

To see what can go wrong we look at another simple 2−dimensional example, Let $M = \mathbb{R}^3$ and construct two smooth vector fields,

$$X(x, y, z) = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z} \tag{21}$$

$$Y(x, y, z) = \frac{\partial}{\partial y} - x\frac{\partial}{\partial z}. \tag{22}$$

We have coordinates $x^j$ with $j = 1, 2, 3$ where $x^1 = x, x^2 = y, x^3 = z$.

Each of these smooth vector fields generates a local 1-parameter group of diffeomorphisms $\varphi : I \times M \to M$ where $I \subset \mathbb{R}^1$. The curves through the point $x \in M$ is $t \to \varphi_t(x)$. For the vector field $X$ the 1−parameter group of diffeomorphisms satisfy,

$$\frac{d(x \circ \varphi))}{dt} = X(x) = X^1(x, y, z) = 1,$$
$$\frac{d(y \circ \varphi)}{dt} = X(y) = X^2(x, y, z) = 0,$$
$$\frac{d(z \circ \varphi)}{dt} = X(z) = X^3(x, y, z) = y.$$

From these conditions we can write down the solution,

$$\varphi_t(x, y, z) = \begin{bmatrix} x + t \\ y \\ z + yt \end{bmatrix}. \tag{23}$$

For the vector field $Y$ let $\eta : I \times M \to M$ be the 1 =parameter group of diffeomorphisms then $\eta$ satisfies,

$$\frac{d(x \circ \eta))}{dt} = Y(x) = Y^1(x, y, z) = 0,$$
$$\frac{d(y \circ \eta)}{dt} = Y(y) = Y^2(x, y, z) = 1,$$
$$\frac{d(z \circ \eta)}{dt} = Y(z) = Y^3(x, y, z) = -x.$$

From these conditions we can write down the solution,

$$\eta_s(x, y, z) = \begin{bmatrix} x \\ y + s \\ z - xs \end{bmatrix}. \tag{24}$$

Do these two $1-$dimensional solutions combine to form a $2-$dimensional surface that has a tangent space that is spanned by the vector fields $X$ and $Y$. We start at a point $(x, y, z) \in M$ and progress along a rectangle of coordinates. First we proceed along $X$ integral curves by $t$ and then along $Y$ by $s$,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \xrightarrow{\varphi_t} \begin{bmatrix} x+t \\ y \\ z+yt \end{bmatrix} \xrightarrow{\eta_s} \begin{bmatrix} x+t \\ y+s \\ (z+yt)-(x+t)s \end{bmatrix} \xrightarrow{\varphi_{-t}} \begin{bmatrix} x \\ y+s \\ z-xs-2ts \end{bmatrix} \tag{25}$$

We compare this when we proceed along $Y$ by $s$ and then $X$ by $t$.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \xrightarrow{\eta_s} \begin{bmatrix} x \\ y+s \\ z-xs \end{bmatrix}. \tag{26}$$

A comparison of equation 25 and equation 26 we see that the $x-$ and $y-$coordinates are the same however the $z-$coordinate is different. In this case there is no integral submanifold.

## 2.1 Distributions

Let $M$ be a differentiable manifold of dimension $m$ and let $N \subset M$ be a $p-$dimensional submanifold. The tangent space $T_x(N)$ is a $p-$dimensional subspace of $T_x(M)$. The two examples discussed earlier both describe two vector fields $X$ and $Y$ that span $p-$dimensional subspace (where $p = 2$) of the original manifold $M = \mathbb{R}^3$. These $p-$planes are an important part of the our discussion so we define the following.

**Definition 2.1.1.** *Let $M$ be a manifold of dimension $m$ and let $p$ be a positive integer $p \leq m$. A $p-$dimensional distribution $\mathcal{D}$ is a choice of a $p-$dimensional subspace of $T_x(M)$ for every $x \in M$. We denote the plane at $x$ by $\mathcal{D}(x)$. The distribution $\mathcal{D}$ is smooth if for every $x \in M$ there is a neighborhood $x \in U \subset M$ and smooth vector fields $X_1, \cdots X_p$ defined on $M$ so that $\mathcal{D}(x)$ is spanned by $X_1(x), \cdots, X_p(x)$ for every $x \in U$.*

Both of the examples discussed above generate smooth $2-$dimensional distributions but only one of these distributions was the tangent bundle of a sub-manifold. The $2-$dimensional manifold is called an integral manifold of the $2-$dimensional distribution.

**Definition 2.1.2.** *A sub-manifold $N \subset M$ with $i : N \to M$ is an integral manifold of the distribution $\mathcal{D}$ if,*

$$di(T_y(N)) = \mathcal{D}(i(y)),$$

*for each $y \in N$.*

This definition just states that the tangent space of $N$ at $y$ corresponds to the distribution $\mathcal{D}(y)$. The key property that insures the existence of integral sub-manifolds is the following.

**Definition 2.1.3.** *A smooth distribution $\mathcal{D}$ is called involutive if $[X,Y] \in \mathcal{D}$ whenever $X$ and $Y$ are smooth vector fields that lie in $\mathcal{D}$. This means that if $X_1, \cdots X_p$ are smooth vector fields that span $\mathcal{D}$ in a neighborhood $U$ then,*

$$[X_i, X_j] = C_{ij}^k X_k.$$

Before we get to the main result we need one preliminary technical proposition. The coordinates described in the following proposition are called flowbox coordinates.

**Proposition 2.1.4.** *Let $M$ be a manifold of dimension $m$ and let $X$ be a smooth vector field on $M$. If $x \in M$ has $X(x) \neq 0$ then there is a coordinate neighborhood $U \subset M$ of $x$ with coordinates $x^1, \cdots x^m$ such that,*

$$X|_U = \frac{\partial}{\partial x^1},$$

*Proof.* Let $x_0 \in M$ and find a coordinate system $(U, \phi)$ with coordinates $y^1, \cdots y^m$ such that $x_0 \in U$ and

$$X_{x_0} = \left.\frac{\partial}{\partial y^1}\right|_{x_0}$$

We simplify the coordinate system by translating so that $y^k(x_0) = 0$. In, perhaps, a smaller, neighborhood $U$ of $x_0$ we have a $1-$parameter family of diffeomorphisms $\varphi_t : U \to \varphi_t(U) \subset M$. We look at the surface $y^1 = 0$ which includes $x_0$ and move along the integral curves of $X$ to form a coordinate system. To do this we define a mapping $\psi : V \subset \mathbb{R}^m \to U$ by

$$\psi(t, a_2, a_3, \cdots a_m) = \varphi_t(\phi^{-1}(0, a_2, \cdots, a_m)).$$

First we note that $\psi(0, \cdots, 0) = x_0$. At $a_2 = a_3 = \cdots = a_m = 0$ we have

$$\psi(t, 0, \cdots, 0) = \varphi_t(\phi^{-1}(0, 0, \cdots, 0)) = \varphi_t(x_0).$$

This is a curve along the curve $X$ and the derivative at $t = 0$ is just $X_{x_0}$. At $t = 0$ we also have

$$\psi(0, a_2, a_3, \cdots a_m) = \phi^{-1}(0, a_2, \cdots, a_m).$$

This means that at $t = 0$ the curves along the coordinates $a_k$ have tangent vector $\frac{\partial}{\partial x^k}$. So, at the point $(0, \cdots, 0)$ the Jacobian of $\psi$ takes the $m-$linearly independent vectors $e_1, \cdots e_m$ to linearly independent vectors $X_{x_0}, \frac{\partial}{\partial x^2}_{x_0}, \cdots \frac{\partial}{\partial x^m}_{x_0}$. This mapping has an invertible Jacobian at $x_0$ and by the inverse mapping theorem the mapping $\psi$ is a diffeomorphism in some neighborhood of $x_0$. The inverse of this diffeomorphism is a coordinate mapping with the desired properties. □

In Theorem 2.2.1 we shall prove that there is an integral sub-manifold of a distribution if and only if the distribution is involutive.

## 2.2 Proof of Frobenius Theorem

The main result is the following.

**Theorem 2.2.1.** *Let $M$ be a manifold of dimension $m$. Let $\mathcal{D}$ be a $p-$dimensional $C^\infty$ distribution. If $\mathcal{D}$ is involutive then for every $x \in M$ there exists an integral sub-manifold $N$ that contains $x$. In fact, there is a coordinate neighborhood $U$ centered at $x$ such that the slices,*

$$x_i = K_i, \text{ for all } i \in p+1, \cdots N,$$

*are integral sub-manifolds of $\mathcal{D}$. If $N$ is a connected integral sub-manifold of $U$ then $i(N) \cap U$ corresponds to one of these slices.*

*Proof.* Let $N$ is a connected integral sub-manifold with $di(N)|_x = \mathcal{D}(i(x))$. For every $x \in N$ there is a neighborhood $U$ of $M$ with coordinates $x^1, \cdots x^m$ so that $N$ is given by $x^{p+1} = \cdots = x^m = 0$. If the $X_i$ are tangent to $N$ then

$$X_k = \sum_{j=1}^{p} X_k^j \frac{\partial}{\partial x^j},$$

where $X_k^j$ is a $p \times p$ invertible matrix. Let $Y = X^{-1}$. Now compute the bracket,

$$\begin{aligned}
[X_k, X_m] &= \sum_{j=1}^{p} \sum_{i=1}^{p} [X_k^j \frac{\partial}{\partial x^j}, X_m^i \frac{\partial}{\partial x^i}] \\
&= \sum_{j=1}^{p} \sum_{i=1}^{p} \left( X_k^j \frac{\partial X_m^i}{\partial x^j} - X_m^j \frac{\partial X_k^i}{\partial x^j} \right) \frac{\partial}{\partial x^i} \\
&= \sum_{j=1}^{p} \sum_{i=1}^{p} \left( X_k^j \frac{\partial X_m^i}{\partial x^j} - X_m^j \frac{\partial X_k^i}{\partial x^j} \right) Y_i^l X_l.
\end{aligned}$$

This means that the distribution is involutive.

We will prove the converse by induction. Let $p = 1$ then we have a single smooth vector field $X_1$ that spans the $1-$dimensional distribution $\mathcal{D}(x)$ at every $x \in M$. The integral curve of $X$ is the $1-$dimensional manifold through $x$ that has tangent space equal to the span of $X$. This proves the theorem for $p = 1$. In this case $[X, X] = 0$ so the involutive assumption is superfluous.

Assume the theorem is true for $p - 1$ and we shall show it is true for $p$. Let $X_1, \cdots X_p$ be smooth vector fields that span the distribution $\mathcal{D}$ in a coordinate neighborhood of $x \in M$.

By Proposition 2.1.4 we can take a neighborhood $U_1$ of $x$ contained in $U$ with coordinates $y^1, \cdots, y^m$ such that,

$$X_1(x) = \left.\frac{\partial}{\partial x^1}\right|_x .$$

We shall form a $p-1$ distribution on the $m-1$ dimensional manifold formed by setting $y^1 = K^1$. Let $W \subset U_1$ be the $m-1$ sub-manifold determined by the condition $y^1 = K^1$ where $K^1$ is a constant. For every $x \in W$ the distribution is determined by $X_1$ and $X_2, \cdots X_p$. We know that $X_1$ does not lie in $W$ so we use this to find projections of $X_k$ along $W$ as follows. Define $Y_k$ for $k = 2, \cdots p$ by,

$$Y_k = X_k - X_k(y^1)X_1. \tag{27}$$

We also define $Y_1 = X_1$ so that

$$\mathcal{D}(x) = span\{X_1(x), \cdots, X_p(x)\} = span\{Y_1(x), \cdots, Y_p(x)\}.$$

Notice, however, that $Y_2, \cdots Y_k$ are tangent to the sub-manifold determined by $y^1 = K^1$ for reasonable choices of $K^1$. To see this we can show that $Y_k$ we compute the derivative of $y^1$ in the direction $Y_k$,

$$Y_k(y^1) = X_k(y^1) - X_k(y^1)X_1(y^1) = 0. \tag{28}$$

The sub-manifold $W$ is one leaf of a foliation and we can define a mapping $\pi : U_1 \to W$ by,

$$\pi(y^1(x), \cdots, y^m(x)) = (y^2(x), \cdots, y^m(x)).$$

This mapping is smooth. On the sub-manifold $W$ we define vector fields $Z_k$ on $W$ for $k = 2, \cdots p$ by,

$$Z_k(w) = Y_k(w).$$

We define a distribution on $W$ by,

$$\mathcal{E}(w) = span < Z_2, \cdots Z_p > .$$

The $Z_k$ span a distribution that is the projection of $\mathcal{D}$ onto $T_x(W)$ for every $w \in W$. To use the induction hypothesis we would like to show that this distribution is involutive. To show this we must compute the commutator of $Z_k$. Let $f : W \to \mathbb{R}$ be a smooth function and extend $f$ to a smooth function $\tilde{f} : U_1 \to \mathbb{R}$ by,

$$\tilde{f}(y^1, \cdots y^m) = f(y^2, \cdots y^m).$$

We compute the commutator for $i, j = 2, \cdots m$,

$$\begin{aligned}[Z_i, Z_j](f) &= [Y_i, Y_j](\tilde{f}) \\ &= C_{ij}^1 Y_1(\tilde{f}) + \sum_{k=2}^{m} C_{ij}^k Y_k(\tilde{f}) \\ &= \sum_{k=2}^{m} C_{ij}^k Y_k(\tilde{f}) \\ &= \sum_{k=2}^{m} C_{ij}^k Z_k(f)\end{aligned}$$

This means that $[Z_i, Z_j]_w \in \mathcal{E}(w)$ for all $w \in W$ and the distribution $\mathcal{E}$ is involutive. By induction there is a $p-1$ dimensional sub-manifold $S \subset W$ with,

$$T_w(S) = \mathcal{E}(w),$$

for all $w \in W$. By the induction hypothesis for any $w \in W$ there is a coordinate system $w^2, \cdots w^m$ on $W$ so that $S$ is the sub-manifold determined by the condition,

$$S = \left\{ w \in W |\, w^{p+1}(w) = K^{p+1}, \cdots w^m(w) = K^m \right\}.$$

By varying $K^j$ we form a foliation in $W$. We meed to extend our $p-1$ sub-manifold $S \subset W$ to a $p$ dimensional sub-manifold $N \subset U_1$. To do this we define convenient coordinates $x^k$. For $x \in U_1$ define

$$x^1(x) = y^1(x), x^2(x) = w^2(\pi(x)), \cdots x^m(x) = w^m(\pi(x)).$$

We can easily extend $S$ along $x^1 = y^1$ by defining,

$$N = \left\{ x \in U_1 |\, (x^2(x), \cdots x^p(x)) \in W \text{ and } x^{p+1} = \cdots x^m = 0 \right\}$$

This is clearly a smooth sub-manifold of $U_1$ but it is not obvious that the vectors $Y_1, \cdots Y_p$ are tangent to $N$. Clear $Y_1$ is tangent to $N$ as any curve that starts in $N$ and is tangent to $Y_1$ remains in $N$. To show that $Y_1$ is always tangent to $N$ note that,

$$Y_1(x^{p+r}) = \frac{\partial}{\partial x^1}\left(x^{p+r}\right) = 0, \text{ for } r = 1, \cdots m-p. \tag{29}$$

The vectors $Y_2, \cdots Y_p$ are more problematic. We must show that that for all $k = 2, \cdots, m$ we have,

$$Y_k(x^{p+r}) = 0, \text{ for } r = 1, \cdots m-p.$$

We know that this holds true on the sub-manifold $W$ as,

$$Y_k(x^{p+r})\big|_{x^1(x)=0} = Z_k(w^{p+r})\big|_{x^1(x)=0} = 0.$$

To show that this is true for general $x$ compute the derivative in the $x^1$ direction,

$$\begin{aligned}\frac{\partial}{\partial x^1}\left(Y_k(x^{p+r})\right) &= Y_1\left(Y_k(x^{p+r})\right)\\ &= Y_1\left(Y_k(x^{p+r})\right) - Y_k\left(Y_1(x^{p+r})\right)\\ &= [Y_1, Y_k](x^{p+r})\\ &= C^1_{1k}Y_1(x^{p+r}) + \sum_{j=2}^{m} C^j_{1k}Y_j(x^{p+r})\\ &= \sum_{j=2}^{m} C^j_{1k}Y_j(x^{p+r})\end{aligned}$$

where we have used that $Y_1, \cdots Y_p$ are involutive and we have used equation 29 again. We get $p-1$ differential equations for the $p-1$ functions $Y_k(x^{p+r})$,

$$\frac{\partial}{\partial x^1}\left(Y_k(x^{p+r})\right) = \sum_{j=2}^{m} C^j_{1k}Y_j(x^{p+r}).$$

The system is a homogeneous linear system that must have a unique solution. However, we know that the $p-1$ functions $Y_k(x^{p+r})$ vanish at $x^1 = 0$ so they must vanish for all $x^1$ for $r = 1, \cdots m-p$ since that is a valid solution and it is unique. We conclude conclude that for every $k$ we have,

$$Y_k(x^{p+r}) = 0, \text{ for } r = 1, \cdots, m-p.$$

This means we can write each $Y_k$ in terms of the basis,

$$Y_k = \sum_{j=1}^{m} Y^j_k \frac{\partial}{\partial x^j}.$$

Apply this derivation to $x^{p+r}$ and we get,

$$Y_k(x^{p+r}) = 0 = Y^{p+r}_k.$$

This means that we can write,

$$Y_k = \sum_{j=1}^{p} Y^j_k \frac{\partial}{\partial x^j}.$$

By equation 27 this means we can write,

$$X_k = \sum_{j=1}^{p} X^j_k \frac{\partial}{\partial x^j}.$$

The manifold $N$ is formed by the equations $x^{p+1} = \cdots = x^m = 0$. This means that $\mathcal{D}(x) = T_x(N)$ for all $x \in U_1$. This proves the case for $p$ and the theorem follows by induction. □

Frobenius theorem insures that for every $x \in M$ there is a submanifold $N \subset M$ so that $x \in N$ and for every $y \in N$ the tangent space $T_y(N)$ coincides with the distribution through $y \in M$. This slicing is called a foliation.

# 3 Frobenius Theorem And Differential Forms

In this section we discuss Frobenius theorem using the language of differential forms. For a more complete discussion see [2].

## 3.1 Introduction

We start with a simple example. Let $M = \mathbb{R}^3$ and define,

$$f(x, y, z) = x^2 + x^2 + z^2 - r^2.$$

The implicit function theorem says that

$$f^{-1}(0) = \left\{ (x, y, z) | \, x^2 + y^2 + z^2 = r^2 = 0 \right\} = S^2(r)$$

is a smooth manifold. If we take the differential of $f$ we get,

$$df = 2xdx + 2ydy + 2zdz.$$

Let $\theta = df$. The mapping $\theta : T_x(M) \to \mathbb{R}$ has a two dimensional kernel,

$$E_x = ker(\theta) \leq T_x(M).$$

Let $v \in E_x$ and let

$$v = \alpha^1 \frac{\partial}{\partial x} + \alpha^2 \frac{\partial}{\partial y} + \alpha^3 \frac{\partial}{\partial z}.$$

The condition that $v \in E_x$ is

$$\begin{aligned} \theta(v) &= 2xdx \left( \alpha^1 \frac{\partial}{\partial x} \right) + 2ydy \left( \alpha^2 \frac{\partial}{\partial y} \right) + 2zdz \left( \alpha^3 \frac{\partial}{\partial z} \right) \\ &= 2 \left( x\alpha^1 + y\alpha^2 + z\alpha^3 \right) = 0. \end{aligned}$$

We can write down solutions for this,

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} y \\ -x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -z \\ y \end{bmatrix}, \begin{bmatrix} z \\ 0 \\ -x \end{bmatrix}.$$

These three vectors are not linearly independent but, as long as $x, y, z$ are not all 0 then we can span the plane $E_x$ by two of these three vectors. These vectors span the tangent spaces of $T_x(M)$. It turns out, because of the topology of $S^2$ there are two non-zero vector fields on the entire surface $S^2$.

These three vectors are written as, We write these three vectors in abstract notation as,

$$y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}, z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}, z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}.$$

The kernel spaces of the smooth differential form $df$ form a integrable distribution. We can find a sub-manifold $N$ (e.g. $N = S^2$ ) that $T_x(N) = kern(df)$. This is the form of Frobenius theorem that we shall discuss.

This simple example shows the basic structure of Frobenius Theorem. A collection of linear independent $1-$forms $\theta^1, \cdots \theta^k$ have joint kernels that form a $m-k$ dimensional sub-space of $T_x(M)$. Can we find a manifold $N$ that has tangent spaces $T_x(N)$ that is the joint kernels of $\theta^1, \cdots, \theta^k$. We call this manifold $N$ the integral manifold of the $1-$forms $\theta^1, \cdots \theta^k$. In our first example we found a manifold, $N = S^2$, which has tangent spaces which are equal to $ker(\theta)$. This is not true for every collection of forms. The $1-$form above is closed, but this is not required, as our next example demonstrates.

**Example 3.1.1.** *Define $\theta = f(x, y, z)dx$ where $f(x, y, z) > 0$ for all $(x, y, z)$. In this case we have*

$$d\theta = \frac{\partial f}{\partial y} dy \wedge dx + \frac{\partial f}{\partial z} dz \wedge dx.$$

*We can write down the kernel,*

$$ker(\theta) = \left\{ a\frac{\partial}{\partial y} + b\frac{\partial}{\partial z} \middle| a, b \in \mathbb{R} \right\}.$$

*There are integral manifolds which consist of planes perpendicular to the $x-$axis.*

## 3.2 Frobenius Theorem

Frobenius Theorem described the conditions required to create a smooth submanifold using the kernels of a collection of $1-$forms. We start with a statement of the theorem.

**Theorem 3.2.1.** *Frobenius Theorem. Let $M$ be a smooth manifold of dimension $m$. Let $\theta^1, \theta^2, \cdots \theta^{m-p} \in \Lambda^1(M)$ be smooth pointwise linearly indepdendent forms. If there exist $1-$forms $\alpha^i_j \in \Lambda^1(M)$ such that,*

$$d\theta^a = \sum_{b=1}^{m-p} \alpha^a_b \wedge \theta^b, \tag{30}$$

*for all $a = 1, \cdots m-p$, then for any $x \in M$ there exists a unique $k-$dimensional manifold*

$$i : N \hookrightarrow M,$$

*such that $x \in N$ and such that $i^*(\theta^a) = 0$ for all $a = 1, \cdots m-k$. Further, for this $x \in M$ there is a coordinate neighborhood $U \subset M$ with coordinates $(x^1, \cdots x^m)$ so that $N$ has coordinates $(x^1, \cdots x^p)$ and*

$$\theta^a = \sum_{b=p+1}^{m} A^a_b dx^b,$$

*so that $\theta^a$ are generated by $dx^a$ for $a = p+1, \cdots m$ and the joint kernel of the $\theta^a$ corresponds to the joint kernel of $dx^a$.*

Recall that $\Lambda^k(M)$ denotes the vector space of smooth $k-$forms on a manifold $M$. Notice that equation 30 is more general than closed, although a collection of closed forms will suffice. Later, we shall see that $d\theta^a$ is in the algebraic ideal generated by the forms $\theta^a$. Before we prove the theorem we prove some preliminary facts. Then we will prove theorem 3.2.1 for the case of $p = 1$.

It is important to note that properties satisfied by the forms $\theta^1, \cdots \theta^{m-p}$ in Theorem 3.2.1 are also satisfied by linear combinations of these forms, as long as there are $m-p$ linearly independent 1−forms. To make this concrete we add the following elementary proposition,

**Proposition 3.2.2.** *Let $\theta^1, \cdots \theta^k \in \Lambda^1(M)$ be 1−forms that satisfy the criteria in Theorem 3.2.1. If $A^a_b$ is a $k \times k$ invertible matrix with smooth entries then the forms $\eta^a$ satisfy the criteria in Theorem 3.2.1 where $\eta^a$ are defined by,*

$$\eta^a = \sum_{b=1}^{k} A^a_b \theta^b \text{ for } a = 1, \cdots k.$$

*Proof.* The forms $\eta^a$ are linearly independent because $A^a_b$ are linearly independent. We take the differential,

$$\begin{aligned} d\eta^a &= d(A^a_b) \wedge \theta^b + A^a_b d\theta^b \\ &= d(A^a_b) \wedge (A^{-1})^b_c \theta^c + A^a_b \alpha^b_d \wedge \theta^d \\ &= d(A^a_b) \wedge (A^{-1})^b_c \eta^c + A^a_b \alpha^b_d \wedge (A^{-1})^d_c \eta^c \\ &= \left( d(A^a_b)(A^{-1})^b_c + A^a_b \alpha^b_d (A^{-1})^d_c \right) \wedge \eta^c \\ &= \beta^a_c \wedge \eta^c \end{aligned}$$

where $\beta^a_c \in \Lambda^1(M)$. □

Another important fact is that the kernels of the collection of differential forms form smooth distributions. In the language of bundles, we will show that the kernels form a sub-bundle of the tangent bundle.

**Theorem 3.2.3.** *Let $M$ be a smooth manifold of dimension $m$. Let $\theta^1, \theta^2, \cdots \theta^{m-p} \in \Lambda^1(M)$ be point-wise linearly independent forms. The set of vectors $X \in T_x(M)$ satisfying $\theta^b(X) = 0$ for $b = 1, \cdots m-p$ form a $m+p$ dimensional differentiable manifold. In fact, it is a $p-$dimensional vector bundle over $M$.*

*Proof.* We set $k = m - p$ so that the forms $\theta^1, \theta^2, \cdots \theta^k$ are linearly independent. Given any $x_0 \in M$ we shall construct a coordinate neighborhood $U$ so that the vectors in the perpendicular space is diffeomorphic to $U \times \mathbb{R}^{m-k}$. Let $(U_1, \varphi)$ be a coordinate neighborhood with coordinates $(y^1, \cdots y^m)$. We can write each $\theta^a$ as,

$$\theta^a = \theta^a_1 dy^1 + \cdots + \theta^a_m dy^m.$$

We write down the matrix, which depends on the value $x \in U$,

$$T(x) = \begin{bmatrix} \theta^1_1 & \theta^1_2 & \cdots & \theta^1_m \\ \theta^2_1 & \theta^2_2 & \cdots & \theta^2_m \\ \vdots & \vdots & \vdots & \vdots \\ \theta^k_1 & \theta^k_2 & \cdots & \theta^k_m \end{bmatrix}$$

For our fixed point $x_0 \in M$ we use Gaussian elimination to change $T(x_0)$. There is a $k \times k$ invertible matrix $R$ so that left-multiplication change $T(x_0)$ into an upper diagonal matrix of the form,

$$\tilde{T}(x_0) = RT(x_0) = \begin{bmatrix} \tilde{\theta}^1_1 & \tilde{\theta}^1_2 & \tilde{\theta}^1_3 & \tilde{\theta}^1_4 & \cdots & \tilde{\theta}^1_{k-1} & \tilde{\theta}^1_k & \cdots & \tilde{\theta}^1_m \\ 0 & \tilde{\theta}^2_2 & \tilde{\theta}^2_3 & \tilde{\theta}^2_4 & \cdots & \tilde{\theta}^2_{k-1} & \tilde{\theta}^2_k & \cdots & \tilde{\theta}^2_m \\ 0 & 0 & \tilde{\theta}^3_3 & \tilde{\theta}^3_4 & \cdots & \tilde{\theta}^3_{k-1} & \tilde{\theta}^3_k & \cdots & \tilde{\theta}^3_m \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \tilde{\theta}^{k-1}_{k-1} & \tilde{\theta}^{k-1}_k & \cdots & \tilde{\theta}^k_m \\ 0 & 0 & 0 & 0 & \cdots & 0 & \tilde{\theta}^k_k & \cdots & \tilde{\theta}^k_m \end{bmatrix}$$

where $\tilde{\theta}^j_j = 1$ for $j = 1, \cdots k$. We use this linear transform $R$ to change coordinate systems as described in Section 1.6.1. We get a new coordinates $(y^1, \cdots y^m)$ so that $\tilde{\theta}$ are coordinates of $\theta^a$ in the $y^j$ coordinates. We get the form above at the point $x_0$. There is a neighborhood $V$ of $x_0 \in V$ such that $\theta^j_j(y) > 0.5$ and all the $0'$s in the sub-diagonal are $< 0.5$. This insures that all the vectors are linearly independent. The new coordinates $(y^1, \cdots y^m)$ are defined on, a possible open subset, of $V$. In these coordinates, given a vector

$$Y = \sum_j Y^j \frac{\partial}{\partial y^j}.$$

This means that $Y \in K_T$ if,

$$\sum_j \tilde{\theta}^a_j Y^j = 0,$$

for all $a$. If $a = 1$ then we have,

$$Y_1 = -\frac{1}{\tilde{\theta}^1_1}\left(\tilde{\theta}^1_2 Y^2 + \cdots + \tilde{\theta}^1_m Y^m\right).$$

We proceed in this way to determine values for $a = 2, 3, \cdots k$,

$$\begin{aligned}
Y_2 &= -\frac{1}{\tilde{\theta}^2_2}\left(\tilde{\theta}^2_1 Y^1 + \tilde{\theta}^2_3 Y^3 + \cdots + \tilde{\theta}^1_m Y^m\right) \\
Y_3 &= -\frac{1}{\tilde{\theta}^3_3}\left(\tilde{\theta}^2_1 Y^1 + \tilde{\theta}^2_2 Y^2 + \tilde{\theta}^3_4 Y^4 + \cdots + \tilde{\theta}^1_m Y^m\right) \\
&\vdots \\
Y_k &= -\frac{1}{\tilde{\theta}^k_k}\left(\tilde{\theta}^2_1 Y^1 + \cdots + \tilde{\theta}^k_{k-1} Y^{k-1} + \tilde{\theta}^k_{k+1} Y^{k+1} + \cdots + \tilde{\theta}^1_m Y^m\right)
\end{aligned}$$

Since $k = m - p$ we have coordinates $y^j$ that form a bundle coordinate system with coordinates, $(y^1, \cdots y^m, Y^{m-p+1}, Y^{m-p+2}, \cdots Y^m)$. All the remaining coordinates $Y^1, \cdots, Y^{m-p}$ are determined. This neighborhood has the topology of $\tilde{U} \times \mathbb{R}^p$. □

**Proposition 3.2.4.** *Let $M$ be a smooth manifold of dimension $m$. Let $\theta^1, \theta^2, \cdots \theta^{m-1} \in \Lambda^1(M)$ be pointwise linearly indepdendent forms. Then for all $x \in M$ there is a coordinate neighborhood $x \in U \subset M$ with coordinates $(x^1, \cdots x^m)$ that satisfy*

$$\theta^a = \sum_{b=2}^m A^a_b dx^b.$$

*Proof.* Given forms $\theta^1, \cdots \theta^{m-1}$ then the perpendicular space is just a line. Theorem 3.2.3 says that this is spanned by a smooth vector field. This smooth vector field has integral curves which are then the integral submanifolds of the distribution.

To get the coordinates $x^j$ we use the Flow box coordinates that we discussed in Proposition 2.1.4. Let $X$ be a smooth vector field that is in the kernel of the $\theta^a$. We pick flow box coordinates with $\frac{\partial}{\partial x^1}$ in the direction of $X$ so it is in the kernel of $\theta^a$. This means that,

$$\theta^a(\frac{\partial}{\partial x^1}) = 0, \text{ for } a = 1, \cdots m - 1.$$

Since $dx^1, \cdots dx^m$ span $T^*(M)$ we have,

$$\theta^a = \sum_{j=1}^{m} A_j^a dx^j = \sum_{b=2}^{m} A_b^a dx^b,$$

where $A_b^a$ is an $(m-1)\times(m-1)$ dimensional invertible matrix and $A_b^a$ is a smooth function of $M$. □

**Proposition 3.2.5.** *Frobenius theorem 3.2.1 is true for any $p$.*

*Proof.* We shall prove the theorem by induction on $p$. The case for $p = 1$ was proved in Proposition 3.2.4. Now assume the theorem is true for $p-1$ we will show it is true for $p$. Let $\theta^1, \cdots \theta^{m-p}$ be linearly independent 1−forms that satisfy,

$$d\theta^a = \sum_{b=1}^{m-p} \alpha_b^a \wedge \theta^b, \text{ for } a = 1, \cdots m-p,$$

where $\alpha_b^a \in \Lambda^1(M)$. We will add another form to this collection and show that the new expanded collection still satisfies the condition of equation 30. Since the theorem is true for $p-1$ this addition collection has $m-(p-1) = m-p+1$ forms and the induction hypothesis is satisfied. We start by adding a 1−form to our collection of one forms.

**Claim 3.2.6.** *If $p > 0$ then there is a function $f : M \to \mathbb{R}$ with $df, \theta^1, \cdots \theta^{m-p}$ are linearly independent.*

Let $U \subset M$ be a coordinate neighborhood with coordinates $(y^1, \cdots y^m)$. Since the forms $\theta^1, \cdots \theta^{m-p}$ cannot span the entire span of forms $dy^1, \cdots dy^{m-p+1}$ there must be some for $dy^k$ that is linearly independent of the $\theta^a$. This means that for every $x \in U$ there is a $k$ so that one can set $f = y^k$.

With $f$ defined as in the claim we have a collection of $m-p+1 = m-(p-1)$ forms given by,

$$\theta^1, \cdots \theta^{m-p}, df. \tag{31}$$

We are assuming the theorem is true for $p-1$ so we must show the forms 31 satisfy the criteria of the theorem. These forms clearly satisfy condition 30 since $d(df) = 0$ and, by assumption,

$$d\theta^a = \sum_{b=1}^{m-p} \alpha_b^a \wedge \theta^b,$$

for $a = 1, \cdots m-1$. Therefore, by Frobenius theorem for $p-1$ we know that for any $x_0 \in M$ there is a neighborhood $V$ and coordinates $(y^1, \cdots y^m)$ that satisfy $x_0 \in V$ and for all for all $y \in V$,

$$\theta^a(y) = \sum_{b=p}^{m} B_b^a(y) dy^b \text{ for } a = 1, \cdots m-p,$$
$$df(y) = \sum_{b=p}^{m} C_b(y) dy^b$$

Notice that the vector fields $\frac{\partial}{\partial y^1} \cdots \frac{\partial}{\partial y^{p-1}}$ are in

$$ker(\theta^1) \cap \cdots ker(\theta^{m-p}) \leq ker(\theta^1) \cap \cdots \cap ker(\theta^{m-p}) \cap ker(df).$$

To prove the theorem we need to add an additional coordinate to this kernel space.

The forms $\theta^a$ and $df$ are non-zero so $C_b(x_0) \neq 0$ for at least one $b$ and we can find, a possible sub, neighborhood of $V$ such that $C_b(x) \neq 0$ for all $x \in V$. We see that we can assume, without a loss of generality, that $C_p(x) \neq 0$ for all $x \in V$. We can then write,

$$dy^p = \frac{1}{C_p} df - \left(\frac{1}{C_p}\right) \left( \sum_{b=p+1}^{m} C_b dy^b \right)$$

We use this formula to eliminate the form $dy^p$ from our collection and replace it with the form $df$. We get the following set of equations,

$$\theta^a = \sum_{b=p+1}^{m} D_b^a dy^b + d^a df \text{ for } a = 1, \cdots m-p,$$
$$df = df$$

where $d^a$ are smooth functions of $C_p$ and $B_p^a$ and $x$. Notice this is the first step in Gaussian elimination. The matrix elements $D_b^a$ are functions of the remaining $C_b$ and $B_b^a$. Notice that $D_b^a$ is now a $(m-p) \times (m-p)$ matrix.

The $m-p+1$ forms $df, dy^{p+1}, \cdots, dy^m$ span the same sub-space of $T^*(M)$ as $\theta^1, \cdots \theta^{m-p}, df$. We write this in matrix form as,

$$\begin{bmatrix} \theta^a \\ df \end{bmatrix} = \begin{bmatrix} D_b^a & d^a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dy^b \\ df \end{bmatrix}. \tag{32}$$

This relationship is invertible so $det(D^a_b(x)) \neq 0$ for all $x \in V$. We can write down the inverse and using this inverse we multiply both sides of 32,

$$\begin{bmatrix} dy^a \\ df \end{bmatrix} = \begin{bmatrix} (D^{-1})^a_b & -(D^{-1})^a_b d^b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta^b \\ df \end{bmatrix} = \begin{bmatrix} (D^{-1})^a_b \theta^a - \left((D^{-1})^a_b d^b\right) df \\ df \end{bmatrix} = \begin{bmatrix} \eta^a + e^a df \\ df \end{bmatrix}$$

We have defined $\eta^a$ and $e^a$ by,

$$\eta^a(y) = (D^{-1})^a_b(y)\theta^b \text{ and } e^a(y) = (D^{-1})^a_b d^b.$$

Using these definitions we have,

$$\begin{bmatrix} dy^a \\ df \end{bmatrix} = \begin{bmatrix} \eta^a + e^a df \\ df \end{bmatrix}$$

There is an $m - p + 1$ dimensional subspace of $T^*(M)$ that has a basis given by the forms $\{df, dy^{p+1}, \cdots, dy^m\}$. There are also basis given by $\{df, \theta^1, \cdots \theta^{m-p}\}$ and $\{\eta^{p+1}, \cdots \eta^m, df\}$.

We can span the sub-space of $T^*(M)$ using the basis $\{df, dy^a\}$ or the basis $\{df, \eta^a\}$. From the matrix equation we get,

$$\eta^a = dy^a - e^a df, \text{ where } a = p + 1, \cdots m. \tag{33}$$

**Claim 3.2.7.** *The 1−forms $\eta^a$ satisfy equation 30 in that,*

$$d\eta^a = \sum_{b=p+1}^{m} E^a_b \eta^b \textit{ for } a = p + 1, \cdots m.$$

This is simply Theorem 3.2.2.

**Claim 3.2.8.** *The functions $e^a(y)$ are function of only $(y^p, \cdots, y^m)$.*

From the computation in Claim 3.2.7 we have the following,

$$d\eta^a = \sum_{b=p+1}^{m} E^a_b \wedge \eta^b,$$

But we also have equation 33,

$$\eta^a = dy^a - e^a df, \text{ where } a = p + 1, \cdots m.$$

We take the differential $d$ of this equation and write everything in terms of the basis $\{\eta^a, df\}$ and equate this with above,

$$d\eta^a = -de^a \wedge df = \gamma^a_b \wedge \eta^b$$

The means that all the terms in $de^a \wedge df$ must contain a term of the form $\beta \wedge \eta^a$. From this we deduce that (note that $df \wedge df = 0$),

$$de^a = \sum_{b=p+1}^{m} C^a_b \eta^b + g^a df, \text{ where } a = p+1, \cdots, m.$$

We can now write this in terms of the $dy^a$ and we have,

$$de^a = \sum_b F^a_b dy^b = \sum_{b=1}^{m} \frac{\partial e^a}{\partial y^b} dy^b = \sum_{b=p+1}^{m} \frac{\partial e^a}{\partial y^b} dy^b, \text{ for } a = p+1, \cdots m.$$

This means that,

$$\frac{\partial e^a}{\partial y^j} = 0, \text{ for } a = p+1, \cdots m, j = 1, \cdots p-1.$$

This concludes claim 3.2.8.

The forms $\eta^a$ satisfy Proposition 3.2.5 on the $m-p+1$ dimensional manifold generated by the coordinates $(y^p, \cdots y^m)$. Using this proposition we can find coordinates $(x^p, \cdots x^m)$ so that,

$$\eta^a = \sum_{b=p+1}^{m} \eta^a_b dx^b \text{ for } a = p+1, \cdots, m.$$

We also have $\frac{\partial}{\partial x^p} \in ker(\eta^{p+1}) \cap \cdots \cap ker(\eta^m)$. Notice that these expressions are true for all $(y^1, \cdots, y^{p-1})$. We define a coordinate system by $x^j = y^j$ for $j = 1, \cdots p$ and $x^{p+1} = y^{p+1}, \cdots, x^m = y^m$. The forms $\eta^{p+1}, \cdots, \eta^m$ are linear combinations of the forms $\theta^1, \cdots, \theta^{m-p}$. This means that our original forms are linear combinations of the $dx^{p+1}, \cdots dx^m$,

$$\theta^a = \sum_{b=p+1}^{m} \theta^a_b dx^b \text{ for } a = p+1, \cdots m$$

The transverse manifold has coordinates $(x^1, \cdots x^p)$ and is determined by the conditions $x^{p+1} = K^{p+1}, \cdots, x^m = K^m$. Thus we have proved the theorem for $p$ assuming $p-1$ and this concludes the theorem. □

**Remark 3.2.9.** *We have two separate proofs of Frobenius theorem. The first uses vector fields and the second uses differential forms. One can also prove the two forms of Frobenius theorem are equivalent using a generalization of the formula in equation 6.*

**Example 3.2.10.** *Frobenius theorem has a close relationship to the implicit function theorem 1.7.2. Let* $G : M \to \mathbb{R}^k$ *so that,*

$$G(p) = \begin{bmatrix} g^1(p) \\ g^2(p) \\ \vdots \\ g^{m-k}(p) \end{bmatrix}$$

*We form the closed forms* $dg^j$ *for* $j = 1, \cdots m-k$*. The intersection of the kernel planes is,*

$$E_p = ker(dg^1) \cap ker(dg^2) \cap \cdots \cap ker(dg^{m-k})$$

*The conditions Frobenius are that* $dg^j$ *are linearly indepdendent which is the same as the rank condition of the implicit function theorem. We look at the Jacobian,*

$$Dg(p) = \begin{bmatrix} \frac{\partial g^1}{\partial x^1} & \frac{\partial g^1}{\partial x^2} & \cdots & \frac{\partial g^1}{\partial x^m} \\ \frac{\partial g^2}{\partial x^1} & \frac{\partial g^2}{\partial x^2} & \cdots & \frac{\partial g^2}{\partial x^m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial g^{m-k}}{\partial x^1} & \frac{\partial g^{m-k}}{\partial x^2} & \cdots & \frac{\partial g^{m-k}}{\partial x^m} \end{bmatrix}$$

*If the* $dg^j$ *are linearly independent then the Jacobian* $Dg$ *has rank* $m-k$*. The other condition is trivial since the forms* $dg^j$ *are closed as* $d\left(dg^j\right) = 0$*. Frobenius then says that there is a sub-manifold* $i : N \to M$ *such that* $i^*(dg^j) = 0$ *for all* $j$ *so that,*

$$dg^j\left(di(X)\right) = dg^j\left(X\right) = X(g^j) = 0,$$

*for all* $j = 1, \cdots m-k$ *so that* $X \in ker(dg^j)$ *for all* $j$*. This means that for all* $z \in N$ *we have*

$$T_z(N) \leq E_p.$$

*We see that the implicit function theorem is really a special case of Frobenius Theorem.*

**Example 3.2.11.** *There are systems that are not Frobenius. Let* $M = \mathbb{R}^3$ *and define a* $1-$*form* $\theta$ *by*

$$\theta = dz - ydx + xdy.$$

*We compute the differential,*

$$d\theta = -dy \wedge dx + dx \wedge dy = 2dx \wedge dy.$$

*We see that* $d\theta \neq \eta \wedge \theta$ *so the distribution is not integrable. In fact, there are no sub-manifolds of* $\mathbb{R}^3$ *with these planes as tangent planes.*

**Example 3.2.12.** *Let $M = \mathbb{R}^3$ and define a $1-$form by,*

$$\theta = \theta_1 dx + \theta_2 dy + \theta_3 dz. \tag{34}$$

*We want to find general conditions where $\theta$ satisfies Frobenius Theorem 3.2.1. The kernels of this $1-$form are given by vectors*

$$X = X^1 \frac{\partial}{\partial x} + X^2 \frac{\partial}{\partial y} + X^3 \frac{\partial}{\partial z},$$

*that have the property that $\theta_i X^i = 0$. Is this system of planes integrable? We compute $d\theta$.*

$$d\theta = \left(\frac{\partial \theta_2}{\partial x} - \frac{\partial \theta_1}{\partial y}\right) dx \wedge dy + \left(\frac{\partial \theta_1}{\partial z} - \frac{\partial \theta_3}{\partial x}\right) dz \wedge dx + \left(\frac{\partial \theta_3}{\partial y} - \frac{\partial \theta_2}{\partial z}\right) dy \wedge dz.$$

*When is $d\theta = \alpha \wedge \theta$ for some $\alpha \in \Lambda^1(M)$. If we extend $\theta$ to a full basis of $T^*(M)$ using forms $\omega_2, \omega_3$ then we can write,*

$$d\theta = p_1(\omega_1 \wedge \omega_2) + p_2(\omega_3 \wedge \theta) + p_3(\theta \wedge \omega_3).$$

*Now we see that the condition that $d\theta = \alpha \wedge \theta$ is equivalent to the condition that $d\theta \wedge \theta = 0$. So the distribution of equation 34 is integrable if and only if*

$$d\theta \wedge \theta = \left(\frac{\partial \theta_2}{\partial x} - \frac{\partial \theta_1}{\partial y}\right) \theta_3 + \left(\frac{\partial \theta_1}{\partial z} - \frac{\partial \theta_3}{\partial x}\right) \theta_2 + \left(\frac{\partial \theta_3}{\partial y} - \frac{\partial \theta_2}{\partial z}\right) \theta_1 = 0.$$

*If the form $\theta$ satisfied Frobenius Theorem (3.2.1 ) then we could, locally, find coordinates $(x^1, x^2, x^3)$ so that,*

$$\theta = A(x) dx^3.$$

*In the classic literature this condition is written, $\boldsymbol{V}_\theta \cdot curl(\boldsymbol{V}_\theta) = 0$.*

*Our example is immediately generalized to 4 or even higher dimension, although the notation gets more complicated since the space of $2-$forms in $\mathbb{R}^4$ is 6 dimensional. We could easily derive a condition so that $3-$planes in $\mathbb{R}^4$ are integrable.*

## 3.3 Differential Forms

The set of alternating forms is a vector space and the product operator $\wedge$ turns the alternating forms into a Graded algebra. We can formally add and multiply forms by real number. A form $\theta$ has order $p$ if $\theta \in \Lambda^p$.

**Definition 3.3.1.** *A graded module $A$ is a module along with a grading that assigns sub-modules $A_k$ to grade $k$. The module $A$ a direct sum of the graded sub-modules,*

$$A = \bigoplus_{n \in N} A_n = A_0 \oplus A_1 \oplus \cdots$$

*A graded module $A$ is a graded algebra if there is a product and the product takes $A_i A_j \subset A_{i+j}$.*

**Definition 3.3.2.** *A sub-space $\mathcal{I} \subset \Omega^*(M)$ is an algebraic ideal if*

1. *$\mathcal{I}$ is a direct sum of homogeneous sub-spaces $\mathcal{I}^k \subset \Omega^k(M)$.*
2. *$\mathcal{I}$ is an ideal under $\wedge$ so that for any $\omega \in \mathcal{I}$ and $\eta \in \Omega^*(M)$ we have $\omega \wedge \eta \in \mathcal{I}$.*

This is called a graded ideal of the graded algebra $\Omega^*(M)$.

**Example 3.3.3.** *Let $\phi^1, \cdots, \phi^s$ is a collection of homogeneous elements. The algebraic ideal generated by these forms is*

$$\langle \phi^1, \cdots, \phi^s \rangle_{alg} = \left\{ \sum_{k=1}^{s} (\gamma^i \wedge \phi^i) \middle| \gamma^1, \cdots, \gamma^s \in \Omega^*(M) \right\}$$

All of the above works for alternating forms on a vector space. To add derivatives we add the differential of a form to the mix. We start with the basic definition.

**Definition 3.3.4.** *An algebraic ideal $\mathcal{I}$ is a differential ideal if $\mathcal{I}$ is closed under exterior product, so that $d\omega \in \mathcal{I}$ for every $\omega \in \mathcal{I}$.*

**Definition 3.3.5.** *Let $M$ be a manifold $M$ and let $\mathcal{D}$ be a distribution. The annihilator of the distribution is the set of forms $\theta \in \Lambda^p(M)$ that have the property that for any $X_1, \cdots X_p$ vector fields that are contained in the distribution $\mathcal{D}$ then*

$$\theta(X_1, X_2, \cdots X_p) = 0.$$

*The forms in the annihilator will annihilate the vectors fields in the distribution.*

Notice that if $\mathcal{D}$ is a distribution then the collection of annihilators form an ideal in $\Lambda^*(M)$. This follows since if $\theta \in \Lambda^p(M)$ annihilates a distribution $\mathcal{D}$ and if $\omega \in \Lambda^k(M)$ then for any $X_1, \cdots X_{p+k} \in \mathcal{D}$ we have,

$$(\omega \wedge \theta)(X_1, \cdots, X_{p+k}) = 0.$$

Frobenius theorem can be stated in our new language. In this language Frobenius theorem is more closely linked to Cartan-Kahler theorem.

**Theorem 3.3.6.** *Let $M$ be a $C^\infty$ differentiable manifold of dimension $m$. Let $\mathcal{I}$ be an ideal generated algebraically by linearly independent $1-$forms $\theta^{m-p+1}, \cdots, \theta^m$. If $\mathcal{I}$ is a differential ideal then through any $x \in M$ there is a integral manifold of $\mathcal{I}$. Further, for any $x \in M$, there is a neighborhood $U$ with $x \in U \subset M$ with coordinates $(y^1, \cdots y^N)$ so that $\mathcal{I}$ is generated by $dx^{m-p+1}, \cdots dx^m$. The coordinates $x^1, \cdots x^p$ are coordinates to a sub-manifold $N$ that any $Y \in T_x(N)$ has $\theta^a(X) = 0$ for all $a = m-p+1, \cdots m$.*

**Example 3.3.7.** *Let $M = \mathbb{R}^3$. In this example we model a PDE system where $u : M = \mathbb{R}^3 \to \mathbb{R}$ and the equations are given by,*

$$\frac{\partial u}{\partial x} = F(x,y,z) \tag{35}$$

$$\frac{\partial u}{\partial y} = G(x,y,z) \tag{36}$$

*If $\gamma : \mathbb{R} \to M$ is a curve that lies in the solution to the system surface then it has form*

$$\begin{aligned}\gamma(t) = \begin{bmatrix} x_0 + at \\ y_0 + bt \\ u(x_0 + at, y_0 + bt) \end{bmatrix} \ \textit{with} \ \frac{d\gamma}{dt}(t) = \begin{bmatrix} a \\ b \\ au_x + bu_y \end{bmatrix} \\ = a\left(\frac{\partial}{\partial x} + F\frac{\partial}{\partial z}\right) + b\left(\frac{\partial}{\partial y} + G\frac{\partial}{\partial z}\right)\end{aligned}$$

*This two dimensional space is the kernel of the following form,*

$$\theta = dz - F(x,y,z)dx - G(x,y,z)dy. \tag{37}$$

*We want to find the differential ideal generated by this single form. We take the differential of $\theta$,*

$$\begin{aligned} d\theta &= -\frac{\partial F}{\partial y}dy \wedge dx - \frac{\partial F}{\partial z}dz \wedge dx - \frac{\partial G}{\partial x}dx \wedge dy - \frac{\partial G}{\partial z}dz \wedge dy \\ &= \left(\frac{\partial F}{\partial y} - \frac{\partial G}{\partial x}\right)dx \wedge dy - \frac{\partial F}{\partial z}dz \wedge dx - \frac{\partial G}{\partial z}dz \wedge dy \end{aligned}$$

*Let's use $\theta$ in equation 37 to eliminate $dz = \theta + Fdx + Gdy$. We get,*

$$\begin{aligned} d\theta &= \left(\frac{\partial F}{\partial y} - \frac{\partial G}{\partial x}\right)dx \wedge dy - \frac{\partial F}{\partial z}(\theta + Fdx + Gdy) \wedge dx - \frac{\partial G}{\partial z}(\theta + Fdx + Gdy) \wedge dy \\ &= \left(\frac{\partial F}{\partial y} - \frac{\partial G}{\partial x}\right)dx \wedge dy - \frac{\partial F}{\partial z}(\theta + Gdy) \wedge dx - \frac{\partial G}{\partial z}(\theta + Fdx+) \wedge dy \\ &= \left(\frac{\partial F}{\partial y} - \frac{\partial G}{\partial x}\right)dx \wedge dy + \left(G\frac{\partial F}{\partial z} - F\frac{\partial G}{\partial z}\right)dx \wedge dy - \theta \wedge \left(\frac{\partial F}{\partial z}dx + \frac{\partial G}{\partial z}dy\right) \end{aligned}$$

*We would like to use Frobenius theorem to find solutions of the system. For $\theta$ to satisfy the theorem $d\theta$ must be in the algebraic ideal generated by $\theta$. This means that all the $dx \wedge dy$ terms must vanish. We arrive at the compatibility condition,*

$$\frac{\partial F}{\partial y} - \frac{\partial G}{\partial x} + G\frac{\partial F}{\partial z} - F\frac{\partial G}{\partial z} = 0. \tag{38}$$

*This equation is not unexpected. Using equations 35 and 36 we have,*

$$\begin{aligned} u_{xy} &= \frac{\partial F}{\partial y} + \frac{\partial F}{\partial z}u_y = \frac{\partial F}{\partial y} + G\frac{\partial F}{\partial z} \\ u_{yx} &= \frac{\partial G}{\partial y} + \frac{\partial G}{\partial z}u_x = \frac{\partial G}{\partial y} + F\frac{\partial F}{\partial z} \end{aligned}$$

*The condition that $u_{xy} = y_{xy}$ is equivalent to equation 38.*

**Example 3.3.8.** *Let $M = \mathbb{R}^3$ and define a PDE system by,*

$$\begin{aligned} \frac{\partial y}{\partial x} &= F(x, y, z) \\ \frac{\partial z}{\partial x} &= G(x, y, z) \end{aligned}$$

*A solution to this system is a curve*

$$\gamma(t) = \begin{bmatrix} at \\ y(at) \\ z(at) \end{bmatrix} \text{ so that } \frac{d\gamma}{dt} = a\left(\frac{\partial}{\partial x} + F\frac{\partial}{\partial y} + G\frac{\partial}{\partial z}\right)$$

*This tangent vector is the kernel of the two forms,*

$$\begin{aligned} \theta^1 &= dy - F(x, y, z)dx \\ \theta^2 &= dz - G(x, y, z)dx \end{aligned}$$

*We compute,*

$$\begin{aligned} d\theta^1 &= -\frac{\partial F}{\partial y}dy \wedge dx - \frac{\partial F}{\partial z}dz \wedge dx \\ &= -\frac{\partial F}{\partial y}\left(\theta^1 + Fdx\right) \wedge dx - \frac{\partial F}{\partial z}\left(\theta^2 + Gdx\right) \wedge dx \\ &= -\frac{\partial F}{\partial y}\theta^1 \wedge dx - \frac{\partial F}{\partial z}\theta^2 \wedge dx \end{aligned}$$

$$
\begin{aligned}
d\theta^2 &= -\frac{\partial G}{\partial y} dy \wedge dx - \frac{\partial G}{\partial z} dz \wedge dx \\
&= -\frac{\partial G}{\partial y} \left(\theta^1 + F dx\right) \wedge dx - \frac{\partial G}{\partial z} \left(\theta^2 + G dx\right) \wedge dx \\
&= -\frac{\partial G}{\partial y} \theta^1 \wedge dx - \frac{\partial G}{\partial z} \theta^2 \wedge dx
\end{aligned}
$$

*Both* $d\theta^1$ *and* $d\theta^2$ *bot satisfy the criteria of Frobenius theorem. This means that there are integral submanifolds. This example is trivial since we know the smooth vector field,*

$$
X(x, y, z) = \frac{\partial}{\partial x} + F \frac{\partial}{\partial y} + G \frac{\partial}{\partial z},
$$

*has integral curves which are solutions to the system.*

For more details see [2] and [1].

# 4 References

## References

[1] R.L. Bryant S.S. Chern R.B. Gardner H.L. Goldschmidt and P.A. Griffiths. *Exterior Differential Systems*. Springer-Verlag, 1991.

[2] Thomas Ivey and J.M. Landsberg. *Cartan for Beginners*. American Mathematical Society, 2003.

[3] John McCleary. *Geometry From a Differentiable Viewpoint*. Cambridge Univerity Press, 1994.

[4] Michael Spivak. *Calculus on Manifolds*. W. A. Benjamin, Inc, 1965.

[5] Shlomo Sternberg. *Lectures on Differential Geometry*. Prentice Hall, Inc, 1964.

[6] Frank Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Scott Foresman and Company, 1970.

# Index