FROBENIUS' THEOREM

BEN A. SHERWOOD

ABSTRACT. Given a manifold M of dimension $n+k$, attach to every $p \in M$ an n-dimensional subspace of the tangent space $T_p(M)$. It is natural to ask if this collection of subspaces is the collection of tangent spaces to a family of submanifolds that cover M. In this paper we prove Frobenius' Theorem, which gives a necessary and sufficient condition for the answer to be yes.

1. MOTIVATION

Consider S^2, the unit 2-sphere in $\mathbb{R}^3$. At each point $x \in S^2$, there is a tangent plane P_x. The plane P_x is a 2-dimensional vector space with x as its zero vector 0_x, and vectors $v_x \in P_x$ are arrows in $\mathbb{R}^3$ with their tail at x. For $v_x \in P_x$, let $\tau(v_x)$ be its parallel translate having its tail at the origin in $\mathbb{R}^3$. Letting TS^2 denote the set of all tangent vectors v_x as x ranges over S^2, the mapping $TS^2 \to S^2 \times \mathbb{R}^3$ given by $v_x \mapsto (x, \tau(v_x))$ is an injection and allows us to topologize TS^2 as a subspace of $S^2 \times \mathbb{R}^3$. Thus TS^2 is a topological space that, as a set, consists of the disjoint union of all the vector spaces P_x for $x \in S^2$.

The same construction applies to n-spheres of arbitrary positive radius. Let S^n_r denote the n-sphere in $\mathbb{R}^{n+1}$ with radius $r > 0$. For each S^n_r, we again construct TS^n_r, a topological space consisting of the disjoint union of all the vector spaces P_x for $x \in S^n_r$. Fixing n, we form the union

$$\Delta = \bigcup_{r \in \mathbb{R},\, r>0} TS^n_r.$$

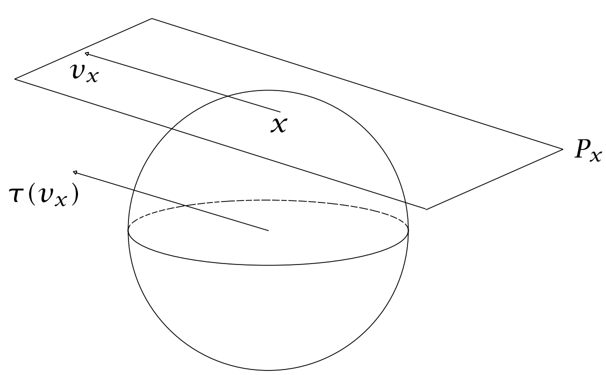

FIGURE 1. P_x is the tangent plane to $x \in S^2$.

We want to think of Δ as a distribution of planes in $\mathbb{R}^{n+1} - \{\mathbf{0}\}$. It is easiest to imagine such a distribution in $\mathbb{R}^3$: picture covering S^2 with tangent planes and then filling out the rest of $\mathbb{R}^3$ with the tangent planes to S^2_r for every $r > 0$.

Conversely, it is not difficult to realize Δ as the collection of tangent spaces to the spheres S^2_r in $\mathbb{R}^3$ for every $r > 0$. Rather than spell out how one might do this, we consider the analogous problem for a different distribution of planes. Construct a new distribution Δ on $\mathbb{R}^3$ by attaching to each $(x, y, z) \in \mathbb{R}^3$ the linear span of $(1, 0, f_x)$ and $(0, 1, f_y)$, where f_x and f_y denote the directional derivatives of a C^1 function $f : \mathbb{R}^2 \to \mathbb{R}$. The question is whether or not Δ is the collection of tangent spaces to a surface, or class of surfaces, in $\mathbb{R}^3$.

We claim Δ is the collection of tangent spaces to the class of surfaces

$$G_r = \{(x, y, f(x, y) + r) : (x, y) \in \mathbb{R}^2\}$$

indexed by $r \in \mathbb{R}$, that is the graph of $f(x, y) + r$. To verify this, we simply construct the tangent space at a point of G_r for each r. To do this, however, we better understand what this space really is.

There is one case where the tangent space to a point is easy, $\mathbb{R}^n$. In this case, we may identify $T\mathbb{R}^n$ with directional derivatives in $\mathbb{R}^n$, so that $T\mathbb{R}^n$ is simply the linear span of $\partial/\partial x_i$ for $i = 1, \ldots, n$. Returning to the problem of constructing TG_r, we can use what we know about $T\mathbb{R}^n$ to calculate a basis for TG_r. In particular, applying the operators $\partial/\partial x$ and $\partial/\partial y$ to $\{(x, y, f(x, y) + r)\}$ shows that $\{(1, 0, f_x), (0, 1, f_y)\}$ is a basis for TG_r for every r. Thus Δ really is the collection of tangent spaces to the class of surfaces given by G_r for each $r \in \mathbb{R}$.

It is certainly not true that every distribution of planes in $\mathbb{R}^3$ is the collection of tangent spaces to the graph of a function $z = f(x, y)$. Consider the distribution Δ given by attaching to each $(x, y, z) \in \mathbb{R}^3$ the linear span of $(1, 0, 0)$ and $(0, 1, x)$. If Δ were the collection of tangent spaces to the graph of a function, then the net change in height upon traversing any closed path in the xy-plane while always moving tangent to Δ would be zero. However this is not the case, as it is straightforward to verify that traversing any rectangle in the xy-plane with side lengths r and s while always moving tangent to Δ yields a net change in height of rs.

Dropping the condition that our distribution Δ be the collection of tangent spaces to the graph of a function, we may ask the more general question of whether or not Δ is the collection of tangent spaces to an arbitrary manifold. While the above examples were chosen so that we could verify easily whether or not they were the collection of tangent spaces to a manifold, in general this question is difficult to answer.

In this paper, we prove a theorem that answers this question. The theorem is known as Frobenius' Theorem, and it says

Theorem 1.0.1. (Frobenius' Theorem) *A distribution Δ on a manifold M is completely integrable if and only if it is involutive.*

Of course, Theorem 1.0.1 looks nothing like what we have been talking about, and the reason is because until now we have only spoken informally about distributions, tangent spaces, and submanifolds. Let us begin, then, by making rigorous the notions underlying the motivation for and statement of Frobenius' Theorem. The most important of these notions is the tangent space at a point of a manifold, and this is where we begin.

2. The Tangent Space at a Point of a Smooth Manifold

We have already mentioned that the tangent space to $\mathbb{R}^n$ is the vector space spanned by $\partial/\partial x_i$ for $i = 1, \dots, n$. In this section, we define the tangent space to an arbitrary manifold M at a point $p \in M$ and then show that we may again identify it with directional derivatives.

Let M be a C^∞ manifold of dimension n. For $p \in M$, denote by $C^\infty(p)$ the algebra of germs of C^∞ functions whose domain of definition includes a neighborhood of p. Recall that a *germ* of a C^∞ function at p is an equivalence class of C^∞ functions whose domain of definition includes a neighborhood of p. The equivalence relation is $f \sim g$ if $f \equiv g$ on any neighborhood of p. $C^\infty(p)$ is a unital algebra over the real numbers and, for a coordinate neighborhood $\{U, \varphi\}$ of p, the map

$$\varphi^* : C^\infty(\varphi(p)) \to C^\infty(p)$$

given by $\varphi^*(f) = f \circ \varphi$ is an isomorphism of the algebra $C^\infty(\varphi(p))$ onto the algebra $C^\infty(p)$. We study the dual vector space homomorphism

$$\varphi_* : C^\infty(p)^* \to C^\infty(\varphi(p))^*$$

when restricted to a certain subspace of $C^\infty(p)^*$, the tangent space.

Definition 2.0.2. The *tangent space* $T_p(M)$ to M at p is the set of all mappings $X_p : C^\infty(p) \to \mathbb{R}$ such that for all $\alpha, \beta \in \mathbb{R}$ and $f, g \in C^\infty(p)$

(1) $X_p(\alpha f + \beta g) = \alpha(X_p f) + \beta(X_p g)$
(2) $X_p(fg) = (X_p f)g(p) + f(p)(X_p g)$ (Leibniz rule)

with the vector space operations in $T_p(M)$ defined by

$$(X_p + Y_p)f = X_p f + Y_p f,$$

$$(\alpha X_p)f = \alpha(X_p f).$$

A *tangent vector to M at p* is any $X_p \in T_p(M)$.

The simplest example of a tangent space is the tangent space $T_p(\mathbb{R}^n)$ to $\mathbb{R}^n$ at a point $p \in \mathbb{R}^n$, however the way it is usually defined looks vastly different from Definition 2.0.2. The usual way $T_p(\mathbb{R}^n)$ is defined goes as follows: write $p = (p_1, \dots, p_n)$ and let $T_p(\mathbb{R}^n)$ consist of all pairs of points $X_p = (p, x)$ corresponding to initial and terminal points of a segment. We give $T_p(\mathbb{R}^n)$ a vector space structure in the following way. Consider the map $f_p : T_p(\mathbb{R}^n) \to V^n$, where V^n denotes the vector space of n-tuples of real numbers, given by $f_p(X_p) = (x_1 - p_1, \ldots, x_n - p_n)$. Define the operations $X_p + Y_p = f_p^{-1}(f(X_p) + f(Y_p))$ and $\alpha X_p = f_p^{-1}(\alpha f_p(X_p))$, where $\alpha \in \mathbb{R}$, on $T_p(\mathbb{R}^n)$. It is straightforward to verify that $T_p(\mathbb{R}^n)$ is a real vector space under these operations and that $f_p : T_p(\mathbb{R}^n) \to V^n$ an isomorphism of vector spaces. Therefore we may give $T_p(\mathbb{R}^n)$ a canonical basis given by $E_{1p} = f_p^{-1}(\mathbf{e}_1), \ldots, E_{np} = f_p^{-1}(\mathbf{e}_n)$, where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ denotes the canonical basis of V^n.

The above definition of $T_p(\mathbb{R}^n)$ looks quite different from that of Definition 2.0.2, however they are actually the same. Recall that for $M = \mathbb{R}^n$ and $p \in M$, any such map $X_p^* : C^\infty(p) \to \mathbb{R}$ satisfying the conditions of Definition 2.0.2 is called a *derivation* at $p \in \mathbb{R}^n$. In particular, the collection $\mathcal{D}(p)$ of derivations at p forms a real vector space with basis $\partial/\partial x_1, \ldots, \partial/\partial x_n$, the directional derivatives evaluated at p. It turns out that $\mathcal{D}(p)$ and $T_p(\mathbb{R}^n)$ as defined in the previous paragraph are isomorphic as vector spaces. Letting

$$X_p = \sum_{i=1}^{n} \alpha_i E_{ip}$$

be an element of $T_p(\mathbb{R}^n)$ and

$$X_p^* f = \sum_{i=1}^{n} \alpha_i \left(\frac{\partial f}{\partial x_i} \right)_p,$$

where $f \in C^\infty(p)$, the map $\mu : T_p(\mathbb{R}^n) \to \mathcal{D}(p)$ given by $\mu(X_p) = X_p^*$ is the desired isomorphism. It is easy to see that μ is an injective homomorphism, however it is non-trivial that μ is a surjection. For a proof that μ is surjective see [1].

Thus the map μ allows us to identify $T_p(\mathbb{R}^n)$ with $\mathcal{D}(p)$, and it is under this identification that the basis vectors $E_{1p}, \ldots, E_{np}$ of $T_p(\mathbb{R}^n)$ are identified with $\partial/\partial x_1, \ldots, \partial/\partial x_n$, the directional derivatives evaluated at p. In particular, our initial definiton of $T_p(\mathbb{R}^n)$ agrees with Definition 2.0.2.

Theorem 2.0.3. *Let $F : M \to N$ be a C^∞ map of manifolds. For $p \in M$, the map $F^* : C^\infty(F(p)) \to C^\infty(p)$ given by $F^*(f) = f \circ F$ is a homomorphism of algebras and induces a dual vector space homomorphism $F_* : T_p(M) \to T_{F(p)}(N)$ defined by $F_*(X_p)f = X_p(F^* f)$, which gives $F_*(X_p)$ as a map $C^\infty(F(p)) \to \mathbb{R}$ while $F_*(T_p(M)) \subseteq T_{F(p)}(N)$. If $F : M \to M$ is the identity, then both F^* and F_* are the identity isomorphism. Furthermore, if $H = G \circ F$ is a composition of C^∞ maps, then $H^* = F^* \circ G^*$ and $H_* = G_* \circ F_*$.*

Proof. The proof is a straightforward verification of definitions and is therefore omitted. □

The homomorphism $F_* : T_p(M) \to T_{F(p)}(M)$ is called the *differential* of F, and the reason for this name will be clear shortly. In light of the above theorem, if $F : M \to N$ is a diffeomorphism of M onto an open set $U \subseteq N$ and $p \in M$, then $F_* : T_p(M) \to T_{F(p)}(N)$ is an isomorphism since $F_*^{-1} \circ F_* : T_p(M) \to T_p(M)$ and $F_* \circ F_*^{-1} : T_{F(p)}(N) \to T_{F(p)}(N)$ are the identity isomorphisms of their corresponding vector spaces. Thus for a smooth manifold M and a coordinate neighborhood $\{U, \varphi\}$ on M, there is an isomorphism $\varphi_* : T_p(M) \to T_{\varphi(p)}(\mathbb{R}^n)$ of the tangent space at each point $p \in U$ onto $T_{\varphi(p)}(\mathbb{R}^n)$. Hence φ_*^{-1} maps $T_{\varphi(p)}(\mathbb{R}^n)$ isomorphically onto $T_p(M)$, where the images

$$E_{ip} = \varphi_*^{-1}(\partial/\partial x_i) \quad i = 1, \ldots, n$$

of the natural basis

$$\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}$$

at each $\varphi(p) \in \varphi(U) \subseteq \mathbb{R}^n$ determine at p a basis $E_{1p}, \ldots, E_{np}$ of $T_p(M)$; these bases are called *coordinate frames*. Notice $\dim T_p(M) = \dim M$ since for each coordinate neighborhood U on M there is a natural basis $E_{1p}, \ldots, E_{np}$ of $T_p(M)$ for every $p \in U$.

Theorem 2.0.4. *Let f be a C^∞ function defined in a neighborhood of p, and $\hat{f} = f \circ \varphi^{-1}$ its expression in local coordinates relative to $\{U, \varphi\}$. Then*

$$E_{ip} f = \left(\frac{\partial \hat{f}}{\partial x_i}\right)_{\varphi(p)}.$$

In particular, if $x_i(q)$ is the i^{th} coordinate function, $X_p x_i$ is the i^{th} component of X_p in the natural basis $E_{1p}, \ldots, E_{np}$, that is,

$$X_p = \sum_{i=1}^{n} (X_p x_i) E_{ip}.$$

Proof. For

$$E_{ip} = \varphi_*^{-1}\left(\frac{\partial}{\partial x_i}\right),$$

we have

$$E_{ip} f = \left(\varphi_*^{-1}\left(\frac{\partial}{\partial x_i}\right)\right) f = \frac{\partial}{\partial x_i}(f \circ \varphi^{-1})\Big|_{x = \varphi(p)}.$$

Now let f to be the i^{th} coordinate function, $f(q) = x_i(q)$, and $X_p = \sum \alpha_j E_{jp}$. Then

$$X_p x_i = \sum_j \alpha_j (E_{jp} x_i) = \sum_j \alpha_j \left(\frac{\partial x_i}{\partial x_j}\right)_{\varphi(p)} = \alpha_i.$$

□

We can use Theorem 2.0.2 to derive a formula for the matrix of the linear map F_* relative to local coordinate systems. Let $F : M \to N$ be a C^∞ map with $\{U, \varphi\}$ and $\{V, \psi\}$ coordinate neighborhoods on M and N, respectively, with $F(U) \subseteq V$. Suppose that in these local coordinates F is given by

$$y_i = F_i(x_1, \ldots, x_n), \quad i = 1, \ldots, m,$$

and that p is a point with coordinates $a = (a_1, \ldots, a_n)$. Then $F(p)$ has y coordinates determined by these functions; let $\partial y_j/\partial x_i$ denote $\partial F_j/\partial x_i$.

Theorem 2.0.5. *Let*

$$E_{ip} = \varphi_*^{-1}\left(\frac{\partial}{\partial x_i}\right)$$

and

$$\tilde{E}_{jF(p)} = \psi_*^{-1}\left(\frac{\partial}{\partial y_j}\right),$$

where $i = 1, \dots, n$ *and* $j = 1, \dots, m$*, be the basis of* $T_p(M)$ *and* $T_{F(p)}(N)$*, respectively, determined by the given coordinate neighborhoods. Then*

$$F_*(E_{ip}) = \sum_{j=1}^{m} \left(\frac{\partial y_j}{\partial x_i}\right)_a \tilde{E}_{jF(p)}, \quad i = 1, \dots, n.$$

In terms of components, if $X_p = \sum \alpha_i E_{ip}$ *maps to* $F_*(X_p) = \sum \beta_j \tilde{E}_{jF(p)}$*, then*

$$\beta_j = \sum_{i=1}^{n} \alpha_i \left(\frac{\partial y_j}{\partial x_i}\right)_a, \quad j = 1, \dots, m.$$

Proof. We have

$$F_*(E_{ip}) = F_* \circ \varphi_*^{-1}\left(\frac{\partial}{\partial x_i}\right)_{\varphi(p)}.$$

By the remarks preceeding the theorem, to compute the components relative to $\tilde{E}_{jF(p)}$ we must apply this vector as an operator on $C^\infty(F(p))$ to the coordinate functions y_j.

$$F_*(E_{ij})y_j = \left(F_* \circ \varphi_*^{-1}\left(\frac{\partial}{\partial x_i}\right)\right)y_j = \frac{\partial}{\partial x_i} y_j(F \circ \varphi^{-1}) = \frac{\partial F_j}{\partial x_i},$$

these derivatives being evaluated at $\varphi(p)$. □

Thus the matrix of the linear transformation F_* is the Jacobian $D\hat{F}$, where $\hat{F}$ is the expression for F in local coordinates around p and $F(p)$. Hence we may think of tangent vectors to a C^∞ manifold as directional derivatives, as was the case for $\mathbb{R}^n$. This is why F_* is called the differential.

3. Preliminary Results on Smooth Manifolds

With the notion of a tangent space to a manifold M at $p \in M$ in hand, we may now develop the theory of smooth manifolds as it relates to Frobenius' Theorem.

3.1. **Vector Fields.**

Over each smooth manifold M there is a natural vector bundle $T(M)$ known as the *tangent bundle* to M. As a set, $T(M)$ is given by

$$T(M) = \bigcup_{p \in M} T_p(M).$$

$T(M)$ can be given a smooth structure so that it is a vector bundle over M with bundle projection

$$\pi : T(M) \longrightarrow M$$

given by $\pi(X_p) = p$. A vector field is simply a section of $T(M)$. More precisely,

Definition 3.1.1. A *vector field* X *of class* C^r *on* M is a section $X : M \longrightarrow T(M)$ whose components in the frames of any local coordinates $\{U, \varphi\}$ are functions of class C^r on U.

In general, a C^∞ map $F : M \to N$ does not pushforward a vector field X on M. If F is not surjective, then no vectors are defined at points in the complement of $F(M)$. If F is not injective, then for a single point in N there are multiple vectors corresponding to push-forwards of X for different points in M. However if we are also given a vector field Y on N, it might be the case that X and Y are related to each other under F. More precisely,

Definition 3.1.2. Let $F : M \to N$ be a C^∞ map and, for $p \in M$, $F_* : T_p(M) \to T_{F(p)}(N)$ the associated homomorphism of tangent spaces. Let Y be a vector field on N such that for each $q \in N$ and $p \in F^{-1}(q) \subseteq M$ we have $F_*(X_p) = Y_q$. Then we say the vector fields X and Y are *F-related* and write $Y = F_*(X)$.

Theorem 3.1.3. *If $F : M \to N$ is a diffeomorphism, then each vector field X on M is F-related to a uniquely determined vector field Y on N.*

Proof. Since F is a diffeomorphism, the associated homomorphism $F_* : T_p(M) \to T_{F(p)}(N)$ is an isomorphism. Given a C^∞-vector field on X on M, then at each point $q \in N$, the vector $Y_q = F_*(X_{F^{-1}(q)})$ is uniquely determined. That Y is a C^∞-vector field follows immediately if we introduce local coordinates and apply Theorem 2.0.4 to the component functions. □

Thus, when $F : M \to N$ is a diffeomorphism, there is a unique pushforward of X by F.

3.2. **One-Parameter Group Actions.**

We now turn our attention to one-parameter group actions. Let $\theta : \mathbb{R} \times M \to M$ satisfy

(1) $\theta_0(p) = p$ for all $p \in M$,
(2) $\theta_t \circ \theta_s(p) = \theta_{t+s}(p) = \theta_s \circ \theta_t(p)$ for all $p \in M$ and $s, t \in \mathbb{R}$.

We think of θ as a C^∞ action of $\mathbb{R}$ on M. In fact, θ defines a C^∞-vector field X on M, which is called the *infinitesimal generator* of θ, according to the following prescription: For each $p \in M$, define $X_p : C^\infty(p) \to \mathbb{R}$ by

$$X_p f = \lim_{\Delta t \to 0} \frac{f(\theta_{\Delta t}(p)) - f(p)}{\Delta t}.$$

To see that X_p is a vector at p and that $p \to X_p$ defines a vector field, let $\{U, \varphi\}$ be a coordinate neighborhood of $p \in M$ and let $I_\delta \times V$ be a neighborhood of $(0, p)$ in $\mathbb{R} \times M$, where $I_\delta = \{t \in \mathbb{R} : -\delta < t < \delta\}$, and V and $\delta > 0$ are chosen such that $\theta(I_\delta \times V) \subset U$. In particular, $V = \theta_0(V)$ is contained in U and contains p. Restricted to the open set $I_\delta \times V$, we may write θ in local coordinates

$$y_1 = h_1(t, x_1, \dots, x_n),$$
$$\vdots$$
$$y_n = h_n(t, x_1, \dots, x_n),$$

or $y = h(t, x)$, where $x = (x_1, \ldots, x_n)$ are the coordinates of $q \in V$ and $y = (y_1, \ldots, y_n)$ of $\theta_t(q)$, its image. The h_i are defined and C^∞ on $I_\delta \times \varphi(V)$ and the range of $h(t, x)$ is in $\varphi(U)$. That θ_0 is the identity and $\theta_{t_1+t_2} = \theta_{t_1} \circ \theta_{t_2}$ is reflected in the conditions

$$h_i(0, x) = x_i \quad \text{and} \quad h_i(t_1 + t_2, x) = h_i(t_1, h(t_2, x))$$

for $i = 1, \ldots, n$. If $\hat{f}(x_1, \ldots, x_n)$ is the local expression for $f \in C^\infty(p)$, then

$$\frac{f(\theta_{\Delta t}(p)) - f(p)}{\Delta t} = \frac{\hat{f}(h(\Delta t, x)) - \hat{f}(x)}{\Delta t}$$

and

$$X_p f = \lim_{\Delta t \to 0} \frac{\hat{f}(h(\Delta t, x)) - \hat{f}(x)}{\Delta t} = \sum_{i=1}^{n} \dot{h}_i(0, x) \left(\frac{\partial \hat{f}}{\partial x_i} \right)_{\varphi(p)},$$

where the dot indicates differentiation with respect to t. This formula holds for all $p \in V$ and implies that on V, $X_p = \sum \dot{h}_i(0, x) E_{ip}$ where $E_i = \varphi_*^{-1}(\partial/\partial x_i)$ and $x = \varphi(p)$, which shows that X is a C^∞-vector field over V. Since every point of M lies in such a neighborhood, X is a C^∞-vector field on M.

If θ is not defined for all time t, then we restrict our attention to certain open subsets of $\mathbb{R} \times M$. In particular, for a C^∞ manifold M let $W \subset \mathbb{R} \times M$ be an open set which satisfies the following condition: for every $p \in M$ there exist real numbers $\alpha(p) < 0 < \beta(p)$ such that $W \cap (\mathbb{R} \times \{p\}) = \{(t, p) : \alpha(p) < t < \beta(p)\}$. Here $W \cap (\mathbb{R} \times \{p\})$ is the time at which the action of $\mathbb{R}$ on M is defined. Denote by $I(p)$ the interval $\alpha(p) < t < \beta(p)$ and by I_δ the interval defined by $|t| < \delta$, so that $W = \bigcup_{p \in M} I(p) \times \{p\}$.

Definition 3.2.1. A *local one-parameter group action*, or *flow*, on a manifold M is a C^∞ map $\theta : W \to M$ which satisfies the following two conditions:

(1) $\theta_0(p) = p$ for all $p \in M$.
(2) If $(s, p) \in W$, then $\alpha(\theta_s(p)) = \alpha(p) - s$, $\beta(\theta_s(p)) = \beta(p) - s$, and moreover for any t such that $\alpha(p) - s < t < \beta(p) - s$, $\theta_{t+s}(p)$ is defined and $\theta_t \circ \theta_s(p) = \theta_{t+s}(p)$.

When $\mathbb{R}$ acts on M, $\theta_t : M \to M$ is a diffeomorphism for each t, with $\theta_t^{-1} = \theta_{-t}$. A similar statement can be made for local actions. Let $V_t \subset M$ be the domain of definition of θ_t, that is, $V_t = \{p \in M : (t, p) \in W\}$.

Theorem 3.2.2. *V_t is an open set for every $t \in \mathbb{R}$ and $\theta_t : V_t \to V_{-t}$ is a diffeomorphism with $\theta_t^{-1} = \theta_{-t}$.*

Proof. Let $p_0 \in V_{t_0}$, so that $(t_0, p_0) \in V$. Since W is open, there is a $\delta > 0$ and a neighborhood V of p_0 such that $\{t : |t - t_0| < \delta\} \times V \subset W$. In particular, $\{t_0\} \times V \subset W$, so that $V \subset V_{t_0}$, so V_{t_0} is open. By (2) of Definition 3.2.1, if $p \in V_t$, then $\alpha(p) < t < \beta(p)$ and $t + (-t)$ lies in the same interval. Thus $\theta_t(p) \in V_{-t}$ and $\theta_{-t} \circ \theta_t(p) = p$. Similarly, $\theta_{-t}(V_{-t}) \subset V_t$ and $\theta_t \circ \theta_{-t}(q) = q$ for any $q \in V_{-t}$. Combining these statements with the fact that θ_t, θ_{-t} are C^∞ on any open subsets of M on which they are defined completes the proof. $\square$

3.3. **The Existence Theorem for Ordinary Differential Equations.** We now introduce the existence theorem for ordinary differential equations. The existence theorem plays a key role in several parts of the development of Frobenius' Theorem, and those results that are independent of future developments will be stated and proved in this section. We leave out the proof of the existence theorem, however, and instead provide a reference.

Theorem 3.3.1. (Existence theorem for ordinary differential equations) *Let U be an open subset of $\mathbb{R}^n$ and I_ε denote the interval $-\varepsilon < t < \varepsilon$, where $t \in \mathbb{R}$. Suppose $f_i(t, x_1, \dots, x_n)$, $i = 1, \dots, n$, are C^r functions on $I_\varepsilon \times U$ where $r \geq 1$. Then for each $x \in U$, there exists a $\delta > 0$ and a neighborhood V of x, where $V \subset U$, such that the following conditions are satisfied:*

(1) *For each $a = (a_1, \dots, a_n) \in V$ there exists an n-tuple of C^{r+1} functions $x(t) = (x_1(t), \dots, x_n(t))$ defined on I_δ and mapping I_δ into U which satisfy the system of first-order differential equations*
$$\frac{dx_i}{dt} = f_i(x, t), \quad i = 1, \dots, n,$$
and the initial conditions
$$x_i(0) = a_i, \quad i = 1, \dots, n.$$
For each a the functions $x(t) = (x_1(t), \dots, x_n(t))$ are uniquely determined in the sense that any other functions $x_1'(t), \dots, x_n'(t)$ satisfying the above two conditions agree with $x(t)$ on an open interval around $t = 0$.

(2) *These functions being uniquely determined by $a = (a_1, \dots, a_n)$ for every $a \in V$, we write them $x_i(t, a_1, \dots, a_n)$, $i = 1, \dots, n$, in which case they are of class C^r in all variables and thus determine a C^r map $I_\delta \times V \to U$.*

Proof. See [3]. □

An important consequence of the existence theorem is the uniqueness of integral curves. Given a vector field X on a manifold M, a curve $t \to F(t)$ defined on an open interval J of $\mathbb{R}$ is an *integral curve* of X if $dF/dt = X_{F(t)}$ on J. If we write F in terms of its coordinate functions
$$F(t) = (x_1(t), \dots, x_n(t)),$$
then the vector equation $\dot{F}(t) = X_{F(t)}$ is satisfied if and only if
$$\frac{dx_i}{dt} = f_i(x_1(t), \dots, x_n(t)), \quad i = 1, \dots, n.$$
Given $x \in F(J)$, the existence theorem implies that for each a in a neighborhood V of x there is a unique integral curve $F(t)$ satisfying $F(0) = a$. $F(t)$ is defined at least for $-\delta < t < \delta$ where $\delta > 0$ is the same for every $a \in V$. Using $F(t, a) = (x_1(t, a), \dots, x_n(t, a))$ to denote an integral curve through V which depends on the initial point a and an overdot for differentiation with respect to t, these equations

become $\dot{x}_i(t,a) = f_i(t,a)$ and $x_i(0,a) = a_i$ for $i = 1,\ldots,n$. By the existence theorem these functions are C^∞ on $I_\delta \times V$.

Theorem 3.3.2. *Let X be a C^∞-vector field on a manifold M. Then for each $p \in M$ there exists a neighborhood V and real number $\delta > 0$ such that there corresponds a C^∞ mapping*

$$\theta^V : I_\delta \times V \to M,$$

satisfying

$$\dot{\theta}^V(t,q) = X_{\theta^V(t,q)}$$

and

$$\theta^V(0,q) = q \quad \text{for all } q \in V.$$

If $F(t)$ is an integral curve of X with $F(0) = q \in V$, then $F(t) = \theta^(t,q)$ for $|t| < \delta$. In particular, this mapping is unique in the sense that if V_1, δ_1 is another such pair for $p \in M$, then $\theta^V = \theta^{V_1}$ on the intersection of their domains.

Proof. For $p \in M$, choose a coordinate neighborhood $\{U,\varphi\}$ and map X to the φ-related vector field $\tilde{X} = \varphi_*(X)$ on $\tilde{U} = \varphi(U) \subseteq \mathbb{R}^n$. Applying the existence theorem gives $F : I_\delta \times \tilde{V} \to \tilde{U}$ defined by $F(t,a) = (x_1(t,a),\ldots,x_n(t,a))$ on a neighborhood $\tilde{V} \subseteq \tilde{U}$ of $\varphi(p)$. We set $V = \varphi^{-1}(\tilde{V})$ and define $\theta^V : I_\delta \times V \to U$ by $\theta^V(t,q) = \varphi^{-1}(F(t,\varphi(q)))$. Since φ and φ^{-1} are diffeomorphisms, we see at once that θ^V satisfies the desired conditions and the final assertion of the theorem is a result of uniqueness of solutions. □

Having defined integral curves and derived a notion of uniqueness for each such curve, we can now determined when a C^∞-vector field is invariant under a given diffeomorphism. We remark that this result plays a significant role in §4.

Theorem 3.3.3. *Let X be a C^∞-vector field on a manifold M and $F : M \to M$ a diffeomorphism. Let $\theta(t,p)$ denote the C^∞ map $\theta : W \to M$ defined by X. Then X is invariant under F if and only if $F(\theta(t,p)) = \theta(t,F(p))$ whenever both sides are defined.*

Proof. Suppose that X is invariant under F. If $\theta_p : I(p) \to M$ is the integral curve of X with $\theta_p(0) = p$, then the diffeomorphism F takes it to an integral curve $F(\theta_p(t))$ of the vector field $F_*(X)$. Since $F_*(X) = X$ and $F(\theta_p(0)) = F(p)$, from uniqueness of integral curves we conclude that $F(\theta_p(t)) = \theta(t,F(p))$. Conversely, suppose $F(\theta(t,p)) = \theta(t,F(p))$. Let $\theta_p(t) = \theta(t,p)$ and let d/dt be the natural basis of $T_0(\mathbb{R})$. By definition, $X_p = \dot{\theta}_p(0) = \theta_{p*}(d/dt)$ and applying the isomorphism $F_* : T_p(M) \to T_{F(p)}(M)$ we have

$$F_*(X_p) = F_* \circ \theta_{p*}(d/dt) = (F \circ \theta_p)_*(d/dt) = \theta_{F(p)*}(d/dt) = X_{F(p)}.$$

□

4. The Lie Algebra of Vector Fields on a Manifold

Denote by $\mathbf{X}(M)$ the set of all C^∞ vector fields defined on a C^∞ manifold M. Then $\mathbf{X}(M)$ is an infinite-dimensional real vector space as well as a module over $C^\infty(M)$. As we shall see, $\mathbf{X}(M)$ is something more than a vector space. First, however, we must introduce a definition.

Definition 4.0.4. A *Lie algebra* is a vector space V over a field F equipped with a binary operation $[\cdot,\cdot] : V \times V \to V$, called the *Lie bracket*, such that the following conditions are satisfied:

(1) $[\cdot,\cdot] : V \times V \to V$ is bilinear over F, that is

$$[\alpha_1 X + \alpha_2 Y, Z] = \alpha_1[X, Z] + \alpha_2[Y, Z]$$

and

$$[X, \alpha_1 Y + \alpha_2 Z] = \alpha_1[X, Y] + \alpha_2[X, Z]$$

for all α_1, α_2 in F and X, Y, Z in V.

(2) $[\cdot,\cdot] : V \times V \to V$ is skew-commutative, that is

$$[X, Y] = -[Y, X]$$

for all X and Y in V.

(3) $[\cdot,\cdot] : V \times V \to V$ satisfies the Jacobi identity, that is

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Now let X and Y be in $\mathbf{X}(M)$. Then, in general, the operator $f \mapsto X_p(Yf)$ defined on $C^\infty(p)$ does not define a vector at p. Hence XY, considered as an operator on C^∞ functions on M, does not in general determine a C^∞-vector field. On the other hand, $Z = XY - YX$ does define a C^∞ vector field on M according to the prescription

$$Z_p f = (XY - YX)_p f = X_p(Yf) - Y_p(Xf).$$

For if $f \in C^\infty(p)$, then Xf and Yf are C^∞ on a neighborhood of p, and this prescription determines a linear map of $C^\infty(p)$ with $\mathbb{R}$. Thus if the Leibniz rule holds for Z_p, then Z_p is an element of $T_p(M)$ at each $p \in M$. So suppose f and g are in $C^\infty(p)$. Then f and g are in $C^\infty(U)$ for some neighborhood U of p. Using the notation $(Xf)_p$ for $X_p f$, we have

$$\begin{aligned}
(XY - YX)_p(fg) &= X_p(Yfg) - Y_p(Xfg) \\
&= X_p(fYg + gYf) - Y_p(fXg + gXf) \\
&= (X_p f)(Yg)_p + f(p)X_p(Yg) + (X_p g)(Yf)_p \\
&\quad + g(p)X_p(Yf) - (Y_p f)(Xg)_p - f(p)Y_p(Xg) \\
&\quad - (Y_p g)(Xf)_p - g(p)(Y_p Xf),
\end{aligned}$$

so that

$$\begin{aligned} Z_p(fg) &= (XY - YX)_p(fg) \\ &= f(p)(XY - YX)_p g + g(p)(XY - YX)_p f \\ &= f(p) Z_p g + g(p) Z_p f. \end{aligned}$$

FInally, if f is C^∞ on any open subset U of M, then so is $(XY - YX)f$, and therefore Z is a C^∞-vector field on M, as claimed. In light of this observation, define the bracket

$$[\cdot, \cdot] : \mathbf{X}(M) \times \mathbf{X}(M) \to \mathbf{X}(M)$$

by $[X, Y] = XY - YX$.

Theorem 4.0.5. $\mathbf{X}(M)$ *with the bracket* $[X, Y]$ *is a Lie algebra over* $\mathbb{R}$.

Proof. Let α and β be real numbers and let X_1, X_2, and Y be C^∞-vector fields on M. It is a simple calculation that $[\alpha X_1 + \beta X_2, Y]f = \alpha[X_1, Y]f + \beta[X_2, Y]f$ for any C^∞ function f, which establishes linearity in the first variable. Skew-commutativity holds by definition, which immediately yields linearity in the second variable. It remains to show that the bracket on $\mathbf{X}(M)$ satisfies the Jacobi identity. Let f be a C^∞ function. Then

$$\begin{aligned} [X, [Y, Z]]f &= X(([Y, Z])f) - [Y, Z](Xf) \\ &= X(Y(Zf)) - X(Z(Yf)) - Y(Z(Xf)) + Z(Y(Xf)). \end{aligned}$$

Permuting cyclically and adding establishes the identity. □

4.1. **The Lie Derivative.** Given a vector field X on M, there is an associated one-parameter group $\theta : W \to M$ generated by X. The isomorphism $\theta_{t*} : T_p(M) \to T_{\theta_t(p)}(M)$ and its inverse allow us to compare the values of vector fields at these two points. To do this, first recall that for each $t \in \mathbb{R}$, $\theta_t : V_t \to V_{-t}$ is a diffeomorphism of the open set V_t, provided V_t is not empty, by Theorem 3.2.2. In particular, for each $p \in M$ there is a neighborhood V and a $\delta > 0$ such that $V \subset V_i$ for $|t| < \delta$. If Y is a second C^∞-vector field on M, we compute at p the rate of change of Y in the direction of X, that is to say along the integral curve of the vector field X passing through p.

Definition 4.1.1. The vector field $L_X Y$, called the *Lie derivative* of Y with respect to X, is defined at each $p \in M$ by either of the following limits.

$$\begin{aligned} (L_X Y)_p &= \lim_{t \to 0} \frac{\theta_{-t*}(Y_{\theta(t,p)}) - Y_p}{t} \\ &= \lim_{t \to 0} \frac{Y_p - \theta_{t*} Y_{\theta(-t,p)}}{t}. \end{aligned}$$

The second equivalence is obtained by replacing t with $-t$. The first expression says first apply the isomorphism θ_{-t*} to $Y_{\theta(t,p)} \in T_{\theta(t,p)}(M)$, taking $T_{\theta(t,p)}(M)$ to $T_p(M)$. Second, in $T_p(M)$ take the difference of this vector and Y_p, multiply by the scalar $1/t$, and then pass to the limit as $t \to 0$. Equivalently, we may identify $Z_p(t) =$

$\theta_{-t*}(Y_{\theta(t,p)}) \in T_p(M)$ with a curve in $\mathbb{R}^n$, so that $(L_X Y)_p = \dot{Z}_p(0)$. Writing the above formula in local coordinates shows that the vector field we obtain is C^∞.

It turns out that if X and Y are vector fields on M, then the Lie derivative of Y with respect to X is simply $[X, Y]$. To prove this, however, first we need a lemma establishing the mean value theorem for integral curves of the vector field X passing through p.

Lemma 4.1.2. *Let X be a C^∞-vector field on M and θ be the corresponding map of $W \subset \mathbb{R} \times M$ onto M. Given $p \in M$ and $f \in C^\infty(U)$, U an open set containing p, we choose $\delta > 0$ and a neighborhood V of p in U such that $\theta(I_\delta \times V) \subset U$. Then there is a C^∞ function $g(t, q)$ defined on $I_\delta \times V$ such that for $q \in V$ and $t \in I_\delta$ we have*

$$f(\theta_t(q)) = f(q) + tg(t, q) \quad \textit{and} \quad X_q f = g(0, q).$$

Proof. There is a neighborhood V of p and $\delta > 0$ such that $\theta_t(p) = \theta(t, p)$ is defined and C^∞ on $I_\delta \times V$ and maps $I_\delta \times V$ into U by Theorem 3.3.2. The function $r(t, q) = f(\theta_t(q)) - f(q)$ is C^∞ on $I_\delta \times V$ and $r(0, q) = 0$. Denote by $\dot{r}(t, q)$ its derivative with respect to t. We define $g(t, q)$ for each fixed q by the formula

$$g(t, q) = \int_0^1 \dot{r}(ts, q)\, ds.$$

This function is also C^∞ on $I_\delta \times V$. By the fundamental theorem of calculus,

$$tg(t, q) = \int_0^1 \dot{r}(ts, q)t\, ds = r(t, q) - r(0, q) = r(t, q),$$

or

$$f(\theta_t(q)) = f(q) + tg(t, q).$$

However

$$g(0, q) = \lim_{t \to 0} g(t, q) = \lim_{t \to 0} \frac{r(t, q)}{t} = \lim_{t \to 0} \frac{f)\theta_t(q)) - f(q)}{t} = X_q f.$$

□

Theorem 4.1.3. *If X and Y are C^∞ vector fields on M, then*

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

Proof. For $p \in M$ and $f \in C^\infty(p)$, we have for some $\delta > 0$

$$\left[\frac{Y_p - \theta_{t*}(Y_{\theta(-t,p)})}{t}\right] f = \frac{Y_p f - Y_{\theta(-t,p)}(f \circ \theta_t)}{t} \tag{1}$$

for $0 < |t| < \delta$. Set $g_t(q) = g(t, q)$ and consider

$$\lim_{t \to 0} \frac{Y_p f - Y_{\theta(-t,p)}(f + tg_t)}{t},$$

which is equal to (1) by the previous lemma. Rewriting yields

$$\lim_{t \to 0} \frac{(Yf)(\theta_t(p)) - (Yf)(p)}{t} - \lim_{t \to 0} Y_{\theta_{t(p)}} g_t.$$

The first term is $X_p(Yf)$. Since $\theta_t(p) \to p$ and $g_t(p) \to X_q f$ as $t \to 0$, $Y_{\theta_t(p)} g_t \to Y_p(Xf)$ and we obtain in the limit as $t \to 0$ in (1)

$$X_p(Yf) - Y_p(Xf) = [X, Y]_p f.$$

Note this also shows $L_X Y$ is C^∞. □

We now prove a theorem that gives an interpretation for a vanishing Lie bracket. In particular, we may regard a vanishing Lie bracket as an action of $\mathbb{R}^2$ on M.

Theorem 4.1.4. *Let X and Y be C^∞-vector fields on a manifold M and let θ and σ denote the local one-parameter group actions they generate. Then $[X, Y] \equiv 0$ on M if and only if for each $p \in M$ there is a $\delta_p > 0$ such that $\sigma_s \circ \theta_t(p) = \theta_t \circ \sigma_s(p)$ for $|t|, |s| < \delta_p$.*

Proof. If $p \in M$, then there exists a neighborhood V of p and $\delta > 0$ such that $\sigma_s \circ \theta_t(q)$ and $\theta_t \circ \sigma_s(q)$ are defined for $q \in V$ and $|s|, |t| < \delta$. If $\sigma_s \circ \theta_t(q) = \theta_t \circ \sigma_s(q)$ on this set, then by Theorem 3.3.3 Y is θ_t-invariant on V for any fixed, small t. Thus

$$[X, Y]_q = (L_X Y)_q = \lim_{t \to 0} \frac{\theta_{-t*}(Y_{\theta(t,q)}) - Y}{t} = \lim_{t \to 0} \frac{Y_q - Y_q}{t} = 0.$$

Conversely, suppose $[X, Y] = 0$ on M. Given $q \in V$, let $Z_q(t) = \theta_{-t*}(Y_{\theta(t,q)})$. We show $\dot{Z}_q(t) = 0$. Let $q' = \theta(t, q)$. Then

$$\begin{aligned} \dot{Z}_p(t) &= \lim_{\Delta t \to 0} \frac{\theta_{-(t+\Delta t, q)*}(Y_{\theta(\Delta t, q')}) - \theta_{-t*}(Y_{q'})}{\Delta t} \\ &= \theta_{-t*}\left(\lim_{\Delta t \to 0} \frac{\theta_{-\Delta t*}(Y_{\theta(\Delta t, q')} - Y_{q'})}{\Delta t} \right) \\ &= \theta_{-t*}((L_X Y)_{q'}) = 0, \end{aligned}$$

since $L_X Y = [X, Y] = 0$. Hence $Z_q(t)$ is constant for $|t| < \delta$ and $q \in V$, so that Y is θ_t-invariant. Thus, by Theorem 3.3.3, $\theta_t \circ \sigma_s = \sigma_s \circ \theta_t$ on V. □

5. Frobenius' Theorem

Let M be a manifold of dimension $n + k$ and assume that to each $p \in M$ is assigned an n-dimensional subspace Δ_p of $T_p(M)$. Suppose that in a neighborhood U of each $p \in M$ there are n linearly independent C^∞-vector fields $X_1, \dots, X_n$ which form a basis of Δ_q for every $q \in U$.

Definition 5.0.5. Set

$$\Delta = \bigcup_{p \in M} \Delta_p.$$

We say that Δ is a C^∞ n-plane *distribution* of dimension n on M and that $X_1, \dots, X_n$ is a *local basis* of Δ.

There is a certain type of distribution with which we will be concerned.

Definition 5.0.6. A distribution Δ on a manifold M is *involutive* if there exists a local basis $X_1, \ldots, X_n$ in a neighborhood of each point such that

$$[X_i, X_j] = \sum_{k=1}^{n} c_{ij}^k X_k,$$

where $1 \le i$ and $j \le n$.

Note that in general the c_{ij}^k will not be constants, but rather C^∞ functions on the neighborhood. Before we state the next definition, let us review the notion of a submanifold.

Roughly speaking, a submanifold N of a differentiable manifold M is a subset that is itself a differentiable manifold. Unfortunately, this definition is ambiguous since it is unclear whether or not N should have the subspace topology of M. Such a submanifold is known as an *imbedded* submanifold, however an imbedded submanifold is not the most general type of submanifold. A more general notion is that of an *immersed* submanifold.

Definition 5.0.7. A submanifold N of a manifold M is an *immersed submanifold* if the image in M of an injective immersion $F : N' \to M$, $N = F(N')$, of a manifold N' into M together with the topology and C^∞ structure which makes $F : N' \to N$ a diffeomorphism.

Notice that it is not necessarily the case that the topology that makes N an immersed submanifold is the subspace topology from M; we give an example that illustrates this point shortly. For the rest of this paper, all submanifolds will be immersed submanifolds.

Definition 5.0.8. If Δ is a C^∞ distribution on a manifold M and N is a connected C^∞ submanifold of M such that for each $q \in N$ we have $T_q(N) \subseteq \Delta_q$, then we say N is an *integral manifold* of Δ.

Definition 5.0.9. A *maximal integral manifold* N of an involutive distribution Δ is a connected integral manifold which contains every connected integral manifold with which it has a point in common.

For an example of an involutive distribution, let $M = \mathbb{R}^{n+k}$ and let Δ be the distribution spanned by $X_i = \partial/\partial x_i$, $i = 1, \ldots, n$. Then Δ is the subspace of dimension n consisting of those vectors parallel to $\mathbb{R}^n$ at each point q of M, and the integral manifolds are the vector subspaces $\mathbb{R}^m$ of $\mathbb{R}^n$. We will soon show this trivial example is true locally for any involutive distribution on a manifold M.

An involutive n-plane distribution Δ on a manifold M defines what is known as a foliated structure, or *foliation*, on M with the maximal integral manifolds as *leaves*. If a foliation is given, the tangent spaces to leaves determine an involutive distribution and conversely.

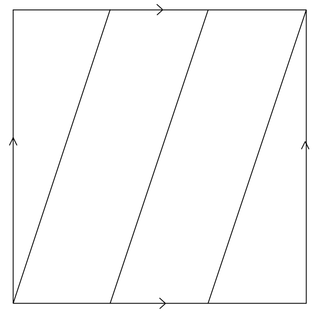

FIGURE 2. A foliation of the torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$.

For example, consider one-dimensional C^∞ foliations of the torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ given by families of parallel lines in $\mathbb{R}^2$ (for an example, see Figure 2). If the slope of a given family of parallel lines is rational, then each leaf is diffeomorphic to S^1. On the other hand, if the slope is irrational, then each leaf is diffeomorphic to $\mathbb{R}$ and can be shown to be dense in T^2. Note that the irrational case is an example where the subspace topology from T^2 disagrees with the topology that makes the family of lines a manifold.

Definition 5.0.10. Let Δ be a C^∞ distribution on a manifold M of dimension $m = n+k$, the dimension of Δ being n. Δ is *completely integrable* if each point $p \in M$ has a coordinate neighborhood $\{U, \varphi\}$ such that if $x_1, \ldots, x_m$ denote the local coordinantes, then the n vectors $E_i = \varphi_*^{-1}(\partial/\partial x_i)$, $i = 1, \ldots, n$, are a local basis on U for Δ.

Note that for a completely integrable distribution Δ on M, there is an n-dimensional integral manifold N through each point $q \in U$ such that $T_q(N) = \Delta_q$, thus $\dim N = n$. If $(a_1, \ldots, a_m)$ denotes the coordinates of q, then an integral manifold through q is the n-slice defined by $x_{n+1} = a_{n+1}, \ldots, x_m = a_m$, that is $N = \varphi^{-1}\{x \in \varphi(U) : x_j = a_j,\ j = n+1, \ldots, m\}$, a slice of U. In particular, Δ is involutive since

$$[E_i, E_j] = \varphi_*^{-1}\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right] = 0,$$

where $1 \leq i$ and $j \leq n$. In this case, we call $\{U, \varphi\}$ *flat* with respect to Δ, so that complete integrability is equivalent to every point having a flat coordinate neighborhood. In particular, any completely integrable distribution is involutive.

In general, however, distributions are not involutive. For example, on $\mathbb{R}^3$ the distribution

$$X_1 = x^3\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3}, \quad X_2 = \frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3}$$

is not involutive since $[X_1, X_2] = -\partial/\partial x_1$, which is not a linear combination of X_1 and X_2.

It is therefore natural to ask, what is a necessary and sufficient condition for a distribution Δ on a manifold M to be involutive? We have already shown being completely integrable is sufficient, and we now show that it is necessary; this is Frobenius' Theorem. First, however, we prove a lemma that will do all the hard work in the proof of Frobenius' Theorem. For the rest of the section, $m = \dim M$ and $n = \dim \Delta$.

Lemma 5.0.11. *Let $X_1, \dots, X_n$ be C^∞ vector fields defined on an open subset $U \subset \mathbb{R}^m$ such that 1) $\{X_{1q}, \dots, X_{nq}\}$ are linearly independent at each $q \in U$ and 2) $[X_i, X_j] = 0$ on U, $1 \le i, j \le n$. Then there is a coordinate neighborhood V, ψ around any $p \in U$ such that $X_i = E_i = \psi_*^{-1}(\partial/\partial x_i)$, $i = 1, \dots, n$, on V.*

Proof. Given $p \in U$, we may assume coordinates $(x_1, \dots, x_m)$ in $\mathbb{R}^m$ such that $p = (0, \dots, 0)$ and $X_{ip} = (\partial/\partial x_i)_0$ for $i = 1, \dots, n$ since we may center p at $(0, \dots, 0)$ and then perform a change of basis on the X_{ip}'s. Let $\theta_t^i(q) = \theta_i(t, q)$ denote the one-parameter group action generated by X_i. Using repeatedly both the existence theorem and continuity, we may assume there exists a neighborhood $W \subset U$ of p and a $\delta > 0$ such that for $|t_i| < \delta$ and for any permutation σ in the symmetric group S_n the mapping $\theta^1_{t_{\sigma(1)}} \circ \cdots \circ \theta^n_{t_{\sigma(n)}}$ is defined on W. Since all $[X_i, X_j] = 0$ on U, Theorem 4.1.4 implies that

$$\theta^{\sigma(1)}_{t_{\sigma(1)}} \circ \cdots \circ \theta^{\sigma(n)}_{t_{\sigma(n)}}(q) = \theta^1_{t_1} \circ \cdots \circ \theta^n_{t_n}(q)$$

for any $q \in W$ and $\sigma \in S_n$. Let $W' \subset W$ be a neighborhood of the origin such that $x = (x_1, \dots, x_n, x_{n+1}, \dots, x_m) \in W'$ implies that $|x_i| < \delta$, $i = 1, \dots, n$. We define a C^∞ mapping $\theta : W' \to \mathbb{R}^m$ by

$$\theta(x_1, \dots, x_n, \dots, x_m) = \theta^1_{x_1} \circ \cdots \circ \theta^n_{x_n}(0, \dots, 0, x_{n+1}, \dots, x_m).$$

Thus we project x to the plane of the last $m-n$ coordinates and then act by the various one-parameter groups, that is, move along the various integral curves of the X_i. This may be done in any order with the same results. Note that $\theta(0, \dots, 0, x_{n+1}, \dots, x_m) = (0, \dots, 0, x_{n+1}, \dots, x_m)$, that is θ is the identity on the $(m-n)$-slice $x_i = 0$, $i = 1, \dots, n$ of W'.

Now consider $\theta_*(\partial/\partial x_i|_a)$, $a \in W'$. For given i and a, let

$$F(t) = \theta(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_m);$$

then $\theta_*(\partial/\partial x_i|_a) = \dot{F}(a_i)$. First, note that for $i > n$ and $a = (0, \dots, 0, a_{n+1}, \dots, a_m)$, $\theta_*(\partial/\partial x_i) = \partial/\partial x_i$. In particular, this holds at $p = (0, \dots, 0)$. For $i \le n$ and any $a \in W'$,

$$F(t) = \theta_t^i[\theta^1_{a_1} \circ \cdots \circ \hat{\theta}^i_t \circ \cdots \circ \theta^n_{a_n}(0, \dots, 0, a_{n+1}, \dots, a_m)],$$

where the hat denotes that θ_t^i is omitted. We have used the independence of order. Let q denote the point in brackets; then $F(t) = \theta_t^i(q)$, so $\dot{F}(a_i) = X_{i\theta_i(a_i,q)} = X_{i\theta(a,q)}$. Thus $\theta_*(\partial/\partial x_i|_a) = X_{i\theta(a)}$, and in particular $\theta_*(\partial/\partial x_i|_0) = X_{ip} = \partial/\partial x_i|_0$ for $i = 1, \dots, n$. It follows that θ is a diffeomorphism on a neighborhood $V' \subset W'$ of the origin by the Inverse Function Theorem and that $V = \theta(V')$, $\psi = \theta^{-1}$ is the desired coordinate neighborhood. □

With Lemma 5.0.11 in hand, we may finally prove Frobenius' Theorem.

Theorem 5.0.12. (Frobenius' Theorem) *A distribution Δ on a manifold M is completely integrable if and only if it is involutive.*

Proof. We have shown that if Δ is completely integrable then it is involutive, and it remains to be shown that if Δ is involutive, then it is completely integrable. First, we show that in a neighborhood U of any $p \in M$ an involutive n-plane distribution Δ has a local basis of vector fields $Y_1, \dots, Y_n$ satisfying $[Y_i, Y_j] \equiv 0$. To do this, we simply use linear algebra. Let $X_1, \dots, X_n$ be an arbitrary local basis defined on a coordinate neighborhood U', φ of p. Relative to the coordinate frames $X_{iq} = \sum_{a=1}^{m} \alpha_{ia}(q) E_{aq}$ we may suppose coordinates so numbered that the matrix $(\alpha_{ij}(q))$, $1 \le i, j \le n$, is non-singular on a neighborhood $U \subset U'$ of p. Let $(\beta_{ij}(q))$ denote its inverse and define a new basis of Δ on U by

$$Y_{iq} = \sum_{j=1}^{n} \beta_{ij}(q) X_{jq} = E_{iq} + \sum_{\nu=n+1}^{m} \gamma_{i\nu}(q) E_{\nu q},$$

where $\gamma_{i\nu} = \sum_{j=1}^{n} \beta_{ij}\alpha_{j\nu}$, $i = 1, \dots, n$; $\nu = n+1, \dots, m$. Since Δ is involutive,

$$[Y_i, Y_j] = \sum_{k=1}^{n} c_{ij}^{k} Y_k = \sum_{k=1}^{n} c_{ij}^{k} (E_k + \sum_{\tau=n+1}^{m} \gamma_{k\tau} E_\tau)$$

on U. On the other hand,

$$[Y_i, Y_j] = [E_i + \sum_{\nu} \gamma_{i\nu} E_\nu, E_j, \sum_{\mu} \gamma_{j\mu} E_\mu] = \sum_{\tau=n+1}^{m} \sigma_{ij}^{\tau} E_\tau.$$

Since $[E_a, E_b] = 0$ for $1 \le a, b \le m$. It is possible for these two linear combinations of $E_1, \dots, E_m$ to be equal only if $c_{ij}^{k} = 0$ on U for $1 \le i, k \le n$, so $Y_1, \dots, Y_n$ is the desired local basis.

Now, given an involutive distribution Δ on a manifold M and $p \in M$, choose a coordinate neighborhood W, σ of p on which there is a local basis $Y_1, \dots, Y_n$ such that $[Y_i, Y_j] \equiv 0$. In $\sigma(W) \subset \mathbb{R}^m$, the vector fields $\tilde{Y}_i = \sigma_*(Y_i)$ have the same property so that by the preceeding lemma there exists a coordinate neighborhood V, ψ of $\sigma(p)$, $V \subset \sigma(W)$ with coordinates $y_1, \dots, y_m$ and with $\tilde{Y}_i = \psi_*^{-1}(\partial/\partial y_i)$ for $i = 1, \dots, n$. Without loss of generality, assume $\psi(V) = B_\varepsilon^m(0)$, an ε-ball in $\mathbb{R}^m$, and $\psi(\sigma(p)) = (0, \dots, 0)$. Define the coordinate neighborhood $\tilde{U}, \varphi$ of p on M by $\tilde{U} = \sigma^{-1}(V)$ with $\varphi = \psi \circ \sigma$. Then $\varphi(p) = (0, \dots, 0)$, $\varphi(\tilde{U}) = B_\varepsilon^m(0)$ and $E_i = \sigma_*^{-1} \circ \psi_*^{-1}(\partial/\partial y_i) = \sigma_*^{-1}(\tilde{Y}_i) = Y_i$ for $i = 1, \dots, n$. Thus Δ is completely integrable. □

6. Acknowledgements

I thank my advisers William Lopes and Kathryn Mann for all the help they gave me during the writing of this paper and throughout the summer.

References

[1] William M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, 2003.

[2] Allen E. Hatcher, *Vector Bundles and K-Theory*, available for download at
`http://www.math.cornell.edu/~hatcher/VBKT/VBpage.html`.
[3] Vladimir I. Arnol'd, *Ordinary Differential Equations*, Springer Textbook, 1992.

Department of Mathematics, The University of Chicago, Chicago, Illinois 60637