ODEs and Integral Flows

Shrivathsa Pandelu

January 20, 2021

Contents

1 Solving a system of equations

Let D be some open set in $\mathbb{R}^{n+1}$ and let $f_1, \ldots, f_n$ be continuous functions defined on D. We seek to obtain differentiable functions $x_i(t), 1 \leq i \leq n$ such that

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = f_i(x_1, \ldots, x_n, t), 1 \leq i \leq n$$

on some interval (t_1, t_2).

Following [1] we shall denote vectors by capital letters, scalars by small letters and the norm of a vector is the usual Pythagorean norm. A vector function of t is a function $X(t) = (x_1(t), \ldots, x_n(t))$ and as usual, it is continuous, differentiable etc. if each of the components is. Denote by F the collection of functions $F = (f_1, \ldots, f_n)$, then we wish to solve

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X, t) \tag{1}$$

on some interval (t_1, t_2). Let $X_0 = (x_{10}, \ldots, x_{n0}), t_0$ be fixed in the given domain D, and suppose F is defined on a region $R = |X - X_0| \leq a, |t - t_0| \leq b$, then we would also like to have a solution satisfying $X(t_0) = X_0$.

Definition 1. *An approximate solution to 1 with error ϵ is a continuous admissible function (i.e., the image should land where F is defined) X defined on some interval around t_0 which is piecewise differentiable such that $X(t_0) = X_0$ and $|X'(t) - F(X(t), t)| \leq \epsilon$ wherever X is differentiable.*

Since R is compact, F is continuous it is bounded by some constant M.

Theorem 1. (*Existence of approximate solutions*) *With a setup as above, given an $\epsilon > 0$ we can construct an approximate solution with error ϵ over the interval $|t - t_0| \leq h = \min(a, b/M)$.*

Proof. The idea is to form an approximate solution as a union of lines (this is somewhat related to how integral of a function is defined using step functions). The constriants on h arise from requiring the lines to lie in R.

Given $\epsilon > 0$ by uniform continuity of F on R, there's a $\delta > 0$ such that if $|X_1 - X_2| < \delta, |t_1 - t_2| < \delta$, then $|F(X_1, t_1) - F(X_2, t_2)| < \epsilon$. We first define our approximate solution to the right of t_0. Obtain numbers $t_0 < t_1 < \cdots < t_m = t_0 + h$ such that

$$|t_{i+1} - t_i| < \min(\delta, \delta/M), i = 0, \ldots, m-1.$$

The approximate solution is defined recursively as

$$X_i(t) = X_{i-1}(t_i) + F(X_{i-1}(t_i), t_i)(t - t_i), t \in [t_i, t_{i+1}], 0 \leq i \leq m-1$$

with $X_{-1}(t_0) = X_0$. The union of these X_is form a continuous function which is piecewise differentiable. We need to verify that this is admissible. By the choice of h, there is a cone starting at X_0 that is contained in R whose boundary is the union of lines with slope M going to the right of t_0 starting from X_0. Each of the line segments X_i is contained in that cone and recursively we see that this solution is admissible. More directly, the end points of the line segments are within $M\delta$ of each other and adding them up gives a distance of $M\delta \times h/\delta \leq b$.

Next, by the choice of δ, for $t_i \leq t \leq t_{i+1}$ we have

$$|X_i(t) - X_i(t_i)| < \delta, t - t_i < \delta$$

therefore $|F(X_i(t), t) - X_i'(t)| = |F(X_i(t), t) - F(X_i(t_i), t_i)| < \epsilon$. Thus we have constructed an approximate solution to the right of t_0. In a similar manner, but going to the left we get an apprximate solution with error ϵ defined on $|t - t_0| \leq h$. $\square$

In order to guarantee exact solutions we will need additional conditions on F, namely Lipschitz continuity with respect to X. This is not too restrictive a condition because on usually deals with smooth F, and being defined on compact sets (or if we are just interested in existence of solutions we may restrict our attention to compact sets using local compactness) it will be Lipschitz.

2 Lipschitz continuity and properties

Definition 2. *A (continuous) vector function $F(X, t)$ defined on some open $D \subseteq \mathbb{R}^{n+1}$ satisfies the Lipschitz condition with respect to X there is a constant k such that $|F(X_1, t) - F(X_2, t)| \leq k|X_1 - X_2|, (X_1, t), (X_2, t) \in D$. The same defintion applies to other domains of definition and when F is scalar.*

Lemma 1. *Let R be a region convex in X, i.e., for each t_0, the subset $t = t_0$ of $\mathbb{R}^n$ is convex (or empty), and suppose the scalar function f has partial derivatives $\frac{\partial f}{\partial x_i}$ defined on R and bounded by some N, then f satisfies the Lipschitz condition on X with constant nN.*

Proof. Because of convexity, we may apply the mean value theorem. For a given $(X_1, t), (X_2, t) \in R$ we have

$$|f(X_1, t) - f(X_2, t)| = |\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_{i1} - x_{i2})| \leq N \sum_{i=1}^{n} |x_{i1} - x_{i2}| \leq nN|X_1 - X_2|$$

where x_{i1}, x_{i2} are the ith components of X_1, X_2 respectively. We can in fact, using Cauchy-Schwartz inequality obtain the better constant $\sqrt{n}N$. $\square$

Lemma 2. *Let R be a region embedded in a larger domain D such that the boundaries of the two have a minimum distance of some $\delta > 0$. Suppose scalar functions $f, \partial f/\partial x_i$ are defined and bounded by M, N respectively on D, then f satisfies the Lipschitz condition on X in R with constant $\max(2M/\delta, nN)$.*

Proof. Let $(X_1, t), (X_2, t) \in R$. If $|X_1 - X_2| \geq \delta$, then

$$\frac{|f(X_1,t) - f(X_2,t)|}{|X_1 - X_2|} \leq \frac{2M}{\delta}.$$

If $|X_1 - X_2| < \delta$, then the line segment between (X_1, t) and (X_2, t) must lie in D. For otherwise (by connectedness of the line), it must include points from the boundaries of R, D which would contradict the definition of δ. Now we may apply the argument of the previous lemma to obtain the constant nN. □

Lemma 3. *A vector function F satisfies the Lipschitz condition on X if and only if each of its components satisfy the Lipschitz condition on X.*

Proof. The proof is quite simple and left to the reader. □

Lemma 4. *If a vector function $X(t)$ is defined and differentiable on some interval (t_1, t_2), then the scalar function $|X(t)|$ is differentiable on the open set $X(t) \neq 0$ and we have*

$$\left|\frac{\mathrm{d}|X|}{\mathrm{d}t}\right| \leq \left|\frac{\mathrm{d}X}{\mathrm{d}t}\right|.$$

Proof. On the open set where $X \neq 0$, we have by chain rule

$$\frac{\mathrm{d}|X|}{\mathrm{d}t} = \frac{\sum x_i \frac{\mathrm{d}x_i}{\mathrm{d}t}}{|X|}.$$

Furthermore,

$$\left|\frac{|X(t+h)| - |X(t)|}{h}\right| \leq \left|\frac{X(t+h) - X(t)}{h}\right|$$

and since the limit exists, the same inequality holds for the derivatives. □

Theorem 2. (*Fundamental inequality*) *Let $F(X,t)$ be continuous in some region R and satisfy the Lipschitz condition on X with constant k. Let $X_1(t), X_2(t)$ be two approximate solutions over an interval $|t - t_0| \leq b$ with errors ϵ_1, ϵ_2 respectively. Set*

$$P(t) = X_1(t) - X_2(t); p(t) = |P(t)|; \epsilon = \epsilon_1 + \epsilon_2.$$

Then

$$|p(t)| \leq e^{k|t-t_0|}|p(t_0)| + \frac{\epsilon}{k}(e^{k|t-t_0|} - 1)$$

for $|t - t_0| \leq b$.

Proof. P is differentiable wherever both X_1, X_2 are differentiable, hence it is differentiable at all but finitely many points. Wherever P is differentiable we by the previous lemma

$$|p'(t)| \leq |P'(t)|$$

and by definition of an approximate solution

$$|P'(t)| \leq \epsilon + |F(X_1(t),t) - F(X_2(t),t)| \leq \epsilon + k|X_1(t) - X_2(t)| = \epsilon + k|p(t)|.$$

Therefore, we have a continuous function $p(t)$ on $|t - t_0| \leq b$ such that

$$|p'(t)| \leq k|p(t)| + \epsilon$$

wherever p is differentiable. Since p is differentiable at all but finitely many points, we can meaningfully talk about integrating p' and the fundamental theorem of calculus applies.

Suppose p is always positive, then to the right of t_0,

$$e^{-kt}(p'(t) - kp(t)) \leq e^{-kt}\epsilon$$

and integrating from t_0 to $t > t_0$ we get the desired inequality. To the left of t_0, use $-p' \leq kp + \epsilon$ instead.

If p is identically 0, then the inequality is obvious. Now, suppose $p(t) \neq 0$ for some $t > t_0$ and that there is a $t_0 < t' < t$ such that $p(t') = 0$. Using continuity, one obtains $t_0 < t_1 < t$ such that $p(t_1) = 0$ and p is non zero on (t_1, t). Now apply the previous case to obtain

$$|p(t)| \leq \frac{\epsilon}{k}(e^{k(t-t_1)} - 1)$$

which is better than the bound in the statement. One does a similar thing for points to the left of t_0. For p being negative, simply use $-p$. □

3 Existence and Uniqueness

From the fundamental inequality, we have the following

Theorem 3. (*Uniqueness*) *Let $F(X,t)$ be continuous and satisfy the Lipschitz condition on X in some neighbourhood of a fixed point (X_0, t_0). Then there can exist at most one solution of $dX/dt = F(X(t), t)$ such that $X(t_0) = X_0$.*

Proof. While the fundamental inequality applied to (closed) intervals around t_0, we can adapt the same proof here as well, since we are dealing with an open neighbourhood of (X_0, t_0). □

Theorem 4. (*Existence*) *Let $F(X,t)$ be continuous and satisfy a Lipschitz condition on X in a region*

$$|X - X_0| \leq b, |t - t_0| \leq a.$$

Let M be the upper bound of $|F|$ in R. Then there exists a solution $X(t)$ of $dX/dt = F(X,t)$ defined over the interval $|t - t_0| \leq h = \min(a, b/M)$.

Proof. The idea is to take the uniform limit of approximate solutions. Let ϵ_n be a sequence of positive numbers decreasing to 0 and let $X_n(t)$ be approximate solutions with error ϵ_n defined over $|t-t_0| \leq h$ passing through X_0 at t_0.

Step 1

Applying the fundamental inequality,

$$|X_n(t) - X_{n+m}(t)| \leq \frac{\epsilon_n + \epsilon_{n+m}}{k}(e^{k|t-t_0|} - 1) \leq \frac{2\epsilon_n}{k}(e^{kh} - 1).$$

Therefore, these X_n converge uniformly to some function X. By uniform convergence, X is continuous and $X(t_0) = X_0$. We need to show that X is differentiable with the desired derivative. We first show that the approximate derivatives of $F(X_n, t)$ converge uniformly to $F(X, t)$.

Step 2

Let R be the region $|X - X_0| \leq Mh, |t - t_0| \leq h$. Given $\epsilon > 0$ by uniform continuity of F, there is a $\delta > 0$ such that whenever $(X_1, t_1), (X_2, t_2) \in R$ and $|X_1 - X_2| < \delta, |t_1 - t_2| < \delta$ we have $|F(X_1, t_1) - F(X_2, t_2)| < \epsilon$.

For this δ, there is an N such that for $n \geq N, |X_n(t) - X(t)| < \delta, |t - t_0| \leq h$. Therefore, for $n \geq N$

$$|F(X,t) - F(X_n, t)| < \epsilon, n \geq N, |t - t_0| \leq h.$$

Thus, $F(X_n, t)$ converges uniformly to $F(X,t)$ on $|t - t_0| \leq h$, hence the integrals $\int_{t_0}^{t} F(X_n, t)dt$ converge to $\int_{t_0}^{t} F(X, t)$ on $|t - t_0| \leq h$.

Step 3

We have shown above that X_n converge uniformly to X, their derivatives converge to $F(X,t)$ uniformly. Given any $t \in |t - t_0| \leq h$ we have

$$\left| \int_{t_0}^{t} [\frac{\mathrm{d}X_n}{\mathrm{d}t} - F(X_n, t)]dt \right| \leq \epsilon_n h$$

therefore, applying the fundamental theorem of calculus (recall that X_n are differentiable at all but finitely many points)

$$|X_n(t) - X_n(t_0) - \int_{t_0}^{t} F(X_n, t)dt| \leq \epsilon_n h$$

We can take the limit of the left side and by uniform convergence, we have

$$|X(t) - X_0 - \int_{t_0}^{t} F(X(t), t)dt| = 0$$

which means that X is differentiable with derivative $F(X, t)$ as required. $\square$

Now that we have shown the existence and uniqueness of solutions, we observe that our approximate solutions as defined above are indeed approximations of the exact solution.

Theorem 5. *Let X be an exact solution of 1 under the assumptions of the previous theorem and $\tilde{X}$ be an approximate solution with error ϵ for $|t - t_0| \leq h$ such that $\tilde{X}(t_0) = X_0$. Then there is a constant N independent of ϵ such that $|\tilde{X}(t) - X(t)| \leq \epsilon N, |t - t_0| \leq h$.*

Proof. Applying the fundamental inequality to $\tilde{X}, X$,

$$|\tilde{X}(t) - X(t)| \leq \frac{\epsilon}{k}(e^{kh} - 1); \text{ i.e., } N = \frac{e^{kh} - 1}{k}.$$ $\square$

Without the Lipschitz condition, the solution need not be unique: consider $dy/dx = y^{1/3}$ with $f(x, y) = y^{1/3}$ which is continuous. There are two solutions passing through $(0, 0)$, namely $y \equiv 0$ and

$$y(x) = \begin{cases} (\frac{2}{3}x)^{3/2} & x \geq 0 \\ 0 & x \leq 0 \end{cases}$$

However, the Lipschitz condition is not required for the existence of solutions. Once can use polynomial approximations of F using the Weierstrass approximation theorem together with the Arzela-Ascoli theorem. Essentially, if P_n are polynomials uniformly converging to F, and if X_n are exact solutions to $dX/dt = P_n$, then one can show that there is a subsequence of X_n converging uniformly to some X and that this X is the required solution. For a proof see [1].

Before moving to the next part, it is good to mention that although above we arrived at an exact solution using Cauchy-Euler approximations, this is not the only way to arrive at a solution. One such way is Picard's theorem which uses a fixed point theorem (contraction theorem). This is similar to Newton's method for finding roots in that one starts with an arbitrary function and generates successive functions which eventually arrive at a solution. Furthermore, there are refinements of the Cauchy-Euler construction, see [1].

4 Smoothness of solutions

In the previous section our solution was essentially a curve passing through a given point (X_0, t_0). Now we look at how it the family of these curves depends on the initial condition X_0.

First we observe that if F is of class $\mathcal{C}^k$, then a solution $X(t)$ defined around some interval around t_0 is of class $\mathcal{C}^{k+1}$. When considering vector fields over manifolds, as we shall later, we can first restrict our attention to some coordinate chart, obtain a smooth solution over an interval and then try to patch up our solutions (or simply deal with maximal intervals where a smooth solution exists).

Theorem 6. *If $F(X, t)$ is continuous and satisfies a Lipschitz condition on X in $|X - \tilde{X}_0| \leq b, |t - t_0| \leq a$, for some fixed $(\tilde{X}_0, t_0)$, then there exists a unique solution $X(X_0, t)$ of*

$$X' = F(X, t), X(t_0) = X_0$$

in the region

$$|X_0 - \tilde{X}_0| \leq b/2, |t - t_0| \leq h', |X - X_0| \leq Mh'$$

where $h' = \min(a, (b/2M))$, M being the bound of F on $|X - \tilde{X}_0| \leq b, |t - t_0| \leq a$.

Proof. The given region contains around (X_0, t_0) the rectangle $|X - X_0| \le b/2, |t - t_0| \le a$ for $|\tilde{X}_0 - X_0| \le b/2$. From the existence theorem, a solution to $X' = F(X,t)$ exists in the interval as mentioned. Note that h' is no less than half of the original h. $\square$

What this means is that there is a smaller neighbourhood of a given $(\tilde{X}_0, t_0)$ where we have a family of solutions passing through every point (X_0, t_0). Note that we are not changing t_0. In the context of vector fields on manifolds, we may take $t_0 = 0$ and the theorem says that there is a flow of a neighbourhood of $(\tilde{X}_0, 0)$, which gives us, in a sense, a time parameterised family of diffeomorphisms which at $t = 0$ gives us the identity morphism and such that each point moves along the vector field determined by F.

Theorem 7. *If in some open neighbourhood of $(\tilde{X}_0, t_0), F(X,t)$ is continuous and satisfies a Lipschitz condition on X for some k, and if a solution $X(X_0, t)$ of*

$$X' = F(X,t), X(t_0) = X_0$$

exists in some rectangle $R, |t - t_0| \le h, |X_0 - \tilde{X}_0| \le l$, then $X(X_0, t)$ is continuous in X_0, t simultaneously in R.

Here we are considering the following function. Given $(X_0, t) \in R$, its image is the value of the curve passing through (X_0, t_0) at time t. The first coordinate gives us a curve, and the second gives us a point on this curve.

Proof. This is a simple application of the fundamental inequality. Fix a point $(X_0^{(1)}, t_1)$. For any $X_0^{(2)}$, we have by the fundamental inequality

$$|X(X_0^{(1)}, t) - X(X_0^{(2)}, t)| \le |X_0^{(1)} - X_0^{(2)}| e^{kh}, |t - t_0| \le h.$$

Now, given $\epsilon > 0$, choose a $\delta > 0$ so that

$$|t - t_1| < \delta \implies |X(X_0^{(1)}, t) - X(X_0^{(1)}, t_1)| < \epsilon$$

Combining the two inequalities and applying the triangle inequality, it is clear that $X(X_0, t)$ is continuous at $(X_0^{(1)}, t_1)$. $\square$

The fundamental inequality shows that the solution is uniformly continuous with respect to the intial value, this means that two curves are never farther apart than their initial points, upto a constant.

Lemma 5. *Let C be a compact subset of $\mathbb{R}^{n+m}, n \ge 1, m \ge 0$ and let f be a scalar function defined on C. Write points of $\mathbb{R}^{n+m}$ as $(x_1, \ldots, x_n, y)$ or (x, y). Suppose f extends to a neighbourhood U of C and has continuous partial derivatives $\partial f / \partial x_i$ in U, then*

$$\frac{f(x+h, y) - f(x,y) - \sum_{i=1}^{n} \partial_i f(x,y) h_i}{|h|}$$

approaches 0 as $h \to 0$ uniformly over C.

Proof. In the proof below, we use $|x|$ to denote the absolute value of real numbers and magnitudes of vectors. For each $p \in C$ there is an $r_p > 0$ such that $B^{n+m}(p, 2r_p) \subseteq U$. Cover C by $\{B^{n+m}(p, r_p)\}_{p \in C}$ and obtain a finite subcover $B^{n+m}(p_1, r_1), \ldots, B^{n+m}(p_k, r_k)$ and let $r = \min\{r_1, \ldots, r_k\}$. Given $p \in C$, it is clear that $B^{n+m}(p, r) \subset U$. So we may define the error function

$$E \colon C \times B(0,r) \to \mathbb{R}$$
$$((x,y), h) \mapsto \begin{cases} \frac{f(x+h,y) - f(x,y) - \sum_{i=1}^{n} \partial_i f(x,y) h_i}{|h|} & h \neq 0 \\ 0 & h = 0 \end{cases}$$

where $B(0,r)$ is the ball of radius r around the origin in $\mathbb{R}^n$. We wish to show that given any $\epsilon > 0$ there is a $\delta > 0$ such that for $|h| < \delta, |E((x,y),h)| < \epsilon \,\forall\, (x,y) \in C$. Fix an $\epsilon > 0$, we know that for each $p = (x,y) \in C$ there is a $\delta_p > 0$ such that the inequality holds at p. We first show that the same δ_p works on a neighbourhood of (x,y).

For two points $p = (x,y), q = (x',y')$ and an $h \in B(0,r), h \neq 0$, write $p_h = (x+h,y), q_h = (x'+h,y')$, then

$$\begin{aligned}
|E(q,h)||h| &\leq |E(p,h)||h| + |f(q_h) - f(q) - f(p_h) + f(p)| + \sum_{i=1}^{n} |\partial_i f(p) - \partial_i f(q)||h_i| \\
&\leq |E(p,h)||h| + |\sum_{i=1}^{n} \partial_i f(z',y')h_i - \sum_{i=1}^{n} \partial_i f(z,y)h_i| + \sum_{i=1}^{n} |\partial_i f(p) - \partial_i f(q)||h_i| \\
&\leq |E(p,h)||h| + \sum_{i=1}^{n} |\partial_i f(z',y') - \partial_i f(z,y)||h_i| + \sum_{i=1}^{n} |\partial_i f(p) - \partial_i f(q)||h_i|
\end{aligned}$$

where z is some point on the line between $x, x+h$ and z' on the line between $x', x'+h$ assuming that these lines are in U. This assumption is justified if the above calculation is done in some suitable ball around (x,y).

Now fix $p_0 \in C$ and $\epsilon > 0$. There is a δ_0 such that for $|h| < \delta_0, E(p_0,h) < \epsilon$. Using uniform continuity of f and its derivatives on C obtain a $\delta > 0$ such that for two points within δ the values of f and its partial derivatives are within ϵ. We can take $4\delta_0 < \delta < r$. Let q be within $\delta/4$ of p_0, then for $|h| < \delta_0$, from the calculation above we have

$$|E(q,h)|h| \leq |E(p_0,h)||h| + n\epsilon|h| + n\epsilon|h| \leq (2n+1)\epsilon|h|.$$

It follows that for a given $\epsilon > 0, p_0 \in C$, there is are $\delta, \delta_0 > 0$ such that for any q within δ of p_0, and $|h| < \delta_0$, we have $|E(q,h)| < \epsilon$. Cover C with such open balls and take a finite subcover. Taking the minimum of all the δ_0s, it follows that $E(p,h)$ converges uniformly to 0 as $h \to 0$ on C. $\square$

Remark. Note that the $\mathbb{R}^m$ factor can be replaced by any other metric space, the presence of the second coordinate doesn't affect the proof in anyway. We can apply mean value theorem because we are dealing with a scalar function, and fixing the second coordinate gives us a $\mathcal{C}^1$ function on some open subset of $\mathbb{R}^n$.

Theorem 8. *Let $F(X,t)$ be continuous and satisfy a Lipschitz condition on X in some region $R : |X - \tilde{X}_0| \leq b, |t - t_0| \leq a$ with upper bound M. Suppose a solution exists on R for all initial points. Write*

$$\begin{aligned}
&F = (f_1, \ldots, f_n) \\
&X(X_0,t) = (x_1(X_0,t), \ldots, x_n(X_0,t)) \\
&X_0 = (x_{10}, \ldots, x_{n0})
\end{aligned}$$

Suppose all partial derivative $\frac{\partial f_i(X,t)}{\partial x_j}, 1 \leq i,j \leq n$ exist and are continuous simultaneously in X, t on R. Then $\frac{\partial x_i(X_0,t)}{\partial x_{j0}}, 1 \leq i,j \leq n$ exist and are continuous in X_0, t simultaneously over R.

Note here that we have assumed a solution exists with any initial point in R, usually we would have to shrink R and then apply the theorem.

Proof. By shrinking R if necessary, we may also assume that F and its derivatives exist on a neighbourhood of R. We will show that $\partial x_i(X_0,t)/\partial x_{10}$ exist for all i, the other derivatives are similar. Take two points $X_0, X_0 + (\Delta x_{10}, 0, \ldots, 0)$. With these points as initial conditions, we have solutions

$$(x_1(X_0,t), \ldots, x_n(X_0,t)) \text{ and } (x_1(X_0,t) + \Delta x_1(X_0,t), \ldots, x_n(X_0,t) + \Delta x_n(X_0,t))$$

respectively. Note that these functions depend on the initial point and time. By the fundamental inequality, we have

$$\Delta = [\Delta x_1(X_0,t)^2 + \cdots + \Delta x_n(X_0,t)^2]^{1/2} \leq e^{kb}|\Delta x_{10}|.$$

Set

$$p_i(X_0, t) = \frac{\Delta x_i(X_0, t)}{\Delta x_{10}}$$

We must show that $\lim_{\Delta x_{10} \to 0} p_i(X_0, t)$ exists. The idea is as follows. Since p_i is a constant times the difference of solutions, differentiating it would give us a relation in terms of f_i. We will show that p_i is an approximate solution to another system of differential equations and that these p_i converge to the solution of that system.

To this end, differentiating Δx_i gives

$$\begin{aligned}\frac{\partial \Delta x_i}{\partial t}(X_0, t) &= f_i(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n, t) - f_i(x_1, \ldots, x_n, t) \\ &= \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x_1, \ldots, x_n, t)\Delta x_j(X_0, t) + \eta_i(x_1, \ldots, x_n, t)\Delta\end{aligned}$$

where η_i is the error function. A couple of things worth mentioning. Our hypothesis is that each f_i has partial derivatives in x_j which are continuous as functions of X, t, since R is compact, f satisfies the hypothesis of the lemma. Secondly, note that all functions on the right are functions of the initial point and time.

By the lemma above, $\eta_i \to 0$ uniformly on R as $\Delta \to 0$. Here although the x_is depend on the initial value and t, we know that as long as Δ is small, η_i is also small, regardless of what $(x_1, \ldots, x_n, t)$ is. However, Δ depends on the initial point and t. This is where the fundamental inequality helps because we can bound Δ uniformly by Δ_{10}.

So, as $\Delta_{10} \to 0, \Delta \to 0$. So, independent of where $(x_1, \ldots, x_n)$ is (which depends on X_0, t), $\eta_i(x_1, \ldots, x_n, t)$ goes to 0 as well. Dividing by Δx_{10} we get

$$\frac{\mathrm{d}p_i}{\mathrm{d}t}(X_0, t) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x_1, \ldots, x_n, t)p_j + \eta_i \frac{\Delta}{\Delta x_{10}}.$$

So, if we fix an initial point $X_0, \Delta x_{10}$, then p_i (now a function of t alone) are an approximate solution of the system

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x_1(t), \ldots, x_n(t), t)q_j \tag{2}$$

where now x_i are treated as a function of time, with initial condition

$$q_i(t_0) = \begin{cases} 1 & i = 1 \\ 0 & i \neq 1 \end{cases}$$

It is easy to see that the derivatives of q_i, which are linear in q with coefficients depending on t, are continuous in q_i, t and satisfy the Lipschitz condition on q_is because the derivatives of f are continuous , hence bounded, on R. Therefore a solution exists with the given initial conditions, moreover it exists on $|t - t_0| \leq a$ because it is a linear system.

Moreover, $p_i(X_0, t)$ satisfy the same system with error less than

$$n \max_i(\eta_i(x_1(X_0, t), \ldots, x_n(X_0, t), t)\frac{\Delta}{\Delta x_{10}})$$

which goes to 0 uniformly, independent of X_0, t, hence can be made arbitrarily small.

Having fixed X_0, because η_is go uniformly to zero over R as $\Delta x_{10} \to 0$ and $\Delta/\Delta x_{10}$ is bounded uniformly on R, it follows that as functions of t on $|t - t_0| \leq a$, the approximate solutions p_i converge to a solution of 2. Therefore, for a fixed X_0, the limit

$$\lim_{\Delta x_{10} \to 0} p_i(t)$$

exists, which means that the $(x_1, \ldots, x_n)$ are differentiable as functions of x_{10}. Note importantly that the limit of these p_is exist *uniformly* on R, i.e., the derivative quotient converges at the same rate at all points of R. We will prove the continuity of these derivatives later. □

We spend some time understanding the system 2. On R, let $X(X_0,t) = (x_1(X_0,t),\dots,x_n(X_0,t))$ be a solution of $dX/dt = F(X,t)$. Then, the partial derivative of x_i with respect to the first coordinate x_{10} of X_0 satsify

$$\frac{\mathrm{d}q_i}{\mathrm{d}t}(X_0,t) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x_1(X_0,t),\dots,x_n(X_0,t))q_i(X_0,t)$$

with initial condition

$$\frac{\mathrm{d}q_i}{\mathrm{d}t}(X_0,t_0) = \begin{cases} 1 & i=1 \\ 0 & i=0 \end{cases}$$

First let us look at the initial condition. Note that it doesn't depend on the initial point X_0. Geometrically, consider two points $X_0, X_0 + (\Delta x_{10}, 0, \dots, 0)$. What we are doing is to take curves passing through these two points at t_0, and looking at the difference of coefficients. As the two curves approach t_0, the first coordinate will depend linearly on the change Δx_{10} while the other coordinates shouldn't change much because the initial points only differ in the first coordinate. This geometric picture is translated into the initial conditions above.

We ask where is the system 2 defined. The system $(q_1,\dots,q_n)$ have time derivatives of the form $a_1q_1 + \cdots + a_nq_n$ where a_i are functions of X_0,t. These q_i are themselves functions of X_0,t, however their dependence on X_0 is not as direct as their dependence on t. For a fixed X_0, this is a system defined on $\mathbb{R}^n \times [t_0 \pm a]$, and that is where our solutions lie. The x_i are continuous, differentiable functions $x_i \colon R \to B$ where $B = |X - \tilde{X}_0| \leq b$. The indirect dependence of the q_i on X_0 is captured in the following.

Theorem 9. *Consider a system of differential equations in which the functions f_i depend upon any number of parameters $\mu_1,\dots,\mu_m$*

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = f_i(x_1,\dots,x_n;\mu_1,\dots,\mu_m;t), 1 \leq i \leq n.$$

If each f_i has partial derivatives with respect to $x_1,\dots,x_n;\mu_1,\mu_2,\dots,\mu_m$ continuous in some $(n+m+1)$ -dimensional region R, then the solutions

$$x_i(x_{10},\dots,x_{n0};\mu_1,\dots,\mu_m;t), 1 \leq i \leq n$$

will have partial derivatives in $\mu_1,\dots,\mu_m$ continuous in all their arguments through whatever part of R the solutions are defined.

Proof. The idea is to simply treat the parameters as constant functions of t. So, along with the given equations, consider $d\mu_j/dt = 0, 1 \leq j \leq m$. This new system of $n+m$ variables has a solutions with initial value $(x_{10},\dots,x_{n0},\mu_1,\dots,\mu_m)$ and is continuous with respect to the initial variables. Moreover, by the theorem above, the solutions are differentiable with respect to each coordinate of the initial condition, in particular the μ_js. □

Continuing with the solutions to 2, applying the theorem above, it follows that q_i are continuous as functions of X_0, therefore the solution $(x_1,\dots,x_n)$ is $\mathcal{C}^1$ with respect to the initial value. We can go further. If the f_i are $\mathcal{C}^2$, then because x_i are $\mathcal{C}^1$ with respect to X_0, q_i are differentiable with respect to $x_{10},\dots,x_{n0}$. From here, it is clear how by increasing the number of variables we can conclude that the solution to $dX/dt = F(X,t)$ is $\mathcal{C}^k$ when F is $\mathcal{C}^k$ with respect to X.

More directly, what we have proved above is that when F is $\mathcal{C}^1$ with respect to the initial conditions, the collection $(x_1,\dots,x_n,q_1,\dots,q_n)$, where x_i, q_j are as above, satisfy a system of differential equations which is $\mathcal{C}^1$ if F is $\mathcal{C}^2$. This is because the x_i have continuous first order partial derivatives. Naturally, a repetition of this process increases the number of variables exponentially, but proves the result for all $\mathcal{C}^k, k < \infty$. The extension to smooth functions is then obvious.

5 Continuous extensions of solutions

This section considers some boundary properties of solutions and can be skipped. Suppose an integral curve is defined in some domain D over the interval $t_0 - l < t_0 < t_0 + m$. We consider what the solution looks like at the end points. Intuitively, they should approach the boundary of D as otherwise we can extend the solution further.

Theorem 10. *Let D be an arbitrary bounded domain (open or closed) in $\mathbb{R}^{n+1}$ and let $F(X,t)$ be continuous in D and satisfy a Lipschitz condition in X locally in D, i.e., aroud each point of D, there is a neighbourhood where the restriction of F is Lipshcitz in X. Suppose X is a solution to $dX/dt = F(X,t)$ passing through (X_0,t_0) defined on the right of t_0 till t_0+m which cannot be extended further. Then, if $p(t)$ is the distance of the point $(X(t),t)$ from the boundary C of D, then*

$$\lim_{t\to t_0+m} p(t) = 0.$$

Note that when F is locally Lipschitz in X and D is compact, F is Lipschitz in X.

Proof. Let $\epsilon > 0$ be given and S be the set of those points in D which have a distance $\geq \epsilon$ from C. It is bounded because D is bounded. Suppose as $t \to t_0 + m$, there are infinitely many points in S, so we get an increasing sequence $t_n \to t_0+m$ such that $P_n = (X(t_n), t_n) \in S$.

Since S is bounded, these P_n have a limit point $\tilde{P} = (\tilde{X}, \tilde{t})$ which is in the interior of D as it can't be in C or in its exterior. Note in particular that $\tilde{t} = t_0 + m$.

Now, there is a rectangle R around $\tilde{P}$ of the form $|X - \tilde{X}| \leq b, |t - \tilde{t}| \leq a$ contained in D. Let M be the bound on $|F|$ in R (finite because F is continuous, R is compact). Furthermore, we can shrink R itself so that F is Lipschitz in X over R.

The idea is to obtain a P_n very close to $\tilde{P}$ in R_1 and then extend the solution to beyond t_0+m. Recall that if we had a rectangle of the form $|X - X_0| \leq b, |t-t_0| \leq a$, then a solution passing through (X_0, t_0) exists on the interval $|t - t_0| \leq h = \min(a, b/M)$.

So, given t_n, to extend it beyond t_0+m, we would like $h > t_0 + m - t_n$. To this end, we first choose P_n such that

$$|\tilde{X} - X(t_n)| \leq b/2, \tilde{t} - t_n \leq a/2.$$

Once we have that, we would have the rectangle

$$|X - X(t_n)| \leq b - |\tilde{X} - X(t_n)|, |t - t_n| \leq a - (\tilde{t} - t_n)$$

contained in R, hence in D and $h = \min(a - (\tilde{t} - t_n), (b - |\tilde{X} - X(t_n)|)/M)$.

Therefore, we initially shrink a so that $a < b/2M$, then $h = a - (\tilde{t} - t_n) > a/2$. Since $\tilde{t} - t_n < a/2$, it is clear that a solution passing through P_n extends beyond t_0+m.

Now we have two solutions passing through P_n defined on a (closed) neighbourhood of t_n, hence they must agree by the uniqueness theorem. Therefore, the original solution can be extended beyond $t_0 + m$ which is contrary to the assumption.

Thus, as $t \to t_0 + m$, the solution must eventually have no points in S. Since $\epsilon > 0$ was arbitrary, the solution must have limit points in the boundary.

Note that since D is bounded, as $t \to t_0 + m$ we always have limit points. □

Remark. What we have shown is that the limit points as $t \to t_0 + m$ (which exist because D is bounded) must all lie in C. It does not mean that there is a single limit point. Moreover, many different solutions may converge to the same limit point in the boundary and this doesn't contradict the uniqueness because the solution is not really defined at the limit point. As an example, restrict the two solutions in the example above on the region $x > 0$ (here f is seen to be locally Lipschitz).

Theorem 11. *If R is the domain $\mathbb{R}^n \times (t_1, t_2)$ and if $F(X,t)$ satisfies a Lipschitz condition uniformly in every subdomain of the type $\mathbb{R}^n \times [t_1', t_2'], t_1 < t_1' < t_2' < t_2$, then the solution of $dX/dt = F(X,t)$ passing through any point of R may be extended throughout the entire open interval (t_1, t_2).*

Proof. Fix a point (X_0, t_0) and look at the maximal interval around it where a solution exists. Let it be (a, b). If $a > t_1$ or $b < t_2$, then one can extend the solution slightly by hypothesis, and patch up the two solutions by uniqueness. Therefore, the solution must exist on (t_1, t_2). Alternately, following [1], one can argue as follows.

Suppose such a solution $X(t)$ cannot be extended beyond some $t_0 < \bar{t} < t_2$. Fix some $\epsilon > 0$ such that $\bar{t} - \epsilon > t_0$, and on $[t_0, \bar{t} - \epsilon]$ consider the approximate solution given by $\tilde{X}(t) \equiv X_0$. The error of approximation is

$$\max_{[t_0, \bar{t}-\epsilon]} |\frac{\mathrm{d}\tilde{X}}{\mathrm{d}t} - F(X_0, t)| \leq \max_{[t_0, \bar{t}]} |F(X_0, t)| = M$$

where a maximum M exists because $F(X_0, t)$ is continuous on the compact set $[t_0, \bar{t}]$.

By the fundamental inequality, we have

$$|X(t) - X_0| \leq \frac{M}{k} e^{k(\bar{t}-t_0)-1}, t_0 \leq t \leq \bar{t} - \epsilon.$$

Note that the bound is independent of ϵ, so X is bounded to the left of $\bar{t}$, hence by the theorem above, it must approach the boundary of R (for we are considering a maximal solution). However, this is impossible as $\bar{t} < t_2$. Therefore, the solution must exist till t_2. Similarly, the solution must exist to the left of t_0 till t_1. □

6 Vector fields

A note to the reader : Since the following sections are based from a variety of sources there may be certain inconsistencies in the notation. Specifically, the flow along a vector field X may be written as X_t or ϕ_t. However, within each proof, I have tried to be consistent with the notation, so it shouldn't be that much of a problem.

Given a smooth manifold M, the tangent vectors at a point are derivations on the algebra of germs of smooth functions at that point. This collection of derivations forms a vector space (of the same dimension as M) and is called the tangent space. The tangent bundle TM is obtained by taking the disjoint union of all tangent spaces. Using the projection map $\pi \colon TM \to M$, we can topologize TM to arrive at a smooth manifold TM such that π is a smooth map.

Since the tangent space at each point is contractible, TM, M are actually homotopy equivalent. We have the inclusion map $i \colon M \to TM$ with inverse given by π and $i \circ \pi$ is homotopic to the identity map via

$$\begin{aligned} F \colon TM \times I &\to TM \\ ((p, v), t) &\mapsto (p, (1-t)v) \end{aligned}$$

(check smoothness locally). Therefore, from the point of homotopy theory TM, M are identical.

A smooth vector field on an open set U is a section $X \colon M \to TU$ of the projection map. Given a curve $c \colon I \to M$, its derivative is a vector field on $c(I)$. When talking about smooth curves on closed intervals we always assume that the curve can be extended (smoothly) to some neighbourhood of the closed interval, that way there is no issue of differentiability at the end points.

The question we are interested is the inverse. Given a vector field, can we find curves, passing through some fixed point say, whose derivatives agree with the given vector field. Note here that the derivative of a curve at a point is a linear map from one tangent space to another, since the tangent bundle of the interval is trivialised, we can talk of the derivative as assiging a single tangent vector to each point on the image of the curve.

Locally, after going to a coordinate chart, we are trying to solve a system of differential equations. Let $p \in M$ and $(U, x_1, \ldots, x_n)$ be a chart around p where the vector field is given by some $a_1 \partial/\partial x_1 + \cdots + a_n \partial/\partial x_n$, then the curve is going to be of the form $(y_1(t), \ldots, y_n(t))$ such that

$$\frac{\mathrm{d}y_i}{\mathrm{d}t} = a_i(y_i(t)).$$

So, locally it is a question in ordinary euclidean space and we know how to solve it. We are then interested in maximal solutions, curves which can go outside coordinate charts. Any solution will be called an integral curve to the vector field. We have the following theorem from [2].

Theorem 12. *Let X be a $\mathcal{C}^\infty$ vector field on a differentiable manifold M. For each $m \in M$ there exists $a(m), b(m) \in \mathbb{R} \cup \{\pm\infty\}$ and a smooth curve*

$$\gamma_m \colon (a(m), b(m)) \to M$$

such that

(*a*) *$0 \in (a(m), b(m))$ and $\gamma_m(0) = m$.*
(*b*) *γ_m is an integral curve of X.*
(*c*) *If $\mu \colon (c, d) \to M$ is a smooth curve satisfying (a), (b), then $(c, d) \subseteq (a(m), b(m))$ and $\mu = \gamma_m|_{(c,d)}$.*

We define the following

Definition 3. *For each $t \in \mathbb{R}$ define a transformation X_t with domain*

$$\mathcal{D}_t = \{m \in M : t \in (a(m), b(m))\}$$

by setting $X_t(m) = \gamma_m(t)$.

(*d*) *For each $m \in M$, there exists an open neighbourhood V of m and an $\epsilon > 0$ such that the map*

$$(t, p) \mapsto X_t(p)$$

is defined and is $\mathcal{C}^\infty$ from $(-\epsilon, \epsilon) \times V \to M$.
(*e*) *$\mathcal{D}_t$ is open for each t.*
(*f*) *$\cup_{t>0} \mathcal{D}_t = M$.*
(*g*) *$X_t \colon \mathcal{D}_t \to \mathcal{D}_{-t}$ is a diffeomorphism with inverse X_{-t}.*
(*h*) *Let s, t be real numbers. Then the domain of $X_s \circ X_t$ is contained in, but not generally equal to, $\mathcal{D}_{s+t}$. However, the domain of $X_s \circ X_t$ is $\mathcal{D}_{s+t}$ in the case in which s, t both have the same sign. Moreover, on the domain of $X_s \circ X_t$, we have $X_s \circ X_t = X_{s+t}$.*

Before going for the proof, let us discuss what the theorem is saying. We have a vector field X. From earlier considerations we know that we can find local solutions. Moreover, by smooth dependence on initial conditions we know that locally the curves depend smoothly on the initial parameter. Now we can consider the maximal possible solution (because we know that solutions exist). The operation X_t sends $\mathcal{D}_t$ forward (or backward) in time along X. And statement h says that going ahead $s + t$ units is the same as going t units and then s units of time. As an example, for a particular vector field on the sphere, we would get a slowly rotating ball, X_t would be the family of rotations parametrised by time.

Proof. We take $(a(m), b(m))$ to be the maximal interval around 0 such that a solution exists passing through m at $t = 0$. It is non empty by the existence theorem (and this proves (f)). Moreover, by the uniqueness theorem (c) follows.

Fix a point m and let $(U, x_1, \ldots, x_n)$ be a coordinate chart around m. We translate everything to U, so we have smooth functions $a_1, \ldots, a_n$ defined on U which we can treat as being defined on $U \times \mathbb{R}$. Now, by using previous theorems and shrinking U to a rectangle if necessary, we know that there is some $\epsilon > 0$ and a neighbourhood V of m where a solution exists on $V \times (-\epsilon, \epsilon)$ and is smoothly dependent on the initial point. Lifting everything back to M proves (d).

Let $t \in (a(m), b(m))$, then $s \mapsto \gamma_m(t + s)$ is an integral curve of X on the maximal domain $(a(m) - t, b(m) - t)$ with initial condition $0 \mapsto \gamma_m(t)$. The interval is maximal because otherwise we could extend $(a(m), b(m))$. It then follows, by (c), that

$$(a(m) - t, b(m) - t) = (a(\gamma_m(t)), b(\gamma_m(t)))$$

and for s in the above interval,

$$\gamma_{\gamma_m(t)}(s) = \gamma_m(t + s).$$

Now, if m is in the domain of $X_s \circ X_t$, then $t \in (a(m), b(m))$ and from the above, $s \in (a(m) - t, b(m) - t)$, i.e., $s + t \in (a(m), b(m))$. Thus, $m \in \mathcal{D}_{s+t}$ and we have $X_s(X_t(m)) = X_{s+t}(m)$. When s, t have the same sign and $m \in \mathcal{D}_{s+t}$, then we immediately have $t \in (a(m), b(m))$, hence $s \in (a(\gamma_m(t)), b(\gamma_m(t)))$, hence m is in the domain of $X_s \circ X_t$.

Parts (e), (g) are trivial for $t = 0$, so we assume $t > 0$ (the negative case is similar). Let $m \in \mathcal{D}_t$. From (d), for every $p \in \gamma_m([0,t])$, there is a neighbourhood V_p of p and an $\epsilon_p > 0$ such that $X_s(q)$ is defined on $(-\epsilon_p, \epsilon_p) \times V_p$. Cover $\gamma_m([0,t])$ by these V_p and obtain a finite subcover whose union is W and take the minimum of the ϵs associated to the finite subcover to get an $\epsilon > 0$ such that $X_s(q)$ is defined on $(-\epsilon, \epsilon) \times W$.

What this means is that we can smoothly move ϵ units of time for every point in W. Now, we need to obtain a neighbourhood around m where we can move ahead t units of time. To this end, obtain an n large enough so that $t/n < \epsilon$. The idea is that we move ahead by t/n units and if we are still in W, then we can move ahead some more.

Note that $X_{t/n}$ is defined on W and is a smooth function to M. Now, set $\alpha_1 = X_{t/n}|_W$ and $W_1 = \alpha_1^{-1}(W)$. Then for $i = 2, \dots, n$ we inductively define

$$\alpha_i = X_{t/n}|_{W_i} \text{ and } W_i = \alpha_i^{-1}(W_{i-1})$$

(alternately one can just take the inverse under $X_{t/n}$ and intersect the inverse with W_{i-1}). Each α_i is $\mathcal{C}^\infty$ and hence W_n is an open subset of W containing m (it contains m as $\gamma_m(t) \in W$), and by part (h),

$$\alpha_1 \circ \cdots \circ \alpha_n|_{W_n} = X_t|_{W_n}.$$

Consequently, $W_n \subseteq \mathcal{D}_t$, hence $\mathcal{D}_t$ is open proving (e).

Note that above we have obtained X_t as a composition of smooth maps locally, hence X_t is a smooth map for all $t \in \mathbb{R}$. By part (h), X_t, X_{-t} are inverses of each other on their respective domains. Therefore, X_t maps $\mathcal{D}_t$ diffeomorphically onto $\mathcal{D}_{-t}$ proving part (g). □

We do not always have the domain of $X_s \circ X_t$ as $\mathcal{D}_{s+t}$, consider for example $M = (0, \infty)$ and the vector field d/dt, then from each point we can move indefinitely to the right, but only for a finite time to the left, therefore the domain of composition of $X_t \circ X_{-t}$ is not the whole space for $t > 0$, but $X_0 = (0, \infty)$.

Definition 4. *A smooth vector field X on M is complete if $\mathcal{D}_t = M$ for all t. In this case, the transformations X_t form a group of transformations of M parametrized by the real numbers called the1-parameter group of X.*

Theorem 13. *Smooth vector fields on compact manifolds are complete.*

Proof. We repeat an argument from the previous theorem. Let M be a compact manifold and X a smooth vector field, around every point $m \in M$ there is a neighbourhood W_m and an $\epsilon_m > 0$ such that a solution to X exists on $(-\epsilon_m, \epsilon_m) \times W_m$. Cover M by W_m and obtain a finite subcover, then it follows that there is an $\epsilon > 0$ such that a solution exists on $(-\epsilon, \epsilon) \times M$. Given $t > 0$ (the negative case is similar), we take an n large enough so that $t/n < \epsilon$. Now each point can be moved ahead t/n units, composing n times we see that $\mathcal{D}_t = M$, i.e., X is complete. □

Given a smooth vector field X on a manifold M, we say that a point $p \in M$ is regular if $X_p \neq 0$ and singular otherwise.

Theorem 14. (*Canonical form for a Regular Vector Field*) *Let X be a smooth vector field on a manifold M and let $p \in M$ be a regular point. There exist coordinates $(x_1, \dots, x_n)$ on a neighbourhood of p where X is equal to $\partial/\partial x_1$.*

Proof. What this means is that the diffeomorphism provided by the chart makes the x_1-lines into the flows of X. Let $(U, y_1, \dots, y_n)$ be a chart near p with p corresponding to origin and by a suitable linear transformation we may assume that $X_p = \partial/\partial y_1$. There is a neighbourhood U_0 of p and an $\epsilon > 0$ such that a flow Φ of X is defined on $(-\epsilon, \epsilon) \times U_0$ and its image lies in U.

Let S be the open subset of $\mathbb{R}^{n-1}$ obtained by projecting U_0 onto the last $n-1$ coordinates and define the smooth map

$$\begin{aligned} \psi \colon (-\epsilon, \epsilon) \times S &\to U \\ (t, q) &\mapsto \Phi(t, (0, q)) \end{aligned}$$

Now,

$$\begin{aligned}(\psi_* \frac{\partial}{\partial t}|_{(t_0,x_0)})f &= \frac{\partial}{\partial t}|_{(t_0,x_0)}(f\circ\psi)\\ &= \frac{\partial}{\partial t}|_{t=t_0} f(\Phi(t,(0,x_0)))\\ &= X_{\psi(t_0,x_0)}f\end{aligned}$$

because Φ is the flow corresponding to X. At the same time we have

$$\psi(0,x_2,\ldots,x_n) = \Phi(0,(0,x_1,\ldots,x_n)) = (0,x_2,\ldots,x_n)$$

therefore,

$$\psi_* \frac{\partial}{\partial x_i}|_{(0,0)} = \frac{\partial}{\partial y_i}|_{(0,0)}$$

hence at the origin, ψ_* maps a basis to a basis, hence applying the inverse function theorem we obtain a diffeomorphism $\psi\colon W\to Y$ where W is a neighbourhood of $(0,0)$ and Y is a neighbourhood of p. The inverse is a chart where the pullback of $\partial/\partial t$ is X thus completing the proof.

Geometrically, we take the horizontal slice of U_0 and send all the points in this slice "upwards" along X (imagining x_1 to be upwards). Then, a neighbourhood of p is parameterised by it's position in the horizontal slice and a time coordinate along the flows of X. Intuitively, ψ takes the horizontal slice and translates it along X (which we are imaginging as not lying on the horizontal slice). □

The interesting things happen around singular points. When p is a regular point, the flow behaves like a steady stream moving in the x_1 direction, but when it is singular, while p is fixed, points around it may swirl around p, asymptotically approach p or go off to infinity and so on. This sort of behaviour is studied under the name of smooth dynamical systems.

Now, let X be a smooth vector field on a manifold M and fix a point $p\in M$. There is a neighbourhood $(-\epsilon,\epsilon)\times V$ of $(0,p)$ and a map $\Phi\colon(-\epsilon,\epsilon)\times V\to M$ which defines the flow corresponding to X, i.e., for each $q\in V$, the curve $t\mapsto\Phi(t,q), t\in(-\epsilon,\epsilon)$ is an integral solution to X passing through q. We study the derivative of Φ.

For a fixed t, denote by j_t the inclusion $V\hookrightarrow\{t\}\times V$. Then, $X_t=\Phi\circ j_t, t\in(-\epsilon,\epsilon)$. By shrinking V if necessary we may assume that V has coordinates and that Φ lands in an open set with the same coordinates (for example, obtain V and consider $\Phi^{-1}(V)$). Then in local coordinates (we don't need explicit coordinates, but it helps to think in terms of coordinates), the derivative of Φ is of the form

$$\begin{bmatrix} X(\Phi(t,q)) & DX_t(q)\end{bmatrix}_{n+1\times n}$$

Fix a time t and suppose $t\pm\delta\subset(-\epsilon,\epsilon)$. Consider the map

$$\begin{aligned}F\colon(-\delta,\delta)\times W &\to M\\ (t',q) &\mapsto \Phi(t+t',q)\end{aligned}$$

where W is some smaller neighbourhood so that F lands in V. This is a well defined, smooth map and equals $\Phi\circ j_t\circ\Phi$ on $(-\delta,\delta)\times V$ because $X_t\circ X_{t'}=X_{t+t'}$ (we need to shrink V so that the composition is defined).

The image of the vector $\partial/\partial t$ at $(0,p)$ under F is going to be $X(\Phi(t,p))$, while the derivative taken along the composition $\Phi j_t\Phi$ is going to be, by chain rule, $DX_t(p)(X(p))$. Therefore, we have

$$(X_t)_*(p)(X(p)) = X(\Phi(t,p))$$

in other words, the push forward of $X(p)$ along the flow X_t is going to be $X(X_t(p))$. Since t,p were arbitrary, the result holds at all other times and points also. Next we look at the push forward of other vectors. In order to do so, we introduce Lie Brackets.

6.1 Lie Bracket

On $\mathbb{R}^n$ we know that the partial derivatives commute on smooth functions. On manifolds, it doesn't make sense to compose two derivatives at a point because derivatives act on functions. Therefore, we must consider vector fields acting on functions to obtain other functions, and in fact this is what happens over $\mathbb{R}^n$ as well, where we consider constant vector fields.

Given two vector fields X, Y on a manifold, and a smooth function $f\colon M \to \mathbb{R}$, we have a smooth function Xf. We can then act Y on Xf to get $Y(Xf)$. In this way we can compose two vector fields, however YX is not itself a vector field. We look at the commutator

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

called the Lie Bracket of X, Y.

It is easy to verify that $[X, Y]$ is another vector field and we have (see [2])

(a) $[X, Y]$ is a smooth vector field on M.
(b) If $f, g \in \mathcal{C}^\infty(M)$, then $[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X$.
(c) $[X, Y] = -[Y, X]$.
(d) (Jacobi identity) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ for all smooth vector fields X, Y, and Z on M.

Suppose we have a coordinate chart $(U, x_1, \ldots, x_n)$ and two vector fields given by

$$X = \sum a_i \frac{\partial}{\partial x_i}, Y = \sum b_i \frac{\partial}{\partial x_i}.$$

Now given a germ f we compute the derivative at a point p as follows

$$\begin{aligned}
[X, Y]_p(f) &= X_p(Yf) - Y_p(Xf) \\
&= \sum_i a_i(p) \frac{\partial}{\partial x_i}\bigg|_p \left(\sum_j b_j \frac{\partial f}{\partial x_j} \right) - \sum_i b_i(p) \frac{\partial}{\partial x_i}\bigg|_p \left(\sum_j a_j \frac{\partial f}{\partial x_j} \right) \\
&= \sum_{i,j} \left(a_i(p) \frac{\partial b_j}{\partial x_i}(p) - b_i(p) \frac{\partial a_j}{\partial x_i}(p) \right) \frac{\partial f}{\partial x_j}
\end{aligned}$$

where we have used the fact that for smooth functions (or more generally $\mathcal{C}^2$ functions) the double derivatives commute (i.e., the lie bracket of the constant vectors in regular euclidean space is zero). Note in particular what happens if Y is a constant vector for example, i.e., if the b_is are constant.

Definition 5. *Let $\varphi\colon M \to N$ be $\mathcal{C}^\infty$. Smooth vector fields X on M and Y on N are called φ-related if $d\varphi \circ X = Y \circ \varphi$. In other words, Y is the push forward of X.*

Theorem 15. *Let $\varphi\colon M \to N$ be $\mathcal{C}^\infty$. Let X and X_1 be smooth vector fields on M and let Y and Y_1 be smooth vector fields on N. If X is φ-related to Y and X_1 is φ-related to Y_1, then $[X, X_1]$ is φ-related to $[Y, Y_1]$.*

Proof. Let $m \in M$ and f a $\mathcal{C}^\infty$ germ at $\varphi(m)$. We need to show that

$$d\varphi([X, X_1]_m)(f) = [Y, Y_1]_{\varphi(m)}(f).$$

Just applying the definitions, we have

$$\begin{aligned}
d\varphi([X, X_1]_m)(f) &= [X, X_1]_m(f \circ \varphi) \\
&= X_m(X_1(f \circ \varphi)) - X_1|_m(X(f \circ \varphi)) \\
&= X_m((d\varphi \circ X_1)(f)) - X_1|_m((d\varphi \circ X)(f)) \\
&= X_m(Y_1(f) \circ \varphi) - X_1|_m(Y(f) \circ \varphi) \\
&= d\varphi(X_m)(Y_1(f)) - d\varphi(X_1|_m)(Y(f)) \\
&= Y_{\varphi(m)}(Y_1(f)) - Y_1|_{\varphi(m)}(Y(f)) \\
&= [Y, Y_1]_{\varphi(m)}(f)
\end{aligned}$$

□

More generally what this means is that the push forward of the lie bracket of two vector fields is the lie bracket of the push forward of the vector fields.

Now, let X, Y be smooth vector fields on M and fix a point $p \in M$. Let V be a neighbourhood of p such that a flow Φ is defined on $(-\epsilon, \epsilon) \times V$ for the vector field X.

Theorem 16. *If $[X, Y] = 0$, then $\Phi_* Y = Y$, i.e., for all t, $X_{t*}(Y) = Y$.*

Proof. We know that at time $t = 0$ $\Phi_* Y = Y$ because the flow is the identity map. The idea is to compute the time derivative of $\Phi_* Y$ and show that it is zero. If we think in terms of coordinates, they are a priori funcctions of time and position and we hope to show that they infact only depend on the position, i.e, $\Phi(t, q)$. This is naturally accomplished with the lie bracket as it is in some sense, the derivative of one vector field along another. However, this requires us to first introduce a time vector. To this end, we first extend Φ to $\tilde{\Phi}$ defined as

$$\begin{aligned} \tilde{\Phi} \colon (-\epsilon, \epsilon) \times V &\to (-\epsilon, \epsilon) \times M \\ (t, q) &\mapsto (t, \Phi(t, q)) \end{aligned}$$

As before we see vector fields on V as vector fields in $(-\epsilon, \epsilon) \times V$ by setting the time coordinate to zero (because the tangent bundle is a product, we use the inclusion map between the two vector fields). We compute the bracket $[\partial/\partial t, \tilde{\Phi}_* Y]$, note that because $\partial/\partial t$ is a constant vector field, this bracket only consists of $\tilde{\Phi}_* Y$ with the coefficients replaced by their time derivatives.

Note that
$$(\tilde{\Phi}_* \frac{\partial}{\partial t})(t, \Phi(t, q)) = \frac{\partial}{\partial t}(t, \Phi(t, q)) + X(\Phi(t, q))$$
i.e., the push forward of the time vector evaluated at the point $(t, \Phi(t, q))$ is the time vector at $(t, \Phi(t, q))$ minus the pushforward of $X(t, q)$. Using $\tilde{\Phi}_* X = X$, we have

$$\begin{aligned} [\frac{\partial}{\partial t}, \tilde{\Phi}_* Y]_{(t, \Phi(t,q))} &= [\tilde{\Phi}_* \frac{\partial}{\partial t} - X, \tilde{\Phi}_* Y]_{(t, \Phi(t,q))} \\ &= \tilde{\Phi}_* [\frac{\partial}{\partial t}, Y]_{(t,q)} - \tilde{\Phi}_* [X, Y]_{(t,q)} \end{aligned}$$

at every point $(t, \Phi(t, q)) \in (-\epsilon, \epsilon) \times M$. Since Y is independent of time, the first term is zero and the second is zero by assumption. So, the bracket $[\partial/\partial t, \tilde{\Phi}_* Y] = 0$ wherever it is defined.

Now, $\tilde{\Phi}_* Y$ is a function of points in the image of $\tilde{\Phi}$ which is an open set because X_t is a diffeomorphism between the open sets $\mathcal{D}_t, \mathcal{D}_{-t}$. Since the time derivative is zero, the coefficients of $\tilde{\Phi}_* Y$ at any (t, q_1) in the image of $\tilde{\Phi}$ is independent of t and equals the value at $(0, q_1)$ which is $Y(q_1)$. It then follows that for every time t, $X_{t*} Y = Y$ as required. □

For another proof, see [3]. To recap, the idea is to show that the pushforward of Y is independent of time, hence the curve so described in $T_{q_1} M$ is constant. Since at $t = 0$, X_t is identity, this means that the pushforward is Y itself.

Note also that above we have made precise the statement that the lie bracket is the derivative of one vector field along another. Specifically we have from the above calculations

$$\left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} (X_{t*} Y) = -[X, Y].$$

Next we show that the flows of commuting vector fields themselves commute. If we have two commuting vector fields X, Y with flows ϕ_t, ψ_t respectively we show that $\phi_t \circ \psi_s = \psi_s \circ \phi_t$ wherever both sides are defined. Geometrically this encloses a parallelogram, going t along X and s along Y is the same as going along Y first and then X. Before that we prove the following more general result, from [3].

Theorem 17. *Suppose $F \colon M \to N$ is a smooth map, X, Y smooth vector fields on M, N respectively with corresponding flows ϕ, ψ. Then X and Y are F-related if and only if for each $t \in \mathbb{R}$, $\psi_t \circ F = F \circ \phi_t$ on the domain of ϕ_t.*

Proof. For a fixed point p in the domain of ϕ_t, the point $\psi_t(F(p))$ is the point in the flow starting at $F(p)$ along Y. Let $\mathcal{D}_p \subset \mathbb{R}$ denote the domain of $\phi^{(p)}$, the flow through p along X.

First suppose X, Y are F-related, define $\gamma\colon \mathcal{D}_p \to N$ as $F \circ \phi^{(p)}$, then

$$\begin{aligned}\gamma'(t) &= F_*(\phi(p)'(t)) \\ &= F_*(X_{\phi^{(p)}(t)}) \\ &= Y_{F(\phi^{(p)}(t))} \\ &= Y_{\gamma(t)}\end{aligned}$$

So, γ is a flow along Y starting at $F(p)$ and by the uniqueness result, the maximal flow through $F(p)$ along Y must be defined on at least $\mathcal{D}_p$ and $\gamma(t) = \psi^{(F(p))}(t)$ as required.

Conversely, assume $\psi^{(F(p)}(t) = F(\phi^{(p)}(t)), t \in \mathcal{D}_p$, then for each $p \in M$ we have

$$\begin{aligned}F_* X_p &= F_*(\phi^{(p)'}(0)) \\ &= (F \circ \phi^{(p)})'(0) \\ &= \psi^{(F(p))'}(0) \\ &= Y_{F(p)}\end{aligned}$$

which shows that X, Y are F-related. □

Theorem 18. *Let M be a smooth manifold, X, Y smooth vector fields on M with flows φ, ψ defined on some neighbourhood of $(0, p), p \in M$. Then X, Y commute if and only if their flows commute, i.e., $\varphi_t \circ \psi_s = \psi_s \circ \varphi_t$ wherever defined.*

Proof. Let us move to a neighbourhood V of p and suppose that both flows are defined on $(-\epsilon, \epsilon) \times V$. Suppose $[X, Y] = 0$, then for a given time t, we know that $\varphi_{t*} Y = Y$, i.e., Y is invariant under φ_t. We consider the map $\varphi_t \circ V \to M$, then it follows that Y on V is φ_t-related to Y on M. By the previous theorem, we know that $\varphi_t \circ \psi_s = \psi_s \circ \varphi_t$ where we are using the same notation ψ_s for flows on V, M because they are flows of the same vector field.

Conversely, suppose the flows commute, then by the previous theorem, we get that Y on V is φ_t-related to Y on M, i.e., $\varphi_{t*} Y = Y$. Hence, the time derivative of $\varphi_{t*} Y$ is zero, which means that the bracket $[X, Y] = 0$ thus the vector fields commute. □

6.2 Distributions and Frobenius Theorem

Definition 6. *Let c be an integer, $1 \leq c \leq d$. A c-dimensional distribution $\mathscr{D}$ on a d-dimensional manifold M s a choice of a c-dimensional subspace $\mathscr{D}(m)$ of M_m for each $m \in M$. $\mathscr{D}$ is smooth if for each $m \in M$ there is a neighbourhood U of m and there are c vector fields $X_1, \ldots, X_c$ of class $\mathcal{C}^\infty$ on U which span $\mathscr{D}$ at each point of U. A vector field X on M is said to belong to (or lie in) the distribution $\mathscr{D}(X \in \mathscr{D})$ if $X_m \in \mathscr{D}(m)$ for each $m \in M$. A smooth distribution $\mathscr{D}$ is called involutive (or completely integrable) if $[X, Y] \in \mathscr{D}$ whenever X, Y are smooth vector fields lying in $\mathscr{D}$.*

Definition 7. *A submanifold (N, ψ) of M is an integral manifold of a distribution $\mathscr{D}$ on M if $d\psi(N_n) = \mathscr{D}(\psi(n))$ for each $n \in N$.*

Here a submanifold refers to a manifold N together with an injective immersion $\psi\colon N \to M$. It is clear that finding an integrable manifold for a given distribution is a generalization of finding integral curves satisfying a given vector field, the only difference being that integral curves are supposed to match the vector field, whereas integrable manifolds need only span the same subspace as the distribution.

Lemma 6. *Let $\mathscr{D}$ be a smooth disctribution on M such that through each point of M there passes an integral manifold of $\mathscr{D}$. Then $\mathscr{D}$ is involutive.*

Proof. Let X, Y be smooth vector fields in $\mathscr{D}$ and let $m \in M$. Let (N, ψ) be an integral manifold of $\mathscr{D}$ through m, and suppose $\psi(n_0) = m$. Since $d\psi\colon N_n \to \mathscr{D}(\psi(n))$ is an isomorphism at each $n \in N$, there exists vector fields $\tilde{X}, \tilde{Y}$ on N such that $d\psi \circ \tilde{X} = X \circ \psi, d\psi \circ \tilde{Y} = Y \circ \psi$. Then, the pushforward of the bracket $[\tilde{X}, \tilde{Y}]$ is the bracket $[X, Y]$, hence $[X, Y]_m \in \mathscr{D}$ and $\mathscr{D}$ is involutive. □

Theorem 19. *(Frobenius) Let $\mathscr{D}$ be a c-dimensional, involutive, $\mathcal{C}^\infty$ disctribution on M^d. Let $p \in M$. Then there exists an integral manifold of $\mathscr{D}$ passing through p.*

Proof. The following proof is taken from [5]. Let $p \in M$ and let (U, φ) be a chart around p. Assume that on U, the vector fields $Y_1, \ldots, Y_r$ form a basis for the given distribution $\mathscr{D}$. Now writing $Y_1, \ldots, Y_r$ in the local coordinates gives us an $r \times n$ matrix of coefficients, by permuting the coordinates if necessary (which amounts to changing the chart) we may assume that at p, the first $r \times r$ minor is non singular, and by continuity, we may shrink U to V so that on V, this minor is non singular.

So, we have the chart $(y_1, \ldots, y_n) = \varphi\colon V \to \mathbb{R}^n$ and a local basis for $\mathscr{D}$ given by $Y_i = \sum a_{ij} \frac{\partial}{\partial y_j}$ (technically, the pushforward of Y_i under φ) with $A(\varphi(q)) = (a_{ij})_{1 \leq i,j \leq r}$ invertible on V. Next, instead of $Y = (Y_1, \ldots, Y_r)$, we look at the collection $X = (X_1, \ldots, X_r) = A^{-1}Y$ which is well defined because A is invertible on all of $\varphi(V)$. Furthermore, since $\mathscr{D}$ is a distribution, each $X_i \in \mathscr{D}$ and since A is invertible, these X_i also form a basis for $\mathscr{D}$.

Since $\mathscr{D}$ is involutive, each bracket $[X_i, X_j] \in \mathscr{D}$, however

$$X_i = \frac{\partial}{\partial x_i} + \sum_{k=r+1}^{n} c_{ik} \frac{\partial}{\partial x_k}, 1 \leq i \leq r$$

for some smooth functions c_{ik} on $\varphi(V)$ therefore,

$$[X_i, X_j] = [\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] + \sum_{k=r+1}^{n} [\frac{\partial}{\partial x_i}, c_{jk} \frac{\partial}{\partial x_k}] + \sum_{k=r+1}^{n} [c_{ik} \frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_j}] + \sum_{k,k'=r+1}^{n} [c_{ik} \frac{\partial}{\partial x_k}, c_{jk'} \frac{\partial}{\partial x_{k'}}].$$

Observe that this is in the span of $\partial/\partial x_k, k \geq r+1$. However, since $\mathscr{D}$ is involutive, it must be a linear combination of $X_1, \ldots, X_r$. But this is possible only if all the brackets are 0, therefore, $X_1, \ldots, X_r$ commute. What we have proved is that given any distribution, we can always find a local basis of *commuting* vector fields.

Let φ_t^i be the flow corresponding to X_i defined in a neighbourhood $(-\epsilon, \epsilon) \times W$ of $(0, p)$ whose image lands in V. Now we define

$$\begin{aligned} \Psi\colon & (-\epsilon, \epsilon)^r \to M \\ & (t_1, \ldots, t_r) \mapsto \varphi_{t_1}^1 \circ \cdots \circ \varphi_{t_r}^r(p) \end{aligned}$$

Since the vector fields X_i commute, we can change the order of composition of the flows. It then follows that the at any $(t_1, \ldots, t_r)$, the derivative of Ψ is injective, because $X_1, \ldots, X_r$ are linearly independent. Thus, Ψ is an immersion and an integral manifold corresponding to the given distribution $\mathscr{D}$. □

A slightly stronger version holds:

Theorem 20. *Let $\mathscr{D}$ be a c-dimensional, involutive, $\mathcal{C}^\infty$ distribution on M^d. Let $m \in M$, then there exists an integral manifold of $\mathscr{D}$ passing through m. Indeed, there exists a cubic coordinate system (U, φ) which is centered at m, with coordinate functions $x_1, \ldots, x_d$ such that the slices $x_i =$ constant, $c+1 \leq i \leq d$ are integral manifolds of $\mathscr{D}$; and if (N, ψ) is a connected integral manifold such that $\psi(N) \subset U$, then $\psi(N)$ lies in one of these slices.*

For a proof, see [2].

References

[1] W. Hurewicz, *Lectures on Ordinary Differential Equations*

[2] F. Warner, *Foundations of Differential Manifolds and Lie Groups*

[3] John M. Lee, *Introduction to Smooth Manifolds*

[4] Notes by Marco Gualtieri, Vector fields

[5] Notes by Andrea Rincon, Frobenius Theorem