DIFFERENTIAL EQUATIONS

SIDDHANT CHAUDHARY

These are my course notes for the course **DIFFERENTIAL EQUATIONS** that I undertook in my fourth semester. Throughout the document the symbol ■ stands for QED.

Contents

Date: January 2021.

1. Ordinary Differential Equations (ODEs)

Definition 1.1. An *ordinary differential equation (ODE)* n is an equation of the form

$$F(t, u(t), u'(t), ..., u^{(n)}(t)) = 0 \tag{$*$}$$

where $F : I \times U \to \mathbb{R}$ is a function, with I an interval in $\mathbb{R}$, U an open subset of $\mathbb{R}^{n+1}$ and the unknown to be found is a function $u : I \to \mathbb{R}$ which satisfies the above equation. The *order* of the given ODE is n.

In this course we assume that the equation in **Definition** 1.1 can be written in the form

$$u^{(n)} = f(t, u, u', ..., u^{(n-1)}) \tag{$**$}$$

Example 1.1. Consider the equation

$$x^2 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 3x\frac{\mathrm{d}y}{\mathrm{d}x} + 4y = 0 \quad , \quad x > 0$$

The above is of the form

$$F(x, y, y', y'') = 0$$

where

$$F(a, b, c, d) = a^2 d + 3ac + 4b$$

If we solve for d in the equation $F = 0$, we get a a DE in the form ($**$).

1.1. Converting an n^{th} order DE to a first order DE.

Consider a DE of the form in ($**$), i.e

$$u^{(n)} = f(t, u, u', ..., u^{(n-1)})$$

Consider a *system* of DEs

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = v_1(t, x_1, ..., x_n)$$
$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = v_2(t, x_1, ..., x_n)$$
$$....$$
$$\frac{\mathrm{d}x_n}{\mathrm{d}t} = v_n(t, x_1, ..., x_n)$$

This is called a *vector valued* differential equation of the first order. Here each x_i is a function on some interval I, and each v_i is a function taking vector valued inputs. We can write the above equation as

$$\frac{\mathrm{d}\boldsymbol{\vec{x}}}{\mathrm{d}t} = \boldsymbol{\vec{v}}(t, \vec{x})$$

where

$$\vec{\boldsymbol{x}} = (x_1, ..., x_n) \quad , \quad \vec{\boldsymbol{v}} = (v_1, ..., v_n)$$

are vector valued functions.

Coming back to our original equation (∗∗), we do the following. Set

$$(***) \qquad \begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}t} &= x_2 \\ \frac{\mathrm{d}x_2}{\mathrm{d}t} &= x_3 \\ &\ldots \\ \frac{\mathrm{d}x_{n-1}}{\mathrm{d}t} &= x_n \\ \frac{\mathrm{d}x_n}{\mathrm{d}t} &= f(t, x_1, x_2, ..., x_{n-1}, x_n) \end{aligned}$$

We see that (∗ ∗ ∗) is equivalent to (∗∗). To show this, suppose

$$\vec{\boldsymbol{\varphi}} = (\varphi_1, ..., \varphi_n)$$

is a solution of (∗ ∗ ∗), where each φ_i is a function of t. Then, φ_1 is a solution of the equation (∗∗). Conversely, if φ is a solution of (∗∗), then

$$\vec{\boldsymbol{\varphi}} = (\varphi, \varphi' ..., \varphi^{(n-1)})$$

is a solution of (∗ ∗ ∗).

Remark 1.0.1. Most of our course will be concerned with such vector valued first order differential equations. So, our standard form will be

$$\frac{\mathrm{d}\vec{\boldsymbol{x}}}{\mathrm{d}t} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}})$$

Infact, as we shall see, it will be *enough* to study equations of the form

$$\frac{\mathrm{d}\vec{\boldsymbol{x}}}{\mathrm{d}t} = \vec{\boldsymbol{v}}(\vec{\boldsymbol{x}})$$

Equations of the above form are called *autonomous differential equations*. These are DEs in which the RHS does not explicitly depend upon the independent variable.

Differential equations often come with initial conditions. These are called *initial value problems (IVP)*. For instance, one could ask the following: Find a function satisfying

$$\frac{\mathrm{d}\vec{\boldsymbol{x}}}{\mathrm{d}t} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}})$$

such that $\vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x_0}}$ for some initial time t_0.

1.2. **Some more terminology.** We will frequently use the dot notation for the derivative. For instance, given a function $\vec{x}$, we have

$$\frac{\mathrm{d}\vec{\boldsymbol{x}}}{\mathrm{d}t} = \dot{\vec{\boldsymbol{x}}}$$

Eventually, we will prove the following statement: if the function $\vec{\boldsymbol{v}}$ is assumed to be $\mathscr{C}^1$, then the IVP

$$\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}}) \quad , \quad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x_0}}$$

has a unique solution in a neighborhood of t_0. We will actually show that there is a *maximal* interval around t_0 on which the solution exists (and since we are working with intervals in $\mathbb{R}$, this interval will be a *maximum*). Let us try to make this more formal. We will show that there is some interval (ω_-, ω_+) containing the point t_0 and some $\vec{\boldsymbol{\varphi}}_{\max} : (\omega_-, \omega_+) \to \mathbb{R}^n$ which is $\mathscr{C}^1$ such that

$$\dot{\vec{\boldsymbol{\varphi}}}_{\max}(t) = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{\varphi}}_{\max}(t))$$

and the interval (ω_-, ω_+) is *maximal* in the following sense: if there is another interval I containing t_0 and a function $\vec{\boldsymbol{\varphi}} : I \to \mathbb{R}^n$ such that

$$\dot{\vec{\boldsymbol{\varphi}}}(t) = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{\varphi}}(t))$$

for each $t \in I$, then

(1) $I \subseteq (\omega_-, \omega_+)$
(2) $\vec{\boldsymbol{\varphi}} = \vec{\boldsymbol{\varphi}}_{\max}|_I$

It should be noted that the function $\vec{\boldsymbol{\varphi}}_{\max}$ depends on *both* t_0 and $\vec{\boldsymbol{x}}_{\mathbf{0}}$, i.e we have some function

$$\vec{\boldsymbol{\psi}}(t_0, x_0, t) = \vec{\boldsymbol{\varphi}}_{\max}(t)$$

The point t_0 is called the *initial time*, and the point $\vec{\boldsymbol{x}}_{\mathbf{0}}$ is called the *initial state*. In general, the $\vec{\boldsymbol{x}}$'s vary in a *state/phase space*, and the t's vary in a *time space*.

1.3. **Integral Curves and Phase Spaces.** This short section will only be about formalising the problem at hand. Let Ω be an open subset of $\mathbb{R}^n$ and let I be an open interval in $\mathbb{R}$. Let $t_0 \in I$ and $\vec{\boldsymbol{x}}_{\mathbf{0}} \in \Omega$. Suppose we have a map

$$\vec{\boldsymbol{v}} : I \times \Omega \to \mathbb{R}^n$$

which is a $\mathscr{C}^1$ map. Then we are interested in the IVP

$$(\dagger) \qquad \dot{\vec{\boldsymbol{x}}} = \frac{\mathrm{d}\vec{\boldsymbol{x}}}{\mathrm{d}t} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}}) \quad , \quad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x}}_{\mathbf{0}}$$

The set Ω is called the *phase space*, and the $\vec{\boldsymbol{x}}$'s are called *states* or *phases*. $I \times \Omega$ is sometimes called the *extended phase space*. A solution is often called an *integral curve*. By a *solution* to this IVP we mean a pair $(J, \vec{\boldsymbol{\varphi}})$ where $J \subseteq I$ is an open interval in $\mathbb{R}$ containing t_0 and

$$\vec{\boldsymbol{\varphi}} : J \to \Omega$$

is a differentiable map such that

$$\dot{\vec{\boldsymbol{\varphi}}}(t) = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{\varphi}}(t)) \quad , \quad \vec{\boldsymbol{\varphi}}(t_0) = x_0$$

for each $t \in J$. The interval J is called an *interval of existence*. Next, suppose $(J, \vec{\boldsymbol{\varphi}})$ is a solution of (†). Consider the map $\vec{\boldsymbol{\psi}}$ given by

$$\vec{\boldsymbol{\psi}}(t) = (t, \vec{\boldsymbol{\varphi}}(t)) = (t, \varphi_1(t), ..., \varphi_n(t))$$

where $(\varphi_1, .., .\varphi_n) = \vec{\boldsymbol{\varphi}}$ (component functions). Let $\hat{\Omega} = I \times \Omega$ be the extended phase space, and let

$$\vec{\boldsymbol{w}} : \hat{\Omega} \to \mathbb{R}^{n+1}$$

be the map given by

$$(s, \vec{\boldsymbol{x}}) \mapsto (1, \vec{\boldsymbol{v}}(s, \vec{\boldsymbol{x}})) = (1, v_1(s, \vec{\boldsymbol{x}}), ..., v_n(s, \vec{\boldsymbol{x}})) \quad , \quad (s, \vec{\boldsymbol{x}}) \in \hat{\Omega}$$

Then, consider the IVP

$$(\bullet) \qquad \dot{\vec{\boldsymbol{\xi}}} = \vec{\boldsymbol{w}}(\vec{\boldsymbol{\xi}}) \quad , \quad \vec{\boldsymbol{\xi}}(t_0) = (t_0, \vec{\boldsymbol{x}}_{\mathbf{0}})$$

Note that

(1) The differential equation (•) is autonomous, i.e the right hand side does not depend upon I.
(2) $\vec{\psi}$ is a solution of (•).
(3) Given a solution $\vec{\psi}$ of (•), say

$$\vec{\psi} = (\psi_0, \psi_1, ..., \psi_n)$$

then we see that $\psi_0(t) = t$ for any $t \in J$ and if $\varphi_i := \psi_i$ then $\vec{\varphi} = (\varphi_1, ..., \varphi_n)$ is a solution of (†).

Note that the DE (•) is an *autonomous* DE. So these observations tell us that to study IVPs of the form (†), it is enough to restrict our attention to *autonomous* DEs, i.e DEs of the form

$$\dot{\vec{x}} = \vec{v}(\vec{x}) \quad , \quad \vec{x}(t_0) = \vec{x_0}$$

1.4. **Autonomous equations when $n = 1$.** Consider the following autonomous IVP (since the dimension is 1, we won't put an arrow):

$$(*) \qquad \dot{x} = v(x) \quad , \quad x(t_0) = x_0$$

where $v : \Omega \to \mathbb{R}$ is a $\mathscr{C}^1$ map on an open interval Ω of $\mathbb{R}$, x_0 is a fixed state in Ω and t_0 a time point.

1.4.1. *Time Reversal.* Let $\varphi : (a, b) \to \Omega$ be a solution of $(*)$. Recall that this implies that $t_0 \in (a, b)$. The map

$$\varphi^{\mathrm{tr}} : (2t_0 - b, 2t_0 - a) \to \Omega$$

given by

$$\varphi^{\mathrm{tr}}(t) = \varphi(2t_0 - t) \quad , \quad 2t_0 - b < t < 2t_0 - a$$

is a solution of the differential equation

$$(*_{\mathrm{tr}}) \qquad \dot{x} = -v(x) \quad , \quad x(t_0) = x_0$$

The IVP $(*_{\mathrm{tr}})$ is called the *time reversal* of $(*)$ (we will see the reason behind the terminology) and the map φ^{tr} is called the *time reversal* of φ.

1.4.2. *State Reversal.* Let $\varphi : (a, b) \to \Omega$ be a solution to $(*)$. Set

$$-\Omega = \{x \in \mathbb{R} \mid -x \in \Omega\}$$

Let

$$v^{\mathrm{sr}} : -\Omega \to \mathbb{R}$$

be the map $v^{\mathrm{sr}}(x) = -v(-x)$. Then v^{sr} is $\mathscr{C}^1$ and $v^{\mathrm{sr}}(-x_0) = -v(x_0)$. Let

$$\varphi^{\mathrm{sr}} : (a, b) \to -\Omega$$

be the map $\varphi^{\mathrm{sr}}(t) = -\varphi(t)$ for $t \in (a, b)$. Then, φ^{sr} is a solution of the IVP

$$(*_{\mathrm{sr}}) \qquad \dot{x} = v^{\mathrm{sr}}(x) \quad , \quad x(t_0) = -x_0$$

The IVP $(*_{\mathrm{sr}})$ is called the *state reversal* of $(*)$ and φ^{sr} is called the *state reversal* of φ.

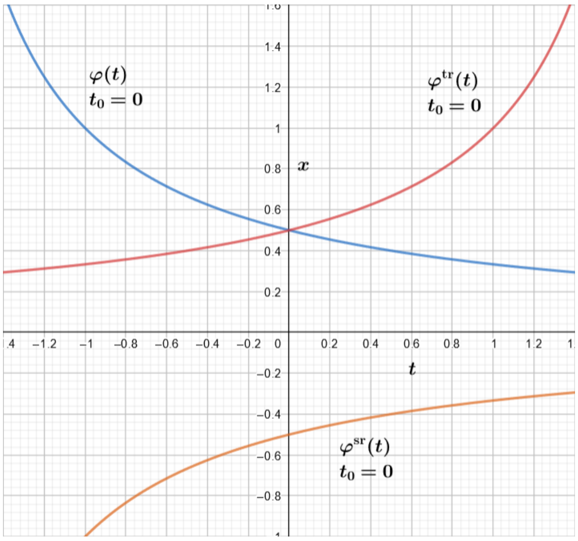


FIGURE 1. The curve in blue is φ, the one in red is φ^{tr} the one in orange is φ^{sr} and the basepoint is $t_0 = 0$. The *time reversal* is like flipping the graph of φ around the axis $t = t_0$. The *state reversal* is like flipping the graph of φ around the $x = 0$ axis.

We now introduce a few more terms related to $(*)$. A state $x \in \Omega$ is said to be *regular* if $v(x) \neq 0$. Otherwise it is called *stationary* or *singular*, i.e x is *stationary* (or *singular*) if $v(x) = 0$. We use the following notation.

$$\Omega^{\text{reg}} := \{x \in \Omega \mid v(x) \neq 0\}$$
$$\Omega^{\text{sing}} := \{x \in \Omega \mid v(x) = 0\}$$

Note that Ω^{reg} is an open subset of Ω, and hence an open subset of $\mathbb{R}$. Also, Ω^{sing} is a closed subset of Ω.

1.4.3. *Solving the equation.* Suppose $x_0 \in \Omega^{\text{reg}}$. Let

$$S_{\text{max}} = (x_m, x_M) \subseteq \Omega^{\text{reg}}$$

be the largest interval containing x_0 which lies in Ω^{reg}. In other words S_{max} is the connected component of the open set Ω^{reg} containing x_0.

Since v is nowhere vanishing on S_{max}, it has a constant sign on it. Let

$$\theta : S_{\text{max}} = (x_m, x_M) \to \mathbb{R}$$

be the function defined by

$$(\star) \qquad \theta(x) = t_0 + \int_{x_0}^{x} \frac{d\xi}{v(\xi)} \quad , \quad x \in S_{\text{max}}$$

Since $v(\xi)$ has a constant sign for $\xi \in S_{\text{max}}$, θ is strictly monotone and hence one-to-one. Moreover, θ is continuous (infact $\mathscr{C}^2$). So let

$$\theta(S_{\text{max}}) = (\omega_-, \omega_+) =: J_{\text{max}}$$

Let $\varphi_{\text{max}} : J_{\text{max}} \to S_{\text{max}}$ be the inverse of θ, i.e $\varphi_{\text{max}} = \theta^{-1}$. The **Inverse Function Theorem** shows that φ_{max} is differentiable and $\mathscr{C}^2$. Now,

$$\theta(\varphi_{\text{max}}(t)) = t$$

and hence

$$\frac{d\theta}{dx}(\varphi_{\max}(t))\dot{\varphi}_{\max}(t) = 1$$

and because $\frac{d\theta}{dx} = \frac{1}{v}$, it follows that

$$\frac{\dot{\varphi}_{\max}(t)}{v(\varphi_{\max}(t))} = 1$$

So, we see that

$$\dot{\varphi}_{\max}(t) = v(\varphi_{\max}(t)) \quad , \quad \varphi_{\max}(t_0) = x_0$$

and hence it follows that $(*)$ has a solution, namely the function $\varphi_{\max} : (\omega_-, \omega_+) \to \Omega$.

1.4.4. *The nature of the solutions.* Having shown that a solution to $(*)$ exists, we will see how solutions to this equation look like.

First, suppose $\varphi : (a, b) \to \Omega^{\text{reg}}$ is a solution of $(*)$. Note that we are requiring φ to take values in Ω^{reg}. We will soon see that every solution of $(*)$ takes values in Ω^{reg}. Let $J = (a, b)$. Because φ is continuous, $\varphi(J)$ is connected and contains x_0. Hence, it follows that $\varphi(J) \subseteq (x_m, x_M) = S_{\max}$ (this is true since $S_{\max}$ is the largest interval in Ω^{reg} containing x_0). Now, we know that $v(\varphi(s)) \neq 0$ for $s \in (a, b) = J$. Also, we know that

$$\dot{\varphi}(s) = v(\varphi(s)) \quad , \quad s \in (a, b)$$

So we have that

$$\frac{\dot{\varphi}(s)}{v(\varphi(s))} = 1 \quad , \quad \forall s \in (a, b)$$

Integrating both sides of the above equation from t_0 to t for some $t \in (a, b) = J$ we get

$$\int_{t_0}^{t} \frac{\dot{\varphi(s)}}{v(\varphi(s))} ds = \int_{t_0}^{t} ds = t - t_0 \quad , \quad t \in J$$

Now use the substitution $\xi = \varphi(s)$ to get

$$\int_{x_0}^{\varphi(t)} \frac{d\xi}{v(\xi)} = t - t_0 \quad , \quad t \in J$$

which we rewrite as

$$t_0 + \int_{x_0}^{\varphi(t)} \frac{d\xi}{v(\xi)} = t \quad , \quad t \in J$$

Now, let θ be the function we defined by equation $(\star)$ in the previous section 1.4.3. The above equation implies that

$$\theta(\varphi(t)) = t$$

Now two things follow from this equation: first is that $J \subseteq (\omega_-, \omega_+)$ since the image of θ is $(\omega_-, \omega_+) = S_{\max}$. Note that $\dot{\varphi} = v \circ \varphi$ which is nowhere vanishing, and hence φ is also one-one. All this implies that φ is an inverse of θ. So, the above equation implies that φ is an inverse of θ and we see that

$$J = (a, b) \subseteq (\omega_-, \omega_+) = S_{\max} \text{ and } \varphi^{-1} = \theta|_J$$

So, it follows that $\varphi = \varphi_{\max}|_{(a,b)}$.

To summarise, if $\varphi : (a, b) \to \Omega^{\text{reg}}$ is a solution of $(*)$, then

$$(**) \qquad \begin{cases} (a, b) \subseteq (\omega_-, \omega_+) \\ \varphi = \varphi_{\max}|_{(a,b)} \end{cases}$$

The statement (∗∗) explains why we used the notation $\varphi_{\max}$. As we mentioned before, we can infact make a stronger statement: if $\varphi : (a, b) \to \Omega$ is a solution of (∗), then φ must take values in Ω^{reg}, and hence (∗∗) has to be true for φ.

1.4.5. *Some Observations.* Let θ and $\varphi_{\max}$ be the functions as in section 1.4.3. θ and $\varphi_{\max}$ being monotone and continuous are homomomorphisms between (x_m, x_M) and (ω_-, ω_+). If $v(x_0) > 0$, they are both strictly increasing, and if $v(x_0) < 0$ they are both strictly decreasing. This gives us the following.

(1) If $v(x_0) > 0$ then

$$\lim_{x\to x_m} \theta(x) = \omega_- \quad , \quad \lim_{x\to x_M} \theta(x) = \omega_+$$

and

$$\lim_{t\to\omega_-} \varphi_{\max}(t) = x_m \quad , \quad \lim_{t\to\omega_+} \varphi_{\max}(t) = x_M$$

(2) If $v(x_0) < 0$ then

$$\lim_{x\to x_m} \theta(x) = \omega_+ \quad , \quad \lim_{x\to x_M} \theta(x) = \omega_-$$

and

$$\lim_{t\to\omega_-} \varphi_{\max}(t) = x_M \quad , \quad \lim_{t\to\omega_+} \varphi_{\max}(t) = x_m$$

Remark 1.0.2. Note that, in our discussion thus far, ω_- and ω_+ are allowed to be $-\infty$ and ∞ respectively. Similar is the case with x_m and x_M.

Suppose, for the sake of definiteness, $v(x_0) > 0$. Hence, in this case point number (1) above applies. We claim that if $x_M \in \Omega$, then $\omega_+ = \infty$. First, note that if $x_M \in \Omega$ then $v(x_M) = 0$, because (x_m, x_M) is the *largest* interval contained in Ω that contains x_0 and v is continuous. Next, let

$$M = \sup\{|\dot{v}(\xi)| \mid \xi \in [x_0, x_M]\}$$

Note that $0 < M < \infty$; M is finite because $[x_0, x_M]$ is a compact set and v is $\mathscr{C}^1$. By the mean value theorem, for each $x \in [x_0, x_M)$ there exists $x^* \in [x, x_M)$ such that

$$v(x) = v(x) - v(x_M) = \dot{v}(x^*)(x - x_M)$$

and using the fact that v does not change sign on (x_m, x_M) we have

$$v(x) = |v(x)| \le M(x_M - x) \quad , \quad x \in [x_0, x_M)$$

This means that

$$\frac{1}{v(\xi)} \ge \frac{1}{M} \cdot \frac{1}{x_M - \xi} \quad , \quad \xi \in [x_0, x_M)$$

Hence for $x \in [x_0, x_M)$ we see that

$$\begin{aligned}\theta(x) = t_0 + \int_{x_0}^{x} \frac{d\xi}{v(\xi)} &\ge t_0 + \frac{1}{M}\int_{x_0}^{x} \frac{d\xi}{x_M - \xi} \\ &= t_0 + \frac{1}{M}\log\frac{x_M - x_0}{x_M - x}\end{aligned}$$

So, it follows that

$$\omega_+ = \lim_{x\to x_M} \theta(x) \ge t_0 + \frac{1}{M}\lim_{x\to x_M} \log\frac{x_M - x_0}{x_M - x} = \infty$$

Lemma 1.1. *Suppose $v(x_0) \ne 0$, or equivalently $x_0 \in \Omega^{\text{reg}}$.*

(1) *If $v(x_0) > 0$ and $x_M \in \Omega$, then $\omega_+ = \infty$.*

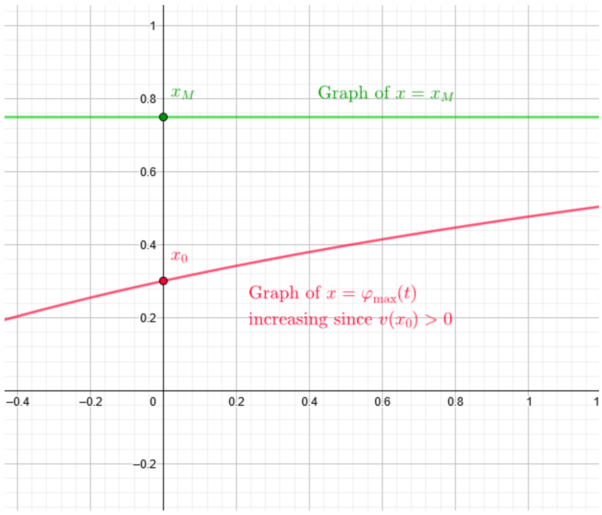


FIGURE 2. $v(x_0) > 0$ and $x_M \in \Omega \implies \omega_+ = \infty$

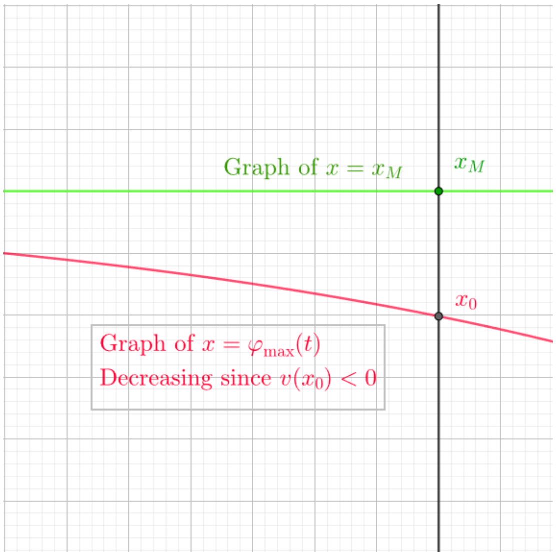


FIGURE 3. $v(x_0) < 0$ and $x_M \in \Omega \implies \omega_- = -\infty$

(2) *If $v(x_0) < 0$ and $x_M \in \Omega$ then $\omega_- = -\infty$.*
(3) *If $v(x_0) < 0$ and $x_m \in \Omega$ then $\omega_+ = \infty$.*
(4) *If $v(x_0) > 0$ and $x_m \in \Omega$ then $\omega_- = -\infty$.*

Proof. We have proved (1) in the above discussion. Part (2) is obtained by applying (1) to the time reversal ($*_{\text{tr}}$) of $\varphi_{\max}$. Part (3) is obtained by applying (1) to the state reversal ($*_{\text{sr}}$) of $\varphi_{\max}$. Part (d) is obtained by applying (2) to the state reversal ($*_{\text{sr}}$) of $\varphi_{\max}$. I have included some pictures to explain things a bit better. See figures 2, 3 and 4 for cases (1), (2) and (3) respectively. ■

Proposition 1.2. *Let $\varphi : (a, b) \to \Omega$ be a solution of ($*$) with $x_0 \in \Omega^{\text{reg}}$. Let $\theta, \varphi_{\max}, x_m, x_M, \omega_-, \omega_+$ etc. be as before. Then*

(1) *φ takes values in Ω^{reg}.*
(2) *$(a, b) \subseteq (\omega-, \omega_+)$.*
(3) *$\varphi = \varphi_{\max}|_{(a,b)}$.*

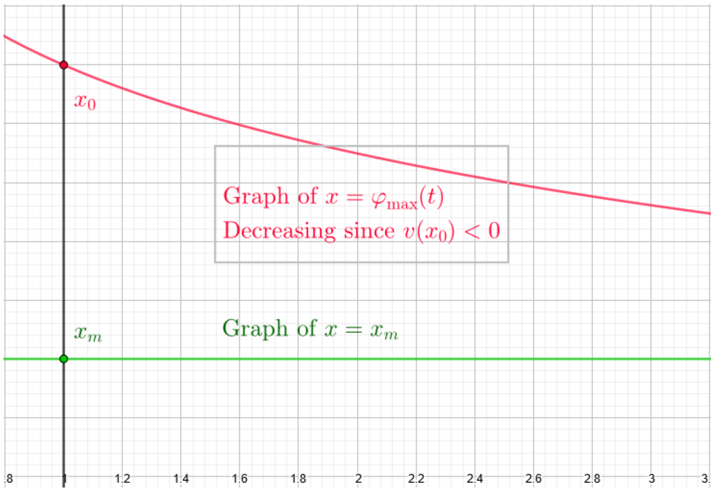


FIGURE 4. $v(x_0) < 0$ and $x_m \in \Omega \implies \omega_+ = \infty$

Proof. Parts (2) and (3) follow from (1) in view of the statement (∗∗) that we proved before. So, it is enough to prove (1). Let (α, β) be the connected component $\varphi^{-1}(\Omega^{\text{reg}})$ containing the point t_0. Now $\varphi|_{(\alpha,\beta)}$ takes values in Ω^{reg}. So, from (∗∗) it follows that $(\alpha, \beta) \subseteq (\omega_-, \omega_+)$ and

$$\varphi(s) = \varphi_{\max}(s) \quad , \quad \forall s \in (\alpha, \beta)$$

It is enough to show that $\alpha = a$ and $\beta = b$. Without loss of generality we assume that $v(x_0) > 0$ (the case $v(x_0) < 0$ can be handled by a state-reversal). Further, we will show that $b = \beta$. The proof mutatis mutandis will show that $a = \alpha$.

Because $(\alpha, \beta) \subseteq (\omega_-, \omega_+)$ and $(\alpha, \beta) \subseteq (a, b)$, we see that $\beta \leq b$ and $\beta \leq \omega_+$. For the sake of contradiction, suppose $\beta < b$ and $\beta < \omega_+$. Now, by taking limits, clearly $\varphi(\beta) = \varphi_{\max}(\beta) \in \Omega^{\text{reg}}$. This contradicts the fact that (α, β) is the largest open interval containing t_0 in $\varphi^{-1}(\Omega^{\text{reg}})$. Therefore, either $\beta = b$ or $\beta = \omega_+$. If $\beta = b$, then we are done. So, suppose $\beta = \omega_+$ and $\beta < b$. Since φ is defined on (a, b), $\varphi(\beta)$ makes sense. Clearly, $v(\varphi(\beta)) = 0$, again by the fact that (α, β) is the largest open interval containing t_0 in $\varphi^{-1}(\Omega^{\text{reg}})$. Moreover, φ is increasing on (α, β) because of the assumption $v(x_0) > 0$. All of this implies that

$$\varphi(\beta) = x_M$$

This means that $x_M \in \Omega$, which then implies that $\omega_+ = \infty$ by **Lemma** 1.1 and hence $\beta = \infty$. But then, $\beta < b$ is a contradiction. So, if $\beta = \omega_+$ then $\beta = b$.

The same argument shows that $\alpha = a$, and hence we are done. ■

Theorem 1.3. *Consider the IVP* (∗). *Then there is a maximal solution* $(J_{\max}, \varphi_{\max})$ *of* (∗) *such that if* (J, φ) *is a solution of* (∗) *then* $J \subseteq J_{\max}$ *and* $\varphi = \varphi_{\max}|_J$.

Proof. Suppose $v(x_0) \neq 0$. Then we have already seen that this is true in **Proposition** 1.2. So, suppose $v(x_0) = 0$. We claim that the maximal solution to (∗) is the constant function

$$\varphi(t) = x_0 \quad , \quad \forall t \in \mathbb{R}$$

In this case $J_{\max} = \mathbb{R}$. It is clear that this constant function is a solution of (∗). Conversely, suppose (J, ψ) is a solution of (∗). Suppose $v(\psi(\tau)) \neq 0$ for some $\tau \in J$. Let $x^* = \psi(\tau)$. Then ψ is a solution of the IVP

$$\dot{x} = v(x) \quad , \quad x(\tau) = x^*$$

So from **Proposition** 1.2, ψ takes values only in Ω^{reg}, which is clearly a contradiction because $x_0 \in J$. ■

1.5. **Picard-Lindelöf Theorem.** In this section, we will prove a very useful and important theorem for existence of solutions to certain ODEs.

1.5.1. *Notation.* Let $A \subseteq \mathbb{R}^n$. A function $\vec{\boldsymbol{f}} : A \to \mathbb{R}^n$ is said to be *Lipschitz* if there is a constant $L > 0$ such that

$$||\vec{\boldsymbol{f}}(\vec{\boldsymbol{x}}) - \vec{\boldsymbol{f}}(\vec{\boldsymbol{x}})|| \leq L||\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|| \quad , \quad \forall \vec{\boldsymbol{x}}, \vec{\boldsymbol{y}} \in A$$

The constant L is called a *Lipschitz constant* for $\vec{\boldsymbol{f}}$.

Suppose $I = [a, b]$ where $a \in (-\infty, \infty)$ and $b \in (a, \infty]$. We say $\vec{\boldsymbol{f}} : I \to \mathbb{R}^n$ is $\mathscr{C}^1$ if $\frac{d}{dt}^+|_{t=a}\vec{\boldsymbol{f}}$ exists, and the resultant function $\dot{\vec{\boldsymbol{f}}} : I \to \mathbb{R}^n$ is continuous, where $\dot{\vec{\boldsymbol{f}}}(a)$ is the one-sided derivative we just mentioned. Similarly, we can make sense of $\mathscr{C}^1$ functions on intervals of the form $(a, b]$ and $[a, b]$. If I is a closed and bounded interval in $\mathbb{R}$, then we denote the set of all continuous maps on I taking values in $\mathbb{R}^n$ by $C(I, n)$. We know from earlier courses that $(C(I, n), ||\cdot||_\infty)$ is a *Banach space*, i.e it is a complete normed vector space.

To prove the existence and uniqueness of solutions to DEs to IVPs on I, it is more convenient to work with a different norm denoted $||\cdot||_w$. Let t_0 be the midpoint of I and $|I| = 2b$. Then, I is of the form $[t_0 - b, t_0 + b]$. Let L be a positive number. Then we define

$$||\cdot||_w = ||\cdot||_{w,L} : C(I, n) \to [0, \infty)$$

by

$$||\vec{\boldsymbol{f}}||_w = \sup_{t\in I}\{e^{-2L|t-t_0|}||\vec{\boldsymbol{f}}(t)||\}$$

It is straightforward to check that this is a norm.

Lemma 1.4. *$||\cdot||_w$ and $||\cdot||_\infty$ are equivalent norms.*

Proof. It can be checked that

$$e^{-L2b} \cdot ||\vec{\boldsymbol{f}}||_\infty \leq ||\vec{\boldsymbol{f}}||_w \leq ||\vec{\boldsymbol{f}}||_\infty \quad , \quad \forall \vec{\boldsymbol{f}} \in C(I, n)$$

■

Theorem 1.5 (**Picard-Lindelöf Theorem**)**.** *Let $\vec{\boldsymbol{a}} \in \mathbb{R}^n$, $t_0 \in \mathbb{R}$ and let*

$$\vec{\boldsymbol{v}} : [t_0 - \alpha, t_0 + \alpha] \times \overline{B}(\vec{\boldsymbol{a}}, r) \to \mathbb{R}^n$$

be a continuous map with upper bound M for $||\vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}})||$. Suppose further that there is a positive constant L such that for each $t \in [t_0 - \alpha, t_0 + \alpha]$, the function $\vec{\boldsymbol{v}}(t, \_\_) : \overline{B}(\vec{\boldsymbol{a}}, r) \to \mathbb{R}^n$ is Lipschitz with Lipschitz constant L. Then the IVP

$$\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}}) \quad , \quad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{a}}$$

has a unique solution defined on $[t_0 - b, t_0 + b]$ where $b = \min\left\{\alpha, \dfrac{r}{M}\right\}$

Proof. Let $I = [t_0 - b, t_0 + b]$ where b is as in the statement of the theorem. Let $||\cdot||_w = ||\cdot||_{w,L}$ where L is the constant in the statement of the theorem. Recall

$$||\vec{\boldsymbol{f}}||_w = \sup_{t\in I}\{e^{-2L|t-t_0|}||\vec{\boldsymbol{f}}(t)||\}$$

Let

$$\begin{aligned} X &= \{\vec{\boldsymbol{f}} \in C(I, n) \mid \vec{\boldsymbol{f}}(I) \subseteq \overline{B}(a, r)\} \\ &= \{\vec{\boldsymbol{f}} \in C(I, n) \mid ||\vec{\boldsymbol{f}}(t) - \vec{\boldsymbol{a}}|| \leq r, t \in I\} \end{aligned}$$

Clearly, X is closed in $C(I, n)$. Therefore $(X, ||\cdot||_\infty)$ is complete. Since $||\cdot||_\infty$ and $||\cdot||_w$ are equivalent, $(X, ||\cdot||_w)$ is also complete.

For $\vec{\boldsymbol{f}} \in X$ and $t \in I$ we define

$$(T\vec{\boldsymbol{f}})(t) = \vec{\boldsymbol{a}} + \int_{t_0}^{t} \vec{\boldsymbol{v}}(s, \vec{\boldsymbol{f}}(s))ds$$

By the fundamental theorem of calculus, the map $t \mapsto (T\vec{\boldsymbol{f}})(t)$ is differentiable on I and hence it is continuous. Moreover, we have

$$\begin{aligned} ||(T\vec{\boldsymbol{f}})(t) - \vec{\boldsymbol{a}}|| &= \left|\left|\int_{t_0}^{t} \vec{\boldsymbol{v}}(s, \vec{\boldsymbol{f}}(s))ds\right|\right| \\ &\leq \left|\int_{t_0}^{t} ||\vec{\boldsymbol{v}}(s, \vec{\boldsymbol{f}}(s))||ds\right| \\ &\leq M\left|\int_{t_0}^{t} ds\right| \\ &\leq M(t - t_0) \\ &\leq Mb \\ &\leq r \quad , \quad \text{since } b = \min\left\{\alpha, \frac{r}{M}\right\} \end{aligned}$$

Notice the absolute value signs above; we have put it for the case $t < t_0$. So, it follows that $T\vec{\boldsymbol{f}} \in X$. Hence we have a map $T : X \to X$. Suppose $\vec{\boldsymbol{f}}, \vec{\boldsymbol{g}} \in X$. Then (again we put an absolute value sign everywhere for the case $t < t_0$)

$$\begin{aligned} e^{-2L|t-t_0|} \cdot ||(T\vec{\boldsymbol{f}})(t) - (T\vec{\boldsymbol{g}})(t)|| &= e^{-2L|t-t_0|}\left|\left|\int_{t_0}^{t} \{\vec{\boldsymbol{v}}(s, \vec{\boldsymbol{f}}(s)) - \vec{\boldsymbol{v}}(s, \vec{\boldsymbol{g}}(s))\}ds\right|\right| \\ &\leq e^{-2L|t-t_0|}\left|\int_{t_0}^{t} ||\vec{\boldsymbol{v}}(s, \vec{\boldsymbol{f}}(s)) - \vec{\boldsymbol{v}}(s, \vec{\boldsymbol{g}}(s))||ds\right| \\ &\leq e^{-2L|t-t_0|} \cdot L\left|\int_{t_0}^{t} ||\vec{\boldsymbol{f}}(s) - \vec{\boldsymbol{g}}(s)||ds\right| \\ &\leq Le^{-2L|t-t_0|}\left|\int_{t_0}^{t} e^{2L|s-t_0|}||\vec{\boldsymbol{f}} - \vec{\boldsymbol{g}}||_w ds\right| \\ &= \frac{L||\vec{\boldsymbol{f}} - \vec{\boldsymbol{g}}||_w e^{-2L|t-t_0|}}{2L}\left|e^{2L|t-t_0|} - 1\right| \\ &= \frac{1}{2}||\vec{\boldsymbol{f}} - \vec{\boldsymbol{g}}||_w(1 - e^{-2L|t-t_0|}) \\ &= \frac{1}{2}||\vec{\boldsymbol{f}} - \vec{\boldsymbol{g}}|| \end{aligned}$$

So, it follows that

$$||T\vec{\boldsymbol{f}} - T\vec{\boldsymbol{g}}||_w \leq \frac{1}{2}||\vec{\boldsymbol{f}} - \vec{\boldsymbol{g}}||_w$$

Hence, T is a contraction mapping with contraction factor $\frac{1}{2}$. By the contraction mapping theorem, it follows that T has a unique fixed point in X, say $\vec{\boldsymbol{\varphi}}$. This fact translates to the equation

$$\varphi(t) = \vec{\boldsymbol{a}} + \int_{t_0}^{t} \vec{\boldsymbol{v}}(s, \vec{\boldsymbol{\varphi}}(s))ds$$

and hence $\vec{\varphi}$ is a solution of the given IVP.

Conversely if $\vec{\psi}$ is a solution of the given IVP in $[t_0-b, t_0+b]$ then by the fundamental theorem of calculus, for $t \in I$ we have

$$\begin{aligned}\vec{\psi}(t) &= \vec{a} + \int_{t_0}^{t} \dot{\vec{\psi}}(s)ds \\ &= \vec{a} + \int_{t_0}^{t} \vec{v}(s, \vec{\psi}(s))ds \end{aligned} \tag{†}$$

Hence for $t \in I$ we have

$$\begin{aligned}|\vec{\psi}(t) - \vec{a}| &= \left| \int_{t_0}^{t} \vec{v}(s, \vec{\psi}(s))ds \right| \\ &\leq M|t - t_0| \\ &\leq Mb \leq r\end{aligned}$$

Thus $\vec{\psi} \in X$. From (†) we see that $T\vec{\psi} = \vec{\psi}$ and hence by uniqueness of the fixed point if follows that $\vec{\psi} = \vec{\varphi}$. This completes our proof. ■

Definition 1.2. Let Ω be a *domain* in $\mathbb{R} \times \mathbb{R}^n$, i.e Ω is a connected open subset of $\mathbb{R}^{n+1}$. A map

$$\vec{v} : \Omega \to \mathbb{R}^n$$

is said to be *locally Lipschitz* with respect to the second variable (or in our case, with respect to the phase) if it is continuous and for each $(t_0, \vec{a}) \in \Omega$, there is a positive number $L = L(t_0, \vec{a})$ and a product set $I \times U$ containing $(t_0, \vec{a})$ as an interior point such that for each $t \in I$, the restriction of $\vec{v}(t, \_\_)$ to U is Lipschitz continuous with Lipschitz constant $L = L(t_0, \vec{a})$. We say it is *uniformly Lipschitz* if $L(t_0, \vec{a})$ does not depend upon the point $(t_0, \vec{a})$.

1.6. **Maximal Intervals of Existence.** Let Ω be a domain in $\mathbb{R} \times \mathbb{R}^n$ and let $\vec{v} : \Omega \to \mathbb{R}^n$ a continuous map such that for each $(t_0, \vec{a}) \in \Omega$ the IVP

$$\dot{\vec{x}} = \vec{v}(t, \vec{x}) \quad , \quad \vec{x}(t_0) = \vec{a} \tag{$*_{t_0,\vec{a}}$}$$

has a solution on some interval I (open, closed or half open) containing t_0 in its interior, and the solution is unique on this interval.

Let us fix the point $(t_0, \vec{a}) \in \Omega$. If I is an interval on which the solution to $(*_{t_0,\vec{a}})$ with t_0 in the interior of I, we call I an *interval of existence* for $(*_{t_0,\vec{a}})$.

Now, suppose I_1, I_2 are open intervals of existence for $(*_{t_0,\vec{a}})$. Let $\vec{\varphi_1}$ and $\vec{\varphi_2}$ be solutions of $(*_{t_0,\vec{a}})$ in I_1, I_2 respectively. From the hypothesis, the set

$$S := \{t \in I_1 \cap I_2 \mid \vec{\varphi_1}(t) = \vec{\varphi_2}(t)\}$$

is open. Indeed, if $\tau \in S$ and $\vec{a}^* = \vec{\varphi_1}(\tau) = \vec{\varphi_2}(\tau)$ then $\vec{\varphi_1}$ and $\vec{\varphi_2}$ are solutions of $(*_{\tau,\vec{a}})$ (i.e the initial timepoint is τ and the initial phase is $\vec{a}^*$), so in a neighborhood of τ, $\vec{\varphi_1}$ and $\vec{\varphi_2}$ agree. On the other hand, S is clearly closed (in $I_1 \cap I_2$) and non-empty, since $t_0 \in S$. Since $I_1 \cap I_2$ is connected this means $S = I_1 \cap I_2$, because the only sets which are both closed in a connected set are the empty set and the set itself. So we see that

$$\vec{\varphi_1} \textit{ and } \vec{\varphi_2} \textit{ agree on } I_1 \cap I_2$$

From the above, the union of all open intervals of existence for $(*_{t_0,\vec{a}})$ is also an interval of existence; we have shown that the solutions agree on the intersection of any two open intervals, hence the solution can be extended to the union of all the intervals.

Moreover, this union will not be a disjoint union because the point t_0 belongs to all the intervals of existence. Now, let

$$J_{\max} = (\omega_-, \omega_+) := \bigcup I$$

where the union is taken over all open intervals of existence of $(*_{t_0,\vec{\boldsymbol{a}}})$. Clearly, it follows that $J_{\max}$ is an open interval of existence of $(*_{t_0,\vec{\boldsymbol{a}}})$.

Now suppose I is an interval of existence of $(*_{t_0,\vec{\boldsymbol{a}}})$ (not necessarily open). If I is open, clearly $I \subseteq J_{\max}$. If I is not open, then it is still true that $I \subseteq J_{\max}$, and we now show this. Without loss of generality, suppose $I = (a, b]$, i.e I has a right boundary point, and let $\vec{\boldsymbol{\varphi}}$ be the solution of $(*_{t_0,\vec{\boldsymbol{a}}})$ on this interval I. Put $a^* = \vec{\boldsymbol{\varphi}}(b)$. From our hypothesis, there is an open interval $I' = (b - \delta_1, b + \delta_2)$ of existence of $(*_{b,\vec{\boldsymbol{a}}^*})$ with $\delta_1, \delta_2 > 0$, and let the solution on the interval I' be $\vec{\boldsymbol{\varphi}}'$. So, we can *extend* the map $\vec{\boldsymbol{\varphi}}$ to the *open* interval $(a, b + \delta_2)$ by making it equal to $\vec{\boldsymbol{\varphi}}'$ on $(b, b + \delta_2)$. By this, $\vec{\boldsymbol{\varphi}}$ becomes a solution to $(*_{t_0,\vec{\boldsymbol{a}}})$ on the open interval $(a, b + \delta_2)$. So, we see that $I \subseteq (a, b + \delta_2) \subseteq J_{\max}$, and hence we are done.

Remark 1.5.1. In all of this, we have assumed that the intervals have non-empty interiors. The degenerate cases where the intervals are singletons are not interesting.

1.7. **The General Existence Theorem.** Now we will prove the following existence theorem in higher dimensions.

Theorem 1.6. *Suppose Ω is a domain in $\mathbb{R} \times \mathbb{R}^n$ and $\vec{\boldsymbol{v}} : \Omega \to \mathbb{R}^n$ a locally Lipschitz continuous function with respect to the second variable. Let $(t_0, \vec{\boldsymbol{a}}_{\mathbf{0}}) \in \Omega$. Then the IVP*

$$(*_{t_0,\vec{\boldsymbol{a}}_{\mathbf{0}}}) \qquad \dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}}) \quad , \quad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{a}}_0$$

has a maximal interval of existence, and is of the form (ω_-, ω_+) with $\omega_- \in [-\infty, \infty)$ and $\omega_+ \in (-\infty, \infty]$. There is a unique solution

$$\vec{\boldsymbol{\varphi}}_0 = \vec{\boldsymbol{\varphi}}_{(t_0,\vec{\boldsymbol{a}}_{\mathbf{0}})} : (\omega_-, \omega_+) \to \mathbb{R}^n$$

of $(*_{t_0,\vec{\boldsymbol{a}}_{\mathbf{0}}})$ on (ω_-, ω_+) and any solution of $(*_{t_0,\vec{\boldsymbol{a}}_{\mathbf{0}}})$ on an interval I containing t_0 is the restriction of $\vec{\boldsymbol{\varphi}}_0$ to I. The variable point $(t, \vec{\boldsymbol{\varphi}}_0(t))$ leaves every compact subset K of Ω as $t \downarrow \omega_-$ and as $t \uparrow \omega_+$.

Proof. I claim that only the last statement needs to be proven, and everything else has been proven. This is so because at the point $(t_0, \vec{\boldsymbol{a}}_{\mathbf{0}})$ we can use the local Lipschitz continuity (w.r.t the second variable) of $\vec{\boldsymbol{v}}$, and apply the **Picard-Lindelöf Theorem** 1.5 to get the existence of a unique solution to $(*_{t_0,\vec{\boldsymbol{a}}_{\mathbf{0}}})$. The existence of maximal intervals follows from the discussion in the subsection 1.6.

So, let $K \subseteq \Omega$ be any compact subset. For each $(t, \vec{\boldsymbol{a}}) \in K$, pick a closed rectangle $[t - 2\alpha(t, \vec{\boldsymbol{a}}), t + 2\alpha(t, \vec{\boldsymbol{a}})] \times \overline{B}(\vec{\boldsymbol{a}}, 2\rho(t, \vec{\boldsymbol{a}}))$ (where $\vec{\boldsymbol{\alpha}}(t, \vec{\boldsymbol{a}})$ and $\rho(t, \vec{\boldsymbol{a}})$ depend on the point $(t, \vec{\boldsymbol{a}})$, and notice the factor of 2) such that this closed rectangle is contained in Ω and in which $\vec{\boldsymbol{v}}$ is uniformly Lipschitz in the second variable; such a rectangle exists because Ω is an open set and $\vec{\boldsymbol{v}}$ is locally Lipschitz.

The sets $(t - \alpha, t + \alpha) \times B(\vec{\boldsymbol{a}}, \rho)$ (where again α and ρ are dependent on $(t, \vec{\boldsymbol{a}})$) form an open cover of K as $(t, \vec{\boldsymbol{a}})$ varies over K. By the compactness of K, we have a finite subcover. So, K is covered by a finite union of closed rectangles

$$K \subseteq \bigcup_{i=1}^{m} [t_i - \alpha_i, t_i + \alpha_i] \times \overline{B}(\vec{\boldsymbol{a}}_{\boldsymbol{i}}, \rho_i)$$

Now, let K' be the union of closed rectangles

$$K' = \bigcup_{i=1}^{m} [t_i - 2\alpha_i, t_i + 2\alpha_i] \times \overline{B}(\vec{\boldsymbol{a}}_{\boldsymbol{i}}, 2\alpha_i)$$

Let $(t, \vec{\boldsymbol{a}}) \in K$. Then, $(t, \vec{\boldsymbol{a}}) \in [t_i - \alpha_i, t_i + \alpha_i] \times \overline{B}(\vec{\boldsymbol{a}}_{\boldsymbol{i}}, \rho_i)$ for some $1 \leq i \leq m$. Then from the triangle inequality we have the following.

$$(\bullet) \qquad [t - \alpha_i, t + \alpha_i] \times \overline{B}(\vec{\boldsymbol{a}}, \rho_i) \subseteq [t_i - 2\alpha_i, t_i + 2\alpha_i] \times \overline{B}(\vec{\boldsymbol{a}}, 2\rho_i)$$

Note that K' is compact being a finite union of closed rectangles. Let

$$M = \sup_{(t,\vec{\boldsymbol{a}}) \in K'} ||\vec{\boldsymbol{v}}(t, \vec{\boldsymbol{a}})||$$
$$\alpha = \min_{1 \leq i \leq m} \alpha_i$$
$$r = \min_{1 \leq i \leq m} \rho_i$$
$$b = \min \left\{ \alpha, \frac{r}{M} \right\}$$

On each rectangle $[t_i - 2\alpha_i, t_i + 2\alpha_i] \times \overline{B}(\vec{\boldsymbol{a}}, 2\rho_i)$, we know that $\vec{\boldsymbol{v}}$ is uniformly Lipschitz. By $(\bullet)$, it is therefore uniformly Lipschitz on $[t - \alpha, t + \alpha] \times \overline{B}(\vec{\boldsymbol{a}}, r)$. So by the **Picard-Lindelöf Theorem** 1.5 we know that if $(\tau, \vec{\boldsymbol{a}}) \in K$ then there is a unique solution $\varphi_{(\tau, \vec{\boldsymbol{a}})}$ on $[\tau - b, \tau + b]$ to the IVP

$$(*_{\tau, \vec{\boldsymbol{a}}}) \qquad \dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}}) \quad , \quad \vec{\boldsymbol{x}}(\tau) = \vec{\boldsymbol{a}}$$

Note that the above equation gives us a family of IVPs as $(\tau, \vec{\boldsymbol{a}})$ varies over K, but note that the constant b does not vary; in other words, as $(\tau, \vec{\boldsymbol{a}})$ varies over K we get a different IVP, but the interval of existence of the solution, which is $[\tau - b, \tau + b]$, has fixed length.

Note that $\vec{\boldsymbol{\varphi}}_0 = \vec{\boldsymbol{\varphi}}_{(t_0, \vec{\boldsymbol{a}}_{\boldsymbol{0}})}$ where $\vec{\boldsymbol{\varphi}}_0$ is an in the statement of the theorem. Suppose $\tau \in (\omega_-, \omega_+)$ and $\vec{\boldsymbol{a}} = \vec{\boldsymbol{\varphi}}_{\boldsymbol{0}}(t)$. From the uniqueness of solutions, clearly $\vec{\boldsymbol{\varphi}}_0 = \vec{\boldsymbol{\varphi}}_{(\tau, \vec{\boldsymbol{a}})}$. Moreover, the maximal interval of existence of $\vec{\boldsymbol{\varphi}}_{(\tau, \vec{\boldsymbol{a}})}$ is therefore also (ω_-, ω_+). So, it follows that if $\tau \in (\omega_-, \omega_+)$ and $(\tau, \vec{\boldsymbol{\varphi}}_0(\tau)) \in K$ then $[\tau - b, \tau + b] \subseteq (\omega_-, \omega_+)$, i.e

$$(\tau, \vec{\boldsymbol{\varphi}}_0(\tau)) \in K \implies \omega_- + b < \tau < \omega_+ - b$$

It follows that if $\tau \in [\omega_+ - b, \omega_+)$ then $(\tau, \vec{\boldsymbol{\varphi}}_0(\tau)) \notin K$ and if $\tau \in (\omega_-, \omega_- + b]$ then $(\tau, \vec{\boldsymbol{\varphi}}_0(\tau)) \notin K$. This proves the claim if both $\omega_- \neq -\infty$ and $\omega_+ = \infty$. If $\omega_- = -\infty$ or $\omega_+ = \infty$, the claim is trivially true because K, being a compact set, is bounded. ∎

Corollary 1.6.1. *If $U \subseteq \mathbb{R}^n$ is a bounded open set and $\Omega = (c, d) \times U$ with (c, d) an open interval in $\mathbb{R}$, then either $\omega_+ = d$ or $\vec{\boldsymbol{\varphi}}_0(t) \to \partial U$ as $t \uparrow \omega_+$, and either $\omega_- = c$ or $\vec{\boldsymbol{\varphi}}_0(t) \to \partial U$ as $t \downarrow \omega_-$.*

Proof. We will only prove the statement for ω_+, and the statement for ω_- will have an analogous proof. Suppose $\omega_+ \neq d$. Then $\omega_+ < d$ and for $\epsilon > 0$ sufficiently small $-\epsilon + \omega_+ \in (c, d)$. Let

$$f : \mathbb{R}^n \to [0, \infty)$$

be the function given by

$$f(x) = \inf_{\vec{\boldsymbol{z}} \in \partial U} ||\vec{\boldsymbol{x}} - \vec{\boldsymbol{z}}||$$

or in simple words, f is the distance from ∂U. Clearly, f is a continuous function. For $\epsilon > 0$, put

$$\Gamma_\epsilon = \{\vec{\boldsymbol{x}} \in U \mid f(\vec{\boldsymbol{x}}) \geq \epsilon\}$$

and let $K_\epsilon = [-\epsilon + \omega_+, \omega_+] \times \Gamma_\epsilon$. Then K_ϵ is compact and for sufficiently small ϵ, K_ϵ is a non-empty subset of Ω.

So by **Theorem** 1.6, $(t, \vec{\varphi}_0(t))$ exits K_ϵ. It cannot exit anywhere on $\{\omega_+\} \times \Gamma_\epsilon$, for $\vec{\varphi}_0(t)$ does not make sense for $t = \omega_+$. Thus, there exists $\tau_\epsilon \in [-\epsilon + \omega_+, \omega_+)$ such that $f(\vec{\varphi}_0(t)) < \epsilon$ for all $t \in [\tau_\epsilon, \omega_+)$. This proves $\vec{\varphi}_0(t) \to \partial U$ as $t \uparrow \omega_+$. This completes the proof. ■

Corollary 1.6.2. *If $\Omega = (c, d) \times \mathbb{R}^n$ with (c, d) an open interval in $\mathbb{R}$ then either $\omega_+ = d$ or $||\vec{\varphi}_0(t)|| \to \infty$ as $t \uparrow \omega_+$, and either $\omega_- = c$ or $||\vec{\varphi}_0(t)|| \to \infty$ as $t \downarrow \omega_-$.*

Proof. Apply **Corollary** 1.6.1 to $(c, d) \times B(\vec{\mathbf{0}}, n)$ where $n \in \mathbb{N}$. If $\omega_+ \neq d$, i.e $\omega_+ < d$ then $\vec{\varphi}_0(t) \to S(\vec{\mathbf{0}}, n)$ as $t \to \omega_+$ (here $S(\vec{\mathbf{0}}, n) = \{x \in \mathbb{R} \mid ||\vec{x}|| = n\}$). Let $n \to \infty$. Clearly, $||\vec{\varphi}_0(t)|| \to \infty$ as $t \uparrow \omega_+$. A similar argument works for ω_-. This completes the proof. ■

1.8. **First Order Linear Equations.** Let $||\cdot||_\circ$ be the operator norm on $\text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m)$ as well as on $M_{m,n}(\mathbb{R})$, the space of $m \times n$ real matrices.

We know that if $A \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m)$ then

$$||A||_\circ = \sup_{||\vec{x}||=1} ||A\vec{x}||$$

Theorem 1.7. *Let $I \subseteq \mathbb{R}$ be an interval (closed, half-open or open, but with non-empty interior), $A : I \to M_n(\mathbb{R})$ and $\vec{g} : I \to \mathbb{R}^n$ continuous maps. Let $(t_0, \vec{a}_0) \in I \times \mathbb{R}^n$. Then the IVP*

$$(*) \qquad \dot{\vec{x}}(t) = A(t)\vec{x}(t) + \vec{g}(t) \quad , \quad \vec{x}(t_0) = \vec{a}_0$$

has a unique solution on I.

Remark 1.7.1. Equations of the above form are called *linear differential equations*. If $\vec{g}(t)$ is identically zero, then it is called a *homogeneous linear differential equation*.

Proof. Note that we can find an increasing sequence of closed intervals

$$I_0 \subset I_1 \subset I_2 \subset ... \subset I_n \subset ...$$

such that $I = \bigcup I_n$. Therefore, it is enough to assume that I is closed, say $I = [c, d]$.

Next, let

$$M := \sup_I ||\vec{g}(t)|| \quad , \quad L := \sup_I ||A(t)||_\circ$$

Note that if we set $\vec{v}(t, x) = A(t)\vec{x}(t) + \vec{g}(t)$ then

$$\begin{aligned} ||\vec{v}(t, x) - \vec{v}(t, y)|| &= ||A(t)\vec{x} - A(t)\vec{y}|| \\ &\leq ||A(t)||_\circ ||\vec{x} - \vec{y}|| \\ &\leq L||\vec{x} - \vec{y}|| \end{aligned}$$

So $\vec{v}$ is uniformly Lipschitz in the second variable on I.

For $\eta > 0$, let $I_\eta = (c - \eta, d + \eta)$. Set $\Omega = I_\eta \times \mathbb{R}^n$. Extend A to I_η by setting $A(t) = A(d)$ for $d \leq t < d + \eta$ and $A(t) = A(c)$ for $c - \eta < t \leq c$. Similarly, extend $\vec{g}$ to I_η by the same recipe. Now A and $\vec{g}$ are continuous on I_η. Again set $\vec{v}(t, \vec{x}) = A(t)\vec{x}(t) + \vec{g}(t)$ for $t \in I_\eta$. Note that M remains the supremum of $||\vec{g}(t)||$ on I_η, and L the supremum of $||A(t)||_0$ on I_η. Clearly the extended $\vec{v}$ is also uniformly Lipschitz in the second variable on I_η.

Let $J_{\max} = (\omega_-, \omega_+)$ be the maximal interval of existence for our (extended) IVP. Recall that $\Omega = I_\eta \times \mathbb{R}^n$. For simplicity consider $t \in [t_0, \omega_+)$ (the case $t \in (\omega_-, t_0]$ can be handled similarly). Let

$$K = ||\vec{\boldsymbol{a}}_0|| + M(\omega_+ - t_0)$$

Let $\vec{\boldsymbol{\varphi}}_0 : J_{\max} \to \Omega$ be the maximal solution of our (extended) IVP. Then for $t \in [t_0, \omega_+)$

$$\vec{\boldsymbol{\varphi}}_0(t) = \vec{\boldsymbol{a}}_0 + \int_{t_0}^{t} A(s)\vec{\boldsymbol{\varphi}}_0(s)ds + \int_{t_0}^{t} \vec{\boldsymbol{g}}(s)ds$$

Therefore

$$\begin{aligned} ||\vec{\boldsymbol{\varphi}}_0(t)|| &\leq ||\vec{\boldsymbol{a}}_0|| + \int_{t_0}^{t} ||A(s)||_0 ||\vec{\boldsymbol{\varphi}}_0(s)||ds + \int_{t_0}^{t} ||\vec{\boldsymbol{g}}(s)||ds \\ &\leq ||\vec{\boldsymbol{a}}_0|| + L\int_{t_0}^{t} ||\vec{\boldsymbol{\varphi}}_0(s)||ds + M\int_{t_0}^{t} ds \\ &= ||\vec{\boldsymbol{a}}_0|| + M(t - t_0) + L\int_{t_0}^{t} ||\vec{\boldsymbol{\varphi}}_0(s)||ds \\ &\leq ||\vec{\boldsymbol{a}}_0|| + M(\omega_+ - t_0) + L\int_{t_0}^{t} ||\vec{\boldsymbol{\varphi}}_0(s)||ds \end{aligned}$$

So from all this, we get

$$(**) \qquad ||\vec{\boldsymbol{\varphi}}_0(t)|| \leq K + L\int_{t_0}^{t} ||\vec{\boldsymbol{\varphi}}_0(s)||ds$$

Let $f(t) = \int_{t_0}^{t} ||\vec{\boldsymbol{\varphi}}_0(s)||ds$. Then (**) amounts to

$$(\dagger) \qquad f'(t) \leq K + Lf(t) \implies f'(t) - Lf(t) \leq K$$

Now we use the trick of multipliying by the integrating factor. So, multiplying both sides above by e^{-Lt} we get

$$e^{-Lt}[f'(t) - Lf(t)] \leq e^{-Lt}K$$

which means

$$\frac{d}{dt}\left\{e^{-Lt}f(t)\right\} \leq e^{-Lt}K$$

So integrating both sides of the above inequality, we get

$$\int_{t_0}^{t} \frac{d}{ds}\{e^{-Ls}f(s)\}ds \leq K\int_{t_0}^{t} e^{-Ls}ds$$

and this implies that

$$e^{-Lt}f(t) - e^{-Lt_0}f(t_0) \leq \frac{K}{-L}(e^{-Lt} - e^{-Lt_0})$$

Because $f(t_0) = 0$, this gives us

$$e^{-Lt}f(t) \leq \frac{K}{L}(e^{-Lt_0} - e^{-Lt})$$

and hence

$$f(t) \leq \frac{K}{L}(e^{L(t-t_0)} - 1)$$

Substitute this inequality back in the inequality (†) and we get

$$\begin{aligned} f'(t) &\leq K + L\frac{K}{L}(e^{L(t-t_0)} - 1) \\ &= Ke^{L(t-t_0)} \\ &\leq Ke^{L}(\omega_+ - t_0) \end{aligned}$$

and all of this means

$$||\vec{\varphi}_0(t)|| \leq Ke^{L(\omega_+ - t_0)}$$

This means that $||\vec{\varphi}_0(t)||$ is bounded on $[t_0, \omega_+)$. So, **Corollary** 1.6.2 implies that $\omega_+ = d + \eta$, because if $\omega_+ < d + \eta$, $||\vec{\varphi}_0(t)||$ will be unbounded. Similarly, $\omega_- = c - \eta$.

This shows that $\vec{\varphi}_0$ exists on I. This is what we had to prove. ■

Remark 1.7.2. $\vec{\varphi}_0(t)$ depends upon the point $(t_0, \vec{a}_0)$ and hence $\vec{\varphi}_0(t) = \vec{\varphi}(t_0, \vec{a}_0, t)$. We will show (if $\vec{v}$ is $\mathscr{C}^1$ in $\vec{x}$) that $\vec{\varphi}$ depends smoothly on $(t_0, \vec{a}_0)$.

1.9. **More on Linear DEs and Variation of Parameters.** For this subsection, let I and A be fixed as in **Theorem** 1.7. Define

$$T : \mathscr{C}^1(I) \to \mathscr{C}^0(I)$$

by $T\vec{f} = \dot{\vec{f}} - A\vec{f}$. Here, for non-negative integer k, $\mathscr{C}^k(I)$ is the set $\mathscr{C}^k(I, \mathbb{R}^n)$, the $\mathbb{R}$-vector space of $\mathscr{C}^k$ maps from I to $\mathbb{R}^n$.

Clearly, T is a linear map. It is also surjective because if $\vec{g} \in \mathscr{C}^0(I)$ we know that the linear DE $(*)$ has a solution. Let $\vec{\varphi}$ be a solution of $(*)$. Then clearly $T\vec{\varphi} = \vec{g}$.

Let

$$S := \mathrm{Ker}(T)$$

Clearly S is the set of solutions of the *homogeneous* linear DE

$$(**) \qquad \dot{\vec{x}} = A\vec{x}$$

on I. Note that automatically S gets a natural structure of an $\mathbb{R}$-vector space. As in problem **5).** of HW-2 we can prove that S has dimension n.

Theorem 1.8. *Let A, I, T, S be as above. Then the following are true.*

1. *$T : \mathscr{C}^1(I) \to \mathscr{C}^0(I)$ is surjective.*
2. *S has dimension n.*

Proof. The proof is summarised in the above discussion. ■

Let $\vec{g} \in \mathscr{C}^0(I)$. We know $T^{-1}(\vec{g})$ is non-empty since T is surjective. Moreover, T is a linear transformation with null space S. So if $\vec{\varphi} \in T^{-1}(\vec{g})$ then

$$T^{-1}(\vec{g}) = S + \vec{\varphi}$$

In this case, $\vec{\varphi}$ is said to be a *particular solution* of $(*)$ and S is the set of solutions of the *homogeneous* linear DE $(**)$.

In summary, the general solution of the linear DE $(*)$ is of the form

$$\vec{\varphi} = c_1\vec{\varphi}_1 + ... + c_n\vec{\varphi}_n + \vec{\varphi}_p$$

where $c_1, ..., c_n$ are arbitrary constants, $\vec{\varphi}_1, ..., \vec{\varphi}_n$ is a basis of S, i.e a set of linearly independent solutions of $(**)$ and $\vec{\varphi}_p$ is a *particular* solution of $(*)$. The term $c_1\vec{\varphi}_1 + ... + c_n\vec{\varphi}_n$ is called the *complementary solution*.

1.9.1. *Variation of Parameters.* Consider the linear DE $(*)$ with A and $\vec{\boldsymbol{g}}$ in $\mathscr{C}^0(I)$. Suppose further that we have a basis $\vec{\boldsymbol{\varphi}}_1, ..., \vec{\boldsymbol{\varphi}}_n$ of S. Let M be the square matrix whose i^{th} column is $\vec{\boldsymbol{\varphi}}_i$, i.e

$$M = \begin{bmatrix} \vec{\boldsymbol{\varphi}}_1 & \dots & \vec{\boldsymbol{\varphi}}_n \end{bmatrix}$$

Now we have that

$$M : I \to \mathrm{GL}_n(\mathbb{R})$$

and the fact that the range of M is $\mathrm{GL}_n(\mathbb{R})$ is actually not hard to see, and is problem **3).** of HW-3. Moreover, since each entry in M is $\mathscr{C}^1$, therefore M is $\mathscr{C}^1$. Also, we have

$$\begin{aligned} \dot{M} &= \begin{bmatrix} \dot{\vec{\boldsymbol{\varphi}}}_1 & \dots & \dot{\vec{\boldsymbol{\varphi}}}_n \end{bmatrix} \\ &= \begin{bmatrix} A\vec{\boldsymbol{\varphi}}_1 & \dots & A\vec{\boldsymbol{\varphi}}_n \end{bmatrix} \\ &= A\begin{bmatrix} \vec{\boldsymbol{\varphi}}_1 & \dots & \vec{\boldsymbol{\varphi}}_n \end{bmatrix} \\ &= AM \end{aligned}$$

and hence

$$(\bullet) \qquad \dot{M} = AM$$

Let $\vec{\boldsymbol{\psi}}$ be the solution of $(*)$. For $t \in I$, since $M(t)$ is invertible we can find scalars $u_1(t), ..., u_n(t)$ such that

$$\vec{\boldsymbol{\psi}}(t) = u_1(t)\vec{\boldsymbol{\varphi}}_1(t) + ... + u_n(t)\vec{\boldsymbol{\varphi}}_n(t)$$

Indeed, set

$$\vec{\boldsymbol{u}}(t) = \begin{bmatrix} u_1(t) \\ \dots \\ u_n(t) \end{bmatrix} = M(t)^{-1}\begin{bmatrix} \vec{\boldsymbol{\psi}}_1(t) \\ \dots \\ \vec{\boldsymbol{\psi}}_n(t) \end{bmatrix} = M(t)^{-1}\vec{\boldsymbol{\psi}}(t)$$

Then clearly $M(t)\vec{\boldsymbol{u}}(t) = \vec{\boldsymbol{\psi}}(t)$, and so the required $u_1(t), ..., u_n(t)$ have been found. Infact, since $M(t)$ is invertible, $\vec{\boldsymbol{u}}(t)$ is actually unique. Since

$$M^{-1} : I \to \mathrm{GL}_n(\mathbb{R})$$

is continuous, it follows that $\vec{\boldsymbol{u}} : I \to \mathbb{R}^n$ given by $\vec{\boldsymbol{u}} = M^{-1}\vec{\boldsymbol{\psi}}$ is also continuous.

So we can translate the problem of finding a particular solution $\vec{\boldsymbol{\psi}}$ of $(*)$ to the following: find continuous $\vec{\boldsymbol{u}} : I \to \mathbb{R}^n$ such that $M\vec{\boldsymbol{u}}$ is a solution of $(*)$. Let $\vec{\boldsymbol{u}}$ be such a function. Then by problem **4).** of HW-3 and equation $(\bullet)$ we see that

$$\begin{aligned} \frac{d}{dt}(M\vec{\boldsymbol{u}}) &= \dot{M}\vec{\boldsymbol{u}} + M\dot{\vec{\boldsymbol{u}}} \\ &= AM\vec{\boldsymbol{u}} + M\dot{\vec{\boldsymbol{u}}} \end{aligned}$$

On the other hand, since $M\vec{\boldsymbol{u}}$ is a solution of $(*)$ we have

$$\frac{d}{dt}(M\vec{\boldsymbol{u}}) = AM\vec{\boldsymbol{u}} + \vec{\boldsymbol{g}}$$

Equating the right sides of the above equations we get

$$M\dot{\vec{\boldsymbol{u}}} = \vec{\boldsymbol{g}}$$

and hence we see that

$$\vec{\boldsymbol{u}} = \int (M^{-1}\vec{\boldsymbol{g}})(t)dt$$

With this $\vec{\boldsymbol{u}}$, a solution of (∗) is

$$\vec{\boldsymbol{\psi}}(t) = M(t)\int_{t_0}^{t}(M^{-1}\vec{\boldsymbol{g}})(s)ds$$

1.10. **One Parameter Group of Transformations.** Let M be any set, and let

$$g : \mathbb{R} \times M \to M$$

be an action of the additive group $\mathbb{R}$ on M. For fixed $t \in \mathbb{R}$, write $g^t : M \to M$ be the map given by the formula

$$g^t(m) = g(t, m)$$

Being a group action, we can regard g as a homomorphism

$$g : \mathbb{R} \to \mathrm{Aut}(M)$$

We have the following two relations, which are immediate.

$$g^t g^s = g^{t+s} \qquad (*)$$

$$g^0 = 1_M$$

Definition 1.3. Any map $g : \mathbb{R} \times M \to M$ satisfying (∗) is called a *one-parameter group of transformations*. This is often denoted $\{g^t\}$. The pair $(M, \{g^t\})$ is called a *phase flow*.

Definition 1.4. Let $(M, \{g^t\})$ be a phase flow. Then M is called the *phase* or *state space* of the flow. A point of M is called a *phase point* or a *state*.

Definition 1.5. Let x be a phase point. The map

$$\varphi = \varphi_x : \mathbb{R} \to M$$

given by

$$\varphi(t) = g^t x$$

is called the *motion* of x under the flow $(M, \{g^t\})$. Its image is called the *phase curve* of x.

Remark 1.8.1. From this point on, whenever the term *manifold* is used, we can just think of it as an open set in $\mathbb{R}^n$.

Definition 1.6. By a *one-parameter group* $\{g^t\}$ *of diffeomorphisms* of a manifold M (which can be thought of as a domain in Euclidean space) is meant a mapping

$$g : \mathbb{R} \times M \to M \quad , \quad g(t, x) = g^t x \quad , \quad t \in \mathbb{R}, x \in M$$

of $\mathbb{R} \times M$ into M such that

(1) g is a $\mathscr{C}^2$ mapping.
(2) The mapping $g^t : M \to M$ is a diffeomorphism for every $t \in \mathbb{R}$.
(3) The family $\{g^t \mid t \in \mathbb{R}\}$ is a one-parameter group of transformations.

Remark 1.8.2. Note that condition (2) above is actually redundant, as condition (3) forces condition (2).

Definition 1.7. A *one-parameter group of linear transformations* in $\mathbb{R}^n$ is a one-parameter group of diffeomorphisms

$$g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

such that $g^t : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation for every $t \in \mathbb{R}$.

Here is a review of what we will be doing.

(1) All one-parameter groups of linear transformations are of the form $g^t = e^{tA}$ where A is a linear transformation. We will (later) define $e^A = \sum \frac{A^m}{m!}$ for any $n \times n$ matrix A.

(2) Consider the IVP

$$\dot{\vec{x}} = \vec{v}(t, \vec{x}) \quad , \quad \vec{x}(\tau) = \vec{a}$$

The solution $\vec{\varphi}_{(\tau,\vec{a})}$ depends upon the point $(\tau, \vec{a})$. Rewrite as

$$\vec{\varphi}_{(\tau,\vec{a})}(t) = \vec{\varphi}(\tau, \vec{a}, t)$$

If $\vec{v}$ is $\mathscr{C}^1$, then $\vec{\varphi}$ varies smoothly with $\tau, \vec{a}$ and t.

1.10.1. *The phase velocity associated with* $\{g^t\}$*.* Fix a one parameter group of diffeomorphisms $\{g^t\}$ on an open set M of $\mathbb{R}^n$. For $\vec{x} \in M$, let

$$\vec{\varphi}_{\vec{x}} : \mathbb{R} \to M$$

be the map $t \mapsto g^t\vec{x}$. As before, we write

$$\dot{\vec{\varphi}}_x = \frac{d}{dt}\vec{\varphi}_{\vec{x}}$$

Note that

$$\dot{\vec{\varphi}}_{\vec{x}} : \mathbb{R} \to \mathbb{R}^n$$

Definition 1.8. The *phase velocity vector* of $\{g^t\}$ at $\vec{x} \in M$ is the vector $\vec{v}(\vec{x})$ given by the formula

$$\vec{v}(\vec{x}) = \dot{\vec{\varphi}}_{\vec{x}}(0) = \lim_{h\to 0}\frac{\vec{\varphi}_{\vec{x}}(h) - \vec{x}}{h}$$

The map $\vec{v} : M \to \mathbb{R}^n$ given by $\vec{x} \mapsto \vec{v}(\vec{x})$ is called the *phase velocity field.*

Theorem 1.9. *Let $\vec{v} : M \to \mathbb{R}^n$ be the phase velocity vector field of a one-parameter group of diffeomorphisms $\{g^t\}$ and let $\vec{x}_0 \in M$. Then $\vec{\varphi}_{\vec{x}_0}$ (as defined above) is the unique solution to the autonomous IVP*

$$\dot{\vec{x}} = \vec{v}(\vec{x}) \quad , \quad \vec{x}(0) = \vec{x}_0$$

Proof. Uniqueness follow from **Theorem** 1.6; because the map g is assumed to be $\mathscr{C}^2$, we see that the map $\vec{v}$ is $\mathscr{C}^1$, and we know that $\mathscr{C}^1$ maps are locally Lipschitz. We have to show that $\vec{\varphi}_{\vec{x}_0}$ is a solution to the given IVP. W e clearly have $\vec{\varphi}_{\vec{x}_0}(0) = \vec{x}_0$, and hence the initial conditions are satisfied. Furthermore, we have the following:

$$\begin{aligned}
\dot{\vec{\varphi}}_{\vec{x}_0}(s) &= \lim_{h\to 0}\frac{\vec{\varphi}_{\vec{x}_0}(s+h) - \vec{\varphi}_{\vec{x}_0}(s)}{h} \\
&= \lim_{h\to 0}\frac{g^{s+h}(\vec{x}_0) - g^s(\vec{x}_0)}{h} \\
&= \lim_{h\to 0}\frac{g^h(g^s(\vec{x}_0)) - g^s(\vec{x}_0)}{h} \\
&= \lim_{h\to 0}\frac{\vec{\varphi}_{g^s(\vec{x}_0)}(h) - g^s(\vec{x}_0)}{h} \\
&= \vec{v}(g^s(\vec{x}_0)) \\
&= \vec{v}(\vec{\varphi}_{\vec{x}_0}(s))
\end{aligned}$$

and this completes the proof. ∎

1.10.2. *One-parameter group of linear transformations.* Here, we will see an interesting example of one-parameter groups.

A *one-parameter group of linear transformations* on $\mathbb{R}^n$ is a one-parameter group $\{g^t\}$ of diffeomorhpisms of $\mathbb{R}^n$ such that each $g^t : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation.

Note that this gives us a group homomorphism

$$g : \mathbb{R} \to \mathrm{GL}_n(\mathbb{R})$$

such that g is $\mathscr{C}^2$, where we are regarding $\mathrm{GL}_n(\mathbb{R})$ as an open subset of $\mathbb{R}^{n^2}$.

Let $\{g^t\}$ be a one-parameter group of linear transformations on $\mathbb{R}^n$. We will compute the phase velocity field. Because $g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is $\mathscr{C}^2$, it follows that the map $\mathbb{R} \to \mathrm{GL}_n(\mathbb{R}) \subset \mathbb{R}^{n^2}$ given by $t \mapsto g^t$ is $\mathscr{C}^2$ (it's a good little exercise to check why this is true), and hence has continuous partial derivatives. Let

$$A = \frac{dg^t}{dt}|_{t=0}$$

and hence $A \in M_n(\mathbb{R})$. Because all norms on $\mathbb{R}^{n^2}$ are equivalent, we get that

$$\lim_{h \to 0} \left|\left| \frac{g^h - I_n}{h} - A \right|\right|_\circ = 0$$

where $||\cdot||_\circ$ is the operator norm on $M_n(\mathbb{R})$ and I_n is the identity matrix/map.

As usual let $||\cdot||$ be the standard Euclidean norm. We have the following.

$$\begin{aligned} \left|\left| \frac{g^h \vec{\boldsymbol{x}} - \vec{\boldsymbol{x}}}{h} - A\vec{\boldsymbol{x}} \right|\right| &= \left|\left| \left(\frac{g^h - I_n}{h} - A \right)(\vec{\boldsymbol{x}}) \right|\right| \\ &\leq \left|\left| \frac{g^h - I_n}{h} - A \right|\right|_\circ \cdot ||x|| \end{aligned}$$

The last quantity goes to 0 as $h \to 0$. So it follows that

$$\lim_{h \to 0} \frac{g^h \vec{\boldsymbol{x}} - \vec{\boldsymbol{x}}}{h} = A\vec{\boldsymbol{x}}$$

and hence by definition, $A\vec{\boldsymbol{x}}$ is the phase velocity at $\vec{\boldsymbol{x}}$ for $\{g^t\}$.

Remark 1.9.1. We will soon show that this forces $g^t = e^{tA}$ for any $t \in \mathbb{R}$, where A is as above. This will show that any one parameter group of linear transformations is given by an exponential.

Remark 1.9.2. Consider the DE

$$\dot{\vec{\boldsymbol{x}}} = A\vec{\boldsymbol{x}}$$

We will show, after exponentials are defined, that $\{e^{tA}\}$ is a one parameter group, with phase velocity field $\vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}) = A\vec{\boldsymbol{x}}$. From **Theorem** 1.9 it will then follow that $\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{a}}}(t) = e^{tA}\vec{\boldsymbol{a}}$ gives us a solution of the given DE with the initial condition $\vec{\boldsymbol{x}}(0) = \vec{\boldsymbol{a}}$.

1.11. **The Exponential Map.** Let $n \in \mathbb{N}$ be a fixed natural number. Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. For $R > 0$, let

$$B_R = \{T \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n) \mid ||T||_\circ \leq R\}$$

Let

$$S_k(T) = \sum_{m=0}^{k} \frac{T^m}{m!}$$

For $0 \le l \le k$ we have

$$\begin{aligned} ||S_k(T) - S_l(T)||_\circ &= \left|\left| \sum_{m=l+1}^{k} \frac{T^m}{m!} \right|\right|_\circ \\ &\le \sum_{m=l+1}^{k} \frac{||T||_\circ^m}{m!} \\ &\le \sum_{m=l+1}^{k} \frac{R^m}{m!} \end{aligned}$$

Since $\sum_{m\ge 0} R^m/m!$ is convergent, the above chain of inequalities shows that $\{S_k\}$ is uniformly Cauchy on B_R. Now because $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$ is a finite dimensional vector space, it is a complete space and hence $\{S_k(T)\}$ is a convergent sequence, and infact it converges uniformly on B_R. So, we conclude that

$$\sum_{m=0}^{\infty} \frac{T^m}{m!} \textit{ converges uniformly on compact subsets of } \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$$

Definition 1.9. Let $T \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$. The *exponential* e^T of T is the element of $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$ given by

$$e^T = \sum_{m=0}^{\infty} \frac{T^m}{m!}$$

Theorem 1.10. *The exponential series for* e^T *converges on compact subsets of* $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$. *Also, if* S *and* T *are commuting elements of* $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$ *then*

$$e^{T+S} = e^T e^S$$

Proof. We only need to prove the second statement of the theorem. Now suppose S and T are commuting linear maps on $\mathbb{K}^n$. Since S and T commute, the binomial theorem applies and we see that

$$\begin{aligned} ||e^{S+T} - S_k(S)S_k(T)||_\circ &= \left|\left| \sum_{m=0}^{\infty} \frac{(S+T)^m}{m!} - \sum_{i=0}^{k} \frac{S^i}{i!} \sum_{j=0}^{k} \frac{T^j}{j!} \right|\right|_\circ \\ &= \left|\left| \sum_{m=0}^{\infty} \frac{(S+T)^m}{m!} - \sum_{m=0}^{k} \sum_{l=0}^{m} \frac{S^l T^{m-l}}{l!(m-l)!} \right|\right|_\circ \\ &= \left|\left| \sum_{m=0}^{\infty} \frac{(S+T)^m}{m!} - \sum_{m=0}^{k} \frac{1}{m!} \sum_{l=0}^{m} \binom{m}{l} S^l T^{m-l} \right|\right|_\circ \\ &= \left|\left| \sum_{m=0}^{\infty} \frac{(S+T)^m}{m!} - \sum_{m=0}^{k} \frac{1}{m!} (S+T)^m \right|\right|_\circ \\ &= \left|\left| \sum_{m=k+1}^{\infty} \frac{(S+T)^m}{m!} \right|\right|_\circ \\ &\le \sum_{m=k+1}^{\infty} \frac{(||S||_\circ + ||T||_\circ)^m}{m!} \end{aligned}$$

(Update: there is an error in the above computation, but it can be suitably fixed). The last term can be made as small as needed by choosing m large enough. So it follows that $S_k(S)S_k(T) \to e^{T+S}$ as $k \to \infty$, and this proves the claim. ■

1.11.1. *Linear DE with constant coefficients.* Suppose $A \in M_n(\mathbb{R})$. We show that e^{tA} is a one-parameter group of linear transformations. Note that the equality

$$e^{(t+s)A} = e^{tA}e^{sA}$$

holds for all $t, s \in \mathbb{R}$ by **Theorem** 1.10. So, we only have to check that the map

$$g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, \quad (t, \vec{\boldsymbol{x}}) \mapsto e^{tA}\vec{\boldsymbol{x}}$$

is $\mathscr{C}^2$. But this is true, and infact g is $\mathscr{C}^\infty$ because of the following: if we fix $t \in \mathbb{R}$, then the map $e^{tA} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map, and hence it is $\mathscr{C}^\infty$. Also, if we fix $\vec{\boldsymbol{x}} \in \mathbb{R}^n$, then the map $e^{tA}\vec{\boldsymbol{x}}$ is given by a power series, and hence it is $\mathscr{C}^\infty$. So, all the partial derivatives of g are $\mathscr{C}^\infty$, and hence g itself is $\mathscr{C}^\infty$.

Now, we show that

$$\frac{de^{tA}}{dt}|_{t=0} = A$$

We have

$$\begin{aligned}\frac{de^{tA}}{dt}|_{t=0} - A &= \lim_{h\to 0}\frac{1}{h}(e^{hA} - I_n) - A \\ &= \lim_{h\to 0}\frac{1}{h}\left(\sum_{m=0}^{\infty}\frac{h^m A^m}{m!} - I_n\right) - A \\ &= \lim_{h\to 0}\sum_{m=2}^{\infty}\frac{h^{m-1}A^m}{m!}\end{aligned}$$

As $h \to 0$, we can assume that $|h| \leq 1$, and hence $|h^{m-1}| \leq |h|$ for $m \geq 2$. So we have

$$\left|\left|\frac{de^{tA}}{dt}|_{t=0} - A\right|\right|_\circ \leq \lim_{h\to 0}|h|\sum_{m=2}^{\infty}\frac{||A||_\circ^m}{m!} \leq \lim_{h\to 0}|h|e^{||A||_0} = 0$$

By definition, it follows that the phase velocity field of $\{e^{tA}\}$ at $\vec{\boldsymbol{x}}$ is $A\vec{\boldsymbol{x}}$. We have actually proven the following theorem.

Theorem 1.11. *Let $A \in M_n(\mathbb{R})$, $\vec{\boldsymbol{a}} \in \mathbb{R}^n$ and $\vec{\boldsymbol{\varphi}} : \mathbb{R} \to \mathbb{R}^n$ be the map given by $\vec{\boldsymbol{\varphi}}(t) = e^{tA}\vec{\boldsymbol{a}}$ and $g^t = e^{tA}$ for $t \in \mathbb{R}$. Then $\vec{\boldsymbol{\varphi}}$ is the unique solution to the IVP $\dot{\vec{\boldsymbol{x}}} = A\vec{\boldsymbol{x}}$, $\vec{\boldsymbol{x}}(0) = \vec{\boldsymbol{a}}$.*

Proof. We have seen that $\{e^{tA}\}$ is a one-parameter group of linear transformations in $\mathbb{R}^n$, and the phase velocity field is given by $\vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}) = A\vec{\boldsymbol{x}}$. Now we can just apply **Theorem** 1.9. ■

Corollary 1.11.1. *All one parameter groups of linear transformations in $\mathbb{R}^n$ are of the form e^{tA} for some $A \in M_n(\mathbb{R})$.*

Proof. Let $\{g^t\}$ be a one-parameter group of linear transformations on $\mathbb{R}^n$. In section 1.10.2, we showed that the phase velocity vector of $\{g^t\}$ at some $\vec{\boldsymbol{x}} \in \mathbb{R}^n$ is given by $\vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}) = A\vec{\boldsymbol{x}}$, where

$$A = \frac{dg^t}{dt}|_{t=0}$$

From **Theorem** 1.9 it follows that the map $\vec{\psi} : \mathbb{R} \to \mathbb{R}^n$ given by $\vec{\psi}(t) = g^t \vec{a}$ is the unique solution of the IVP $\dot{\vec{x}} = A\vec{x}$, $\vec{x}(0) = \vec{a}$. However by **Theorem** 1.11 we also know that the map $\vec{\varphi}(t) = e^{tA}\vec{a}$ is also a solution of this IVP. So we see that $\vec{\psi} = \vec{\varphi}$, and hence $e^{tA} = g^t$ for all $t \in \mathbb{R}$. ■

1.12. **Jordan Canonical Forms.** A *Jordan block matrix* is a square matrix of the form

$$J = J(\lambda) = J_n(\lambda) = J = \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda \end{bmatrix} \tag{1.1}$$

i.e the main diagonal consists of λ, and the *super-diagonal* consists of only 1. The rest of the entries are 0.

Remark 1.11.1. The following are true of the matrix $J(\lambda)$ above.

(1) λ is an eigenvalue of $J(\lambda)$, and $\{\vec{e}_1\}$ is a basis for the corresponding eigenspace.
(2) Over an algebraically closed field, for any linear transformation $T : V \to V$ on a finite dimensional vector space we can find a basis $\mathcal{B} = \{v_1, ..., v_n\}$ of V such that the matrix of T with respect to $v_1, ..., v_n$ looks something like the below block-matrix form.

$$A = \begin{bmatrix} J_1 & & & \\ & J_2 & 0 & \\ & 0 & \dots & \\ & & & J_k \end{bmatrix}$$

Here each J_i is a Jordan block matrix.
(3) The number of Jordan blocks corresponding to an eigenvalue λ is the geometric multiplicity (i.e the dimension of the corresponding eigenspace) of λ.
(4) The Jordan form above for T is canonical, i.e upto permutation of the blocks any two Jordan forms for T are the same.
(5) The Jordan decomposition given above decomposes V into a direct sum

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

with V_i corresponding to the Jordan block J_i and J_i can be regarded as a linear operator on V_i. The sets $\mathcal{B} \cap V_i$ for $1 \leq i \leq k$ partitions $\mathcal{B}$ into disjoint sets, and for each i, $\mathcal{B} \cap V_i$ is a basis of V_i.
(6) Suppose the field is $\mathbb{C}$. The basis $\mathcal{B}$ can be chosen in a way such that the following holds: if $\mathcal{B} \cap V_i = \{\vec{u}^i_1, ..., \vec{u}^i_{r_i}\}$, then the conjugates $\overline{\vec{u}^i_j}$ for $j = 1, ..., r_i$ are also in $\mathcal{B}$ and form an ordered basis for some V_l in the given decomposition. If the eigenvalue corresponding to V_i is real then $V_l = V_k$, otherwise $V_l \neq V_k$.

1.13. **Real Jordan Canonical Forms.** These are also called *real canonical forms.* Let $T \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^n)$, and let $T_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n$ be its complexification. Now, we can find a basis $\mathcal{B}$ of $\mathbb{C}^n$ such that the matrix of $T_{\mathbb{C}}$ with respect to the basis $\mathcal{B}$ is in Jordan canonical form

$$\begin{bmatrix} J_1 & & & \\ & J_2 & 0 & \\ & 0 & ... & \\ & & & J_t \end{bmatrix}$$

Suppose the Jordan decomposition decomposes $\mathbb{C}^n$ into a direct sum

$$\mathbb{C}^n = V_1 \oplus \cdots V_t$$

Let V be one of the V_k and without loss of generality assume $V = V_1$. Let J be the corresponding Jordan block. Assume the eigenvalue corresponding to J is not real, i.e

$$\lambda \notin \mathbb{R}$$

Let $\mathcal{B} \cap V = \{\vec{\boldsymbol{u}}_1, ..., \vec{\boldsymbol{u}}_r\}$ where the ordering of the subscripts is the same as the ordering in $\mathcal{B}$. We immediately see that

$$\begin{aligned} T_{\mathbb{C}}(\vec{\boldsymbol{u}}_1) &= \lambda \vec{\boldsymbol{u}}_1 \\ T_{\mathbb{C}}(\vec{\boldsymbol{u}}_i) &= \vec{\boldsymbol{u}}_{i-1} + \lambda \vec{\boldsymbol{u}}_i, \quad , \quad i = 2, ..., r \end{aligned}$$

If $\vec{\boldsymbol{w}}_i = \overline{\vec{\boldsymbol{u}}_i}$ for $1 \leq i \leq r$ then as mentioned in point number (6) in **Remark** 1.11.1 we see that $\{\vec{\boldsymbol{w}}_1, ..., \vec{\boldsymbol{w}}_r\}$ is an ordered basis for some $V_l \neq V$. Without loss of generality, we assume $V_l = V_2$ and let $W = V_2$. We see that

$$\begin{aligned} T_{\mathbb{C}}(\vec{\boldsymbol{w}}_1) &= \overline{\lambda} \vec{\boldsymbol{w}}_1 \\ T_{\mathbb{C}}(\vec{\boldsymbol{w}}_i) &= \vec{\boldsymbol{w}}_{i-1} + \overline{\lambda} \vec{\boldsymbol{w}}_i, \quad , \quad i = 2, ..., r \end{aligned}$$

Also, we see that $\{\vec{\boldsymbol{u}}_1, ..., \vec{\boldsymbol{u}}_r, \vec{\boldsymbol{w}}_1, ..., \vec{\boldsymbol{w}}_r\}$ is a basis for $V \oplus W$.

Now, for $1 \leq j \leq r$, let

$$\vec{\boldsymbol{c}}_j = \frac{1}{2}(\vec{\boldsymbol{u}}_j + \vec{\boldsymbol{w}}_j) \quad , \quad \vec{\boldsymbol{d}}_j = \frac{1}{2i}(\vec{\boldsymbol{u}}_j - \vec{\boldsymbol{w}}_j)$$

and we immediately see that for each $1 \leq j \leq r$

$$\vec{\boldsymbol{u}}_j = \vec{\boldsymbol{c}}_j + i\vec{\boldsymbol{d}}_j \quad , \quad \vec{\boldsymbol{w}}_j = \vec{\boldsymbol{c}}_j - i\vec{\boldsymbol{d}}_j$$

and the thing to note is that each $\vec{\boldsymbol{c}}_j, \vec{\boldsymbol{d}}_j \in \mathbb{R}^n$. Also, note that the span of $\{\vec{\boldsymbol{u}}_1, ..., \vec{\boldsymbol{u}}_r, \vec{\boldsymbol{w}}_1, ..., \vec{\boldsymbol{w}}_r\}$ is equal to the span of $\{\vec{\boldsymbol{c}}_1, ..., \vec{\boldsymbol{c}}_r, \vec{\boldsymbol{d}}_1, ..., \vec{\boldsymbol{d}}_r\}$, and hence the set $\{\vec{\boldsymbol{c}}_1, ..., \vec{\boldsymbol{c}}_r, \vec{\boldsymbol{d}}_1, ..., \vec{\boldsymbol{d}}_r\}$ is a basis of $V \oplus W$.

If $\lambda = a + ib$, then it can be shown that

$$T\vec{\boldsymbol{c}}_j = \begin{cases} a\vec{\boldsymbol{c}}_j - b\vec{\boldsymbol{d}}_j & j = 1 \\ \vec{\boldsymbol{c}}_{j-1} + a\vec{\boldsymbol{c}}_j - b\vec{\boldsymbol{d}}_j & 2 \leq j \leq r \end{cases}$$

and that

$$T\vec{\boldsymbol{d}}_j = \begin{cases} b\vec{\boldsymbol{c}}_j + a\vec{\boldsymbol{d}}_j & j = 1 \\ \vec{\boldsymbol{d}}_{j-1} + b\vec{\boldsymbol{c}}_j + a\vec{\boldsymbol{d}}_j & 2 \leq j \leq r \end{cases}$$

So, if we modify our basis so that the $2r$ members are written in the order

$$(*) \qquad \vec{\boldsymbol{c}}_1, \vec{\boldsymbol{d}}_1, \vec{\boldsymbol{c}}_2, \vec{\boldsymbol{d}}_2, ..., \vec{\boldsymbol{c}}_r, \vec{\boldsymbol{d}}_r$$

and if we do this for every pair of eigenvalues which are not real, then the matrix of T with respect to this basis will look as follows: the blocks corresponding to the elements as in (∗) will be of the form:

$$J = J(\lambda, \overline{\lambda}) = \begin{bmatrix} M & I_2 & & & 0 \\ & M & I_2 & & \\ & & M & \ddots & \\ & & & \ddots & I_2 \\ & & & & M \end{bmatrix} \tag{1.2}$$

where M is the 2×2 matrix

$$M = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \tag{1.3}$$

When $\lambda \in \mathbb{R}$, we have a simple Jordan block of the form (1.1). If $\lambda \notin \mathbb{R}$, we have a block of the form (1.2). These blocks are called *real Jordan blocks* or *real canonical blocks*.

1.14. **Revisting Exponentials.** First, we will look at a lemma that is useful for computing exponentials via Jordan decompositions.

Lemma 1.12. *Let $A \in M_n(\mathbb{R})$ and $\Gamma \in \mathrm{GL}_n(\mathbb{R})$. Then the following hold.*

(1) $e^{\Gamma A \Gamma^{-1}} = \Gamma e^A \Gamma^{-1}$. *So, change of basis respects exponentiation in some sense.*

(2) *If*

$$A = \begin{bmatrix} A_1 & 0 & 0 & \cdots & 0 \\ & A_2 & 0 & \cdots & 0 \\ & & A_3 & \cdots & 0 \\ & & & \ddots & 0 \\ & & & & A_t \end{bmatrix}$$

then

$$e^A = \begin{bmatrix} e^{A_1} & 0 & 0 & \cdots & 0 \\ & e^{A_2} & 0 & \cdots & 0 \\ & & e^{A_3} & \cdots & 0 \\ & & & \ddots & 0 \\ & & & & e^{A_t} \end{bmatrix}$$

Proof. For (1), observe that for every $N \in \mathbb{N}$ we have

$$\sum_{m=0}^{N} \frac{(\Gamma A \Gamma^{-1})^m}{m!} = \Gamma \left(\sum_{m=0}^{N} \frac{A^m}{m!} \right) \Gamma^{-1}$$

Taking limits on both sides as $N \to \infty$, we get what we want to prove.

For (2), note that

$$A = \begin{bmatrix} A_1 & 0 & 0 & \cdots & 0 \\ & 0 & 0 & \cdots & 0 \\ & & 0 & \cdots & 0 \\ & & & \ddots & 0 \\ & & & & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ & A_2 & 0 & \cdots & 0 \\ & & 0 & \cdots & 0 \\ & & & \ddots & 0 \\ & & & & 0 \end{bmatrix} + \ldots + \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ & 0 & 0 & \cdots & 0 \\ & & 0 & \cdots & 0 \\ & & & \ddots & 0 \\ & & & & A_t \end{bmatrix}$$

and clearly, the matrices in the above sum commute with each other. So keeping **Theorem** 1.10 in mind, we can assume without loss of generality that

$$A_2 = ... = A_t = 0$$

Let the size of the block A_i be $r_i \times r_i$. Put

$$Z_N = \sum_{m=0}^{N} \frac{A_1^m}{m!} - e^{A_1}$$

and let

$$E = \begin{bmatrix} e^{A_1} & & & \cdots & \\ & I_{r_2} & & \cdots & \\ & & I_{r_3} & \cdots & \\ & & & \ddots & \\ & & & & I_{r_t} \end{bmatrix}$$

Then

$$\sum_{m=0}^{N} \frac{A^m}{m!} - E = \begin{bmatrix} Z_N & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}$$

which is easy to see. Moreover, it is easy to see that the operator norm of the RHS above is $||Z_N||_\circ$, which easily follows from the definition of the operator norm. So, it follows that

$$\sum_{m=0}^{\infty} \frac{A^m}{m!} = E$$

which means that $e^A = E$. And this is what we had to show. ∎

1.15. **Structure of solutions of homogeneous linear DEs.** Let A be a constant $n \times n$ matrix over $\mathbb{R}$, and consider the DE

$$\dot{\vec{x}} = A\vec{x}$$

Let the real Jordan form of A be

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{bmatrix}$$

There is some $\Gamma \in \mathrm{GL}_n(\mathbb{R})$ such that

$$A = \Gamma J \Gamma^{-1}$$

As we saw in section 1.13, the blocks J_i for $1 \leq i \leq t$ are either of the form (1.1) if $\lambda \in \mathbb{R}$, or they are of the form (1.2) if $\lambda \notin \mathbb{R}$. Here λ's are the eigenvalues of A.

Now suppose $\lambda_i \in \mathbb{R}$ and let J_i be the corresponding Jordan block, and let it's size be $r_i \times r_i$. So we see that

$$J_i = \lambda_i I_{r_i} + B$$

where B is the $r_i \times r_i$ matrix whose super-diagonal consists only of 1's. Since $\lambda_i I_{r_i}$ and B commute, we see that

$$\begin{aligned} e^{tJ_i} = e^{t\lambda_i I_{r_i} + tB} &= e^{t\lambda_i I_{r_i}} e^{tB} \\ &= e^{t\lambda_i} I_{r_i} e^{tB} \\ &= e^{t\lambda_i} I_{r_i} \begin{bmatrix} 1 & t & t^2/2! & \dots & t^{r_i-1}/(n-1)! \\ & 1 & t & \dots & t^{r_i-2}/(r_i-2)! \\ & & 1 & \dots & t^{r_i-3}/(r_i-3)! \\ & & & \dots & \dots \\ & & & \dots & 1 \end{bmatrix} \end{aligned}$$

where in the last step, we have used **Problem 2** of **QUIZ-1**.

On the other hand, if $\lambda_i \notin \mathbb{R}$, then the corresponding Jordan block J_i is of the form (1.2), i.e

$$J_i = \begin{bmatrix} M_i & I_2 & & & 0 \\ & M_i & I_2 & & \\ & & M_i & \ddots & \\ & & & \ddots & I_2 \\ & & & & M_i \end{bmatrix}$$

where M_i is of the form (1.3) with $a = a_i, b = b_i$ being the real and imaginary parts of λ. In that case, by problem **4).** of **HW-5** we see that

$$e^{tJ_i} = \begin{bmatrix} e^{tM_i} & te^{tM_i} & \frac{t^2}{2!}e^{tM_i} & \cdots & \frac{t^{(r_i-1)}}{(r_i-1)!}e^{tM_i} \\ & e^{tM_i} & te^{tM_i} & \cdots & \frac{t^{(r_i-2)}}{(r_i-2)!}e^{tM_i} \\ & & e^{tM_i} & \cdots & \frac{t^{(r_i-3)}}{(r_i-3)!}e^{tM_i} \\ & & & \ddots & \\ & & & \cdots & te^{tM_i} \\ & & & \cdots & e^{tM_i} \end{bmatrix}$$

and

$$e^{tM_i} = \begin{bmatrix} e^{ta_i}\cos(tb_i) & -e^{ta_i}\sin(tb_i) \\ e^{ta_i}\sin(tb_i) & e^{ta_i}\cos(tb_i) \end{bmatrix}$$

where a_i, b_i are the real and imaginary parts of λ respectively.

Now, from **Theorem** 1.11 we know that the solutions of $\dot{\vec{x}} = A\vec{x}$ are of the form $e^{tA}\vec{a}$ (where $\vec{a} \in \mathbb{R}^n$ is fixed but arbitrary). More precisely, the solution of the IVP $\dot{\vec{x}} = A\vec{x}$, $\vec{x}(0) = \vec{a}$ is $\vec{\varphi}(t) = e^{tA}\vec{a}$. So using **Lemma** 1.12, we see that the solutions of the DE are of the form $\vec{\varphi} : \mathbb{R} \to \mathbb{R}^n$ where $\vec{\varphi} = (\varphi_1, ..., \varphi_n)$ and $\varphi_i(t)$ is a linear combination of $\{t^j e^{t\lambda_i} \mid 1 \leq j \leq r_i - 1, \lambda \in \mathbb{R}\}$, $\{t^j e^{ta_i}\cos(tb_i) \mid 1 \leq j \leq r_i - 1, \lambda_i = a_i + ib_i, b_i \neq 0\}$ and $\{t^j e^{ta_i}\sin(tb_i) \mid 1 \leq j \leq r_i - 1, \lambda_i = a_i + ib_i, b_i \neq 0\}$ as i ranges from 1 to t.

Remark 1.12.1. It is not being claimed that *every* linear combination is possible for each entry independent of the other entries, because in that case we will have n^2 degrees of freedom. But, we know that the space of solutions of the DE has dimension n. Moreover, two distinct Jordan blocks J_i, J_j may have the same associated eigenvalue (real of complex).

1.16. **Scalar n^{th} order Linear DEs (homogeneous).** Consider the DE

$$(*) \qquad y^{(n)} + a_0 y^{(n-1)} + ... + a_{n-1}y' + a_n y = 0$$

We know that this is equivalent to the vector valued DE

$$(**) \qquad \dot{\vec{x}} = A\vec{x}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ & 0 & 1 & \dots & 0 \\ & & 0 & \ddots & \vdots \\ & & & \ddots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}$$

i.e the super-diagonal of A is all 1's, and the last row is the vector $-(a_0, ..., a_{n-1})$. If φ is a solution of $(*)$, then $\vec{\varphi} = (\varphi, \varphi', ..., \varphi^{(n-1)})$ is a solution of $(**)$ and conversely if $\vec{\varphi} = (\varphi_1, ..., \varphi_n)$ is a solution of $(**)$ then $\varphi = \varphi_1$ is a solution of $(*)$. By problem **1).** of **HW-5**, the characteristic polynomial of the DE $(*)$ is (upto sign) equal to the characteristic polynomial of the matrix A.

Let

$$\sigma(A) = \{\lambda \in \mathbb{C} \mid \lambda \text{ is an eigenvalue of } A\}$$

and let

$$S(A) = \sigma(A)/R$$

where R is the equivalence relation on $\sigma(A)$ given by

$$\lambda R \lambda' \iff \lambda' \in \{\lambda, \overline{\lambda}\}$$

Note that

$$\sigma(A) = \sigma_1(A) \cup \sigma_2(A)$$

where $\sigma_1(A) = \sigma(A) \cap \mathbb{R}$ and $\sigma_2(A) = \sigma(A) - \sigma_1(A)$. It is also easy to see that

$$S(A) = S_1(A) \sqcup S_2(A)$$

where $S_1(A) = \sigma_1(A)/R$ and $S_2(A) = \sigma_2(A)/R$, and $\sqcup$ represents a disjoint union. By problem **2).** of **HW-5** there is a one-to-one correspondence between (complex) Jordan blocks of A and elements of $\sigma(A)$. Our discussion on real Jordan blocks then tells us that there is a one-to-one correspondence between $S(A)$ and the number of real Jordan blocks.

Let $s = [\lambda]$ denote the equivalence class of $\lambda \in \sigma(A)$. Let the corresponding (real) Jordan block be $\tilde{J}_s$. There is a well-defined multiplicity of s, namely the multiplicity of λ representing s, as a root of the characteristic polynomial. Let this number be r_s. This is half the size of $\tilde{J}_s$ if $s \in S_2(A)$ otherwise it is the size of $\tilde{J}_s$.

If $s \in S_2(A)$ let a_s and $b_s > 0$ be real numbers such that $a_s + ib_s$ represents s, i.e $s = [a_s + ib_s]$. Since we are insisting $b_s > 0$, this is well-defined.

From the discussion in section 1.15 it follows that the solutions of $(*)$ are in the linear span of

$$Q = \bigcup_{s \in S_1(A)} \bigcup_{j=0}^{r_s - 1} \{t^j e^{t\lambda_s}\} \cup \bigcup_{s \in S_2(A)} \bigcup_{j=0}^{r_s-1} \{t^j e^{a_s t} \cos b_s t, t^j e^{a_s t} \sin b_s t\}$$

Since the span of solutions of $(**)$ has dimension n over $\mathbb{R}$ (using a homework problem), it follows that the elements in Q (which are n in number) are linearly independent and form a basis for the space of solutions of $(*)$.

2. Manifolds

An *n-dimensional manifold* M is a Hausdorff, second countable topological space together with data $\mathscr{A}$, called an *atlas*. $\mathscr{A}$ is a collection of pairs of the form (U, φ) (such a pair is called a *chart*) with U an open subset of M, $\varphi : U \to \varphi(U)$ a homeomorphism, where $\varphi(U) \subset \mathbb{R}^n$ is an open set, such that the following hold.

(1) $M = \bigcup_{(U,\varphi)\in\mathscr{A}} U$

(2) If (U, φ) and (V, ψ) are two charts in $\mathscr{A}$ then the map θ given by the commutative diagram below is a diffeomorphism.

$$\begin{array}{ccccc} & & U \cap V & & \\ & \swarrow^{\text{via } \psi} & & \searrow^{\text{via } \varphi} & \\ \psi(U \cap V) & & \xrightarrow[\theta]{\sim} & & \varphi(U \cap V) \end{array}$$

To represent this situation in a more geometrical way, we have the following picture.

2.1. The Tangent Bundle.

Let M be an n-dimensional manifold. For any $\boldsymbol{p} \in M$, we use the notation $T_{\boldsymbol{p}}$ to denote the space $\mathscr{D}\mathrm{er}_{\boldsymbol{p}}$, i.e the space of all $\mathbb{R}$-linear maps $D : \mathscr{C}^\infty_{\boldsymbol{p}} \to \mathbb{R}$ which satisfy

$$D(fg) = f(\boldsymbol{p})Dg + g(\boldsymbol{p})Df$$

for all $f, g \in \mathscr{C}^\infty(M)$. Here, $\mathscr{C}^\infty_{\boldsymbol{p}}$ is the germ of such functions, as defined in the notes on vector fields.

Put

$$T(M) := \bigcup_{\boldsymbol{p}\in M} T_{\boldsymbol{p}}$$

The above union is a disjoint union because $T_{\boldsymbol{a}} \cap T_{\boldsymbol{b}} = \phi$ if $\boldsymbol{a} \neq \boldsymbol{b}$. So, we have a map $\omega : T(M) \to M$ such that $\omega(D) = \boldsymbol{p}$ if $D \in T_{\boldsymbol{p}}$. Here, $T_{\boldsymbol{p}} = \omega^{-1}(\boldsymbol{p})$. The *tangent bundle* of M is the pair $(T(M), \omega)$.

Example 2.1. Suppose M is an open subset of $\mathbb{R}^n$ and $\vec{\boldsymbol{v}} = \dot{\vec{\boldsymbol{\gamma}}}$ is a velocity vector of some path γ passing through $\vec{\boldsymbol{p}}$ with $\gamma(0) = \vec{\boldsymbol{p}}$. Then $f \mapsto \left.\dfrac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f \circ \gamma)$ is a derivation, and all derivations look like this for such M.

2.2. DEs over Manifolds.

In this section, we will try to make sense of differential equations over manifolds.

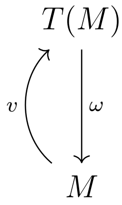
Given a $\mathscr{C}^1$ path $\gamma : I \to M$ with I an open interval, and a point $t \in I$, define $\dot\gamma(t)$ to be the derivation on $\mathscr{C}^\infty_{\gamma(t)}$ given by

$$(\dot\gamma(t))(f) = \left.\frac{\mathrm{d}}{\mathrm{d}s}(f \circ \gamma)(s)\right|_{s=t} \qquad (f \in \mathscr{C}^\infty_{\gamma(t)})$$

It is clear that the right hand side above can be evaluated by taking a representation for a germ, and the choice of the representative will not matter in the calculation. So, $\dot\gamma(t) \in T_{\gamma(t)}(M)$ for $t \in I$. So, we have a map $\dot\gamma : I \to T(M)$. Since γ is $\mathscr{C}^1$, it is not hard to see that $\dot\gamma$ is continuous. Also, the following diagram commutes.

$$\begin{array}{ccc} & & T(M) \\ & \nearrow^{\dot\gamma} & \downarrow{\scriptstyle\omega} \\ I & \xrightarrow[\gamma]{} & M \end{array}$$

Definition 2.1. A *vector field* v on M is a map $v : M \to T(M)$ such that for every $p \in M$, $v(p) \in T_p(M)$, i.e $v(p)$ is a derivation on $\mathscr{C}_p^\infty$ at p. Equivalently, v is a *vector field* if v is such that $\omega \circ v = 1_M$.



If v is a vector field on M, then we have a DE of the form

$$\dot{x} = v(x)$$

A solution is a $\mathscr{C}^1$ path $\varphi : I \to M$ on an open interval I such that $\dot{\varphi}(t) = v(\varphi(t))$ for $t \in I$, where the equality is an equality of points in $T(M)$.

If $p \in M$ and $\tau \in \mathbb{R}$ we can talk about the IVP

$$\dot{x} = v(x) \quad , \quad x(\tau) = p$$

where the meaning of a solution φ to this IVP is clear.

2.2.1. *The non-autonomous case.* Let Ω be an open subset of $\mathbb{R} \times M$. Let $\pi : \Omega \to M$ be the second projection. We regard Ω as the extended phase space. Let $v : \Omega \to T(M)$ be a continuous map such that $\omega \circ v = \pi$ (this is very similar to the vector field definition). The corresponding DE here is

$$\dot{x} = v(t, x)$$

A solution is a $\mathscr{C}^1$ map $\varphi : I \to M$ on an open interval I such that $(t, \varphi(t)) \in \Omega$ for all $t \in I$ and $\dot{\varphi}(t) = v(t, \varphi(t))$ for $t \in I$. Ofcourse, we can have an IVP

$$\dot{x} = v(t, x) \quad , \quad x(\tau) = p$$

where $(\tau, p) \in \Omega$, and the meaning of a solution to the IVP is also clear.

2.3. **First Integrals.** Let v be a smooth vector field on a manifold M. A *first integral* for v is a smooth function f on an open set in M such that f is not constant on any open subset of its domain and $v(f) = 0$. The notation is: $v(p)$ is a derivation for each $p \in M$, and $v(f)$ is the function on the domain of f that maps a point p to $v(p)(f)$.

If M is a domain in $\mathbb{R}^n$, and if we interpret $\vec{\boldsymbol{v}}$ to be a smooth map from M to $\mathbb{R}^n$ where $\vec{\boldsymbol{v}} = (v_1, ..., v_n)$ then as proven in the supplementary notes on vector fields we have

$$\vec{\boldsymbol{v}}(f) = \sum_i v_i \frac{\partial f}{\partial x_i} = \langle \nabla(f), \vec{\boldsymbol{v}} \rangle$$

So, if f is a first integral for $\vec{\boldsymbol{v}}$, then $\nabla(f)$ and $\vec{\boldsymbol{v}}$ are orthogonal. Let f be a first integral for $\vec{\boldsymbol{v}}$, $\vec{\boldsymbol{p}}$ a point in the domain of f and let $c = f(\vec{\boldsymbol{p}})$. Let S be the hypersurface $S = f^{-1}(f(\vec{\boldsymbol{p}})) = f^{-1}(c)$. Since $\nabla f(\vec{\boldsymbol{p}})$ is orthogonal to the level set S, it follows that $\vec{\boldsymbol{v}}(\vec{\boldsymbol{p}})$ is tangential to S at $\vec{\boldsymbol{p}}$. Moreover, f is constant along any phase curve of $\vec{\boldsymbol{v}}$.

Example 2.2. Look at **Example 1.1.1** in Lecture 14.

2.4. **Integral Hypersurfaces.** We are interested in solving the DE

$$\dot{x} = v(x)$$

on a manifold M, where $v : M \to T(M)$ is a $\mathscr{C}^1$ vector field.

Temporarily, we assume that M is an open subset of $\mathbb{R}^n$. Using vector notations, we can write

$$\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}) \tag{2.1}$$

Let the components of $\vec{\boldsymbol{v}}$ be $(v_1, ..., v_n)$. Let $g : M \to \mathbb{R}$ be a first integral for $\vec{\boldsymbol{v}}$ so that

$$\sum_{i=1}^{n} v_i \frac{\partial g}{\partial x_i} \equiv 0 \qquad \text{on } M$$

Now if I is an open interval of existence of (2.1) and $\vec{\boldsymbol{\varphi}} : I \to M$ is a solution of (2.1) then we claim that $g \circ \varphi : I \to \mathbb{R}$ is a constant function. This is actually a consequence of the chain rule, because

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\{(g \circ \vec{\boldsymbol{\varphi}})(t)\} &= \sum_{i=1}^{n} \left(\frac{\partial g}{\partial x_i}(\vec{\boldsymbol{\varphi}}(t)) \right) \dot{\varphi}_i(t) \\
&= \sum_{i=1}^{n} \left(\frac{\partial g}{\partial x_i}(\vec{\boldsymbol{\varphi}}(t)) \right) v_i(\vec{\boldsymbol{\varphi}}(t)) \\
&= \vec{\boldsymbol{v}}(\vec{\boldsymbol{\varphi}}(t))(g) \\
&= 0
\end{aligned}$$

since g is a first integral for $\vec{\boldsymbol{v}}$.

Now, suppose $c \in g(M)$, where g is as above. Let $S \subseteq M$ be the hypersurface $g = c$. Such surfaces are called *integral hypersurfaces*, since they come from first integrals. Suppose $\vec{\boldsymbol{p_0}} \in S$. Let $t_0 \in \mathbb{R}$, and let $\vec{\boldsymbol{\varphi}} : I \to M$ be a solution of the IVP

$$\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}) \quad , \quad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{p_0}}$$

where I is an open interval of existence. Above we have seen that $g \circ \vec{\boldsymbol{\varphi}}$ is a constant function, which means that

$$g(\vec{\boldsymbol{\varphi}}(t)) = g(\vec{\boldsymbol{\varphi}}(t_0)) = g(\vec{\boldsymbol{p_0}}) = c$$

for all $t \in I$. So, $\vec{\boldsymbol{\varphi}}(t) \in S$ for all $t \in I$. So we have proved the following result.

Proposition 2.1. *Let $\vec{\boldsymbol{\varphi}} : I \to M$ be a solution of the IVP* (2.1) *and suppose for some $t_0 \in I$, $\vec{\boldsymbol{\varphi}}(t_0) \in S$. Then $\vec{\boldsymbol{\varphi}}(t) \in S$ for all $t \in I$.*

2.5. $n-1$ **first integrals.** Now suppose we have $n-1$ first integrals $f_1, ..., f_{n-1}$ for $\vec{\boldsymbol{v}}$. Let

$$\vec{\boldsymbol{f}} = (f_1, ..., f_{n-1})$$

Let $\vec{\boldsymbol{c}} = (c_1, ..., c_{n-1}) \in \mathbb{R}^{n-1}$ be in the image of $\vec{\boldsymbol{f}}$, and let

$$C = \vec{\boldsymbol{f}}^{-1}(\vec{\boldsymbol{c}}) \subseteq M$$

Assume that

$$\operatorname{rank}(\vec{\boldsymbol{f}}'(\vec{\boldsymbol{p}})) = n-1 \qquad \forall \vec{\boldsymbol{p}} \in C$$

In such a case, we say that $f_1, ..., f_{n-1}$ are *functionally independent*. Then from the Implicit Function Theorem, we know that C is a one-dimensional manifold. For instance, if the last $n-1$ columns of $J\vec{\boldsymbol{f}}(\vec{\boldsymbol{p}})$ are linearly independent at some point $\vec{\boldsymbol{p}} = (p_1, ..., p_n)$

then there is an open interval $(p_1-\epsilon, p_1+\epsilon)$ and a map $\vec{\varphi}:(p_1-\epsilon,p_1+\epsilon)\to\mathbb{R}^{n-1}$ such that

$$(x,\varphi_1(x),...,\varphi_{n-1}(x))\in C \qquad x\in(p_1-\epsilon,p_1+\epsilon)$$

and such that $\varphi_i(p_1)=p_{i+1}$ for $i=1,...,n-1$.

Let $\vec{p_0}\in C$. Let $\vec{\varphi}:J\to M$ be a maximal solution of the IVP (with $t_0\in J$ fixed)

$$\dot{\vec{x}}=\vec{v}(\vec{x}) \quad , \quad \vec{x}(t_0)=\vec{p_0}$$

From **Proposition** 2.1 we see that

$$\vec{\varphi}(t)\in C \qquad \forall t\in J$$

Indeed, if S_i is the hypersurface $f_i^{-1}(c_i)$, then **Proposition** 2.1 implies that $\vec{\varphi}(t)\in S_i\ \forall t\in J$, and hence $\vec{\varphi}(t)\in\cap_{i=1}^{n-1}S_i=C$.

Again, note that we are interested in the DE

$$\dot{\vec{x}}=\vec{v}(\vec{x})$$

and so it is enough to concentrate on the regular locus of $\vec{v}$, because if $\vec{v}(\vec{p_0})=0$, then the only solution of $\dot{\vec{x}}=\vec{v}(\vec{x})$, $\vec{x}(t_0)=\vec{p_0}$ is the constant solution $\vec{\varphi}=\vec{p_0}$. So from now on, we assume that $\vec{v}$ is nowhere vanishing on M.

Let $\vec{f}$, $\vec{c}=(c_1,...,c_{n-1})$ and $C=\vec{f}^{-1}(\vec{c})$ be as in the beginning of this section. From ANA2, the space of velocity vectors at a point $\vec{p}\in C$ is the null space of $\vec{f}'(\vec{p})$. Moreover, we know that $\vec{v}(f_i)\equiv 0$ on M for $i=1,...,n-1$, since the f_i's are first integrals. This means that

$$\sum_{i=1}^{n}v_1(\vec{p})\frac{\partial f_i}{\partial x_1}(\vec{p})+\cdots+v_n(\vec{p})\frac{\partial f_i}{\partial x_n}(\vec{p})=0$$

for each $1\le i\le n-1$, where $\vec{v}=(v_1,...,v_n)$. This implies that $\vec{v}(\vec{p})$ lies in the null space of $\vec{f}'(\vec{p})$. Now, as remarked before, C can be locally parametrized because it is a one-dimensional manifold, which follows from the Implicit Function Theorem. Thus, if $\vec{p_0}\in C$, there is an open neighborhood of $\vec{p}_0$ in C which is homeomorphic to $(-\epsilon,\epsilon)$ in $\mathbb{R}$. So, C can be described locally as

$$\vec{\gamma}(t)=(\gamma_1(\lambda),...,\gamma_n(\lambda))$$

in parametric form: C is locally

$$\begin{aligned}x_1&=\gamma_1(\lambda)\\x_2&=\gamma_2(\lambda)\\&\vdots\\x_n&=\gamma_n(\lambda)\end{aligned}$$

Moreover, $\frac{\mathrm{d}\vec{\gamma}}{\mathrm{d}\lambda}(\lambda)$ does not vanish since $\vec{\gamma}$ is a diffeomorphism to C. Since the null space of $\vec{f}'(\vec{p_0})$ is one-dimensional (since its rank is $n-1$), and since neither $\vec{v}(\vec{p_0})$ nor $\vec{\gamma}'(\lambda_0)$ vanish (where λ_0 is such that $\vec{\gamma}(\lambda_0)=\vec{p_0}$) therefore each is a non-zero multiple of the other, because $\vec{\gamma}'(\lambda_0)$ is a velocity vector passing through $\vec{p_0}$.

In general we therefore get a nowhere vanishing function u such that

$$\vec{\gamma}'(\lambda)=u(\lambda)\vec{v}(\vec{\gamma}(\lambda)) \qquad , \qquad \lambda\in I$$

for some interval I. It is easy to see that $u(\lambda)$ is continuous, being the ratio of two continuous functions.

Now fix $\vec{\boldsymbol{p_0}} \in C$ as above. Again, consider the IVP

$$\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}) \quad , \quad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{p_0}}$$

and suppose λ_0 is in the domain of $\vec{\boldsymbol{\gamma}}$ such that $\vec{\boldsymbol{\gamma}}(t_0) = \vec{\boldsymbol{p_0}}$. Let

$$t = t_0 + \int_{\lambda_0}^{\lambda} u(y)\, \mathrm{d}y \tag{2.2}$$

n Since $u(y)$ does not vanish anywhere on I, the above function is a monotone function of λ, i.e t is a monotone function of λ. It can be inverted, and hence we see that

$$\lambda = \lambda(t)$$

which we obtain by solving for λ in the equation (2.2). Consider the function

$$\vec{\boldsymbol{\varphi}}(t) = \vec{\boldsymbol{\gamma}}(\lambda(t))$$

Then we see that

$$\begin{aligned}
\dot{\vec{\boldsymbol{\varphi}}}(t) &= \frac{\mathrm{d}\lambda}{\mathrm{d}t}\frac{\mathrm{d}\vec{\boldsymbol{\gamma}}}{\mathrm{d}\lambda} \\
&= \frac{\mathrm{d}\lambda}{\mathrm{d}t}(u(\lambda(t))\vec{\boldsymbol{v}}(\vec{\boldsymbol{\gamma}}(\lambda(t)))) \\
&= \frac{\mathrm{d}\lambda}{\mathrm{d}t}(u(\lambda(t))\vec{\boldsymbol{v}}(\vec{\boldsymbol{\varphi}}(t))) \\
&= \frac{1}{\mathrm{d}t/\mathrm{d}\lambda}(u(\lambda(t))\vec{\boldsymbol{v}}(\vec{\boldsymbol{\varphi}}(t))) \\
&= \frac{1}{u(\lambda(t))}(u(\lambda(t))\vec{\boldsymbol{v}}(\vec{\boldsymbol{\varphi}}(t))) \\
&= \vec{\boldsymbol{v}}(\vec{\boldsymbol{\varphi}}(t))
\end{aligned}$$

and hence $\vec{\boldsymbol{\varphi}}$ is a solution of our IVP.

2.6. **DEs on Compact Manifolds.** The main result of this section will be the following.

Proposition 2.2. *Let M be a compact manifold, v a $\mathscr{C}^1$ vector field on M and x_0 a point on M. Then the maximal interval of existence for the IVP*

$$\dot{x} = v(x) \quad , \quad x(0) = x_0$$

is $\mathbb{R}$.

Proof. Let $\varphi_a : J(a) \to M$ be the maximal solution for the IVP

$$\dot{x} = v(x) \qquad , x(0) = a$$

Soon, we will prove that there exists a neighborhood $(-\epsilon_a, \epsilon_a) \times U_a$ in $\mathbb{R} \times M$ of $(0, a)$ where U_a is an open neighborhood of a in M such that for all $b \in U_a$, $(-\epsilon_a, \epsilon_a) \subseteq J(b)$.

The U_a's cover M. Since M is compact, there is a finite open cover $U_{a_1}, ..., U_{a_n}$. Let $\epsilon = \min\{\epsilon_{a_1}, ..., \epsilon_{a_n}\}$. Then $(-\epsilon, \epsilon) \subseteq J(b)$ for all $b \in M$. It follows that the IVP $\dot{x} = v(x)$, $x(0) = a$ has a solution in $(-\epsilon, \epsilon)$ for all $a \in M$, and ϵ is independent of a. By Problem 7 of the mid-term exam, we are done. ∎

2.7. **The Logistic Equation.** Consider the DE

$$\frac{\mathrm{d}y}{\mathrm{d}t} = y(1-y)$$

and let $y(0) = y_0$. So, we have that $v(y) = y(1-y)$. When $y_0 = 0, 1$, then the solutions by uniqueness are $y \equiv 0, y \equiv 1$ respectively. Suppose now that y_0 is a regular point, i.e $y_0(1-y_0) \neq 0$. Solving this DE, we get

$$\ln\left|\frac{y}{y-1}\right| = t + \ln\left|\frac{y_0}{y_0-1}\right|$$

which gives us

$$t = \ln\left|\frac{y_0-1}{y_0} \cdot \frac{y}{y-1}\right|$$

Now, if y is a solution to the IVP, we see that $y(t) \notin \{0,1\}$ for all t in the maximal interval of existence, and this is an easy consequence of uniqueness of solutions, since we assumed that $y_0 \notin \{0,1\}$. Now suppose C is one of the connected components of $\mathbb{R} \setminus \{0,1\}$. From what we just remarked, it follows that if $y_0 \in C$, then $y(t) \in C$ for all t in the maximal interval of existence (because y is continuous). So, it follows that

$$\frac{y_0-1}{y-1} > 0 \quad , \quad \frac{y}{y_0} > 1$$

and so we can write

$$t = \ln\left\{\frac{y_0-1}{y_0}\frac{y}{y-1}\right\}$$

and so we get

$$y(t) = \frac{y_0e^t}{y_0e^t - y_0 + 1} \quad , \quad t \in J(y_0)$$

where $J(y_0)$ is the maximal interval of existence. Note that the above formula works even when $y_0 = 0$ or $y_0 = 1$. Consider the following.

(1) Note that if $y_0 \in (0,1)$, then $J(y_0) = \mathbb{R}$. This can be seen by noting that in this case, $y_0e^t - y_0 + 1 > 0$ for all $t \in \mathbb{R}$.
(2) Now suppose $y_0 > 1$. Then
$$y_0e^t - y_0 + 1 = 0$$
has a solution (in t), namely
$$t_\infty = \ln\frac{y_0-1}{y_0}$$
and so in this case $J(y_0) = (t_\infty, \infty)$ and also $t_\infty < 0$.
(3) Finally suppose $y_0 < 0$. Once again, $\frac{y_0-1}{y_0} > 0$, and in fact $\frac{y_0-1}{y_0} > 1$. Also, the equation
$$y_0e^t - y_0 + 1 = 0$$
has $t_\infty = \ln\frac{y_0-1}{y_0}$ as a solution. Therefore $J(y_0) = (t_\infty, \infty)$ and in this case $t_\infty > 0$.

In cases (2) and (3) above, our solution does not extend to all of $\mathbb{R}$. The problem is that $\mathbb{R}$ is not compact. So in the next section, we will try to *compactify* $\mathbb{R}$, and study the same equation over that compact set.

2.8. **The logistic equation on a circle.** See the instructor notes for lectures 17 and 18.

2.9. **Change of coordinates.** Let Ω be open in $\mathbb{R}^{n+1}$ and assume that $\Omega = I \times U$ where I is an open interval in $\mathbb{R}$ and U is open in $\mathbb{R}^n$. Let $\vec{\boldsymbol{F}} : U \xrightarrow{\sim} W$ be a diffeomorphism, where W is an open subset of $\mathbb{R}^n$, and suppose that $\vec{\boldsymbol{F}}$ is atleast $\mathscr{C}^2$. Let $\Omega' = I \times W$. Let $\vec{\boldsymbol{v}} : \Omega \to \mathbb{R}^n$ be a $\mathscr{C}^1$ vector field. Let

$$\vec{\boldsymbol{G}} = \vec{\boldsymbol{F}}^{-1}$$

Let $\vec{\boldsymbol{w}} : \Omega' \to \mathbb{R}^n$ be the map

$$\vec{\boldsymbol{w}}(t, \vec{\boldsymbol{y}}) = \vec{\boldsymbol{F}}'(\vec{\boldsymbol{G}}\vec{\boldsymbol{y}})\vec{\boldsymbol{v}}(t, \vec{\boldsymbol{G}}\vec{\boldsymbol{y}}) \tag{2.3}$$

where J is the Jacobian. Then $\vec{\boldsymbol{w}}$ is a $\mathscr{C}^1$ map, being a composite of $\mathscr{C}^1$ maps.

Proposition 2.3. *A map $\vec{\boldsymbol{\varphi}} : I \to U$ on an open interval I is a solution of the DE $\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}})$ if and only if the map $\vec{\boldsymbol{\psi}} := \vec{\boldsymbol{F}} \circ \vec{\boldsymbol{\varphi}}$ is a solution of $\dot{y} = \vec{\boldsymbol{w}}(t, \vec{\boldsymbol{y}})$.*

Proof. This is a straightforward application of the chain rule. Since $\vec{\boldsymbol{F}}$ and $\vec{\boldsymbol{G}}$ are inverse functions, we see that

$$(\vec{\boldsymbol{F}}')^{-1}(\vec{\boldsymbol{x}}) = \vec{\boldsymbol{G}}'(\vec{\boldsymbol{F}}(\vec{\boldsymbol{x}}))$$

for any $\vec{\boldsymbol{x}} \in U$. So by equation (2.3) it follows that

$$\vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}}) = \vec{\boldsymbol{G}}'(\vec{\boldsymbol{F}}(\vec{\boldsymbol{x}}))\vec{\boldsymbol{w}}(t, \vec{\boldsymbol{F}}(\vec{\boldsymbol{x}}))$$

and this establishes a symmetry between $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$. So, it is enough to prove one direction of the proposition. Suppose $\vec{\boldsymbol{\varphi}} : I \to U$ is a solution of the DE $\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}})$ and let $\vec{\boldsymbol{\psi}} = \vec{\boldsymbol{F}} \circ \vec{\boldsymbol{\varphi}}$, where I is some open interval. Then we have the following.

$$\begin{aligned}
\dot{\vec{\boldsymbol{\psi}}}(t) &= \vec{\boldsymbol{F}}'(\vec{\boldsymbol{\varphi}}(t))\dot{\vec{\boldsymbol{\varphi}}}(t) \\
&= \vec{\boldsymbol{F}}'(\vec{\boldsymbol{\varphi}}(t))\vec{\boldsymbol{v}}(t, \vec{\boldsymbol{\varphi}}(t)) \\
&= \vec{\boldsymbol{F}}'(\vec{\boldsymbol{G}}(\vec{\boldsymbol{\psi}}(t)))\vec{\boldsymbol{v}}(t, \vec{\boldsymbol{G}}(\vec{\boldsymbol{\psi}}(t))) \\
&= \vec{\boldsymbol{w}}(t, \vec{\boldsymbol{\psi}}(t))
\end{aligned}$$

and hence $\vec{\boldsymbol{\psi}}$ is a solution of the DE $\dot{\vec{\boldsymbol{y}}} = \vec{\boldsymbol{w}}(t, \vec{\boldsymbol{y}})$. This completes the proof. ■

2.10. **Estimates.** Let $\vec{\boldsymbol{v}} : \Omega \to \mathbb{R}^n$ be a continuous function which is Lipschitz in $\vec{\boldsymbol{x}}$ with Lipschitz constant L on a domain Ω contained in $\mathbb{R} \times \mathbb{R}^n$. We will consider the DE

$$\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}})$$

and we will be interested in the behavior of this DE as we vary the initial conditions, i.e the initial time point and the initial phase. Let $\vec{\boldsymbol{\xi}} = (\tau, \vec{\boldsymbol{a}})$ be a point in Ω. The symbol $(\Delta)_{(\tau, \vec{\boldsymbol{a}})}$ or $(\Delta)_{\vec{\boldsymbol{\xi}}}$ will denote the IVP

$$\dot{\vec{\boldsymbol{x}}} = \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{x}}) \quad , \quad \vec{\boldsymbol{x}}(\tau) = \vec{\boldsymbol{a}} \tag{$(\Delta)_{(\tau, \vec{\boldsymbol{a}})}$}$$

Also, $J(\vec{\boldsymbol{\xi}})$ or $J(\tau, \vec{\boldsymbol{a}})$ will denote the maximal interval of existence for solutions of $(\Delta)_{(\tau, \vec{\boldsymbol{a}})}$, and $\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}$ or $\vec{\boldsymbol{\varphi}}_{(\tau, \vec{\boldsymbol{a}})}$ will denote the solution of $(\Delta)_{(\tau, \vec{\boldsymbol{a}})}$.

Definition 2.2. We say that $\vec{\boldsymbol{\varphi}}$ is an ϵ-*approximate* solution of $(\Delta)_{(\tau, \vec{\boldsymbol{a}})}$ on an interval I if $(t, \vec{\boldsymbol{\varphi}}(t)) \in \Omega$ for all $t \in I$ and

$$||\dot{\vec{\boldsymbol{\varphi}}}(t) - \vec{\boldsymbol{v}}(t, \vec{\boldsymbol{\varphi}}(t))|| < \epsilon \qquad \forall\ t \in I$$

Now suppose $\vec{\varphi}$ and $\vec{\psi}$ are $\mathscr{C}^1$ functions on an interval I, with $\vec{\varphi}$ and ϵ_1-approximate, and $\vec{\psi}$ an ϵ_2-approximate solution. Suppose further that we have a specified point $\tau_0 \in I$ and that $||\vec{\varphi}(\tau_0) - \vec{\psi}(t_0)|| \leq \delta$. From problem **8)** of HW-4 we have the following *fundamental estimate*:

$$||\vec{\varphi}(t) - \vec{\psi}(t)|| \leq \delta e^{L|t-t_0|} + \frac{\epsilon_1 + \epsilon_2}{L}(e^{L|t-t_0|} - 1) \tag{FE}$$

for all $t \in I$.

Lemma 2.4. *Suppose $\vec{v}$ is bounded, $M < \infty$ an upper bound for $\vec{v}$, and $\vec{\xi}_0 = (\tau_0, \vec{a}_0)$ a point in Ω. Let $[c, d]$ be an interval of existence for $(\Delta)_{\vec{\xi}_0}$ such that $[c, d] \times \{\vec{a}_0\} \subset \Omega$. Let $\vec{\varphi}_0 : (c, d) \to \mathbb{R}^n$ be the solution to $(\Delta)_{\vec{\xi}_0}$ on $[c, d]$. Then*

$$||\vec{\varphi}_0(t) - \vec{a}_0|| \leq \frac{M}{L}(e^{L|t-\tau_0|} - 1) \qquad (t \in [c, d])$$

Proof. Let $\vec{\psi} : [c, d] \to \mathbb{R}^n$ be the constant map $\vec{\psi} = \vec{a}_0$. By hypothesis, we see that $(t, \vec{\psi}(t)) \in \Omega$ for all $t \in [c, d]$. So, for $t \in [c, d]$ we have

$$||\dot{\vec{\psi}}(t) - \vec{v}(t, \vec{\psi}(t))|| = ||\vec{v}(t, \psi(t))|| \leq M$$

Thus $\vec{\psi}$ is an M-approximate solution of $(\Delta)_{\vec{\xi}_0}$. On the other hand $\vec{\varphi}_0$ is an exact solution. So the fundamental estimate (FE) with $\delta = 0$, $\epsilon_1 = 0$ and $\epsilon_2 = M$ gives the reuslt. ∎

2.11. **Continuity with respect to initial conditions.** For this section, we fix a solution

$$\vec{\varphi} : [c, d] \to \mathbb{R}^n$$

of the DE $(\Delta)_{\vec{\xi}}$. For each $\delta > 0$, let

$$U_\delta = \{(\tau, \vec{a}) \in \mathbb{R}^{n+1} \mid \tau \in [c, d],\ ||\vec{a} - \vec{\varphi}(\tau)|| < \delta\}$$

Lemma 2.5. *There exists $\delta_1 > 0$ such that the closure $\overline{U_{\delta_1}}$ of U_{δ_1} in $\mathbb{R}^{n+1}$ is a compact subset of Ω.*

Proof. The map $f : [c, d] \times \mathbb{R}^n \to [c, d] \times \mathbb{R}^n$ given by

$$f(t, \vec{a}) = (t, \vec{a} + \vec{\varphi}(t))$$

is a homeomorphism with inverse g given by

$$g(t, \vec{a}) = (t, \vec{a} - \vec{\varphi}(t))$$

This means that $U_\delta = f([c, d] \times B(\vec{0}, \delta))$ is an open subset of $[c, d] \times \mathbb{R}^n$ for every $\delta > 0$.

Next, let $\Omega' = f^{-1}(\Omega \cap [c, d] \times \mathbb{R}^n)$. Then Ω' is open in $[c, d] \times \mathbb{R}^n$ and contains $[c, d] \times \{0\}$. If $d(t, \vec{a})$ is the distance between $(t, \vec{a})$ and the closed subset $[c, d] \times \mathbb{R}^n \setminus \Omega'$ of $[c, d] \times \mathbb{R}^n$, then d is continuous on $[c, d] \times \mathbb{R}^n$. Since $K = [c, d] \times \{0\}$ is compact, the infimum of d on K is attained on K and is a positive number η. Pick $\delta_1 < \eta$. Then $[c, d] \times \overline{B}(\vec{0}, \delta_1) \subseteq \Omega'$. It follows that $\overline{U_{\delta_1}} = f([c, d] \times \overline{B}(\vec{0}, \delta_1))$ is compact and is contained in $\Omega \cap [c, d] \times \mathbb{R}^n$. ∎

Theorem 2.6. *Let the notations be as above. Then, there is some $\delta > 0$ such that the following are true.*

1. *$U_\delta \subseteq \Omega$.*
2. *For every $\vec{\xi} = (\tau, \vec{a}) \in U_\delta$ the solution $\vec{\varphi}_{\vec{\xi}}$ of $(\Delta)_{\vec{\xi}}$ exists on $[c, d]$.*
3. *The map $(t, \tau, \vec{a}) \mapsto \vec{\varphi}_{(\tau, \vec{a})}(t)$ is uniformly continuous on $V = [c, d] \times U_\delta$.*

Proof. Let δ_1 be as given by **Lemma** 2.5, i.e $\overline{U_{\delta_1}}$ is compact and contained in Ω. Let

$$D = \{\delta \mid 0 < \delta < e^{-L(d-c)}\delta_1\}$$

First, we show that (1) and (2) are true for every δ in D. Let

$$U = \bigcup_{\delta \in D} U_\delta$$

It can be checked that $U = U_{\delta_m}$, where $\delta_m = e^{-L(d-c)}\delta_1$.

For $\vec{\boldsymbol{\xi}} \in U$, the fundamental estimate (FE) (with $\epsilon_1 = \epsilon_2 = 0$) gives

$$||\vec{\boldsymbol{\varphi}}(t) - \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}(t)|| < \delta_1$$

for all $t \in [c,d] \cap J(\vec{\boldsymbol{\xi}})$, where as before, $J(\xi)$ is the maximal interval of existence of the DE $(\Delta)_{\vec{\boldsymbol{\xi}}}$. We want to argue that $[c,d] \subseteq J(\vec{\boldsymbol{\xi}})$. Note that we have just shown that $(t, \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}(t)) \in U_{\delta_1}$ for all t in $J(\vec{\boldsymbol{\xi}}) \cap [c,d]$. Since $(t, \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}(t))$ must exit the compact set $[c,d] \times \overline{U_{\delta_1}}$, the above inequality forces it to exit at $\{c\} \times \mathbb{R}^n$ and $\{d\} \times \mathbb{R}^n$. Thus $[c,d] \cap J(\vec{\boldsymbol{\xi}}) = [c,d]$. This proves (1) and (2).

Now, let $\vec{\boldsymbol{F}} : [c,d] \times U \to \mathbb{R}^n$ be the map given by the formula

$$F(t, \tau, \vec{\boldsymbol{a}}) = \vec{\boldsymbol{\varphi}}_{(\tau, \vec{\boldsymbol{a}})}(t)$$

We have to show that $\vec{\boldsymbol{F}}$ is continuous. Since U is a subset of $\overline{U_{\delta_1}}$ which is compact, $\vec{\boldsymbol{v}}$ is bounded on U. Let M be the supremum of $||\vec{\boldsymbol{v}}||$ on the compact set $\overline{U_{\delta_1}}$. Since $\overline{U_{\delta_1}}$ is compact, $M < \infty$.

Let $\vec{\boldsymbol{\xi}}_0 = (\tau_0, \vec{\boldsymbol{a}}_0) \in U$, and let us examine the continuity of $\vec{\boldsymbol{F}}$ at $(s, \vec{\boldsymbol{\xi}}_0) \in [c,d] \times U$. Since U is open in $[c,d] \times \mathbb{R}^n$, there exists a rectangle $W = [\alpha, \beta] \times B(\vec{\boldsymbol{a}}_0, r)$ in U containing $\vec{\boldsymbol{\xi}}_0$, and hence for every $\vec{\boldsymbol{\xi}} = (\tau, \vec{\boldsymbol{a}}) \in W$, the line segment $[\alpha, \beta] \times \{\vec{\boldsymbol{a}}\}$ lies in $U \subseteq \Omega$. By **Lemma** 2.4, we see that

$$(\dagger) \qquad ||\vec{\boldsymbol{a}} - \vec{\boldsymbol{\varphi}}_{(\tau, \vec{\boldsymbol{a}})}(\tau_0)|| \leq \frac{M}{L}(e^{L|\tau - \tau_0|} - 1) \qquad (\vec{\boldsymbol{\xi}} = (\tau, \vec{\boldsymbol{a}}) \in W)$$

We claim that $\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}} \to \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}_0}$ uniformly on $[c,d]$ as $\vec{\boldsymbol{\xi}} \to \vec{\boldsymbol{\xi}}_0$. We may assume that $\vec{\boldsymbol{\xi}}$ approaches $\vec{\boldsymbol{\xi}}_0$ through points in W. By the fundamental estimate (FE) we get that

$$\begin{aligned}
||\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}_0}(t) - \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}(t)|| &\leq ||\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}_0}(\tau_0) - \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}(\tau_0)||e^{L(d-c)} \\
&= ||\vec{\boldsymbol{a}}_0 - \vec{\boldsymbol{\varphi}}_{(\tau, \vec{\boldsymbol{a}})(\tau_0)}||e^{L(d-c)} \\
&\leq ||\vec{\boldsymbol{a}}_0 - \vec{\boldsymbol{a}}||e^{L(d-c)} + ||\vec{\boldsymbol{a}} - \vec{\boldsymbol{\varphi}}_{(\tau, \vec{\boldsymbol{a}})}(\tau_0)||e^{L(d-c)} \\
&\leq ||\vec{\boldsymbol{a}}_0 - \vec{\boldsymbol{a}}||e^{L(d-c)} + \frac{M}{L}(e^L|\tau - \tau_0| - 1)e^{L(d-c)}
\end{aligned}$$

where in the last step we have used (†). Now, the expression that we have obtained, namely

$$\vec{\boldsymbol{h}}(\vec{\boldsymbol{\xi}}) = h(\tau, \vec{\boldsymbol{a}}) = ||\vec{\boldsymbol{a}}_0 - \vec{\boldsymbol{a}}||e^{L(d-c)} + \frac{M}{L}(e^{L|\tau - \tau_0|} - 1)e^{L(d-c)}$$

is a continuous function of $\vec{\boldsymbol{\xi}}$, which is independent of $t \in [c,d]$. Moreover, $h(\vec{\boldsymbol{\xi}}) \to 0$ as $\vec{\boldsymbol{\xi}} \to \vec{\boldsymbol{\xi}}_0$. So, it follows that $\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}} \to \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}_0}$ uniformly on $[c,d]$ as $\vec{\boldsymbol{\xi}} \to \vec{\boldsymbol{\xi}}_0$.

We now show that $\vec{\boldsymbol{F}} : [c,d] \times U \to \mathbb{R}^n$ is continuous. Let $\vec{\boldsymbol{\varphi}}_0 = \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}_0}$. Since $\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}$ converges uniformly on $[c,d]$ to $\vec{\boldsymbol{\varphi}}_0$ as $\vec{\boldsymbol{\xi}} \to \vec{\boldsymbol{\xi}}_0$, therefore given $\epsilon > 0$ we can find $\eta_1 > 0$ such that

$$||\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}(s) - \vec{\boldsymbol{\varphi}}_0(s)|| < \epsilon \qquad (s \in [c,d])$$

whenever $||\vec{\boldsymbol{\xi}} - \vec{\boldsymbol{\xi}}_0|| < \eta_1$. Now $\vec{\boldsymbol{\varphi}}_0$ is uniformly continuous on the compact set $[c, d]$, and hence there exists $\eta_2 > 0$ such that

$$||\vec{\boldsymbol{\varphi}}_0(t) - \vec{\boldsymbol{\varphi}}_0(s)|| < \epsilon$$

whenever $|s - t| < \eta_2$. Since

$$\begin{aligned} ||\vec{\boldsymbol{F}}(t, \vec{\boldsymbol{\xi}}) - F(\vec{\boldsymbol{s}}, \vec{\boldsymbol{\xi}}_0)|| &= ||\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}(t) - \vec{\boldsymbol{\varphi}}_0(s)|| \\ &\leq ||\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}}(t) - \vec{\boldsymbol{\varphi}}_0(t)|| + ||\vec{\boldsymbol{\varphi}}_0(t) - \vec{\boldsymbol{\varphi}}_0(s)|| \end{aligned}$$

it follows that

$$||\vec{\boldsymbol{F}}(t, \vec{\boldsymbol{\xi}}) - \vec{\boldsymbol{F}}(s, \vec{\boldsymbol{\xi}}_0)|| < 2\epsilon$$

whenever $||\vec{\boldsymbol{\xi}} - \vec{\boldsymbol{\xi}}_0|| < \eta_1$ and $|t - s| < \eta_2$. Thus $\vec{\boldsymbol{F}}$ is continuous on $[c, d] \times U$.

It remains to show that $\vec{\boldsymbol{F}}$ is uniformly continuous on U_δ for any $\delta \in D$. Now, $\overline{U_\delta} \subset U = U_{\delta_m}$. To see this, let

$$f : [c, d] \times \mathbb{R}^n \to [c, d] \times \mathbb{R}^n$$

be $f(t, \vec{\boldsymbol{a}}) = (t, \vec{\boldsymbol{a}} + \vec{\boldsymbol{\varphi}}(t))$ Recall that f is a homeomorphism, and

$$U_\delta = f([c, d] \times B(\vec{\boldsymbol{0}}, \delta))$$

and

$$\overline{U_\delta} = f([c, d] \times \overline{B}(\vec{\boldsymbol{0}}, \boldsymbol{\delta}))$$

Since $\delta < \delta_m = e^{-L(d-c)\delta_1}$, therefore

$$[c, d] \times \overline{B}(\vec{\boldsymbol{0}}, \boldsymbol{\delta}) \subseteq [c, d] \times B(\vec{\boldsymbol{0}}, \delta_m)$$

Hence $\overline{U_\delta} \subseteq U_{\delta_m} = U$. Thus F is defined on $\overline{U_\delta}$, and $\overline{U_\delta}$ is compact. It follows that $\vec{\boldsymbol{F}}$ is uniformly continuous on U_δ. ■

2.12. **Topological Straightening.** Let us fix $\vec{\boldsymbol{\xi}}_0 = (\tau_0, \vec{\boldsymbol{a}}_0) \in \Omega$ and we fix an interval of existence $I = [\tau_0 - c, \tau_0 + c]$ for the solution $\vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}_0}$ of $(\Delta)_{\vec{\boldsymbol{\xi}}_0}$.

Let U_δ be the set of points $\vec{\boldsymbol{\xi}} = (\tau, \vec{\boldsymbol{a}} + \vec{\boldsymbol{\varphi}}_{\vec{\boldsymbol{\xi}}_0}(t)) \in I \times \mathbb{R}^n$ such that $||\vec{\boldsymbol{a}}|| < \delta$. We showed in **Lemma** 2.5 that there exists $\delta_1 > 0$ such that $\overline{U_{\delta_1}} \subset \Omega$ and $\overline{U_{\delta 1}}$ is compact. Let δ_m be defined by the formula

$$\delta_m = e^{-L(2c)}\delta_1$$

We have seen that if $\vec{\boldsymbol{\xi}} \in U_{\delta_m}$ then $[\tau_0 - c, \tau_0 + c] \subseteq J(\vec{\boldsymbol{\xi}})$. Note that $U_{\delta_m} \subseteq U_{\delta_1}$ since $\delta_m < \delta_1$.

If $\vec{\boldsymbol{a}} \in \mathbb{R}^n$ is such that $||\vec{\boldsymbol{a}} - \vec{\boldsymbol{a}}_0|| < \delta_m$, then $(\tau_0, \vec{\boldsymbol{a}}) \in U_{\delta_m} \subset U_{\delta_1} \subset \Omega$. Moreover, by **Theorem** 2.6 part (2), I is an interval of existence for $\vec{\boldsymbol{\varphi}}_{(\tau_0, \vec{\boldsymbol{a}})}$. By the fundamental estimate (FE) it is then easy to see that $(t, \vec{\boldsymbol{\varphi}}_{(\tau_0, \vec{\boldsymbol{a}})}(t)) \in U_{\delta_1}$ for all $t \in I$.

So, we have a map $\boldsymbol{\Phi} : I \times B(\vec{\boldsymbol{a}}_0, \delta_m) \to U_{\delta_1}$ given by

$$\boldsymbol{\Phi}(t, \vec{\boldsymbol{a}}) = (t, \vec{\boldsymbol{\varphi}}_{(\tau_0, \vec{\boldsymbol{a}})}(t))$$

Note that

$$\boldsymbol{\Phi}(t, \vec{\boldsymbol{a}}) = (t, \vec{\boldsymbol{F}}(t, \tau_0, \vec{\boldsymbol{a}}))$$

So the image of $\boldsymbol{\Phi}$ is the graph of $\vec{\boldsymbol{F}}|_S$ where $S = I \times \{\tau_0\} \times B(\vec{\boldsymbol{a}}_{\boldsymbol{0}}, \delta_m)$. Also note that $\boldsymbol{\Phi}$ is continuous. Below, we assume that $\delta \in (0, \delta_m)$.

(1) By the uniqueness of solutions of $(\Delta)_{\vec{\boldsymbol{\xi}}_0}$, we see that $\boldsymbol{\Phi}$ is a one-one map.

(2) Let $R_\delta^\circ = (\tau_0 - c, \tau_0 + c) \times B(\vec{a}_0, \delta)$, $R_\delta = I \times B(\vec{a}_0, \delta)$ and $\overline{R_\delta} = I \times \overline{B}(\vec{a}_0, \delta)$. Set

$$V_\delta^\circ = \mathbf{\Phi}(R_\delta^\circ), \quad V_\delta = \mathbf{\Phi}(R_\delta), \quad \text{and } \overline{V}_\delta = \mathbf{\Phi}(\overline{R_\delta}) \tag{2.4}$$

Since $\mathbf{\Phi}$ is a one-one continuous map, we can apply a result on the invariance of domain for $\mathbb{R}^{n+1}$ to conclude that V_δ° is an open subset of $\mathbb{R}^{n+1}$ and that $\mathbf{\Phi}|_{R_\delta^\circ}$ is a homeomorphism between R_δ° and V_δ°. Also, since $\mathbf{\Phi}$ is injective and $\overline{R_\delta}$ is compact, the restriction $\mathbf{\Phi}|_{\overline{R_\delta}}$ is a homeomorphism. Hence, $\mathbf{\Phi}|_{R_\delta}$ is also a homeomorphism.

Some nice pictures are given in section 2 of the Lecture 21 notes.

2.13. **Differentiability with respect to initial phase.** We now assume that $\vec{v}$ is $\mathscr{C}^1$ in $\vec{x}$. So, the all partial derivatives $\frac{\partial \vec{v}}{\partial x_i}$ exist on Ω and are continuous in $(t, \vec{x})$. Clearly, being $\mathscr{C}^1$ implies that $\vec{v}$ is locally Lipschitz in $\vec{x}$.

2.13.1. *The equation of variations.* Let

$$D_2(t, \vec{x}) = \begin{bmatrix} \frac{\partial v_1}{\partial x_1} & \cdots & \frac{\partial v_1}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial v_n}{\partial x_1} & \cdots & \frac{\partial v_n}{\partial x_n} \end{bmatrix}$$

Define the map

$$A : R_{\delta_m} \to M_n(\mathbb{R})$$

by

$$A(t, \vec{x}) = D_2(\mathbf{\Phi}(t, \vec{x}))$$

Clearly, for each fixed t and $\vec{x}$, $A(t, \vec{x})$ is a linear map. We now consider the *equation of variations*

$$\dot{\vec{z}} = A(t, \vec{x})\vec{z} \quad , \quad \vec{z}(\tau_0) = \vec{e}_j \tag{2.5}$$

where $\vec{e}_1, ..., \vec{e}_n$ is the standard basis of $\mathbb{R}^n$.

Let $\vec{\zeta}$ be the unique solution of the linear IVP (2.5). We know that I is an interval of existence for $\vec{\zeta}$ (since this is a linear IVP). Note that $\vec{\zeta}$ depends upon $\vec{x}$. Therefore, we think of $\vec{\zeta}$ as a function of $\vec{x}$ and write $\vec{\zeta}(t, \vec{x})$.

2.13.2. *Differentiability with respect to x.* Since $\overline{U_{\delta_1}}$ is compact and in Ω, and since $\vec{v}$ is $\mathscr{C}^1$, therefore D_2 is bounded on $\overline{U_{\delta_1}}$. Let $0 < M < \infty$ be such that

$$||D_2(t, \vec{a})|| \le M \qquad (t, \vec{a}) \in U_{\delta_1}$$

Recall that $\vec{F} : I \times \delta_m \to \mathbb{R}^n$ is the map $(t, \vec{\xi}) \mapsto \vec{\varphi}_{\vec{\xi}}(t)$. Note that if $\vec{x}$ and $\vec{x} + h\vec{e}_j$ both lie in $B(\vec{a}_0, \delta_m)$ then for $t \in I$, by the fundamental estimate (FE) we have

$$||\vec{F}(t, \tau_0, \vec{x} + h\vec{e}_j) - \vec{F}(t, \tau_0, \vec{x})|| \le |h| e^{2Lc} \tag{2.6}$$

This follows from the fundamental estimate and the fact that $\vec{\varphi}_{\tau_0, \vec{a}}$ is an ϵ-approximate solution with $\epsilon = 0$.

Theorem 2.7. *Suppose, as above, $\vec{v}$ is $\mathscr{C}^1$ in $\vec{x}$. Let $\tau_0 \in I$ be a fixed initial time point. Then $\vec{F}(t, \tau_0, \vec{x})$ is $\mathscr{C}^1$ as a function of $(t, \vec{x})$ on $R_{\delta_m}^\circ$.*

Proof. See **Theorem 3.2.3** in Lecture 22. ■