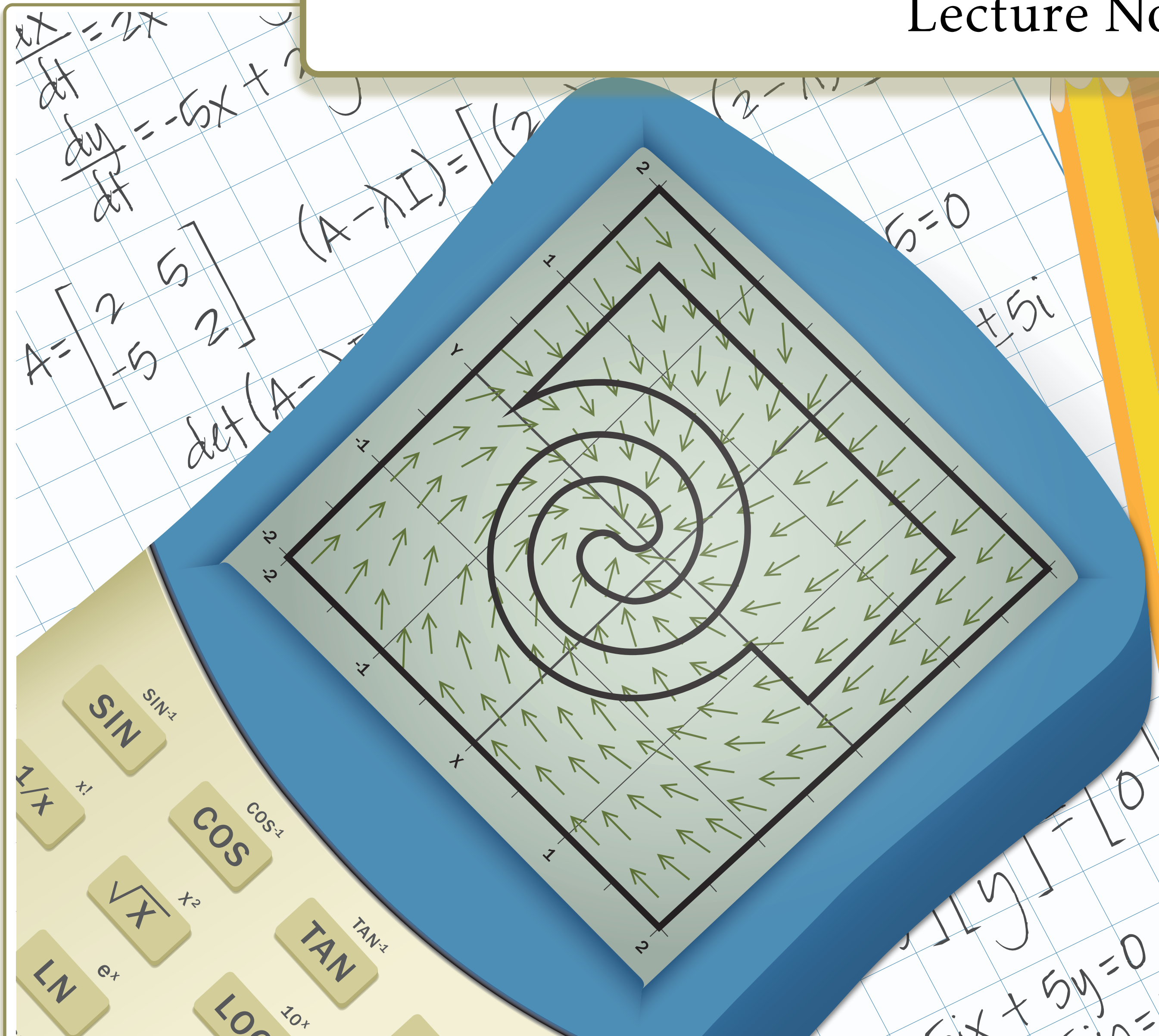
UDGRP Winter 2024
Lecture Notes



DYNAMICAL SYSTEMS

ARITRABHA MAJUMDAR (BMAT2311)

Undergraduate Directed Group Reading Project Winter 2024

Stability Analysis of Ordinary Differential Equations

Aritrabha Majumdar (BMAT2311)

Indian Statistical Institute, Banglaore

Contents

1 Linear Systems

1.1 Autonomous System

An autonomous System of Ordinary Differential Equations is a system which does NOT explicitly depend on the independent variables. It is of the form

$$\frac{d}{dt}x(t) = f(x(t)) \ ; \quad x \in \mathbb{R}^n$$

Solutions are invariant under horizontal translations.

Proof: Say, $x_1(t)$ is a solution of the ODE $\frac{dx}{dt} = f(x)$, $x(0) = x_0$.
Then $x_2(t) = x_1(t - t_0)$ solves $\frac{dx}{dt} = f(x)$, $x(t_0) = x_0$.
Now we set $s = t - t_0$ which essentially gives $x_2(t) = x_1(s)$ and $ds = dt$. Thus,

$$\frac{d}{dt}x_2(t) = \frac{d}{ds}x_1(s) = f(x_1(s)) = f(x_2(t))$$

And for the initial condition, we have $x_2(t_0) = x_1(t_0 - t_0) = x_0$

An autonomous system of two first order differential equations has the form

$$\frac{dx}{dt} = f(x, y)$$

$$\frac{dy}{dt} = g(x, y)$$

If the system is linear, we can express it in the given format

$$\frac{dx}{dt} = ax + by$$

$$\frac{dy}{dt} = cx + dy$$

For which we can write

$$\dot{\mathbf{x}} = \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = A\mathbf{x} \ ; \quad (a, b, c, d) \in \mathbb{R}^4$$

1.2 Uncoupled System

An uncoupled system of Ordinary Differential Equations is a system in which differential equation of one of the dependent variables is independent of the others. Clearly in this case, the matrix A is (NOT always) be diagonal.

$$\frac{dx}{dt} = ax \implies x = c_1 e^{at}$$

$$\frac{dy}{dt} = by \implies y = c_2 e^{bt}$$

$$\dot{\mathbf{x}} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mathbf{x} \implies \mathbf{x} = \begin{bmatrix} e^{at} & 0 \\ 0 & e^{bt} \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \mathbf{C} e^{At}$$

After a bit careful examination, it is evident that the solutions of this differential equation lies on $\mathbb{R}^2$ and they have the form $y = kx^{\frac{b}{a}}$, where $k = \frac{c_2^{\frac{1}{b}}}{c_1^{\frac{1}{a}}}$.

Phase Plane: While trying to describe the motion of the particle governed by the provided differential equations, we can draw the solution curves in the plane $\mathbb{R}^n$, and this is known as the *Phase Plane*. Clearly, in the above uncoupled system, $\mathbb{R}^2$ is the *Phase Plane*.

Phase Portrait: The set of all solution curves drawn in the Phase space is known as *Phase Portrait*.

Dynamical Systems: A dynamical system governed by $\dot{\mathbf{x}} = A\mathbf{x}$ is a function $\phi : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}^n$ and it is given by $\phi(\mathbf{C}, t) = \mathbf{C}e^{At}$. Geometrically, it describes motion of points in phase plane along the solution curves.

Equilibrium Point: For $c_1 = c_2 = 0$, $\mathbf{x}(t) = 0 \ \forall t \in \mathbb{R}$ and the origin is referred to as an *equilibrium point* of a system of Differential Equations.

The function $f(\mathbf{x}) = A\mathbf{x}$ defines a mapping $f : \mathbb{R}^n \mapsto \mathbb{R}^n$; which defines a vector field on $\mathbb{R}^n$. If we draw each vector along with its initial points, then we get a pictorial representation of the vector field. It is an interesting observation that at each point in the *phase space*, the solution curve is tangent to the vectors in the vector field. Actually, it is pretty obvious, as at time t, the velocity vector $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$ is tangent to the solution curve.

We observe this for $\dot{\mathbf{x}} = I\mathbf{x}$

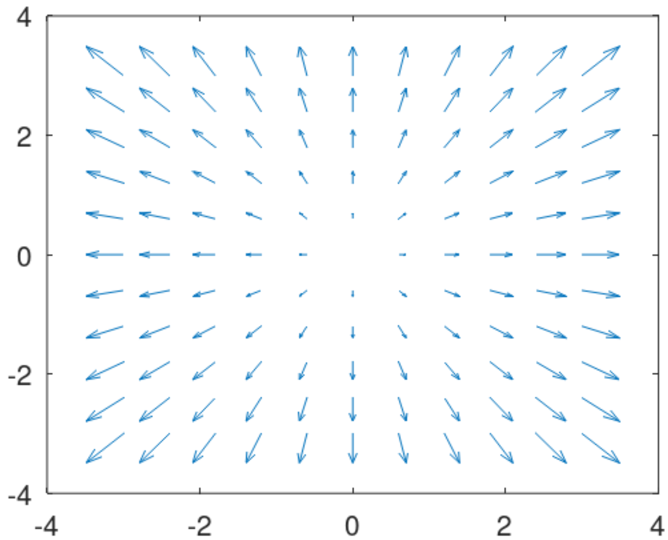


Figure 1: Vector field representation for $\dot{\mathbf{x}} = I\mathbf{x}$

Asymptotic Stability of Origin: Here, we look at

$$\lim_{t\to\infty}(x(t), y(t)) = \lim_{t\to\infty}(c_1e^{at}, c_2e^{bt})$$

If $a < 0$ and $b < 0$, then this limit goes to $(0, 0)$. Otherwise, most of the solutions diverge to infinity,
Roughly speaking, an equilibrium (x_0, y_0) is asymptotically stable if every trajectory $(x(t), y(t))$ beginning from an initial condition near (x_0, y_0) stays near (x_0, y_0) for $t > 0$, and

$$\lim_{t\to\infty}(x(t), y(t)) = (x_0, y_0)$$

The equilibrium is unstable if there are trajectories with initial conditions arbitrarily close to the equilibrium that move far away from that equilibrium.
Later on, we will discuss about this in greater detail.

Invariance of the Axes: There is another observation that we can make for uncoupled systems. Suppose that the initial condition for an uncoupled system lies on the x axis; that is, suppose $y_0 = 0$, then the solution $(x(t), y(t)) = (x_0e^{at}, 0)$ also lies on the x axis $\forall$ time. Similarly, if the initial condition lies on the y axis, then the solution $(0, y_0e^{bt})$ lies on the y axis $\forall$ time.

1.3 Diagonalization

Theorem: If *eigenvalues* $\lambda_1, \lambda_2, ..., \lambda_n$ of a matrix A are *real* and *distinct*, then any set of corresponding *eigenvectors* $\{v_1, v_2, ...v_n\}$ forms a basis of $\mathbb{R}^n$. The matrix $P = \begin{bmatrix} v_1 & v_2 & ... & v_n \end{bmatrix}$ is invertible and

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

This *theorem* can be used to reduce the linear system $\dot{\mathbf{x}} = A\mathbf{x}$ to an uncoupled linear system. To do so, we first define the change of coordinates $\mathbf{x} = P\mathbf{y}$. So we have,

$$\dot{\mathbf{y}} = P^{-1}\dot{\mathbf{x}} = P^{-1}A\mathbf{x} = P^{-1}AP\mathbf{y}$$

$$\implies \dot{\mathbf{y}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \mathbf{y}$$

$$\implies \mathbf{y}(t) = \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \mathbf{y}(0)$$

$$\implies P^{-1}\mathbf{x}(\mathbf{t}) = \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} P^{-1}\mathbf{x}(0)$$

$$\implies \mathbf{x}(t) = P \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} P^{-1}\mathbf{x}(0)$$

Stable, Unstable and Center Subspace

It is evident that the solution is stable $\forall\, t \in \mathbb{R}$ iff all *eigenvalues* are negative. Keeping this in mind, we consider $\{v_1, \dots, v_k\}$ to be the *eigenvectors* corresponding to *negative eigenvalues*, and $\{v_{k+1}, \dots, v_n\}$ to be the *eigenvectors* corresponding to *positive eigenvalues*.
Then we denote the *stable subspace of the Linear System* by

$$E^S = span\{v_1, \dots, v_k\}$$

and the *unstable subspace of the Linear System* by

$$E^U = span\{v_{k+1}, \dots, v_n\}$$

If we have *pure imaginary eigenvalues*, then we also get a *center subspace*, namely E^C.

1.4 Matrix Norm

Here, while performing all the calculations, we consider L^2 norm.
We define the *norm* of a matrix A to be

$$||A|| = \max_{x \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{||Ax||}{||x||} = \max_{||x||=1} ||Ax||$$

Some Properties:

- $||A|| \geq 0$; $||A|| = 0 \iff A = 0$.
- $||\lambda A|| \leq |\lambda| \cdot ||A||$, $\lambda \in \mathbb{R}, A \in \mathbb{R}^n$.
- $||A + B|| \leq ||A|| + ||B||$.
- $||Ax|| \leq ||A|| \cdot ||x||$.
- $||AB|| \leq ||A|| \cdot ||B||$.
- $||A^k|| \leq ||A||^k$, $k \in \mathbb{N} \cup \{0\}$
- $\|T^{-1}\| \geq \frac{1}{\|T\|}$

Now, $||Ax|| \geq 0 \ \forall \ ||x|| \leq 1$; hence $||A|| \geq 0$.
Now $A = 0 \implies ||Ax|| = 0 \ \forall x \in \mathbb{R}^n \implies ||A|| = 0$.
Say, the $(i, j)^{\text{th}}$ entry of A is non zero. Hence $||Ax|| = \sqrt{a_{ij}^2 x_j^2} > 0$. ($x_j \neq 0$) Similarly by induction, we can show that if k elements of A are non-zero, then $||Ax|| > 0$. Hence if

$||A|| = 0$, then $A = 0$.
$\therefore ||A|| \geq 0$ and $||A|| = 0 \iff A = 0$. ∎

On the other hand,

$$||\lambda A|| = \max_{||x|| \leq 1} ||\lambda Ax|| = \max_{||x|| \leq 1} |\lambda| \cdot ||Ax|| = |\lambda| \max_{||x|| \leq 1} ||Ax|| = |\lambda| \cdot ||A|| \;\; ; \;\; \lambda \in \mathbb{R} \;\; \blacksquare$$

Again,

$$||A+B|| = \max_{||x|| \leq 1} ||(A+B)x|| \leq \max_{||x|| \leq 1} (||Ax||+||Bx||) \leq \left(\max_{||x|| \leq 1} ||Ax|| + \max_{||x|| \leq 1} ||Bx|| \right) = ||A||+||B|| \;\; \blacksquare$$

Again,

$$||A|| = \max_{x \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{||Ax||}{||x||} \implies \frac{||Ax||}{||x||} \leq ||A|| \implies ||Ax|| \leq ||A|| \cdot ||x|| \;\; \blacksquare$$

Moreover,

$$||AB|| = \max_{||x|| \leq 1} ||ABx|| \leq ||A|| \max_{||x|| \leq 1} ||Bx|| = ||A|| \cdot ||B|| \;\; \blacksquare$$

We also observe

$$||A^k|| \leq ||A|| \cdot ||A^{k-1}|| \leq \cdots \leq ||A||^k \;\; \blacksquare$$

And lastly,

$$1 = \|TT^{-1}\| \leq \|T\| \cdot \|T^{-1}\| \implies \|T^{-1}\| \geq \frac{1}{\|T\|} \;\; \blacksquare$$

Limit of a Linear Operator: A sequence of linear operators $\{T_k\}_{k \geq 1} \subseteq \mathcal{L}(\mathbb{R}^n)$ is said to converge to a limiting linear operator $T \in \mathcal{L}(\mathbb{R}^n)$ as $k \to \infty$ if for every $\varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall k \geq N$, $||T_k - T|| < \varepsilon$.

Show that for each $t \in \mathbb{R}$, the solution of $\dot{\mathbf{x}} = A\mathbf{x}$ is a continuous function of the initial condition.

Proof: Say, the solution obtained is given by $\phi(t, \mathbf{x_0}) = \mathbf{x_0} e^{At}$. For a fixed t, we take the matrix norm to be defined analogously as L^2 norm. We also define $\delta := \frac{\varepsilon}{||e^{At}||}$. Now for $||\mathbf{y_0} - \mathbf{x_0}|| < \delta$, we have $||\phi(t, \mathbf{y_0}) - \phi(t, \mathbf{x_0})|| \leq ||e^{At}|| \cdot ||\mathbf{y_0} - \mathbf{x_0}|| < \varepsilon$ ∎

1.5 Exponentials of Operators

Say we have a given $T \in \mathcal{L}(\mathbb{R}^n)$ and a given $t_0 \in \mathbb{R}$. Say $||T|| = a$

$$\left\| \frac{T^k t^k}{k!} \right\| \leq \frac{||T^k|| \cdot |t|^k}{k!} \leq \frac{||T||^k \cdot t_0^k}{k!} = \frac{a^k \cdot t_0^k}{k!}$$

$\forall\ |t| < t_0$. Now

$$\sum_{k=0}^{\infty} \frac{(at_0)^k}{k!} = e^{at_0}$$

So, by *Weierstrass M test*, the sum $\sum_{k=0}^{\infty} \frac{T^k t^k}{k!}$ converges *uniformly* and *absolutely*. So, now we define the matrix exponential as

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} \quad ; \quad t \in \mathbb{R}$$

Note that $\|e^{At}\| \leq e^{\|T\| \cdot |t|}$

Theorem 1: If $P, T \in \mathcal{L}(\mathbb{R}^n)$ and $S = PTP^{-1}$, then $e^S = Pe^T P^{-1}$.

Proof: According to the definition,

$$e^S = \sum_{k=1}^{\infty} \frac{S^k}{k!} = \sum_{k=1}^{\infty} \frac{(PTP^{-1})^k}{k!} = P \sum_{k=1}^{\infty} \frac{T^k}{k!} P^{-1} = Pe^T P^{-1} \quad \blacksquare$$

Theorem 2: If $S, T \in \mathcal{L}(\mathbb{R}^n)$ and they *commute*, then $e^{S+T} = e^S e^T$.

Proof: If S and T *commute*, then

$$(S+T)^n = \sum_{k=0}^{n} \binom{n}{k} S^k T^{n-k}$$

Therefore

$$e^{S+T} = \sum_{n=0}^{\infty} \frac{(S+T)^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k+j=n} \frac{n!}{k!j!} S^k T^j = \left(\sum_{k=0}^{\infty} \frac{S^k}{k!} \right) \cdot \left(\sum_{j=0}^{\infty} \frac{T^j}{j!} \right) = e^S e^T \quad \blacksquare$$

Theorem 3: if $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, then $e^A = e^a \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}$

Proof: we take $z = a + ib = re^{i\theta}$; under standard notations. Now we can write

$$A^2 = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix} \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix} = \begin{bmatrix} r^2\cos 2\theta & -r^2\sin 2\theta \\ r^2\sin 2\theta & r^2\cos 2\theta \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(z^2) & -\mathrm{Im}(z^2) \\ \mathrm{Im}(z^2) & \mathrm{Re}(z^2) \end{bmatrix}$$

Thus by induction,$A^k = \begin{bmatrix} \mathrm{Re}(z^k) & -\mathrm{Im}(z^k) \\ \mathrm{Im}(z^k) & \mathrm{Re}(z^k) \end{bmatrix}$. Now we have

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \begin{bmatrix} \mathrm{Re}(\frac{z^k}{k!}) & -\mathrm{Im}(\frac{z^k}{k!}) \\ \mathrm{Im}(\frac{z^k}{k!}) & \mathrm{Re}(\frac{z^k}{k!}) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(e^z) & -\mathrm{Im}(e^z) \\ \mathrm{Im}(e^z) & \mathrm{Re}(e^z) \end{bmatrix}$$

Now $e^z = e^{a+ib} = e^a(\cos b + i \sin b)$, so we have $\mathrm{Re}(e^z) = e^a \cos b$ and $\mathrm{Im}(e^z) = e^a \sin b$.

$$\therefore e^A = e^a \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix} \quad \blacksquare$$

Note: *If $a = 0$, this matrix represents anticlockwise rotation by b degrees.*

Theorem 4: If $A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$, then $e^A = e^a \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$

Proof: $A = aI + \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} = aI + B$. Clearly aI and B commute. Moreover $B^k = 0 \; \forall k \geq 2 \implies e^B = I + B$. So we can hereby conclude

$$e^A = e^{aI+B} = e^a e^B = e^a \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \quad \blacksquare$$

Theorem 5: If $A = PDP^{-1}$, where D is *diagonal*, then $\det(e^A) = e^{\mathrm{trace}(D)}$.

Proof: $e^A = Pe^DP^{-1} \implies \det(e^A) = \det(Pe^DP^{-1}) = \det(e^D)$ As D is *diagonal* we can write,

$$e^A = e^{\mathrm{trace}(D)} = e^{\mathrm{trace}(P^{-1}AP)} = e^{\mathrm{trace}(P^{-1}PA)} = e^{\mathrm{trace}(A)} \quad \blacksquare$$

Theorem 6: If $\mathbf{x}$ is an *eigenvector* of T with *eigenvalue* λ, then $\mathbf{x}$ is also an *eigenvector* of e^T with *eigenvalue* e^λ

Proof: $T^2\mathbf{x} = T(\lambda\mathbf{x}) = \lambda(T\mathbf{x}) = \lambda^2\mathbf{x}$. Thus by induction, $T^k\mathbf{x} = \lambda^k\mathbf{x}$. Now we have

$$e^T\mathbf{x} = \left(\sum_{k=0}^{\infty} \frac{T^k}{k!} \right) \mathbf{x} = \left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right) \mathbf{x} = e^\lambda \mathbf{x} \quad \blacksquare$$

Theorem 7: $T \in \mathcal{L}(\mathbb{R}^n)$ and $E \subset \mathbb{R}^n$ is T invariant; then show E is also e^T invariant.

Proof: Clearly if $\mathbf{v} \in E$, then it is a linear combination of the basis vectors, so is $\frac{\mathbf{v}}{k!}$, $k \in \mathbb{N} \cup \{0\}$.
On the other hand, it is trivial that $E \supseteq T(E) \supseteq T^2(E) \supseteq \cdots \supseteq T^k(E) \supseteq \dots$
Say, we have

$$\mathbf{v} = \bigcap_{k \in \mathbb{N} \cup \{0\}} T^k(E) \qquad \text{(We define } T^0(E) = E\text{)}$$

Now,

$$e^T(\mathbf{v}) = \left(\sum_{k=0}^{\infty} \frac{T^k}{k!} \right)(\mathbf{v}) = \sum_{k=0}^{\infty} T^k \left(\frac{\mathbf{v}}{k!} \right)$$

We know

$$T^k \left(\frac{\mathbf{v}}{k!} \right) \in E \ \forall k \in \mathbb{N} \cup \{0\}$$

These altogether concludes

$$e^T(\mathbf{v}) \in E \implies e^T(E) \subseteq E \ \blacksquare$$

1.6 The Fundamental Theorem for Linear Systems

Here, our aim is to establish the fact that for $\mathbf{x_0} \in \mathbb{R}^n$, the initial value problem

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

$$\mathbf{x}(\mathbf{0}) = \mathbf{x_0}$$

has a unique solution $\forall t \in \mathbb{R}$ which is given by

$$\mathbf{x}(t) = \mathbf{x_0} e^{At}$$

Lemma: Let A be a square matrix. Then

$$\frac{d}{dt} e^{At} = A e^{At}$$

Proof:

$$\frac{d}{dt} e^{At} = \lim_{h \to 0} \frac{e^{A(t+h)} - e^{At}}{h} = e^{At} \lim_{h \to 0} \frac{e^{Ah} - I}{h} = e^{At} \lim_{h \to 0} \left(A + \sum_{k=1}^{\infty} \frac{A^{k+1} h^k}{(k+1)!} \right) = A e^{At}$$

Note: *Here, we can place the limit inside the summation because* $|h| \leq 1$

If $\mathbf{x}(t)$ has the mentioned form, then we can easily observe

$$\mathbf{x}'(t) = \frac{d}{dt}\mathbf{x_0}e^{At} = \mathbf{x_0}Ae^{At} = A\mathbf{x}(t)$$

Now to show that this is the only solution, we consider $\mathbf{x}(t)$ to be any solution of the provided initial value problem. Now we fix $\mathbf{y}(t) = e^{-At}\mathbf{x}(t)$. Now we differentiate both side to obtain

$$\mathbf{y}'(t) = -Ae^{-At}\mathbf{x_0} + e^{-At}\mathbf{x}'(t) = -Ae^{-At}\mathbf{x_0} + e^{-At}A\mathbf{x}(t) = 0$$

Setting $t = 0$, we obtain $\mathbf{y}(0) = \mathbf{x_0}$, and this suffices the proof of uniqueness. ∎

1.7 Linear Systems in $\mathbb{R}^2$

In this section, we describe various *phase portraits* of the equation

$$\dot{\mathbf{x}} = A\mathbf{x} \quad , \quad \mathbf{x} \in \mathbb{R}^2$$

Say, $\mathbf{v}$ is an *eigenvector* of A with *eigenvalue* λ. Now, $\mathbf{x} = A\mathbf{v}$, where a is a *scalar*. Hence

$$\dot{\mathbf{x}} = A(a\mathbf{v}) = a\lambda v$$

The derivative is a multiple of $\mathbf{v}$ and hence points along the line determined by $\mathbf{v}$. As $\lambda > 0$, the derivative points in the direction of $\mathbf{v}$ when a is positive and in the opposite direction when a is negative.

We consider $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ and we draw the *vector field* and a couple of solutions *(go to next page)*. Notice that the picture looks like a ***source*** with arrows coming out from the origin. Hence we call this type of picture a source or sometimes an ***unstable node***.

If $A = \begin{bmatrix} -1 & -1 \\ 0 & -2 \end{bmatrix}$, then both the *eigenvalues* are *negative*. We call this kind of picture a ***sink*** or sometimes a ***stable node***.

If $A = \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix}$, then one *eigenvalue* is positive, and the other is negative. Then, we reverse the arrows on one line (corresponding to the negative eigenvalue) in **Figure 2**. This is known as a ***Saddle***

Suppose the eigenvalues are purely imaginary. That is, suppose the eigenvalues are $\pm ib$. For example, let $A = \begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix}$. Consider the eigenvalue $2i$ and its eigenvector $\begin{bmatrix} 1 \\ 2i \end{bmatrix}$. The real and imaginary parts of $\vec{v}e^{i2t}$ are

$$\text{Re}\begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{i2t} = \begin{bmatrix} \cos(2t) \\ -2\sin(2t) \end{bmatrix}, \quad \text{Im}\begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{i2t} = \begin{bmatrix} \sin(2t) \\ 2\cos(2t) \end{bmatrix}$$

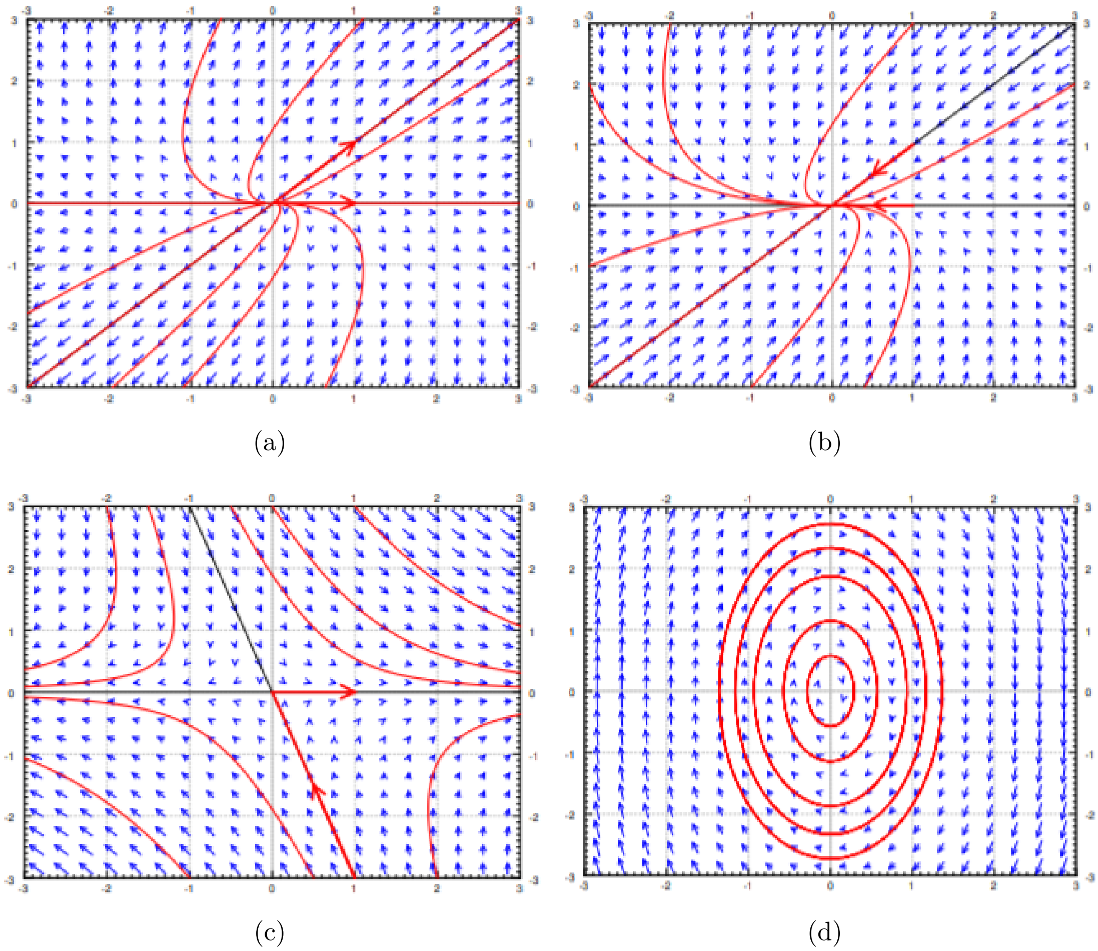
We can take any linear combination of them to get other solutions, which one we take depends on the initial conditions. Now note that the real part is a *parametric equation* for an *ellipse*. Same with the imaginary part and in fact any linear combination of the two. This is what happens in general when the *eigenvalues* are *purely imaginary*. So when the eigenvalues are purely imaginary, we get ellipses for the solutions. This type of picture is sometimes called a ***center***.

Now suppose the complex eigenvalues have a positive real part. For example, let $A = \begin{bmatrix} 1 & 1 \\ -4 & 1 \end{bmatrix}$. We take $1 + 2i$ and its eigenvector $\begin{bmatrix} 1 \\ 2i \end{bmatrix}$, and find the real and imaginary of $\vec{v}e^{(1+2i)t}$ are

$$\operatorname{Re}\begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{(1+2i)t} = e^t \begin{bmatrix} \cos(2t) \\ -2\sin(2t) \end{bmatrix} \quad \operatorname{Im}\begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{(1+2i)t} = e^t \begin{bmatrix} \sin(2t) \\ 2\cos(2t) \end{bmatrix}$$

Note the e^t in front of the solutions. This means that the solutions grow in magnitude while spinning around the origin. Hence we get a ***spiral source.***

Finally suppose the complex eigenvalues have a negative real part. Here we get a e^{-t} in front of the solution. This means that the solutions shrink in magnitude while spinning around the origin. Hence we get a ***spiral sink***.



(a) (b)

(c) (d)

Figure 2: (a) Source (b) Sink (c) Saddle (d) Center

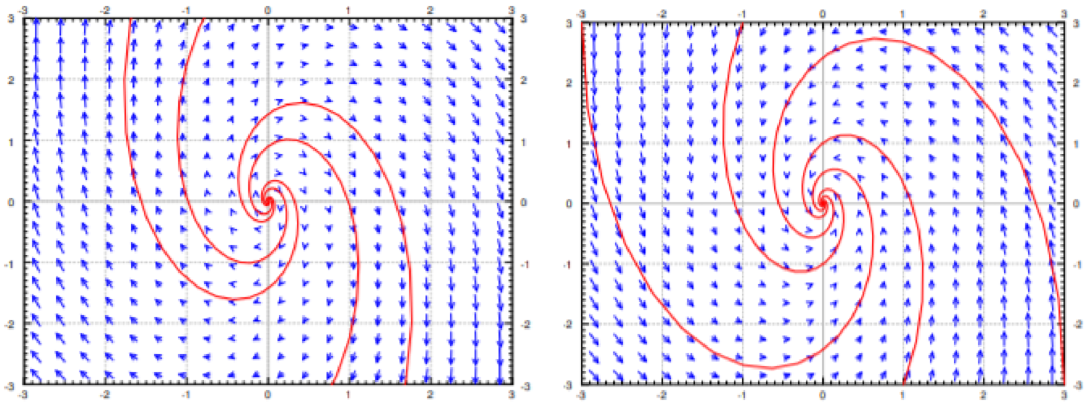
1.8 System with Complex *eigenvalues*

If $A \in GL_{2n}(\mathbb{R})$ and has complex *eigenvalues*, they occur as *conjugate pairs.* The following Theorem gives us an insight about this.

Theorem: If $A \in GL_{2n}(\mathbb{R})$ has $2n$ distinct *complex eigenvalues*, $\lambda_j = a_j + ib_j$ and $\overline{\lambda_j} = a_j - ib_j$, $\forall j = 1(1)n$ with corresponding *eigenvectors* $\mathbf{w}_j = \mathbf{u}_j + i\mathbf{v}_j$ and $\overline{\mathbf{w}_j} = \mathbf{u}_j - i\mathbf{v}_j$; then $\{\mathbf{u_1}, \mathbf{v_1}, \ldots, \mathbf{u_n}, \mathbf{v_n}\}$ forms a *basis* for $\mathbb{R}^{2n}$. Moreover the matrix $P = \begin{bmatrix}\mathbf{v_1} & \mathbf{u_1} & \ldots & \mathbf{v_n} & \mathbf{u_n}\end{bmatrix}$ is *invertible* and

$$P^{-1}AP = \operatorname{diag}\begin{bmatrix} a_j & -b_j \\ b_j & a_j \end{bmatrix}$$

is a $2n \times 2n$ matrix with 2×2 blocks across the diagonal.



(a) (b)

Figure 3: (a) Spiral Source (b)Spiral Sink

Proof: We say, if V is a real vector space, its *complexification* $V^{\mathbb{C}}$ is the complex vector space consisting of elements $x + iy$ where $x, y \in V$. If $T : V \to W$, its *complexification* $T^{\mathbb{C}} : V^{\mathbb{C}} \to W^{\mathbb{C}}$ is defined by

$$T^{\mathbb{C}}(x + iy) = Tx + iTy$$

Clearly $T^{\mathbb{C}}$ has same *eigenvalues* as of T. So we have $\mathbf{w} = \mathbf{u} + i\mathbf{v}$ and $\overline{\mathbf{w}} = \mathbf{u} - i\mathbf{v}$ in $V^{\mathbb{C}}$ with *eigenvalues* λ and $\overline{\lambda}$. Clearly

$$\mathbf{u} = \frac{\mathbf{w} + \overline{\mathbf{w}}}{2} \quad , \quad \mathbf{v} = \frac{\mathbf{w} - \overline{\mathbf{w}}}{2i}$$

Clearly, $\mathbf{u}$ and $\mathbf{v}$ are *linearly independent*, and they form a *basis* for V. Now we want to compute the matrix of T with respect to this new *basis*. So we compute

$$T^{\mathbb{C}}(\mathbf{w}) = \lambda\mathbf{w} = (a + ib)(\mathbf{u} + i\mathbf{v}) = (a\mathbf{u} - b\mathbf{v}) + i(a\mathbf{v} + b\mathbf{u})$$

Moreover we also have

$$T^{\mathbb{C}}(\mathbf{w}) = T\mathbf{u} + iT\mathbf{v}$$

So, on comparison, we have

$$T\mathbf{v} = a\mathbf{v} + b\mathbf{u} = \begin{bmatrix}\mathbf{v} & \mathbf{u}\end{bmatrix}\begin{bmatrix}a \\ b\end{bmatrix} \quad , \quad T\mathbf{u} = a\mathbf{u} - b\mathbf{v} = \begin{bmatrix}\mathbf{v} & \mathbf{u}\end{bmatrix}\begin{bmatrix}-b \\ a\end{bmatrix}$$

So, clearly in the *basis* $\{\mathbf{v}, \mathbf{u}\}$, the matrix of T is $\begin{bmatrix}a & -b \\ b & a\end{bmatrix}$ and hereby we can conclude the matrix $P = \begin{bmatrix}\mathbf{v_1} & \mathbf{u_1} & \dots & \mathbf{v_n} & \mathbf{u_n}\end{bmatrix}$ is *invertible* and

$$P^{-1}AP = \operatorname{diag}\begin{bmatrix}a_j & -b_j \\ b_j & a_j\end{bmatrix}$$

is a 2×2 matrix with blocks across the diagonal. ■

If we use $P = \begin{bmatrix}\mathbf{u_1} & \mathbf{v_1} & \dots & \mathbf{u_n} & \mathbf{v_n}\end{bmatrix}$, then we have

$$P^{-1}AP = \text{diag}\begin{bmatrix} a_j & b_j \\ -b_j & a_j \end{bmatrix}$$

So, now we have the solution of the initial value problem

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

as

$$\mathbf{x}(t) = P \text{ diag } e^{a_j t}\begin{bmatrix} \cos b_j t & -\sin b_j t \\ \sin b_j t & \cos b_j t \end{bmatrix} P^{-1}\mathbf{x_0}$$

1.9 Multiple *eigenvalues*

Till now, we have only dealt with those systems which have *distinct eigenvalues*. Now, we want to solve the system where A has *multiple eigenvalues*.

Definition:

Let λ be a *eigenvalue* of a $n \times n$ matrix A with multiplicity $m \leq n$. Then for $k = 1(1)m$, any non-zero solution $\mathbf{v}$ of

$$(A - \lambda I)^k \mathbf{v} = 0$$

is known as a *generalised eigenvector* of A.

Theorem: If $T \in \mathcal{L}(V)$ with *real eigenvalues*, then there is *only one way* of writing T as $S + N$, where S is *diagonalizable*, N is *nilpotent*, and $SN = NS$.

Proof: Let E_k be the *generalised eigenspace* of T, $\forall k = 1(1)m$. We define $T_k = T|_{E_k}$. Now we have

$$V = \bigoplus_{k=1}^{m} E_k \quad , \quad T = \bigoplus_{k=1}^{m} T_k$$

Note that S and N both commute with S and N, hence both of them commute with $T = S + N$ as well. So we have E_k is *invariant* under S and N. Now we say $S_k = \lambda_k I \in \mathcal{L}(E_k)$ and $N_k = T_k - S_k$. If we can show $S|_{E_k} = S_k$, it will then conclude $N|_{E_k} = N_k$, and thus we can show the *uniqueness*.
Enough to show $S|_{E_k} - S_k = \mathbf{0}$.
Now, it is given that S is *diagonalizable*, so is $S|_{E_k}$, then $S|_{E_k} - \lambda_k I$ is also *diagonalizable*.

Hence $S|_{E_k} - S_k$ is diagonalizable.
Now, on the other hand, $S|_{E_k} - S_k = T|_{E_k} - N|_{E_k} - T_k + N_k = N_k - N|_{E_k}$. Here, $N|_{E_k}$ commutes with T_k and $\lambda_k I$; so it also commutes with N_k. Using *Binomial Theorem*, we can hereby conclude that $N_k - N|_{E_k}$ is *nilpotent*.
So, $S|_{E_k} - S_k$ is a *nilpotent* and *diagonal* matrix, i.e $S|_{E_k} - S_k = \mathbf{0}$ ■

So, now we have the solution of the initial value problem

$$\dot{\mathbf{x}} = A\mathbf{x} \ \ , \ \ \mathbf{x}(\mathbf{0}) = \mathbf{x_0}$$

as

$$\mathbf{x}(t) = P \text{ diag}\left[e^{\lambda_j t}\right] P^{-1}\left[I + Nt + \cdots + \frac{N^k t^k}{k!}\right]\mathbf{x_0}$$

If λ is an*eigenvalue* with multiplicity n, then the solution of the *initial value problem* is

$$\mathbf{x}(t) = e^{\lambda t}\left[I + Nt + \cdots + \frac{N^k t^k}{k!}\right]\mathbf{x_0}$$

Under the light of this theorem, we can right the theorem discussed in the previous section in a newly tailored way.

Theorem: If $A \in GL_{2n}(\mathbb{R})$ has $2n$ *complex eigenvalues*, $\lambda_j = a_j + ib_j$ and $\overline{\lambda_j} = a_j - ib_j$, $\forall j = 1(1)n$ with corresponding *eigenvectors* $\mathbf{w}_j = \mathbf{u}_j + i\mathbf{v}_j$ and $\overline{\mathbf{w}_j} = \mathbf{u}_j - i\mathbf{v}_j$; then $\{\mathbf{u_1}, \mathbf{v_1}, \ldots, \mathbf{u_n}, \mathbf{v_n}\}$ forms a *basis* for $\mathbb{R}^{2n}$. Moreover the matrix $P = \begin{bmatrix}\mathbf{v_1} & \mathbf{u_1} & \ldots & \mathbf{v_n} & \mathbf{u_n}\end{bmatrix}$ is *invertible*, $A = S + N$ and

$$P^{-1}SP = \text{diag}\begin{bmatrix} a_j & -b_j \\ b_j & a_j \end{bmatrix}$$

is a $2n \times 2n$ matrix with 2×2 blocks across the diagonal, the matrix N is *nilpotent* of order $k \leq 2n$.

So, now we have the solution of the initial value problem

$$\dot{\mathbf{x}} = A\mathbf{x} \ \ , \ \ \mathbf{x}(\mathbf{0}) = \mathbf{x_0}$$

as

$$\mathbf{x}(t) = P \text{ diag } e^{a_j t}\begin{bmatrix} \cos b_j t & -\sin b_j t \\ \sin b_j t & \cos b_j t \end{bmatrix} P^{-1}\left[I + Nt + \cdots + \frac{N^k t^k}{k!}\right]\mathbf{x_0}$$

1.10 Stability Theory

Say, for am matrix A has *generalised eigenvalues* $\lambda_j = a_j + ib_j$ and *generalised eigenvector* $\mathbf{v}_j = \mathbf{u}_j + i\mathbf{w}_j$. Then the *stable subspace*, *unstable subspace* and *central subspace* is given by

$$E^S = \text{Span}\{\mathbf{u}_j, \mathbf{v}_j \mid a_j < 0\}$$

$$E^U = \text{Span}\{\mathbf{u}_j, \mathbf{v}_j \mid a_j > 0\}$$
$$E^C = \text{Span}\{\mathbf{u}_j, \mathbf{v}_j \mid a_j = 0\}$$

Solutions in E^S tend to approach $\mathbf{x(0)}$ as $t \to \infty$; and solutions in E^U tend to approach $\mathbf{x(0)}$ as $t \to -\infty$.
The set of mappings $e^{At} : \mathbb{R}^n \to \mathbb{R}^n$ may be regarded as the movement of points $\mathbf{x_0} \in \mathbb{R}^n$ along the trajectories.

Hyperbolic flow: If all *eigenvalues* of A has non-zero real parts, then the *flow* $e^{At} : \mathbb{R}^n \to \mathbb{R}^n$ is called *hyperbolic flow*, and the corresponding linear system is known as *hyperbolic linear system*.

A subspace $E \subset \mathbb{R}^n$ is said to be *invariant with respect to the flow* if $e^{At}E \subset E$, $\forall t \in \mathbb{R}$.

Lemma: Let E be a *generalised eigenspace* of matrix A with respect to its *generalised eigenvalue* λ. Show that $AE \subset E$.

Proof: Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be basis of *generalised eigenvectors* for E. Then for $\mathbf{v} \in E$

$$\mathbf{v} = \sum_{k=1}^{n} c_k \mathbf{v}_k \implies A\mathbf{v} = \sum_{k=1}^{n} c_k A\mathbf{v}_k$$

Now each of the $\mathbf{v}_k$s being *generalised eigenvectors*, we say

$$\mathbf{V}_k = (A - \lambda I) \quad \mathbf{V}_k \in \text{Ker}(A - \lambda I)^{j-1} \subset E$$

Thus, by induction, $A\mathbf{v}_k = \lambda \mathbf{v}_k + \mathbf{V}_k \in E$, so does their linear combination. Hence $AE \subset E$ ■

Clearly, according to the definition $\mathbb{R}^n = E^S \oplus E^U \oplus E^C$.
For $\mathbf{x_0} \in E^S$,

$$\mathbf{x_0} = \sum_{k=1}^{n_s} c_k \mathbf{V}_k \quad \{\mathbf{V}_k\}_{k=1}^{n_s} \subset B \text{ is a } \textit{basis} \text{ for the stable subspace } E^S$$

Now,

$$e^{At}\mathbf{x_0} = \sum_{k=1}^{n_s} c_k e^{At} \mathbf{V}_k$$

As $A^k \mathbf{V}_j \in E^S$, then $e^{At}\mathbf{x_0} \in E^S$, $\forall t \in \mathbb{R}$.

So, E^S is *invariant* with respect to the *flow*, so is E^U and E^C. ■

Sink (or Source): If all *eigenvalues* has negative (or positive) real part, then the *origin* is known as *sink (or source)* of the linear system.

Theorem: The following statements are equivalent

(a) $\forall\ \mathbf{x_0} \in \mathbb{R}^n \lim_{t\to\infty} e^{At}\mathbf{x_0} = 0$ and for $\mathbf{x_0} \neq \mathbf{0}$, $\lim_{t\to-\infty} e^{At}\mathbf{x_0} = \infty$

(b) All *eigenvalues* of A has negative real part.

(c) There are positive constants $a, c, m, M \in \mathbb{R}$ such that $\forall \mathbf{x_0} \in \mathbb{R}^n$

$$|e^{At}\mathbf{x_0}| \leq Me^{-ct}\mathbf{x_0} \quad t \geq 0$$

$$|e^{At}\mathbf{x_0}| \geq me^{-at}\mathbf{x_0} \quad t \leq 0$$

Proof: Here we use the fact that any solution of the linear system is the linear combination of functions of the form $t^k e^{at} \cos bt$ or $t^k e^{at} \sin bt$.
Say, one of the *eigenvalues* has positive real part. For that particular *eigenvalue* $\forall \mathbf{x_0} \neq 0$ $\lim_{t\to\infty} e^{At}\mathbf{x_0} = \infty$ and for $\mathbf{x_0} \in \mathbb{R}^n$, $\lim_{t\to-\infty} e^{At}\mathbf{x_0} = 0$, contradicting (a). If one of the *eigenvalues* has a zero real part, then the solution is of the form $t^k \cos bt$ or $t^k \sin bt$, and again clearly $\forall \mathbf{x_0} \in \mathbb{R}^n \lim_{t\to\infty} e^{At}\mathbf{x_0} \neq 0$ So, we can say (a) $\implies$ (b). ∎
sin and cos being periodic function, for *eigenvalues* with a negative *real* part, We can give them a bound as described in (c). So (b) $\implies$ (c). ∎
Using squeeze theorem on the relation obtained at (c) and by taking $t \to \infty$, we get $\forall \mathbf{x_0} \in \mathbb{R}^n \lim_{t\to\infty} e^{At}\mathbf{x_0} = 0$, and the second inequality in part (c) gives us $\mathbf{x_0} \neq \mathbf{0}$, $\lim_{t\to-\infty} e^{At}\mathbf{x_0} = \infty$. Hence (c) $\implies$ (a). ∎

In the similar fashion, we can devise another theorem, with similar proof.

Theorem: The following statements are equivalent

(a) $\forall\ \mathbf{x_0} \in \mathbb{R}^n \lim_{t\to-\infty} e^{At}\mathbf{x_0} = 0$ and for $\mathbf{x_0} \neq \mathbf{0}$, $\lim_{t\to\infty} e^{At}\mathbf{x_0} = \infty$

(b) All *eigenvalues* of A has positive real part.

(c) There are positive constants $a, c, m, M \in \mathbb{R}$ such that $\forall \mathbf{x_0} \in \mathbb{R}^n$

$$|e^{At}\mathbf{x_0}| \leq Me^{ct}\mathbf{x_0} \quad t \leq 0$$

$$|e^{At}\mathbf{x_0}| \geq me^{at}\mathbf{x_0} \quad t \geq 0$$

1.11 Nonhomogeneous Linear Systems

In this section, we are concerned about differential equations of the type

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}(t)$$

Where A is a $n \times n$ matrix and $\mathbf{b}(t)$ is a vector valued function.

Fundamental Matrix Solution: A fundamental matrix solution of $\dot{\mathbf{x}} = A\mathbf{x}$ is any *nonsingular* $n \times n$ matrix function $\mathbf{\Phi}(t)$ that satisfies $\mathbf{\Phi'}(t) = A\mathbf{\Phi}(t)$, $\forall t \in \mathbb{R}$.

Once we find a *Fundamental Matrix Solution* for the *homogeneous* system , we can find the solution to the corresponding *nonhomogeneous* system.

Theorem: If $\mathbf{\Phi}(t)$ is a *fundamental matrix solution*, then the solution of the *nonhomogeneous* system and the initial condition $\mathbf{x}(\mathbf{0}) = x_0$ is *unique*, and is given by

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{\Phi}^{-1}(\mathbf{0})\mathbf{x_0} + \int_0^t \mathbf{\Phi}(t)\mathbf{\Phi}^{-1}(\tau)\mathbf{b}(\tau)\, d\tau$$

Proof: We differentiate $\mathbf{x}(t)$ as defined above.

$$\dot{\mathbf{x}} = \mathbf{\Phi'}(t)\mathbf{\Phi}^{-1}(\mathbf{0})\mathbf{x_0} + \mathbf{\Phi}(t)\mathbf{\Phi}^{-1}(t)\mathbf{b}(t) + \int_0^t \mathbf{\Phi'}(t)\mathbf{\Phi}^{-1}(\tau)\mathbf{b}(\tau)\, d\tau$$

$$\implies \dot{\mathbf{x}} = A\left(\mathbf{\Phi}(t)\mathbf{\Phi}^{-1}(\mathbf{0})\mathbf{x_0} + \int_0^t \mathbf{\Phi}(t)\mathbf{\Phi}^{-1}(\tau)\mathbf{b}(\tau)\, d\tau\right) + \mathbf{b}(t)$$

$$\therefore \dot{\mathbf{x}} = A\mathbf{x}(t) + \mathbf{b}(t) \; \blacksquare$$

With $\mathbf{\Phi}(t) = e^{At}$, the solution of the *nonhomogeneous* linear system looks like

$$\mathbf{x}(t) = e^{At}\mathbf{x_0} + e^{At}\int_0^t e^{-A\tau}\mathbf{b}(\tau)\, d\tau$$

2 Nonlinear Systems: Local Theory

2.1 Some Preliminary Concepts and Definitions

Differentiability: The function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *differentiable* at $\mathbf{x_0} \in \mathbb{R}^n$ if there exists a linear transformation $D\mathbf{f}(\mathbf{x_0}) \in \mathcal{L}(\mathbb{R}^n)$ that satisfies

$$\lim_{||h|| \to 0} \frac{||\mathbf{f}(\mathbf{x_0} + \mathbf{h}) - \mathbf{f}(\mathbf{x_0}) - D\mathbf{f}(\mathbf{x_0})\mathbf{h}||}{||\mathbf{h}||} = 0$$

The linear transformation $D\mathbf{f}(\mathbf{x_0})$ is the derrivative of $\mathbf{f}$ at $\mathbf{x_0}$. Now we look into a theorem that enables us to compute derrivative in coordinates.

Theorem: Consider a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ differentiable at $a \in \mathbb{R}^n$. Then all the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist at a. In particular, for f differentiable at a, we have,

$$(Df)(a) = J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{m \times n}$$

Proof: Without loss of generality, we take $m = 1$, and let $a = (a_1, a_2, \ldots, a_n)$. Fix an arbitrary index $i \in \{1, 2, \ldots, n\}$. We define $\eta_i : [a_i - \epsilon, a_i + \epsilon] \to \mathbb{R}^n$, defined by

$$\eta_i(t) = (a_1, \ldots, a_{i-1}, t, a_{t+1}, \ldots, a_n) = a + (t - a_i)e_i$$

As $\mathcal{O}_n$ is open and η_i is continuous, we can find ϵ small such that $f([[a_i - \epsilon, a_i + \epsilon]]) \subseteq \mathbb{R}^n$. Evidently, η_i is differentiable and $(D\eta_i) = [0, \ldots, 1, \ldots, 0]^t = e_i^t$ over $[a_i - \epsilon, a_i + \epsilon]$. Now, by the definition of partial derivatives, $D(f \circ \eta_i)(a_i) = f_{x_i}(a)$.

Again, by chain rule, as f is differentiable at a, $D(f \circ \eta_i)(a_i) = f_{x_i}(a)$ exists, and

$$\begin{aligned} D(f \circ \eta_i)(a_i) &= Df(\eta_i(a_i)) \cdot D\eta_i(a_i) \\ \implies f_{x_i}(a) &= Df(a) \cdot e_i^t = [Df(a)]_i \end{aligned}$$

As the index i was arbitrary to begin with, this completes the proof. ■

Continuity: Suppose V_1 and V_2 be two *normed* linear spaces with respective norms $||.||_1$ and $||.||_2$. Then $\mathbf{f} : V_1 \to V_2$ is continuous at $\mathbf{x_0} \in V_1$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $\mathbf{x} \in V_1$ and $||\mathbf{x} - \mathbf{x_0}||_1 \delta$ implies $||\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x_0})||_2 < \varepsilon$. $\mathbf{f}$ is said to be continuous on $E \subseteq V_1$ if it is continuous $\forall$ points in E, and we write $\mathbf{f} \in \mathcal{C}(E)$.

$\mathcal{C}^1(E)$ **Functions:** If the function $\mathbf{f} : E \to \mathbb{R}^n$ is differentiable on E, then we say $\mathbf{f} \in \mathcal{C}^1(E)$.
The following theorem, almost analogous to the previous one, helps us to decide whether a function belongs to $\mathcal{C}^1(E)$.

Theorem: Suppose E is an open subset of $\mathbb{R}^n$ and $\mathbf{f} : E \to \mathbb{R}^n$. Then $\mathbf{f} \in \mathcal{C}^1(E)$ *iff* $\frac{\partial f_i}{\partial x_j}$ exists $\forall i, j = 1(1)n$, and they are *continuous*.

Remarks: Higher order derrivatives can be defined in a similar fashion. And the similar notion holds for the conditionof $f \in \mathcal{C}^k$.
A function $\mathbf{f} : R \to \mathbb{R}^n$is said to be *analytic*, if each of its components are *analytic*, i.e for $j = 1(1)n$ and $\mathbf{x_0} \in E$, $f_j(x)$ has a *taylor series* which converges to $f_j(x)$ in some neighborhood of $\mathbf{x_0}$ in E.

2.2 The Fundamental Existence Uniqueness Theorem

In this section, our primary focus will revolve around *Piccard's Classical Method of Successive Approximations*. We will establish the *existence*, *uniqueness*, *Continuity* and *Differentiability* of the soution of the intial value problem for given *intial condition* and *parameters* under the hypothesis that $f \in \mathcal{C}^1(E)$.

Definition: Suppose $\mathbf{f} \in \mathcal{C}(E)$, where E is an *open* subset of $\mathbb{R}^n$. Then $\mathbf{x}(t)$ is a solution of the differential equation on an interval I if $\mathbf{x}(t)$ is differentiable on I and if $\forall\ t \in I$, $\mathbf{x}(t) \in E$ and $\mathbf{x}'(t) = \mathbf{f}(\mathbf{x}(t))$.

Locally Lipschitz: The function $\mathbf{f}$ is said to be *locally lipschitz* on E if for each $\mathbf{x_0} \in E$, there is an ε-neighborhood of $\mathbf{x_0}$, $N_\varepsilon(\mathbf{x_0}) \subset E$ and a constant $K > 0$ such that $\mathbf{x}, \mathbf{y} \in N_\varepsilon(\mathbf{x_0})$

$$||\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})|| \leq K||\mathbf{x} - \mathbf{y}||$$

Lemma: If $\mathbf{f} : E \to \mathbb{R}^n$ where E is an open subset of $\mathbb{R}^n$ and $\mathbf{f} \in \mathcal{C}^1(E)$, then $\mathbf{f}$ is *locally lipschitz* on E.

Proof: Since E is open, $\exists \varepsilon > 0$ for given $\mathbf{x_0} \in E$, such that $N_\varepsilon(\mathbf{x_0}) \subset E$. Now we define

$$K = \max_{||\mathbf{x}-\mathbf{x_0}||<\varepsilon/2} ||D\mathbf{f}(\mathbf{x})||$$

We say that the $\varepsilon/2$ neighborhood around $\mathbf{x_0}$ as N_0. Now, for $\mathbf{x}, \mathbf{y} \in N_0$, we set $\mathbf{u} = \mathbf{x} - \mathbf{y}$. So for $0 \le s \le 1$, we have $\mathbf{x} + s\mathbf{u} \in N_0$, since N_0 is *convex*. We define $\mathbf{F} : [0,1] \mapsto \mathbb{R}^n$ by

$$\mathbf{F}(s) = \mathbf{f}(\mathbf{x} + s\mathbf{u}) \implies \mathbf{F}'(s) = D\mathbf{f}(\mathbf{x} + s\mathbf{u})\mathbf{u}$$

Therefore, now we have

$$|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})| = |\mathbf{F}(1) - \mathbf{F}(0)| = \left| \int_0^1 \mathbf{F}'(s)\, ds \right| \le \int_0^1 |D\mathbf{f}(\mathbf{x} + s\mathbf{u})\mathbf{u}|\ ds$$

$$\le \int_0^1 ||D\mathbf{f}(\mathbf{x} + s\mathbf{u})\mathbf{u}||\, |\mathbf{u}|\ ds \le K|\mathbf{u}| = K|\mathbf{y} - \mathbf{x}|$$

Complete Space: Let V be a normed linear space. Then a sequence $\mathbf{u}_k \subset V$ is called a *Cauchy Sequence* if $\forall\ \varepsilon > 0$ there is an N such that $k, m \ge N$ implies that

$$||\mathbf{u}_k - \mathbf{u}_m|| < \varepsilon$$

The space V is caleed *Complete* if every Cauchy Sequence in V converges to some element in V.
The space $\mathcal{C}(I)$ is *complete normed linear space*, as a sequence of functions is *uniformly convergent* if and only if it is a *Cauchy Sequence.*

Theorem: Let E be an *open subset* of $\mathbb{R}^n$ containing $\mathbf{x_0}$ and assume $\mathbf{f} \in \mathcal{C}^1(E)$. Then $\exists a > 0$ such that the *initial value problem*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

$$\mathbf{x}(\mathbf{0}) = \mathbf{x_0}$$

has a *unique solution* in the interval $[-a, a]$.

Proof: Since $\mathbf{f} \in \mathcal{C}^1(E)$, it follows from the lemma proven above, we can say, $\exists \varepsilon > 0$ such that $N_\varepsilon(\mathbf{x_0}) \subset E$ and a constant $K > 0$ such that $\forall \mathbf{x}, \mathbf{y} \in N_\varepsilon(\mathbf{x_0})$,

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \le K|\mathbf{x} - \mathbf{y}|$$

We set $b = \varepsilon/2$. Then the continuous function $\mathbf{f}(x)$ is bounded on the *compact set*

$$N_0 = \{x \in \mathbb{R}^n \colon |\mathbf{x} - \mathbf{x_0}| \leq b\}$$

Let

$$M = \max_{\mathbf{x} \in N_0} |\mathbf{f}(\mathbf{x})|$$

Now, we use Piccard's Successive Approximations. We assume $\exists a \in \mathbb{R}^+$ such that $\mathbf{u_k}(t)$ is defined and *continuous* on $[-a, a]$, as well as it satisfies

$$\max_{[-a,a]} |\mathbf{u_k}(t) - \mathbf{x_0}| \leq b$$

It certainly follow that $\mathbf{f}(\mathbf{u_k}(t))$ is *defined* and *continuous* on $[-a, a]$ and therefore that

$$\mathbf{u_{k+1}}(t) = \mathbf{x_0} + \int_0^t \mathbf{f}(\mathbf{u_k})(s)\, ds$$

is defined and continuous on $[a, -a]$. aand satisfies

$$|\mathbf{u_{k+1}}(t) - \mathbf{x_0}| \leq \int_0^t |\mathbf{f}(\mathbf{u_k}(s))|\, ds \leq Ma, \quad , \quad \forall t \in [-a, a]$$

Thus, by choosing $0 < a < b/M$, it follows from *induction* that $\mathbf{u_k}(t)$ is defined and continiuous.

Now, since $\forall t \in [-a, a]$ and $\forall k \in \mathbb{N} \cup \{0\} := \mathbb{N}_0$ we have $\mathbf{u}_k(t) \in \mathbb{N}_0$, it follows from teh *Lipschitz Condition* satissfied by $\mathbf{f}$ that $\forall t \in [-a, a]$

$$\begin{aligned}
|\mathbf{u}_2(t) - \mathbf{u}_1(t)| &\leq \int_0^t |\mathbf{f}(\mathbf{u}_1(s)) - \mathbf{f}(\mathbf{u_0}(s))|\, ds \\
&\leq K \int_0^t |\mathbf{u_1}(s) - \mathbf{u_0}(s)|\, ds \\
&\leq Ka \max_{[-a,a]} |\mathbf{u_1}(t) - \mathbf{x_0}| \\
&\leq Kab
\end{aligned}$$

And then assuming that

$$\max_{[-a,a]} |\mathbf{u}_j(t) - \mathbf{u}_{j-1}(t)| \leq (Ka)^{j-1} b$$

for some integer $j \geq 2$, it follows that $\forall\ t \in [-a, a]$

$$\begin{aligned}
|\mathbf{u}_{j+1}(t) - \mathbf{u}_j(t)| &\leq \int_0^t |\mathbf{f}\left(\mathbf{u}_j(s)\right) - \mathbf{f}\left(\mathbf{u}_{j-1}(s)\right)|\, ds \\
&\leq K \int_0^t |\mathbf{u}_j(s) - \mathbf{u}_{j-1}(s)|\, ds \\
&\leq Ka \max_{[-a,a]} |\mathbf{u}_j(t) - \mathbf{u}_{j-1}(t)| \\
&\leq (Ka)^j b.
\end{aligned}$$

Thus, it follows by induction that our assumption holds for $j = 2, 3, \dots$ Setting $\alpha = Ka$ and choosing $0 < a < 1/K$, we see that for $m > k \geq N$ and $t \in [-a, a]$

$$
\begin{aligned}
|\mathbf{u}_m(t) - \mathbf{u}_k(t)| &\leq \sum_{j=k}^{m-1} |\mathbf{u}_{j+1}(t) - \mathbf{u}_j(t)| \\
&\leq \sum_{j=N}^{\infty} |\mathbf{u}_{j+1}(t) - \mathbf{u}_j(t)| \\
&\leq \sum_{j=N}^{\infty} \alpha^j b = \frac{\alpha^N}{1-\alpha} b
\end{aligned}
$$

This last quantity approaches zero as $N \to \infty$. Therefore, $\forall\ \varepsilon > 0$ there exists an N such that $m, k \geq N$ implies that

$$\|\mathbf{u}_m - \mathbf{u}_k\| = \max_{[-a,a]} |\mathbf{u}_m(t) - \mathbf{u}_k(t)| < \varepsilon$$

i.e., $\{\mathbf{u}_k\}$ is a Cauchy sequence of continuous functions in $\mathcal{C}([-a, a])$. It follows from the above theorem that $\mathbf{u}_k(t)$ converges to a continuous function $\mathbf{u}(t)$ uniformly $\forall\ t \in [-a, a]$ as $k \to \infty$. And then taking the limit of both sides of equation defining the successive approximations, we see that the continuous function

$$\mathbf{u}(t) = \lim_{k \to \infty} \mathbf{u}_k(t)$$

satisfies the integral equation

$$\mathbf{u}(t) = \mathbf{x}_0 + \int_0^t \mathbf{f}(\mathbf{u}(s)) ds$$

$\forall\ t \in (-a, a]$. We have used the fact that the integral and the limit can be interchanged since the limit in continuous and by the fundamental
theorem of calculus, the right differentiable and

$$\mathbf{u}'(t) = \mathbf{f}(\mathbf{u}(t))$$

$\forall\ t \in [-a, a]$. Furthermore, $\mathbf{u}(0) = \mathbf{x}_0$ and from (4) it follows that $\mathbf{u}(t) \in N_\varepsilon(\mathbf{x}_0) \subset E\ \forall\ t \in [-a, a]$. Thus $\mathbf{u}(t)$ is a solution of the initia value problem on $[-a, a]$. It remains to show that it is the only solution.

Let $\mathbf{u}(t)$ and $\mathbf{v}(t)$ be two solutions of the initial value problem on $[-a, a]$. Then the continuous function $|\mathbf{u}(t) - \mathbf{v}(t)|$ achieves its maximum at some point $t_1 \in [-a, a]$. It follows that

$$\begin{aligned}
\|\mathbf{u}-\mathbf{v}\| &= \max_{[-a,a]} |\mathbf{u}(t)-\mathbf{v}(t)| \\
&= \left| \int_0^{t_1} \mathbf{f}(\mathbf{u}(s)) - \mathbf{f}(\mathbf{v}(s)) ds \right| \\
&\leq \int_0^{|t_1|} |\mathbf{f}(\mathbf{u}(s)) - \mathbf{f}(\mathbf{v}(s))| ds \\
&\leq K \int_0^{|t_1|} |\mathbf{u}(s)-\mathbf{v}(s)| ds \\
&\leq Ka \max |\mathbf{u}(t)-\mathbf{v}(t)| \\
&\leq Ka \|\mathbf{u}-\mathbf{v}\|
\end{aligned}$$

But $Ka < 1$ and this last inequality can only be satisfied if $\|\mathbf{u}-\mathbf{v}\| = 0$. Thus, $\mathbf{u}(t) = \mathbf{v}(t)$ on $[-a, a]$. We have shown that the successive approximations converge uniformly to a unique solution of the initial value problem on the interval $[-a, a]$ where a is any number satisfying $0 < a < \min\left(\frac{b}{M}, \frac{1}{K}\right)$.

Remark: Exactly the same method of proof shows that the initial value problem

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\
\mathbf{x}(t_0) &= \mathbf{x}_0
\end{aligned}$$

has a unique solution on some interval $[t_0 - a, t_0 + a]$.

2.3 Dependence on Initial Conditions and Parameters

In this section we investigate the dependence of the solution of the initial value problem

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\
\mathbf{x}(0) &= \mathbf{y}
\end{aligned}$$

on the initial condition $\mathbf{y}$. If the differential equation depends on a parameter $\boldsymbol{\mu} \in \mathbb{R}^{\boldsymbol{m}}$, i.e., if the function $\mathbf{f}(\mathbf{x})$ in initial value problem is replaced by $\mathbf{f}(\mathbf{x}, \boldsymbol{\mu})$, then the solution $\mathbf{u}(t, \mathbf{y}, \boldsymbol{\mu})$ will also depend on the parameter $\boldsymbol{\mu}$. Roughly speaking, the dependence of the solution $\mathbf{u}(t, \mathbf{y}, \boldsymbol{\mu})$ on the initial condition $\mathbf{y}$ and the parameter μ is as continuous as the function f. In order to establish this type of continuous dependence of the solution on initial conditions and parameters, we first establish a result due to T.H. Gronwall.

Gronwall's Lemma: Suppose that $g(t)$ is a continuous real valued function that satisfies $g(t) \geq 0$ and

$$g(t) \leq C + K \int_0^t g(s)ds$$

$\forall\ t \in [0, a]$ where C and K are positive constants. It then follows that $\forall\ t \in [0, a]$,

$$g(t) \leq Ce^{Kt}$$

Proof: Let $G(t) = C + K \int_0^t g(s)ds$ for $t \in [0, a]$. Then $G(t) \geq g(t)$ and $G(t) > 0\ \forall\ t \in [0, a]$. It follows from the fundamental theorem of calculus that

$$G'(t) = Kg(t)$$

and therefore that

$$\frac{G'(t)}{G(t)} = \frac{Kg(t)}{G(t)} \leq \frac{KG(t)}{G(t)} = K$$

$\forall\ t \in [0, a]$. And this is equivalent to saying that

$$\frac{d}{dt}(\log G(t)) \leq K$$

or

$$\log G(t) \leq Kt + \log G(0)$$

or

$$G(t) \leq G(0)e^{Kt} = Ce^{Kt}$$

$\forall\ t \in [0, a]$, which implies that $g(t) \leq Ce^{Kt}\ \forall\ t \in [0, a]$.

Theorem: Let E be an open subset of $\mathbb{R}^n$ containing $\mathbf{x}_0$ and assume that $\mathbf{f} \in \mathcal{C}^1(E)$. Then there esists an $a > 0$ and $a\delta > 0$ such that $\forall\ \mathbf{y} \in N_\delta(\mathbf{x}_0)$ the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\ \mathbf{x}(0) &= \mathbf{y} \end{aligned}$$

has a unique solution $\mathbf{u}(t, \mathbf{y})$ with $\mathbf{u} \in C^{\mathbf{1}}(G)$ where $G = [-a, a] \times N_\delta(\mathbf{x}_0) \subset \mathbb{R}^{n+1}$; furthermore, for each $\mathbf{y} \in N_\delta(\mathbf{x}_0), \mathbf{u}(t, \mathbf{y})$ is a twice continiously differentiable function of t for $t \in [-a, a]$.

Proof: Since $\mathbf{f} \in \mathcal{C}^1(E)$, it follows from the lemma in Section 2.2 that there is an ε-neighborhood $N_\varepsilon(\mathbf{x}_0) \subset E$ and a constant $K > 0$ such that $\forall\ \mathbf{x}$ and $\mathbf{y} \in N_\varepsilon(\mathbf{x}_0)$,

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq K|\mathbf{x} - \mathbf{y}|$$

As in the proof of the fundamental existence theorem, let $N_0 = \{\mathbf{x} \in \mathbb{R}^n\ \ |\mathbf{x} - \mathbf{x}_0| \leq \varepsilon/2\}$, let M_0 be the maximum of $|\mathbf{f}(\mathbf{x})|$ on N_0 and let M_1 be the maximum of $\|D\mathrm{f}(\mathbf{x})\|$ on N_0. Let $\delta = \varepsilon/4$, and for $\mathbf{y} \in N_\delta(\mathbf{x}_0)$ define the successive approximations $\mathbf{u}_k(t, \mathbf{y})$ as

$$\begin{aligned} \mathbf{u}_0(t, \mathbf{y}) &= \mathbf{y} \\ \mathbf{u}_{k+1}(t, \mathbf{y}) &= \mathbf{y} + \int_0^t \mathbf{f}\left(\mathbf{u}_k(s, \mathbf{y})\right) ds \end{aligned}$$

Assume that $\mathbf{u}_k(t, \mathbf{y})$ is defined and continuous $\forall\ (t, \mathbf{y}) \in G = [-\mathrm{a}, \mathrm{a}\ |\times N_\delta(\mathbf{x}_0)$ and that $\forall\ \mathrm{y} \in N_\delta(\mathbf{x}_0)$

$$\|\mathbf{u}_k(t, \mathbf{y}) - \mathbf{x}_0\| < \varepsilon/2$$

where $\|\cdot\|$ denotes the maximum over all $t \in [-a, a]$. This is clearly satisfied for $k = 0$. And assuming this is true for k, it follows that $\mathbf{u}_{k+1}(t, \mathbf{y})$, defined by the above successive approximations, is continuous on G. This follows since a continuous function of a continuous function is continuous and since the above integral of the continuous function $\mathbf{f}(\mathbf{u}_k(s, \mathbf{y}))$ is continuous in t by the fundamental theorem of calculus and also in $\mathbf{y}$. We also have

$$\|\mathbf{u}_{k+1}(t, \mathbf{y}) - \mathbf{y}\| \leq \int_0^t |\mathbf{f}\left(\mathbf{u}_k(s, \mathbf{y})\right)| ds \leq M_0 a$$

for $t \in [-a, a]$ and $\mathbf{y} \in N_\delta(\mathbf{x}_0) \subset N_0$. Thus, for $t \in [-a, a]$ and $\mathbf{y} \in N_\delta(\mathbf{x}_0)$ with $\delta = \varepsilon/4$, we have

$$\begin{aligned}\|\mathbf{u}_{k+1}(t,\mathbf{y})-\mathbf{x}_0\| &\le \|\mathbf{u}_{k+1}(t,\mathbf{y})-\mathbf{y}\|+\|\mathbf{y}-\mathbf{x}_0\|\\ &\le M_0a+\varepsilon/4<\varepsilon/2\end{aligned}$$

provided $M_0a<\varepsilon/4$, i.e., provided $a<\varepsilon/(4M_0)$. Thus, the above induction hypothesis holds $\forall\ k=1,2,3,\ldots$ and $(t,\mathbf{y})\in G$ provided $a<\varepsilon/(4M_0)$.

We next show that the successive approximations $\mathbf{u}_k(t,\mathbf{y})$ converge uniformly to a continuous function $\mathbf{u}(t,\mathbf{y})\ \forall\ (t,\mathbf{y})\in G$ as $k\to\infty$. As in the proof of the fundamental existence theorem,

$$\begin{aligned}\|\mathbf{u}_2(t,\mathbf{y})-\mathbf{u}_1(t,\mathbf{y})\| &\le Ka\,\|\mathbf{u}_1(t,\mathbf{y})-\mathbf{y}\|\\ &\le Ka\,\|\mathbf{u}_1(t,\mathbf{y})-\mathbf{x}_0\|+Ka\,\|\mathbf{y}-\mathbf{x}_0\|\\ &\le Ka(\varepsilon/2+\varepsilon/4)\le Ka\varepsilon\end{aligned}$$

for $(t,\mathbf{y})\in G$. And then it follows exactly as in the proof of the fundamental existence theorem in Section 2.2 that

$$\|\mathbf{u}_{k+1}(t,\mathbf{y})-\mathbf{u}_k(t,\mathbf{y})\|\le(Ka)^k\varepsilon$$

for $(t,\mathbf{y})\in G$ and consequently that the successive approximations converge uniformly to a continuous function $\mathbf{u}(t,\mathbf{y})$ for $(t,\mathbf{y})\in G$ as $k\to\infty$ provided $a<1/K$. Furthermore, the function $\mathbf{u}(t,\mathbf{y})$ satisfies

$$\mathbf{u}(t,\mathbf{y})=\mathbf{y}+\int_0^t\mathbf{f}(\mathbf{u}(s,\mathbf{y}))ds$$

for $(t,\mathbf{y})\in G$ and also $\mathbf{u}(0,\mathbf{y})=\mathbf{y}$. And it follows from the inequality that $\mathbf{u}(t,\mathbf{y})\in N_{\varepsilon/2}(\mathbf{x}_0)\ \forall\ (t,\mathbf{y})\in G$. Thus, by the fundamental theorem of calculus and the chain rule, it follows that

$$\dot{\mathbf{u}}(t,\mathbf{y})=\mathbf{f}(\mathbf{u}(t,\mathbf{y}))$$

and that

$$\ddot{\mathbf{u}}(t,\mathbf{y})=D\mathbf{f}(\mathbf{u}(t,\mathbf{y}))\dot{\mathbf{u}}(t,\mathbf{y})$$

$\forall\,(t,\mathbf{y}) \in G$; i.e., $\mathbf{u}(t,\mathbf{y})$ is a twice continuously differentiable function of t which satisfies the initial value problem $\forall\,(t,\mathbf{y}) \in G$. The uniqueness of the solution $\mathbf{u}(t,\mathbf{y})$ follows from the fundamental theorem in the previous Section.

We now show that $\mathbf{u}(t,\mathbf{y})$ is a continuously differentiable function of $\mathbf{y}$ $\forall\,(t,\mathbf{y}) \in [-a,a] \times N_{\delta/2}(\mathbf{x}_0)$. In order to do this, fix $\mathbf{y}_0 \in N_{\delta/2}(\mathbf{x}_0)$ and choose $\mathbf{h} \in \mathbb{R}^n$ such that $|\mathbf{h}| < \delta/2$. Then $\mathbf{y}_0 + \mathbf{h} \in N_\delta(\mathbf{x}_0)$. Let $\mathbf{u}(t,\mathbf{y}_0)$ and $\mathbf{u}(t,\mathbf{y}_0+\mathbf{h})$ be the solutions of the initial value problem with $\mathbf{y} = \mathbf{y}_0$ and with $\mathbf{y} = \mathbf{y}_0 + \mathbf{h}$ respectively. It then follows that

$$
\begin{aligned}
|\mathbf{u}(t,\mathbf{y}_0+\mathbf{h}) - \mathbf{u}(t,\mathbf{y}_0)| &\leq |\mathbf{h}| + \int_0^t |\mathbf{f}(\mathbf{u}(s,\mathbf{y}_0+\mathbf{h})) - \mathbf{f}(\mathbf{u}(s,\mathbf{y}_0))|\,ds \\
&\leq |\mathbf{h}| + K\int_0^t |\mathbf{u}(s,\mathbf{y}_0+\mathbf{h}) - \mathbf{u}(s,\mathbf{y}_0)|\,ds
\end{aligned}
$$

$\forall\, t \in [-a,a]$. Thus, it follows from Gronwall's Lemma that

$$
|\mathbf{u}(t,\mathbf{y}_0+\mathbf{h}) - \mathbf{u}(t,\mathbf{y}_0)| \leq |\mathbf{h}|e^{K|t|}
$$

$\forall\, t \in [-a,a]$. We next define $\Phi(t,\mathbf{y}_0)$ to be the fundamental matrix solution of the initial value problem

$$
\begin{aligned}
\dot{\Phi} &= A(t,\mathbf{y}_0)\,\Phi \\
\Phi(0,\mathbf{y}_0) &= I
\end{aligned}
$$

with $A(t,\mathbf{y}_0) = D\mathbf{f}(\mathbf{u}(t,\mathbf{y}_0))$ and I the $n \times n$ identity matrix. The existence and continuity of $\Phi(t,\mathbf{y}_0)$ on some interval $[-a,a]$ follow from the method of successive approximations. It then follows from the initial value problems for $\mathbf{u}(t,\mathbf{y}_0)$, $\mathbf{u}(t,\mathbf{y}_0+\mathbf{h})$ and $\Phi(t,\mathbf{y}_0)$ and Taylor's Theorem,

$$
\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{u}_0) = D\mathbf{f}(\mathbf{u}_0)(\mathbf{u} - \mathbf{u}_0) + \mathbf{R}(\mathbf{u},\mathbf{u}_0)
$$

where $|\mathbf{R}(\mathbf{u},\mathbf{u}_0)| / |\mathbf{u} - \mathbf{u}_0| \to 0$ as $|\mathbf{u} - \mathbf{u}_0| \to 0$, that

$$
\begin{aligned}
|\mathbf{u}(t,\mathbf{y}_0)-\mathbf{u}(t,\mathbf{y}_0+\mathbf{h})+\Phi(t,\mathbf{y}_0)\mathbf{h}| \leq \int_0^t & |\mathbf{f}(\mathbf{u}(s,\mathbf{y}_0)) \\
&- \mathbf{f}(\mathbf{u}(s,\mathbf{y}_0+\mathbf{h})) + D\mathbf{f}(\mathbf{u}(s,\mathbf{y}_0))\Phi(s,\mathbf{y}_0)\mathbf{h}\mid ds \\
&\leq \int_0^t \|D\mathbf{f}(\mathbf{u}(s,\mathbf{y}_0))\| \, |\mathbf{u}(s,\mathbf{y}_0)-\mathbf{u}(s,\mathbf{y}_0+\mathbf{h})+\Phi(s,\mathbf{y}_0)\mathbf{h}| \, ds \\
&+ \int_0^t |\mathbf{R}(\mathbf{u}(s,\mathbf{y}_0+\mathbf{h}),\mathbf{u}(s,\mathbf{y}_0))| \, ds
\end{aligned}
$$

Since $|\mathbf{R}(u,u_0)|/|u-u_0| \to 0$ as $|\mathbf{u}-\mathbf{u}_0| \to 0$ and since $\mathbf{u}(s,\mathbf{y})$ is continuols on G, it follows that given any $\varepsilon_0 > 0$, there exists a $\delta_0 > 0$ such that
$|\mathbf{h}| < \delta_0$ then $|\mathbf{R}(\mathbf{u}(s,\mathbf{y}_0),\mathbf{u}(s,\mathbf{y}_0+\mathbf{h}))| < \varepsilon_0 |\mathbf{u}(s,\mathbf{y}_0)-\mathbf{u}(s,\mathbf{y}_0+\mathbf{h})| \;\forall\; s \in [-a,a]$. Thus, if we let

$$g(t) = |\mathbf{u}(t,\mathbf{y}_0)-\mathbf{u}(t,\mathbf{y}_0+\mathbf{h})+\Phi(t,\mathbf{y}_0)\mathbf{h}|$$

It then follows from the conclusion obtained by *Gronwall's Lemma* and the inequality aforededuced that $\forall\; t \in [-a,a], \mathbf{y}_0 \in N_{\delta/2}(\mathbf{x}_0)$ and $|\mathbf{h}| < \min(\delta_0, \delta/2)$ we have

$$g(t) \leq M_1 \int_0^t g(s)ds + \varepsilon_0 |\mathbf{h}| a e^{Ka}$$

Hence, it follows from Gronwall's Lemma that for any given $\varepsilon_0 > 0$

$$g(t) \leq \varepsilon_0 |\mathbf{h}| a e^{Ka} e^{M_1 a}$$

$\forall\; t \in [-a,a]$ provided $|\mathbf{h}| < \min(\delta_0, \delta/2)$. Thus,

$$\lim_{|\mathbf{h}|\to 0} \frac{|\mathbf{u}(t,\mathbf{y}_0)-\mathbf{u}(t,\mathbf{y}_0+\mathbf{h})+\Phi(t,\mathbf{y}_0)\mathbf{h}|}{|\mathbf{h}|} = 0$$

uniformly $\forall\; t \in [-a,a]$. Therefore,

$$\frac{\partial \mathbf{u}}{\partial \mathbf{y}}(t,\mathbf{y}_0) = \Phi(t,\mathbf{y}_0)$$

$\forall\, t \in [-a,a]$ where $\Phi(t,\mathbf{y}_0)$ is the fundamental matrix solution of the initial value problem (5) which is continuous in t and in $\mathbf{y}_0$ $\forall\; t \in [-a,a]$ and $\mathbf{y}_0 \in N_{\delta/2}(\mathbf{x}_0)$. This completes the proof of the theorem.

Some Remarks

1. A similar proof shows that if $\mathbf{f} \in C^r(E)$ then the solution $\mathbf{u}(t, \mathbf{y})$ of the initial value problem is in $C^r(G)$ where G is defined as in the above theorem. And if $\mathbf{f}(\mathbf{x})$ is a (real) analytic function for $\mathbf{x} \in E$ then $\mathbf{u}(t, \mathbf{y})$ is analytic in the interior of G.

2. If $\mathbf{x}_0$ is an equilibrium point of *initialvalueproblem*, i.e., $f(x_0) = 0$ so that $\mathbf{u}(t, \mathbf{x}_0) = \mathbf{x}_0 \; \forall \; t \in \mathbb{R}$, then

$$\Phi(t, \mathbf{x}_0) = \frac{\partial \mathbf{u}}{\partial \mathbf{x}_0}(t, \mathbf{x}_0)$$

satisfies

$$\begin{aligned} \dot{\Phi} &= D\mathbf{f}(\mathbf{x}_0)\Phi \\ \Phi(0, \mathbf{x}_0) &= I \end{aligned}$$

And according to the Fundamental Theorem for Linear Systems

$$\Phi(t, \mathbf{x}_0) = e^{D\mathbf{f}(\mathbf{x}_0)t}$$

3. It follows from the continuity of the solution $\mathbf{u}(t, \mathbf{y})$ of the initial value problem that for each $t \in [-a, a]$

$$\lim_{\mathbf{y} \to \mathbf{x}_0} \mathbf{u}(t, \mathbf{y}) = \mathbf{u}(t, \mathbf{x}_0)$$

It follows from the inequality that this limit is uniform $\forall \; t \in [-a, a]$.

Now we arrive at two interesting results on the basis of this theorem.

Corollary: Under the hypothesis of the above theorem,

$$\Phi(t, \mathbf{y}) = \frac{\partial \mathbf{u}}{\partial \mathbf{y}}(t, \mathbf{y})$$

for $t \in [-a, a]$ and $\mathbf{y} \in N_\delta(\mathbf{x}_0)$ if and only if $\Phi(t, \mathbf{y})$ is the fundamental matrix solution of

$$\begin{aligned} \dot{\Phi} &= D\mathbf{f}[\mathbf{u}(t, \mathbf{y})]\Phi \\ \Phi(0, \mathbf{y}) &= I \end{aligned}$$

for $t \in [-a, a]$ and $\mathbf{y} \in N_\delta(\mathbf{x}_0)$.

Theorem: Let E be an open subset of $\mathbb{R}^{n+m}$ containing the point $(\mathbf{x}_0, \boldsymbol{\mu}_0)$ where $\mathbf{x}_0 \in \mathbb{R}^n$ and $\boldsymbol{\mu}_0 \in \mathbb{R}^m$ and assume that $\mathbf{f} \in \mathcal{C}^1(E)$. It then follows that there exists an $a > 0$ and a $\delta > 0$ such that $\forall\ \mathbf{y} \in N_\delta\left(\mathbf{x}_0\right)$ and $\boldsymbol{\mu} \in N_\delta\left(\boldsymbol{\mu}_0\right)$, the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}) \\ \mathbf{x}(0) &= \mathbf{y} \end{aligned}$$

has a unique solution $\mathbf{u}(t, \mathbf{y}, \boldsymbol{\mu})$ with $\mathbf{u} \in \mathcal{C}^1(G)$ where $G = [-a, a] \times\ N_\delta\left(\mathbf{x}_0\right) \times N_\delta\left(\boldsymbol{\mu}_0\right)$.

This theorem follows immediately from the previous theorem by replacing the vectors $\mathbf{x}_0, \mathbf{x}, \dot{\mathbf{x}}$ and $\mathbf{y}$ by the vectors $\left(\mathbf{x}_0, \boldsymbol{\mu}_0\right), (\mathbf{x}, \boldsymbol{\mu}), (\dot{\mathbf{x}}, \mathbf{0})$ and $(\boldsymbol{y}, \boldsymbol{\mu})$ or it can be proved directly using *Gronwall's Lemma* and the method of successive approximations.

2.4 The Maximal Interval of Existence

The fundamental existence-uniqueness theorem established that if $\mathbf{f} \in \mathcal{C}^1(E)$ then the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned}$$

has a unique solution defined on some interval $(-a, a)$. In this section we show that initial value problem has a unique solution $\mathbf{x}(t)$ defined on a maximal interval of existence (α, β). Furthermore, if $\beta < \infty$ and if the limit

$$\mathbf{x}_1 = \lim_{t \to \beta^-} \mathbf{x}(t)$$

exists then $\mathbf{x}_1 \in \partial E$, the boundary of E. The boundary of the open set $E, \partial E = \bar{E} \setminus E$ where $\bar{E}$ denotes the closure of E. On the other hand, if the above limit exists and $\mathbf{x}_1 \in E$, then $\beta = \infty, \mathbf{f}\left(\mathbf{x}_1\right) = \mathbf{0}$ and $\mathbf{x}_1$ is an equilibrium point of the *initial value problem*. Now we look into the following lemmas and theorems to understand the underlying concepts in a greater detail.

Lemma: Let E be an open subset of $\mathbb{R}^n$ containing $\mathbf{x}_0$ and suppose $\mathbf{f} \in \mathcal{C}^1(E)$. Let $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ be solutions of the initial value problem on the intervals I_1 and I_2. Then $0 \in I_1 \cap I_2$ and if I is any open interval containing 0 and contained in $I_1 \cap I_2$, it follows that $\mathbf{u}_1(t) = \mathbf{u}_2(t)\ \forall\ t \in I$.

Proof: Since $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ are solutions of the initial value problem on I_1 and I_2 respectively, it follows from Definition 1 in Section 2.2 that $0 \in I_1 \cap I_2$. And if I is an open interval containing 0 and contained in $I_1 \cap I_2$, then the fundamental existence-uniqueness theorem in Section 2.2 implies that $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ on some open interval $(-a, a) \subset I$. Let I^* be the union of all such open intervals contained in I. Then I^* is the largest open interval contained in I on which $\mathbf{u}_1(t) = \mathbf{u}_2(t)$. Clearly $I^* \subset I$ and if I^* is a proper subset of I, then one of the endpoints t_0 of I^* is contained in $I \subset I_1 \cap I_2$. It follows from the continuity of $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ on I that

$$\lim_{t \to t_0} \mathbf{u}_1(t) = \lim_{t \to t_0} \mathbf{u}_2(t)$$

Call this common limit $\mathbf{u}_0$. It then follows from the uniqueness of solutions that $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ on some interval $I_0 = (t_0 - a, t_0 + a) \subset I$. Thus, $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ on the interval $I^* \cup I_0 \subset I$ and I^* is a proper subset of $I^* \cup I_0$. But this contradicts the fact that I^* is the largest open interval contained in I on which $\mathbf{u}_1(t) = \mathbf{u}_2(t)$. Therefore, $I^* = I$ and we have $\mathbf{u}_1(t) = \mathbf{u}_2(t)\ \forall\ t \in I$.

Theorem: Let E be an open subset of $\mathbb{R}^n$ and assume that $\mathbf{f} \in \mathcal{C}^1(E)$. Then for each point $\mathbf{x}_0 \in E$, there is a maximal interval J on which the initial value problem has a unique solution, $\mathbf{x}(t)$; i.e., if the initial value problem has a solution $\mathbf{y}(t)$ on an interval I then $I \subset J$ and $\mathbf{y}(t) = \mathbf{x}(t)\ \forall\ t \in I$. Furthermore, the maximal interval J is open; i.e., $J = (\alpha, \beta)$.

Proof: By the fundamental existence-uniqueness theorem in Section 2.2, the initial value problem has a unique solution on some open interval $(-a, a)$. Let (α, β) be the union of all open intervals I such that initial value problem has a solution on I. We define a function $\mathbf{x}(t)$ on (α, β) as follows: Given $t \in (\alpha, \beta), t$ belongs to some open interval I such that initial value problem has a solution $\mathbf{u}(t)$ on I; for this given $t \in (\alpha, \beta)$, define $\mathbf{x}(t) = \mathbf{u}(t)$. Then $\mathbf{x}(t)$ is a well-defined function of t since if $t \in I_1 \cap I_2$ where I_1 and I_2 are any two open intervals such that initial value problem has solutions $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ on I_1 and I_2 respectively, then by the lemma $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ on the open interval $I_1 \cap I_2$. Also, $\mathbf{x}(t)$ is a solution of initial value problem on (α, β) since each point $t \in (\alpha, \beta)$ is contained in some open interval I on which the initial value problem has a unique solution $\mathbf{u}(t)$ and since $\mathbf{x}(t)$ agrees with $\mathbf{u}(t)$ on I. The fact that J is open follows from the fact that any solution of initial value problem on an interval $(\alpha, \beta]$ can be uniquely continued to a solution on an interval $(\alpha, \beta + a)$ with $a > 0$ as in the proof of Theorem 2 below.

Theorem: Let E be an open subset of $\mathbb{R}^n$ containing $\mathbf{x}_0$, let $\mathbf{f} \in \mathcal{C}^1(E)$, and let (α, β) be the maximal interval of existence of the solution $\mathbf{x}(t)$ of the initial value problem. Assume that $\beta < \infty$. Then given any compact set $K \subset E$, there exists a $t \in (\alpha, \beta)$ such that $\mathbf{x}(t) \notin K$.

Proof: Since f is continuous on the compact set K, there is a positive number M such that $|\mathbf{f}(\mathbf{x})| \leq M \ \forall \ \mathbf{x} \in K$. Let $\mathbf{x}(t)$ be the solution of the initial value problem on its maximal interval of existence (α, β) and assume that $\beta < \infty$ and that $\mathbf{x}(t) \in K \ \forall \ t \in (\alpha, \beta)$. We first show that $\lim_{t \to \beta^-} \mathbf{x}(t)$ exists. If $\alpha < t_1 < t_2 < \beta$ then

$$|\mathbf{x}(t_1) - \mathbf{x}(t_2)| \leq \int_{t_1}^{t_2} |\mathbf{f}(\mathbf{x}(s))| ds \leq M |t_2 - t_1|$$

Thus as t_1 and t_2 approach β from the left, $|\mathbf{x}(t_2) - \mathbf{x}(t_1)| \rightarrow 0$ which, by the Cauchy criterion for convergence in $\mathbb{R}^n$ (i.e., the completeness of $\mathbb{R}^n$) implies that $\lim_{t \to \beta^-} \mathbf{x}(t)$ exists. Let $\mathbf{x}_1 = \lim_{t \to \beta^-} \mathbf{x}(t)$. Then $\mathbf{x}_1 \in K \subset E$ since K is compact. Next define the function $\mathbf{u}(t)$ on $(\alpha, \beta]$ by

$$\mathbf{u}(t) = \begin{cases} \mathbf{x}(t) & \text{for } t \in (\alpha, \beta) \\ \mathbf{x}_1 & \text{for } t = \beta \end{cases}$$

Then $\mathbf{u}(t)$ is differentiable on $(\alpha, \beta]$. Indeed,

$$\mathbf{u}(t) = \mathbf{x}_0 + \int_0^t \mathbf{f}(\mathbf{u}(s)) ds$$

which implies that

$$\mathbf{u}'(\beta) = \mathbf{f}(\mathbf{u}(\beta))$$

i.e., $\mathbf{u}(t)$ is a solution of the initial value problem on $(\alpha, \beta]$. The function $\mathbf{u}(t)$ is called the continuation of the solution $\mathbf{x}(t)$ to $(\alpha, \beta]$. Since $\mathbf{x}_1 \in E$, it follows from the fundamental existence-uniqueness theorem in Section 2.2 that the initial value problem $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ together with $\mathbf{x}(\beta) = \mathbf{x}_1$ has a unique solution $\mathbf{x}_1(t)$ on some interval $(\beta - a, \beta + a)$. By the above lemma, $\mathbf{x}_1(t) = \mathbf{u}(t)$ on $(\beta - a, \beta)$ and $\mathbf{x}_1(\beta) = \mathbf{u}(\beta) = \mathbf{x}_1$. So if we define

$$\mathbf{v}(t) = \begin{cases} \mathbf{u}(t) & \text{for} \quad t \in (\alpha, \beta] \\ \mathbf{x}_1(t) & \text{for} \quad t \in [\beta, \beta + a) \end{cases}$$

then $\mathbf{v}(t)$ is a solution of the initial value problem on $(\alpha, \beta + a)$. But this contradicts the fact that (α, β) is the maximal interval of existence for
the initial value problem. Hence, if $\beta < \infty$, it follows that there exists a $t \in (\alpha, \beta)$ such that $\mathbf{x}(t) \notin K$.

If (α, β) is the maximal interval of existence for the initial value problem then $0 \in (\alpha, \beta)$ and the intervals $[0, \beta)$ and $(\alpha, 0]$ are called the right and left maximal intervals of existence respectively. Essentially the same proof yields the following result.

Theorem: Let E be an open subset of $\mathbb{R}^n$ containing $\mathbf{x}_0$, let $\mathbf{f} \in \mathcal{C}^1(E)$, and let $[0, \beta)$ be the right maximal interval of existence of the solution $\mathbf{x}(t)$ of the initial value problem. Assume that $\beta < \infty$. Then given any compact set $K \subset E$, there exists a $t \in (0, \beta)$ such that $\mathbf{x}(t) \notin K$.

Corollary: Under the hypothesis of the above theorem, if $\beta < \infty$ and if $\lim_{t \to \beta^-} \mathbf{x}(t)$ exists then $\lim_{t \to \beta^-} \mathbf{x}(t) \in \dot{E}$.

Proof: If $\mathbf{x}_1 = \lim_{t \to \beta^-} \mathbf{x}(t)$, then the function

$$\mathbf{u}(t) = \begin{cases} \mathbf{x}(t) & \text{for} \quad t \in [0, \beta) \\ \mathbf{x}_1 & \text{for} \quad t = \beta \end{cases}$$

is continuous on $[0, \beta]$. Let K be the image of the compact set $[0, \beta]$ under the continuous map $\mathbf{u}(t)$; i.e.,

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{u}(t) \text{ for some } t \in [0, \beta]\}$$

Then K is compact. Assume that $\mathbf{x}_1 \in E$. Then $K \subset E$ and it follows from Theorem 3 that there exists a $t \in (0, \beta)$ such that $\mathbf{x}(t) \notin K$. This is a contradiction and therefore $\mathbf{x}_1 \notin E$. But since $\mathbf{x}(t) \in E \;\forall\; t \in [0, \beta)$, it follows that $\mathbf{x}_1 = \lim_{t \to \beta^-} \mathbf{x}(t) \in \bar{E}$. Therefore $\mathbf{x}_1 \in \bar{E} \sim E$; i.e., $\mathbf{x}_1 \in \dot{E}$.

Corollary: Let E be an open subset of $\mathbb{R}^n$ containing $\mathbf{x}_0$, let $\mathbf{f} \in \mathcal{C}^1(E)$, and let $[0, \beta)$ be the right maximal interval of existence of the solution $\mathbf{x}(t)$ of the initial value problem. Assume that there exists a compact set $K \subset E$ such that

$$\{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y} = \mathbf{x}(t) \text{ for some } t \in [0, \beta)\} \subset K$$

It then follows that $\beta = \infty$; i.e. the initial value problem has a solution $\mathbf{x}(t)$ on $[0, \infty)$.

Proof: This corollary is just the contrapositive of the statement of the aforementioned theorem.

We next prove the following theorem which strengthens the result on uniform convergence with respect to initial conditions.

Theorem: Let E be an open subset of $\mathbb{R}^n$ containing $\mathbf{x}_0$ and let $\mathbf{f} \in \mathcal{C}^1(E)$. Suppose that the initial value problem has a solution $\mathbf{x}\left(t, \mathbf{x}_0\right)$ defined on a closed interval $[a, b]$. Then there exists a $\delta > 0$ and a positive constant K such that $\forall\ \mathbf{y} \in N_\delta\left(\mathbf{x}_0\right)$ the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\ \mathbf{x}(0) &= \mathbf{y} \end{aligned} \tag{2}$$

has a unique solution $\mathbf{x}(t, \mathbf{y})$ defined on $[a, b]$ which satisfies

$$\left|\mathbf{x}(t, \mathbf{y}) - \mathbf{x}\left(t, \mathbf{x}_0\right)\right| \leq \left|\mathbf{y} - \mathbf{x}_0\right| e^{K|t|}$$

and

$$\lim_{\mathbf{y} \to \mathbf{x}_0} \mathbf{x}(t, \mathbf{y}) = \mathbf{x}\left(t, \mathbf{x}_0\right)$$

uniformly $\forall\ t \in [a, b]$.

Remark: If in Theorem 4 we have a function $\mathbf{f}(\mathbf{x}, \boldsymbol{\mu})$ depending on a parameter $\boldsymbol{\mu} \in \mathbb{R}^m$ which satisfies $\mathbf{f} \in \mathcal{C}^1(E)$ where E is an open subset of $\mathbb{R}^{n+m}$ containing $(\mathbf{x}_0, \boldsymbol{\mu}_0)$, it can be shown that if for $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ the initial value problem has a solution $\mathbf{x}\left(t, \mathbf{x}_0, \boldsymbol{\mu}_0\right)$ defined on a closed interval $a \leq t \leq b$, then there is a $\delta > 0$ and a $K > 0$ such that $\forall\ \mathbf{y} \in N_\delta\left(\mathbf{x}_0\right)$ and $\boldsymbol{\mu} \in N_\delta\left(\boldsymbol{\mu}_0\right)$ the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}) \\ \mathbf{x}(0) &= \mathbf{y} \end{aligned}$$

has a unique solution $\mathbf{x}(t, \mathbf{y}, \boldsymbol{\mu})$ defined for $a \leq t \leq b$ which satisfies

$$\left|\mathbf{x}(t, \mathbf{y}, \boldsymbol{\mu}) - \mathbf{x}\left(t, \mathbf{x}_0, \boldsymbol{\mu}_0\right)\right| \leq \left[\left|\mathbf{y} - \mathbf{x}_0\right| + \left|\boldsymbol{\mu} - \boldsymbol{\mu}_0\right|\right] e^{K|t|}$$

and

$$\lim_{(\mathbf{y},\boldsymbol{\mu})\to(\mathbf{x}_0,\boldsymbol{\mu}_0)} \mathbf{x}(t,\mathbf{y},\boldsymbol{\mu}) = \mathbf{x}\left(t,\mathbf{x}_0,\boldsymbol{\mu}_0\right)$$

uniformly $\forall\ t \in [a,b]$.
In order to prove this theorem, we first establish the following lemma.

Lemma: Let E be an open subset of $\mathbb{R}^n$ and let A be a compact subset of E. Then if $\mathbf{f} : E \to \mathbb{R}^n$ is locally Lipschitz on E, it follows that $\mathbf{f}$ satisfies a Lipschitz condition on A.

Proof: Let M be the maximal value of the continuous function f on the compact set A. Suppose that $\mathbf{f}$ does not satisfy a Lipschitz condition on A. Then for every $K > 0$, we can find $\mathbf{x}, \mathbf{y} \in A$ such that

$$|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})| > K|\mathbf{y} - \mathbf{x}|$$

In particular, there exist sequences $\mathbf{x}_n$ and $\mathbf{y}_n$ in A such that

$$\left|\mathbf{f}\left(\mathbf{y}_n\right) - \mathbf{f}\left(\mathbf{x}_n\right)\right| > n\left|\mathbf{y}_n - \mathbf{x}_n\right| \tag{*}$$

for $n = 1, 2, 3, \ldots$. Since A is compact, there are convergent subsequences, call them $\mathbf{x}_n$ and $\mathbf{y}_n$ for simplicity in notation, such that $\mathbf{x}_n \to \mathbf{x}^*$ and $\mathbf{y}_n \to \mathbf{y}^*$ with $\mathbf{x}^*$ and $\mathbf{y}^*$ in A. It follows that $\mathbf{x}^* = \mathbf{y}^*$ since $\forall\ n = 1, 2, 3, \ldots$

$$|\mathbf{y}^* - \mathbf{x}^*| = \lim_{n\to\infty}\left|\mathbf{y}_n - \mathbf{x}_n\right| \leq \frac{1}{n}\left|\mathbf{f}\left(\mathbf{y}_n\right) - \mathbf{f}\left(\mathbf{x}_n\right)\right| \leq \frac{2M}{n}$$

Now, by hypotheses, there exists a neighborhood N_0 of $\mathbf{x}^*$ and a constant K_0 such that $\mathbf{f}$ satisfies a Lipschitz condition with Lipschitz constant K_0 $\forall$ $\mathbf{x}$ and $\mathbf{y} \in N_0$. But since $\mathbf{x}_n$ and $\mathbf{y}_n$ approach $\mathbf{x}^*$ as $n \to \infty$, it follows that $\mathbf{x}_n$ and $\mathbf{y}_n$ are in N_0 for n sufficiently large; i.e., for n sufficiently large

$$\left|\mathbf{f}\left(\mathbf{y}_n\right) - \mathbf{f}\left(\mathbf{x}_n\right)\right| \leq K\left|\mathbf{y}_n - \mathbf{x}_n\right|.$$

But for $n \geq K$, this contradicts the above inequality (*) and this establishes the lemma.

Proof of Theorem: Since $[a, b]$ is compact and $\mathbf{x}(t, \mathbf{x}_0)$ is a continuous function of $t, \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{x}(t, \mathbf{x}_0)$ and $a \leq t \leq b\}$ is a compact subset of E. And since E is open, there exists an $\varepsilon > 0$ such that the compact set

$$A = \{\mathbf{x} \in \mathbb{R}^n || \mathbf{x} - \mathbf{x}(t, \mathbf{x}_0) | \leq \varepsilon \text{ and } a \leq t \leq b\}$$

is a subset of E. Since $\mathbf{f} \in \mathcal{C}^1(E)$, $\mathbf{f}$ is locally Lipschitz in E; and then by the above lemma, $\mathbf{f}$ satisfies a Lipschitz condition

$$|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})| \leq K|\mathbf{y} - \mathbf{x}|$$

$\forall\ \mathbf{x}, \mathbf{y} \in A$. Choose $\delta > 0$ so small that $\delta \leq \varepsilon$ and $\delta \leq \varepsilon e^{-K(b-a)}$. Let $\mathbf{y} \in N_\delta(\mathbf{x}_0)$ and let $\mathbf{x}(t, \mathbf{y})$ be the solution of the initial value problem on its maximal interval of existence (α, β). We shall show that $[a, b] \subset (\alpha, \beta)$. Suppose that $\beta \leq b$. It then follows that $\mathbf{x}(t, \mathbf{y}) \in A\ \forall\ t \in (\alpha, \beta)$ because if this were not true then there would exist a $t^* \in (\alpha, \beta)$ such that $\mathbf{x}(t, \mathbf{x}_0) \in A$ for $t \in (\alpha, t^*]$ and $\mathbf{x}(t^*, \mathbf{y}) \in \dot{A}$. But then

$$\begin{aligned}
|\mathbf{x}(t, \mathbf{y}) - \mathbf{x}(t, \mathbf{x}_0)| &\leq |\mathbf{y} - \mathbf{x}_0| + \int_0^t | \mathbf{f}(\mathbf{x}(s, \mathbf{y})) - f(\mathbf{x}(s, \mathbf{x}_0) | \, ds \\
&\leq |\mathbf{y} - \mathbf{x}_0| + K \int_0^t |\mathbf{x}(s, \mathbf{y}) - \mathbf{x}(s, \mathbf{x}_0)| \, ds
\end{aligned}$$

$\forall\ t \in (\alpha, t^*]$. And then by Gronwall's Lemma, it follows that

$$|\mathbf{x}(t^*, \mathbf{y}) - \mathbf{x}(t^*, \mathbf{x}_0)| \leq |\mathbf{y} - \mathbf{x}_0| \, e^{K|t^*|} < \delta e^{K(b-a)} < \varepsilon$$

since $t^* < \beta \leq b$. Thus $\mathbf{x}(t^*, \mathbf{y})$ is an interior point of A, a contradiction Thus, $\mathbf{x}(t, \mathbf{y}) \in A$ $\forall\ t \in (\alpha, \beta)$. But then (α, β) is not the maximal interval of existence of $\mathbf{x}(t, \mathbf{y})$, a contradiction. Thus $b - \beta$ It is similarly proved that $\alpha < a$. Hence, $\forall\ \mathbf{y} \in N_\delta(\mathbf{x}_0)$, the initial value problem has a unique solution defined on $[a, b]$. Furthermore, we assume that there is a $t^* \in [a, b)$ such that $\mathbf{x}(t, \mathbf{y}) \in A\ \forall\ t \in [a, t^*]$. and $\mathbf{x}(t^*, \mathbf{y}) \in \dot{A}$, a repeat of the above argument based on Gronwall's Lemma leads to a contradiction and shows that $\mathbf{x}(t, \mathbf{y}) \in A\ \forall\ t \in [a, b]$ and hence that

$$|\mathbf{x}(t, \mathbf{y}) - \mathbf{x}(t, \mathbf{x}_0)| \leq |\mathbf{y} - \mathbf{x}_0| \, e^{K|t|}$$

$\forall\ t \in [a, b]$. It then follows that

$$\lim_{\mathbf{y}\to\mathbf{x}_0} \mathbf{x}(t,\mathbf{y}) = \mathbf{x}\left(t,\mathbf{x}_0\right)$$

uniformly $\forall\ t \in [a,b]$.

2.5 The Flow defined by a Differential Equation

We have already defined the flow, $e^{At} : \mathbb{R}^n \to \mathbb{R}^n$, of the linear system

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

The mapping $\phi_t = e^{At}$ satisfies the following basic properties $\forall\ \mathbf{x} \in \mathbb{R}^n$:

(i) $\boldsymbol{\phi}_0(\mathbf{x}) = \mathbf{x}$

(ii) $\phi_s\left(\phi_t(\mathbf{x})\right) = \phi_{s+t}(\mathbf{x})\ \forall\ s,t \in \mathbb{R}$

(iii) $\phi_{-t}\left(\phi_t(\mathbf{x})\right) = \phi_t\left(\boldsymbol{\phi}_{-t}(\mathbf{x})\right) = \mathbf{x}\ \forall\ t \in \mathbb{R}$.

Observe that these properties follow either from the definitions, or from the properties we have already proved.
In this section, we define the flow, ϕ_t, of the nonlinear system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

and show that it satisfies these same basic properties. In the following definition, we denote the maximal interval of existence (α,β) of the solution $\phi\left(t,\mathbf{x}_0\right)$ of the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned}$$

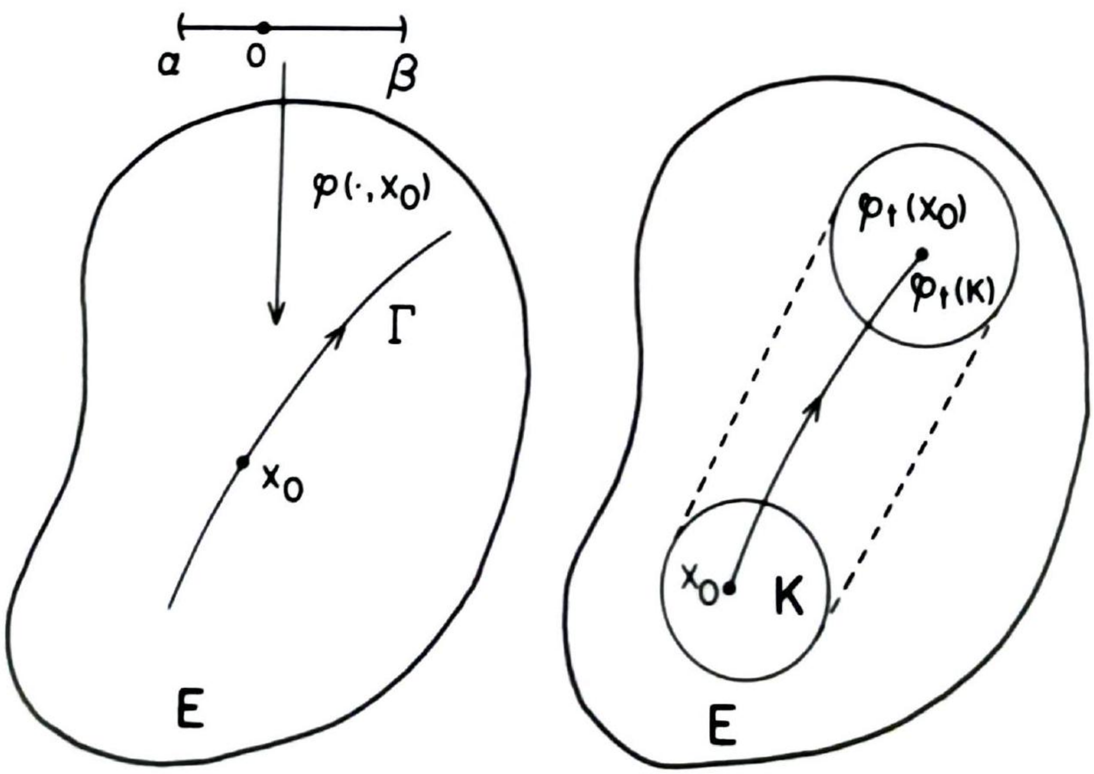
by $I\left(\mathbf{x}_0\right)$ since the endpoints α and β of the maximal interval generally depend on $\mathbf{x}_0$.

Definition: Let E be an open subset of $\mathbb{R}^n$ and let $\mathbf{f} \in \mathcal{C}^1(E)$. For $\mathbf{x}_0 \in E$, let $\phi(t, \mathbf{x}_0)$ be the solution of the initial value problem defined on its maximal interval of existence $I(\mathbf{x}_0)$. Then for $t \in I(\mathbf{x}_0)$, the set of mappings ϕ_t defined by

$$\phi_t(\mathbf{x}_0) = \phi(t, \mathbf{x}_0)$$

is called the flow of the differential equation or the flow defined by the differential equation; $\boldsymbol{\phi}_t$ is also referred to as the flow of the vector field $\mathbf{f}(\mathbf{x})$.

If we think of the initial point $\mathbf{x}_0$ as being fixed and let $I = I(\mathbf{x}_0)$, then the mapping $\phi(\cdot, \mathbf{x}_0) : I \to E$ defines a solution curve or trajectory of the system the concerned differential equation through the point $\mathbf{x}_0 \in E$. As usual, the mapping $\phi(\cdot, \mathbf{x}_0)$ is identified with its graph in $I \times E$ and a trajectory is visualized as a motion along a curve Γ through the point $\mathbf{x}_0$ in the subset E of the phase space $\mathbb{R}^{\boldsymbol{n}}$; cf. Figure 1. On the other hand, if we think of the point $\mathbf{x}_0$ as varying throughout $K \subset E$, then the flow of the differential equation, $\phi_t : K \to E$ can be viewed as the motion of all the points in the set K.



(a) (b)

Figure 4: (a) A trajectory Γ of the system (b)The flow ϕ_t of the system

If we think of the differential equation as describing the motion of a flund. then a trajectory of the concerned differential equation describes the motion of an individual particle in the fluid while the flow of the differential equation describes the mostion of the entire fluid.

We now show that the basic properties (i) (iii) of linear flows are also satisficd by nonlinear flows. But first we extend Theorem 1 of Section 2.3, establishing that $\phi(t, \mathbf{x}_0)$ is a locally smooth function, to a global result. Using the same notation as in Definition 1, let us define the set $\Omega \subset \mathbb{R} \times E$ as

$$\Omega = \{(t, \mathbf{x}_0) \in \mathbb{R} \times E \mid t \in I(\mathbf{x}_0)\}$$

Theorem: Let E be an open subset of $\mathbb{R}^n$ and let $\mathbf{f} \in \mathcal{C}^1(E)$. Then Ω is an open subset of $\mathbb{R} \times E$ and $\phi \in \mathcal{C}^1(\Omega)$.

Proof: If $(t_0, \mathbf{x}_0) \in \Omega$ and $t_0 > 0$, then according to the definition of the set Ω, the solution $\mathbf{x}(t) = \boldsymbol{\phi}(t, \mathbf{x}_0)$ of the initial value problem is defined on $[0, t_0]$. Thus, as in the proof of Theorem 2 in Section 2.4, the solution $\mathbf{x}(t)$ can be extended to an interval $[0, t_0 + \varepsilon]$ for some $\varepsilon > 0$; i.e., $\phi(t, \mathbf{x}_0)$ is defined on the closed interval $[t_0 - \varepsilon, t_0 + \varepsilon]$. It then follows from Theorem 4 in Section 2.4 that there exists a neighborhood of $\mathbf{x}_0, N_\delta(\mathbf{x}_0)$, such that $\boldsymbol{\phi}(t, \mathbf{y})$ is defined on $\mid t_0 - \varepsilon, t_0 + \varepsilon] \times N_\delta(\mathbf{x}_0)$; i.e., $(t_0 - \varepsilon, t_0 + \varepsilon) \times N_\delta(\mathbf{x}_0) \subset \Omega$. Therefore, Ω is open in $\mathbb{R} \times E$. It follows from Theorem 4 in Section 2.4 that $\phi \in \mathcal{C}^1(G)$ where $G = (t_0 - \varepsilon, t_0 + \varepsilon) \times N_\delta(\mathbf{x}_0)$. A similar proof holds for $t_0 \leq 0$, and since $(t_0, \mathbf{x}_0)$ is an arbitrary point in Ω, it follows that $\phi \in \mathcal{C}^1(\Omega)$.

Remark: This theorem can be generalized to show that if f $\in C^r(E)$ with $r \geq 1$, then $\phi \in C^r(\Omega)$ and that if $\mathbf{f}$ is analytic in E, then ϕ is analytic in Ω.

Theorem: Let E be an open set of $\mathbb{R}^n$ and let $\mathbf{f} \in \mathcal{C}^1(E)$. Then $\forall\ \mathbf{x}_0 \in E$, if $t \in I(\mathbf{x}_0)$ and $s \in I(\phi_t(\mathbf{x}_0))$, it follows that $s + t \in I(\mathbf{x}_0)$ and

$$\phi_{s+t}(\mathbf{x}_0) = \phi_s(\phi_t(\mathbf{x}_0)).$$

Proof: Suppose that $s > 0, t \in I(\mathbf{x}_0)$ and $s \in I(\phi_t(\mathbf{x}_0))$. Let the maximal interval $I(\mathbf{x}_0) = (\alpha, \beta)$ and define the function $\mathbf{x} : (\alpha, s + t] \rightarrow E$ by

$$\mathbf{x}(r) = \begin{cases} \phi(r, \mathbf{x}_0) & \text{if } \alpha < r \leq t \\ \phi(r - t, \phi_t(\mathbf{x}_0)) & \text{if } t \leq r \leq s + t \end{cases}$$

Then $\mathbf{x}(r)$ is a solution of the initial value problem on $(\alpha, s + t]$. Hence $s + t \in I(\mathbf{x}_0)$ and by uniqueness of solutions

$$\phi_{s+t}\left(\mathbf{x}_0\right)=\mathbf{x}(s+t)=\phi\left(s,\phi_t\left(\mathbf{x}_0\right)\right)=\phi_s\left(\phi_t\left(\mathbf{x}_0\right)\right)$$

If $s=0$ the statement of the theorem follows immediately. And if $s<0$, then we define the function $\mathbf{x}:[s+t,\beta)\rightarrow E$ by

$$\mathbf{x}(t)=\begin{cases}\phi\left(r,\mathbf{x}_0\right) & \text{if } t\leq r<\beta\\ \phi\left(r-t,\phi_t\left(\mathbf{x}_0\right)\right) & \text{if } s+t\leq r\leq t\end{cases}$$

Then $\mathbf{x}(r)$ is a solution of the initial value problem on $[s+t,\beta)$ and the last statement of the theorem follows from the uniqueness of solutions as above.

Theorem: Under the hypotheses of the first theorem of this section , if $(t,\mathbf{x}_0)\in\Omega$ then there exists a neighborhood U of $\mathbf{x}_0$ such that $\{t\}\times U\subset\Omega$. It then follows that the set $V=\phi_t(U)$ is open in E and that

$$\phi_{-t}\left(\phi_t(\mathbf{x})\right)=\mathbf{x}\ \forall\ \mathbf{x}\in U$$

and

$$\boldsymbol{\phi}_t\left(\boldsymbol{\phi}_{-t}(\mathbf{y})\right)=\mathbf{y}\ \forall\ \mathbf{y}\in V$$

Proof: If $(t,\mathbf{x}_0)\in\Omega$ then it follows as in the proof of Theorem 1 that there exists a neighborhood of $\mathbf{x}_0,U=N_\delta\left(\mathbf{x}_0\right)$, such that $(t-\varepsilon,t+\varepsilon)\times U\subset\Omega$; thus, $\{t\}\times U\subset\Omega$. For $\mathbf{x}\in U$, let $\mathbf{y}=\phi_t(\mathbf{x})\ \forall\ t\in I(\mathbf{x})$. Then
$-t\in I(\mathbf{y})$ since the function $\mathbf{h}(s)=\boldsymbol{\phi}(s+t,\mathbf{y})$ is a solution of the concerned differential equation on $|-t,0|$ that satisfies $\mathbf{h}(-t)=\mathbf{y}$; i.e., $\boldsymbol{\phi}_-$, is defined on the set $V=\phi_t(U)$. It then follows from the previous theorem that $\boldsymbol{\phi}_1\left(\boldsymbol{\phi}_1(\mathbf{x})\right)=\phi_0(\mathbf{x})=\mathbf{x}\ \forall\ \mathbf{x}\in U$ and that $\phi_t\left(\phi_t(\mathbf{y})\right)=\phi_0(\mathbf{y})=\mathbf{y}\ \forall\ \boldsymbol{y}\in V$. It remains to prove that V is open. Let $V^*\supset V$ be the maximal subset of E on which ϕ_{-t} is defined. V^* is open because Ω is open and $\phi_-:V^*\rightarrow E$ is continuous because ϕ is continuous. Therefore, the inverse image of the open set U under the continuous map ϕ_{-t}, i.e., $\phi_t(U)$, is open in E. Thus, V is open in E.

Later we intend show that the time along each trajectory of the concerned differential equation can be rescaled, without affecting the phase portrait of the concerned differential equation, so that $\forall\ \mathbf{x}_0\in E$, the solution $\phi\left(t,\mathbf{x}_0\right)$ of the initial value problem is defined $\forall$ $t\in\mathbb{R}$; i.e., $\forall\ \mathbf{x}_0\in E,I\left(\mathbf{x}_0\right)=(-\infty,\infty)$. This rescaling avoids some of the complications found in stating the above theorems. Once this rescaling has been made, it follows that $\Omega=\mathbb{R}\times E,\phi\in\mathcal{C}^1(\mathbb{R}\times E),\phi_t\in\mathcal{C}^1(E)\ \forall\ t\in\mathbb{R}$, and properties (i)-(iii) for the flow of the nonlinear system the concerned differential equation hold $\forall\ t\in\mathbb{R}$ and $\mathbf{x}\in E$ just as for the linear flow e^{At}.From now on, it will be assumed that this rescaling has been made so that $\forall\ \mathbf{x}_0\in E,\phi\left(t,\mathbf{x}_0\right)$ is defined $\forall\ t\in\mathbb{R}$; i.e., we shall assume throughout the remainder of this chapter that the flow of the nonlinear system the concerned differential equation $\phi_t \in \mathcal{C}^1(E)\ \forall\ t \in \mathbb{R}$.

> **Definition:** Let E be an open subset of $\mathbb{R}^n$, let $\mathbf{f} \in \mathcal{C}^1(E)$, and let $\phi_t : E \to E$ be the flow of the differential equation defined $\forall\ t \in \mathbb{R}$. Then a set $S \subset E$ is called invariant with respect to the flow ϕ_t if $\phi_t(S) \subset S\ \forall\ t \in \mathbb{R}$ and S is called positively (or negatively) invariant with respect to the flow ϕ_t if $\phi_t(S) \subset S\ \forall\ t \geq 0$ (or $t \leq 0$).

We have already showed that the stable, unstable and center subspaces of the linear system $\dot{\mathbf{x}} = A\mathbf{x}$ are invariant under the linear flow $\phi_t = e^{At}$. A similar result will be established for the nonlinear flow ϕ_t of the concerned differential equation.

2.6 Linearization

A good place to start analyzing the nonlinear system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

is to determine the equilibrium points of the concerned differential equation and to describe the behavior of the concerned differential equation near its equilibrium points. In the next two sections it is shown that the local behavior of the nonlinear system the concerned differential equation near a hyperbolic equilibrium point $\mathbf{x}_0$ is qualitatively determined by the behavior of the linear system

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

with the matrix $A = D\mathbf{f}\left(\mathbf{x}_0\right)$, near the origin. The linear function $A_{\mathbf{x}} = \mathrm{D}\ \mathbf{f}\left(\mathbf{x}_0\right)\mathbf{x}$ is called the linear part of $\mathbf{f}$ at $\mathbf{x}_0$.

> **Definition:** A point $\mathbf{x}_0 \in \mathbb{R}^n$ is called an equilibrium point or critical point of the concerned differential equation if $\mathbf{f}\left(\mathbf{x}_0\right) = \mathbf{0}$. An equilibrium point $\mathbf{x}_0$ is called a hyperbolic equilibrium point of the concerned differential equation if none of the eigenvalues of the matrix $D\mathbf{f}\left(\mathbf{x}_0\right)$ have zero real part. The linear system with the matrix $A = D\mathbf{f}\left(\mathbf{x}_0\right)$ is called the linearization of the concerned differential equation at $\mathbf{x}_0$.

If $\mathbf{x}_0 = \mathbf{0}$ is an equilibrium point of the concerned differential equation, then $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ and, by Taylor's Theorem,

$$\mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{0})\mathbf{x} + \frac{1}{2}D^2\mathbf{f}(\mathbf{0})(\mathbf{x},\mathbf{x}) + \cdots$$

It follows that the linear function $D\mathbf{f}(\mathbf{0})\mathbf{x}$ is a good first approximation to the nonlinear function $\mathbf{f}(\mathbf{x})$ near $\mathbf{x} = \mathbf{0}$ and it is reasonable to expect that the behavior of the nonlinear system the concerned differential equation near the point $\mathbf{x} = \mathbf{0}$ will be approximated by the behavior of its linearization at $\mathbf{x} = \mathbf{0}$. Later it will be shown that this is indeed the case if the matrix $D\mathbf{f}(\mathbf{0})$ has no zero or pure imaginary eigenvalues.

Note that if $\mathbf{x}_0$ is an equilibrium point of the concerned differential equation and $\boldsymbol{\phi}_t : E \to \mathbb{R}^n$ is the flow of the differential equation, then $\phi_t(\mathbf{x}_0) = \mathbf{x}_0 \ \forall\ t \in \mathbb{R}$. Thus, $\mathbf{x}_0$ is called a fixed point of the flow ϕ_t; it is also called a zero, a critical point, or a singular point of the vector field $\mathbf{f} : E \to \mathbb{R}^n$. We next give a rough classification of the equilibrium points of the concerned differential equation according to the signs of the real parts of the eigenvalues of the matrix $D\mathbf{f}(\mathbf{x_0})$.

Definition: An equilibrium point $\mathbf{x}_0$ of the concerned differential equation is called a sink if all of the eigenvalues of the matrix $D\mathbf{f}(\mathbf{x}_0)$ have negative real part; it is called a source if all of the eigenvalues of $D\mathbf{f}(\mathbf{x}_0)$ have positive real part; and it is called a saddle if it is a hyperbolic equilibrium point and $D\mathbf{f}(\mathbf{x}_0)$ has at least one eigenvalue with a positive real part and at least one with a negative real part.

Later we shall see that if $\mathbf{x}_0$ is a hyperbolic equilibrium point of the concerned differential equation then the local behavior of the nonlinear system the concerned differential equation is topologically equivalent to the local behavior of the linear system ; i.e., there is a continuous one-to-one map of a neighborhood of $\mathbf{x}_0$ onto an open set U containing the origin, $H : N_\varepsilon(\mathbf{x}_0) \to U$, which transforms the concerned differential equation into the linear system, maps trajectories of the concerned differential equation in $N_\varepsilon(\mathbf{x}_0)$ onto trajectories of the linear system in the open set U, and preserves the orientation of the trajectories by time, i.e., H preserves the direction of the flow along the trajectories.

2.7 Stable Manifold Theorem

The stable manifold theorem is one of the most important results in the local qualitative theory of ordinary differential equations. The theorem shows that near a hyperbolic equilibrium point $\mathbf{x}_0$, the nonlinear system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

has stable and unstable manifolds S and U tangent at $\mathbf{x}_0$ to the stable and unstable subspaces E^s and E^u of the linearized system

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

where $A = D\mathbf{f}(\mathbf{x}_0)$. Furthermore, S and U are of the same dimensions as E^s and E^u, and if ϕ_t is the flow of the nonlinear system, then S and U are positively and negatively invariant under ϕ_t respectively and satisfy

$$\lim_{t\to\infty} \phi_t(\mathbf{c}) = \mathbf{x}_0 \ \forall \ \mathbf{c} \in S$$

and

$$\lim_{t\to-\infty} \phi_t(\mathbf{c}) = \mathbf{x}_0 \ \forall \ \mathbf{c} \in U$$

We first illustrate these ideas with an example and then make them more precise by proving the stable manifold theorem. It is assumed that the equilibrium point $\mathbf{x}_0$ is located at the origin throughout the remainder of this section. If this is not the case, then the equilibrium point $\mathbf{x}_0$ can be translated to the origin by the affine transformation of coordinates $\mathbf{x} \to \mathbf{x} - \mathbf{x}_0$.

Definition: Let X be a metric space and let A and B be subsets of X. A homeomorphism of A onto B is a continuous one-to-one map of A onto B, $h : A \to B$, such that $h^{-1} : B \to A$ is continuous. The sets A and B are called homeomorphic or topologically equivalent if there is a homeomorphism of A onto B. If we wish to emphasize that h maps A onto B, we write $h : A \twoheadrightarrow B$.

Defnition: An n-dimensional differentiable manifold, M (or a manifold of class $\mathcal{C}^k$), is a connected metric space with an open covering $\{U_\alpha\}$, i.e., $M = \bigcup_\alpha U_\alpha$, such that

1. $\forall \ \alpha, U_\alpha$ is homeomorphic to the open unit ball in $\mathbb{R}^n, B = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x}| < 1\}$, i.e., $\forall \ \alpha$ there exists a homeomorphism of U_α onto $B, \mathbf{h}_\alpha : U_\alpha \to B$

2. if $U_\alpha \cap U_\beta \neq \emptyset$ and $\mathbf{h}_\alpha : U_\alpha \to B, \mathbf{h}_\beta : U_\beta \to B$ are homeomorphisms, then $\mathbf{h}_\alpha(U_\alpha \cap U_\beta)$ and $\mathbf{h}_\beta(U_\alpha \cap U_\beta)$ are subsets of $\mathbb{R}^n$ and the map

$$\mathbf{h} = \mathbf{h}_\alpha \circ \mathbf{h}_\beta^{-1} : \mathbf{h}_\beta(U_\alpha \cap U_\beta) \to \mathbf{h}_\alpha(U_\alpha \cap U_\beta)$$

is differentiable (or of class $\mathcal{C}^k$) and $\forall \ \mathbf{x} \in \mathbf{h}_\beta(U_\alpha \cap U_\beta)$, the Jacobian determinant $\det D\mathbf{h}(\mathbf{x}) \neq 0$.

The manifold M is said to be analytic if the maps $h = h_\alpha \circ h_\beta^{-1}$ are analytic. The pair $(U_\alpha, \mathbf{h}_\alpha)$ is called a chart for the manifold M and the set of all charts is called an atlas for M. The differentiable manifold M is called orientable if there is an atlas with $\det D\mathbf{h}_\alpha \circ \mathbf{h}_\beta^{-1}(\mathbf{x}) > 0 \ \forall\ \alpha, \beta$ and $\mathbf{x} \in \mathbf{h}_\beta\left(U_\alpha \cap U_\beta\right)$.

The Stable Manifold Theorem: Let E be an open subset of $\mathbb{R}^n$ containing the origin, let $\mathbf{f} \in \mathcal{C}^1(E)$, and let $\boldsymbol{\phi}_t$ be the flow of the nonlinear system . Suppose that $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ and that D $\mathbf{f}(\mathbf{0})$ has k eigenvalues with negative real part and $n-k$ eigenvalues with positive real part. Then there exists a k-dimensional differentiable manifold S tangent to the stable subspace E^s of the linear system at 0 such that $\forall$ $t \geq 0, \phi_t(S) \subset S$ and $\forall\ \mathbf{x}_0 \in S$,

$$\lim_{t \to \infty} \phi_t\left(\mathbf{x}_0\right) = \mathbf{0}$$

and there exists an $n-k$ dimensional differentiable manifold U tangent to the unstable subspace E^u of the linear system at $\mathbf{0}$ such that $\forall\ t \leq 0, \phi_t(U) \subset U$
and $\forall\ \mathbf{x}_0 \in U$,

$$\lim_{t \to -\infty} \phi_t\left(\mathbf{x}_{\mathbf{0}}\right) = \mathbf{0}$$

Before proving this theorem, we remark that if $\mathbf{f} \in \mathcal{C}^1(E)$ and $\mathbf{f}(\mathbf{0}) = 0$, then the system can be written as

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{F}(\mathbf{x}) \tag{3}$$

where $A = D\mathbf{f}(\mathbf{0}), \mathbf{F}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - A\mathbf{x}, \mathbf{F} \in \mathcal{C}^1(E), \mathbf{F}(\mathbf{0}) = \mathbf{0}$ and $D\mathbf{F}(\mathbf{0}) = 0$. This in turn implies that $\forall\ \varepsilon > 0$ there is a $\delta > 0$ such that $|\mathbf{x}| \leq \delta$ and $|\mathbf{y}| \leq \delta$ imply that

$$|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})| \leq \varepsilon |\mathbf{x} - \mathbf{y}| \tag{4}$$

Furthermore, there is an $n \times n$ invertible matrix C such that

$$B = C^{-1}AC = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

where the eigenvalues $\lambda_1, \ldots, \lambda_k$ of the $k \times k$ matrix P have negative real part and the eigenvalues $\lambda_{k+1}, \ldots, \lambda_n$ of the $(n-k) \times (n-k)$ matrix Q have positive real part. We can choose $\alpha > 0$ sufficiently small that for $j = 1, \ldots, k$,

$$\operatorname{Re}\left(\lambda_j\right) < -\alpha < 0$$

Letting $\mathbf{y} = C^{-1}\mathbf{x}$, the system then has the form

$$\dot{\mathbf{y}} = B\mathbf{y} + \mathbf{G}(\mathbf{y})$$

where $\mathbf{G}(\mathbf{y}) = C^{-1}\mathbf{F}(C\mathbf{y}) \in \mathcal{C}^1(\tilde{E})$ where $\tilde{E} = C^{-1}(E)$ and $\mathbf{G}$ satisfies the Lipschitz-type condition above.

It will be shown in the proof that there are $n-k$ differentiable functions $\psi_j\left(y_1, \ldots, y_k\right)$ such that the equations

$$y_j = \psi_j\left(y_1, \ldots, y_k\right), \quad j = k+1, \ldots, n$$

define a k-dimensional differentiable manifold $\tilde{S}$ in $\mathbf{y}$-space. The differentiable manifold S in x-space is then obtained from $\tilde{S}$ under the linear transformation of coordinates $\mathbf{x} = C\mathbf{y}$.

Proof: Consider the system $\dot{\mathbf{y}} = B\mathbf{y} + \mathbf{G}(\mathbf{y})$. Let

$$U(t) = \begin{bmatrix} e^{Pt} & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad V(t) = \begin{bmatrix} 0 & 0 \\ 0 & e^{Qt} \end{bmatrix}$$

Then $\dot{U} = BU, \dot{V} = BV$ and

$$e^{Bt} = U(t) + V(t)$$

It is not difficult to see that with $\alpha > 0$ chosen as in the penultimate system, we can choose $K > 0$ sufficiently large and $\sigma > 0$ sufficiently small that

$$\|U(t)\| \leq Ke^{-(a+\sigma)t} \;\forall\; t \geq 0$$

and

$$\|V(t)\| \leq Ke^{\sigma t} \;\forall\; t \leq 0$$

Next consider the integral equation

$$\mathbf{u}(t,\mathbf{a}) = U(t)\mathbf{a} + \int_0^t U(t-s)\mathbf{G}(\mathbf{u}(s,\mathbf{a}))ds - \int_t^\infty V(t-s)\mathbf{G}(\mathbf{u}(s,\mathbf{a}))ds$$

If $\mathbf{u}(t,\mathbf{a})$ is a continuous solution of this integral equation, then it is a solution of the differential equation considered initially in the proof. We now solve this integral equation by the method of successive approximations. Let

$$\mathbf{u}^{(0)}(t,\mathbf{a}) = 0$$

and

$$\begin{aligned}\mathbf{u}^{(j+1)}(t,\mathbf{a}) = U(t)\mathbf{a} &+ \int_0^t U(t-s)\mathbf{G}\left(\mathbf{u}^{(j)}(s,\mathbf{a})\right)ds \\ &- \int_t^\infty V(t-s)\mathbf{G}\left(\mathbf{u}^{(j)}(s,\mathbf{a})\right)ds \qquad (*)\end{aligned}$$

Assume that the induction hypothesis

$$\left|\mathbf{u}^{(j)}(t,\mathbf{a}) - \mathbf{u}^{(j-1)}(t,\mathbf{a})\right| \le \frac{K|\mathbf{a}|e^{-\alpha t}}{2^{j-1}}$$

holds for $j = 1, 2, \ldots, m$ and $t \ge 0$. It clearly holds for $j = 1$ provided $t \ge 0$. Then using the Lipschitz-type condition (4) satisfied by the function $\mathbf{G}$ and the above estimates on $\|U(t)\|$ and $\|V(t)\|$, it follows from the induction hypothesis that for $t \ge 0$

$$\begin{aligned}\left|\mathbf{u}^{(m+1)}(t,\mathbf{a}) - \mathbf{u}^{(m)}(t,\mathbf{a})\right| &\le \int_0^t \|U(t-s)\|\varepsilon\left|\mathbf{u}^{(m)}(s,\mathbf{a}) - \mathbf{u}^{(m-1)}(s,\mathbf{a})\right|ds \\ &+ \int_t^\infty \|V(t-s)\|\varepsilon\left|\mathbf{u}^{(m)}(s,\mathbf{a}) - \mathbf{u}^{(m-1)}(s,\mathbf{a})\right|ds \\ &\le \varepsilon\int_0^t Ke^{-(\alpha+\sigma)(t-s)}\frac{K|\mathbf{a}|e^{-\alpha s}}{2^{m-1}}ds \\ &+ \varepsilon\int_t^\infty Ke^{\sigma(t-s)}\frac{K|\mathbf{a}|e^{-\alpha s}}{2^{m-1}}ds \\ &\le \frac{\varepsilon K^2|\mathbf{a}|e^{-\alpha t}}{\sigma 2^{m-1}} + \frac{\varepsilon K^2|\mathbf{a}|e^{-\alpha t}}{\sigma 2^{m-1}} \\ &< \left(\frac{1}{4} + \frac{1}{4}\right)\frac{K|\mathbf{a}|e^{-\alpha t}}{2^{m-1}} = \frac{K|\mathbf{a}|e^{-\alpha t}}{2^m}\end{aligned}$$

provided $\varepsilon K/\sigma < 1/4$; i.e., provided we choose $\varepsilon < \frac{\sigma}{4K}$. In order that the condition hold for the function $\mathbf{G}$, it suffices to choose $K|\mathbf{a}| < \delta/2$; i.e., we choose $|\mathbf{a}| < \frac{\delta}{2K}$. It then follows by induction that (9) holds $\forall\ j = 1, 2, 3, \ldots$ and $t \geq 0$. Thus, for $n > m > N$ and $t \geq 0$,

$$\begin{aligned} \left|\mathbf{u}^{(n)}(t,\mathbf{a}) - \mathbf{u}^{(m)}(t,\mathbf{a})\right| &\leq \sum_{j=N}^{\infty} \left|\mathbf{u}^{(j+1)}(t,\mathbf{a}) - \mathbf{u}^{(j)}(t,\mathbf{a})\right| \\ &\leq K|\mathbf{a}| \sum_{j=N}^{\infty} \frac{1}{2^j} = \frac{K|\mathbf{a}|}{2^{N-1}} \end{aligned}$$

This last quantity approaches zero as $N \to \infty$ and therefore $\left\{\mathbf{u}^{(j)}(t,\mathbf{a})\right\}$ is a Cauchy sequence of continuous functions. So, we now know that

$$\lim_{j\to\infty} \mathbf{u}^{(j)}(t,\mathbf{a}) = \mathbf{u}(t,\mathbf{a})$$

uniformly $\forall\ t \geq 0$ and $|\mathbf{a}| < \delta/2K$. Taking the limit of both sides of (*), it follows from the uniform convergence that the continuous function $\mathbf{u}(t,\mathbf{a})$ satisfies the integral equation and hence the differential equation. It follows by induction and the fact that $\mathbf{G} \in \mathcal{C}^1(\tilde{E})$ that $\mathbf{u}^{(j)}(t,\mathbf{a})$ is a differentiable function of a for $t \geq 0$ and $|\mathbf{a}| < \delta/2K$. Thus, it follows from the uniform convergence that $\mathbf{u}(t,\mathbf{a})$ is a differentiable function of a for $t \geq 0$ and $|\mathbf{a}| < \delta/2K$. The last estimate implies that

$$|\mathbf{u}(t,\mathbf{a})| \leq 2K|\mathbf{a}|e^{-\alpha t}$$

for $t \geq 0$ and $|\mathbf{a}| < \delta/2K$.

It is clear from the integral equation that the last $n-k$ components of the vector a do not enter the computation and hence they may be taken as zero. Thus, the components $u_j(t,\mathbf{a})$ of the solution $\mathbf{u}(t,\mathbf{a})$ satisfy the initial conditions

$$u_j(0,\mathbf{a}) = a_j \text{ for } j = 1,\ldots,k$$

and

$u_j(0,\mathbf{a}) = -\left(\int_0^\infty V(-s)\mathbf{G}\left(\mathbf{u}\left(s,a_1,\ldots,a_k,\mathbf{0}\right)\right)ds\right)_j$ for $j = k+1,\ldots,n$.

For $j = k+1,\ldots,n$ we define the functions

$$\psi_j\left(a_1,\ldots,a_k\right) = u_j\left(0,a_1,\ldots,a_k,0,\ldots,0\right)$$

Then the initial values $y_j = u_j\,(0, a_1, \ldots, a_k, 0, \ldots, 0)$ satisfy

$$y_j = \psi_j\,(y_1, \ldots, y_k) \text{ for } j = k+1, \ldots, n$$

according to the definition. These equations then define a differentiable manifold $\tilde{S}$ for $\sqrt{y_1^2 + \cdots + y_k^2} < \delta/2K$. Furthermore, if $\mathbf{y}(t)$ is a solution of the differential equation with $\mathbf{y}(0) \in \tilde{S}$, i.e., with $\mathbf{y}(0) = \mathbf{u}(0, \mathbf{a})$, then

$$\mathbf{y}(t) = \mathbf{u}(t, \mathbf{a})$$

It follows that if $\mathbf{y}(t)$ is a solution of (6) with $\mathbf{y}(0) \in \tilde{S}$, then $\mathbf{y}(t) \in \tilde{S}\ \forall\ t \geq 0$ and it follows from the estimate (11) that $\mathbf{y}(t) \to 0$ as $t \to \infty$. It can also be shown that if $\mathbf{y}(t)$ is a solution of (6) with $\mathbf{y}(0) \notin \tilde{S}$ then $\mathbf{y}(t) \nrightarrow 0$ as $t \to \infty$,

$$\frac{\partial \psi_j}{\partial y_i}(\mathbf{0}) = 0$$

for $i = 1, \ldots, k$ and $j = k+1, \ldots, n$; i.e., the differentiable manifold $\tilde{S}$ is tangent to the stable subspace $E^s = \{\mathbf{y} \in \mathbb{R}^n \mid y_1 = \cdots y_k = 0\}$ of the linear system $\dot{\mathbf{y}} = B\mathbf{y}$ at 0 .

The existence of the unstable manifold $\tilde{U}$ of (6) is established in exactly the same way by considering the differential system with $t \to -t$, i.e.,

$$\dot{\mathbf{y}} = -B\mathbf{y} - \mathbf{G}(\mathbf{y})$$

The stable manifold for this system will then be the unstable manifold $\tilde{U}$ for the differential system. Note that it is also necessary to replace the vector $\mathbf{y}$ by the vector $(y_{k+1}, \ldots, y_n, y_1, \ldots, y_k)$ in order to determine the $n-k$ dimensional manifold $\tilde{U}$ by the above process. This completes the proof of the Stable Manifold Theorem.

Definition: Let ϕ_t be the flow of the nonlinear system . The global stable and unstable manifolds of the nonlinear system of our concern at $\mathbf{0}$ are defined by $W^s(\mathbf{0}) = \bigcup_{t \leq 0} \phi_t(S)$ and $W^u(\mathbf{0}) = \bigcup_{t \geq 0} \phi_t(U)$ respectively; $W^s(\mathbf{0})$ and $W^u(\mathbf{0})$ are also referred to as the global stable and unstable manifolds of the origin respectively. It can be shown that the global stable and unstable manifolds $W^s(\mathbf{0})$ and $W^u(\mathbf{0})$ are unique and that they are invariant with respect to the flow ϕ_t; furthermore, $\forall$ $\mathbf{x} \in W^s(\mathbf{0}), \lim_{t \to \infty} \phi_t(\mathbf{x}) = \mathbf{0}$ and $\forall\ \mathbf{x} \in W^u(\mathbf{0}), \lim_{t \to -\infty} \phi_t(\mathbf{x}) = \mathbf{0}$.

It can be shown that in a small neighborhood, N, of a hyperbolic critical point at the origin, the local stable and unstable manifolds, S and U, of the concerned system at the origin are given by

$$S = \{\mathbf{x} \in N \mid \phi_t(\mathbf{x}) \to \mathbf{0} \text{ as } t \to \infty \text{ and } \phi_t(\mathbf{x}) \in N \text{ for } t \geq 0\}$$

and

$$U = \{\mathbf{x} \in N \mid \phi_t(\mathbf{x}) \to \mathbf{0} \text{ as } t \to -\infty \text{ and } \phi_t(\mathbf{x}) \in N \text{ for } t \leq 0\}$$

respectively.

It follows from the upper bound of $\mathbf{u}(t, \mathbf{a})$ in the proof of the stable manifold theorem that if $\mathbf{x}(t)$ is a solution of the differential equation (6) with $\mathbf{x}(0) \in S$, i.e., if $\mathbf{x}(t) = C\mathbf{y}(t)$ with $\mathbf{y}(0) = \mathbf{u}(0, \mathbf{a}) \in \tilde{S}$, then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that if $|\mathbf{x}(0)| < \delta$ then

$$|\mathbf{x}(t)| \leq \varepsilon e^{-\alpha t}$$

$\forall\ t \geq 0$. Just as in the proof of the stable manifold theorem, α is any positive number that satisfies $\mathrm{Re}\,(\lambda_j) < -\alpha$ for $j = 1, \ldots, k$ where λ_j, $j = 1, \ldots, k$ are the eigenvalues of $D\mathbf{f}(\mathbf{0})$ with negative real part. This result shows that solutions starting in S, sufficiently near the origin, approach the origin exponentially fast as $t \to \infty$.

Corollary: Under the hypotheses of the Stable Manifold Theorem, if S and U are the stable and unstable manifolds of the system at the origin and if $\mathrm{Re}\,(\lambda_j) < -\alpha < 0 < \beta < \mathrm{Re}\,(\lambda_m)$ for $j = 1, \ldots, k$ and $m = k+1, \ldots, n$, then given $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\mathbf{x}_0 \in N_\delta(\mathbf{0}) \cap S$ then $|\phi_t(\mathbf{x}_0)| \leq \varepsilon e^{-\alpha t}$ $\forall\ t \geq 0$ and if $\mathbf{x}_0 \in N_\delta(\mathbf{0}) \cap U$ then $|\phi_t(\mathbf{x}_0)| \leq \varepsilon e^{\beta t}$ $\forall\ t \leq 0$.

2.8 The Hartman-Grobman Theorem

The Hartman-Grobman Theorem is another very important result in the local qualitative theory of ordinary differential equations. The theorem shows that near a hyperbolic equilibrium point $\mathbf{x_0}$, the nonlinear system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

has the same qualitative structure as the linear system

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

with $A = D\mathbf{f}(\mathbf{x}_0)$. Throughout this section we shall assume that the equilibrium point $\mathbf{x}_0$ has been translated to the origin.

Definition: Two autonomous systems of differential equations are said to be topologically equivalent in a neighborhood of the origin or to have the same qualitative structure near the origin if there is a homeomorphism H mapping an open set U containing the origin onto an open set V containing the origin which maps trajectories of one in U onto trajectories of the other in V and preserves their orientation by time in the sense that if a trajectory is directed from $\mathbf{x}_1$ to $\mathbf{x}_2$ in U, then its image is directed from $H(\mathbf{x}_1)$ to $H(\mathbf{x}_2)$ in V. If the homeomorphism H preserves the parameterization by time, then the systems are said to be topologically conjugate in a neighborhood of the origin.

The Hartman-Grobman Theorem: Let E be an open subset of $\mathbb{R}^n$ containing the origin, let $\mathbf{f} \in \mathcal{C}^1(E)$, and let ϕ_t be the flow of the nonlinear system. Suppose that $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ and that the matrix $A = \mathrm{Df}(\mathbf{0})$ has no eigenvalue with zero real part. Then there exists a homeomorphism H of an open set U containing the origin onto an open set V containing
the origin such that for each $\mathbf{x}_0 \in U$, there is an open interval $I_0 \subset \mathbb{R}$ containing zero such that $\forall\ \mathbf{x}_0 \in U$ and $t \in I_0$

$$H \circ \phi_t(\mathbf{x}_0) = e^{At} H(\mathbf{x}_0);$$

i.e., H maps trajectories of the non-linear system near the origin onto trajectories of the linear system near the origin and preserves the parameterization by time.

Outline of the Proof: Consider the nonlinear system with $\mathbf{f} \in \mathcal{C}^1(E)$, $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ and $A = D\mathbf{f}(\mathbf{0})$.

1. Suppose that the matrix A is written in the form

$$A = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

where the eigenvalues of P have negative real part and the eigenvalues of Q have positive real part.
2. Let ϕ_t be the flow of the nonlinear system and write the solution

$$\mathbf{x}(t, \mathbf{x}_0) = \phi_t(\mathbf{x}_0) = \begin{bmatrix} \mathbf{y}(t, \mathbf{y}_0, \mathbf{z}_0) \\ \mathbf{z}(t, \mathbf{y}_0, \mathbf{z}_0) \end{bmatrix}$$

where

$$\mathbf{x}_0 = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{z}_0 \end{bmatrix} \in \mathbb{R}^n$$

$\mathbf{y}_0 \in E^s$, the stable subspace of A and $\mathbf{z}_0 \in E^u$, the unstable subspace of A.
3. Define the functions

$$\tilde{\mathbf{Y}}(\mathbf{y}_0, \mathbf{z}_0) = \mathbf{y}(1, \mathbf{y}_0, \mathbf{z}_0) - e^P \mathbf{y}_0$$

and

$$\tilde{\mathbf{z}}(\mathbf{y}_0, \mathbf{z}_0) = \mathbf{z}(1, \mathbf{y}_0, \mathbf{z}_0) - e^Q \mathbf{z}_0$$

Then $\tilde{\mathbf{Y}}(\mathbf{0}) = \tilde{\mathbf{Z}}(\mathbf{0}) = D\tilde{\mathbf{Y}}(\mathbf{0}) = D\tilde{\mathbf{Z}}(\mathbf{0}) = 0$. And since $\mathbf{f} \in \mathcal{C}^1(E)$, $\tilde{\mathbf{Y}}(\mathbf{y}_0, \mathbf{z}_0)$ and $\tilde{\mathbf{Z}}(\mathbf{y}_0, \mathbf{z}_0)$ are continuously differentiable. Thus,

$$\left\| D\tilde{\mathbf{Y}}(\mathbf{y}_0, \mathbf{z}_0) \right\| \leq a$$

and

$$\left\| D\tilde{\mathbf{Z}}(\mathbf{y}_0, \mathbf{z}_0) \right\| \leq a$$

on the compact set $|\mathbf{y}_0|^2 + |\mathbf{z}_0|^2 \leq s_0^2$. The constant a can be taken as small as we like by choosing s_0 sufficiently small. We let $\mathbf{Y}(\mathbf{y}_0, \mathbf{z}_0)$ and $\mathbf{Z}(\mathbf{y}_0, \mathbf{z}_0)$ be smooth functions which are equal to $\tilde{\mathbf{Y}}(\mathbf{y}_0, \mathbf{z}_0)$ and $\tilde{\mathbf{Z}}(\mathbf{y}_0, \mathbf{z}_0)$ for $|\mathbf{y}_0|^2 + |z_0|^2 \leq (s_0/2)^2$ and zero for $|\mathbf{y}_0|^2 + |\mathbf{z}_0|^2 \geq s_0^2$. Then by the mean value theorem

$$|\mathbf{Y}(\mathbf{y}_0, \mathbf{z}_0)| \leq a\sqrt{|\mathbf{y}_0|^2 + |\mathbf{z}_0|^2} \leq a(|\mathbf{y}_0| + |\mathbf{z}_0|)$$

and

$$|\mathbf{Z}(\mathbf{y}_0, \mathbf{z}_0)| \leq a\sqrt{|\mathbf{y}_0|^2 + |\mathbf{z}_0|^2} \leq a(|\mathbf{y}_0| + |\mathbf{z}_0|)$$

$\forall$ $(\mathbf{y}_0, \mathbf{z}_0) \in \mathbb{R}^n$. We next let $B = e^P$ and $C = e^Q$. Then assuming that we have carried out the normalization, we have

$$b = \|B\| < 1 \quad \text{and} \quad c = \left\|C^{-1}\right\| < 1$$

4. For

$$\mathbf{x} = \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} \in \mathbb{R}^n$$

define the transformations

$$L(\mathbf{y}, \mathbf{z}) = \begin{bmatrix} B\mathbf{y} \\ C\mathbf{z} \end{bmatrix}$$

and

$$T(\mathbf{y}, \mathbf{z}) = \begin{bmatrix} B\mathbf{y} + \mathbf{Y}(\mathbf{y}, \mathbf{z}) \\ C\mathbf{z} + \mathbf{Z}(\mathbf{y}, \mathbf{z}) \end{bmatrix}$$

i.e. $L(\mathbf{x}) = e^A\mathbf{x}$ and locally $T(\mathbf{x}) = \phi_1(\mathbf{x})$.

Lemma: There exists a homeomorphism H of an open set U containing the origin onto an open set V containing the origin such that

$$H \circ T = L \circ H.$$

Proof: We establish this lemma using the method of successive approximations. For $\mathbf{x} \in \mathbb{R}^n$, let

$$H(\mathbf{x}) = \begin{bmatrix} \Phi(\mathbf{y}, \mathbf{z}) \\ \Psi(\mathbf{y}, \mathbf{z}) \end{bmatrix}$$

Then $H \circ T = L \circ H$ is equivalent to the pair of equations

$$\begin{aligned} B\Phi(\mathbf{y},\mathbf{z}) &= \Phi(B\mathbf{y}+\mathbf{Y}(\mathbf{y},\mathbf{z}),C\mathbf{z}+\mathbf{Z}(\mathbf{y},\mathbf{z})) \\ C\Psi(\mathbf{y},\mathbf{z}) &= \Psi(B\mathbf{y}+\mathbf{Y}(\mathbf{y},\mathbf{z}),C\mathbf{z}+\mathbf{Z}(\mathbf{y},\mathbf{z})) \end{aligned}$$

First of all, define the successive approximations for the second equation by

$$\Psi_0(\mathbf{y},\mathbf{z}) = \mathbf{z}$$

$$\Psi_{k+1}(\mathbf{y},\mathbf{z}) = C^{-\mathbf{1}}\Psi_k(B\mathbf{y}+\mathbf{Y}(\mathbf{y},\mathbf{z}),C\mathbf{z}+\mathbf{Z}(\mathbf{y},\mathbf{z}))$$

It then follows by an easy induction argument that for $k=0,1,2,\ldots$, the $\Psi_k(\mathbf{y},\mathbf{z})$ are continuous and satisfy $\Psi_k(\mathbf{y},\mathbf{z})=\mathbf{z}$ for $|\mathbf{y}|+|\mathbf{z}|\geq 2s_0$. We next prove by induction that for $j=1,2,\ldots$

$$|\Psi_j(\mathbf{y},\mathbf{z})-\Psi_{j-1}(\mathbf{y},\mathbf{z})| \leq Mr^j(|\mathbf{y}|+|\mathbf{z}|)^\delta$$

where $r=c[2\max(a,b,c)]^\delta$ with $\delta\in(0,1)$ chosen sufficiently small so that $r<1$ (which is possible since $c<1$) and $M=ac\left(2s_0\right)^{1-\delta}/r$. First of all for $j=1$

$$\begin{aligned} |\Psi_1(\mathbf{y},\mathbf{z})-\Psi_0(\mathbf{y},\mathbf{z})| &= C^{-1}\Psi_0(B\mathbf{y}+\mathbf{Y}(\mathbf{y},\mathbf{z}),C\mathbf{z}+\mathbf{Z}(\mathbf{y},\mathbf{z}))-\mathbf{z}\ | \\ &= \left|C^{-1}(C\mathbf{z}+\mathbf{Z}(\mathbf{y},\mathbf{z}))-\mathbf{z}\right| \\ &= \left|C^{-1}\mathbf{Z}(\mathbf{y},\mathbf{z})\right| \leq \left\|C^{-1}\right\| |\mathbf{Z}(\mathbf{y},\mathbf{z})| \\ &\leq ca(|\mathbf{y}|+|\mathbf{z}|) \leq Mr(|\mathbf{y}|+|\mathbf{z}|)^\delta \end{aligned}$$

since $\mathbf{Z}(\mathbf{y},\mathbf{z})=0$ for $|\mathbf{y}|+|\mathbf{z}|\geq 2s_0$. And then assuming that the induction hypothesis holds for $j=1,\ldots,k$ we have

$$\begin{aligned} |\Psi_{k+1}(\mathbf{y},\mathbf{z})-\Psi_k(\mathbf{y},\mathbf{z})| &= |\ C^{-1}\Psi_k(B\mathbf{y}+\mathbf{Y}(\mathbf{y},\mathbf{z}),C\mathbf{z}+\mathbf{Z}(\mathbf{y},\mathbf{z})) \\ &\quad -C^{-1}\Psi_{k-1}(B\mathbf{y}+\mathbf{Y}(\mathbf{y},\mathbf{z}),C\mathbf{z}+\mathbf{Z}(\mathbf{y},\mathbf{z}))\ | \\ &\leq \left\|C^{-1}\right\| |\Psi_k('')-\Psi_{k-1}('')| \\ &\leq cMr^k[|B\mathbf{y}+\mathbf{Y}(\mathbf{y},\mathbf{z})|+|C\mathbf{z}+\mathbf{Z}(\mathbf{y},\mathbf{z})|]^\delta \\ &\leq cMr^k[b|\mathbf{y}|+2a(|\mathbf{y}|+|\mathbf{z}|)+c|\mathbf{z}|]^\delta \\ &\leq cMr^k[2\max(a,b,c)]^\delta(|\mathbf{y}|+|\mathbf{z}|)^\delta \\ &= Mr^{k+1}(|\mathbf{y}|+|\mathbf{z}|)^\delta \end{aligned}$$

Thus, just as in the proof of the fundamental theorem of non-linear systems and the stable manifold theorem, $\Psi_k(\mathbf{y}, \mathbf{z})$ is a Cauchy sequence of continuous functions which converges uniformly as $k \to \infty$ to a continuous function $\Psi(\mathbf{y}, \mathbf{z})$. Also, $\Psi(\mathbf{y}, \mathbf{z}) = \mathbf{z}$ for $|\mathbf{y}|+|\mathbf{z}| \geq 2s_0$. Taking limits in (4) shows that $\Psi(\mathbf{y}, \mathbf{z})$ is a solution of the second equation.

The first equation in the proof above can be written as

$$B^{-1}\Phi(\mathbf{y}, \mathbf{z}) = \mathbf{\Phi}\left(B^{-1}\mathbf{y} + \mathbf{Y}_1(\mathbf{y}, \mathbf{z}), C^{-1}\mathbf{z} + \mathbf{Z}_1(\mathbf{y}, \mathbf{z})\right)$$

where the functions $\mathbf{Y}_1$ and $\mathbf{Z}_1$ are defined by the inverse of T (which exists if the constant a is sufficiently small, i.e., if s_0 is sufficiently small) as follows:

$$T^{-1}(\mathbf{y}, \mathbf{z}) = \begin{bmatrix} B^{-1}\mathbf{y} + \mathbf{Y}_1(\mathbf{y}, \mathbf{z}) \\ C^{-1}\mathbf{z} + \mathbf{Z}_1(\mathbf{y}, \mathbf{z}) \end{bmatrix}$$

Then equation can be solved for $\Phi(\mathbf{y}, \mathbf{z})$ by the method of successive approximations exactly as above with $\Phi_0(\mathbf{y}, \mathbf{z}) = \mathbf{y}$ since $b = \|B\| < 1$. We therefore obtain the continuous map

$$H(\mathbf{y}, \mathbf{z}) = \begin{bmatrix} \Phi(\mathbf{y}, \mathbf{z}) \\ \Psi(\mathbf{y}, \mathbf{z}) \end{bmatrix}.$$

And it follows that H is a homeomorphism of $\mathbb{R}^n$ onto $\mathbb{R}^n$.
5. We now let H_0 be the homeomorphism defined above and let L^t and T^t be the one-parameter families of transformations defined by

$$L^t\left(\mathbf{x}_0\right) = e^{At}\mathbf{x}_0 \quad \text{and} \quad T^t\left(\mathbf{x}_0\right) = \phi_t\left(\mathbf{x}_0\right).$$

Define

$$H = \int_0^1 L^{-s} H_0 T^s ds$$

It then follows using the above lemma that there exists a neighborhood of the origin for which

$$
\begin{aligned}
L^t H &= \int_0^1 L^{t-s} H_0 T^{s-t} ds T^t \\
&= \int_{-t}^{1-t} L^{-s} H_0 T^s ds T^t \\
&= \left[\int_{-t}^{0} L^{-s} H_0 T^s ds + \int_0^{1-t} L^{-s} H_0 T^s ds \right] T^t \\
&= \int_0^1 L^{-s} H_0 T^s ds T^t = H T^t
\end{aligned}
$$

since by the above lemma $H_0 = L^{-1} H_0 T$ which implies that

$$
\begin{aligned}
\int_{-t}^{0} L^{-s} H_0 T^s ds &= \int_{-t}^{0} L^{-s-1} H_0 T^{s+1} ds \\
&= \int_{1-t}^{1} L^{-s} H_0 T^s ds
\end{aligned}
$$

Thus, $H \circ T^t = L^t H$ or equivalently

$$
H \circ \phi_t (\mathbf{x}_0) = e^{At} H (\mathbf{x}_0)
$$

and it can be shown that H is a homeomorphism on $\mathbb{R}^n$. This completes the outline of the proof of the *Hartman Grobman Theorem*.

Theorem: Let E be an open subset of $\mathbb{R}^n$ containing the point $\mathbf{x}_0$, let $\mathbf{f} \in C^2(E)$, and let $\boldsymbol{\phi}_t$ be the flow of the nonlinear system. Suppose that $\mathbf{f}(\mathbf{x}_0) = \mathbf{0}$ and that all of the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the matrix $A = D\mathbf{f}(\mathbf{x}_0)$ have negative (or positive) real part. Then there exists a $\mathcal{C}^1$-diffeomorphism H of a neighborhood U of $\mathbf{x}_0$ onto an open set V containing the origin such that for each $\mathbf{x} \in U$ there is an open interval $I(\mathbf{x}) \subset \mathbb{R}$ containing zero such that $\forall$ $\mathbf{x} \in U$ and $t \in I(\mathbf{x})$

$$
H \circ \phi_t(\mathbf{x}) = e^{At} H(\mathbf{x})
$$

2.9 Saddles, Nodes, Foci and Centers

In this section, we are not going to explore intricate theoretical details, but just a quick review of the *topological* definition of these key terms.

Center: The origin is called a *center* for the non linear system if there exists a $\delta > 0$ such that every solution curve of the non linear system in the deleted neighborhood $N_\delta(\mathbf{0}) \setminus \{\mathbf{0}\}$ is a *closed curve* with $\mathbf{0}$ in its *interior*.

Center-focus: The origim is known as a *center-focus* for the non-linear system if there exists a sequence of closed curves Γ_n, with Γ_{n+1} in the interior of Γ_n such that $\Gamma_n \to \mathbf{0}$ as $n \to \infty$ and such that every trajectory between Γ_n and Γ_{n+1} spirals towards Γ_n or Γ_{n+1} as $t \to \pm\infty$.

Stable focus: The origin is known as a *stable focus* for the non-linear system if there exist a $\delta > 0$ such that for $0 < r_0 < \delta$ and $\theta_0 \in \mathbb{R}$, $r(t, r_0, \theta_0) \to 0$ and $|\theta(t, r_0, \theta_0)| \to \infty$ as $t \to \infty$.

Unstable focus: The origin is known as a *unstable focus* for the non-linear system if there exist a $\delta > 0$ such that for $0 < r_0 < \delta$ and $\theta_0 \in \mathbb{R}$, $r(t, r_0, \theta_0) \to 0$ and $|\theta(t, r_0, \theta_0)| \to \infty$ as $t \to -\infty$.

Stable node: The origin is known as a *stable node* for the non linear system if there exists a $\delta > 0$ such that for $0 < r_0 < \delta$ and $\theta_0 \in \mathbb{R}$, $r(t, r_0, \theta_0) \to 0$ as $t \to \infty$ and $\lim_{t\to\infty} \theta(t, r_0, \theta_0)$ exists.

Unstable node: The origin is known as a *unstable node* for the non linear system if there exists a $\delta > 0$ such that for $0 < r_0 < \delta$ and $\theta_0 \in \mathbb{R}$, $r(t, r_0, \theta_0) \to 0$ as $t \to -\infty$ and $\lim_{t\to-\infty} \theta(t, r_0, \theta_0)$ exists.

Proper Node: The origin is known as a *proper node* if if it is a *node* and every ray through the origin is tangent to some trajectory of the non-linear system.

Topological saddle: The origin is a *topological saddle* for a non-linear system if there exists two trajectories Γ_1 and Γ_2 which approach $\mathbf{0}$ as $t \to \infty$ and two trajectories Γ_3 and Γ_4 which approaches $\mathbf{0}$ as $t \to -\infty$ and if there exists a $\delta > 0$ such that all other trajectories which start in the deleted neighborhood of the *origin* leave the δ-neighborhood as $t \to \pm\infty$. The trajectories $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ are known as *separatrices*.

3 Nonlinear Systems: Global Theory

We have seen that any nonlinear system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

with $\mathbf{f} \in \mathcal{C}^1(E)$ and E an open subset of $\mathbb{R}^n$, has a unique solution $\phi_t(\mathbf{x}_0)$, passing through

a point $x_0 \in E$ at time $t = 0$ which is defined for all $t \in I(\mathbf{x}_0)$, the maximal interval of existence of the solution. Furthermore, the flow ϕ_t of the system satisfies (i) $\phi_0(\mathbf{x}) = \mathbf{x}$ and (ii) $\phi_{t+s}(\mathbf{x}) = \phi_t(\phi_s(\mathbf{x}))$. for all $\mathbf{x} \in E$ and the function $\phi(t, \mathbf{x}) = \phi_t(\mathbf{x})$ defines a $\mathcal{C}^1$-map $\phi : \Omega \to E$ where $\Omega = \{(t, \mathbf{x}) \in \mathbb{R} \times E \mid t \in I(\mathbf{x})\}$.
In this chapter we define a dynamical system as a $\mathcal{C}^1$-map $\phi : \mathbb{R} \times E \to E$ which satisfies (i) and (ii) above. We first show that we can rescale the time in any $\mathcal{C}^1$-system (eg: nonlinear system) so that for all $\mathbf{x} \in E$, the maximal interval of existence $I(\mathbf{x}) = (-\infty, \infty)$. Thus any $\mathcal{C}^1$-system (eg: nonlinear system), after an appropriate rescaling of the time, defines a dynamical system $\boldsymbol{\phi} : \mathbb{R} \times E \to E$ where $\phi(t, \mathbf{x}) = \phi_t(\mathbf{x})$ is the solution of the mentioned nonlinear system with $\phi_0(\mathbf{x}) = \mathbf{x}$. We next consider limit sets and attractors of dynamical systems. Besides equilibrium points and periodic orbits, a dynamical system can have homoclinic loops or separatrix cycles as well as strange attractors as limit sets. We study periodic orbits in some detail and give the Stable Manifold Theorem for periodic orbits as well as several examples which illustrate the general theory in this chapter. Determining the nature of limit sets of nonlinear systems with $n \geq 3$ is a challenging problem which is the subject of much mathematical research at this time.

3.1 Dynamical Systems and Global Existence Theorems

A dynamical system gives a functional description of the solution of a physical problem or of the mathematical model describing the physical problem. For example, the motion of the undamped pendulum is a dynamical system in the sense that the motion of the pendulum is described by its position and velocity as functions of time and the initial conditions.

Mathematically speaking, a dynamical system is a function $\phi(t, \mathbf{x})$, defined for all $t \in \mathbb{R}$ and $\mathbf{x} \in E \subset \mathbb{R}^n$, which describes how points $\mathbf{x} \in E$ move with respect to time. We require that the family of maps $\phi_t(\mathbf{x}) = \phi(t, \mathbf{x})$ have the properties of a flow have already been defined.

Definition: A dynamical system on E is a $\mathcal{C}^1$-map

$$\phi : \mathbb{R} \times E \to E$$

where E is an open subset of $\mathbb{R}^{\boldsymbol{n}}$ and if $\phi_t(\mathbf{x}) = \boldsymbol{\phi}(t, \mathbf{x})$, then ϕ_t satisfies

(i) $\phi_0(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in E$, and,

(ii) $\phi_t \circ \phi_s(\mathbf{x}) = \phi_{t+s}(\mathbf{x})$ for all $s, t \in \mathbb{R}$ and $\mathbf{x} \in E$.

Remark: It follows from definition that for each $t \in \mathbb{R}, \phi_t$ is a $\mathcal{C}^1$ map of E into E which has a $\mathcal{C}^1$-inverse, ϕ_{-t}; i.e., ϕ_t with $t \in \mathbb{R}$ is a one-parameter family of diffeomorphisms on E that forms a commutative group under composition.

It is easy to see that if A is an $n \times n$ matrix then the function $\phi(t, \mathbf{x}) = e^{At}\mathbf{x}$ defines a dynamical system on $\mathbb{R}^n$ and also, for each $\mathbf{x}_0 \in \mathbb{R}^n$, $\boldsymbol{\phi}(t, \mathbf{x}_0)$ is the solution of the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} \\ \mathbf{x}(0) &= \mathbf{x}_0. \end{aligned}$$

In general, if $\boldsymbol{\phi}(t, \mathbf{x})$ is a dynamical system on $E \subset \mathbb{R}^{\boldsymbol{n}}$, then the function

$$\mathbf{f}(\mathbf{x}) = \left.\frac{d}{dt}\boldsymbol{\phi}(t, \mathbf{x})\right|_{t=0}$$

defines a $\mathcal{C}^1$-vector field on E and for each $\mathbf{x}_0 \in E, \phi(t, \mathbf{x}_0)$ is the solution of the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\ \mathbf{x}(0) &= \mathbf{x}_0. \end{aligned}$$

Furthermore, for each $\mathbf{x}_0 \in E$, the maximal interval of existence of $\phi(t, \mathbf{x}_0)$, $I(\mathrm{x}_0) = (-\infty, \infty)$. Thus, each dynamical system gives rise to a $\mathcal{C}^1$-vector field $\mathbf{f}$ and the dynamical system describes the solution set of the differential equation defined by this vector field. Conversely, given a differential equation with $\mathbf{f} \in \mathcal{C}^1(E)$ and E an open subset of $\mathbb{R}^n$, the solution $\phi(t, \mathbf{x}_0)$ of the initial value problem with $\mathbf{x}_0 \in E$ will be a dynamical system on E if and only if for all $\mathbf{x}_0 \in E, \phi(t, \mathbf{x}_0)$ is defined for all $t \in \mathbb{R}$; i.e., if and only if for all $\mathbf{x}_0 \in E$, the maximal interval of existence $I(\mathbf{x}_0)$ of $\phi(t, \mathbf{x}_0)$ is $(-\infty, \infty)$. In this case we say that $\phi(t, \mathbf{x}_0)$ is the dynamical system on E defined by the mentioned initial value problem.

The next theorem shows that any $\mathcal{C}^1$-vector field $\mathbf{f}$ defined on all of $\mathbb{R}^n$ leads to a dynamical system on $\mathbb{R}^n$. While the solutions $\phi(t, \mathbf{x}_0)$ of the original system may not be defined for all $t \in \mathbb{R}$, the time t can be rescaled along trajectories of the original system to obtain a topologically equivalent system for which the solutions are defined for all $t \in \mathbb{R}$.

Before stating this theorem, we generalize the notion of topological equivalent systems for a neighborhood of the origin.

Definition: Suppose that $\mathbf{f} \in \mathcal{C}^1(E_1)$ and $\mathbf{g} \in \mathcal{C}^1(E_2)$ where E_1 and E_2 are open subsets of $\mathbb{R}^n$. Then the two autonomous systems of differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

and

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$$

are said to be topologically equivalent if there is a homeomorphism $H : E_1 \to E_2$ which maps trajectories of the first differential equation onto trajectories of the second one and preserves their orientation by time. In this case, the vector fields $\mathbf{f}$ and $\mathbf{g}$ are also said to be topologically equivalent. If $E = E_1 = E_2$ then the two systems are said to be topologically equivalent on E and the vector fields $\mathbf{f}$ and $\mathbf{g}$ are said to be topologically equivalent on E.

Global Existence Theorem: For $\mathbf{f} \in \mathcal{C}^1(\mathbb{R}^{\boldsymbol{n}})$ and for each $\mathbf{x}_0 \in \mathbb{R}^{\boldsymbol{n}}$, the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \frac{\mathbf{f}(\mathbf{x})}{1 + |\mathbf{f}(\mathbf{x})|} \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned}$$

has a unique solution $\mathbf{x}(t)$ defined for all $t \in \mathbb{R}$, i.e., (3) defines a dynamical system on $\mathbb{R}^{\boldsymbol{n}}$; furthermore, (3) is topologically equivalent to (1) on $\mathbb{R}^{\boldsymbol{n}}$.

Remark: The original system and the modified one in the theorem are topologically equivalent on $\mathbb{R}^{\boldsymbol{n}}$ since the time t along the solutions $\mathbf{x}(t)$ of (1) has simply been rescaled according to the formula

$$\tau = \int_0^t [1 + |\mathbf{f}(\mathrm{x}(s))|]\, ds$$

i.e., the homeomorphism H is simply the identity on $\mathbb{R}^n$. The solution $\mathbf{x}(t)$ of (1), with respect to the new time τ, then satisfies

$$\frac{d\mathbf{x}}{d\tau} = \frac{d\mathbf{x}}{dt} \Big/ \frac{d\tau}{dt} = \frac{\mathbf{f}(\mathbf{x})}{1 + |\mathbf{f}(\mathbf{x})|}$$

i.e., $\mathbf{x}(t(\tau))$ is the solution of the modified system where $t(\tau)$ is the inverse of the strictly increasing function $\tau(t)$ defined by the rescalation. The function $\tau(t)$ maps the maximal interval of existence (α, β) of the solution $\mathbf{x}(t)$ of the original system one-to-one and onto $(-\infty, \infty)$, the maximal interval of existence of the modified system.

Proof: It is not difficult to show that if $f \in \mathcal{C}^1(\mathbb{R}^n)$ then the function

$$\frac{\mathbf{f}}{1+|\mathbf{f}|} \in \mathcal{C}^1(\mathbb{R}^n)$$

For $\mathbf{x}_0 \in \mathbb{R}^{\boldsymbol{n}}$, let $\mathbf{x}(t)$ be the solution of the modified initial value problem on its maximal interval of existence $(\ \alpha, \beta\ )$. So, $\mathbf{x}(t)$ satisfies the integral equation **(Verify!!)**

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_0^t \frac{\mathbf{f}(\mathbf{x}(s))}{1+|\mathbf{f}(\mathbf{x}(s))|} ds$$

for all $t \in (\alpha, \beta)$ and since $|\mathbf{f}(\mathbf{x})|/(1+|\mathbf{f}(\mathbf{x})|) \leq 1$, it follows that

$$|\mathbf{x}(t)| \leq |\mathbf{x}_0| + \int_0^{|t|} ds = |\mathrm{x}_0| + |t|$$

for all $t \in (\alpha, \beta)$. Suppose that $\beta < \infty$. Then

$$|\mathbf{x}(t)| \leq |\mathbf{x}_0| + \beta$$

for all $t \in [0, \beta)$; i.e., for all $t \in [0, \beta)$, the solution of the modified system through the point x_0 at time $t = 0$ is contained in the compact set

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x}| \leq |\mathbf{x}_0| + \beta\} \subset \mathbb{R}^n.$$

But then, we know that $\beta = \infty$, a contradiction. Therefore, $\beta = \infty$. A similar proof shows that $\alpha = -\infty$. Thus, for all $\mathbf{x}_0 \in \mathbb{R}^{\boldsymbol{n}}$, the maximal interval of existence of the solution $\mathbf{x}(t)$ of the modified initial value problem, $(\alpha, \beta) = (-\infty, \infty)$.

An Interesting Example to Note: For $x_0 > 0$ the initial value problem

$$\begin{aligned} \dot{x} &= \frac{1}{2x} \\ x(0) &= x_0 \end{aligned}$$

has the unique solution $x(t) = \sqrt{t + x_0^2}$ defined on its maximal interval of existence $I(x_0) = (-x_0^2, \infty)$. The function $f(x) = 1/(2x) \in \mathcal{C}^1(E)$ where $E = (0, \infty)$. We have $x(t) \to 0 \in \dot{E}$ as $t \to -x_0^2$. The related initial value problem

$$\begin{aligned} \dot{x} &= \frac{1/2x}{1 + (1/2x)} = \frac{1}{2x+1} \\ x(0) &= x_0 \end{aligned}$$

has the unique solution

$$x(t) = -\frac{1}{2} + \sqrt{t + (x_0 + 1/2)^2}$$

defined on its maximal interval of existence $I(x_0) = \left(-(x_0 + 1/2)^2, \infty\right)$. We see that in this case $I(x_0) \neq \mathbb{R}$.

However, a slightly more subtle rescaling of the time along trajectories of the original initial value problem does lead to a dynamical system equivalent to the original one even when E is a proper subset of $\mathbb{R}^{\boldsymbol{n}}$. This idea is due to Vinograd.

Theorem: Suppose that E is an open subset of $\mathbb{R}^n$ and that $\mathbf{f} \in \mathcal{C}^1(E)$. Then there is a function $\mathbf{F} \in \mathcal{C}^l(E)$ such that

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$$

defines a dynamical system on E and such that the new dynamical system is topologically equivalent to the original one on E.

Proof: First of all, as in global existence theorem, the function

$$\mathbf{g}(\mathbf{x}) = \frac{\mathbf{f}(\mathbf{x})}{1 + |\mathbf{f}(\mathbf{x})|} \in \mathcal{C}^1(E)$$

$|\mathbf{g}(\mathbf{x})| \leq 1$, the original system and the modified one are topologically equivalent on E. Furthermore, solutions $\mathbf{x}(t)$ of the modified system satisfy

$$\int_0^t |\dot{\mathbf{x}}(t')|\, dt' = \int_0^t |\mathbf{g}(\mathbf{x}(t'))|\, dt' \leq |t|$$

i.e., for finite t, the trajectory defined by $\mathbf{x}(t)$ has finite arc length. Let (α, β) be the maximal interval of existence of $\mathbf{x}(t)$ and suppose that $\beta < \infty$. Then since the arc length of the half-trajectory defined by $\mathbf{x}(t)$ for $t \in (0, \beta)$ is finite, the half-trajectory defined by $\mathbf{x}(t)$ for $t \in [0, \beta)$ must have a limit point

$$\mathbf{x}_1 = \lim_{t \to \beta^-} \mathbf{x}(t) \in \dot{E}$$

Now define the closed set $K = \mathbb{R}^{\boldsymbol{n}} \setminus E$ and let

$$G(\mathbf{x}) = \frac{d(\mathbf{x}, K)}{1 + d(\mathbf{x}, K)}$$

where $d(\mathbf{x}, \mathbf{y})$ denotes the distance between $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^{\boldsymbol{n}}$ and

$$d(\mathbf{x}, K) = \inf_{y \in K} d(\mathbf{x}, \mathbf{y})$$

i.e., for $\mathbf{x} \in E, d(\mathbf{x}, K)$ is the distance of $\mathbf{x}$ from the boundary ∂E of E. Then the function $G \in \mathcal{C}^l\left(\mathbb{R}^n\right), 0 \leq G(\mathbf{x}) \leq 1$ and for $\mathbf{x} \in K, G(\mathbf{x}) = 0$. Let $\mathbf{F}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) G(\mathbf{x})$. Then $\mathbf{F} \in \mathcal{C}^1(E)$ and the system, $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, is topologically equivalent to our initial modification on E since we have simply rescaled the time along trajectories of that initially modified system; i.e., the homeomorphism H is simply the identity on E. Furthermore, the system $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ has a bounded right-hand side and therefore its trajectories have finite arc-length for finite t. To prove that the modification by Vinograd defines a dynamical system on E, it suffices to show that all half-trajectories of the aforementioned modification which (a) start in E, (b) have finite arc length s_0, and (c) terminate at a limit point $\mathbf{x}_1 \in \dot{E}$ are defined for all $t \in [0, \infty)$. Along any solution $\mathbf{x}(t)$ of that modification, $\frac{ds}{dt} = |\dot{x}(t)|$ and hence

$$t = \int_0^s \frac{ds'}{\left|\mathbf{F}\left(\mathbf{x}\left(t\left(s'\right)\right)\right)\right|}$$

where $t(s)$ is the inverse of the strictly increasing function $s(t)$ defined by

$$s = \int_0^t \left|\mathbf{F}\left(\mathbf{x}\left(t'\right)\right)\right| dt'$$

for $s > 0$. But for each point $\mathbf{x} = \mathbf{x}(t(s))$ on the half-trajectory we have

$$G(\mathbf{x}) = \frac{d(\mathbf{x}, K)}{1 + d(\mathbf{x}, K)} < d(\mathbf{x}, K) = \inf_{\mathbf{y} \in K} d(\mathbf{x}, \mathbf{y}) \leq d\left(\mathbf{x}, \mathbf{x}_1\right) \leq s_0 - s$$

And therefore since $0 < |g(x)| \leq 1$, we have

$$t \geq \int_0^s \frac{ds'}{s_0 - s'} = \log \frac{s_0 - s}{s_0}$$

and hence $t \to \infty$ as $s \to s_0$; i.e., the half-trajectory defined by $\mathbf{x}(t)$ is defined for all $t \in [0, \infty)$; i.e., $\beta = \infty$. Similarly, it can be shown that $\alpha = -\infty$ and hence, the modified system defines a dynamical system on E which is topologically equivalent to the unmodified original system on E.

For $\mathbf{f} \in \mathcal{C}^1(E), E$ an open subset of $\mathbb{R}^n$, the second theorem implies that there is no loss in generality in assuming that the original system defines a dynamical system $\phi\left(t, \mathbf{x}_0\right)$ on E. Throughout the remainder of this book we therefore make this assumption; i.e., we assume that for all $\mathbf{x}_0 \in E$, the maximal interval of existence $I\left(x_0\right) = (-\infty, \infty)$. In the next section, we then go on to discuss the limit sets of trajectories $\mathbf{x}(t)$ of the original as $t \to \pm\infty$. However, we first present two more global existence theorems which are of some interest.

Theorem: Suppose that $\mathbf{f} \in \mathcal{C}^1\left(\mathbb{R}^n\right)$ and that $\mathbf{f}(\mathbf{x})$ satisfies the global Lipschitz condition

$$|\mathbf{f}(\mathbf{x})| - \mathbf{f}(\mathbf{y})| \leq M|\mathbf{x} - \mathbf{y}|$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{\boldsymbol{n}}$. Then for $\mathbf{x}_0 \in \mathbb{R}^{\boldsymbol{n}}$, the initial value problem (1) has a unique solution $\mathbf{x}(t)$ defined for all $t \in \mathbb{R}$.

Proof: Let $\mathbf{x}(t)$ be the solution of the original initial value problem on its maximal interval of existence (α, β). Then using the fact that $d|\mathbf{x}(t)|/dt \leq |\dot{\mathbf{x}}(t)|$ and the triangle inequality,

$$\begin{aligned}
\frac{d}{dt}\left|\mathbf{x}(t) - \mathbf{x}_0\right| \leq |\dot{\mathbf{x}}(t)| &= |\mathbf{f}(\mathbf{x}(t))| \\
&\leq \left|\mathbf{f}(\mathbf{x}(t)) - \mathbf{f}\left(\mathbf{x}_0\right)\right| + \left|\mathbf{f}\left(\mathbf{x}_0\right)\right| \\
&\leq M\left|\mathbf{x}(t) - \mathbf{x}_0\right| + \left|\mathbf{f}\left(\mathbf{x}_0\right)\right|
\end{aligned}$$

Thus, if we assume that $\beta < \infty$, then the function $g(t) = |\mathbf{x}(t) - \mathrm{x}_0|$ satisfies

$$g(t) = \int_0^t \frac{dg(s)}{ds}\, ds \leq \left|\mathbf{f}\left(\mathbf{x}_0\right)\right| \beta + M \int_0^t g(s)\, ds$$

for all $t \in (0, \beta)$. It then follows from Gronwall's Lemma that

$$|\mathbf{x}(t) - \mathbf{x}_0| \leq \beta \left|\mathbf{f}\left(\mathbf{x}_0\right)\right| e^{M\beta}$$

for all $t \in [0, \beta)$; i.e., the trajectory of the original system through the point $\mathbf{x}_0$ at time $t = 0$ is contained in the compact set

$$K = \left\{\mathbf{x} \in \mathbb{R}^{\mathbf{n}} \middle| |\mathbf{x} - \mathbf{x}_0| \leq \beta |\mathbf{f}\left(\mathbf{x}_0\right) | \, e^{M\beta}\right\} \subset \mathbb{R}^n.$$

But then by one of the corollaries we have already proven, it follows that $\beta = \infty$, a contradiction. Therefore, $\beta = \infty$ and it can similarly be shown that $\alpha = -\infty$. Thus, for all $\mathbf{x}_0 \in \mathbb{R}^n$, the maximal interval of existence of the solution $\mathbf{x}(t)$ of the initial value problem, $I\left(\mathrm{x}_0\right) = (-\infty, \infty)$.

If $\mathbf{f} \in \mathcal{C}^1(M)$ where M is a compact subset of $\mathbb{R}^{\boldsymbol{n}}$, then $\mathbf{f}$ satisfies a global Lipschitz condition on M and we have a result similar to the above theorem for $\mathbf{x}_0 \in M$. This result has been extended to compact manifolds by Chillingworth.

Theorem: Let M be a compact manifold and let $\mathbf{f} \in C^{\mathbf{1}}(M)$. Then for $\mathbf{x}_0 \in M$, the initial value problem has a unique solution $\mathbf{x}(t)$ defined for all $t \in \mathbb{R}$.

3.2 Limit Sets

Consider the autonomous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

with $\mathbf{f} \in \mathcal{C}^1(E)$ where E is an open subset of $\mathbb{R}^{\boldsymbol{n}}$. In the previous section, we saw that there is no loss in generality in assuming that the nonlinear system defines a dynamical system $\phi(t, \mathbf{x})$ on E. For $\mathbf{x} \in E$, the function $\phi(\cdot, \mathbf{x}) : \mathbb{R} \to E$ defines a solution curve, trajectory, or orbit of the nonlinear system through the point $\mathbf{x}_0$ in E. If we identify the function $\phi(\cdot, \mathbf{x})$ with its graph, we can think of a trajectory through the point $\mathrm{x}_0 \in E$ as a motion along the curve

$$\Gamma_{\mathbf{x}_0} = \left\{\mathbf{x} \in E \mid \mathbf{x} = \boldsymbol{\phi}\left(t, \mathbf{x}_0\right), t \in \mathbb{R}\right\}$$

defined by the nonlinear system. We shall also refer to $\Gamma_{\mathbf{x}_0}$ as the trajectory of the nonlinear system through the point $\mathbf{x}_0$ at time $t = 0$. If the point $\mathbf{x}_0$ plays no role in the discussion, we simply denote the trajectory by Γ and draw the curve Γ in the subset E of the phase space $\mathbb{R}^{\boldsymbol{n}}$ with an arrow indicating the direction of the motion along Γ with increasing time. By the positive half-trajectory through the point $\mathbf{x}_0 \in E$, we mean the motion along the curve

$$\Gamma^{+}_{\mathbf{x}_0} = \{\mathbf{x} \in E \mid \mathbf{x} = \phi\left(t, \mathbf{x}_0\right), t \geq 0\}$$

$\Gamma^{-}_{\mathbf{x}_0}$, is similarly defined. Any trajectory $\Gamma = \Gamma^{+} \cup \Gamma^{-}$.

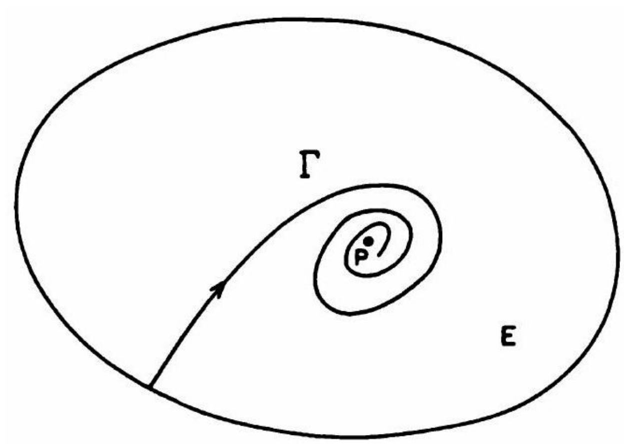


Figure 5: A trajectory of Γ of the initial value problem which approaches the ω limit point $\mathbf{p} \in E$ as $t \to \infty$.

Definition: A point $\mathbf{p} \in E$ is an ω-limit point of the trajectory $\phi(\cdot, \mathbf{x})$ of the linear system if there is a sequence $t_n \to \infty$ such that

$$\lim_{n \to \infty} \phi\left(t_n, \mathbf{x}\right) = \mathbf{p}$$

Similarly, if there is a sequence $t_n \to -\infty$ such that

$$\lim_{n \to \infty} \phi\left(t_n, \mathbf{x}\right) = \mathbf{q}$$

and the point $\mathbf{q} \in E$, then the point $\mathbf{q}$ is called an α-limit point of the trajectory $\phi(\cdot, \mathbf{x})$ of the initial nonlinear system. The set of all ω-limit points of a trajectory Γ is called the ω-limit set of Γ and it is denoted by $\omega(\Gamma)$. The set of all α-limit points of a trajectory Γ is called the α-limit set of Γ and it is denoted by $\alpha(\Gamma)$. The set of all limit points of $\Gamma, \alpha(\Gamma) \cup \omega(\Gamma)$ is called the limit set of Γ.

Theorem: The α and ω-limit sets of a trajectory Γ of the initial nonlinear system, $\alpha(\Gamma)$ and $\omega(\Gamma)$, are closed subsets of E and if Γ is contained in a compact subset of $\mathbb{R}^n$, then $\alpha(\Gamma)$ and $\omega(\Gamma)$, are non-empty, connected, compact subsets of E.

Proof: It follows from Definition 1 that $\omega(\Gamma) \subset E$. In order to show that $\omega(\Gamma)$ is a closed subset of E, we let $\mathbf{p}_n$ be a sequence of points in $\omega(\Gamma)$ with $\mathbf{p}_n \to \mathbf{p} \in \mathbb{R}^n$ and show that $\mathbf{p} \in \omega(\Gamma)$. Let $\mathbf{x}_0 \in \Gamma$. Then since $\mathbf{p}_n \in \omega(\Gamma)$, it follows that for each $n \in \mathbb{N}$, there is a sequence $t_k^{(n)} \to \infty$ as $k \to \infty$ such that

$$\lim_{k\to\infty} \boldsymbol{\phi}\left(t_k^{(n)} \cdot \mathbf{x}_0\right) = \mathbf{p}_n$$

Furthermore, we may assume that $t_k^{(n+1)} > t_k^{(n)}$ since otherwise we can choose a subsequence of $t_k^{(n)}$ with this property. The above equation implies that $\forall\ n \geq 2$, there is a sequence of integers $K(n) > K(n-1)$ such that for $k \geq K(n)$,

$$\left|\phi\left(t_k^{(n)} \cdot \mathbf{x}_0\right) - \mathbf{p}_n\right| < \frac{1}{n}$$

Let $t_n = t_{K(n)}^{(n)}$. Then $t_n \to \infty$ and by the triangle inequality,

$$|\boldsymbol{\phi}\left(t_n, \mathbf{x}_0\right) - \mathbf{p}| \leq |\boldsymbol{\phi}\left(t_n, \mathbf{x}_0\right) - \mathbf{p}_n| + |\mathbf{p}_n - \mathbf{p}| \leq \frac{1}{n} + |\mathbf{p}_n - \mathbf{p}| \to 0$$

as $n \to \infty$. Thus $\mathbf{p} \in \omega(\Gamma)$.
If $\Gamma \subset K$, a compact subset of $\mathbb{R}^n$. and $\boldsymbol{\phi}\left(t_n, \mathbf{x}_0\right) \to \mathbf{p} \in \omega(\Gamma)$, then $\mathbf{p} \in K$ since $\phi\left(t_n, \mathbf{x}_0\right) \in \Gamma \subset K$ and K is compact. Thus, $\omega(\Gamma) \subset K$ and therefore $\omega(\Gamma)$ is compact since a closed subset of a compact set is compact. Furthermore, $\omega(\Gamma) \neq 0$ since the sequence of points $\phi\left(n, \mathbf{x}_0\right) \in K$ contains a convergent subsequence which converges to a point in $\omega(\Gamma) \subset K$. Finally, suppose that $\omega(\Gamma)$ is not connected. Then there exist two nonempty, disjoint, closed sets A and B such that $\omega(\Gamma) = A \cup B$. Since A and B are both bounded, they are a finite distance δ apart where the distance from A to B

$$d(A, B) = \inf_{\mathbf{x}\in A.\mathbf{y}\in B} |\mathbf{x} - \mathbf{y}|$$

Since the points of A and B are ω-limit points of Γ, there exists arbitrarily large t such that $\phi\left(t, \mathbf{x}_0\right)$ are within $\delta/2$ of A and there exists arbitrarily large t such that the distance of $\phi\left(t, \mathbf{x}_0\right)$ from A is greater than $\delta/2$. Since the distance $d\left(\phi\left(t, \mathbf{x}_0\right), A\right)$ of $\phi\left(t, \mathbf{x}_0\right)$ from A is a continuous function of t, it follows that there must exist a sequence $t_n \to \infty$ such that $d\left(\phi\left(t_n, \mathbf{x}_0\right), A\right) = \delta/2$. Since $\{\boldsymbol{\phi}\left(t_n, \mathbf{x}_0\right)\} \subset K$ there is a subsequence converging to a point $\mathbf{p} \in \omega(\Gamma)$ with $d(\mathbf{p}, A) = \delta/2$. But, then $d(\mathbf{p}, B) \geq d(A, B) - d(\mathbf{p}, A) = \delta/2$ which implies that $\mathbf{p} \notin A$ and $\mathbf{p} \notin B$; i.e., $\mathbf{p} \notin \omega(\Gamma)$, a contradiction. Thus, $\omega(\Gamma)$ is connected. A similar proof serves to establish these same results for $\alpha(\Gamma)$.

Theorem: If $\mathbf{p}$ is an ω-limit point of a trajectory Γ of the initial nonlinear system, then all other points of the trajectory $\phi(\cdot, \mathbf{p})$ of the initial nonlinear system through the point $\mathbf{p}$ are also ω-limit points of Γ; i.e., if $\mathbf{p} \in \omega(\Gamma)$ then $\Gamma_{\mathbf{p}} \subset \omega(\Gamma)$ and similarly if $\mathbf{p} \in \alpha(\Gamma)$ then $\Gamma_p \subset \alpha(\Gamma)$.

Proof: Let $\mathbf{p} \in \omega(\Gamma)$ where Γ is the trajectory $\phi(\cdot, \mathbf{x}_0)$ of the initial nonlinear system through the point $\mathbf{x}_0 \in E$. Let $\mathbf{q}$ be a point on the trajectory $\phi(\cdot, \mathbf{p})$ of the initial nonlinear system through the point $\mathbf{p}$; i.e., $\mathbf{q} = \boldsymbol{\phi}(\bar{t}, \mathbf{p})$ for some $\bar{t} \in \mathbb{R}$. Since $\mathbf{p}$ is an ω-limit point of the trajectory $\phi(\cdot, \mathbf{x}_0)$, there is a sequence $t_n \to \infty$ such that $\phi(t_n, \mathbf{x}_0) \to \mathbf{p}$. Thus we have.

$$\phi\left(t_n + \tilde{t}, \mathbf{x}_0\right) = \phi\left(\tilde{t}, \phi\left(t_n, \mathbf{x}_0\right)\right) \to \phi(\tilde{t}, \mathbf{p}) = \mathbf{q}$$

And since $t_n + \tilde{t} \to \infty$, the point $\mathbf{q}$ is an ω-limit point of $\phi(\cdot, \mathbf{x}_0)$. A similar proof holds when $\mathbf{p}$ is an α-limit point of Γ and this completes the proof of the theorem.

It follows from this theorem that $\forall$ points $\mathbf{p} \in \omega(\Gamma), \phi_t(\mathbf{p}) \in \omega(\Gamma)\ \forall\ t \in \mathbb{R}$; i.e., $\phi_t(\omega(\Gamma)) \subset \omega(\Gamma)$. Thus, according to definition, we have the following result.

Corollary: $\alpha(\Gamma)$ and $\omega(\Gamma)$ are invariant with respect to the flow ϕ_t of the initial nonlinear system.

The α - and ω-limit sets of a trajectory Γ of the initial nonlinear system are thus closed invariant subsets of E. In the next definition, a neighborhood of a set A is any open set U containing A and we say that $\mathbf{x}(t) \to A$ as $t \to \infty$ if the distance $d(\mathrm{x}(t), A) \to 0$ as $t \to \infty$.

3.3 Attractors

A closed invariant set $A \subset E$ is called an attracting set of the initial nonlinear system if there is some neighborhood U of A such that $\forall\ \mathbf{x} \in U, \phi_t(\mathbf{x}) \in U\ \forall\ t \geq 0$ and $\phi_t(\mathbf{x}) \to A$ as $t \to \infty$. An attractor of the initial nonlinear system is an attracting set which contains a dense orbit.

Note that any equilibrium point $\mathbf{x}_0$ of the initial nonlinear system is its own α and ω-limit set since $\phi(t, \mathbf{x}_0) = \mathbf{x}_0\ \forall\ t \in \mathbb{R}$. And if a trajectory Γ of the initial nonlinear system has a unique ω-limit point x_0, then by the above Corollary, x_0 is an equilibrium point of the initial nonlinear system. A stable node or focus, is the ω-limit set of every trajectory in some neighborhood of the point; and a stable node or focus of the initial nonlinear system is an attractor of the initial nonlinear system. However, not every ω limit set of a trajectory of the initial nonlinear system is an attracting set of the initial nonlinear system; for example, a saddle $\mathbf{x}_0$ of a planar system is the ω-limit set of three trajectories in a neighborhood $N(\mathrm{x}_0)$, but no other trajectories through points in $N(\mathbf{x}_0)$ approach x_0 as $t \to \infty$.

If $\mathbf{q}$ is any regular point in $\alpha(\Gamma)$ or $\omega(\Gamma)$ then the trajectory through $\mathbf{q}$ is called a limit orbit of Γ. Thus, by the second theorem , we see that $\alpha(\Gamma)$ and $\omega(\Gamma)$ consist of equilibrium points and limit orbits of the initial nonlinear system. We now consider some specific examples of limit sets and attractors.

Circular Attractor

Consider the system

$$\dot{x} = -y + x\left(1 - x^2 - y^2\right)$$
$$\dot{y} = x + y\left(1 - x^2 - y^2\right).$$

In polar coordinates, we have

$$\dot{r} = r\left(1 - r^2\right)$$
$$\dot{\theta} = 1.$$

We see that the origin is an equilibrium point of this system; the flow spirals around the origin in the counter-clockwise direction; it spirals outward for $0 < r < 1$ since $\dot{r} > 0$ for $0 < r < 1$; and it spirals inward for $r > 1$ since $\dot{r} < 0$ for $r > 1$. The counter-clockwise flow on the unit circle describes a trajectory Γ_0 of the initial nonlinear system since $\dot{r} = 0$ on $r = 1$. The trajectory through the point $(\cos\theta_0, \sin\theta_0)$ on the unit circle at $t = 0$ is given by $\mathbf{x}(t) = (\cos(t + \theta_0), \sin(t + \theta_0))^T$. The phase portrait for this system is shown in the figure. The trajectory Γ_0 is called a stable limit cycle.

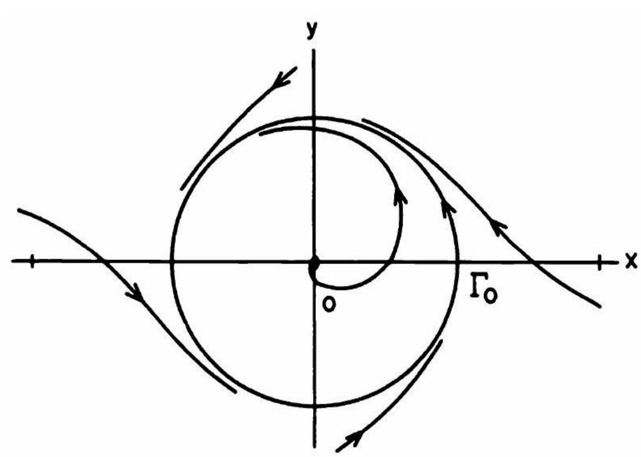


Figure 6: A stable limit cycle Γ_0 which is an attractor of the initial nonlinear system.

Spherical Attractor

The system

$$\begin{aligned}\dot{x} &= -y + x\left(1 - z^2 - x^2 - y^2\right)\\ \dot{y} &= x + y\left(1 - z^2 - x^2 - y^2\right)\\ \dot{z} &= 0\end{aligned}$$

has the unit two-dimensional sphere S^2 together with that portion of the z-axis outside S^2 as an attracting set. Each plane $z = z_0$ is an invariant set and for $|z_0| < 1$ the ω-limit set of any trajectory not on the z-axis is a stable cycle on S^2.

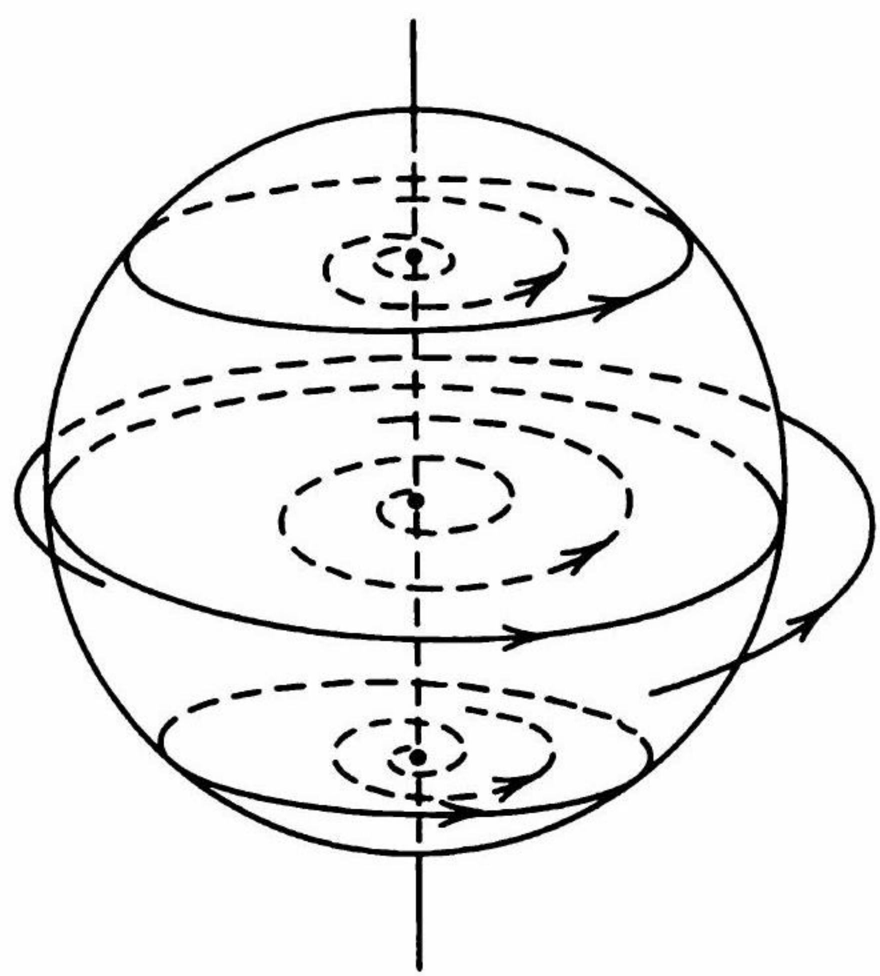


Figure 7: A dynamical system with S^2 as attracting set

Cylindrical Attractor

The system

$$
\begin{aligned}
\dot{x} &= -y + x\left(1 - x^2 - y^2\right) \\
\dot{y} &= x + y\left(1 - x^2 - y^2\right) \\
\dot{z} &= \alpha
\end{aligned}
$$

has the z-axis and the cylinder $x^2 + y^2 = 1$ as invariant sets. The cylinder is an attracting set.

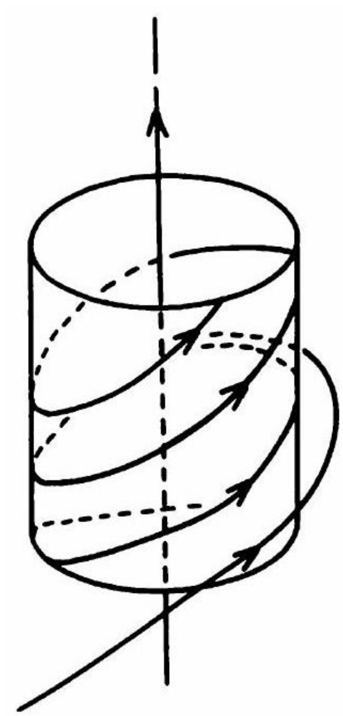


Figure 8: A dynamical system with cylinder as a attracting set

Toroidal Attractor

If in the previous example we identify the points $(x, y, 0)$ and $(x, y, 2\pi)$ in the planes $z = 0$ and $z = 2\pi$, we get a flow in $\mathbb{R}^3$ with a two-dimensional invariant torus T^2 as an attracting set. The z-axis gets mapped onto an unstable cycle Γ. And if α is an irrational multiple of π then the torus T^2 is an attractor and it is the ω-limit set of every trajectory except the cycle Γ.

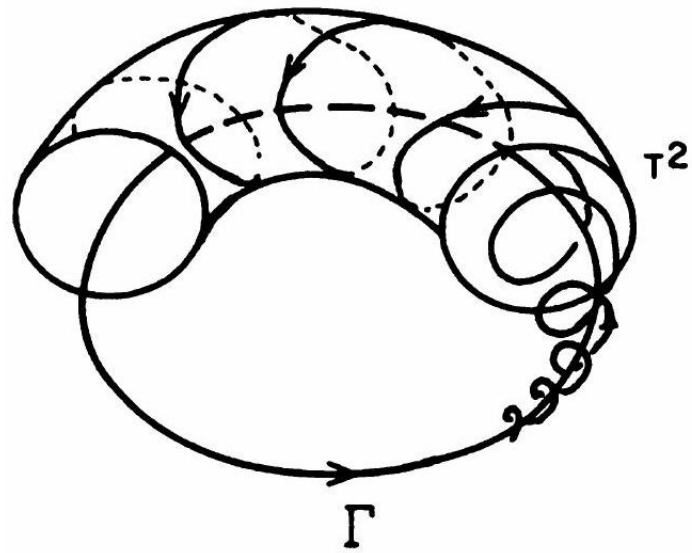


Figure 9: A dynamical system with an invariant torus as an attracting set.

Lorenz System

The original work of Lorenz in 1963 as well as the more recent work of *Sparrow* indicates that for certain values of the parameters σ, ρ and β, the system

$$\begin{aligned}\dot{x} &= \sigma(y - x)\\ \dot{y} &= \rho x - y - xz\\ \dot{z} &= -\beta z + xy\end{aligned}$$

has a strange attracting set. For example for $\sigma = 10, \rho = 28$ and $\beta = 8/3$, a single trajectory of this system is shown in the figure along with a "branched surface" S. The attractor A of this system is made up of an infinite number of branched surfaces S which are interleaved and which intersect; however, the trajectories of this system in A do not intersect but move from one branched surface to another as they circulate through the apparent branch. The numerical results and the related theoretical work indicate that the closed invariant set A contains

(i) a countable set of periodic orbits of arbitrarily large period,

(ii) an uncountable set of nonperiodic motions and

(iii) a dense orbit.

The attracting set A having these properties is referred to as a strange attractor.

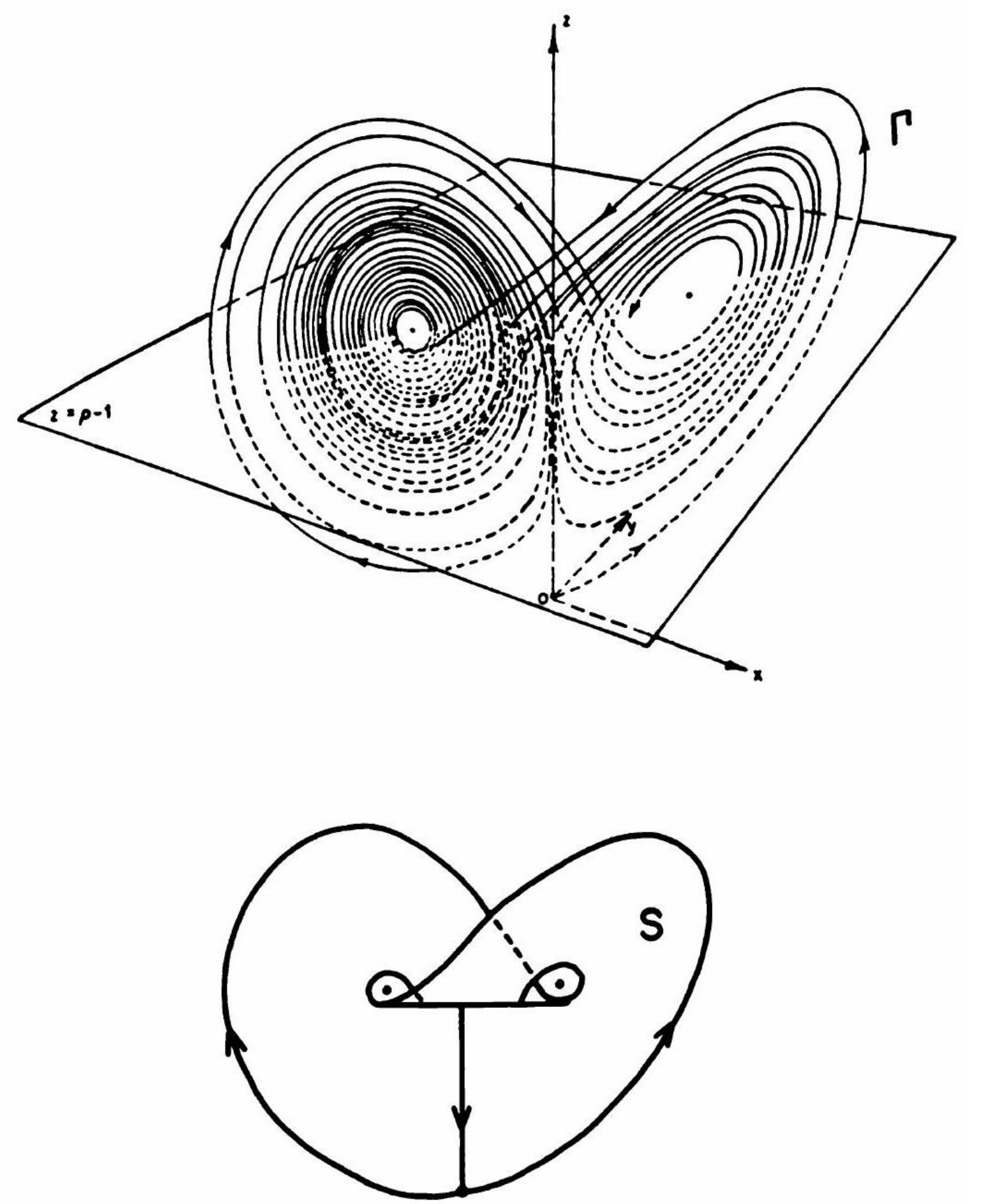


Figure 10: A trajectory Γ of the Lorenz system and the corresponding branched surface S.

Halvorsen Attractor

This is another famous strange attractor, governed by the differential equations

$$\dot{x} = ax - 4y - 4z - y^2$$
$$\dot{y} = ay - 4z - 4x - z^2$$
$$\dot{z} = az - 4x - 4y - x^2$$

For different values of the parameter a, different results are obtained.

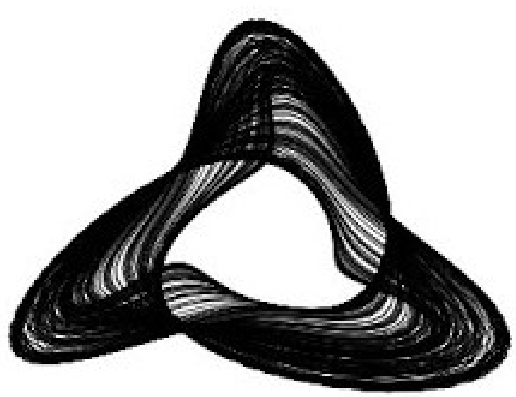


Figure 11: Halvorsen Attractor

Note: Plotting Attractors in GeoGebra

GeoGebra is a great tool to observe all the attractors and dynamical systems we have discussed, it gives us robust customization and free choice of number of observed particles and parameters.

```
##Parameters, you can modify it according to the equation##
d = 10
b = 8/3
p = 28

##System of differential equations: Lorenz attractor, you can go for
   the others as well##
x'(t,x,y,z) = d * (y - x)
y'(t,x,y,z) = x * (p - z) - y
z'(t,x,y,z) = x * y - b * z

##Initial Condition##
x0 = 1
y0 = 1
z0 = 1

```

```
    ##Numerical solution##
    NSolveODE({x', y', z'}, 0, {x0, y0, z0}, 20)

    ##Note##
    # The command NSolveODE() creates three curves
    # containing the numerical solution of the system
    # per variable (x,y and z) and they are plotted
    # against time in the 2D graphic view.

    ##Calculate length of solution 1##
    len = Length(numericalIntegral1)

    ##Define points from the solution##
    L_1 = Sequence( (y(Point(numericalIntegral1, i)), y(Point(
       numericalIntegral2, i)), y(Point(numericalIntegral3, i))), i, 0, 1,
       1 / len )

    ##Draw curve##
    f = Polyline(L_1)

    ##Finally, you need to hide numericalIntegra1, numericalIntegra2,
       numericalIntegra3, and L_1##
```

References

[1] Lawrence Perko, *Differential Equation and Dynamical Systems.* Springer, 1998.

[2] M. W. Hirsch, S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra* Academic Press, 1974.