Introduction to Ordinary Differential Equations, Math 462/512, Fall 2020

Jens Lorenz

November 25, 2020

Department of Mathematics and Statistics,
UNM, Albuquerque, NM 87131

Contents

1 Overview

1.1 Constant–Coefficient Systems $x' = Ax$ and Stability of a Fixed Point

Qualitative properties of systems

$$x' = Ax, \quad x(t) \in \mathbb{R}^n \ ,$$

for constant real matrices A.

Phase plane diagrams for $n = 2$.

The matrix exponential e^{At}.

Complex and real Jordan normal form.

Estimates of $|e^{At}|$ as $t \to \infty$.

Asymptotic stability of $x = 0$ for systems

$$x' = Ax + Q(x)$$

if all eigenvalues of A have negative real parts and $|Q(x)| \le C|x|^2$.

1.2 First Order Systems $x' = A(t)x + b(t)$

Fundamental matrices $\Phi(t)$ satisfy

$$\Phi'(t) = A(t)\Phi(t), \quad det\,\Phi(t) \neq 0 \ .$$

Liouville's theorem:

$$\frac{d}{dt}\,det\,\Phi(t) = tr\,A(t) \cdot det\,\Phi(t)$$

The inhomogeneous initial value problem

$$x' = A(t)x + b(t), \quad x(0) = x_0$$

has the solution

$$x(t) = \Phi_0(t)x_0 + \Phi_0(t)\int_0^t \Phi_0^{-1}(s)b(s)\,ds$$

where $\Phi_0(t)$ is the normalized fundamental matrix, i.e., $\Phi_0(0) = I$.

Difficulties with localization: Assume that, for every fixed t_0, the norm of $e^{A(t_0)t}$ decays exponentially as $t \to \infty$. Will all solutions $x(t)$ of the system $x' = A(t)x$ decay as $t \to \infty$? The answer is *No*, in general.

1.3 Floquet Theory

Consider a linear system

$$x' = A(t)x, \quad x(t) \in \mathbb{R}^n \ ,$$

where the matrix function $A(t) \in \mathbb{R}^{n \times n}$ has period $T > 0$, i.e., $A(t+T) \equiv A(t)$.

The normalized fundamental matrix can be written in the form

$$\Phi_0(t) = Q(t)e^{Bt}$$

where $B \in \mathbb{C}^{n\times n}$ is a constant matrix and $Q(t) \in \mathbb{C}^{n\times n}$ has period T. The matrix

$$M := \Phi_0(T) = e^{BT} \in \mathbb{R}^{n\times n}$$

is called the monodromy matrix of the system $x' = A(t)x$. The eigenvalues μ_j of M are called the Floquet multipliers. If $|\mu_j| < 1$ for all j then $x(t) \to 0$ as $t \to \infty$ for every solution of the system $x' = A(t)x$.

1.4 Basic Existence and Uniqueness Results

Let $f : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ denote a continuous function. If the function $f(x, t)$ satisfies a Lipschitz condition,

$$|f(x,t) - f(y,t)| \le L|x - y| \quad \text{for all} \quad x, y \in \mathbb{R} \quad \text{and} \quad 0 \le t \le T \ ,$$

then the initial value problem

$$x' = f(x,t), \quad x(0) = a \ ,$$

has a unique solution $x(t) = x(t, a)$. Our proof uses the iteration

$$x_{k+1}(t) = a + \int_0^t f(x_k(s), s))\, ds \quad \text{for} \quad k = 0, 1, \ldots$$

starting with $x_0 \equiv 0$. The completeness of the Banach space $\Big(C[0, T], |\cdot|_\infty\Big)$ is important.

One often considers autonomous IVPs

$$x' = f(x), \quad x(0) = a \ , \tag{1.1}$$

where $f \in C^1(\mathbb{R}^n)$. In this case the function $f(x)$ is locally Lipschitz. One obtains a local solution in a maximal time interval $0 \le t < t^*$ where t^* depends on the initial value a. If t^* is finite then

$$\lim_{t \to t^*-} |x(t)| = \infty \ ,$$

i.e., the solution blows up as t approaches t^*.

(This result for ODEs is quite different if one considers PDEs. In the PDE case, some spatial norms of the solution may blow up while others remain bounded as t approaches t^*.)

Let $x(t) = x(t, a)$ denote the solution of the IVP (1.1). If $f \in C^1$ then we will prove that the function $a \to x(t, a)$ is differentiable and the equation

$$x_t(t, a) = f(x(t, a))$$

can be differentiated in a. One obtains that

$$x_{at}(t,a) = f_x(x(t,a))x_a(t,a), \quad x_a(0,a) = I \ .$$

This says that the matrix function $\Phi_0(t) = x_a(t,a)$ is the normalized fundamental matrix of the linear ODE system $\phi' = A(t)\phi$ where $A(t) = f_x(x(t,a))$.

1.5 Asymptotic Stability of Periodic Orbits

Let $u(t) \in \mathbb{R}^N$ denote a periodic solution of the nonlinear system $x' = f(x)$, thus $u(t) \equiv u(t+T)$ where $T > 0$ and $u(t)$ is not constant. To analyze the stability of the corresponding orbit

$$\gamma = \{u(t) \ : \ t \in \mathbb{R}\}$$

we consider the linear system

$$x' = Df(u(t))x$$

where $Df(x) = f'(x)$ is the Jacobian of $f(x)$. The results of Chapter 4 indicate that the monodromy matrix

$$M = \Phi_0(T) = e^{BT}$$

plays an important role. From $u'(t) = f(u(t))$ one obtains that $u''(t) = Df(u(t))u'(t)$. This shows that the function $x(t) := u'(t)$ solves the linear system $x' = Df(u(t))x$ and one obtains that

$$u'(0) = u'(T) = \Phi_0(T)u'(0) \ .$$

The number $\mu_1 = 1$ is an eigenvalue of the monodromy matrix $M = \Phi_0(T)$ and Theorem 4.2 is not directly applicable.

Instead, one defines the Poincaré map corresponding to the orbit γ and some cross–section. Then Theorem 14.6, a theorem about the stability of a fixed point of a map, can be applied if the Floquet multipliers $\mu_2, \ldots, \mu_n$ are all strictly less than 1 in absolute value.

1.6 Introduction to Bifurcations of Equilibria

We consider scalar ODEs depending on a real parameter λ:

$$x' = f(x, \lambda) \ .$$

A simple example is the equation

$$x' = \lambda + x^2 \ .$$

For $\lambda < 0$ there are two fixed points,

$$x_{1,2}(\lambda) = \pm\sqrt{-\lambda} \ ,$$

which collide at $\lambda = 0$. This is an example of a saddle–node bifurcation.

For the equation

$$x' = x(\lambda - x)$$

a transcritical bifurcation occurs at $\lambda = 0$: The two solution branches

$$x_1(\lambda) = 0, \quad x_2(\lambda) = \lambda$$

cross each other at $\lambda = 0$ and an exchange of stability occurs at $\lambda = 0$.

The equation

$$x' = x(\lambda - x^2)$$

gives an example of a supercritical pitchfork bifurcation. If one perturbs this equation by $h \in \mathbb{R}$ and considers

$$x' = h + x(\lambda - x^2)$$

then the pitchfork bifurcation gets perturbed. We will discuss the 3D surface of fixed point,

$$x^* = x^*(\lambda, h) \ .$$

A so-called cusp–catastrophe occurs.

1.7 Hopf Bifurcation

The simplest example where a Hopf bifurcation occurs is given by the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A(\lambda) \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{where} \quad A(\lambda) = \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix}$$

and λ is a real parameter. The eigenvalues of $A(\lambda)$ are

$$\mu_{1,2}(\lambda) = \lambda \pm i \ .$$

For $\lambda < 0$ the origin is an asymptotically stable fixed point. For $\lambda > 0$ the origin is an unstable fixed point. Thus, at $\lambda = 0$ a change of stability occurs.

An important observation is that for $\lambda = 0$ the system has a branch of periodic orbits given by

$$\begin{pmatrix} x \\ y \end{pmatrix}(t) = a \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

where $a > 0$ is the amplitude of the orbit.

More generally, let $f : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N$ denote a smooth function and consider the ODE system

$$x' = f(x, \lambda), \quad x(t) \in \mathbb{R}^N \ ,$$

depending on the real parameter λ. Assume that $x^*(\lambda) \in \mathbb{R}^N$ denotes a branch of fixed points, depending smoothly on λ, i.e.,

$$f(x^*(\lambda), \lambda) = 0 \ .$$

Let

$$A(\lambda) = Df(x^*(\lambda), \lambda) = f_x(x^*(\lambda), \lambda)) \in \mathbb{R}^{N \times N}$$

denote the Jacobian of $f(\cdot, \lambda)$ at the fixed point $x^*(\lambda)$. Assume that all eigenvalues $\mu_j(\lambda)$ of $A(\lambda)$ satisfy

$$\operatorname{Re} \mu_j(\lambda) < 0 \quad \text{for} \quad \lambda < \lambda_0 \ .$$

Thus the fixed point $x^*(\lambda)$ is asymptotically stable for $\lambda < \lambda_0$. Also, assume that for $\lambda = \lambda_0$ a pair of eigenvalues are purely imaginary,

$$\mu_{1,2} = \pm ib, \quad b > 0 \ ,$$

and assume that

$$\frac{d}{d\lambda} \operatorname{Re} \mu_{1,2}(\lambda_0) > 0 \ .$$

The fixed point $x^*(\lambda)$ is unstable for $\lambda_0 < \lambda \leq \lambda_0 + \varepsilon$ and one can expect a branch of periodic orbits to occur for λ near λ_0. If the branch of periodic orbits occurs for $\lambda_0 < \lambda \leq \lambda_0 + \varepsilon$, then a supercritical Hopf bifurcation occurs at $\lambda = \lambda_0$; the periodic orbits are stable.

On the other hand, if the branch of periodic orbits occurs for $\lambda_0 - \varepsilon \leq \lambda < \lambda_0$, then a subcritical Hopf bifurcation occurs at $\lambda = \lambda_0$; the periodic orbits are unstable.

We will prove this result for systems in two dimensions.

1.8 The Poincaré–Bendixson Theorem

1.9 Hamiltonian Systems

1.10 The Energy Method and the Mathematical Pendulum

1.11 Planetary Motion: The Two Body Problem

1.12 The Stable Manifold Theorem

2 Constant Coefficient Systems $x' = Ax$ and Stability of a Fixed Point

Let $A \in \mathbb{R}^{n\times n}$ denote a real $n \times n$ matrix and let $x_0 \in \mathbb{R}^n$. The initial–value problem (IVP)

$$x' = Ax, \quad x(0) = x_0 \ ,$$

has the unique solution

$$x(t) = x(t, x_0) = e^{At}x_0, \quad t \in \mathbb{R} \ ,$$

where

$$e^{At} = \sum_{j=0}^{\infty} \frac{1}{j!} A^j t^j \ .$$

This result is theoretically useful, but what does it tell you about the behavior of the solution?

In Section 2.1 we consider examples for $n = 2$ and use elementary techniques to obtain solutions and to discuss their properties.

For general n, the formula $x(t) = e^{At}x_0$ is useful if one can transform A to diagonal form,

$$T^{-1}AT = \Lambda = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

since $A = T\Lambda T^{-1}$ yields that

$$e^{At} = Te^{\Lambda t}T^{-1} \quad \text{where} \quad e^{\Lambda t} = \begin{pmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_n t} \end{pmatrix} .$$

If A has complex eigenvalues then the transformation matrix T cannot be real. In ODE theory one tries to avoid non–real transformations since they may not be applicable to nonlinear extensions of linear equations, like

$$x' = Ax + Q(x)$$

where $Q : \mathbb{R}^n \to \mathbb{R}^n$ is nonlinear.

We will consider the transformation of real matrices $A \in \mathbb{R}^{n\times n}$ to complex and real Jordan normal form.

2.1 Examples in 2D

Example 1: The system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$$

has the general solution

$$x(t) = x_0 e^{-t}, \quad y(t) = y_0 e^{2t} \ .$$

A sketch of the phase plane:

The origin is a fixed point for any linear system $x' = Ax$. In Example 1 the origin is a saddle point.

Example 2: Consider the initial–value problem

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} -1 & -2 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix}(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \ . \qquad (2.1)$$

Obtain that

$$y(t) = y_0 e^{-2t}$$

and

$$x' = -x - 2y_0 e^{-2t}, \quad x(0) = x_0 \ .$$

Recall that the IVP

$$x' = ax + g(t), \quad x(0) = x_0 \ ,$$

has the solution

$$x(t) = x_0 e^{at} + \int_0^t e^{a(t-s)} g(s)\, ds \ .$$

One obtains the solution of (2.1):

$$\begin{pmatrix} x \\ y \end{pmatrix}(t) = \begin{pmatrix} x_0 e^{-t} \\ y_0 e^{-2t} \end{pmatrix} + \begin{pmatrix} -2y_0(e^{-t} - e^{-2t}) \\ 0 \end{pmatrix} \tag{2.2}$$

$$= (x_0 - 2y_0)e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y_0 e^{-2t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \tag{2.3}$$

An important process to obtain the general solution of a system $x' = Ax$ is to transform A to diagonal form, if possible. Let us apply this process to (2.1). The matrix

$$A = \begin{pmatrix} -1 & -2 \\ 0 & -2 \end{pmatrix}$$

has the eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -2$. Also,

$$A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad A \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} .$$

Denote the eigenvectors by

$$t^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad t^{(2)} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

and define the matrix T with columns $t^{(j)}$:

$$T = \left(t^{(1)} \; t^{(2)} \right) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} .$$

We have

$$AT = T\Lambda \quad \text{with} \quad \Lambda = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} .$$

Therefore,

$$T^{-1}AT = \Lambda$$

is a similarity transformation of A to diagonal form.

If one uses the new variable

$$\begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix}$$

determined by the transformation

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = T \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix}$$

one obtains the system

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix}' = \Lambda \begin{pmatrix} \xi \\ \eta \end{pmatrix} ,$$

thus

$$\xi(t) = \xi_0 e^{-t} \quad \text{and} \quad \eta(t) = \eta_0 e^{-2t} \; .$$

Therefore,

$$\begin{pmatrix} x \\ y \end{pmatrix}(t) = \xi_0 e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \eta_0 e^{-2t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \; .$$

From

$$\begin{pmatrix} \xi_0 \\ \eta_0 \end{pmatrix} = T^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \; ,$$

obtain that

$$\xi_0 = x_0 - 2y_0, \quad \eta_0 = y_0 \; ,$$

thus

$$\begin{pmatrix} x \\ y \end{pmatrix}(t) = (x_0 - 2y_0) e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y_0 e^{-2t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \; .$$

This solution formula agrees with (2.3).

A sketch of the phase plane:

Note: If the initial vector $(x_0, y_0)^T$ is a multiple of the eigenvector $t^{(1)}$, then the solution decays like e^{-t}:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \alpha e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \; .$$

If the initial vector $(x_0, y_0)^T$ is a multiple of the eigenvector $t^{(2)}$, then the solution decays like e^{-2t}:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \beta e^{-2t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \; .$$

If

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \alpha t^{(1)} + \beta t^{(2)}$$

then the solution is

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \alpha e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta e^{-2t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} .$$

If $\alpha \neq 0$ then the corresponding orbit becomes tangential to the eigenvector $t^{(1)}$, i.e., to the x–axis, as $t \to \infty$ and the orbit approaches the origin.

Example 3: Consider the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{2.4}$$

The matrix A has the eigenvalues $\lambda_{1,2} = \pm i$. On can transform A to diagonal form,

$$T^{-1}AT = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} ,$$

but the transformation matrix T is not a real matrix. To solve the IVP for the system (2.4) it is easy to check that the system yields

$$x' = -y, \quad y' = x ,$$

thus

$$x'' + x = 0, \quad y'' + y = 0 .$$

The initial condition

$$x(0) = x_0, \quad y(0) = y_0$$

yields the solution

$$\begin{aligned} x(t) &= x_0 \cos t - y_0 \sin t , && (2.5) \\ y(t) &= y_0 \cos t + x_0 \sin t . && (2.6) \end{aligned}$$

It follows that (with $c = \cos t, s = \sin t$)

$$\begin{aligned} x^2(t) + y^2(t) &= (x_0 c - y_0 s)^2 + (y_0 c + x_0 s)^2 \\ &= x_0^2 + y_0^2 , \end{aligned}$$

thus all orbits are circle around the origin. The solutions $(x(t), y(t))^T$ have the period 2π and encircle the origin counterclockwise since

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

Let us consider another method to obtain the solution of the IVP for the system (2.4). In general, if $A \in \mathbb{C}^{n \times n}$, then the IVP

$$x' = Ax, \quad x(0) = x_0 \ ,$$

has the solution

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

where

$$e^{At} = \sum_{j=0}^{\infty} \frac{1}{j!} A^j t^j \ .$$

Let us use this formula for

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \ .$$

It is easy to check that $A^2 = -I$, thus

$$A^{2k} = (-1)^k I \quad \text{and} \quad A^{2k+1} = (-1)^k A \ .$$

Therefore,

$$\begin{aligned}
e^{At} &= \sum_{j=0}^{\infty} \frac{1}{j!} A^j t^j \quad (j = 2k \text{ or } j = 2k+1) \\
&= \sum_{k=0}^{\infty} \frac{1}{(2k)!} A^{2k} t^{2k} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} A^{2k+1} t^{2k+1} \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} t^{2k} I + \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} t^{2k+1} A \\
&= \cos t \, I + \sin t \, A \\
&= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}
\end{aligned}$$

The solution formula

$$\begin{pmatrix} x \\ y \end{pmatrix} (t) = e^{At} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

yields that

$$\begin{pmatrix} x \\ y \end{pmatrix} (t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} x_0 + \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} y_0 \ .$$

This agrees with (2.5), (2.6).

Example 4: Consider the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{2.7}$$

where $a, b \in \mathbb{R}, b \neq 0$. We have

$$A = aI + bA_0$$

where A_0 is the matrix called A in Example 3. Since the matrices A_0 and I commute under multiplication, one obtains that

$$\begin{aligned} e^{At} &= e^{aI+bA_0} \\ &= e^{at}e^{A_0bt} \\ &= e^{at}\begin{pmatrix} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{pmatrix} \end{aligned}$$

If $a < 0$ the orbits spiral towards the origin as $\to \infty$. If $a > 0$ the orbits spiral towards infinity as $\to \infty$.

Example 5: Consider the IVP

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix}(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} . \tag{2.8}$$

The eigenvalues of A are $\lambda_{1,2} = \pm 2i$. Instead of using the complex transformation

$$T^{-1}AT = \begin{pmatrix} 2i & 0 \\ 0 & -2i \end{pmatrix}$$

we use the real transformation

$$P^{-1}AP = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \quad \text{with} \quad P = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} .$$

Using the result of Example 4 one obtains that

$$e^{At} = P\begin{pmatrix} \cos(2t) & -\sin(2t) \\ \sin(2t) & \cos(2t) \end{pmatrix}P^{-1} = \begin{pmatrix} \cos(2t) & -2\sin(2t) \\ \frac{1}{2}\sin(2t) & \cos(2t) \end{pmatrix} .$$

The solution of the IVP (2.8) is

$$\begin{pmatrix} x \\ y \end{pmatrix}(t) = \begin{pmatrix} \cos(2t)x_0 - 2\sin(2t)y_0 \\ \frac{1}{2}\sin(2t)x_0 + \cos(2t)y_0 \end{pmatrix}$$

It is not difficult to check that

$$\frac{1}{4}x^2(t) + y^2(t) = \frac{1}{4}x_0^2 + y_0^2 = const =: c^2 .$$

The orbits are ellipses with equations

$$\left(\frac{x}{2c}\right)^2 + \left(\frac{y}{c}\right)^2 = 1 \quad \text{for} \quad c > 0 .$$

Example 6: Consider the IVP

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix}(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} . \tag{2.9}$$

The matrix A has the eigenvalue $\lambda = -1$, which is geometrically simple, but algebraically double.

We write

$$A = -I + J \quad \text{with} \quad J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} .$$

It is easy to check that $J^2 = 0$. Therefore,

$$e^{At} = e^{-t}e^{Jt} = e^{-t}(I + tJ) = e^{-t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} .$$

The solution of the IVP (2.9) is

$$x(t) = e^{-t}x_0 + e^{-t}ty_0, \quad y(t) = e^{-t}y_0 ,$$

i.e.,

$$\begin{pmatrix} x \\ y \end{pmatrix}(t) = e^{-t} \begin{pmatrix} x_0 + ty_0 \\ y_0 \end{pmatrix}$$

Let

$$P(t) := \Big(x(t), y(t)\Big) = e^{-t}t\Big(\frac{x_0}{t} + y_0, \frac{y_0}{t}\Big) \quad \text{for} \quad t > 0 .$$

Of course, the points $P(t)$ approach the origin as $t \to \infty$. We claim that the orbits are tangential to the x–axis at the origin. The tangent vector at $P(t)$ is

$$P'(t) = e^{-t}t\Big(\frac{y_0 - x_0}{t} - y_0, \frac{y_0}{t}\Big) .$$

For $t \to \infty$ and $y_0 \neq 0$ this vector approaches

$$(-y_0, 0) .$$

At the origin, the orbits are tangential to the eigenvector $(1, 0)^T$ of A.

A sketch of the phase plane:

2.2 Solutions of $x' = Ax$ Corresponding to Complex Eigenvalues

Consider the IVP

$$x' = Ax, \quad x(0) = x_0 \ ,$$

where $A \in \mathbb{R}^{n\times n}, x(t) \in \mathbb{R}^n$. Assume that

$$\lambda = a + ib, \quad a, b \in R, \quad b > 0 \ ,$$

is an eigenvalue of A. (We know that $\bar{\lambda} = a - ib$ also is an eigenvalue of A.) Let $w = u + iv, u, v \in \mathbb{R}^n$, denote an eigenvector of A to the eigenvalue $\lambda = a + ib$, thus

$$\begin{aligned} A(u+iv) &= (a+ib)(u+iv) = (au-bv) + i(av+bu) \\ A(u-iv) &= (a-ib)(u-iv) = (au-bv) - i(av+bu) \end{aligned}$$

One obtains that

$$Au = au - bv \tag{2.10}$$

$$Av = av + bu \tag{2.11}$$

Let us assume that the initial vector $x(0) = x_0$ lies in $span\{u, v\}$, i.e.,

$$x_0 = \alpha_0 v + \beta_0 u, \quad \alpha_0, \beta_0 \in \mathbb{R} \ .$$

We claim that the solution

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

lies in $span\{u, v\}$ for all $t \in \mathbb{R}$, i.e.,

$$x(t) = \alpha(t) v + \beta(t) u \tag{2.12}$$

for real valued functions $\alpha(t), \beta(t)$. Indeed, if $x(t)$ solves $x' = Ax$ and has the form (2.12) then

$$Ax = x' = \alpha' v + \beta' u$$

and

$$\begin{aligned} Ax &= \alpha A v + \beta A u \\ &= \alpha(av + bu) + \beta(au - bv) \\ &= (a\alpha - b\beta) v + (b\alpha + a\beta) u \ , \end{aligned}$$

thus

$$\alpha' = a\alpha - b\beta, \quad \beta' = b\alpha + a\beta \ .$$

In matrix form,

$$\begin{pmatrix} \alpha'(y) \\ \beta'(t) \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} . \tag{2.13}$$

The converse is also easy to check: If $\alpha(t), \beta(t)$ solve (2.13) then

$$x(t) = \alpha(t)v + \beta(t)u$$

solves $x' = Ax$. We have proved:

Lemma 2.1 *Let $\lambda_{1,2} = a \pm ib$ denote a pair of complex conjugate eigenvalues of $A \in \mathbb{R}^{n\times n}$ and let*

$$Aw = (a+ib)w, \quad A\bar{w} = (a-ib)\bar{w} \ .$$

Let $w = u + iv$ where $u, v \in \mathbb{R}^n$. A vector function

$$x(t) = \alpha(t)v + \beta(t)u$$

with $\alpha, \beta \in C^1(\mathbb{R})$ solves $x' = Ax$ if and only if (2.13) holds.

2.3 Part of the Transformation to Complex and Real Jordan Normal Form

2.3.1 Use of One Complex Eigenvalue

Let $A \in \mathbb{R}^{n\times n}$ have the complex eigenvalues

$$\lambda = a + ib \quad \text{and} \quad \bar{\lambda} = a - ib$$

and let

$$Aw = \lambda w \quad \text{and} \quad A\bar{w} = \bar{\lambda}\bar{w} \ . \tag{2.14}$$

If $(w|\bar{w}) \in \mathbb{C}^{n\times 2}$ denotes the matrix with columns w and $\bar{w}$ then

$$\begin{aligned} A(w|\bar{w}) &= (Aw|A\bar{w}) \\ &= (\lambda w|\bar{\lambda}\bar{w}) \\ &= (w|\bar{w}) \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} . \end{aligned}$$

This shows that eigenvectors w and $\bar{w}$ play a role when one transforms A to a diagonal matrix or to complex Jordan normal form.

Use of a Real Transformation: As above, let w and $\bar{w}$ denote eigenvectors of A, i.e., (2.14 holds. Also, as above, let $\lambda = a + ib$. Recall (2.10), (2.11), thus

$$Au = au - bv \quad \text{and} \quad Av = av + bu \ .$$

Therefore,

$$\begin{aligned} A(v|u) &= (Av|Au) \qquad (2.15)\\ &= (av+bu|au-bv) \qquad (2.16)\\ &= (v|u)\begin{pmatrix} a & -b \\ b & a \end{pmatrix}. \qquad (2.17) \end{aligned}$$

This shows that the vectors u and v play a role when one transforms A to real Jordan normal form.

2.3.2 A Complex Double Eigenvalue

Let $A \in \mathbb{R}^{n\times n}$ denote a real $n \times n$ matrix with complex eigenvalue λ, which is geometrically simple, but algebraically double.

The Complex View: There exist non–trivial vectors $w_1, w_2 \in \mathbb{C}^n$ with

$$\begin{aligned} (A-\lambda I)w_1 &= 0 \qquad (2.18)\\ (A-\lambda I)w_2 &= w_1 \qquad (2.19) \end{aligned}$$

(The vector w_1 is an eigenvector of A to the eigenvalue λ. The vector w_2 is a generalized eigenvector, satisfying $(A-\lambda I)^2 w_2 = 0$. The vectors w_1 and w_2 form a basis of the generalized eigenspace of A to the eigenvalue λ.)

The equations (2.18) and (2.19) yield in matrix form:

$$\begin{aligned} A(w_1|w_2) &= (Aw_1|Aw_2)\\ &= (\lambda w_1|w_1+\lambda w_2)\\ &= (w_1|w_2)\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}. \end{aligned}$$

Similarly,

$$\begin{aligned} (A-\bar\lambda I)\bar w_1 &= 0\\ (A-\bar\lambda I)\bar w_2 &= \bar w_1 \end{aligned}$$

In matrix form:

$$A(\bar w_1|\bar w_2) = (\bar w_1|\bar w_2)\begin{pmatrix} \bar\lambda & 1 \\ 0 & \bar\lambda \end{pmatrix}.$$

Taken together, one obtains

$$A(w_1|w_2|\bar w_1|\bar w_2) = (w_1|w_2|\bar w_1|\bar w_2)\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \bar\lambda & 1 \\ 0 & 0 & 0 & \bar\lambda \end{pmatrix}.$$

This is a part of the similarity transformation of A to complex Jordan normal form.

The Real View: Let

$$w_1 = u_1 + iv_1, \quad w_2 = u_2 + iv_2$$

with $u_1, v_1, u_2, v_2 \in \mathbb{R}^n$. Since $Aw_1 = (a+ib)w_1$ obtain that

$$A(v_1|u_1) = (v_1|u_1) \begin{pmatrix} a & -b \\ b & a \end{pmatrix} .$$

(See (2.17).)

From

$$Aw_2 = (a+ib)w_2 + w_1$$

obtain that

$$\begin{aligned} A(u_2 + iv_2) &= (a+ib)(u_2+iv_2) + u_1 + iv_1 \\ &= au_2 - bv_2 + u_1 + i(bu_2 + av_2 + v_1) \ , \end{aligned}$$

thus

$$\begin{aligned} Au_2 &= au_2 - bv_2 + u_1 \\ Av_2 &= bu_2 + av_2 + v_1 \ . \end{aligned}$$

In matrix form,

$$A(v_2|u_2) = (v_1|u_1|v_2|u_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ a & -b \\ b & a \end{pmatrix} .$$

Setting

$$T = (v_1|u_1|v_2|u_2)$$

one obtains that

$$AT = T \begin{pmatrix} a & -b & 1 & 0 \\ b & a & 0 & 1 \\ 0 & 0 & a & -b \\ 0 & 0 & b & -a \end{pmatrix} .$$

This shows that the real and imaginary parts of eigenvectors and generalized eigenvectors play a role when one transforms A to real Jordan normal form.

2.4 An Estimate of e^{At}

The following is Theorem 14.9 proved in Appendix I.

Theorem 2.1 *Let $A \in \mathbb{C}^{n\times n}$ and assume that*

$$\mathrm{Re}\,\lambda_j < -\alpha < 0 \quad \textit{for all} \quad \lambda_j \in \sigma(A) \ .$$

Then there exists a constant $K \geq 1$ so that

$$|e^{At}| \leq K e^{-\alpha t} \quad \textit{for all} \quad t \geq 0 \ .$$

2.5 The System $x' = Ax + b(t)$

Let $A \in \mathbb{R}^{n\times n}$ and let $b : \mathbb{R} \to \mathbb{R}^n$ denote a continuous function. The IVP

$$x' = Ax + b(t), \quad x(0) = x_0 \ ,$$

has the solution

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}b(s)\,ds \ . \tag{2.20}$$

2.6 Application: Asymptotic Stability of a Fixed Point

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ denote a C^1–map and consider the system

$$x' = f(x) \quad \text{for} \quad x(t) \in \mathbb{R}^n \ .$$

If $f(x)$ is a nonlinear function, then the ODE system $x' = f(x)$ is called nonlinear. In Chapter 5 we will give local and global existence and uniqueness results for the IVP

$$x' = f(x), \quad x(0) = x_0 \ .$$

We denote the solution of the IVP by $x(t) = x(t, x_0)$.

A point $x^* \in \mathbb{R}^n$ is called a fixed point of the system $x' = f(x)$ if $f(x^*) = 0$.

Definition: A fixed point x^* is called stable if for all $\varepsilon > 0$ there exists $\delta > 0$ so that

$$|x_0 - x^*| \leq \delta \quad \text{implies} \quad |x(t, x_0) - x^*| \leq \varepsilon \quad \text{for all} \quad t \geq 0 \ .$$

A fixed point is called asymptotically stable if x^* is stable and there exists $\delta > 0$ so that

$$|x_0 - x^*| \leq \delta \quad \text{implies} \quad \lim_{t\to\infty} x(t, x_0) = x^* \ .$$

An important theorem:

Theorem 2.2 *Let $f : \mathbb{R}^n \to \mathbb{R}^n, f \in C^2$ and let $f(x^*) = 0$. Assume that every eigenvalue λ of the Jacobian $A = f'(x^*)$ satisfies $\mathrm{Re}\,\lambda < 0$. Then the fixed point x^* is asymptotically stable.*

Proof: Without loss of generality, we may assume that $x^* = 0$. We write the system $x' = f(x)$ in the form

$$x' = Ax + Q(x)$$

with $A = f'(0)$ and

$$|Q(x)| \le C|x|^2 \quad \text{for all} \quad x \in \mathbb{R}^n \quad \text{with} \quad |x| \le 1 \; .$$

(See Lemma 2.3.)

Choose $\alpha > 0$ so that

$$\text{Re}\,\lambda_j < -\alpha < 0 \quad \text{for all} \quad \lambda_j \in \sigma(A) \; .$$

By Theorem 2.1 there exists a constant $K \ge 1$ so that

$$|e^{At}| \le Ke^{-\alpha t} \quad \text{for all} \quad t \ge 0 \; .$$

We claim: If $\varepsilon > 0$ is sufficiently small and $0 < |x_0| \le \varepsilon$ then the solution of the nonlinear equation $x' = Ax + Q(x)$ satisfies

$$|x(t)| < 2Ke^{-\alpha t}|x_0| \quad \text{for} \quad t \ge 0 \; . \tag{2.21}$$

To prove this, let $0 < |x_0| = \varepsilon$ be so small that

$$2K\varepsilon \le 1$$

and

$$\varepsilon < \frac{\alpha}{4K^2C} \; . \tag{2.22}$$

We claim that the solution of the nonlinear equation satisfies

$$|x(t)| < 2K\varepsilon e^{-\alpha t} \quad \text{for all} \quad t \ge 0 \; .$$

If this does not hold, then let $T > 0$ denote the smallest time where equality occurs. We then have (using 2.20):

$$\begin{aligned} 2K\varepsilon e^{-\alpha T} &= |x(T)| \\ &\le K\varepsilon e^{-\alpha T} + \int_0^T |e^{A(T-s)}||Q(x(s))|\,ds \end{aligned}$$

Here

$$|e^{A(T-s)}| \le Ke^{-\alpha(T-s)} \quad \text{for} \quad 0 \le s \le T$$

and

$$|Q(x(s))| \le C4K^2\varepsilon^2 e^{-2\alpha s} \quad \text{for} \quad 0 \le s \le T \; .$$

Obtain:

$$
\begin{aligned}
K\varepsilon e^{-\alpha T} &\le \int_0^T |e^{A(T-s)}||Q(x(s))|\,ds \\
&\le C4K^2\varepsilon^2 \cdot Ke^{-\alpha T}\int_0^T e^{-\alpha s}\,ds \\
&\le 4CK^3\varepsilon^2 e^{-\alpha T}\cdot \frac{1}{\alpha}
\end{aligned}
$$

This yields that

$$1 \le \frac{4CK^2\varepsilon}{\alpha}\ ,$$

thus

$$\frac{\alpha}{4CK^2} \le \varepsilon\ .$$

This contradicts (2.22).

Since the estimate $|x(t)| < 2K\varepsilon e^{-\alpha t}, t \ge 0$, holds for all small $\varepsilon > 0$, it is not difficult to complete the proof of asymptotic stability of the fixed point $x^* = 0$. ◇

Lemma 2.2 *Let* $\phi : [0,1] \to \mathbb{R}, \phi \in C^2$. *Assume that* $\phi(0) = \phi'(0) = 0$ *and let*

$$|\phi''(s)| \le c_0 \quad \textit{for} \quad 0 \le s \le 1\ .$$

Then the estimes

$$|\phi(1)| \le \frac{c_0}{2}$$

holds.

Proof: For $0 \le t \le 1$ we have

$$\phi'(t) = \int_0^t \phi''(s)\,ds, \quad |\phi'(t)| \le tc_0\ .$$

Therefore,

$$\phi(1) = \int_0^1 \phi'(t)\,dt, \quad |\phi(1)| \le c_0\int_0^1 t\,dt = \frac{c_0}{2}\ .$$

◇

Lemma 2.3 *Let* $q : \mathbb{R}^n \to \mathbb{R}, q \in C^2$. *Assume that*

$$q(0) = 0, \quad \nabla q(0) = 0$$

and let

$$q''(x) = \Big(D_iD_jq(x)\Big)_{1\le i,j\le n}$$

denote the Hessian of q *at* $x \in \mathbb{R}^n$*. (Here* $D_j = \partial/\partial x_j$*.) Assume*

$$|q''(x)| \le c_0 \quad \textit{for} \quad |x| \le 1 \ .$$

Then the estimates

$$|q(x)| \le \frac{c_0}{2}\,|x|^2 \quad \textit{for} \quad |x| \le 1$$

holds.

Proof: Fix $x \in \mathbb{R}^n$ with $|x| \le 1$ and set

$$\phi(t) = q(tx) \quad \text{for} \quad 0 \le t \le 1 \ .$$

We have $\phi(0) = q(0) = 0$ and

$$\phi'(t) = x \cdot (\nabla q)(tx) \ ,$$

thus $\phi'(0) = 0$. Also,

$$\phi'(t) = x \cdot (\nabla q)(tx) = \sum_j x_j (D_j q)(tx)$$

yields that

$$\phi''(t) = \sum_i \sum_j x_i x_j (D_i D_j q)(tx) = x^T q''(tx) x = \langle x, q''(tx)x \rangle \ .$$

Therefore, for $0 \le t \le 1$,

$$|\phi''(t)| \le c_0 |x|^2 \ .$$

Applying the previous lemma we obtain that

$$|q(x)| = |\phi(1)| \le \frac{c_0}{2}\,|x|^2 \ .$$

$\diamond$

Figures for Chapter 2

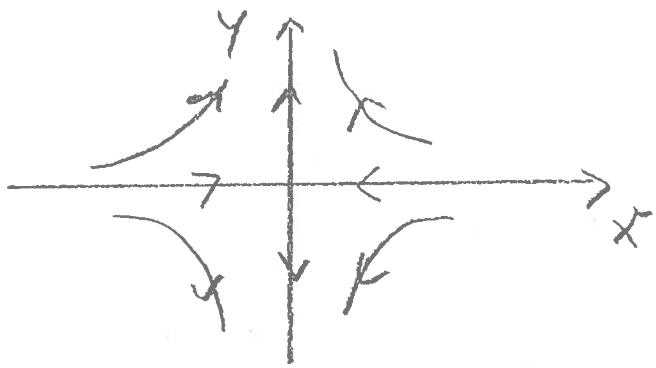

Figure 2.1

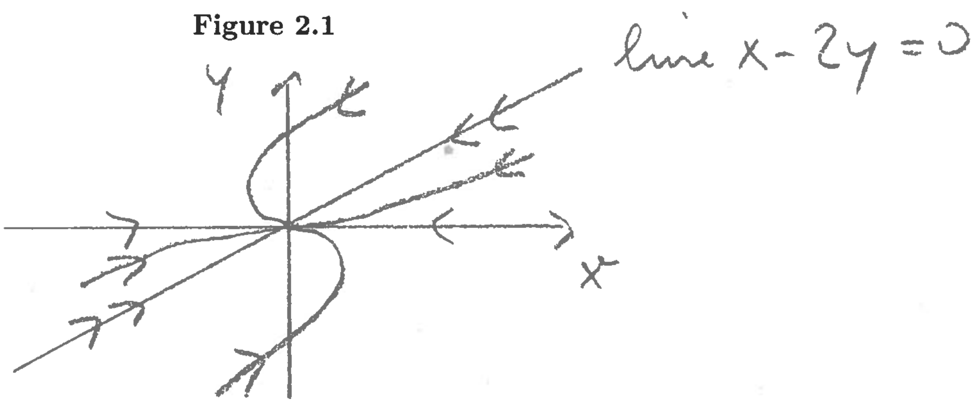

Figure 2.2

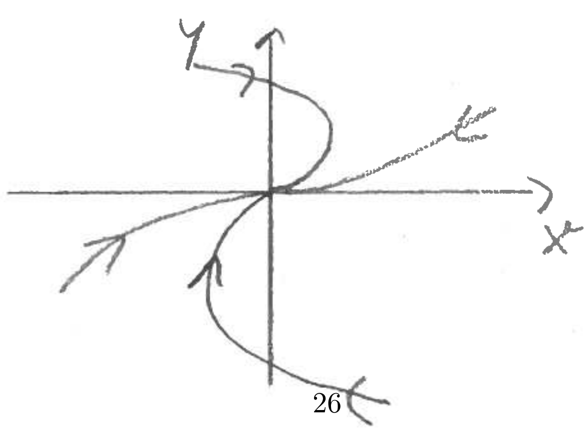

Figure 2.3

3 First Order Linear Systems $x' = A(t)x + b(t)$

3.1 Fundamental Matrices; Liouville's Theorem

Let $A(t)$ denote a continuous matrix function taking values in $\mathbb{R}^{n\times n}$ for $0 \le t \le T$.

Consider the IVP

$$x' = A(t)x, \quad x(0) = a \ , \tag{3.1}$$

where $a \in \mathbb{R}^n$ is a given vector. Theorem 5.2 implies that the IVP has a unique solution $x(t) = x(t, a)$.

Let $a^1, \ldots, a^n$ denote a basis of $\mathbb{R}^n$ and let $x(t, a^j)$ denote the solution of the IVP

$$x' = A(t)x, \quad x(0) = a^j \quad \text{for} \quad j = 1, \ldots, n \ .$$

We form the matrix $\Phi(t) \in \mathbb{R}^{n\times n}$ whose j–th column is $x(t, a^j)$:

$$\Phi(t) = \Big(x(t, a^1)| \ldots |x(t, a^n)\Big), \quad 0 \le t \le T \ . \tag{3.2}$$

It is easy to check that

$$\Phi'(t) = A(t)\Phi(t), \quad \Phi(0) = \Big(a^1| \ldots |a^n\Big) \ .$$

Definition: *A matrix function $\Phi(t) \in \mathbb{R}^{n\times n}$ is called a fundamental matrix for the system $x' = A(t)x$ in $0 \le t \le T$ if $\Phi \in C^1[0, T]$ and*

$$\Phi'(t) = A(t)\Phi(t), \quad det\ \Phi(t) \neq 0 \quad for \quad 0 \le t \le T \ .$$

The normalized fundamental matrix $\Phi_0(t)$ is the fundamental matrix satisfying $\Phi_0(0) = I$.

Liouville's Theorem, which we prove below, implies that a matrix function $\Phi \in C^1[0, T]$ is a fundamental matrix if the differential equation $\Phi'(t) = A(t)\Phi(t)$ holds and if det $\Phi(t_0) \neq 0$ for some $0 \le t_0 \le T$.

Theorem 3.1 *(Liouville) Let $\Phi'(t) = A(t)\Phi(t)$ for $0 \le t \le T$. Then we have*

$$\frac{d}{dt}\, det\ \Phi(t) = tr\ A(t) \cdot det\ \Phi(t) \ , \tag{3.3}$$

and, therefore,

$$det\ \Phi(t) = det\ \Phi(0) \cdot \exp\Big(\int_0^t tr\ A(s)\, ds\Big) \ . \tag{3.4}$$

Proof: We use the notation

$$\Phi(t) = \Big(\phi_{ij}(t)\Big)_{1\le i,j\le n}$$

and have

$$\det\,\Phi(t) = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\,\phi_{1\sigma_1}(t)\cdots\phi_{n\sigma_n}(t)\ .$$

Here S_n denotes the group of all permutations of the set $\{1,2,\ldots,n\}$.

Consider the derivative of one term:

$$\frac{d}{dt}\Big(\phi_{1\sigma_1}(t)\cdots\phi_{n\sigma_n}(t)\Big)\ .$$

This derivative is the sum of n terms, where in each term exactly one factor is differentiate:

$$\begin{aligned}\frac{d}{dt}\Big(\phi_{1\sigma_1}\phi_{2\sigma_2}\cdots\phi_{n\sigma_n}\Big) &= \phi'_{1\sigma_1}\phi_{2\sigma_2}\ldots\phi_{n\sigma_n}+\phi_{1\sigma_1}\phi'_{2\sigma_2}\ldots\phi_{n\sigma_n}+\ldots\\ &\quad+\phi_{1\sigma_1}\phi_{2\sigma_2}\ldots\phi'_{n\sigma_n}\end{aligned}$$

Multiplying by $\operatorname{sgn}\sigma$ and summing over all $\sigma\in S_n$ one obtains:

$$\begin{aligned}\frac{d}{dt}\det\,\Phi(t) &= \begin{vmatrix}\phi'_{11} & \phi'_{12} & \ldots & \phi'_{1n}\\ \phi_{21} & \phi_{22} & \ldots & \phi_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \phi_{n1} & \phi_{n2} & \ldots & \phi_{nn}\end{vmatrix} + \begin{vmatrix}\phi_{11} & \phi_{12} & \ldots & \phi_{1n}\\ \phi'_{21} & \phi'_{22} & \ldots & \phi'_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \phi_{n1} & \phi_{n2} & \ldots & \phi_{nn}\end{vmatrix} + \ldots\\ &\quad+\begin{vmatrix}\phi_{11} & \phi_{12} & \ldots & \phi_{1n}\\ \phi_{21} & \phi_{22} & \ldots & \phi_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \phi'_{n1} & \phi'_{n2} & \ldots & \phi'_{nn}\end{vmatrix}\end{aligned}$$

Consider the first term on the right–hand side of the above equation. Using that $\Phi' = A\Phi$ one obtains that

$$\phi'_{1j} = \sum_{k=1}^{n} a_{1k}\phi_{kj}\quad\text{for}\quad j=1,\ldots,n\ .$$

Therefore,

$$\begin{vmatrix}\phi'_{11} & \phi'_{12} & \ldots & \phi'_{1n}\\ \phi_{21} & \phi_{22} & \ldots & \phi_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \phi_{n1} & \phi_{n2} & \ldots & \phi_{nn}\end{vmatrix} = \sum_{k=1}^{n} a_{1k}\begin{vmatrix}\phi_{k1} & \phi_{k2} & \ldots & \phi_{kn}\\ \phi_{21} & \phi_{22} & \ldots & \phi_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \phi_{n1} & \phi_{n2} & \ldots & \phi_{nn}\end{vmatrix}\ .$$

For $k=2,\ldots,n$ the determinant in the above sum is zero since the first row agrees with the k–th row. Therefore,

$$\begin{vmatrix}\phi'_{11} & \phi'_{12} & \ldots & \phi'_{1n}\\ \phi_{21} & \phi_{22} & \ldots & \phi_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \phi_{n1} & \phi_{n2} & \ldots & \phi_{nn}\end{vmatrix} = a_{11}\det\,\Phi\ .$$

The other terms in the sum–formula for $\frac{d}{dt}\det\,\Phi$ are treated in the same way. One obtains that

$$\frac{d}{dt}\,\det\,\Phi(t) = (a_{11} + \ldots + a_{nn})\det\,\Phi\ .$$

$\diamond$

Recall the following elementary result:

Lemma 3.1 *Let $\alpha \in C^1[0,T], \beta \in C[0,T]$ denote scalar functions with*

$$\frac{d}{dt}\,\alpha(t) = \beta(t)\alpha(t), \quad 0 \le t \le T\ .$$

Then

$$\alpha(t) = \alpha(0)\exp\Big(\int_0^t \beta(s)\,ds\Big), \quad 0 \le t \le T\ . \tag{3.5}$$

Note that Lemma 3.1 yields formula (3.4) of Liouville's Theorem.

A simple implication of formula (3.5) is that either $\alpha(t) \equiv 0$ or $\alpha(t) \neq 0$ for all $0 \le t \le T$. Therefore, if the matrix function $\Phi(t)$ is constructed as in (3.2), where $a^1, \ldots, a^n$ are linearly independent, then $\det \Phi(t) \neq 0$ for $0 \le t \le T$ and $\Phi(t)$ is a fundamental matrix for the system $x' = A(t)x$.

3.2 Inhomogeneous Linear Systems

Let $A(t) \in \mathbb{R}^{n\times n}$ and $b(t) \in \mathbb{R}^n$ depend continuously on $0 \le t \le T$.

Lemma 3.2 *Let $\Phi(t)$ denote a fundamental matrix for the system $x' = A(t)x$. The IVP*

$$x' = A(t)x + b(t), \quad x(0) = a\ ,$$

has the solution

$$x(t) = \Phi(t)\Phi(0)^{-1}a + \Phi(t)\int_0^t \Phi(s)^{-1}b(s)\,ds, \quad 0 \le t \le T\ . \tag{3.6}$$

Proof: Let $x(t)$ be defined by (3.6). Then $x(0) = a$ and, since $\Phi' = A\Phi$,

$$\begin{aligned} x'(t) &= A(t)\Phi(t)\Phi(0)^{-1}a + A(t)\Phi(t)\int_0^t \Phi(s)^{-1}b(s)\,ds + \Phi(t)\Phi(t)^{-1}b(t) \\ &= A(t)x(t) + b(t) \end{aligned}$$

$\diamond$

3.3 Difficulties with Localization

First let $A \in \mathbb{R}^{n\times n}$ denote a constant matrix and assume that all eigenvalues of A have negative real parts. We know that all solutions $x(t)$ of the equation $x' = Ax$ decay exponentially as $t \to \infty$. Now consider a variable coefficient system

$$x' = A(t)x \quad \text{for} \quad t \geq 0$$

and assume that for every fixed $0 \leq t_0 < \infty$ all eigenvalues of the matrix $A(t_0)$ have negative real parts. Can one conclude that the solutions of the variable coefficients system $x' = A(t)x$ decay to 0 as $t \to \infty$? The answer is *No*, in general, as the following example shows.

Example: Consider the system $x' = A(t)x$ where

$$A(t) = U(t)A_0U^T(t), \quad A_0 = \begin{pmatrix} \alpha & \beta \\ 0 & \gamma \end{pmatrix}, \quad U(t) = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

with real constants α, β, γ and

$$c = \cos t, \quad s = \sin t \ .$$

It is clear that the eigenvalues of $A(t)$ are α and γ.

Let us assume that

$$\alpha < 0 \quad \text{and} \quad \gamma < 0 \ .$$

Then, if we freeze the coefficient matrix $A(t)$ at any time t_0 and consider the constant coefficient system

$$x' = A(t_0)x \ ,$$

the solution will decay exponentially.

To understand the variable coefficient equation, introduce a new variable $y(t)$ by

$$x(t) = U(t)y(t), \quad x' = Uy' + U'y \ .$$

The equation $x' = Ax$ becomes

$$Uy' + U'y = AUy = UA_0y \ ,$$

or

$$y' = (A_0 - U^TU')y \ .$$

It is easy to compute that

$$U^TU' = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} -s & -c \\ c & -s \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \ .$$

Therefore, one obtains that $y' = By$ with

$$B = A_0 - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \alpha & \beta + 1 \\ -1 & \gamma \end{pmatrix} .$$

The eigenvalues λ of B satisfy

$$\begin{aligned} 0 &= \det(B - \lambda I) \\ &= (\alpha - \lambda)(\gamma - \lambda) + \beta + 1 \\ &= \lambda^2 - (\alpha + \gamma)\lambda + \alpha\gamma + \beta + 1 \end{aligned}$$

thus

$$\lambda_{1,2} = \frac{1}{2}(\alpha + \gamma) \pm \sqrt{\frac{1}{4}(\alpha + \gamma)^2 - \alpha\gamma - \beta - 1} .$$

Therefore, assuming that $-\beta$ is sufficiently large, the matrix B has a positive eigenvalue. Then the function $y(t)$ typically grows exponentially and the function $x(t) = U(t)y(t)$ also typically grows exponentially.

4 Floquet Theory

In this chapter we consider a variable coefficient linear system

$$x' = A(t)x, \quad x(t) \in \mathbb{R}^n \ , \tag{4.1}$$

where $A(t) \in \mathbb{R}^{n\times n}$ is a continuous, periodic matrix function, $A(t+T) \equiv A(t)$. Here $T > 0$ is the period. We will show that the normalized fundamental matrix $\Phi_0(t)$ can be written in the form

$$\Phi_0(t) = Q(t)e^{Bt}, \quad t \in \mathbb{R} \ ,$$

where the matrix function $Q(t)$ has period T and where $B \in \mathbb{C}^{n\times n}$ is constant. (The matrix B is not unique.)

The matrix

$$M := \Phi_0(T) = e^{BT}$$

is called the monodromy matrix of the system $x' = A(t)x$. The eigenvalues of M are

$$\mu_j = e^{\eta_j T}, \quad j = 1, 2, \ldots, n \ ,$$

where the η_j are the eigenvalues of B. The μ_j are called the Floquet multipliers and the η_j are the Floquet exponents of the system $x' = A(t)x$.

We will prove: If $|\mu_j| < 1$ for $j = 1, 2, \ldots, n$ then all solutions $x(t)$ of (4.1) decay to zero as $t \to \infty$.

Later we will consider nonlinear systems $x' = f(x)$ with periodic solutions, $u(t+T) \equiv u(t)$. The Floquet multipliers of the linearized system

$$y' = Df(u(t))y$$

play a role in the stability analysis of the periodic solution $u(t)$.

4.1 Auxiliary Results from Linear Algebra

Lemma 4.1 *For $A \in \mathbb{C}^{n\times n}$ we have*

$$det\, e^A = e^{tr\, A} \ .$$

Proof: The matrix e^{At} is the normalized fundamental matrix of the system $x' = Ax$. By Theorem 3.1 we have

$$\frac{d}{dt}\det e^{At} = (\mathrm{tr}\, A)\det e^{At} \ ,$$

therefore

$$\det e^{At} = e^{(\mathrm{tr}\, A)t} \ .$$

For $t = 1$ obtain that $\det e^A = e^{\mathrm{tr}\, A}$. $\diamond$

Lemma 4.2 *Let $R \in \mathbb{C}^{n\times n}$ be nilpotent, i.e., $R^n = 0$. Then*

$$B = \sum_{j=1}^{n-1} \frac{(-1)^{j+1}}{j}\, R^j$$

satisfies

$$e^B = I + R \; .$$

Proof: For $z \in \mathbb{C}, |z| < 1$, let

$$f(z) = \log(1+z) \; ,$$

thus

$$\begin{aligned} f'(z) &= \frac{1}{1+z} \\ &= \sum_{k=0}^{\infty} (-1)^k z^k \\ f(z) &= \log(1+z) \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1}\, z^{k+1} \\ &= \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}\, z^j \end{aligned}$$

We have

$$e^{f(z)} = e^{\log(1+z)} = 1 + z \quad \text{for} \quad |z| < 1 \; .$$

Since $R^n = 0$ we have

$$B = \sum_{j=1}^{n-1} \frac{(-1)^{j+1}}{j}\, R^j = \log(I+R) \; .$$

Since the series for $\log(I+R)$ is finite we can substitute R for z in the identity

$$e^{\log(1+z)} = 1 + z$$

and obtain

$$e^B = I + R \; .$$

Lemma 4.3 *Let $A \in \mathbb{C}^{n\times n}$ with $det\, A \neq 0$. Then there exists a (non–unique) matrix $B \in \mathbb{C}^{n\times n}$ with*

$$A = e^B \; .$$

Proof: Using Schur's Lemma and blocking we obtain a transformation of A to blockdiagonal form,

$$S^{-1}AS = \mathrm{diag}\,(A_j)$$

where

$$A_j = \lambda(I_k + R_k), \quad R_k^k = 0, \quad \lambda \neq 0 \ .$$

By the previous lemma there exists $B_j \in \mathbb{C}^{k\times k}$ with

$$e^{B_j} = I_k + R_k \ .$$

Also, since $\lambda \neq 0$, there exists $\beta \in \mathbb{C}$ with $e^\beta = \lambda$. We then have

$$e^{\beta I_k + B_j} = e^\beta e^{B_j} = \lambda(I_k + R_k) = A_j \ .$$

Since every block A_j can be written in this way, the claim follows. $\diamond$

Example: Let

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \ .$$

We want to construct a matrix B with $e^B = A$.

The matrix A has the eigenvalues $\pm i$ and there exists a nonsingular matrix $S \in \mathbb{C}^{2\times 2}$ with

$$S^{-1}AS = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} =: \Lambda \ .$$

If we set

$$D = \begin{pmatrix} \pi i/2 & 0 \\ 0 & -\pi i/2 \end{pmatrix}$$

then

$$e^D = \begin{pmatrix} e^{\pi i/2} & 0 \\ 0 & e^{-\pi i/2} \end{pmatrix} = \Lambda \ .$$

We set

$$B = SDS^{-1}$$

and obtain

$$e^B = Se^D S^{-1} = S\Lambda S^{-1} = A \ .$$

Remark 1: The matrix B can be obtained explicitly: We have $D = \frac{\pi}{2}\,\Lambda$, thus

$$B = SDS^{-1} = \frac{\pi}{2}\, S\Lambda S^{-1} = \frac{\pi}{2}\, A \ ,$$

thus

$$B = \begin{pmatrix} 0 & -\pi/2 \\ \pi/2 & 0 \end{pmatrix} \ .$$

Remark 2: The matrix B with $e^B = A$ is non–unique. Instead of the matrix D given above we can choose

$$D = \begin{pmatrix} \pi i/2 + 2\pi i k_1 & 0 \\ 0 & -\pi i/2 + 2\pi i k_2 \end{pmatrix}$$

where $k_1, k_2 \in \mathbb{Z}$ are arbitrary. One obtains $e^D = \Lambda$ and the matrix $B = SDS^{-1}$ satisfies $e^B = A$.

Lemma 4.4 *Let $A, B \in \mathbb{C}^{n\times n}$ and let $A = e^B$. Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of B, listed by their algebraic multiplicities, i.e.,*

$$det\,(B - zI) = \Pi_{j=1}^n (\lambda_j - z) \ .$$

Then the numbers

$$e^{\lambda_1}, \ldots, e^{\lambda_n}$$

are the eigenvalues of A, listed by their algebraic multiplicities.

Proof: By Schur's Theorem there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ so that

$$U^* B U = \Lambda + R$$

where $\Lambda = \mathrm{diag}\,(\lambda_j)$ and R is strictly upper triangular. We have

$$e^{\Lambda + R} = I + (\Lambda + R) + \frac{1}{2}\,(\Lambda + R)^2 + \ldots = e^\Lambda + \tilde{R}$$

where $\tilde{R}$ is strictly upper triangular. To obtain this, note that

$$(\Lambda + R)^2 = \Lambda^2 + \Lambda R + R\Lambda + R^2$$

where

$$\Lambda R + R\Lambda + R^2$$

is strictly upper triangular. The same argument works for all powers of $\Lambda + R$. One obtains that A is similar to

$$e^\Lambda + \tilde{R} = \mathrm{diag}\,(e^{\lambda_j}) + \tilde{R} \ .$$

$\diamond$

4.2 Linear Systems with Periodic Matrix Functions

We consider systems

$$x' = A(t)x$$

where $A(t+T) = A(t)$ for all $t \in \mathbb{R}$ for some fixed $T > 0$. As before, $A : \mathbb{R} \to \mathbb{R}^{n\times n}$ is continuous.

The example

$$A(t) = \begin{pmatrix} \cos t & 0 \\ 0 & 1 \end{pmatrix}, \quad \Phi_0(t) = \begin{pmatrix} e^{\sin t} & 0 \\ 0 & e^t \end{pmatrix}$$

shows that we cannot expect fundamental matrices to be periodic.

Theorem 4.1 *Let $\Phi_0(t)$ denote the normalized fundamental matrix of the system $x' = A(t)x$. There exists a matrix function $Q \in C^1(\mathbb{R})$ with $Q(t+T) \equiv Q(t)$ and there exists a constant matrix B so that*

$$\Phi_0(t) = Q(t)e^{Bt} \ .$$

Proof: Set

$$V(t) = \Phi_0(t+T)\Phi_0^{-1}(T) \ .$$

Then $V(0) = I$ and

$$\begin{aligned} V'(t) &= A(t+T)\Phi_0(t+T)\Phi_0^{-1}(T) \\ &= A(t)V(t) \end{aligned}$$

It follows that $V(t) = \Phi_0(t)$. In other words,

$$\Phi_0(t+T) = \Phi_0(t)\Phi_0(T) \ . \tag{4.2}$$

Now let

$$\Phi_0(T) = e^{BT} \tag{4.3}$$

and set

$$Q(t) = \Phi_0(t)e^{-Bt} \ .$$

It is then clear that

$$\Phi_0(t) = Q(t)e^{Bt}$$

and

$$\begin{aligned} Q(t+T) &= \Phi_0(t+T)e^{-BT}e^{-Bt} \quad \text{(use (4.2) and (4.3))} \\ &= \Phi_0(t)e^{-Bt} \\ &= Q(t) \end{aligned}$$

This proves the theorem. $\diamond$

Definition: *The matrix* $M = \Phi_0(T) = e^{BT}$ *is called the monodromy matrix of the system* $x' = A(t)x$ *with respect to the period* T*. The eigenvalues* μ_j *of* $\Phi_0(T)$ *are called the Floquet multipliers of the system* $x' = A(t)x$*. If* η_j *denotes the eigenvalues of* B*, then* $\mu_j = e^{T\eta_j}$ *are the Floquet multipliers. The numbers* η_j *are called Floquet exponents of the system. (The* η_j *are not uniquely determined. One can always add* $2\pi i n_j/T$ *to* η_j *where* n_j *is an integer.)*

Transformation to a constant coefficient system: Consider the system

$$x' = A(t)x, \quad x(0) = x_0 \ .$$

The solution is

$$x(t) = Q(t)e^{Bt}x_0 \ .$$

Set

$$y(t) = e^{Bt}x_0 \ .$$

We then have

$$x(t) = Q(t)y(t)$$

and

$$y' = By, \quad y(0) = x_0 \ .$$

We can express this as follows: If we introduce the variable $y(t)$ by $x(t) = Q(t)y(t)$ then the variable $y(t)$ solves the constant coefficient system

$$y' = By, \quad y(0) = x_0 \ .$$

Modulo a periodic change of variables, $x(t) = Q(t)y(t)$, we obtain a constant coefficient system for the variable $y(t)$.

Theorem 4.2 *Assume that the Floquet multipliers* μ_j *satisfy* $|\mu_j| < 1, j = 1, 2, \ldots, n$*. Then all solutions* $x(t)$ *of the system* $x' = A(t)x$ *converge to zero as* $t \to \infty$*.*

Proof: This is clear since $y(t) \to 0$ as $t \to \infty$ and the periodic matrix function $Q(t)$ is bounded.

Example: The following simple example illustrates the non–uniqueness of B. Consider the scalar IVP

$$x' = (1 + \cos t)x, \quad x(0) = 1 \ .$$

Here $n = 1$ and $A(t) = 1 + \cos t$ has the period $T = 2\pi$. Obtain

$$\int_1^{x(t)} \frac{dx}{x} = \int_0^t (1 + \cos s)\, ds = t + \sin t \ ,$$

thus

$$x(t) = e^{t+\sin t} \ .$$

In this case the normalized fundamental matrix is the scalar function

$$\Phi_0(t) = e^{\sin t} e^t \ .$$

The monodromy matrix is

$$\Phi_0(2\pi) = e^{2\pi} = e^{2\pi(1+ik)}, \quad k \in \mathbb{Z} \ .$$

Clearly, we have

$$\Phi_0(t) = e^{\sin t} e^t = Q(t) e^{Bt}$$

where

$$Q(t) = e^{\sin t}$$

has period 2π and $B = 1$. However, we can also write

$$\Phi_0(t) = e^{\sin t - ikt} e^{(1+ik)t} \quad \text{where} \quad k \in \mathbb{Z} \ .$$

Here

$$Q(t) = e^{\sin t - ikt}$$

has period 2π and $B = 1 + ik$.

4.3 A First Result on Periodic Solutions of Nonlinear Systems

Let $f : \mathbb{R}^n \to \mathbb{R}^n$, $f \in C^1$, and consider the system

$$x' = f(x) \ .$$

Assume that $u(t)$ is a solution with minimal period $T > 0$. In particular, $u(t)$ is not constant. It follows that

$$f(u(0)) \neq 0 \ .$$

Let

$$A(t) = Df(u(t)) \ .$$

The matrix function $A(t)$ has the period T, but it is possible that T is not the minimal period of the function $A(t)$. This will not play any role for the following, however.

The system

$$y' = A(t) y$$

is obtained by linearization about $u(t)$. With

$$\Phi_0(t)$$

we denote the normalized fundamental matrix and with

$$M = \Phi_0(T) = e^{BT}$$

we denote the monodromy matrix of the linear system $y' = A(t)y$.

Theorem 4.3 *The monodromy matrix M has the eigenvalue $\mu_1 = 1$ with corresponding eigenvector $f(u(0)) = u'(0)$.*

Proof: From $u'(t) = f(u(t))$ obtain that

$$u''(t) = A(t)u'(t) \ .$$

Therefore,

$$\begin{aligned} u'(0) &= u'(T) \\ &= \Phi_0(T)u'(0) \\ &= Mu'(0) \ . \end{aligned}$$

$\diamond$

5 Basic Existence and Uniqueness Results

5.1 Example of Non–Uniqueness

Let

$$f(x) = 0 \quad \text{for} \quad x < 0 \quad \text{and} \quad f(x) = \sqrt{x} \quad \text{for} \quad x \geq 0 \ .$$

Consider the IVP

$$x' = f(x), \quad x(0) = 0 \ .$$

Clearly, $x \equiv 0$ is a solution. Ansatz for another solution:

$$x(t) = ct^\alpha \ .$$

One obtains a solution for

$$\alpha = 2, \quad c = \frac{1}{4} \ .$$

In fact, if $t_0 > 0$, then the function

$$x(t) = \begin{cases} 0 & \text{for} \quad 0 \leq t \leq t_0 \\ \frac{1}{4}(t - t_0)^2 & \text{for} \quad t > t_0 \end{cases}$$

is a solution.

The example shows the importance of a Lipschitz condition for the function $f(x)$ to obtain uniqueness of a solution of the IVP

$$x' = f(x), \quad x(0) = x_0 \ .$$

5.2 Completeness of $(C[0,T], |\cdot|_\infty)$

The result shown here is important to prove the existence of a solution of the IVP

$$x' = f(x,t), \quad x(0) = a \ ,$$

where $f : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ is a continuous function which is Lipschitz w.r.t. the state variable x.

Let $C[0,T]$ denote the vector space of all continuous functions $x : [0,T] \to \mathbb{R}^n$. Let $|\cdot|$ denote the Euclidean norm on $\mathbb{R}^n$. The maximum norm on $C[0,T]$ is defined by

$$|x|_\infty = \max\{|x(t)| \ : \ 0 \leq t \leq T\} \ .$$

Theorem 5.1 *The normed space* $\Big(C[0,T], |\cdot|_\infty\Big)$ *is complete.*

Proof: Let $x_j = x_j(t)$ denote a Cauchy sequence in $C[0,T]$ w.r.t. $|\cdot|_\infty$. Thus, for every $\varepsilon > 0$ there exists $J(\varepsilon) \in \mathbb{N}$ so that

$$|x_j - x_k|_\infty \le \varepsilon \quad \text{for} \quad j > k \ge J(\varepsilon) \ .$$

Since $\mathbb{R}^n$ is complete w.r.t. $|\cdot|$ one obtains that for every $0 \le t \le T$ the sequence $x_j(t)$ converges,

$$x_j(t) \to x(t) \in \mathbb{R}^n \quad \text{as} \quad j \to \infty \ .$$

We claim that the limit function $x(t)$ is continuous.

Given $\varepsilon > 0$ there exists $k \in \mathbb{N}$ so that for all $j > k$:

$$|x_j(t) - x_k(t)| \le \frac{\varepsilon}{3} \quad \text{for all} \quad t \in [0,T] \ .$$

Since $x_j(t) \to x(t)$ as $j \to \infty$ obtain that

$$|x(t) - x_k(t)| \le \frac{\varepsilon}{3} \quad \text{for all} \quad t \in [0,T] \ .$$

Fix any $t_0 \in [0,T]$. We will prove that the limit function $x(t)$ is continuous in t_0. By assumption, x_k is continuous in t_0. Therefore, for given $\varepsilon > 0$ there exists $\delta > 0$ so that

$$|x(t_0) - x(t)| \le \frac{\varepsilon}{3} \quad \text{for} \quad |t_0 - t| \le \delta, \quad t \in [0,T] \ .$$

Obtain for $|t_0 - t| \le \delta$:

$$|x(t_0) - x(t)| \le |x(t_0) - x_k(t_0)| + |x_k(t_0) - x_k(t)| + |x_k(t) - x(t)| \le \varepsilon \ .$$

This shows that $x \in C[0,T]$. Also, given $\varepsilon > 0$ there exists $J(\varepsilon)$ so that for $j > k \ge J(\varepsilon)$:

$$|x_j(t) - x_k(t)| \le \varepsilon \quad \text{for all} \quad t \in [0,T] \ .$$

For $j \to \infty$ obtain that

$$|x(t) - x_k(t)| \le \varepsilon \quad \text{for all} \quad t \in [0,T]$$

for $k \ge J(\varepsilon)$. This proves that $x_k \to x$ w.r.t. $|\cdot|_\infty$. $\diamond$

5.3 IVPs: Existence and Uniqueness of Solutions

The proof of existence will be based on the following lemma.

Lemma 5.1 *Let $T > 0$ and let $L \ge 0$. Let $\eta_k : [0,T] \to [0,\infty)$ denote a sequence of non–negative continuous functions satisfying the estimates*

$$\eta_{k+1}(t) \le L \int_0^t \eta_k(s)\, ds \quad \textit{for} \quad 0 \le t \le T, \quad k = 0, 1, \ldots \tag{5.1}$$

Then we have

$$\eta_k(t) \le \frac{(Lt)^k}{k!} \max_{0\le s\le t} \eta_0(s) \quad \textit{for} \quad 0 \le t \le T \ . \tag{5.2}$$

In particular,

$$|\eta_k|_\infty \le \frac{(LT)^k}{k!} |\eta_0|_\infty \to 0 \quad \textit{as} \quad k \to \infty \ .$$

Proof: We use induction in k. For $k = 0$ the estimate (5.2) is trivial. Set

$$M(t) = \max_{0\le s\le t} \eta_0(s) \quad \text{for} \quad 0 \le t \le T$$

and assume that (5.2) holds for some k, thus

$$\eta_k(t) \le \frac{(Lt)^k}{k!} M(t) \quad \text{for} \quad 0 \le t \le T \ .$$

Using (5.1) obtain:

$$\begin{aligned} \eta_{k+1}(t) &\le L \cdot \frac{L^k}{k!} M(t) \int_0^t s^k \, ds \\ &= \frac{(Lt)^{k+1}}{(k+1)!} M(t) \end{aligned}$$

for $0 \le t \le T$. $\diamond$

5.3.1 The Case Where f Is Globally Lipschitz

Theorem 5.2 *Let $f : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ denote a continuous function and assume that there exists $L > 0$ so that*

$$|f(x,t) - f(y,t)| \le L|x-y| \quad \textit{for all} \quad 0 \le t \le T \quad \textit{and for all } x, y \in \mathbb{R}^n \ .$$

Then, for every $a \in \mathbb{R}^n$, the initial value problem

$$x' = f(x,t), \quad x(0) = a \ , \tag{5.3}$$

has a unique solution $x \in C^1[0,T]$.

Proof: *Existence:* We define a sequence of function $x_k \in C[0,T]$ with values in $\mathbb{R}^n$ as follows: Let $x_0(t) \equiv a$ and let

$$x_{k+1}(t) = a + \int_0^t f(x_k(s), s) \, ds \quad \text{for} \quad 0 \le t \le T, \quad k = 0, 1, 2, \dots \ . \tag{5.4}$$

We have

$$x_{k+2}(t) = a + \int_0^t f(x_{k+1}(s), s) \, ds \quad \text{for} \quad 0 \le t \le T$$

and obtain that

$$x_{k+2}(t) - x_{k+1}(t) = \int_0^t \Big(f(x_{k+1}(s), s) - f(x_k(s), s) \Big)\, ds \quad \text{for} \quad 0 \le t \le T \ .$$

Therefore,

$$|x_{k+2}(t) - x_{k+1}(t)| \le L \int_0^t |x_{k+1}(s) - x_k(s)|\, ds \quad \text{for} \quad 0 \le t \le T \ .$$

We set

$$\eta_k(t) = |x_{k+1}(t) - x_k(t)| \quad \text{for} \quad 0 \le t \le T$$

and obtain that

$$\eta_{k+1}(t) \le L \int_0^t \eta_k(s)\, ds \quad \text{for} \quad 0 \le t \le T, \quad k = 0, 1, 2, \ldots$$

Lemma 5.1 yields the estimates

$$|\eta_k|_\infty \le \frac{(LT)^k}{k!} |\eta_0|_\infty, \quad k = 0, 1, 2, \ldots$$

where

$$|\eta_k|_\infty = \max\{|\eta_k(t)| \ : \ 0 \le t \le T\} \ .$$

Since

$$\sum_{k=0}^\infty \frac{(LT)^k}{k!} < \infty$$

it follows that the sequence $x_k = x_k(t)$ is a Cauchy sequence in $C[0,T]$ w.r.t. the maximum norm. By Theorem 5.1 there exists $x \in C[0,T]$ with $|x_k - x|_\infty \to 0$ as $k \to \infty$. Taking the limit $k \to \infty$ in equation (5.4) one obtains that

$$x(t) = a + \int_0^t f(x(s), s)\, ds \quad \text{for} \quad 0 \le t \le T \ .$$

This implies that $x \in C^1[0,T]$ and x solves the IVP (5.3).

Uniqueness: Assume that the functions $x, y \in C^1[0,T]$ both solve the IVP (5.3). For $q = x - y$ obtain:

$$q(t) = x(t) - y(t) = \int_0^t \Big(f(x(s), s) - f(y(s), s) \Big)\, ds \ ,$$

thus

$$|q(t)| \le L \int_0^t |q(s))\, ds \quad \text{for} \quad 0 \le t \le T \ .$$

Applying Lemma 5.1 with $\eta_k(t) = |q(t)|$ for all k yields the estimates

$$|q(t)| \le \frac{(LT)^k}{k!} |q|_\infty \quad \text{for} \quad 0 \le t \le T \ .$$

Taking the limit $k \to \infty$ one obtains that $q \equiv 0$, thus $x \equiv y$. $\diamond$.

5.3.2 The Case Where f Is Locally Lipschitz

Theorem 5.3 *Let $f, f_x \in C(\mathbb{R}^n \times [0,T])$ and let $R > 0$. Consider the IVP*

$$x' = f(x,t), \quad x(0) = a \ , \tag{5.5}$$

where $|a| \le R$. The IVP has a unique solution in some interval $0 \le t \le \tau_R$ where $0 < \tau_R \le T$ depends on R, but is independent of $a \in B_R(0)$.

Proof: *Existence:* Choose a cut–off function $\phi_R \in C^\infty[0,\infty)$ which takes values in $[0,1]$ and satisfies

$$\phi_R(r) = \begin{cases} 1 & \text{for} \quad 0 \le t \le R+1 \\ 0 & \text{for} \quad t \ge R+2 \end{cases}$$

and set

$$\tilde{f}(x,t) = f(x,t)\phi_R(|x|) \ .$$

Set

$$M_R = \max\{|f(x,t)| \ : \ |x| \le R+1, 0 \le t \le T\}$$

and note that the function $\tilde{f}(x,t)$ is globally Lipschitz in x since there exists a constant $L_R > 0$ with

$$|\tilde{f}_x(x,t)| \le L_R \quad \text{for} \quad x \in \mathbb{R}^n \quad \text{and} \quad 0 \le t \le T \ .$$

By Theorem 5.2 the IVP

$$x'(t) = \tilde{f}(x,t), \quad x(0) = a \ ,$$

has a unique solution $x \in C^1[0,T]$. It is clear that $x(t)$ solves (5.5) in $0 \le t \le \tau$ as long as $|x(t)| \le R+1$ in $0 \le t \le \tau$.

Set

$$\tau_R = \min\{T, \frac{1}{M_R}\} \ .$$

We claim that $|x(t)| < R+1$ for $0 \le t < \tau_R$. If this does not hold then there exists t_0 with $0 < t_0 < \tau_R$ and

$$|x(t_0)| = R+1 \quad \text{and} \quad |x(t)| < R+1 \quad \text{for} \quad 0 \le t < t_0 \ .$$

Using that $|x(0)| = |a| \le R$ one obtains:

$$
\begin{aligned}
1 &\leq |x(t_0) - x(0)| \\
&= \Big| \int_0^{t_0} x'(t)\, dt \Big| \\
&\leq \int_0^{t_0} |x'(t)|\, dt \\
&= \int_0^{t_0} |f(x(t), t)|\, dt \\
&\leq t_0 M_R
\end{aligned}
$$

This implies that $t_0 \geq 1/M_R \geq \tau_R$, which contradicts $0 < t_0 < \tau_R$.

Uniqueness: Assume that $x(t)$ and $y(t)$ solve the IVP

$$x' = f(x,t), \quad x(0) = x_0$$

for $0 \leq t \leq \tau$. Choose $R > 0$ so that

$$|x(t)| \leq R \quad \text{and} \quad |y(t)| \leq R \quad \text{for} \quad 0 \leq t \leq \tau \ .$$

There exists $L > 0$ with

$$|f_x(x,t)| \leq L \quad \text{for} \quad |x| \leq R \quad \text{and} \quad 0 \leq t \leq \tau \ .$$

One then obtains that

$$|x(t) - y(t)| \leq L \int_0^t |x(s) - y(s)|\, ds \quad \text{for} \quad 0 \leq t \leq \tau \ .$$

As in the proof of Theorem 5.2 it follows that $x \equiv y$ in $[0, \tau]$. $\diamond$

5.4 The Maximal Interval of Existence and Possible Blow–Up

Let $f : \mathbb{R}^n \to \mathbb{R}^n$, $f \in C^1$, and let $a \in \mathbb{R}^n$. Consider the IVP

$$x' = f(x), \quad x(0) = a \ . \tag{5.6}$$

Set $R_0 = |a|$ and

$$M_0 = \max\{|f(x)| \ : \ |x| \leq R_0 + 1\} \quad \text{and} \quad \tau_0 = 1/M_0 \ .$$

By Theorem 5.3 the IVP (5.6) has a unique solution $x(t)$ defined for $0 \leq t \leq \tau_0$.

(If $M_0 = 0$ then, trivially, $x(t) \equiv a$ solves the IVP for $0 \leq t < \infty$. In the following we assume that $M_j > 0$.)

Set

$$a^{(1)} = x(\tau_0), \quad R_1 = |a^{(1)}|$$

and

$$M_1 = \max\{|f(x)| \ : \ |x| \leq R_1 + 1\}, \quad \tau_1 = 1/M_1 \ .$$

The initial value problem

$$x' = f(x), \quad x(\tau_0) = a^{(1)}$$

has a unique solution defined for $\tau_0 \le t \le \tau_0 + \tau_1$. Clearly, if we put the solution in $[0, \tau_0]$ and the solution in $[\tau_0, \tau_0 + \tau_1]$ together, we obtain a solution $x \in C^1[0, \tau_0 + \tau_1]$.

This process can be continued inductively. Will one always obtain a solution defined for $0 \le t < \infty$) ?

The simple example

$$x' = x^2, \quad x(0) = 1 \ ,$$

with solution

$$x(t) = \frac{1}{1-t} \quad \text{for} \quad 0 \le t < 1 \ ,$$

shows that the answer is *No*, in general.

Theorem 5.4 *Let $f : \mathbb{R}^n \to \mathbb{R}^n, f \in C^1$, and let $a \in \mathbb{R}^n$. Consider the IVP*

$$x' = f(x), \quad x(0) = a \ . \tag{5.7}$$

There are two cases:

Case 1: *A unique solution exists for $0 \le t < \infty$.*

Case 2: *A unique solution exists for $0 \le t < t^*$ where $0 < t^* < \infty$, but does not exist for $0 \le t \le t^*$. In this case,*

$$|x(t)| \to \infty \quad as \quad t \to t^*_- \ , \tag{5.8}$$

i.e., the solution blows up as t approaches t^*.

Proof: Using the process outlined above, one obtains a sequence of positive numbers $\tau_0, \tau_1, \ldots$ and, for every n, a unique solution $x(t)$ of (5.7) defined for

$$0 \le t \le \tau_0 + \tau_1 + \ldots + \tau_n \ .$$

If $\sum_{j=0}^\infty \tau_j = \infty$, then we have Case 1.

Now assume that

$$\sum_{j=0}^\infty \tau_j =: t^* < \infty \ .$$

In this case, we obtain a unique solution $x(t)$ for $0 \le t < t*$, where t^* is finite.

It remains to prove (5.8). Clearly, $\tau_j \to 0$, thus $M_j = 1/\tau_j \to \infty$ where

$$M_j = \max\{|f(x)| \ : \ |x| \le |a^{(j)}| + 1\}$$

and

$$a^{(j)} = x(t_j), \quad t_j = \tau_0 + \tau_1 + \ldots + \tau_{j-1} \quad \text{for} \quad j = 1, 2, \ldots$$

Since $M_j \to \infty$ one obtains that

$$|a^{(j)}| = |x(t_j)| \to \infty \quad \text{as} \quad j \to \infty \; . \tag{5.9}$$

Suppose (5.8) does not hold. Then there exists $R > 0$ and a sequence $\tilde{t}_j \to t^*_-$ with

$$|x(\tilde{t}_j)| \le R \quad \text{for all} \quad j \; .$$

Set

$$\tilde{a}^{(j)} = x(\tilde{t}_j) \; ,$$

thus

$$|\tilde{a}^{(j)}| \le R \quad \text{for all} \quad j \; .$$

Apply Theorem 5.3 to the IVP

$$x' = f(x), \quad x(\tilde{t}_j) = \tilde{a}^{(j)}$$

and obtain a unique solution in an interval $[\tilde{t}_j, \tilde{t}_j + \Delta]$, where $\Delta > 0$ depends on R, but not on j. Since $\tilde{t}_j \to t^*_-$ one obtains that

$$\tilde{t}_j + \Delta > t^*$$

for large j. Thus, one obtains a solution beyond t^*. This contradicts (5.9). It follows that $|x(t)| \to \infty$ as $t \to t^*_-$. $\diamond$

A Sufficient Condition for All–Time Existence.

Lemma 5.2 *(Gronwall's Lemma) Let $g : [0,T] \to [0,\infty)$ denote a non–negative, continuous function satisfying the estimate*

$$g(t) \le C + L \int_0^t g(s)\, ds \quad \textit{for} \quad 0 \le t \le T \; ,$$

where C and L are positive constants. Then

$$g(t) \le Ce^{Lt} \quad \textit{for} \quad 0 \le t \le T \; .$$

Proof: Set

$$G(t) = C + L \int_0^t g(s)\, ds \quad \text{for} \quad 0 \le t \le T \; ,$$

thus $g(t) \le G(t)$ and $G(t) \ge C > 0$. We have $G'(t) = Lg(t)$ and

$$\frac{G'(t)}{G(t)} = \frac{Lg(t)}{G(t)} \le \frac{LG(t)}{G(t)} = L \; ,$$

thus

$$\frac{d}{dt} \ln G(t) = \frac{G'(t)}{G(t)} \le L \; .$$

It follows that

$$\ln \frac{G(t)}{G(0)} \le Lt \ ,$$

thus

$$G(t) \le G(0) e^{Lt} = C e^{Lt} \ .$$

The estimate $g(t) \le Ce^{Lt}$ follows since $g(t) \le G(t)$. $\diamond$

Theorem 5.5 *Let $f : \mathbb{R}^n \to \mathbb{R}^n, f \in C^1$, and assume that*

$$\langle x, f(x) \rangle \le C + L|x|^2 \quad \textit{for all} \quad x \in \mathbb{R}^n \ ,$$

where C and L are positive constants. Then the IVP

$$x' = f(x), \quad x(0) = a \ ,$$

has a unique solution, which exists for $0 \le t < \infty$. Here the initial value $a \in \mathbb{R}^n$ is arbitrary.

Proof: Let $0 \le t < t^*$ denote the maximal interval of existence of the initial value problem. Set

$$g(t) = |x(t)|^2 \ .$$

We have

$$\begin{aligned} g'(t) &= 2\langle x(t), x'(t) \rangle \\ &= 2\langle x(t), f(x(t)) \rangle \\ &\le 2C + 2L|x(t)|^2 \\ &= 2C + 2Lg(t) \end{aligned}$$

Applying Gronwall's Lemma in every interval $[0, T]$ with $0 < T < t^*$ yields the bound

$$|x(t)|^2 = g(t) \le 2Ce^{2Lt} \quad \text{for} \quad 0 \le t < t^* \ .$$

Since blow–up cannot occur, the solution exists for $0 \le t < \infty$. $\diamond$

Example: (The Lorenz system)
(Edward Lorenz, 1917–2008, meteorologist)
For $x \in \mathbb{R}^3$ let

$$f(x) = \begin{pmatrix} \sigma(x_2 - x_1) \\ rx_1 - x_2 - x_1 x_3 \\ x_1 x_2 - b x_3 \end{pmatrix} = Ax + Q(x) \ .$$

Here σ, r, b are real parameters, $A \in \mathbb{R}^{3\times 3}$ and

$$Q(x) = \begin{pmatrix} 0 \\ -x_1 x_3 \\ x_1 x_2 \end{pmatrix} .$$

The system $x' = f(x)$ has been studied extensively in chaos theory. We have

$$\langle x, Q(x) \rangle = x_2(-x_1 x_3) + x_3 x_1 x_2 = 0 \quad \text{for all} \quad x \in \mathbb{R}^3 .$$

It follows that

$$\langle x, f(x) \rangle = \langle x, Ax \rangle \leq |A||x|^2 .$$

By Theorem 5.5 the solutions of the Lorenz system cannot blow–up, thus they exists for $0 \leq t < \infty$.

5.5 Differentiation w.r.t. the Initial Value

Let $f : \mathbb{R}^n \to \mathbb{R}^n, f \in C^1$. Consider the IVP

$$x' = f(x), \quad x(0) = a + h ,$$

where $a, h \in \mathbb{R}^n$. Assume that the solution $x(t, a+h)$ exists in the interval $0 \leq t \leq T$ if $|h| \leq \delta_0$ where $\delta_0 > 0$.

Thus we have

$$x_t(t, a+h) = f(x(t, a+h)) \quad \text{for} \quad 0 \leq t \leq T \quad \text{and} \quad x(0, a+h) = a+h \tag{5.10}$$

for $|h| \leq \delta_0$. Differentiating the function

$$b \to x_t(t, b)$$

formally at $b = a$ one obtains that

$$x_{at}(t, a) = Df(x(t, a)) x_a(t, a), \quad x_a(0, a) = I \in \mathbb{R}^{n \times n} . \tag{5.11}$$

Here $Df(x) \in \mathbb{R}^{n \times n}$ is the Jacobian of f at $x \in \mathbb{R}^n$.

Thus, if the formal differentiation process can be justified, then the matrix function

$$\Phi(t) := x_a(t, a) \in \mathbb{R}^{n \times n}$$

is the normalized fundamental matrix of the linear equation

$$\phi'(t) = Df(x(t, a)) \phi(t), \quad 0 \leq t \leq T . \tag{5.12}$$

In the following, we denote by $\Phi(t) \in \mathbb{R}^{n \times n}$ the normalized fundamental matrix for the linear system 5.12).

We will prove that the function $b \to x(t, b)$ from $B(a, \delta_0) \to \mathbb{R}^n$ is differentiable at $b = a$ and that its derivative at $b = a$ equals $\Phi(t)$. In other words, we will prove:

$$\lim_{h\to 0} \frac{1}{|h|}\Big(x(t,a+h) - x(t,a) - \Phi(t)h\Big) = 0 \quad \text{for} \quad 0 \le t \le T \ . \tag{5.13}$$

An Auxiliary Result:

Lemma 5.3 *Consider the linear, variable coefficient IVP*

$$q'(t) = A(t)q(t) + r(t), \quad 0 \le t \le T \ ,$$

where $A(t) \in \mathbb{R}^{n\times n}$ *and* $r(t) \in \mathbb{R}^n$ *are continuous. Let* $|A(t)| \le L$ *for* $0 \le t \le T$. *A solution* $q \in C^1[0,T]$ *satisfies the estimate*

$$|q(t)| \le C e^{(L+\frac{1}{2})t} \quad \textit{for} \quad 0 \le t \le T$$

with

$$C^2 = |q(0)|^2 + \int_0^T |r(s)|^2\, ds \ .$$

Proof: We have

$$\begin{aligned}
\frac{d}{dt}|q(t)|^2 &= 2\langle q(t), q'(t)\rangle \\
&\le 2L|q(t)|^2 + 2|q(t)||r(t)| \\
&\le (2L+1)|q(t)|^2 + |r(t)|^2
\end{aligned}$$

Therefore,

$$\begin{aligned}
|q(t)|^2 &\le |q(0)|^2 + \int_0^t |r(s)|^2\, ds + (2L+1)\int_0^t |q(s)|^2\, ds \\
&\le C^2 + (2L+1)\int_0^t |q(s)|^2\, ds
\end{aligned}$$

The estimate for $q(t)$ follows from Gronwall's Lemma, Lemma 5.2. $\diamond$

Next, we prove that the function $b \to x(t,b)$ from $B(a,\delta_0) \to \mathbb{R}^n$ satisfies the Lipschitz estimate

$$|x(t,a+h) - x(t,a)| \le C|h| \quad \text{for} \quad 0 \le t \le T \quad \text{and} \quad |h| \le \delta_0$$

where $C > 0$ is a constant.

Let us recall the Mean–Value Theorem: For $\alpha, \beta \in \mathbb{R}^n$ we have

$$f(\beta) - f(\alpha) = \Big(\int_0^1 Df(s\beta + (1-s)\alpha)\, ds\Big)(\beta - \alpha) \ .$$

We have

$$\begin{aligned} x_t(t, a+h) &= f(x(t, a+h)) \\ x_t(t, a) &= f(x(t, a)) \end{aligned}$$

thus, using the Mean Value Theorem,

$$x_t(t, a+h) - x_t(t, a) = \Big(\int_0^1 Df(sx(t, a+h) + (1-s)x(t, a))\, ds \Big) (x(t, a+h) - x(t, a)) \, .$$

Set

$$A(t, h) = \int_0^1 Df(sx(t, a+h) + (1-s)x(t, a))\, ds \, .$$

Since $f \in C^1$ there exists $L > 0$ with

$$|A(t, h)| \le L \quad \text{for} \quad 0 \le t \le T, \quad |h| \le \delta_0 \, .$$

For

$$w(t, h) = x(t, a+h) - x(t, a)$$

obtain that

$$w_t(t, h) = A(t, h) w(t, h), \quad w(0, h) = h \, ,$$

thus

$$\frac{d}{dt} |w(t, h)|^2 \le 2L |w(t, h)|^2 \, .$$

This yields that

$$|w(t, h)|^2 \le |h|^2 + 2L \int_0^t |w(\tau, h)|^2 \, d\tau \, .$$

Using Gronwall's Lemma we obtain the estimate

$$|w(t, h)| \le e^{LT} |h| \quad \text{for} \quad 0 \le t \le T \, .$$

This proves the Lipschitz bound:

$$|x(t, a+h) - x(t, a)| \le e^{LT} |h| \quad \text{for} \quad 0 \le t \le T, \quad |h| \le \delta_0 \, . \tag{5.14}$$

To prove (5.13) we set

$$q(t, h) = x(t, a+h) - x(t, a) - \Phi(t) h \quad \text{for} \quad 0 \le t \le T \quad \text{and} \quad |h| \le \delta_0 \, .$$

Here, by definition, $\Phi(t)$ is the normalized fundamental matrix for the ODE

$$\phi'(t) = Df(x(t,a))\phi(t) \ .$$

Thus, $\Phi(t)$ satisfies

$$\Phi'(t) = Df(x(t,a))\Phi(t), \quad \Phi(0) = I \ .$$

We note that $q(0,h) = 0$ and must prove that

$$\lim_{h\to 0} \frac{1}{|h|}\, q(t,h) = 0 \ .$$

We have

$$\begin{aligned} q_t(t,h) &= x_t(t,a+h) - x_t(t,a) - \Phi'(t)h \\ &= f(x(t,a+h)) - f(x(t,a)) - Df(x(t,a))\Phi(t)h \\ &= Df(x(t,a))\Big(x(t,a+h) - x(t,a)\Big) + R(t,h) - Df(x(t,a))\Phi(t)h \end{aligned}$$

where

$$\begin{aligned} R(t,h) &= f(x(t,a+h)) - f(x(t,a)) - Df(x(t,a))\Big(x(t,a+h) - x(t,a)\Big) \\ &= \Big(\int_0^1 [Df(sx(t,a+h) + (1-s)x(t,a)) - Df(x(t,a))]\, ds\Big)\Big(x(t,a+h) - x(t,a)\Big) \end{aligned}$$

From the equation for q_t obtain:

$$q_t(t,h) = Df(x(t,a))q(t,h) + R(t,h) \ . \tag{5.15}$$

Using that

$$|x(t,a+h) - x(t,a)| \le C|h|$$

we obtain

$$|R(t,h)| \le C|h| \int_0^1 H(s,t,h)\, ds$$

with

$$H(s,t,h) = \Big| Df(sx(t,a+h) + (1-s)x(t,a)) - Df(x(t,a)) \Big|$$

for $0 \le s \le 1, 0 \le t \le T, |h| \le \delta_0$. Note that $H(s,t,h)$ is uniformly continuous on the compact set

$$[0,1] \times [0,T] \times \bar{B}(a,\delta_0)$$

and

$$H(s,t,0) \equiv 0 \ .$$

It follows that for all $\varepsilon > 0$ there exists $\delta > 0$ so that

$$\int_0^1 H(s,t,h)\,ds \le \varepsilon \quad \text{for} \quad 0 \le t \le T \quad \text{and} \quad |h| \le \delta \ .$$

Therefore, for all $\varepsilon > 0$ there exists $\delta > 0$ so that

$$|R(t,h) \le \varepsilon |h| \quad \text{for} \quad 0 \le t \le T \quad \text{and} \quad |h| \le \delta \ .$$

From (5.15) and Lemma 5.3 obtain that

$$|q(t,h)| \le C_T \max_{0 \le t \le T} |R(t,h)| \ .$$

It follows that given $\varepsilon > 0$ there exists $\delta > 0$ so that for $0 \le t \le T$:

$$\frac{1}{|h|} \Big| x(t,a+h) - x(t,a) - \Phi(t)h \Big| \le \varepsilon$$

if $|h| \le \delta$.

This proves that

$$\lim_{h \to 0} \frac{1}{|h|} \Big(x(t,a+h) - x(t,a) - \Phi(t)h \Big) = 0 \ .$$

We have proved:

Theorem 5.6 *Let $f : \mathbb{R}^n \to \mathbb{R}^n, f \in C^1$. Assume that the IVP*

$$x' = f(x), \quad x(0) = a + h$$

has the unique solution $x(t, a+h)$ for $0 \le t \le T$ where $a, h \in \mathbb{R}^n, |h| \le \delta_0$ for some $\delta_0 > 0$. The function

$$b \to x(t,b)$$

from $B(a, \delta_0)$ to $\mathbb{R}^n$ is differentiable at $b = a$ and the Jacobian

$$x_a(t,a) =: \Phi(t)$$

is the normalized fundamental matrix of the linear system

$$\phi'(t) = DF(x(t,a))\phi(t)$$

for $0 \le t \le T$.

The following simple example shows that the differentiability w.r.t. the initial condition may fail if the function f is not differentiable.

Example: Let $f(x) = |x|$ for $x \in \mathbb{R}$. The IVP

$$x' = |x|, \quad x(0) = x_0$$

has the solution

$$x(t) = x_0 e^t \quad \text{for} \quad x_0 \ge 0$$

and the solution

$$x(t) = x_0 e^{-t} \quad \text{for} \quad x_0 < 0 \ .$$

Denote the solution by $\phi(x_0, t)$. The function $x_0 \to \phi(x_0, t)$ is not differentiable at $x_0 = 0$ if $t > 0$.

6 Asymptotic Stability of Periodic Orbits

6.1 Basics and Terminology

Let $f : \mathbb{R}^N \to \mathbb{R}^N$ denote a C^1–map with

$$|f_x(x)| \le L \quad \text{for all} \quad x \in \mathbb{R}^N \ .$$

The solution $x(t)$ of the IVP

$$x' = f(x), \quad x(0) = a$$

is denoted by $x(t) = \phi(t, a)$. Note that

$$|\phi(t,a) - \phi(t,b)| \le e^{Lt}|a - b| \quad \text{for} \quad t \ge 0 \ . \tag{6.1}$$

To show this, set

$$\delta(t) = \phi(t,a) - \phi(t,b) \ ,$$

and note that

$$\delta'(t) = f(\phi(t,a)) - f(\phi(t,b)) \ .$$

Then apply the mean–value theorem (see Theorem 14.3) to obtain

$$|\delta'(t)| \le L|\delta(t)|$$

and use that

$$\frac{d}{dt}\,|\delta(t)|^2 = 2\delta(t) \cdot \delta'(t) \le 2L|\delta(t)|^2 \ .$$

Then apply Theorem 14.1 to obtain the estimate (6.1).

We assume that $u(t)$ is a periodic solution of the differential system $x' = f(x)$ and denote the minimal period of $u(t)$ by $T > 0$. In particular, we assume that $u(t)$ is not constant.

If

$$x_0 := u(0) = u(T)$$

then

$$u(t) = \phi(t, x_0) \equiv \phi(t + T, x_0) \ .$$

Set

$$A(t) = Df(u(t)) \in \mathbb{R}^{N \times N}$$

and note that $A(t) \equiv A(t + T)$

Lemma 6.1 *The linear system*

$$x'(t) = Df(u(t))x(t) \tag{6.2}$$

has the normalized fundamental matrix

$$\Phi_0(t) = D\phi(t, x_0)$$

where $D\phi(t,a) = \phi_a(t,a)$.

Proof: From

$$\frac{d}{dt}\phi(t,a) = f(\phi(t,a)), \quad \phi(0,a) = a$$

obtain:

$$\frac{d}{dt}\phi_a(t,a) = Df(\phi(t,a))\phi_a(t,a), \quad \phi_a(0,a) = I \ .$$

Taking $a = x_0$ proves the claim. $\diamond$

Remark: The lemma says that the linearized solution $D\phi(t, x_0)$ is the solution of the linearized equation,

$$x' = Df(u(t))x \ .$$

Roughly, linearizing commutes with solving.

We denote the monodromy matrix of the linear system $x' = Df(u(t))x$ by

$$M := \Phi_0(T) = \phi_a(T, x_0)$$

and denote the eigenvalues of M by $\mu_1, \ldots, \mu_N$. These are the Floquet multipliers of the linear system (6.2).

Lemma 6.2 $\mu_1 = 1$ *is a Floquet multiplier of (6.2).*

Proof: From $u'(t) = f(u(t))$ obtain $u''(t) = Df(u(t))u'(t)$. Thus, $u'(t)$ solves the linear system $x' = Df(u(t))x$. Therefore,

$$u'(0) = u'(T) = Mu'(0) \ .$$

If $u'(0) = 0$ then $0 = f(u(0))$, and $u(t)$ would be constant. This contradicts our assumption that $u(t)$ is a periodic solution. $\diamond$

The set

$$\gamma = \{u(t) \mid 0 \le t \le T\}$$

is the orbit determined by the periodic solution $u(t)$. We define the ε–neighborhood of γ by

$$U_\varepsilon = U_\varepsilon(\gamma) = \{x \in \mathbb{R}^N \mid dist(x, \gamma) < \varepsilon\} \ .$$

Here

$$dist(x, \gamma) = \min\{|x - y| \ : \ y \in \gamma\} \ .$$

Definition: *a) The orbit γ is stable if for every $\varepsilon > 0$ there exists $\delta > 0$ so that for all $a \in U_\delta$ we have*

$$\phi(t, a) \in U_\varepsilon \quad \textit{for all} \quad t \geq 0 \ .$$

b) The orbit γ is called asymptotically stable if γ is stable and there exists $\delta > 0$ so that for all $a \in U_\delta$ we have

$$dist(\phi(t, a), \gamma) \to 0 \quad \text{as} \quad t \to \infty \ .$$

We will prove the following result in Section 6.5:

Theorem 6.1 *If $|\mu_j| < 1$ for $j = 2, \ldots, N$ then the orbit γ is asymptotically stable.*

6.2 The Return Time

Set

$$v^N := f(x_0)/|f(x_0)| \in \mathbb{R}^N, \quad |v^N| = 1 \ .$$

The hyperplane

$$H = \{x \in \mathbb{R}^N \mid \langle x - x_0, v^N \rangle = 0\}$$

crosses the orbit γ orthogonally at $u(0) = x_0$. See Figure 6.1.

Let $v^1, v^2, \ldots, v^N$ denote an ONB of $\mathbb{R}^N$. We set

$$V = (v^1 | \ldots | v^{N-1}) \in \mathbb{R}^{N \times (N-1)} \quad \text{and} \quad \tilde{V} = (V | v^N) \in \mathbb{R}^{N \times N} \ .$$

We have

$$V^T V = I_{N-1} \quad \text{and} \quad \tilde{V}^T \tilde{V} = I_N \ .$$

The hyperplane H can also be written as

$$H = \{x \in \mathbb{R}^N \ : \ x = x_0 + V\alpha, \alpha \in \mathbb{R}^{N-1}\} \ .$$

Consider any $a \in \mathbb{R}^N$. Then $\phi(t,a) \in H$ iff

$$\langle \phi(t,a) - x_0, v^N \rangle = 0 \ .$$

We define the map $F : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ by

$$F(t,a) = \langle \phi(t,a) - x_0, v^N \rangle \ .$$

It is then clear that $F(0,x_0) = F(T,x_0)$ and

$$\begin{aligned} F(T,x_0) &= \langle \phi(T,x_0) - x_0, v^N \rangle \\ &= \langle 0, v^N \rangle \\ &= 0 \end{aligned}$$

Also,

$$F_t(t,a) = \langle \phi_t(t,a), v^N \rangle \ ,$$

thus

$$\begin{aligned} F_t(T,x_0) &= \langle \phi_t(T,x_0), v^N \rangle \\ &= \langle f(x_0), v^N \rangle \\ &> 0 \ . \end{aligned}$$

Apply the Implicit Function Theorem (see Theorem 16.1) to the map

$$F : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$$

where $\mathbb{R}$ is the state space, $\mathbb{R}^N$ is the parameter space, and

$$F(T,x_0) = 0$$

is the unperturbed equation. We now perturb x_0 and obtain a return time $t = t^*(a) \sim T$ for $a \sim x_0$. See Figure 6.2.

More precisely, obtain that there exist $r > 0, \tau > 0$ so that, for every $a \in B_r(x_0)$, the equation

$$F(t,a)=0$$

has a unique solution $t=t^*(a)\in(T-\tau,T+\tau)$. The function $a\to t^*(a)$ is C^1 on $B_r(x_0)$. The time $t^*(a)$ is called the return time of the orbit starting at $a\in B_r(x_0)$.

Note that for $a\in B_r(x_0)$:

$$\phi\Big(t^*(a),a\Big)\in H \quad \text{and} \quad t^*(a)\sim T\ .$$

Also, $t^*(x_0)=T$.

6.3 The Poincaré Map Corresponding to the Orbit γ

We define the map

$$P:B_r(x_0)\to\mathbb{R}^N$$

by

$$P(a)=\phi\Big(t^*(a),a\Big)$$

and note that $P(a)\in H\subset\mathbb{R}^N$ for all $a\in B_r(x_0)$.

We have

$$P(x_0)=x_0$$

and

$$P'(a)=\phi_t\Big(t^*(a),a\Big)\nabla t^*(a)+\phi_a(t^*(a),a)\in\mathbb{R}^{N\times N}\ .$$

Evaluating this equation at $a=x_0$ one finds that

$$P'(x_0)=f(x_0)\nabla t^*(x_0)+M \tag{6.3}$$

where $M=\phi_a(T,x_0)$ is the monodromy matrix of the system $x'=Df(u(t))x$.

One often restricts P to

$$N_r=B_r(x_0)\cap H$$

and then calls the map

$$P|_{N_r}:N_r\subset H\to H$$

the Poincaré map corresponding to the orbit γ for the cross section H. See Figure 6.3.

6.4 The Poincaré Map in V–Coordinates

Recall the matrices

$$V = (v^1|\ldots|v^{N-1}) \in \mathbb{R}^{N\times(N-1)} \quad \text{and} \quad \tilde{V} = (V|v^N) \in \mathbb{R}^{N\times N}$$

from Section 6.2.

For each $a \in N_r = B_r(x_0)\cap H$ there exists a unique $\alpha \in \mathbb{R}^{N-1}$ with

$$a = x_0 + V\alpha, \quad \alpha \in B_r(\mathbb{R}^{N-1}) \ .$$

Here we use the notation

$$B_r(\mathbb{R}^{N-1}) = \{\alpha \in \mathbb{R}^{N-1} \ : \ |\alpha| < r\} \ .$$

Since $P(a) \in H$ there exists a unique β so that

$$P(a) = x_0 + V\beta, \quad \beta \in \mathbb{R}^{N-1} \ .$$

The map $\alpha \to \beta$ expresses the map P, restricted to $N_r = B_r(x_0)\cap H$, in the coordinates determined by the orthonormal basis vectors $v^1,\ldots,v^{N-1}$ of H. Explicitly,

$$\begin{aligned} \beta &= V^T(P(a) - x_0) \\ &= V^T(P(x_0 + V\alpha) - x_0) \end{aligned}$$

Therefore,

$$\tilde{P}(\alpha) = V^T(P(x_0+V\alpha) - x_0) \quad \text{for} \quad \alpha \in B_r(\mathbb{R}^{N-1})$$

is the Poincaré map in V–coordinates.

Clearly, $\tilde{P}(0) = 0$ since $P(x_0) = x_0$. Also,

$$\tilde{P}'(0) = V^T P'(x_0) V = V^T M V \in \mathbb{R}^{(N-1)\times(N-1)} \ .$$

Here we have used (6.3) and $V^T f(x_0) = 0$.

For the following recall that the Floquet multipliers $\mu_1 = 1, \mu_2, \ldots, \mu_N$ are the eigenvalues of the monodromy matrix $M = \phi_a(T, x_0)$.

Theorem 6.2 *Let $\mu_1 = 1, \mu_2, \ldots, \mu_N$ denote the Floquet multipliers of the system $x' = Df(u(t))x$, i.e., the μ_j are the eigenvalues of the monodromy matrix $M = \Phi_0(T) = \phi_a(T, x_0)$. Then the spectrum of $\tilde{P}(0) = V^T M V$ is*

$$\sigma(\tilde{P}(0)) = \{\mu_2, \ldots, \mu_N\} \ .$$

Proof: Note that $Mv^N = v^N$ and $V^T M v^N = V^T v^N = 0$. Therefore, the claim follows from

$$\tilde{V}^T M \tilde{V} = \begin{pmatrix} V^T \\ v^{NT} \end{pmatrix} M \Big(V | v^N \Big) = \begin{pmatrix} V^T M V & 0 \\ * & 1 \end{pmatrix} \ .$$

$\diamond$

6.5 Stability and Asymptotic Stability of Periodic Orbits

We assume $|\mu_j| < 1$ for $j = 2, \ldots, N$. Thus, the matrix $\tilde{P}'(0)$ has a spectral radius strictly less than one.

Using Theorem 14.6 one obtains:

Theorem 6.3 *Under the above assumption:*
a) For all $\varepsilon > 0$ there exists $\delta > 0$ so that

$$|\alpha| < \delta \quad \textit{implies} \quad |\tilde{P}^n(\alpha)| < \varepsilon \quad \textit{for all} \quad n \geq 1 \; .$$

b) There exists $\delta > 0$ so that

$$|\alpha| < \delta \quad \textit{implies} \quad \tilde{P}^n(\alpha) \to 0 \quad \textit{as} \quad n \to \infty \; .$$

We first translate this result into a result for the mapping

$$P_{N_r} : N_r \to H \quad \text{where} \quad N_r = B_r(x_0) \cap H \; .$$

Theorem 6.4 *Under the above assumption:*
a) For all $\varepsilon > 0$ there exists $\delta > 0$ so that

$$a \in N_\delta \quad \textit{implies} \quad P^n(a) \in N_\varepsilon \quad \textit{for all} \quad n \geq 1 \; .$$

b) There exists $\delta > 0$ so that

$$a \in N_\delta \quad \textit{implies} \quad P^n(a) \to x_0 \quad \textit{as} \quad n \to \infty \; .$$

We now show asymptotic stability of the orbit γ. Let $\varepsilon > 0$ be given and let $\delta > 0$ be chosen so that the following holds:

$$a \in N_\delta \quad \text{implies} \quad P^n(a) \in N_\varepsilon \quad \text{for all} \quad n \in \mathbb{N} \quad \text{and} \quad P^n(a) \to 0 \quad \text{as} \quad n \to \infty \, .$$

We claim that for $a \in N_\delta$:

$$dist(\phi(t,a), \gamma) \to 0 \quad \text{as} \quad t \to \infty \; . \tag{6.4}$$

To this end, consider the sequence

$$a^{(n)} = P^n(a) \in N_\varepsilon, \quad n = 0, 1, \ldots$$

We have

$$\begin{aligned} a^{(0)} &= a \\ a^{(1)} &= P(a) = \phi(t^*(a), a) \\ a^{(2)} &= P^2(a) = \phi(t^*(a) + t^*(P(a)), a) \end{aligned}$$

etc.

Define

$$t_n = t^*(a) + t^*(P(a)) + \ldots + t^*(P^{n-1}(a)) \ .$$

We then have

$$a^{(n)} = P^n(a) = \phi(t_n, a) \in N_\varepsilon \quad \text{for all} \quad n \ge 0 \ .$$

Recall that

$$T - \tau < t^*(b) < T + \tau \quad \text{for all} \quad b \in N_r$$

where $\tau > 0$ is small. In the following, we assume without loss of generality that $0 < \tau \le T/2$. Then we have $t_n \ge nT/2$, thus $t_n \to \infty$ as $n \to \infty$.

Let $t > 0$ be large. We write

$$t = t_n + \tilde{t} \quad \text{with} \quad 0 \le \tilde{t} \le T + \tau \ .$$

We then have

$$\phi(t, a) = \phi(t_n + \tilde{t}, a) = \phi(\tilde{t}, P^n(a))$$

with $P^n(a) = a^{(n)} \in N_\varepsilon$. Therefore,

$$\begin{aligned}
|\phi(t,a) - u(\tilde{t})| &= |\phi(\tilde{t}, a^{(n)}) - \phi(\tilde{t}, x_0)| \\
&\le e^{L\tilde{t}} |a^{(n)} - x_0| \\
&\le e^{3LT/2} |a^{(n)} - x_0| \\
&= C|a^{(n)} - x_0|
\end{aligned}$$

As $t \to \infty$ we have $n \to \infty$ and, therefore, $a^{(n)} \to x_0$ by Theorem 6.4. This proves (6.4).

We have shown:

Theorem 6.5 *Let $f : \mathbb{R}^N \to \mathbb{R}^N, f \in C^1$, and let $|f'(x)| \le L$ for all $x \in \mathbb{R}^N$. Denote the solution of the IVP*

$$x' = f(x), \quad x(0) = a \ ,$$

by $\phi(t, a)$. Let $u(t)$ denote a solution of the equation $x' = f(x)$ with period $T > 0$,

$$u(t) \equiv u(t + T) \ .$$

Let $\gamma = \{u(t) \ : \ 0 \le t \le T\}$ denote the orbit determined by $u(t)$ and set $x_0 := u(0) = u(T)$. Set

$$H(x_0) := \{x \in \mathbb{R}^N \ : \ \langle x - x_0, f(x_0) \rangle = 0\}$$

and

$$N_\delta(x_0) := B_\delta(x_0) \cap H(x_0) \ .$$

Let M denote the monodromy matrix for the linear system

$$x' = A(t)x \quad \textit{where} \quad A(t) = f'(u(t)) \ .$$

The vector $u'(0) = f(x_0)$ satisfies

$$Mf(x_0) = f(x_0) \ ,$$

thus $\mu_1 = 1$ is an eigenvalue of M. Assume that $|\mu_j| < 1$ for $j = 2, \ldots, N$. Then the following holds: Given $\varepsilon > 0$ there exists $\delta > 0$ so that

a) $a \in N_\delta(x_0)$ implies $\phi(t,a) \in U_\varepsilon$ for all $t \geq 0$.

b) $a \in N_\delta(x_0)$ implies $dist(\phi(t,a), \gamma) \to 0$ as $t \to \infty$.

Let $a \in \mathbb{R}^N$ denote an arbitrary point close to γ and let $y_0 \in \gamma$ denote the point on γ closest to a,

$$dist(a, \gamma) = |y_0 - a| \ .$$

Given $\varepsilon > 0$ determine $\delta = \delta(y_0) > 0$ as in the previous theorem, where x_0 is replaced by y_0. One can show (see Remarks below) that there exists $\delta_0 > 0$, depending on $\varepsilon > 0$ but not on y_0, so that

$$\delta(y_0) \geq \delta_0 > 0 \quad \text{for all} \quad y_0 \in \gamma \ . \tag{6.5}$$

Then, if $a \in U_{\delta_0}$ is arbitrary, it follows that $\phi(t,a) \in U_\varepsilon$ for $t \geq 0$ and

$$dist(\phi(t,a), \gamma) \to 0 \quad \text{as} \quad t\infty \ .$$

This completes the proof that γ is asymptotically stable.

Remarks: Recall that $\Phi_0(t) = Q(t)e^{Bt}$ denotes the normalized fundamental matrix of the system $x' = A(t)x$ where $A(t) = f'(u(t))$. For any $y_0 \in \gamma$ there exists $t_0 \in [0, T]$ with $y_0 = u(t_0)$. Consider the linear system

$$y' = A(t + t_0)y \ .$$

Since

$$\Phi_0'(t + t_0) = A(t + t_0)\Phi_0(t + t_0)$$

one obtains that

$$\Psi(t) = \Phi(t + t_0)(\Phi_0(t_0))^{-1}$$

is the normalized fundamental matrix of the system $y' = A(t + t_0)y$. The corresponding monodromy matrix is

$$\begin{aligned} \Psi(T) &= Q(T + t_0)e^{BT + Bt_0}\Big(Q(t_0)e^{Bt_0}\Big)^{-1} \\ &= Q(t_0)e^{BT}(Q(t_0))^{-1} \\ &= Q(t_0)M(Q(t_0))^{-1} \end{aligned}$$

Thus, the monodromy matrix $\Psi(T)$ of the system $y' = A(t + t_0)y$ is similar to the monodromy matrix M of the system $x' = A(t)x$. Thus, the Floquet multipliers μ_j are independent of $y_0 \in \gamma$. Since $|Q(t_0)||(Q(t_0))^{-1}| \leq const.$ one obtains existence of $\delta_0 > 0$ with (6.5).

6.6 Stability Check for $N = 2$

Consider a system $x' = f(x)$ where $x(t) \in \mathbb{R}^2$ and assume that $u(t)$ is a periodic solution, $u(t) \equiv u(t + T)$. Let $M = \Phi_0(T) \in \mathbb{R}^{2\times 2}$ denote the monodromy matrix of the linear system $x' = Df(u(t))x$ and let $\mu_1 = 1$ and μ_2 denote the eigenvalues of M. We know that

$$det\, M = \mu_1\mu_2 = \mu_2 \ .$$

By Liouville's Theorem (Theorem 3.1) we have

$$\begin{aligned} \mu_2 &= det\, M \\ &= det\, \Phi_0(T) \\ &= \exp\Big(\int_0^T tr\, Df(u(s))\, ds\Big) \\ &= \exp\Big(\int_0^T (\nabla \cdot f)(u(s))\, ds\Big) \end{aligned}$$

Therefore, if

$$\int_0^T (\nabla \cdot f)(u(s))\, ds < 0$$

then $0 < \mu_2 < 1$ and the orbit γ corresponding to $u(t)$ is asymptotically stable. If

$$\int_0^T (\nabla \cdot f)(u(s))\, ds > 0$$

then $\mu_2 > 1$ and γ is unstable.

Example: (Problem 1 on page 219, Perko)
Consider the 2D system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \Big(1 - \frac{x^2}{4} - y^2\Big)\begin{pmatrix} x \\ y \end{pmatrix} =: \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} \ .$$

It is easy to check that

$$u(t) = \begin{pmatrix} 2\cos(2t) \\ \sin(2t) \end{pmatrix}$$

is a solution of period $T = \pi$. The corresponding orbit γ is the ellipse

$$\left(\frac{x}{2}\right)^2 + y^2 = 1 \ . \tag{6.6}$$

We have

$$\begin{aligned} f_{1x} &= 1 - \frac{x^2}{4} - y^2 - \frac{x^2}{2} \\ f_{2y} &= 1 - \frac{x^2}{4} - y^2 - 2y^2 \end{aligned}$$

If one evaluates the divergence of f at $u(t)$, then the term $1 - x^2/4 - y^2$ drops out since $u(t)$ moves along the ellipse (6.6) and one obtains that

$$(\nabla \cdot f)(u(t)) = -\frac{1}{2}\, u_1^2(t) - 2u_2^2(t) < 0 \ .$$

The orbit γ is asymptotically stable.

Figures for Chapter 6

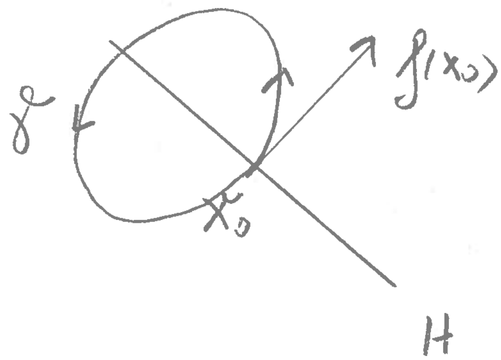

Figure 6.1

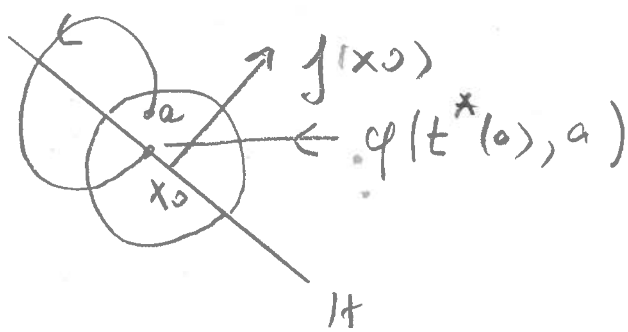

Figure 6.2

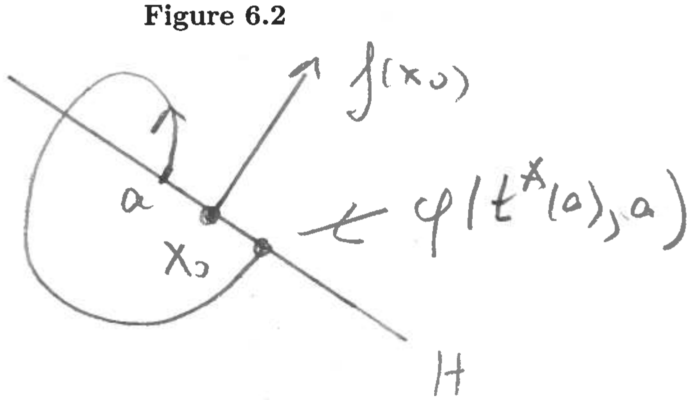

Figure 6.3

7 Introduction to Bifurcations of Equilibria

7.1 Examples

Saddle–Node Bifurcation: Consider the ODE

$$x' = r + x^2$$

where r is a real parameter. For $r < 0$ there are two fixed points,

$$x_1(r) = -\sqrt{-r}, \quad x_2(r) = \sqrt{-r} \ .$$

The fixed point $x_1(r)$ is stable; $x_2(r)$ is unstable. At $r = 0$ the two fixed points collide. There is no fixed point for $r > 0$.

For systems of equations, a stable fixed point often is a node, an unstable fixed point often is a saddle. Suppose a parameter r changes from $r < r_0$ to $r > r_0$ and at $r = r_0$ a saddle collides with a node. Here it is assumed that the saddle and the node both exist for $r < r_0$, for example. Then one says that a saddle–node bifurcation occurs at $r = r_0$.

Transcritical Bifurcation: Consider the ODE

$$x' = x(r - x)$$

For every parameter value $r \in \mathbb{R}$ the equation has two fixed points, the trivial fixed point $x_1(r) = 0$ and the fixed point $x_2(r) = r$. The trivial fixed point is stable for $r < 0$ and unstable for $r > 0$. The fixed point $x_2(r) = r$ is unstable for $r < 0$ and stable for $r > 0$. Thus, at $r = 0$, where the two branches of fixed points cross each other, an exchange of stability from one branch to the other occurs.

Supercritical Pitchfork Bifurcation: Consider the ODE

$$x' = x(r - x^2)$$

For every parameter $r \in \mathbb{R}$ the equation has the trivial fixed point $x_1(r) = 0$. The trivial fixed point is stable for $r < 0$, but loses its stability for $r > 0$. For $r > 0$ two new fixed points occur:

$$x_2(r) = \sqrt{r} \quad \text{and} \quad x_3(r) = -\sqrt{r} \ .$$

Both of them are stable. When r crosses from $r < 0$ to $r > 0$, then the trivial stable fixed point $x_1(r) = 0$ is replaced by the two stable fixed points $x_{2,3}(r) = \pm\sqrt{r}$.

Remarks on symmetry: Note that the nonlinear function

$$f(x) = rx - x^3$$

obeys the rule

$$f(-x) = -f(x) \ .$$

The equation has Z_2–symmetry. If S is the operator defined by

$$Sx(t) = -x(t)$$

then

$$S\frac{d}{dt} = \frac{d}{dt}S \quad \text{and} \quad Sf = fS \ .$$

Therefore, if $x(t)$ is a solution of the equation $x' = f(x)$ then $Sx(t) = -x(t)$ is also a solution.

One says that the group $Z_2 = \{id, S\}$ acts on functions $x(t)$.

Subcritical Pitchfork Bifurcation: Consider

$$x' = x(r + x^2)$$

For every r the equation has the trivial fixed point $x_1(r) = 0$, which is stable for $r < 0$ and unstable for $r > 0$. If $r < 0$ there are the additional unstable fixed points

$$x_2(r) = \sqrt{-r} \quad \text{and} \quad x_3(r) = -\sqrt{-r} \ .$$

When r crosses from $r < 0$ to $r > 0$, the trivial branch $x_1(r) = 0$ loses its stability, but no new stable fixed points occur. The pitchfork bifurcation is subcritical, i.e., the pitchfork occurs for parameter values r below the critical value $r = 0$.

Hysteresis Phenomenon If a subcritical pitchfork bifurcation occurs, additional nonlinear terms may stabilize the dynamics.

For example, consider

$$x' = rx + x^3 - x^5 \ .$$

This example can be used to illustrate the hysteresis phenomenon.

A simpler equation with hysteresis is

$$x' = r + x - x^3 = f(x, r) \ .$$

Sketch the function

$$g(x) = -x + x^3$$

with zeros at $-1, 0, 1$. The function $g(x)$ attains a local maximum at $x_0 = -1/\sqrt{3}$ and

$$M := g(x_0) = 2/(3\sqrt{3}) \ .$$

For

$$-M < r < M$$

the fixed point equation

$$r = x - x^3$$

has three solutions. The middle fixed point is unstable, the outer two fixed points are stable. Sketch the function $r = -x + x^3$ and switch coordinates to obtain the fixed points x^* over the parameter r. Saddle–node bifurcations occur at $r = -M$ and $r = M$. The stable branches can be used to illustrate the hysteresis phenomenon when r changes slowly between $-M - \varepsilon$ and $M + \varepsilon$.

7.2 Perturbation of the Supercritical Pitchfork

What happens to a bifurcation if the equation is perturbed? As an example, consider the equation

$$x' = h + rx - x^3 \tag{7.1}$$

with parameters h and r. If $h = 0$ then a supercritical pitchfork bifurcation occurs at $r = 0$. The parameter h unfolds this singularity. It breaks the symmetry $f(-x) = -f(x)$ of the nonlinear term.

It is interesting to make sketches of the fixed points $x^* = x^*(h, r)$ were h is fixed and $-\infty < r < \infty$. For $h = 0$ a pitchfork occurs. Consider the following four curves. They have no tangent at $(x, r) = (0, 0)$:

$$\begin{array}{lllllllll}
\Gamma_1 & : & x(r) = 0 & \text{for} & r \le 0 & \text{and} & x(r) = \sqrt{r} & \text{for} & r \ge 0 \\
\Gamma_2 & : & x(r) = 0 & \text{for} & r \ge 0 & \text{and} & x(r) = -\sqrt{r} & \text{for} & r \ge 0 \\
\Gamma_3 & : & x(r) = 0 & \text{for} & r \le 0 & \text{and} & x(r) = -\sqrt{r} & \text{for} & r \ge 0 \\
\Gamma_4 & : & x(r) = 0 & \text{for} & r \ge 0 & \text{and} & x(r) = \sqrt{r} & \text{for} & r \ge 0
\end{array}$$

The pair Γ_1, Γ_2 as well as the pair Γ_3, Γ_4 make up the pitchfork which occurs for $h = 0$.

We claim: If $h > 0$ then the pair Γ_1, Γ_2 gets perturbed. The perturbed curves Γ_{1h}, Γ_{2h} are smooth. The curve Γ_{2h} has a saddle–node.

If $h < 0$ then the pair Γ_3, Γ_4 gets perturbed. The perturbed curves Γ_{3h}, Γ_{4h} are smooth. The curve Γ_{4h} has a saddle–node.

To see this, we solve the fixed point equation

$$h + rx - x^3 = 0$$

for r and obtain

$$r = r(x) = x^2 - \frac{h}{x} \ .$$

Sketch the graphs of

$$r = x^2, \quad r = -\frac{h}{x}, \quad r = x^2 - \frac{h}{x}$$

for $x > 0$ and for $x < 0$. The cases $h > 0$ and $h < 0$ are different. Then switch the two graphs (for $h > 0$ and for $h < 0$) of $r = x^2 - \frac{h}{x}$ and obtain x as a function of r.

7.3 The Saddle–Nodes of the Perturbed Pitchfork

Fix $h > 0$. We want to determine the saddle–node (r, x) on the curve Γ_{2h}. We have $r > 0 > x$ and

$$r = x^2 - \frac{h}{x} \quad \text{and} \quad 0 = \frac{dr}{dx} = 2x + \frac{h}{x^2} \; .$$

Obtain

$$\begin{aligned} x &= -(h/2)^{1/3} \\ r &= (h/2)^{2/3} + (2/h)^{1/3} h \\ &= (h/2)^{2/3} + 2(h/2)^{2/3} \\ &= 3(h/2)^{2/3} \end{aligned}$$

Thus, for $h > 0$ fixed, a saddle–node bifurcation occurs at

$$r_{sn} = 3(h/2)^{2/3} \; .$$

The saddle–node is

$$(r_{sn}, x_{sn}) = \Big(3(h/2)^{2/3}, -(h/2)^{1/3}\Big) \; .$$

The equation

$$r = 3(h/2)^{2/3}$$

can also be expressed as

$$h = 2(r/3)^{3/2} \; .$$

One obtains that

$$x_{sn} = -(h/2)^{1/3} = -(r_{sn}/3)^{1/2} \; . \tag{7.2}$$

Similarly, fix $h < 0$. A saddle–node (r, x) lies on Γ_{4h}. We have $r > 0$ and $x > 0$. The saddle–node is

$$(r_{sn}, x_{sn}) = \Big(3(-h/2)^{2/3}, (-h/2)^{1/3}\Big) \; .$$

The equation

$$r = 3(-h/2)^{2/3}$$

can also be expressed as

$$h = -2(r/3)^{3/2} \; .$$

The points in the (r, h)–plane, where a saddle–node bifurcation (w.r.t. r) occurs, lie on the cusp

$$h = \pm 2(r/3)^{3/2} \; . \tag{7.3}$$

7.4 Stability of Fixed Points of the Perturbed Pitchfork

Fix r and h. Which fixed points x of the equation (7.1) are stable, which are unstable? Let

$$f(x,r,h) = h + rx - x^3, \quad f_x(x,r,h) = r - 3x^2 \ .$$

A fixed point x is a solution of the equation

$$h + rx - x^3 = 0 \ .$$

It is stable if

$$f_x(x,r,h) = r - 3x^2 < 0$$

and is unstable if

$$f_x(x,r,h) = r - 3x^2 > 0 \ .$$

In the following we fix $h > 0$ and let x be a fixed point, i.e.,

$$h + rx - x^3 = 0 \ ,$$

thus $x \neq 0$ and

$$r = x^2 - \frac{h}{x} \ .$$

It follows that

$$f_x(x,r,h) = r - 3x^2 = -2x^2 - \frac{h}{x} =: q(x) \ .$$

We have $q(x) = 0$ if $x^3 = -\frac{h}{2}$, i.e.,

$$x = -\Big(\frac{h}{2}\Big)^{1/3} = x_{sn} \ .$$

Sketching the function $q(x)$ makes it clear that

$$q(x) < 0 \quad \text{for} \quad x > 0 \quad \text{and} \qquad \text{for} \quad x < x_{sn}$$

and

$$q(x) > 0 \quad \text{for} \quad x_{sn} < x < 0 \ .$$

It follows that for $r \le 0$ all fixed points x are stable. These fixed points x are positive, $x > 0$.

If $r > 0$ then the fixed points $x > 0$ and the fixed points x with

$$x < x_{sn} < 0$$

are stable whereas the fixed points x with

$$x_{sn} < x < 0$$

are unstable.

A saddle–node bifurcation occurs at

$$(r_{sn}, x_{sn}) = \left(3(h/2)^{2/3}, -(h/2)^{1/3}\right) .$$

7.5 Bifurcation Diagrams for Fixed r

Consider the equation

$$f(x, r, h) = h + rx - x^3 = 0$$

for fixed r.

Case 1: $r \le 0$: The fixed point equation requires that

$$f(x, r, h) = h + rx - x^3 = 0 ,$$

thus

$$h = -rx + x^3 .$$

The function $x \to -rx + x^3$ is strictly monotonically inceasing for $r \le 0$. For every $h \in \mathbb{R}$ there exists a unique $x \in \mathbb{R}$ so that $h = -rx + x^3$. Inverting the function $h(x) = -rx + x^3$ one obtains the fixed point $x^* = x^*(h)$ and the bifurcation diagram.

One can sketch the function $h = h(x) = -rx + x^3$ and then switch the axes to obtain the fixed points $x^* = x^*(h)$ as a function of h for $r \le 0$ fixed.

Note that

$$f_x(x, r, h) = r - 3x^2 < 0$$

for $r \le 0$ and $x \neq 0$. Also, if $h = r = 0$, then the fixed point $x^* = 0$ is stable for the equation $x' = -x^3$. Therefore, if $r \le 0$ then all fixed points $x^* = x^*(h)$ are stable.

Case 2: $r > 0$: The fixed point equation is

$$h = -rx + x^3 .$$

The function

$$h(x) = -rx + x^3$$

attains a local maximum at

$$x_0 = -\sqrt{r/3}$$

and a local minimum at $x_1 = \sqrt{r}3$. The maximum is

$$M := 2\Big(\frac{r}{3}\Big)^{3/2}$$

and the minimum is $-M$.

For $-M < h < M$ the equation

$$h + rx - x^3 = 0$$

has three solutions. Two are stable fixed points, the middle one is unstable. For $h = M$ and $h = -M$ two fixed points collide; a saddle–node bifurcation occurs.

One can sketch the function $h = h(x) = -rx + x^3$ and then switch the axes to obtain the branch of fixed points $x^* = x^*(h)$ for $r > 0$ fixed. One obtains a hysteresis loop.

Saddle–nodes for fixed $r > 0$ occur if $h_x = -r + 3x^2 = 0$, i.e., at

$$x_{0,1} = \pm\sqrt{r/3}\ .$$

The corresponding values of h are

$$h(\pm\sqrt{r/3}) = \pm M = \pm 2\Big(\frac{r}{3}\Big)^{3/2}\ .$$

This agrees with the cusp equation (7.3) derived above.

7.6 The 3D Surface of Fixed points

Figure 3.6.5 in Strogatz shows the surface of fixed points

$$x^* = x^*(r,h)\ .$$

One can recognize the hysteresis loops. The saddle–nodes occur for parameters (r,h) on the cusp (7.3). The stable state of the system can change discontinuously if the parameters (r, h) cross the cusp. This may correspond to a catastrophe.

It is more difficult to recognize the pitchfork bifurcation with respect to r at $h = 0$ and the perturbed pitchfork bifurcations, which occur for fixed $h \neq 0$ as a function of $r > 0$.

8 Hopf Bifurcation

Eberhard Hopf (1902–1983)

8.1 Three Model Problems

Example 1: A degenerate Hopf bifurcation. Let

$$A(\lambda) = \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

where λ is a real parameter. The eigenvalues of $A(\lambda)$ are

$$\mu_{1,2}(\lambda) = \lambda \pm i\ .$$

For all $\lambda \in \mathbb{R}$ the ODE system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A(\lambda) \begin{pmatrix} x \\ y \end{pmatrix} \tag{8.1}$$

has the fixed point 0. For $\lambda < 0$ we have $\mathrm{Re}\,\mu_{1,2}(\lambda) < 0$, thus the origin is asymptotically stable. For $\lambda > 0$ the origin is unstable. For $\lambda = 0$ the system (8.1) has a branch of stable periodic orbits given by the solutions

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = a \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

where $a > 0$ is the amplitude.

A bifurcation diagram for Example 1:

Example 2: A supercritical Hopf bifurcation. Consider the ODE system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A(\lambda) \begin{pmatrix} x \\ y \end{pmatrix} - \sqrt{x^2 + y^2} \begin{pmatrix} x \\ y \end{pmatrix}\ . \tag{8.2}$$

As in Example 1, the origin is an asymptotically stable fixed point for $\lambda < 0$, which becomes unstable for $\lambda > 0$. What happens to the branch of periodic orbits?

We write the system in polar coordinates: For $r = \sqrt{x^2 + y^2}$ obtain

$$
\begin{aligned}
rr' &= \frac{1}{2}(r^2)' \\
&= xx' + yy' \\
&= \lambda x^2 - xy - rx^2 + yx + \lambda y^2 - ry^2 \\
&= \lambda r^2 - r^3
\end{aligned}
$$

thus $r' = \lambda r - r^2 = r(\lambda - r)$. Also, for $\theta = \arctan(y/x)$ obtain

$$
\begin{aligned}
\theta' &= \frac{1}{1 + y^2/x^2} \cdot \frac{y'x - yx'}{x^2} \\
&= \frac{1}{r^2}\,(xy' - yx') \\
&= \frac{1}{r^2}\,(x^2 + \lambda xy - rxy - \lambda xy + y^2 + rxy) \\
&= 1
\end{aligned}
$$

In polar coordinates the system (8.2) reads

$$
r' = r(\lambda - r), \quad \theta' = 1 \ . \tag{8.3}
$$

For $\lambda > 0$ one obtains the periodic solutions

$$
r(t) \equiv \lambda, \quad \theta(t) = t \bmod 2\pi \ .
$$

The function $g(r) = \lambda r - r^2$ satisfies

$$
g'(\lambda) = -\lambda < 0 \quad \text{for} \quad \lambda > 0 \ .
$$

Therefore, the periodic orbits are asymptotically stable.

A bifurcation diagram for Example 2:

When λ changes from $\lambda < 0$ to $\lambda > 0$ the asymptotically stable state $x = y = 0$ is replaced by a stable periodic solution of small amplitude, $a = \lambda$. Thus, when λ passes the bifurcation point at $\lambda = 0$, a soft transition occurs.

Example 3: A subcritical Hopf bifurcation. In Example 2 we subtracted the quadratic term

$$\sqrt{x^2+y^2}\begin{pmatrix} x \\ y \end{pmatrix}$$

on the right side of the ODE system. We will now add this term and consider the equation

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A(\lambda)\begin{pmatrix} x \\ y \end{pmatrix} + \sqrt{x^2+y^2}\begin{pmatrix} x \\ y \end{pmatrix} . \tag{8.4}$$

As in Examples 1 and 2, the origin is an asymptotically stable fixed point for $\lambda < 0$, which becomes unstable for $\lambda > 0$. In polar coordinates the system (8.4) reads

$$r' = r(\lambda + r), \quad \theta' = 1 .$$

For $\lambda < 0$ one obtains the periodic solutions

$$r(t) \equiv -\lambda, \quad \theta(t) = t \bmod 2\pi$$

with amplitude $a = -\lambda$. For $g(r) = \lambda r + r^2$ we have

$$g'(-\lambda) = -\lambda > 0 \quad \text{for} \quad \lambda < 0 .$$

Therefore, the periodic orbits are unstable.

As λ changes from $\lambda < 0$ to $\lambda > 0$ one has a sharp (or catastrophic) transition: For $\lambda < 0$ the origin is a stable state. For $\lambda > 0$ the origin becomes unstable, and every non–zero solution blows up in finite time.

A bifurcation diagram for Example 3:

8.2 Hopf Bifurcation in 2D: Setup in Polar Coordinates

Consider the following system in polar coordinates:

$$\begin{aligned} \frac{dr}{dt} &= \lambda r + s_1(r,\theta,\lambda) \\ \frac{d\theta}{dt} &= 1 + s_2(r,\theta,\lambda) \end{aligned}$$

where the s_j are smooth functions, which are 2π–periodic in θ. We assume that $r_0 > 0, \lambda_0 > 0$ and

$$s_{1,2} \ : \ [-r_0, r_0] \times \mathbb{R} \times [-\lambda_0, \lambda_0] \to \mathbb{R} \ .$$

Assume the following estimates for all arguments

$$(r, \theta, \lambda) \in [-r_0, r_0] \times \mathbb{R} \times [-\lambda_0, \lambda_0] \ :$$

$$\begin{aligned} |s_1| + |s_{1\lambda}| &\leq Cr^2 \\ |s_2| + |s_{2\lambda}| + |s_{1r}| &\leq C|r| \\ |s_{2r}| &\leq C \end{aligned}$$

In the following, we will assume that λ_0 and r_0 are small in absolute value. Eliminate t to obtain

$$\frac{dr}{d\theta} = \lambda r + s(r, \theta, \lambda) \tag{8.5}$$

where $|s(r, \theta, \lambda)| \leq Cr^2$.

Details: We have

$$\begin{aligned} \frac{dr}{d\theta} &= \frac{\lambda r + s_1}{1 + s_2} \\ &= \frac{\lambda r(1 + s_2)}{1 + s_2} + \frac{s_1 - \lambda r s_2}{1 + s_2} \\ &= \lambda r + s \end{aligned}$$

where

$$|s| \leq 2(|s_1| + |\lambda||r||s_2) \leq Cr^2 \ .$$

Similarly, one can derive the bounds

$$\begin{aligned} |s_\lambda| &\leq Cr^2 \\ |s_r| &\leq C|r| \end{aligned}$$

Consider the ODE (8.5). Given an initial condition

$$r(0) = a \quad \text{where} \quad |a| < a_0 \ ,$$

we denote the solution of (8.5) by

$$r(\theta; \lambda, a) \ .$$

For $|a| < a_0$ we obtain the Poincaré map P depending on the parameter λ:

$$(\lambda, a) \to P(\lambda, a) := r(2\pi; \lambda, a) \ .$$

If $0 < a < a_0$ and we have

$$P(\lambda, a) = a \ ,$$

then

$$r(2\pi; \lambda, a) = a = r(0; \lambda, a) \ .$$

If this holds then the function $\theta \to r(\theta; \lambda, a)$ determines a periodic orbit γ_a. The number $a > 0$ can be thought of as the (approximate) amplitude of the orbit.

This motivates to consider the equation

$$r(2\pi; \lambda, a) = a, \quad |a| < a_0 \ , \tag{8.6}$$

to determine a solution $\lambda = \lambda^*(a)$.

We will use the implicit function theorem to show that, for $|a| < a_0$, the equation (8.6) has a unique smooth solution $\lambda^*(a)$ with $\lambda^*(0) = 0$ if $a_0 > 0$ is sufficiently small.

If

$$\frac{d\lambda^*}{da}(a) > 0 \quad \text{for} \quad 0 < a < a_0$$

then we obtain a branch of periodic orbits for positive λ values. We will show that the orbits are asymptotically stable: We have a supercritical Hopf bifurcation.

If

$$\frac{d\lambda^*}{da}(a) < 0 \quad \text{for} \quad 0 < a < a_0$$

then we obtain a branch of periodic orbits for negative λ values. We will show that the orbits are unstable: We have a subcritical Hopf bifurcation.

8.3 Hopf Bifurcation in 2D: Application of the Implicit Function Theorem

We have

$$\frac{dr}{d\theta} = \lambda r + s(r, \theta, \lambda) \quad \text{and} \quad r(0) = a \ .$$

Therefore,

$$r(2\pi; \lambda, a) = e^{2\pi\lambda} a + \int_0^{2\pi} e^{\lambda(2\pi - \tau)} s(r(\tau; \lambda, a), \tau, \lambda) \, d\tau \ .$$

Remark: Since $s|(r, \theta, \lambda)| \leq Cr^2$ one can show that the absolute value of the integral can be bounded by Ca^2 if $|a| < a_0$ and a_0 is sufficiently small.

We rewrite the equation

$$r(2\pi;\lambda,a)=a$$

as

$$(1-e^{-2\pi\lambda})a+\int_0^{2\pi} e^{-\lambda\tau}s(r(\tau;\lambda,a),\tau,\lambda)\,d\tau=0\ .$$

For $a\neq 0$ divide by a and then extend the left–hand side smoothly to $a=0$. This leads us to define the function $h(\lambda,a)$ by

$$h(\lambda,a)=1-e^{-2\pi\lambda}+\frac{1}{a}\int_0^{2\pi} e^{-\lambda\tau}s(r(\tau;\lambda,a),\tau,\lambda)\,d\tau \quad \text{for} \quad 0<|a|<a_0$$

and

$$h(\lambda,a)=1-e^{-2\pi\lambda} \quad \text{for} \quad a=0\ .$$

Using that $s=\mathcal{O}(r^2)$ and $|r|\leq C|a|$ one obtains that the integral is $\mathcal{O}(a^2)$. Therefore, $h(\lambda,a)$ is a smooth function.

Since

$$h(0,0)=0 \quad \text{and} \quad h_\lambda(0,0)=2\pi\neq 0$$

the implicit function theorem gives us a unique smooth function $\lambda^*(a)$ with

$$h(\lambda^*(a),a)=0 \quad \text{for} \quad |a|<a_0 \quad \text{and} \quad \lambda^*(0)=0\ .$$

For $a>0$ we obtain a corresponding periodic orbit γ_a for the original equation (8.5) with $\lambda=\lambda^*(a)$.

8.4 Stability or Instability of the Periodic Orbits

Fix $\bar{a}>0$ and set $\bar{\lambda}=\lambda^*(\bar{a})$. We want to understand the stability or instability of the periodic orbit $\gamma_{\bar{a}}$.

To this end, consider the Poincaré map

$$a\to P(\bar{\lambda},a)\ .$$

We first show that $P_a(\bar{\lambda},a)>0$. We have

$$P(\bar{\lambda},a)=r(2\pi;\lambda,a)\ ,$$

thus

$$P_a(\bar{\lambda},a)=r_a(2\pi;\lambda,a)=e^{2\pi\lambda}+\mathcal{O}(|a|)\ .$$

Here we have used that $|s_r|\leq C|r|$. The equation

$$P_a(\bar{\lambda},a)=e^{2\pi\lambda}+\mathcal{O}(|a|)$$

implies that $P_a(\bar{\lambda}, a) > 0$ if a_0 is sufficiently small.

If

$$P_a(\bar{\lambda}, a)|_{a=\bar{a}} < 1 \tag{8.7}$$

then the orbit is asymptotically stable. This holds since the condition (8.7) implies that $\bar{a}$ is a stable fixed point of the map $a \to P(\bar{\lambda}, a)$.

For the same reason, if

$$P_a(\bar{\lambda}, a)|_{a=\bar{a}} > 1$$

then the orbit $\gamma_{\bar{a}}$ is unstable.

We now relate the sign of

$$1 - P_a(\bar{\lambda}, a)|_{a=\bar{a}}$$

to the sign of

$$\frac{d\lambda^*}{da}(\bar{a}) \ .$$

We have

$$a = P(\lambda^*(a), a) \ ,$$

thus

$$1 = P_\lambda(\lambda^*(a), a)\frac{d\lambda^*}{da}(a) + P_a(\lambda^*(a), a) \ .$$

At $a = \bar{a}$:

$$1 - P_a(\bar{\lambda}, \bar{a}) = P_\lambda(\bar{\lambda}, \bar{a})\frac{d\lambda^*}{da}(\bar{a}) \ .$$

We have to show that

$$P_\lambda(\bar{\lambda}, \bar{a}) > 0 \quad \text{for} \quad \bar{a} > 0 \ .$$

From

$$P(\lambda, a) = r(2\pi; \lambda, a)$$

we have

$$P_\lambda(\lambda, a) = r_\lambda(2\pi; \lambda, a) \ .$$

Recall that

$$r(2\pi; \lambda, a) = e^{2\pi\lambda}a + \int_0^{2\pi} e^{\lambda(2\pi-\tau)}s(r(\tau; \lambda, a), \tau, \lambda)\, d\tau \ .$$

We can also write this as

$$r(2\pi; \lambda, a) = e^{2\pi\lambda}a + Int(\lambda, a)$$

where $Int(\lambda, a)$ denotes the integral. Therefore,

$$r_\lambda(2\pi; \lambda, a) = 2\pi e^{2\pi\lambda} a + Int_\lambda(\lambda, a) \ .$$

Using that

$$|s| + |s_\lambda| + |s_r r_\lambda| = \mathcal{O}(a^2)$$

it follows that

$$r_\lambda(2\pi; \lambda, a) > 0 \quad \text{for} \quad 0 < a < a_0 \ .$$

Summary: There are three cases:

Case 1: $\frac{d\lambda^*}{da}(a) > 0$ for $0 < a < a_0$: **A supercritical Hopf bifurcation.**
In this case we have

$$0 < \lambda^*(a) \quad \text{for} \quad 0 < a < a_0 \ .$$

The original system has a periodic orbit of (approximate) amplitude a for $0 < \lambda < \lambda_1$. Each orbit is asymptotically orbitally stable.

Case 2: $\frac{d\lambda^*}{da}(0) < 0$ for $0 < a < a_0$: **A subcritical Hopf bifurcation.**
In this case we have

$$0 > \lambda^*(a) \quad \text{for} \quad 0 < a < a_0 \ .$$

The original system has a periodic orbit of (approximate) amplitude a for $0 > \lambda > -lam_2$. Each orbit is orbitally unstable.

Case 3: Neither Case 1 nor Case 2 holds. In this case $\frac{d\lambda^*}{da}(0) = 0$.

The simplest example for this case is the system (8.1), which reads in polar coordinates

$$r' = \lambda r, \quad \theta' = 1 \ .$$

The periodic solution of amplitude $a > 0$ is

$$r(t) \equiv a, \quad \theta(t) = t \bmod 2\pi$$

for $\lambda = \lambda^*(a) = 0$. Clearly, in this case $\frac{d\lambda^*}{da}(0) = 0$. Since the parameter λ does not change along the branch of periodic solutions, the equation (8.1) leads to a degenerate Hopf bifurcation.

The following example shows that the condition $\frac{d\lambda^*}{da}(0) = 0$ of Case 3 can also occur for non–degenerate branches.

Example 4: The system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A(\lambda) \begin{pmatrix} x \\ y \end{pmatrix} - (x^2 + y^2) \begin{pmatrix} x \\ y \end{pmatrix} \tag{8.8}$$

reads in polar coordinates

$$r' = r(\lambda - r^2), \quad \theta' = 1 \ .$$

If $\lambda = a^2$ then the constant function

$$\theta \to r(\theta; \lambda, a) \equiv a$$

determines a periodic orbit for $a > 0$. Thus we have

$$\lambda^*(a) = a^2, \quad \frac{d\lambda^*}{da}(0) = 0 \ .$$

Since λ changes with a, the branch of periodic orbits is non–degenerate.

For fixed $a > 0$ the function

$$\phi(r) = r(a^2 - r^2) \quad \text{with} \quad \phi'(r) = a^2 - r^2 - 2r^2$$

satisfies

$$\phi'(a) = -2a^2 < 0 \ .$$

This yields that the orbits given by $r(\theta; \lambda, a) \equiv a$ are stable.

8.5 General Theory in 2D

Let $\Lambda = [-\lambda_0, \lambda_0]$ and let

$$f : \mathbb{R}^2 \times \Lambda \to \mathbb{R}^2$$

denote a smooth function with $|Df(x, y, \lambda)| \le C$ and $f(0, 0, \lambda) \equiv 0$.

The ODE system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = f(x, y, \lambda) = \begin{pmatrix} f_1(x, y, \lambda) \\ f_2(x, y, \lambda) \end{pmatrix} \tag{8.9}$$

has the fixed point $(x, y) = 0$ for $-\lambda_0 \le \lambda \le \lambda_0$.

Set

$$A(\lambda) = Df(0, 0, \lambda) \in \mathbb{R}^{2\times 2}$$

where

$$Df = \begin{pmatrix} f_{1x} & f_{1y} \\ f_{2x} & f_{2y} \end{pmatrix} \ .$$

Denote the eigenvalues of $A(\lambda)$ by $\mu_{1,2}(\lambda)$.

Assume that

$$\mu_{1,2}(\lambda) = \alpha(\lambda) \pm i\beta(\lambda), \quad \alpha(\lambda), \beta(\lambda) \in \mathbb{R} \ ,$$

were $\alpha(\lambda) < 0$ for $-\lambda_0 \le \lambda < 0$. Also, assume

$$\alpha(0) = 0, \quad \frac{d\alpha}{d\lambda}(0) > 0, \quad \beta(0) > 0 \ .$$

We will show that under these assumptions a Hopf–bifurcation occurs at $\lambda = 0$.

We can write the given system in the form

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A(\lambda) \begin{pmatrix} x \\ y \end{pmatrix} + Q(x,y,\lambda) \tag{8.10}$$

where

$$\begin{aligned} |Q| + |Q_\lambda| &\le C(x^2+y^2) \\ |Q_x| + |Q_y| &\le C(|x|+|y|) \end{aligned}$$

for all arguments (x,y,λ).

We make three pretransformations and will then use polar coordinates:

a) A linear change of variables: There exists a nonsingular matrix $V(\lambda) \in \mathbb{R}^{2\times 2}$ so that

$$V^{-1}(\lambda)A(\lambda)V(\lambda) = \begin{pmatrix} \alpha(\lambda) & -\beta(\lambda) \\ \beta(\lambda) & \alpha(\lambda) \end{pmatrix} .$$

For simplicity of notation, we assume that the system (8.10) already has the form

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} \alpha(\lambda) & -\beta(\lambda) \\ \beta(\lambda) & \alpha(\lambda) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + Q(x,y,\lambda) . \tag{8.11}$$

b) Rescaling of time: Introduce new variables

$$\tilde{x}(\tau), \quad \tilde{y}(\tau)$$

by

$$x(t) = \tilde{x}(\beta(\lambda)t), \quad y(t) = \tilde{y}(\beta(\lambda)t) .$$

Then, with $\beta = \beta(\lambda)$:

$$x' = \beta\tilde{x}', \quad y' = \beta\tilde{y}'$$

and the equation

$$x' = \alpha x + \beta y + Q_1$$

becomes

$$\beta\tilde{x}' = \alpha\tilde{x} - \beta\tilde{y} + Q_1 .$$

We can divide by β and obtain

$$\tilde{x}' = \tilde{\alpha}\tilde{x} - \tilde{y} + \tilde{Q}_1 .$$

Note: $\tilde{\alpha}(\lambda) = \alpha(\lambda)/\beta(\lambda)$ and $\alpha(0) = 0$ yields that $\tilde{\alpha}'(0) = \alpha'(0)/\beta(0) > 0$.

Thus, using the new variables, $\beta(\lambda)$ becomes 1. For simplicity of notation, we assume that the given system has the form

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} \alpha(\lambda) & -1 \\ 1 & \alpha(\lambda) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + Q(x,y,\lambda) \ . \tag{8.12}$$

c) Rescaling of α: Let $\tilde{\lambda} = \alpha(\lambda)$. Since $\frac{d\alpha}{d\lambda}(0) > 0$ we can express λ as a function of $\tilde{\lambda}$. Then, dropping the tilde–notation, the system (8.12) becomes

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + Q(x,y,\lambda) \tag{8.13}$$

where $|Q(x,y,\lambda)| \leq C(x^2+y^2)$.

After these pretransformations, we introduce polar coordinates:

$$r = \sqrt{x^2+y^2}, \quad \theta = \arctan(y/x) \ .$$

Obtain:

$$\begin{aligned} rr' &= xx' + yy' \\ &= \lambda x^2 - xy + xQ_1 + yx + \lambda y^2 + yQ_1 \\ &= \lambda r^2 + r\cos\theta \ Q_1 + r\sin\theta \ Q_2 \end{aligned}$$

This yields that

$$r' = \lambda r + s_1(r,\theta,\lambda)$$

where $|s_1(r,\theta,\lambda)| \leq C_1 r^2$.

Also,

$$\begin{aligned} \theta' &= \frac{1}{1+y^2/x^2} \cdot \frac{y'x - yx'}{x^2} \\ &= \frac{1}{r^2}\Big(x^2 + \lambda xy + xQ_1 - \lambda xy + y^2 - yQ_2\Big) \\ &= \frac{1}{r^2}(r^2 + xQ_1 - yQ_2) \\ &= 1 + \frac{1}{r}(\cos\theta \ Q_1 - \sin\theta \ Q_2) \\ &= 1 + s_2(r,\theta,\lambda) \end{aligned}$$

where $|s_2(r,\theta,\lambda)| \leq C_2|r|$. $|Q(x,y,\lambda)| \leq C(x^2+y^2)$.

The system that we have obtained,

$$\begin{aligned} \frac{dr}{dt} &= \lambda r + s_1(r,\theta,\lambda) \\ \frac{d\theta}{dt} &= 1 + s_2(r,\theta,\lambda) \end{aligned}$$

satisfies the conditions that were required in Sections 8.2, 8.3, and 8.4.

This, after all, implies that a Hopf bifurcation occurs under the assumptions that we made for the system (8.9).

9 The Poincaré–Bendixson Theorem; Uncoupled Oscillators

9.1 General Concepts Related to Asymptotic Behavior

Let $f : \mathbb{R}^N \to \mathbb{R}^N, f \in C^1, |Df(x)| \le L$ for all $x \in \mathbb{R}^N$. We denote the solution of

$$x' = f(x), \quad x(0) = x_0 \ ,$$

by $\phi(t, x_0)$. Recall: If $a, b \in \mathbb{R}^N$ then

$$|\phi(t,a) - \phi(t,b)| \le e^{L|t|}|a-b| \quad \text{for} \quad t \in \mathbb{R} \ . \tag{9.1}$$

The point sets

$$\gamma(x_0) = \{\phi(t,x_0) \ : \ t \in \mathbb{R}\} \quad \text{and} \quad \gamma^+(x_0) = \{\phi(t,x_0) \ : \ t \ge 0\}$$

are called the orbit and the positive semi–orbit of the solution $t \to \phi(t, x_0)$.

The set

$$\omega(x_0) = \{y \in \mathbb{R}^N \ : \ \phi(t_n, x_0) \to y \text{ for some sequence } t_n \to \infty\}$$

is called the omega–limit set of the positive semi–orbit $\gamma^+(x_0)$.

Examples:
1) If $f(u^*) = 0$ and $\phi(t, x_0) \to u^*$ as $t \to \infty$ then $\omega(x_0) = \{u^*\}$.
2) If $= \phi(t, x_0)$ is T–periodic, then

$$\omega(x_0) = \{\phi(t,x_0) \ : \ 0 \le t \le T\} = \gamma^+(x_0) = \gamma(x_0) \ .$$

3) Assume that u^* is an equilibrium of the ODE $x' = f(x)$ and assume that there are two homoclinic orbits $\mathcal{O}_1$ and $\mathcal{O}_2$ for u^*. Then it is possible that

$$\omega(x_0) = \mathcal{O}_1 \cup \mathcal{O}_2 \cup \{u^*\}$$

for some $x_0 \in \mathbb{R}^N$. See **Figure 9.1.**

4) Let $S^1 = \mathbb{R} \mod 1$. Consider the 2–torus $\mathbb{T}^2 = S^1 \times S^1$ and consider the dynamical system defined by

$$\theta_1' = \omega_1 \tag{9.2}$$
$$\theta_2' = \omega_2 \tag{9.3}$$

where ω_1 and ω_2 are positive frequencies. Given initial conditions

$$\theta_j(0) = \theta_{j0} \quad \text{for} \quad j = 1, 2$$

obtain the solutions

$$\theta_j(t) = (\theta_{j0} + \omega_j t) \quad \text{mod } 1 \ .$$

The solution of the system (9.2), (9.3) with initial value $\theta_0 = (\theta_{10}, \theta_{20})$ is

$$\phi(t, \theta_0) = \begin{pmatrix} (\theta_{10} + \omega_1 t) & \text{mod } 1 \\ (\theta_{20} + \omega_2 t) & \text{mod } 1 \end{pmatrix}$$

One can prove: If the solution $\phi(t, \theta_0)$ is periodic for some initial vector θ_0 then the quotient

$$\frac{\omega_2}{\omega_1}$$

is a rational number and all solutions of the system (9.2), (9.3) are periodic.

If ω_2/ω_1 is irrational then, for all initial data θ_0, the omega–limit set of the semi–orbit $\gamma^+(\theta_0)$ equals the whole torus,

$$\omega(\theta_0) = \mathbb{T}^2 \ .$$

Furthermore, the following holds if ω_2/ω_1 is irrational: Let

$$R = [a_1, b_1] \times [a_2, b_2] \subset \mathbb{T}^2 \ ,$$

i.e., R is a rectangle in the torus $\mathbb{T}^2$. Let $\theta(t)$ denote a solution of the system (9.2), (9.3) and let $T > 0$. Let

$$S(T, R) = \{ t \ : \ 0 \le t \le T, \phi(t) \in R \} \ .$$

Then

$$\lim_{T\to\infty} \frac{1}{T} \ measure\Big(S(T, R) \Big) = area(R) \ ,$$

i.e., the long–time average, which the orbit of any solution spends in R, equals the area of R.

Definition: *A topological space X is called connected if X cannot be written as the union of two disjoint non–empty open subsets of X.*

By definition, a subset M of a topological space X is called closed if its complement $X \setminus M$ is open. Therefore, X is connected if and only if X cannot be written as the union of two disjoint, non–empty, closed subsets of X. In other words, X is disconnected if X can be written in the form $X = M_1 \cup M_2$ where M_1, M_2 are disjoint, non–empty, closed subsets of X.

Theorem 9.1 *As above, let $\gamma^+(x_0)$ denote the positive semi–orbit of $\phi(t, x_0)$. If $\gamma^+(x_0)$ is bounded then $\omega(x_0)$ is non–empty, compact, and connected.*

Proof: It is clear that $\omega(x_0)$ is non–empty and bounded.

Let us show that $\omega(x_0)$ is closed. Set $x(t) = \phi(t, x_0)$. Let $q_n \in \omega(x_0)$ and $q_n \to q$. There exists $t_n > n$ with

$$|q_n - x(t_n)| < \frac{1}{2n} \ .$$

Then, for all large n:

$$|q - x(t_n)| \le |q - q_n| + |q_n - x(t_n)| \le \frac{1}{n} \; .$$

Therefore, $x(t_n) \to q$, which yields that $q \in \omega(x_0)$. This shows that $\omega(x_0)$ is closed.

Suppose that $\omega(x_0)$ is not connected. Then there exist two non–empty sets M_1, M_2, which are closed as subsets of $\omega(x_0)$, with

$$M_1 \cup M_2 = \omega(x_0), \quad M_1 \cap M_2 = \emptyset \; .$$

Since $\omega(x_0)$ is a closed subset of $\mathbb{R}^N$, the sets M_j are also closed subsets of $\mathbb{R}^N$. Since they are bounded, they are compact. Therefore,

$$dist(M_1, M_2) =: \delta > 0 \; .$$

There are time sequences $s_n \to \infty$ and $t_n \to \infty$ with

$$dist(x(s_n), M_1) \to 0 \quad \text{and} \quad dist(x(t_n), M_2) \to 0 \; .$$

For all large n:

$$dist(x(s_n), M_1) \le \frac{\delta}{4}, \quad dist(x(t_n), M_1) \ge \frac{3\delta}{4} \; .$$

By continuity, there exists ξ_n between s_n and t_n with

$$dist(x(\xi_n), M_1) = \frac{\delta}{2} \; .$$

It then follows that

$$dist(x(\xi_n), M_2) \ge \frac{\delta}{2} \; .$$

Since the sequence $x(\xi_n)$ is bounded, there exists a convergent subsequence:

$$x(\xi_{n_k}) \to r \; .$$

The point r lies in $\omega(x_0)$ but neither lies in M_1 nor in M_2. This contradiction proves connectedness of $\omega(x_0)$. $\diamond$

9.2 Periodic Orbits in 2D

Theorem 9.2 (Poincaré–Bendixson) *Let $f : \mathbb{R}^2 \to \mathbb{R}^2, f \in C^1, |Df(x)| \le L$ for all $x \in \mathbb{R}^2$. Let $x_0 \in \mathbb{R}^2$ and assume:*

a) $\gamma^+(x_0)$ is bounded.

b) If $a \in \omega(x_0)$ then $f(a) \ne 0$.

Under these assumptions, the set $\omega(x_0)$ is a periodic orbit. Furthermore,

i) Either $\omega(x_0) = \gamma^+(x_0)$;

ii) or $\omega(x_0) = cl(\gamma^+(x_0)) \setminus \gamma^+(x_0)$.

In the second case, the set $\omega(x_0)$ is the limit cycle of the positive semi–orbit $\gamma^+(x_0)$. Case i) occurs if and only if $\gamma^+(x_0) \cap \omega(x_0)$ is not empty.

In the above theorem, $cl(\gamma^+(x_0))$ denotes the closure of the set $\gamma^+(x_0)$ as a subset on $\mathbb{R}^2$.

An essential tool for proving the Poincaré–Bendixson Theorem is the Jordan Curve Theorem.

Remark: The Jordan Curve Theorem is due to Camille Jordan (1838–1922). C. Jordan is also known for the transformation of matrices to Jordan normal form.

Definition: Let

$$S^1 = \{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 = 1\}$$

denote the unit circle in $\mathbb{R}^2$. A set $\Gamma \subset \mathbb{R}^2$ is called a **Jordan curve** if there exists a continuous mapping

$$h : S^1 \to \Gamma$$

which is one–to–one, onto, so that $\phi^{-1} : \Gamma \to S^1$ is also continuous. In other words, a Jordan curve Γ is a subset of $\mathbb{R}^2$ which is homeomorphic to the unit circle S^1.

Theorem 9.3 (Jordan Curve Theorem) *If $\Gamma \subset \mathbb{R}^2$ is a Jordan curve, then one can write*

$$\mathbb{R}^2 \setminus \Gamma = A \cup B$$

where A and B are non–empty, disjoint, open, connected subsets of $\mathbb{R}^2$. The set A is bounded and is called the interior of Γ; the set B is unbounded and called the exterior.

The Jordan Curve Theorem seems rather obvious, but turns out to be difficult to prove. One can prove it using tools of algebraic topology.

Before proving the Poincaré–Bendixson Theorem, we will prove auxiliary lemmas.

In the following we will assume:

a) $f : \mathbb{R}^2 \to \mathbb{R}^2, f \in C^1, |Df(x)| \leq L$ for all $x \in \mathbb{R}^2$.
b) $x_0 \in \mathbb{R}^2$ and $\gamma^+(x_0)$ is bounded.
c) For all $a \in \omega(x_0)$ we have $f(a) \neq 0$.

Notation: Let $a \in \mathbb{R}^2$ and let $f(a) \neq 0$. Let $L = L(a, f(a))$ denote the straight line through a orthogonal to $f(a)$. For $\Delta > 0$ consider the piece of line

$$L_\Delta = L_\Delta(a) := L \cap B_\Delta(a) \ .$$

See **Figure 9.2.**

If $\Delta > 0$ is small enough, then

$$\langle f(y), f(a) \rangle > 0 \quad \text{for all} \quad y \in L_\Delta \ . \tag{9.4}$$

The line L divides the plane into two half–planes. If (9.4) holds and if $y \in L_\Delta$ then the vectors $f(y)$ and $f(a)$ point into the same half–plane.

In the following lemma, we track points near a to the line L using trajectories of the equation $x' = f(x)$.

Lemma 9.1 (Tracking Lemma) *Let $a \in \mathbb{R}^2$ and let $f(a) \neq 0$. There exist $\varepsilon_0 > 0$ and $\delta_0 > 0$ so that the following holds: For all $x \in B_{\delta_0}(a)$ there is a unique time $t^*(x)$ with*

$$\phi(t^*(x), x) \in L, \quad |t^*(x)| < \varepsilon_0 \ .$$

The function $x \to t^*(x)$ is C^1 and satisfies

$$t^*(x) = 0 \quad \text{for all} \quad x \in L \cap B_{\delta_0}(a) \ .$$

Furthermore, by choosing δ_0 small enough, we may assume that all points

$$y = \phi(t^*(x), x), \quad x \in B_{\delta_0}(a) \ ,$$

lie on a line–piece L_Δ with (9.4).

See **Figure 9.3.**
Proof: We know that $\phi(t, x) \in L$ is equivalent to

$$h(t, x) := \langle \phi(t, x) - a, f(a) \rangle = 0 \ .$$

For the function $h(t, x)$ we have $h(0, a) = 0$ and

$$h_t(0, a) = \langle \phi_t(0, a), f(a) \rangle = |f(a)|^2 > 0 \ .$$

The existence, uniqueness, and smoothness of t^* follows from the Implicit Function Theorem. $\diamond$

The following lemma contains a part of the Poincaré–Bendixson Theorem.

Lemma 9.2 *Under the assumptions of the Poincaré–Bendixson Theorem, the following conditions are equivalent:*

a) The solution $\phi(t, x_0)$ is periodic.
b) $\gamma^+(x_0) = \omega(x_0)$
c) $\gamma^+(x_0) \cap \omega(x_0) \neq \emptyset$

Proof:
a) implies b): Let $\phi(t, x_0)$ have the period T. Let $y \in \gamma^+(x_0)$. There exists $t \geq 0$ with

$$y = \phi(t, x_0) = \phi(t + nT, x_0) \quad \text{for all} \quad n \in \mathbb{N} \ .$$

This yields that $y \in \omega(x_0)$. Thus we have shown that $\gamma^+(x_0) \subset \omega(x_0)$.

Conversely, let $y \in \omega(x_0)$, thus

$$y = \lim_{n \to \infty} \phi(t_n, x_0)$$

for a sequence $t_n \to \infty$. By assumption a) the function $\phi(t, x_0)$ is periodic, which implies that $\gamma^+(x_0)$ is a closed subset of $\mathbb{R}^2$. Therefore, $y \in \gamma^+(x_0)$. We have shown that a) implies b).

b) implies c) is trivial.

c) implies a): Let $a \in \gamma^+(x_0) \cap \omega(x_0)$, thus $f(a) \neq 0$ by assumption. We will use the line segment $L_\Delta = L \cap B_\Delta(a)$ introduced above.

Since $a \in \gamma^+(x_0)$ there exists $t_0 \geq 0$ with $a = \phi(t_0, x_0)$. Then, given $\varepsilon > 0$, there exists $t_1 > 0$ so that

$$x_1 := \phi(t_1, a) \in L_\Delta, \quad |x_1 - a| < \varepsilon \ .$$

To obtain this result we have used that $a \in \omega(x_0)$ and have used the Tracking Lemma.

If $x_1 = a$ then it follows that the function $\phi(t, a)$ has the period t_1 and a) is shown.

Now suppose that $x_1 \in L_\Delta$, but $x_1 \neq a$. See **Figure 9.4.** Let $\mathcal{O}$ denote the piece of orbit given by $\phi(t, a)$ for $0 \leq t \leq t_1$ and let Γ denote the Jordan curve consisting of $\mathcal{O}$ and the line segment on L from x_1 to a. It is then clear that the orbit $\phi(t, x_1), t > 0$, lies inside Γ and cannot approach a. This contradicts the assumption that $a \in \omega(x_0)$. Therefore, $x_1 \neq a$ leads to a contradiction. This proves the lemma. $\diamond$

Remark: The implication

$$\gamma^+(x_0) \cap \omega(x_0) \neq \emptyset \implies \gamma^+(x_0) = \omega(x_0) \tag{9.5}$$

does not hold in $\mathbb{R}^N$ for $N \geq 3$. There can be a point $P \in \gamma^+(x_0)$ which can be approximated by $\phi(t_n, x_0)$ as $t_n \to \infty$ and $\omega(x_0) = \{P\}$ is possible. See **Figure 9.5.** The implication (9.5) also does not hold for ODEs on $\mathbb{T}^2$.

To complete the proof of Theorem 9.2 it remains to consider the case

$$\gamma^+(x_0) \cap \omega(x_0) = \emptyset \tag{9.6}$$

and to prove: If (9.6) holds then $\omega(x_0)$ is a periodic orbit which is the limit cycle of $\gamma^+(x_0)$. The positive semi–orbit $\gamma^+(x_0)$ spirals towards the periodic orbit $\omega(x_0)$ as $t \to \infty$.

We start with the following simple observation:

Lemma 9.3 *Let the assumptions of Theorem 9.2 hold. If $a \in \omega(x_0)$ then all all points $\phi(t, a), t \in \mathbb{R}$, lie in $\omega(x_0)$, i.e., $\gamma(a) \subset \omega(x_0)$.*

Proof: Since $a \in \omega(x_0)$ there exists a sequence of times $t_n \to \infty$ so that

$$x_n := \phi(t_n, x_0) \to a \quad \text{as} \quad n \to \infty \ .$$

Let $t \in \mathbb{R}$ be arbitrary. Using (9.1) we have

$$\phi(t + t_n, x_0) = \phi(t, x_n) \to \phi(t, a) \quad \text{as} \quad n \to \infty$$

and $\phi(t, a) \in \omega(x_0)$ follows. $\diamond$

Lemma 9.4 *Assume that $\gamma(a)$ is a periodic orbit of the system $x' = f(x)$. Consider the line piece $L_\Delta = L \cap B_\Delta(a)$ where L is the straight line through a orthogonal to $f(a)$. If $\Delta > 0$ is sufficiently small then the orbit $\gamma(a)$ intersects L_Δ only in a.*

Proof: Suppose $\gamma(a)$ intersects L_Δ also in $b \neq a$. See **Figure 9.6.** Using the Jordan Curve Theorem, one obtains a contradiction. $\diamond$

Lemma 9.5 *Let the assumptions of Theorem 9.2 hold and let $a \in \omega(x_0)$, thus $f(a) \neq 0$. As in the previous lemma, consider the line piece $L_\Delta = L \cap B_\Delta(a)$ where $\Delta > 0$ is sufficiently small. If $b \in L_\Delta \cap \omega(x_0)$ then $b = a$.*

Proof: Since $a, b \in \omega(x_0)$ the orbit $\gamma^+(x_0)$ approaches a and b for large times. Using the Jordan Curve Theorem one obtains that $a \neq b$ is not possible. See **Figure 9.7.** $\diamond$

Lemma 9.6 *Let the assumptions of Theorem 9.2 hold and let $a \in \omega(x_0)$, thus $\gamma(a) \subset \omega(x_0)$ by Lemma 9.3. If $\gamma(a)$ is a periodic orbit then $\gamma(a) = \omega(x_0)$.*

Proof: Suppose that $\omega(x_0) \setminus \gamma(a) \neq \emptyset$. Recall that $\omega(x_0)$ is connected by Theorem 9.1. Since the periodic orbit $\gamma(a)$ is closed, the set $\omega(x_0) \setminus \gamma(a)$ is not closed. There exists a point $Q \in \gamma(a)$ which is a limit of points $b_n \in \omega(x_0) \setminus \gamma(a)$.

Consider the line piece $L_\Delta(Q)$ through the point Q. See **Figure 9.8.** Use the Tracking Lemma to obtain that there exist points $a_n \in L_\Delta(Q)$ and small times t_n with $|t_n| < \varepsilon$ so that $\phi(t_n, a_n) = b_n$. Since $b_n \notin \gamma(Q)$ we have $a_n \notin \gamma(Q)$, thus $a_n \neq Q$. This contradicts the previous lemma. It follows that $\gamma(Q) = \gamma(a) = \omega(x_0)$ is a periodic orbit. $\diamond$

Completion of the Proof of Theorem 9.2.

Let the assumptions of Theorem 9.2 hold and assume that

$$\gamma^+(x_0) \cap \omega(x_0) = \emptyset \ .$$

We must show that $\omega(x_0)$ is a periodic orbit.

Let $a \in \omega(x_0)$. We will prove that $\gamma(a)$ is periodic. (Then we have by Lemma 9.6 that $\gamma(a) = \omega(x_0)$ and are done.)

We know that $\omega(x_0)$ is a closed and bounded set by Theorem 9.1. Since $\gamma^+(a) \subset \omega(x_0)$ there exists a point $P \in \omega(x_0)$ and a sequence of times $t_n \to \infty$ with $\phi(t_n, a) \to P$ as $n \to \infty$. Since $P \in \omega(x_0)$ we have $f(P) \neq 0$ and consider the line piece

$$L_\Delta = L \cap B_\Delta(P)$$

through P. For large n there are times $t_n^* \to \infty$ so that

$$a_n := \phi(t_n^*, a) \to P, \quad a_n \in L_\Delta \cap \gamma^+(a) \ .$$

See **Figure 9.9**.

Since $a \in \omega(x_0)$ we have $\gamma^+(a) \subset \omega(x_0)$ (by Lemma 9.3), thus $a_n \in L_\Delta \cap \omega(x_0)$. Also, $P \in L_\Delta \cap \omega(x_0)$. By Lemma 9.5 it follows that $a_n = a_{n+1} = P$, thus $\gamma(a)$ is periodic. If $a \in \omega(x_0)$ then an evolution sketched in the phase diagram in Figure 9.9 cannot occur.

Summary: Under the assumptions of Theorem 9.2 there are only two cases:
Case 1: $\gamma^+(x_0) = \gamma(x_0) = \omega(x_0)$ is a periodic orbit.
Case 2: $\gamma^+(x_0) \cap \omega(x_0) = \emptyset$ and $\omega(x_0)$ is a periodic orbit.
In Case 2 we have

$$\omega(x_0) \subset \Big(cl(\gamma^+(x_0)) \setminus \gamma^+(x_0) \Big) \ .$$

Also, if $Q \in cl(\gamma^+(x_0))$ then there exists $t_n \geq 0$ so that $\phi(t_n, x_0) \to Q$ as $n \to \infty$. If the sequence t_n is bounded, then $Q \in \gamma^+(x_0)$. Thus, if

$$Q \in \Big(cl(\gamma^+(x_0)) \setminus \gamma^+(x_0) \Big)$$

and $\phi(t_n, x_0) \to Q$ then t_n is unbounded. It follows that there is a subsequence t_k^* of t_n so that $t_k^* \to \infty$ and $\phi(t_k^*, x_0) \to Q$. Therefore, $Q \in \omega(x_0)$. This proves that

$$\omega(x_0) = cl(\gamma^+(x_0)) \setminus \gamma^+(x_0)$$

holds in Case 2.

Theorem 9.2 yields the following result:

Theorem 9.4 *Let $f : \mathbb{R}^2 \to \mathbb{R}^2, f \in C^1$. Assume that $\Omega \subset \mathbb{R}^2$ is a closed bounded set and $f(a) \neq 0$ for all $a \in \Omega$. Also, assume that $x_0 \in \Omega$ implies that $\gamma^+(x_0) \subset \Omega$. Then, for every $x_0 \in \Omega$, the positive semi–orbit $\gamma^+(x_0)$ either is a periodic orbit or spirals towards the periodic orbit $\omega(x_0)$ as $t \to \infty$.*

See **Figure 9.10**.

Remarks: In Theorem 9.2 we made the assumption that $|Df(x)|$ is bounded on $\mathbb{R}^2$. In Theorem 9.4 we only assumed that $f \in C^1$. Since Ω is bounded there exists $R > 0$ with $\Omega \subset B_R(0)$. Choose a C^∞–function $\chi : [0, \infty) \to [0, 1]$ with $\chi(r) = 1$ for $0 \leq r \leq R$ and $\chi(r) = 0$ for $r \geq R + 1$. Set

$$\tilde{f}(x) = \chi(|x|) f(x) \quad \text{for} \quad x \in \mathbb{R}^2 \ .$$

Clearly, $|D\tilde{f}(x)|$ is bounded on $\mathbb{R}^2$ and Theorem 9.2 applies to the system $x' = \tilde{f}(x)$. Since $\tilde{f}(x) = f(x)$ for $x \in \Omega$ the dynamics of the two systems $x' = \tilde{f}(x)$ and $x' = f(x)$ agrees on Ω. Therefore, Theorem 9.3 follows from Theorem 9.2.

Example: Consider the system

$$r' = r(1 - r^2) + \lambda r \cos\theta, \quad \theta' = 1 \ ,$$

in polar coordinates. For $\lambda = 0$ the solution

$$r(t) = 1, \quad \theta(t) = t \bmod 2\pi$$

is a periodic solution.

Exercise: Show that the system has a periodic solution for $|\lambda| < 1$.
Hint: Consider a region of the form

$$\Omega = [r_{min}, r_{max}] \times (\mathbb{R} \bmod 2\pi)$$

and apply Theorem 9.4.

9.3 Two Uncoupled Oscillators

9.3.1 Density of Every Orbit

Consider the system

$$\theta_1' = \omega_1, \quad \theta_2' = \omega_2$$

where $\omega_j > 0$ are constant frequencies and $\theta_j(t) \in S^1 = \mathbb{R} \mod 1$. Given initial conditions

$$\theta_j(0) = \alpha_j \quad \text{for} \quad j = 1, 2$$

the solutions is

$$\theta(t) = \begin{pmatrix} \theta_1(t) \\ \theta_2(t) \end{pmatrix} = \begin{pmatrix} (\alpha_1 + \omega_1 t) \mod 1 \\ (\alpha_2 + \omega_2 t) \mod 1 \end{pmatrix}$$

Theorem 9.5 *If $c = \omega_2/\omega_1$ is irrational then every orbit is dense in the torus $S^1 \times S^1$.*

Proof: We first assume the initial conditions

$$\theta_1(0) = 0, \quad \theta_2(0) = \beta \in [0, 1) .$$

Let $T = 1/\omega_1$. We have

$$\theta_1(T) = (0 + \omega_1 T) \mod 1 = 0$$

and

$$\theta_2(T) = (\beta + \omega_2 T) \mod 1 = (\beta + c) \mod 1 .$$

We define the Poincar'e map $P : S^1 \to S^1$ by

$$P\beta = (\beta + c) \mod 1$$

and obtain that

$$P^n\beta = (\beta + nc) \mod 1 = \theta_2(nT) \quad \text{for} \quad n \in \mathbb{N} .$$

We claim that the sequence $P^n\beta$ is dense in S^1.

Note that if $P^n\beta = P^m\beta$ then

$$\beta + nc = \beta + mc + k$$

for some integer k. Since c is irrational one obtains that $m = n$. It follows that the points $P^n\beta, n \in \mathbb{N}$, are all distinct.

Define the distance function on S^1:

$$dist(r,s) = \min_{k=0,-1,1} |r - s + k| \ .$$

We have

$$dist(Pr, Ps) = dist(r,s) \ .$$

The sequence $P^n\beta$ has an accumulation point $Q \in S^1$. Given $\varepsilon > 0$ there exist $m > 0, k > 0$ so that

$$dist(P^k\beta, Q) < \varepsilon/2, \quad dist(P^{k+m}\beta, Q) < \varepsilon/2 \ .$$

It follows that

$$dist(P^k\beta, P^{m+k}\beta) < \varepsilon \ ,$$

thus

$$dist(\beta, P^m\beta) < \varepsilon \ .$$

Summary: Given $\varepsilon > 0$ there exists $m = m(\varepsilon) \in \mathbb{N}$ with

$$dist(\beta, P^m\beta) < \varepsilon \ .$$

We note that the map P^m maps the arc from β to $P^m\beta$ onto the arc from $P^m\beta$ to $P^{2m}\beta$. This can be continued by applying P^m repeatedly. One obtains that for every $Q \in S^1$ there exists $l \in \mathbb{N}$ with

$$dist(P^{lm}\beta, Q) < \varepsilon \quad \text{where} \quad m = m(\varepsilon) \ .$$

This proves the claim that the sequence $P^n\beta$ is dense in S^1.

We now show that every orbit $\theta(t)$ is dense in $S^1 \times S^1$. Let $Q = (q_1, q_2) \in S^1 \times S^1$ and let $\varepsilon > 0$ be given. We will show that there exists a time $t \geq 0$ so that

$$\theta_1(t) = q_1, \quad dist(\theta_2(t), q_2) < \varepsilon \ .$$

Since $\theta_1(t) = (\alpha_1 + \omega_1 t) \mod 1$ there exists a time $t_1 \geq 0$ with $\theta_1(t_1) = q_1$. Set $T = 1/\omega_1$ and note that

$$\theta_1(t_1 + nT) = \theta_1(t_1) = q_1 \ .$$

Now consider the sequence

$$\begin{aligned} \theta_2(t_1 + nT) &= \theta_2(\alpha_2 + t_1 + nc) \\ &= (\beta + nc) \mod 1 \end{aligned}$$

with $\beta = (\alpha_2 + t_1) \mod 1$.

The theorem now follows since the sequence $P^n\beta = (\beta + nc) \mod 1$ is dense in S^1. $\diamond$

9.3.2 Time–Average Equals Space–Average for the Poincaré Map

Theorem 9.6 *Let $c \in \mathbb{R} \setminus \mathbb{Q}$. Let $P : S^1 \to S^1$ denote the map*

$$\beta \to P\beta = (\beta + c) \mod 1 \; .$$

Let $I \subset [0,1)$ denote an interval of length l. For $N \in \mathbb{N}$ let

$$M(I,N) = \Big\{ n \in \{0,1,\ldots,N\} \;:\; P^n\beta \in I \Big\} \; .$$

Then

$$\lim_{N\to\infty} \frac{1}{N+1} \#M(I,N) = l = length(I) \; . \tag{9.7}$$

Proof: If $f : \mathbb{R} \to \mathbb{C}$ is any function let

$$A_N f = \frac{1}{N+1} \sum_{j=0}^{N} f(P^j\beta)$$

denote the average value of f on the orbit part $P^j\beta, j = 0,1,\ldots,N$.

With χ_I we denote the characteristic function of I,

$$\chi_I(x) = 1 \quad \text{for} \quad x \in I, \quad \chi_I(x) = 0 \quad \text{for} \quad x \notin I \; .$$

Since

$$\#M(I,N) = \sum_{j=0}^{N} \chi_I(P^j\beta)$$

we must prove that

$$A_N \chi_I \to l \quad \text{as} \quad N \to \infty \; .$$

We will use Fourier analysis. For $k \in \mathbb{Z}$ let

$$f_k(x) = e^{2\pi i k x} \; .$$

Note the orthogonality

$$(f_j, f_k)_{L_2} = \int_0^1 \overline{f_j(x)} f_k(x)\, dx = \int_0^1 e^{2\pi i (k-j)x}\, dx = \delta_{jk} \quad \text{for all} \quad j,k \in \mathbb{Z} \; .$$

The Fourier series representation of $\chi(x)$ is

$$\chi_I(x) = \sum_{k=-\infty}^{\infty} \alpha_k f_k(x)$$

with

$$\alpha_k = \int_0^1 \chi_I(x) e^{-2\pi i k x}\, dx \; .$$

In particular, for $k = 0$ obtain that

$$\alpha_0 = \int_0^1 \chi_I(x)\, dx = l \ .$$

Let us compute $A_N f_k$. For $k = 0$ we have $f_0 \equiv 1$, thus

$$A_N f_0 = 1 \quad \text{for all} \quad N \in \mathbb{N} \ .$$

Now let $k \in \mathbb{Z}, k \neq 0$. We have

$$f_k(P^j \beta) = e^{2\pi i k(\beta + jc)} = e^{2\pi i k \beta} q^j \quad \text{with} \quad q = e^{2\pi i k c} \ .$$

Note that $|q| = 1$ and $q \neq 1$ since c is irrational. Using the geometric sum formula obtain that

$$\sum_{j=0}^{N} f_k(P^j \beta) = e^{2\pi i k \beta} \, \frac{1 - q^{N+1}}{1 - q} \ .$$

It follows that

$$\Big| \sum_{j=0}^{N} f_k(P^j \beta) \Big| \leq \frac{2}{|1 - q|} \ .$$

Here

$$q = q(k) = e^{2\pi i k c}$$

depends on k, but is independent of N. It follows that

$$A_N f_k \to 0 \quad \text{as} \quad N \to \infty \quad \text{for} \quad k \neq 0 \ .$$

Together with the result for $k = 0$ obtain that

$$\lim_{N \to \infty} \Big(\frac{1}{N+1} \sum_{j=0}^{N} f_k(P^j \beta) \Big) = \delta_{k0}, \quad k \in \mathbb{Z} \ . \tag{9.8}$$

In the following equations we exchange two limit processes. This needs to be justified, however.

$$
\begin{aligned}
\lim_{N\to\infty} \frac{1}{N+1}\,\#M(I,N) &= \lim_{N\to\infty} \Big(\frac{1}{N+1} \sum_{j=0}^{N} \chi_I(P^j\beta) \Big) \\
&= \lim_{N\to\infty} \Big(\frac{1}{N+1} \sum_{j=0}^{N} \sum_{k=-\infty}^{\infty} \alpha_k f_k(P^j\beta) \Big) \\
&= \sum_{k=-\infty}^{\infty} \alpha_k \Big(\lim_{N\to\infty} \frac{1}{N+1} \sum_{j=0}^{N} f_k(P^j\beta) \Big) \\
&= \sum_{k=-\infty}^{\infty} \alpha_k \delta_{k0} \\
&= \alpha_0 \\
&= l
\end{aligned}
$$

We now give a rigorous proof. It will be based on the **Stone–Weierstrass Theorem** stated below.

A function $p : \mathbb{R} \to \mathbb{C}$ is a trigonometric polynomial (of period 1) if $p(x)$ has the form

$$
p(x) = \sum_{k=-K}^{K} \alpha_k e^{2\pi i k x}
$$

where K is finite and $\alpha_k \in \mathbb{C}$. Let $\mathbb{T}$ denote the vector space of all trigonometric polynomials.

Theorem 9.7 *(Stone–Weierstrass) Let $f : \mathbb{R} \to \mathbb{C}$ denote a continuous function of period 1. Then, for all $\varepsilon > 0$, there exists $p_\varepsilon \in \mathbb{T}$ so that*

$$
|f - p_\varepsilon|_\infty \le \varepsilon \ .
$$

Recall the definition of the average,

$$
A_N f = \frac{1}{N+1} \sum_{j=0}^{N} f(P^j\beta)
$$

where $P^j\beta = (\beta + jc) \mod 1$, c irrational.

Lemma 9.7 *Let $f : \mathbb{R} \to \mathbb{C}$ denote a continuous function of period 1. Then*

$$
A_N f \to \int_0^1 f(x)\,dx \quad as \quad N \to \infty \ .
$$

Proof: Let $\varepsilon > 0$ be given and let $p_\varepsilon \in \mathbb{T}$ satisfy $|f - p_\varepsilon|_\infty \le \varepsilon$. We have

$$\begin{aligned}\Big|A_N f - \int_0^1 f dx\Big| &\leq \Big|A_N f - A_N p_\varepsilon\Big| + \Big|A_N p_\varepsilon - \int_0^1 p_\varepsilon dx\Big| + \Big|\int_0^1 p_\varepsilon dx - \int_0^1 f dx\Big| \\ &\leq 2\varepsilon + \Big|A_N p_\varepsilon - \int_0^1 p_\varepsilon dx\Big|\end{aligned}$$

Because of (9.8) we have

$$A_N e^{2\pi ikx} \to \delta_{k0} \quad \text{as} \quad N \to \infty$$

and obtain that

$$A_N p_\varepsilon \to \int_0^1 p_\varepsilon(x)\, dx \quad \text{as} \quad N \to \infty \ .$$

Therefore, if $N \geq N(\varepsilon)$, then

$$|A_N f - \int_0^1 f\, dx| \leq 3\varepsilon \ .$$

This proves the lemma. $\diamond$

For $\varepsilon > 0$ there are continuous, 1–periodic functions $g_\varepsilon(x)$ and $f_\varepsilon(x)$ with

$$g_\varepsilon(x) \leq \chi_I(x) \leq f_\varepsilon(x) \quad \text{for} \quad x \in S^1$$

and

$$\int_0^1 g_\varepsilon\, dx = l - \varepsilon \quad \text{and} \quad \int_0^1 f_\varepsilon\, dx = l + \varepsilon$$

where

$$l = \int_0^1 \chi_I(x)\, dx \ .$$

We then have for all $N \in \mathbb{N}$:

$$A_N g_\varepsilon \leq A_N \chi_I \leq A_N f_\varepsilon \ .$$

If $N \geq N(\varepsilon)$ then, using the previous lemma, we obtain that

$$l - 2\varepsilon \leq A_N g_\varepsilon \leq A_N \chi_I \leq A_N f_\varepsilon \leq l + 2\varepsilon \ .$$

This proves that

$$|A_N \chi_I - l| \leq 2\varepsilon \ .$$

This completes the proof of Theorem 9.6. $\diamond$

9.3.3 Time–Average Equals Space–Average for the Uncoupled Oscillators

Recall that

$$\Theta(t) = \Big((\omega_1 t) \mod 1, (\beta + \omega_2 t) \mod 1\Big)$$

where ω_2/ω_1 is irrational. Let R denote a rectangle in the torus $S^1 \times S^1$, i.e., $R \subset [0,1) \times [0,1)$. With $\chi_R(\theta)$ we denote the characteristic function of R.

We claim that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \chi_R(\theta(t))\, dt = area(R) \; . \tag{9.9}$$

For $k = (k_1, k_2) \in \mathbb{Z}^2$ let

$$f_k(x,y) = e^{2\pi i k_1 x} e^{2\pi i k_2 y} \; .$$

Obtain that

$$f_k(\theta(t)) = e^{2\pi i k_2 \beta} e^{2\pi i (k_1\omega_1 + k_2\omega_2)t} \; .$$

Note that for $k = (0,0)$ we have $f_k = f_0 \equiv 1$, thus

$$\frac{1}{T} \int_0^T f_k(\theta(t))\, dt = 1 \quad \text{for} \quad k = (0,0) \; .$$

Now assume that $k \in \mathbb{Z}^2, k \neq (0,0)$ and note that

$$k_1\omega_1 + k_2\omega_2 \neq 0$$

since ω_2/ω_1 is irrational.

One obtain that

$$\int_0^T e^{2\pi i (k_1\omega_1 + k_2\omega_2)t}\, dt = \frac{1}{2\pi i (k_1\omega_1 + k_2\omega_2)} \Big(e^{2\pi i (k_1\omega_1 + k_2\omega_2)T} - 1 \Big) \; .$$

The absolute value of the integral has a bound that does not depend on $T > 0$. Therefore,

$$\frac{1}{T} \int_0^T f_k(\theta(t))\, dt \to 0 \quad \text{as} \quad T \to \infty \quad \text{for} \quad k \neq (0,0) \; .$$

Since

$$\frac{1}{T} \int_0^T f_k(\theta(t))\, dt \to \delta_{k_1 0}\delta_{k_2 0} \quad \text{as} \quad T \to \infty$$

the proof of (9.9) can be completed in a similar way as the proof of Theorem 9.6.

Figures for Chapter 9

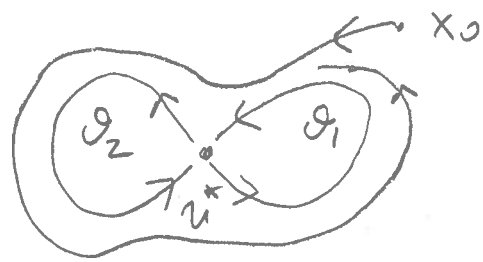

Figure 9.1

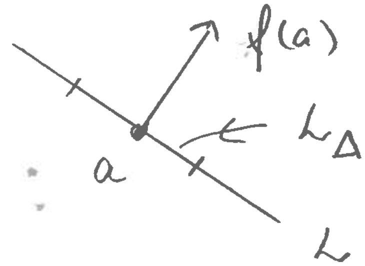

Figure 9.2

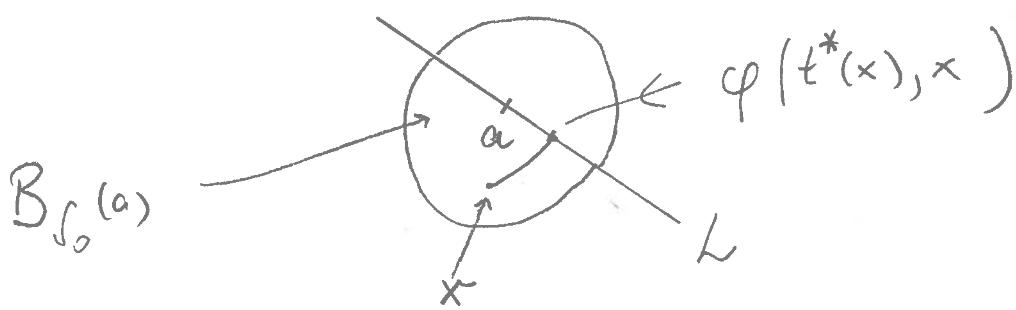

Figure 9.3

Figures for Chapter 9

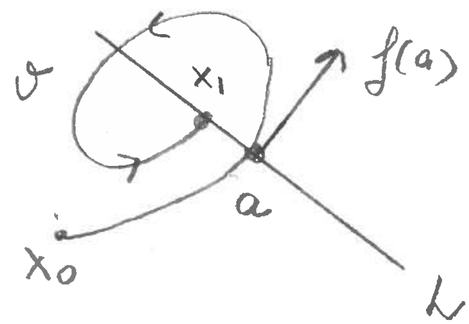

Figure 9.4

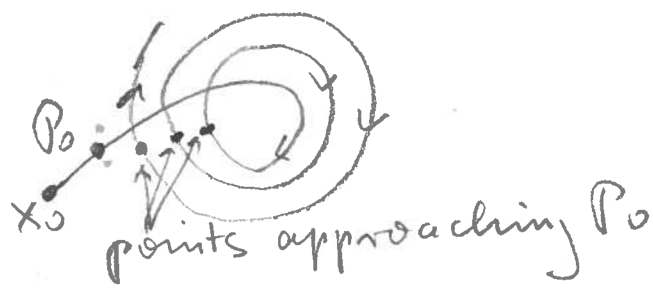

Figure 9.5

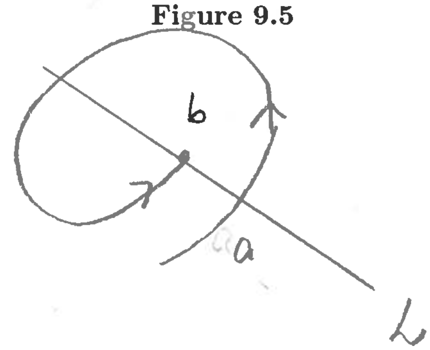

Figure 9.6

Figures for Chapter 9

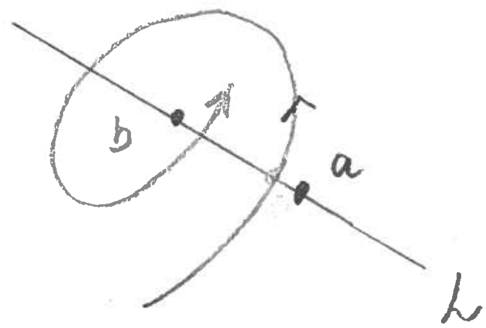

Figure 9.7

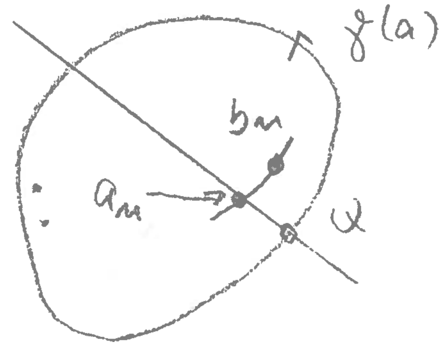

Figure 9.8

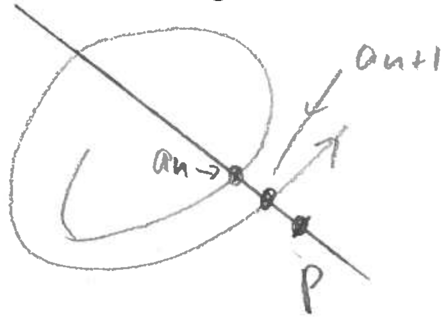

Figure 9.9

Figures for Chapter 9

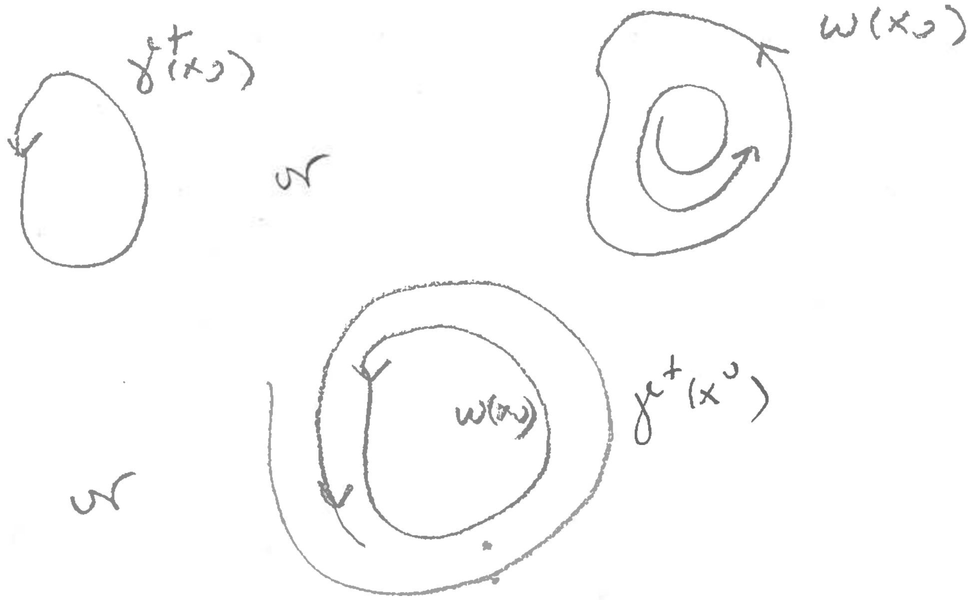

Figure 9.10

10 Hamiltonian Systems; Stability via Liapunov

10.1 Notations

We use the notation

$$q = (q_1, \ldots, q_N)^T$$

for the generalized position coordinates and

$$p = (p_1, \ldots, p_N)^T$$

for the generalized momenta.

If $H(q,p)$ denotes the Hamiltonian function of a system, then the dynamical equations

$$\begin{aligned} q' &= H_p(q,p) \\ p' &= -H_q(q,p) \end{aligned}$$

describe the evolution of the system.

Let

$$x = \begin{pmatrix} q \\ p \end{pmatrix}, \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} .$$

With these notations, the Hamiltonian system can be written as

$$x' = JH'(x) \tag{10.1}$$

where

$$H'(x) = (\nabla H(x))^T = \Big(\frac{\partial H}{\partial q_1}, \ldots, \frac{\partial H}{\partial q_N}, \frac{\partial H}{\partial p_1}, \ldots, \frac{\partial H}{\partial p_N}\Big)^T (x) .$$

10.2 An Example of a Hamiltonian System

Consider Newton's equation

$$mq''(t) = F(q)$$

where $q(t) \in \mathbb{R}$ is the position of the point mass m at time t and $F(q)$ is a scalar force acting on the point mass.

The function

$$U(q) = -\int_0^q F(\xi)\, d\xi$$

is the potential energy corresponding to the force $F(q)$. The kinetic energy is

$$\frac{m}{2}\,(q')^2 = \frac{1}{2m}\,p^2$$

where $p = mq'$ denotes the momentum. The Hamiltonian is

$$H(q,p) = \frac{1}{2m} p^2 + U(q) \; .$$

The Hamiltonian system is

$$\begin{aligned} q' = & \quad H_p & = \frac{1}{m} p \\ p' = & \; -H_q & = -dU/dq \end{aligned}$$

One obtains that

$$mq'' = p' = -dU/dq = F(q) \; ,$$

which agrees with Newton's equation.

If one assumes Hooke's law,

$$F(q) = -kq \; ,$$

then $U(q) = \frac{k}{2} q^2$ and

$$H(q,p) = \frac{1}{2m} p^2 + \frac{k}{2} q^2 \; .$$

10.3 Energy Conservation

Consider any Hamiltonian system $x' = JH'(x)$. Since $J \in \mathbb{R}^{2N \times 2N}$ is skew symmetric, we have

$$\langle a, Ja \rangle = 0 \quad \text{for all} \quad a \in \mathbb{R}^{2N} \; .$$

Let $x(t)$ denote a solution of the equation $x' = JH'(x)$. Then we have

$$\begin{aligned} \frac{d}{dt} H(x(t)) &= \nabla H(x)\, x' \\ &= \nabla H(x)\, J (\nabla H(x))^T \\ &= \langle H'(x), JH'(x) \rangle \\ &= 0 \end{aligned}$$

Energy is constant along any orbit.

10.4 The Pendulum Equation as Hamiltonian

Newton's equation reads:

$$ml\alpha'' = -mg \sin \alpha$$

or

$$\alpha'' + \frac{g}{l} \sin \alpha = 0 \; .$$

Here

$$g = 9.81 \, \frac{meter}{sec^2}$$

is the acceleration die to gravity. Using

$$\begin{aligned} q &= l\alpha \\ p &= ml\alpha' \end{aligned}$$

we have

$$E_{kin} = \frac{1}{2m} \, p^2, \quad E_{pot} = mgl(1 - \cos\alpha) \ .$$

The Hamiltonian is

$$H(q,p) = \frac{1}{2m} \, p^2 + mgl(1 - \cos(q/l)) \ .$$

The Hamiltonian equations become

$$q' = p/m, \quad p' = -mg\sin(q/l) \ .$$

One derives

$$q'' + g\sin(q/l) = 0$$

which is equivalent to Newton's equation.

10.5 Celestial Mechanics: The Two Body Problem as Hamiltonian

Let $\mathbf{r}_1 \in \mathbb{R}^3$ and $\mathbf{r}_2 \in \mathbb{R}^3$ denote the position vectors of two bodies of mass m_1 and m_2. Newton's equations are:

$$\begin{aligned} m_1 \mathbf{r}_1'' &= G \, \frac{m_1 m_2}{|\mathbf{r}_2 - \mathbf{r}_1|^3} \, (\mathbf{r}_2 - \mathbf{r}_1) =: F \\ m_2 \mathbf{r}_2'' &= G \, \frac{m_1 m_2}{|\mathbf{r}_2 - \mathbf{r}_1|^3} \, (\mathbf{r}_1 - \mathbf{r}_2) = -F \end{aligned}$$

Here

$$G = 6.67 \cdot 10^{-11} \, \frac{\text{Newton meter}^2}{\text{kg}^2}$$

is the gravitational constant. We want to obtain a Hamiltonian system equivalent to Newton's equations.

Let $\mathbf{p}_j = m_j \mathbf{r}_j'$ denote the momenta. The kinetic energy is

$$\begin{aligned} E_{kin}(\mathbf{p_1}, \mathbf{p_2}) &= \frac{1}{2}\Big(m_1 |\mathbf{r}_1'|^2 + m_2 |\mathbf{r}_2'|^2 \Big) \\ &= \frac{1}{2}\Big(\frac{1}{m_1} |\mathbf{p}_1|^2 + \frac{1}{m_2} |\mathbf{p}_2|^2 \Big) \end{aligned}$$

We have

$$\nabla_{\mathbf{p}_1} E_{kin} = \frac{1}{m_1}\,\mathbf{p}_1^T, \quad \nabla_{\mathbf{p}_2} E_{kin} = \frac{1}{m_2}\,\mathbf{p}_2^T \ . \tag{10.2}$$

The potential energy is

$$E_{pot}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{G m_1 m_2}{|\mathbf{r}_1 - \mathbf{r}_2|} \ .$$

Note: If $\mathbf{r} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$ then

$$\begin{aligned} \frac{\partial}{\partial x_j}\frac{1}{|\mathbf{r}|} &= \frac{\partial}{\partial x_j}\left(x_1^2 + x_2^2 + x_3^2\right)^{-1/2} \\ &= -\frac{x_j}{|\mathbf{r}|^3} \end{aligned}$$

thus

$$\nabla \frac{1}{|\mathbf{r}|} = -|\mathbf{r}|^{-3}\,\mathbf{r}^T \ .$$

Let

$$\mathbf{r}_m = \left(x_1^{(m)}, x_2^{(m)}, x_3^{(m)}\right)^T \quad \text{for} \quad m = 1, 2 \ .$$

Then we have for $j = 1, 2, 3$:

$$\frac{\partial}{\partial x_j^{(1)}}\,\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = -\frac{x_j^{(1)} - x_j^{(2)}}{|\mathbf{r}_1 - \mathbf{r}_2|^3}$$

and

$$\frac{\partial}{\partial x_j^{(2)}}\,\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = -\frac{x_j^{(2)} - x_j^{(1)}}{|\mathbf{r}_1 - \mathbf{r}_2|^3} \ .$$

Therefore,

$$\nabla_{\mathbf{r}_1} E_{pot} = G m_1 m_2 |\mathbf{r}_1 - \mathbf{r}_2|^{-3}\,(\mathbf{r}_1 - \mathbf{r}_2)^T = -F \tag{10.3}$$

and

$$\nabla_{\mathbf{r}_2} E_{pot} = G m_1 m_2 |\mathbf{r}_1 - \mathbf{r}_2|^{-3}\,(\mathbf{r}_2 - \mathbf{r}_1)^T = F \ . \tag{10.4}$$

Let

$$q = \left(x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_1^{(2)}, x_2^{(2)}, x_3^{(2)}\right)^T$$

denote the coordinate vector and let

$$p = \left(p_1^{(1)}, p_2^{(1)}, p_3^{(1)}, p_1^{(2)}, p_2^{(2)}, p_3^{(2)}\right)^T$$

denote the momenta vector. The Hamiltonian function is

$$H(q,p) = E_{kin}(p) + E_{pot}(q) \ .$$

Because of (10.2) the Hamiltonian equation

$$q' = H_p$$

requires that

$$\mathbf{r}_1' = \frac{1}{m_1}\,\mathbf{p_1} \quad \text{and} \quad \mathbf{r}_2' = \frac{1}{m_2}\,\mathbf{p_2} \ .$$

Because of (10.3) and (10.4) the Hamiltonian equation

$$p' = -H_q$$

requires that

$$\mathbf{p_1}' = F \quad \text{and} \quad \mathbf{p_2}' = -F \ .$$

Clearly, the Hamiltonian system

$$q' = H_p, \quad p' = -H_q$$

is equivalent to Newton's equations.

10.6 The Jacobian of a Hamiltonian Function

Consider the Hamiltonian equation

$$x' = f(x) \quad \text{with} \quad f(x) = JH'(x), \quad H'(x) = (\nabla H(x))^T \ .$$

The Jacobian of $f(x)$ is

$$f'(x) = JH''(x)$$

where

$$H''(x) = \Big(D_i D_j H(x)\Big)_{i,j=1\ldots,2N}$$

is the Hessian of H. Since $H''(x)$ is symmetric we have

$$f'(x) = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

with symmetric $N \times N$ matrices A and C. Therefore,

$$f'(x) = \begin{pmatrix} B^T & C \\ -A & -B \end{pmatrix} \ .$$

In particular, it follows that

$$tr\,(f'(x)) = 0 \quad \text{for all} \quad x = \begin{pmatrix} q \\ p \end{pmatrix} \ .$$

We will see that this implies conservation of phase volume.

10.7 Stability of Fixed Points

Let x^* denote a fixed point of the Hamiltonian system $x' = JH'(x) =: f(x)$. We then have

$$\nabla H(x^*) = 0 \ .$$

Let $\lambda_1, \ldots, \lambda_{2N}$ denote the eigenvalues of the Jacobian $f'(x^*)$ listed according to their algebraic multiplicity.

We then have

$$\sum \lambda_j = tr\,(f'(x^*)) = 0 \ .$$

Therefore, if there exists λ_j with $\mathrm{Re}\,\lambda_j < 0$ there also exists λ_k with $\mathrm{Re}\,\lambda_k > 0$.

Lemma 10.1 *If x^* is a stable equilibrium of the Hamiltonian system $x' = JH'(x)$ then*

$$\mathrm{Re}\,\lambda_j = 0 \quad \textit{for all} \quad \lambda_j \in \sigma(f'(x^*)) \ .$$

10.8 Evolution of Phase Volume

We first recall the general transformation formula for integrals.

Theorem 10.1 *Let $A, B \subset \mathbb{R}^N$ denote open, bounded sets and let $S : A \to B$ be a C^1 bijection with $det\, S'(x) > 0$ for all $x \in A$. Let $h \in L_1(B)$. Then we have*

$$\int_B h(y)\,dy = \int_A h(S(x))\,det\,S'(x)\,dx \ .$$

In the theorem, the mapping $S : A \to B$ can be thought of as a coordinate transformation and $S'(x) \in \mathbb{R}^{N\times N}$ denotes its Jacobian.

In the special case where $h \equiv 1$ we obtain

$$vol(B) = \int_A det\,S'(x)\,dx \ . \tag{10.5}$$

Let $f : \mathbb{R}^N \to \mathbb{R}^N, f \in C^1, |f_x(x)| \le C$. Consider the initial value problem

$$x' = f(x), \quad x(0) = x_0 \ ,$$

with solution $x(t) = \phi(t, x_0)$.

Let $V_0 \subset \mathbb{R}^N$ denote an open bounded set and define

$$V(t) = \{\phi(t,x) \ : \ x \in V_0\} \ .$$

Intuitively, the mapping

$$S : \left\{ \begin{array}{ccc} V_0 & \to & V(t) \\ x & \to & \phi(t,x) \end{array} \right.$$

describes how the volume V_0 evolves under the flow determined by the map $x \to \phi(t,x)$ for fixed t.

We apply formula (10.5) and obtain

$$
\begin{aligned}
vol(V(t)) &= \int_{V_0} det\, S'(x)\, dx \\
&= \int_{V_0} det\, \phi_x(t,x)\, dx
\end{aligned}
$$

Recall the following formulas:

$$
\begin{aligned}
\phi_t(t,x) &= f(\phi(t,x)), \quad \phi(0,x) = x \\
\phi_{xt}(t,x) &= f'(\phi(t,x))\phi_x(t,x), \quad \phi_x(0,x) = I \\
\frac{d}{dt}\, det\, \phi_x(t,x) &= tr f'(\phi(t,x)) \cdot det\, \phi_x(t,x)
\end{aligned}
$$

The last equation follows from Liouville's Theorem, Theorem 3.1: If $\Phi'(t) = A(t)\Phi(t)$ then

$$\frac{d}{dt}\, det\, \Phi(t) = tr\, A(t) \cdot det\, \Phi(t) \ .$$

Therefore,

$$det\, \phi_x(t,x) = \exp\Big(\int_0^t tr f'(\phi(\tau,x))\, d\tau \Big) \ .$$

Lemma 10.2 *Assume that $tr f'(x) = 0$ for all $x \in \mathbb{R}^N$. Then phase volume is preserved:*

$$vol(V(t)) = vol(V(0)), \quad t \in \mathbb{R} \ .$$

In the Hamiltonian case, the condition $tr f'(x) = 0$ is fulfilled as was noted in Section 10.6. Therefore, the flow of a Hamiltonian system preserves phase volume.

The reason for $tr\, f'(x) = 0$ in the Hamiltonian case is simple:

We have

$$x = \begin{pmatrix} q \\ p \end{pmatrix}, \quad f(x) = \begin{pmatrix} H_p(q,p) \\ -H_q(q,p) \end{pmatrix},$$

thus

$$f'(x) = \begin{pmatrix} H_{pq}(q,p) & H_{pp}(q,p) \\ -H_{qq}(q,p) & -H_{qp}(q,p) \end{pmatrix} \ .$$

Since $H_{pq} = H_{qp}$ it follows that $tr\, f'(x) = 0$.

The following result is called **Poincaré's Recurrence Theorem**.

Theorem 10.2 *Let $\Omega \subset \mathbb{R}^N$ denote a bounded open set and let the mapping $F : \Omega \to \Omega$ be bijective and volume preserving. Consider a set $U = B_\varepsilon(y) \subset \Omega$. There exists $x \in U$ and $n \in \mathbb{N}$ so that $F^n(x) \in U$. In words: Think of applying the map F as evolving by one time step. Given any $y \in \Omega$, there exists a point x arbitrarily close to y whose future state $F^n(x)$ will return arbitrarily close to y in finite time.*

Proof: The sets

$$F^n(U), \quad n = 1, 2, 3 \ldots$$

cannot be disjoint. There exists $j > k \geq 1$ so that

$$F^j(U) \cap F^k(U) \neq \emptyset \; .$$

Apply F^{-k} to obtain that

$$F^n(U) \cap U \neq \emptyset \quad \text{for} \quad n = j - k \geq 1 \; .$$

Let $\xi \in F^n(U) \cap U$. There exists $x \in U$ with $F^n(x) = \xi$. Therefore, $x \in U$ and $F^n(x) \in U$. $\diamond$

Note: Let $x(t) = \phi(t, x_0)$ denote the solution of an initial value problem $x' = f(x), x(0) = x_0$. Let $T > 0$ denote some fixed time. If $tr f'(x) = 0$ for all x, then the map F defined by $F(x_0) = \phi(T, x_0)$ is volume preserving and bijective. Also, $F^n(x) = \phi(nT, x)$. Poincaré's Recurrence Theorem applies if one can find a bounded set Ω which is invariant under F.

10.9 Symplectic Matrices

Symplectic matrices and canonical transformation play an important role in the theoretical analysis of Hamiltonian systems.

A matrix $S \in \mathbb{R}^{2N \times 2N}$ is called symplectic if

$$S^T J S = J \quad \text{where} \quad J = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix} \; .$$

An equivalent condition is that S leaves the bilinear form

$$\langle x, Jy \rangle$$

invariant, i.e., S is symplectic if and only if

$$\langle Sx, JSy \rangle = \langle x, Jy \rangle \quad \text{for all} \quad x, y \in \mathbb{R}^{2N} \; .$$

The equivalence is clear since $\langle Sx, JSy \rangle = \langle x, S^T JSy \rangle$.

Lemma 10.3 *If $\phi(t, x)$ denotes the solution of a Hamiltonian system $x' = JH'(x)$ then each matrix $\phi_x(t, x)$ is symplectic.*

Proof: We have

$$\phi_t(t,x) = JH'(\phi(t,x)), \quad \phi(0,x) = x \ .$$

Therefore,

$$\phi_{xt}(t,x) = JH''(\phi(t,x))\phi_x(t,x), \quad \phi_x(0,x) = I \ .$$

Abbreviate

$$\Phi(t) = \phi_x(t,x), \quad A(t) = H''(\phi(t,x)) \quad \text{where} \quad x \in \mathbb{R}^{2N} \text{ is fixed} \ .$$

Then $A(t)$ is a symmetric matrix and

$$\Phi'(t) = JA(t)\Phi(t), \quad \Phi(0) = I \ .$$

Now let

$$M(t) = \Phi(t)^T J \Phi(t) \ .$$

Then we have $M(0) = J$ and

$$\begin{aligned} M' &= (\Phi')^T J \Phi + \Phi^T J \Phi' \\ &= \Phi^T A(-J) J \Phi + \Phi^T J^2 A \Phi \\ &= 0 \end{aligned}$$

In the last equation we have used that $J^2 = -I$. $\diamond$

10.10 Properties of Symplectic Matrices

The set of all symplectic matrices $S \in \mathbb{R}^{2N \times 2N}$ is denote by $Sp(N)$. Properties 1) and 2) of the following lemma say that the set of matrices $Sp(N)$ forms a group under matrix multiplication.

Lemma 10.4 *1) If $S_1, S_2 \in Sp(N)$ then $S_1 S_2 \in Sp(N)$.*
2) If $S \in Sp(N)$ then S is nonsingular and $S^{-1} \in Sp(N)$.
3) If $S \in Sp(N)$ then $S^T \in Sp(N)$.
4) If $S \in Sp(N)$ then S is similar to S^{-1}.

Proof: Property 1) holds since $S_1^T J S_1 = J$ and $S_2^T J S_2 = J$ yields that $(S_1 S_2)^T J S_1 S_2 = J$. To show 2) first note that $S^T J S = J$ implies that $|det\,(S)| = 1$. (In fact, one can show that $det\,(S) = 1$. See Lax [Linear Algebra].)

From $S^T J S = J$ one obtains that $J = (S^{-1})^T J S^{-1}$. This says that $S^{-1} \in Sp(N)$.

To show 3) take the inverse of $S^T J S = J$ and use that $J^{-1} = -J$. Obtain:

$$S^{-1} J (S^{-1})^T = J \ .$$

Therefore, $(S^{-1})^T$ is symplectic. By 2) it follows that S^T is symplectic.

To show 4), note that $S^T JS = J$ implies that $J^{-1}S^T J = S^{-1}$. Thus, $S^T \sim S^{-1}$. For any square matrix A we have $A \sim A^T$. (This is a nontrivial result of linear algebra; A and A^T of the same Jordan canonical form.) Therefore, $S \sim S^{-1}$. ⋄

Remark: The group $Sp(N)$ is a Lie–group. One defines a Lie–group as a group, which is also a manifold, and the operations of multiplication,

$$(S_1, S_2) \to S_1 S_2 \ ,$$

and taking an inverse,

$$S \to S^{-1}$$

are smooth.

10.11 Canonical Transformations of Hamiltonian Systems

We first consider a general coordinate transformation applied to a general system $x' = f(x)$ where $f : \mathbb{R}^n \to \mathbb{R}^n$.

Let $R : \mathbb{R}^n \to \mathbb{R}^n$ be 1-1 and onto and assume that $R \in C^1$ and $det\,(R'(y)) \neq 0$ for all $y \in \mathbb{R}^n$. Let $x = R(y)$. For $y = y(t)$ obtain

$$\begin{aligned} R'(y)y' &= f(R(y)) \\ y' &= (R'(y))^{-1} f(R(y)) \end{aligned}$$

Now consider a Hamiltonian system

$$x' = JH'(x), \quad x = \begin{pmatrix} q \\ p \end{pmatrix} \in \mathbb{R}^{2N}$$

and a coordinate transformation $x = R(y)$.

Let

$$K(y) = H(R(y))$$

denote the Hamiltonian function in y–coordinates. The Hamiltonian system determined by K is

$$y' = JK'(y) \tag{10.6}$$

Here

$$\begin{aligned} \nabla K(y) &= \nabla H(R(y)) R'(y) \\ K'(y) &= (R'(y))^T H'(R(y)) \end{aligned}$$

Thus the Hamiltonian system (10.6) reads

$$y' = J(R'(y))^T H'(R(y)) \tag{10.7}$$

Using the y–coordinates in the given Hamiltonian system $x' = JH'(x)$ one obtains

$$y' = (R'(y))^{-1} JH'(R(y)) \tag{10.8}$$

The two systems (10.7) and (10.8) are identical if and only if

$$(R'(y))^{-1} J = J(R'(y))^T \quad \text{for all} \quad y \in \mathbb{R}^{2N} \ .$$

This is equivalent to

$$J = R'(y) J (R'(y))^T \quad \text{for all} \quad y \in \mathbb{R}^{2N} \ .$$

Since the matrix $R'(y)$ is symplectic if and only if $(R'(y))^T$ is symplectic one obtains that the coordinate transformation $x = R(y)$ leads to the Hamiltonian system corresponding to the Hamiltonian function $K(y) = H(R(y))$ if and only if the Jacobian $R'(y)$ is symplectic for all $y \in \mathbb{R}^{2N}$.

Definition: A coordinate transformation $x = R(y)$ applied to the Hamiltonian system $x' = JH'(x)$ is called a canonical transformation if $R'(y)$ is symplectic for all $y \in \mathbb{R}^{2N}$.

Theorem 10.3 *If a canonical transformation $x = R(y)$ is applied to the Hamiltonian system $x' = JH'(x)$ then one obtains the Hamiltonian system $y' = JK'(y)$ with $K(y) = H(R(y))$.*

10.12 Hamiltonian Systems in Action–Angle Variables

Remark: The vector $J \in \mathbb{R}^N$ introduced below has nothing to do with the matrix $J \in \mathbb{R}^{2N \times 2N}$ used above.

Let $H(q, p)$ denote a Hamiltonian function depending on the coordinate vector $q \in \mathbb{R}^N$ and the vector of momenta, $p \in \mathbb{R}^N$. The corresponding Hamiltonian system is

$$q' = H_p, \quad p' = -H_q \ .$$

Assume we can make a canonical transformation to so–called action–angle variables

$$J \in \mathbb{R}^N, \quad \alpha \in \mathcal{T}^N = (\mathbb{R} \bmod 2\pi)^N$$

and obtain the transformed Hamiltonian function $\tilde{H}(\alpha, J)$. The transformed Hamiltonian system is

$$\alpha' = \tilde{H}_J, \quad J' = -\tilde{H}_\alpha \ .$$

The new variables α and J are called action–angle variables if $\tilde{H}_\alpha = 0$. If this is the case then

$$J' = 0, \quad \alpha' = \tilde{H}_J \ .$$

Given the initial conditions

$$J(0) = J_0, \quad \alpha(0) = \alpha_0$$

one obtains

$$J(t) \equiv J_0, \quad \alpha(t) = (\omega t + \alpha_0) \bmod 2\pi$$

where $\omega = \tilde{H}_J(J_0)$. The action variable $J(t) = J_0$ is constant in time and the angle variable $\alpha(t)$ evolves on an N–dimensional torus.

If one can introduce actions–angle variables for a Hamiltonian system, then the evolution of all trajectories takes place on a torus.

10.13 The Harmonic Oscillator in Action–Angle Variables

Recall the equation for a harmonic oscillator

$$q'' + \omega^2 q = 0, \quad \omega^2 = k/m \ .$$

Its Hamiltonian function is

$$H(q,p) = \frac{1}{2m} p^2 + \frac{k}{2} q^2$$

where $p = mq'$.

Introduce new variables J and α by

$$q = \sqrt{\frac{2}{m\omega} J} \sin\alpha, \quad p = \sqrt{2m\omega J} \cos\alpha \ .$$

We will check below that this is a canonical transformation.

The Hamiltonian function becomes

$$\begin{aligned} H(q,p) &= \frac{1}{2m} 2m\omega J \cos^2\alpha + \frac{k}{2} \frac{2}{m\omega} \sin^2\alpha \\ &= \omega J \cos^2\alpha + \frac{k}{m\omega} J \sin^2\alpha \quad (\text{use } k = m\omega^2) \\ &= \omega J \\ &=: \tilde{H}(\alpha, J) \end{aligned}$$

The corresponding Hamiltonian system is

$$\alpha' = \tilde{H}_J = \omega, \quad J' = -\tilde{H}_\alpha = 0$$

with solution

$$J(t) \equiv J_0, \quad \alpha(t) = (\omega t + \alpha_0) \bmod 2\pi \ .$$

For the harmonic oscillator equation one obtains the solution

$$q(t) = \sqrt{\frac{2}{m\omega} J_0} \, \sin(\omega t + \alpha_0) \ .$$

This is in agreement with the general solution of the harmonic oscillator equation:

$$q(t) = A \sin(\omega t + \alpha_0) \ .$$

Let us check that the coordinate transformation

$$(\alpha, J) \to R(\alpha, J) = (q, p)$$

is canonical. We have

$$R(\alpha, J) = \begin{pmatrix} \sqrt{\frac{2}{m\omega} J} \, \sin\alpha \\ \sqrt{2m\omega J} \, \cos\alpha \end{pmatrix} \ ,$$

thus

$$\begin{aligned} R'(\alpha, J) &= \begin{pmatrix} \sqrt{\frac{2}{m\omega} J} \, \cos\alpha & \frac{1}{2} \sqrt{\frac{2}{m\omega}} \frac{1}{\sqrt{J}} \, \sin\alpha \\ -\sqrt{2m\omega J} \, \sin\alpha & \frac{1}{2} \sqrt{2m\omega} \frac{1}{\sqrt{J}} \, \cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} a\cos\alpha & b\sin\alpha \\ -\frac{1}{b} \, \sin\alpha & \frac{1}{a} \, \cos\alpha \end{pmatrix} =: S \end{aligned}$$

With $c = \cos\alpha, s = \sin\alpha$ obtain:

$$\begin{aligned} S^T J S &= \begin{pmatrix} ac & -s/b \\ bs & c/a \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} ac & bs \\ -s/b & c/a \end{pmatrix} \\ &= \begin{pmatrix} s/b & ac \\ -c/a & bs \end{pmatrix} \begin{pmatrix} ac & bs \\ -s/b & c/a \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = J \end{aligned}$$

This shows that the transformation $(q, p) = R(\alpha, J)$ is canonical.

10.14 Stability via Liapunov

Example 1 Consider the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} -y^3 \\ x^3 \end{pmatrix} =: f(x, y) \ .$$

The point

$$P = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is the only fixed point. Is it stable? We have

$$Df(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} ,$$

thus the eigenvalues of $Df(P)$ can not be used for stability analysis.

Consider the function

$$L(x,y) = x^4 + y^4$$

and let

$$v(t) = L(x(t), y(x))$$

where $(x(t), y(t))$ is a solution of the system. We have

$$v'(t) = (L_x x' + L_y y')(t) = \Big(4x^3(-y^3) + 4y^3 x^3\Big)(t) = 0 .$$

This implies that

$$x^4(t) + y^4(t) = x_0^4 + y_0^4 \quad \text{for all} \quad t \geq 0 .$$

We claim that this implies stability of the fixed point $P = (0,0)$.

Recall:

Definition: Consider the system $x' = f(x)$ and denote the solution with initial condition $x(0) = x_0$ by $\phi(t, x_0)$. A fixed point x^* is called stable if for all $\varepsilon > 0$ there exists $\delta > 0$ so that

$$|x_0 - x^*| \leq \delta \quad \text{implies} \quad |\phi(t, x_0) - x^*| \leq \varepsilon \quad \text{for all} \quad t \geq 0 .$$

To return to the example, assume that

$$x_0^2 + y_0^2 \leq \delta^2 .$$

From

$$x^4(t) + y^4(t) = x_0^4 + y_0^4 \quad \text{for all} \quad t \geq 0$$

obtain that, for $t \geq 0$,

$$x^2(t) \leq \sqrt{x_0^4 + y_0^4} \leq x_0^2 \leq \delta^2$$

and, similarly, $y^2(t) \leq \delta^2$. Therefore,

$$x^2(t) + y^2(t) \leq 2\delta^2 = \varepsilon^2$$

if we choose $\delta = \varepsilon/\sqrt{2}$. This proves that the origin is a stable fixed point of the given system. The function $L(x,y)$ is an example of a Liapunov function for the fixed point $P = 0$.

In the following, we assume that

$$f:\mathbb{R}^n\to\mathbb{R}^n,\quad f\in C^1,\quad |Df(x)|\le C\quad \text{for all}\quad x\in\mathbb{R}^n\ ,$$

and we assume that $f(0)=0$. The solution of the IVP

$$x'=f(x),\quad x(0)=x_0\ ,$$

is denoted by $\phi(t,x_0)$.

Let

$$L:\mathbb{R}^n\to\mathbb{R},\quad L\in C^1\ ,$$

and let $v(t)=L(\phi(t,x_0))$. We have

$$v'(t)=\nabla L(\phi(t,x_0))f(\phi(t,x_0)=:\dot{L}(\phi(t,x_0))$$

where

$$\dot{L}(x)=\nabla L(x)f(x)=\sum_{j=1}^n\frac{\partial L}{\partial x_j}(x)f_j(x)\ .$$

Definition: Let $L:\mathbb{R}^n\to\mathbb{R}, L\in C^1$.

a) L is called a Liaponov function (corresponding to $f(x)$ and the fixed point $x^*=0$) if

$$L(0)=0,\quad L(x)>0\quad\text{and}\quad \dot{L}(x)\le 0\quad\text{for all}\quad x\in\mathbb{R}^n\setminus\{0\}\ .$$

b) L is called a strict Liaponov function (corresponding to $f(x)$ and the fixed point $x^*=0$) if

$$L(0)=0,\quad L(x)>0\quad\text{and}\quad \dot{L}(x)<0\quad\text{for all}\quad x\in\mathbb{R}^n\setminus\{0\}\ .$$

Theorem 10.4 *Consider that system $x'=f(x)$ where $f(0)=0$ and assume that there exists a corresponding Liapunov function. Then the fixed point $x^*=0$ is stable.*

Proof: We have to show that for all $\varepsilon>0$ there exists $\delta>0$ so that

$$|x_0|\le\delta\quad\text{implies}\quad|\phi(t,x_0)|\le\varepsilon\quad\text{for all}\quad t\ge 0\ .$$

For given $\varepsilon>0$ set

$$m_\varepsilon:=\min\{L(x)\ :\ |x|=\varepsilon\}>0\ .$$

Since $L(0)=0$ there exists $\delta>0$ with $0<\delta<\varepsilon$ and

$$L(x)<m_\varepsilon\quad\text{for}\quad|x|\le\delta\ .$$

Let $|x_0|\le\delta$. We claim that $|\phi(t,x_0)|<\varepsilon$ for all $t\ge 0$. Set

$$v(t)=L(\phi(t,x_0))\ .$$

We have $v(0) = L(x_0)$, thus $v(0) < m_\varepsilon$. Also, $v'(t) \leq 0$ for all $t \geq 0$, thus $v(t) < m_\varepsilon$ for all $t \geq 0$. If $|\phi(t, x_0)| < \varepsilon$ does not hold for all $t \geq 0$ then there exists $t_1 > 0$ with

$$|\phi(t_1, x_0)| = \varepsilon \ .$$

This would imply that

$$v(t_1) = L(\phi(t_1, x_0) \geq m_\varepsilon \ .$$

This contradicts the estimate $v(t) < m_\varepsilon$ for all $t \geq 0$. $\diamond$

Theorem 10.5 *Consider that system $x' = f(x)$ where $f(0) = 0$ and assume that there exists a corresponding strict Liapunov function. Then the fixed point $x^* = 0$ is asymptotically stable.*

Proof: By the previous theorem, the fixed point $x^* = 0$ is stable. We must show that there exists $\delta > 0$ so that $|x_0| \leq \delta$ implies that $\phi(t, x_0) \to 0$ as $t \to \infty$.

Let $x_0 \in \mathbb{R}^n, x_0 \neq 0$, thus $\phi(t, x_0) \neq 0$ for all t. Set $v(t) = L(\phi(t, x_0))$ and note that $v'(t) < 0$ for all t.

Let t_k denote a sequence of positive times with $t_k \to \infty$ and assume that $\phi(t_k, x_0) \to y_0$. We claim that $y_0 = 0$. Suppose $y_0 \neq 0$. Since $v'(t) < 0$ for all t we have

$$L(\phi(t_k, x_0)) > L(y_0) \quad \text{for all} \quad k \in \mathbb{N} \ .$$

Fix some $t > 0$ and consider $\phi(t, y_0)$. We have

$$L(\phi(t, y_0)) < L(y_0) \ .$$

Here we used the assumption that $y_0 \neq 0$. By continuity there exists $\varepsilon > 0$ so that

$$L(\phi(t, y)) < L(y_0) \quad \text{for} \quad |y - y_0| < \varepsilon \ .$$

For large k we have

$$|\phi(t_k, x_0) - y_0| < \varepsilon$$

and can use the above estimate for $y = \phi(t_k, x_0)$ to obtain that

$$L(\phi(t_k + t, x_0)) < L(y_0) \ .$$

This contradicts the estimate

$$L(\phi(t_k + t, x_0)) > L(y_0)$$

which follows from

$$L(\phi(t_j, x_0)) > L(y_0) \quad \text{for all} \quad j \in \mathbb{N} \ .$$

This contradiction shows that $y_0 = 0$.

Thus, if $t_k \to \infty$ and if $\phi(t_k, x_0)$ converges as $k \to \infty$, then

$$\phi(t_k, x_0) \to 0 \quad \text{as} \quad k \to \infty \ .$$

Recall: Given $\varepsilon > 0$ there exists $\delta > 0$ so that $|x_0| \le \delta$ implies $|\phi(t, x_0)| \le \varepsilon$ for $t \ge 0$. We use this for $\varepsilon = 1$. If $|x_0| \le \delta$ then $|\phi(t, x_0)| \le 1$ for $t \ge 0$. We claim that

$$\phi(t, x_0) \to 0 \quad \text{as} \quad t \to \infty \quad \text{if} \quad |x_0| \le \delta \ .$$

If this does not hold then there exists $\varepsilon > 0$ so that for all $T > 0$ there exists $t \ge T$ with

$$|\phi(t, x_0)| \ge \varepsilon \ .$$

We use this for $T = k \in \mathbb{N}$ and obtain $t_k \ge k$ with

$$|\phi(t_k, x_0)| \ge \varepsilon \quad \text{for all} \quad k \in \mathbb{N} \ .$$

However, the sequence $\phi(t_k, x_0)$ is bounded and, therefore, has a convergent subsequence,

$$\phi(t_{k_n}, x_0) \to y_0 \ .$$

But we have proved that $y_0 = 0$ and obtain a contradiction. $\diamond$

Remark: Assume existence of a strict Liapunov function (corresponding to f and $x^* = 0$) and assume that every semi–orbit $\phi(t, x_0), t \ge 0$, is bounded. Then the arguments of the proof given above show that $\phi(t, x_0) \to 0$ as $t \to \infty$, i.e., the fixed point $x^* = 0$ is globally attracting.

Example 2 Consider the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} -x^3 - y^3 \\ x^3 - y^3 \end{pmatrix} =: f(x, y) \ .$$

Let

$$L(x, y) = x^4 + y^4 \ .$$

We have

$$\begin{aligned} \dot{L} &= \nabla L \cdot f \\ &= 4x^3(-x^3 - y^3) + 4y^3(x^3 - y^3) \\ &= -4x^6 - 4y^6 \end{aligned}$$

Thus $\dot{L}(x, y) < 0$ for $(x, y) \ne 0$. The function $L(x, y)$ is a strict Liapunov function.

We claim that every semi–orbit

$$(x(t), y(t)), \quad t \ge 0 \ ,$$

is bounded. We have

$$x^4(t) + y^4(t) \le x_0^4 + y_0^4 =: C^4 \quad \text{for} \quad t \ge 0 \; .$$

This yields that $x^4(t) \le C^4, x^2(t) \le C^2$. Obtain

$$x^2(t) + y^2(t) \le 2C^2 \quad \text{for} \quad t \ge 0 \; .$$

Every semi–orbit is bounded; therefore, the fixed point $(0, 0)$ is globally attracting.

11 The Energy Method and the Mathematical Pendulum

Many nonlinear ODEs cannot be solved explicitly by integrations. However, if they can be solved explicitly, the solution formula may be useful and interesting.

The energy method is an integration technique which can be applied to scalar IVPs of the form

$$x'' + f(x) = 0, \quad x(0) = x_0, \quad x'(0) = x_1 \tag{11.1}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a smooth function.

In Section 11.2 we will apply the energy technique to the equations for the Mathematical Pendulum. One obtains a connection to complete elliptic integrals studied in complex variables.

11.1 Solution of $x'' + f(x) = 0$ by Integration

Let $F(x) = \int_0^x f(\xi)\,d\xi$, thus $F'(x) = f(x)$. Multiply the ODE $x'' + f(x) = 0$ by x' to obtain

$$\frac{d}{dt}\Big(\frac{1}{2}(x')^2 + F(x)\Big) = 0 \ .$$

If the initial data $x(0) = x_0, x'(0) = x_1$ are satisfied, then one obtains

$$\frac{1}{2}\,(x'(t))^2 + F(x(t)) = C_0$$

with

$$C_0 = \frac{1}{2}\,x_1^2 + F(x_0) \ .$$

Therefore,

$$(x'(t))^2 = 2C_0 - 2F(x(t))$$

which yields that

$$\begin{aligned} \frac{dx}{dt} &= \pm\sqrt{2}\,\sqrt{C_0 - F(x)} \\ \frac{dx}{\sqrt{C_0 - F(x)}} &= \pm\sqrt{2}\,dt \end{aligned}$$

thus

$$\int_{x_0}^{x(t)} \frac{d\xi}{\sqrt{C_0 - F(\xi)}} = \pm\sqrt{2}\,t \ .$$

If the integral

$$\int \frac{d\xi}{\sqrt{C_0 - F(\xi)}}$$

can be obtained explicitly, one may obtain a useful formula for solutions of the ODE $x'' + f(x) = 0$.

11.2 The Mathematical Pendulum

The pendulum equation reads

$$\alpha'' + \omega^2 \sin\alpha = 0 \tag{11.2}$$

where $\alpha = \alpha(t)$ is the angular displacement of the pendulum and
$\omega = g/l$. We consider the ODE (11.2) with initial condition

$$\alpha(0) = \alpha_0, \quad \alpha'(0) = 0$$

where $0 < \alpha_0 < \pi$. By sketching the phase plane diagram it is clear that the solution is periodic and α_0 is its maximal amplitude. If the period is T then

$$\alpha(t) > 0 \quad \text{for} \quad 0 \le t < \frac{T}{4} \quad \text{and} \quad \alpha(\frac{T}{4}) = 0 \ .$$

11.2.1 The Linearized Equation

If $|\alpha(t)|$ is small one may replace $\sin\alpha$ by α to obtain the linear IVP

$$\alpha'' + \omega^2\alpha = 0, \quad \alpha(0) = \alpha_0, \quad \alpha'(0) = 0 \ .$$

The solution is

$$\alpha(t) = \alpha_0 \cos(\omega t) \ .$$

Its period

$$T_0 = \frac{2\pi}{\omega}$$

is independent of the amplitude α_0.

11.2.2 Application of the Energy Method

The energy method applied to the nonlinear problem yields

$$\frac{1}{2}\alpha'^2 - \omega^2 \cos\alpha = -\omega^2 \cos\alpha_0 \ .$$

Therefore,

$$\alpha'(t) = -\sqrt{2}\,\omega\Big(\cos\alpha(t) - \cos\alpha_0\Big)^{1/2} \ .$$

This yields

$$\frac{d\alpha}{(\cos\alpha - \cos\alpha_0)^{1/2}} = -\sqrt{2}\,\omega\, dt$$

and

$$\int_{\alpha(t)}^{\alpha_0} \frac{d\theta}{(\cos\theta - \cos\alpha_0)^{1/2}} = \sqrt{2}\,\omega t \ .$$

We use the identity

$$\cos x = 1 - 2\sin^2(x/2)$$

to obtain

$$\int_{\alpha(t)}^{\alpha_0} \frac{d\theta}{(\sin^2(\alpha_0/2) - \sin^2(\theta/2))^{1/2}} = 2\omega t \ .$$

Using the notation

$$k = \sin(\alpha_0/2)$$

we have

$$\int_{\alpha(t)}^{\alpha_0} \frac{d\theta}{(k^2 - \sin^2(\theta/2))^{1/2}} = 2\omega t \ .$$

If T is the period of the oscillation, then $\alpha(T/4) = 0$. This yields the following relation between the period T and the amplitude α_0:

$$\int_{0}^{\alpha_0} \frac{d\theta}{(k^2 - \sin^2(\theta/2))^{1/2}} = \frac{\omega T}{2} \ . \tag{11.3}$$

Define a new variable ϕ by

$$\sin(\theta/2) = k\sin\phi \ .$$

Then $\theta = 0$ corresponds to $\phi = 0$ and $\theta = \alpha_0$ corresponds to $\phi = \pi/2$. Further, note that

$$\cos(\theta/2)d\theta = 2k\cos\phi\, d\phi.$$

Therefore,

$$\begin{aligned} \frac{d\theta}{(k^2 - \sin^2(\theta/2))^{1/2}} &= \frac{2k\cos\phi\, d\phi}{\cos(\theta/2)(k^2 - k^2\sin^2\phi)^{1/2}} \\ &= \frac{2d\phi}{\cos(\theta/2)} \\ &= \frac{2d\phi}{(1 - k^2\sin^2\phi)^{1/2}} \end{aligned}$$

The relation (11.3) becomes

$$T = \frac{4}{\omega}\int_0^{\pi/2} \frac{d\phi}{(1 - k^2\sin^2\phi)^{1/2}}$$

where

$$k = \sin(\alpha_0/2)\ .$$

The function

$$K(k) = \int_0^{\pi/2} \frac{d\phi}{(1-k^2\sin^2\phi)^{1/2}}, \quad 0 \le k < 1\ ,$$

is called the complete elliptic integral of the first kind.

Thus we have shown the relation

$$T = T(\alpha_0) = \frac{4}{\omega}\, K(\sin(\alpha_0/2))$$

between the amplitude α_0 and the period T of the pendulum.

Discussion: Clearly, for $\alpha_0 = 0$ one obtains

$$T(0) = \frac{4}{\omega}\, K(0) = \frac{4}{\omega}\,\frac{\pi}{2} = \frac{2\pi}{\omega}\ .$$

This agrees with the period

$$T_0 = \frac{2\pi}{\omega}$$

of the solution of the linearized equation.

If $\alpha_0 \to \pi$ then $k = \sin(\alpha_0/2) \to 1$ and $K(k) \to \infty$.

The period $T(\alpha_0)$ approaches ∞ as $\alpha_0 \to \pi$. The initial state

$$\alpha(0) = \pi, \quad \alpha'(0) = 0$$

is an unstable fixed point of the pendulum equation

$$\alpha'' + \omega^2 \sin\alpha = 0$$

written as a first order system. Using the notations

$$x = \alpha, \quad y = \alpha'\ ,$$

the system reads

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} y \\ -\omega^2 \sin x \end{pmatrix} =: f(x,y)\ .$$

We have $f(\pi,0) = (0,0)^T$ and

$$f'(x,y) = \begin{pmatrix} 0 & 1 \\ -\omega^2\cos x & 0 \end{pmatrix}, \quad A := f'(\pi,0) = \begin{pmatrix} 0 & 1 \\ \omega^2 & 0 \end{pmatrix}\ .$$

The matrix A has the eigenvalues $\lambda_{1,2} = \pm\omega$. The unstable equilibrium point

$$\alpha(0) = \pi, \quad \alpha'(0) = 0$$

is a saddle point.

12 Planetary Motion: The Two Body Problem

Kepler's three laws of planetary motion were explained by Newton using Newton's second law and a central gravitational field which decays like $1/r^2$.

Kepler's Laws:

1. The orbit of a planet is an ellipse with the Sun at one of the two foci.

2. A line segment joining a planet and the Sun sweeps out equal areas during equal intervals of time.

3. The square of a planet's orbital period T is proportional to the cube of the semi–major a of its orbit.

Remark: Correct is

$$\frac{a^3}{T^2} = G\,\frac{M+m}{4\pi^2} \sim G\,\frac{M}{4\pi^2}$$

where M is the mass of the sun, m is the mass of the planet, and

$$G = 6.674 * 10^{-11}\,\frac{meter^3}{kg\ sec^2} = 6.674 * 10^{-11}\,\frac{Newton\ meter^2}{kg^2}$$

is the gravitational constant. The unit of force

$$1\,Newton = 1\,\frac{kg\ meter}{sec^2}$$

is about the weight of an apple.

Nicolaus Copernicus, 1473–1543
Galileo Galilei, 1564–1642
Johannes Kepler, 1571– 1630
Isaac Newton, 1643–1727
Pierre–Simon Marquis de Laplace, 1749–1825
Friedrich Wilhelm Bessel, 1784–1846

Mechanics was considered a model for all sciences, the starting point of the scientific revolution.

12.1 The Two Body Problem: Reduction to One Body in a Central Field

Denote the masses by m_1 (sun) and m_2 (planet). Denote the gravitational constant by G. The body position vectors are $\mathbf{r}_1$ and $\mathbf{r}_2$. Further, let

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1\ .$$

The vector $\mathbf{r}$ is directed from m_1 to m_2, i.e., from the sun to the planet.

We then have (with $r = |\mathbf{r}|$):

$$\begin{aligned} m_1\ddot{\mathbf{r}}_1 &= Gm_1m_2\,\frac{\mathbf{r}}{r^3} =: \mathbf{F} \\ m_2\ddot{\mathbf{r}}_2 &= -\mathbf{F} \end{aligned}$$

Set $M = m_1 + m_2$. Then

$$\mathbf{R} = \frac{1}{M}(m_1\mathbf{r}_1 + m_2\mathbf{r}_2)$$

is the position vector of the center of mass.

It follows that

$$\ddot{\mathbf{R}} = 0 \ .$$

We may assume that we have chosen a coordinate system so that $\mathbf{R}(t) \equiv 0$.

Let

$$m = \frac{m_1 m_2}{m_1 + m_2} \ .$$

We have

$$m_1 m_2 \ddot{\mathbf{r}}_1 = m_2 \mathbf{F} \quad \text{and} \quad m_1 m_2 \ddot{\mathbf{r}}_2 = -m_1 \mathbf{F} \ .$$

Therefore,

$$m_1 m_2 \ddot{\mathbf{r}} = -(m_1 + m_2)\mathbf{F} = -M\mathbf{F} \ .$$

Dividing by M we obtain

$$m\ddot{\mathbf{r}} = -k\,\frac{\mathbf{r}}{r^3}, \quad k = Gm_1 m_2 \ .$$

12.2 One Body in a Central Field

The equation

$$m\ddot{\mathbf{r}} = -k\,\frac{\mathbf{r}}{r^3}$$

describes the motion of a body in an attracting central force field which decays like r^{-2}.

Angular Momentum. The vector

$$\mathbf{L} = \mathbf{r} \times m\dot{\mathbf{r}}$$

is the angular momentum. Since $\ddot{\mathbf{r}}$ has the same direction as $\mathbf{r}$ it follows that $\dot{\mathbf{L}} = 0$, thus

$$\mathbf{L}(t) = \mathbf{L}_0 = \mathbf{r}_0 \times m\dot{\mathbf{r}}_0$$

is a constant vector given in terms of the initial data

$$\mathbf{r}(0) = \mathbf{r}_0, \quad \dot{\mathbf{r}}(0) = \dot{\mathbf{r}}_0 \ .$$

We assume that $\mathbf{L}_0$ is a non–zero vector of length

$$|\mathbf{L}_0| = l_0 \ .$$

(One can also consider the case $\mathbf{L}_0 = 0$. If $L_0 = 0$ then $\mathbf{r}((t)$ and $\dot{\mathbf{r}}(t)$ are parallel at all time. This leads to motion in a straight line and possible collision of the body with the center.)

Since $\mathbf{r}(t)$ is orthogonal to $\mathbf{L}_0$ at all time, the motion takes place in the plane through the origin orthogonal to $\mathbf{L}_0$.

We choose polar coordinates in this plane so that

$$\mathbf{r} = r(c,s,0)^T \quad \text{with} \quad c = \cos\theta, \quad s = \sin\theta \ .$$

It then follows that

$$\dot{\mathbf{r}} = \dot{r}(c,s,0)^T + r(-s,c,0)^T\dot{\theta} \ ,$$

thus

$$\mathbf{L} = \mathbf{r} \times m\dot{\mathbf{r}} = mr^2\dot{\theta}\mathbf{e}_3 \ .$$

(Note that

$$\begin{pmatrix} c \\ s \\ 0 \end{pmatrix} \times \begin{pmatrix} -s \\ c \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{e}_3$$

since the two vectors on the left side of the equation are orthogonal and both have length one.)

Therefore, assuming that $\dot{\theta} > 0$,

$$l_0 = mr^2\dot{\theta} \ .$$

The geometric interpretation of the time–independence of

$$\frac{1}{2}r^2\frac{d\theta}{dt}$$

is Kepler's second law.

Conservation of Energy. We define

$$\begin{aligned} E_{kin}(t) &= \frac{m}{2}\,|\dot{\mathbf{r}}(t)|^2 \\ E_{pot}(t) &= -\frac{k}{r(t)} \end{aligned}$$

and

$$E(t) = E_{kin}(t) + E_{pot}(t) \ .$$

Obtain

$$\begin{aligned} \dot{E}_{kin}(t) &= m\langle \dot{\mathbf{r}}(t), \ddot{\mathbf{r}}(t)\rangle \\ &= -\frac{k}{r^3(t)}\,\langle \dot{\mathbf{r}}(t), \mathbf{r}(t)\rangle \end{aligned}$$

Here (ignoring the trivial third dimension)

$$\mathbf{r} = r \begin{pmatrix} c \\ s \end{pmatrix} , \quad \dot{\mathbf{r}} = \dot{r} \begin{pmatrix} c \\ s \end{pmatrix} + r \begin{pmatrix} -s \\ c \end{pmatrix} \dot{\theta} \ ,$$

thus

$$\langle \dot{\mathbf{r}}(t), \mathbf{r}(t) \rangle = r(t)\dot{r}(t)$$

and

$$\dot{E}_{kin}(t) = -\frac{k\dot{r}(t)}{r^2(t)} \ .$$

Also, $E_{pot}(t) = -\frac{k}{r(t)}$ implies that

$$\dot{E}_{pot}(t) = \frac{k\dot{r}(t)}{r^2(t)} \ .$$

One obtains that $E'(t) \equiv 0$, thus

$$E(t) \equiv E_0 = \frac{m}{2} \, |\dot{\mathbf{r}}_0|^2 - \frac{k}{r_0} \ .$$

The energy $E(t)$ is a constant in time with a value determined by the initial data.

Below we will assume that $E_0 < 0$. The cases $E_0 = 0$ and $E_0 > 0$ can be treated similarly, leading to a parabolic and a hyperbolic orbit.

From

$$\dot{\mathbf{r}} = \dot{r} \begin{pmatrix} c \\ s \end{pmatrix} + r \begin{pmatrix} -s \\ c \end{pmatrix} \dot{\theta}$$

it follows that

$$|\dot{\mathbf{r}}(t)|^2 = \dot{r}^2(t) + r^2(t)\dot{\theta}^2(t) \ ,$$

thus

$$E(t) = \frac{m}{2} \left(\dot{r}^2(t) + r^2(t)\dot{\theta}^2(t) \right) - \frac{k}{r(t)} \equiv E_0 \ .$$

We now use conservation of momentum to eliminate $\dot{\theta}(t)$ from this equation. We have $l \equiv mr^2(t)\dot{\theta}(t)$, thus

$$mr^2(t)\dot{\theta}^2(t) = \frac{l^2}{mr^2(t)}$$

and obtain the fundamental relation

$$E_0 = \frac{m}{2} \, \dot{r}^2(t) + \frac{l^2}{2mr^2(t)} - \frac{k}{r(t)} \ ,$$

which is a differential equation for $r(t)$. In this equation, the constants $l = mr_0^2\dot{\theta}_0$ and E_0 are determined by the initial data.

12.3 The Equation of an Ellipse in Cartesian and Polar Coordinates

Apollonius of Perga (240 BC – 190 BC) defined an ellipse as the curve where a plane intersects a circular cone. The French mathematician Dandelin (1794 – 1847) gave a great proof that ellipses can also be defined in another way: Let F_1 and F_2 denote two points in the plane and let P denote all points in the plane with

$$|F_1 - P| + |F_2 - P| = 2a \ .$$

Then the points P lie on the ellipse with foci F_1 and F_2. See the article *Dandelin spheres* in Wikipedia.

An Ellipse in Cartesian Coordinates. Let $0 < c < a$ and let

$$F_1 = (c, 0), \quad F_2 = (-c, 0) \ .$$

Consider all points $P = (x, y)$ with

$$|F_1 - P| + |F_2 - P| = 2a \ .$$

By definition, these points form the ellipse with foci F_1, F_2 and major semi–axis a. We want to derive the equation satisfied by (x, y). Let

$$\begin{aligned} d_1^2 &= |F_1 - P|^2 = (x - c)^2 + y^2 \\ d_2^2 &= |F_2 - P|^2 = (x + c)^2 + y^2 \end{aligned}$$

The equation

$$d_1 + d_2 = 2a$$

is equivalent to

$$d_2^2 = (2a - d_1)^2 \ ,$$

thus

$$(x + c)^2 + y^2 = 4a^2 - 4ad_1 + (x - c)^2 + y^2 \ .$$

This is equivalent to

$$4cx - 4a^2 = -4ad_1 \ ,$$

or

$$a^2 - cx = ad_1 \ .$$

Squaring yields

$$a^4 - 2a^2cx + c^2x^2 = a^2(x^2 - 2cx + c^2 + y^2)$$

or

$$a^4 + c^2x^2 = a^2x^2 + a^2c^2 + a^2y^2 \ .$$

One obtains the equivalent condition

$$a^2(a^2 - c^2) = x^2(a^2 - c^2) + a^2y^2 \ .$$

If we define $b > 0$ by $b^2 = a^2 - c^2$ then we obtain

$$1 = \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 \ . \tag{12.1}$$

This is the equation of an ellipse with semi–axes $a \geq b > 0$ and foci

$$F_1 = (c, 0), \quad F_2 = (-c, 0)$$

where $c^2 = a^2 - b^2$.

The **area of the ellipse** given by (12.1) is

$$A = \pi ab \ .$$

Details: Solve (12.1) for y to obtain

$$y = \pm b\sqrt{1 - \frac{x^2}{a^2}} \ ,$$

thus

$$\begin{aligned}
\frac{1}{2}A &= b\int_{-a}^{a} \sqrt{1 - \frac{x^2}{a^2}}\, dx \\
&= ab\int_{-1}^{1} \sqrt{1 - \xi^2}\, d\xi \\
&= ab\int_{-\pi/2}^{\pi/2} \cos^2\theta\, d\theta \\
&= \frac{\pi}{2}\, ab
\end{aligned}$$

An Ellipse in Polar Coordinates. Denote polar coordinates centered at F_1 by (r, θ). Let $P = (x, y)$ denote a point on the ellipse with polar coordinates (r, θ). The distances of P to the foci are

$$d_1 = |F_1 - P| = r, \quad d_2 = |F_2 - P| = 2a - r \ .$$

If $\phi = \pi - \theta$ then the cosine theorem in the triangle F_1PF_2 yields

$$(2a - r)^2 = 4c^2 + r^2 - 4cr\cos\phi \ .$$

Therefore,

$$4a^2 - 4ar + r^2 = 4c^2 + r^2 - 4cr\cos\phi \ .$$

This yields

$$a^2 - c^2 = r(a - c\cos\phi) \ .$$

Introduce the eccentricity $\varepsilon = c/a$ of the ellipse to obtain

$$r(\theta) = \frac{a(1-\varepsilon^2)}{1+\varepsilon\cos\theta} \; . \tag{12.2}$$

This is the equation of an ellipse in polar coordinates where the center of the polar coordinate system is the point $F_1 = (c, 0)$ with $c = \varepsilon a$. The point F_1 is one of the two foci, $(c, 0)$ and $(-c, 0)$, of the ellipse.

We note that

$$r(0) = a(1-\varepsilon) = a - c \; . \tag{12.3}$$

Summary: In Cartesian coordinates the equation (12.1) describes an ellipse with semi–axes $a \geq b > 0$. If

$$c^2 = a^2 - b^2, \quad 0 \leq c < a \; ,$$

then the ellipse has the foci at

$$F_1 = (c, 0) \quad \text{and} \quad F_2 = (-c, 0) \; .$$

The eccentricity of the ellipse is

$$\varepsilon = \frac{c}{a}, \quad 0 \leq \varepsilon < 1 \; .$$

In polar coordinates centered at $F_1 = (c, 0)$ the ellipse is given by

$$r(\theta) = \frac{a(1-\varepsilon^2)}{1+\varepsilon\cos\theta}, \quad 0 \leq \theta \leq 2\pi \; .$$

12.4 Derivation of Kepler's Orbit by Conservation of Energy and Angular Momentum

Recall the dynamical equations

$$\begin{aligned} \frac{dr}{dt} &= \pm\Big(\frac{2E_0}{m} + \frac{2k}{mr} - \frac{l^2}{m^2r^2}\Big)^{1/2} \\ \frac{d\theta}{dt} &= \frac{l}{mr^2} \end{aligned}$$

The equations yield

$$\frac{dr}{d\theta} = \pm r^2 \Big(A + \frac{B}{r} - \frac{1}{r^2}\Big)^{1/2} \tag{12.4}$$

with

$$\begin{aligned} A &= \frac{2E_0}{m} \cdot \frac{m^2}{l^2} = \frac{2mE_0}{l^2} \\ B &= \frac{2k}{m} \cdot \frac{m^2}{l^2} = \frac{2mk}{l^2} \end{aligned}$$

We will solve the differential equation (12.4) for $r = r(\theta)$ using separation of variables. Formally:

$$\frac{dr}{\pm r^2\left(A + \frac{B}{r} - \frac{1}{r^2}\right)^{1/2}} = d\theta\ .$$

In the indefinite integral

$$Int = \int \frac{dr}{\pm r^2\left(A + \frac{B}{r} - \frac{1}{r^2}\right)^{1/2}}$$

use the substitution

$$r = \frac{1}{u}, \quad dr = -\frac{du}{u^2}, \quad -\frac{dr}{r^2} = du$$

to obtain

$$Int = \int \pm(A + Bu - u^2)^{-1/2}\, du, \quad u = 1/r\ .$$

The integral can be evaluated in terms of arccos. Note that the identity

$$\alpha = \arccos(\cos\alpha)$$

yields

$$\arccos'(y) = \pm(1 - y^2)^{-1/2}\ .$$

Set

$$h(u) = \arccos\left(\frac{2u - B}{q}\right)$$

where the constant q is to be determined. We have

$$\begin{aligned} h'(u) &= \pm\frac{2}{q}\Big(1 - q^{-2}(4u^2 - 4uB + B^2)\Big)^{-1/2} \\ &= \pm 2\Big(q^2 - 4u^2 + 4uB - B^2\Big)^{-1/2} \\ &= \pm\Big(\frac{1}{4}(q^2 - B^2) + uB - u^2\Big)^{-1/2} \end{aligned}$$

If

$$\frac{1}{4}(q^2 - B^2) = A$$

then

$$h'(u) = \pm(A + Bu - u^2)^{-1/2}\ .$$

Thus, if we let

$$q = \sqrt{B^2 + 4A}$$

then

$$Int = \int \pm(A + Bu - u^2)^{-1/2}\, du = \arccos\Big(\frac{2u - B}{q}\Big) + const\ .$$

Recalling the substitution $r = 1/u$ we obtain

$$\theta - \theta_0 = \arccos\Big(\frac{1}{q}(\frac{2}{r} - B)\Big)\ ,$$

thus

$$\frac{1}{r} = \frac{B}{2}\Big(1 + \frac{q}{B}\cos(\theta - \theta_0)\Big)\ .$$

We now define ε and a by

$$\varepsilon = \frac{q}{B}, \quad a(1 - \varepsilon^2) = \frac{2}{B}$$

and find that

$$r(\theta) = \frac{a(1 - \varepsilon^2)}{1 + \varepsilon\cos(\theta - \theta_0)}\ .$$

Using the condition $r(0) = a(1 - \varepsilon)$ (see (12.3)) we obtain $\theta_0 = 0$, thus

$$r(\theta) = \frac{a(1 - \varepsilon^2)}{1 + \varepsilon\cos\theta}\ .$$

Since this equation agrees with (12.2), the orbit of the planet is an ellipse. We have derived Kepler's first law using Newton's second law and the law of gravitational attraction.

It is interesting to relate the semi–axis a to the energy E_0. We have

$$\varepsilon^2 - 1 = \frac{4A}{B^2} = \frac{2l^2 E_0}{mk^2}$$

and

$$a(1 - \varepsilon^2) = \frac{2}{B} = \frac{l^2}{mk}\ ,$$

thus

$$a = \frac{k}{2|E_0|}\ .$$

12.5 Derivation of Kepler's Orbit by Lagrange's Equations

Recall the energies

$$\begin{aligned} E_{kin} &= \frac{m}{2}\left(\dot r^2 + r^2\dot\theta^2\right) \\ E_{pot} &= -\frac{k}{r} \end{aligned}$$

and the size of the angular momentum

$$l = mr^2\dot{\theta}\ .$$

Obtain $r^2\dot{\theta}^2 = \frac{l^2}{m^2r^2}$ and

$$\begin{aligned} E_0 &\equiv E_{kin} + E_{pot} \\ &= \frac{m}{2}\,\dot{r}^2 + \frac{m}{2}\,r^2\dot{\theta}^2 - \frac{k}{r} \\ &= \frac{m}{2}\,\dot{r}^2 + \frac{l^2}{2m}\,\frac{1}{r^2} - \frac{k}{r} \end{aligned}$$

Differentiation and division by $\dot{r}$ yields that

$$m\ddot{r} - \frac{l^2}{m}\,\frac{1}{r^3} + \frac{k}{r^2} = 0\ . \tag{12.5}$$

Let us derive this equation using the Lagrange function

$$\begin{aligned} L(r,\theta,\dot{r},\dot{\theta}) &= E_{kin} - E_{pot} \\ &= \frac{m}{2}\,\dot{r}^2 + \frac{m}{2}\,r^2\dot{\theta}^2 + \frac{k}{r} \end{aligned}$$

We have

$$\begin{aligned} L_{\dot{r}} &= m\dot{r} \\ L_r &= mr\dot{\theta} - \frac{k}{r^2} \\ L_{\dot{\theta}} &= mr^2\dot{\theta} \\ L_\theta &= 0 \end{aligned}$$

The Lagrange equations are

$$\begin{aligned} \frac{d}{dt}\,L_{\dot{r}} - L_r &= 0 \\ \frac{d}{dt}\,L_{\dot{\theta}} - L_\theta &= 0 \end{aligned}$$

This yields that

$$\begin{aligned} m\ddot{r} - mr\dot{\theta}^2 + \frac{k}{r^2} &= 0 \\ mr^2\dot{\theta} &= l \end{aligned}$$

As above, one obtains the equation (12.5) for $r = r(t)$.

From $l = mr^2\frac{d\theta}{dt}$ obtain that

$$\frac{d}{dt} = \frac{l}{mr^2}\frac{d}{d\theta} ,$$

thus

$$\begin{aligned} \frac{dr}{dt} &= \frac{l}{mr^2}\frac{dr}{d\theta} \\ \frac{d^2r}{dt^2} &= \frac{l}{mr^2}\frac{d}{d\theta}\left(\frac{l}{mr^2}\frac{dr}{d\theta}\right) \\ &= \frac{l^2}{m^2r^4}\frac{d^2r}{d\theta^2} - \frac{2l^2}{m^2r^5}\left(\frac{dr}{d\theta}\right)^2 \end{aligned}$$

Using these equations in (12.5) yields that

$$\frac{l^2}{mr^4}\frac{d^2r}{d\theta^2} - \frac{2l^2}{mr^5}\left(\frac{dr}{d\theta}\right)^2 - \frac{l^2}{mr^3} + \frac{k}{r^2} = 0 .$$

Multiply by $\frac{mr^2}{l^2}$ to obtain

$$\frac{1}{r^2}\frac{d^2r}{d\theta^2} - \frac{2}{r^3}\left(\frac{dr}{d\theta}\right)^2 - \frac{1}{r} + \frac{mk}{l^2} = 0 . \qquad (12.6)$$

Set

$$u(\theta) = \frac{1}{r(\theta)}$$

and obtain that

$$\begin{aligned} \frac{du}{d\theta} &= -\frac{1}{r^2}\frac{dr}{d\theta} \\ \frac{d^2u}{d\theta^2} &= -\frac{1}{r^2}\frac{d^2r}{d\theta^2} + \frac{2}{r^3}\left(\frac{dr}{d\theta}\right)^2 \end{aligned}$$

Therefore, if equation (12.6) is expressed as an equation for $u(\theta)$ one obtains that

$$u''(\theta) + u(\theta) = \frac{mk}{l^2} .$$

12.6 Negative E_0 Leads to an Ellipse

Set

$$\gamma = \frac{mk}{l_0^2} .$$

One can check the dimensionality of γ:

$$[\gamma] = \frac{1}{length} .$$

The general solution of the equation

$$u'' + u = \gamma$$

is

$$u(\theta) = \gamma + \alpha \cos\theta + \beta \sin\theta \ .$$

If one requires that

$$r(\theta) = \frac{1}{u(\theta)}$$

has a minimum at $\theta = 0$ then $\beta = 0$. This yields that

$$u(\theta) = \gamma(1 + \varepsilon \cos\theta) \ .$$

We claim that $\varepsilon^2 < 1$ if and only if $E_0 < 0$.

Recall the energy equation

$$\frac{m}{2}\dot{r}^2 + \frac{l_0^2}{2m}\frac{1}{r^2} - \frac{k}{r} = E_0 \ .$$

Assume that

$$r(\theta) = \frac{1}{\gamma} \cdot \frac{1}{1 + \varepsilon c} \quad \text{where} \quad c = \cos\theta \ .$$

Obtain

$$\frac{dr}{d\theta} = \frac{1}{\gamma} \cdot \frac{\varepsilon s}{(1 + \varepsilon c)^2} \quad \text{where} \quad s = \sin\theta \ .$$

We have

$$\begin{aligned}
\dot{r}(t) &= \frac{dr}{d\theta}\dot{\theta} \\
&= \frac{1}{\gamma} \cdot \frac{\varepsilon s}{(1+\varepsilon c)^2} \cdot \frac{l_0}{mr^2} \\
&= \frac{l_0}{m\gamma} \cdot \frac{\varepsilon s}{(1+\varepsilon c)^2} \cdot \gamma^2 (1+\varepsilon c)^2 \\
&= \frac{l_0 \gamma}{m} \varepsilon s
\end{aligned}$$

thus

$$\dot{r}^2 = \frac{l_0^2 \gamma^2}{m^2} \cdot \varepsilon^2 s^2 \ .$$

Since

$$\gamma = \frac{mk}{l_0^2}$$

obtain that

$$m\frac{l_0^2\gamma^2}{m^2} = k\gamma \ .$$

Obtain

$$\begin{aligned}
\frac{m}{2}\,\dot{r}^2 + \frac{l_0^2}{2m}\frac{1}{r^2} - \frac{k}{r} &= \frac{1}{2}\,k\gamma\cdot\Big(\varepsilon^2 s^2 + (1+\varepsilon c)^2\Big) - k\gamma(1+\varepsilon c) \\
&= k\gamma\Big(\frac{1}{2} + \frac{\varepsilon^2}{2} - 1\Big)
\end{aligned}$$

Using the energy equation, we have

$$\varepsilon^2 - 1 = \frac{2E_0}{k\gamma} \ .$$

This shows that $\varepsilon^2 < 1$ is equivalent to $E_0 < 0$.

12.7 Time Dependence

The Period T. Recall that $mr^2\dot{\theta} = l$, thus

$$\frac{1}{2}r^2 d\theta = \frac{l}{2m}\,dt \ .$$

Integration over one period yields the area of the ellipse,

$$\pi ab = \frac{l}{2m}\,T \ ,$$

thus

$$T = \frac{2\pi abm}{l} \ .$$

Using that

$$b^2 = a^2(1-\varepsilon^2) = \frac{al^2}{mk}$$

one obtains

$$T^2 = 4\pi^2 a^3 \frac{m}{k} \ .$$

Kepler's Third Law. To apply the result to the two body problem, we set

$$m = \frac{m_1 m_2}{m_1 + m_2}, \quad k = Gm_1 m_2 \ .$$

This yields

$$T^2 = \frac{4\pi^2 a^3}{G(m_1+m_2)} \ .$$

Kepler's Equation. Let u denote the eccentric anomaly, the angle at the center. Let $\mu = 2\pi t/T$ denote the mean anomaly. Kepler's equation is

$$\mu = u - \varepsilon \sin u \ .$$

Derivation: Denote the area SP_0P by $A(t)$. Then we have

$$A(0) = 0, \quad A(T) = \pi ab$$

and $A'(t) = const$ by Kepler's second law. Therefore,

$$A(t) = \frac{\pi abt}{T} \ .$$

Let $B(t)$ denote the area SP_0Q. Then we have

$$B(t) = \frac{a}{b}\, A(t) = \frac{\pi a^2 t}{T} \ .$$

One can also express $B(t)$ as the difference between the area of a circular sector and a triangle. This yields

$$B(t) = \frac{1}{2}\, a^2 u - \frac{1}{2} \varepsilon a^2 \sin u \ .$$

One then obtains Kepler's equation.

Solution of Kepler's Equation in Terms of Bessel Functions. Let $u(\mu) \in [0, \pi]$ denote the solution of Kepler's equation for $0 \le \mu \le \pi$. We have

$$u(0) = u(\pi) - \pi = 0$$

and write

$$u(\mu) - \mu = \sum_{n=1}^{\infty} c_n \sin(n\mu) \ .$$

One obtains

$$\begin{aligned}
\frac{\pi}{2} c_j &= \int_0^\pi (u(\mu) - \mu) \sin(j\mu)\, d\mu \\
&= \frac{1}{j} \int_0^\pi (u'(\mu) - 1) \cos(j\mu)\, d\mu \\
&= \frac{1}{j} \int_0^\pi u'(\mu) \cos(j\mu)\, d\mu
\end{aligned}$$

Note that the function $u(\mu)$ is strictly increasing since

$$\mu = u(\mu) - \varepsilon \sin(u(\mu))$$

implies

$$1 = u'(\mu) - \varepsilon \cos(u(\mu))\, u'(\mu) \ ,$$

thus

$$u'(\mu) > 0 \ .$$

Denote the inverse function of $u(\mu)$ by $\mu(v)$. We then have

$$u(\mu(v)) \equiv v, \quad u'(\mu(v))\mu'(v) \equiv 1 \ .$$

In the above integral we make the substitution

$$\mu = \mu(v), \quad 0 \le v \le \pi \ .$$

Since

$$u'(\mu)d\mu = dv$$

we obtain

$$\frac{\pi}{2}c_j = \frac{1}{j}\int_0^\pi \cos(j\mu(v))\, dv \ .$$

So far, we have not used Kepler's equation. Note that

$$\mu = u(\mu) - \varepsilon \sin(u(\mu))$$

and $u(\mu(v)) = v$. In Kepler's equation, replace μ by $\mu(v)$ to obtain that

$$\mu(v) = v - \varepsilon \sin v \ .$$

This yields

$$c_j = \frac{2}{\pi j}\int_0^\pi \cos\Big(j\varepsilon \sin(v) - jv\Big)\, dv \ .$$

Bessel defined

$$J_j(x) = \frac{1}{\pi}\int_0^\pi \cos\Big(x \sin(v) - jv\Big)\, dv \quad \text{for} \quad j = 0, 1, \ldots$$

One then obtains

$$c_j = \frac{2}{j}J_j(j\varepsilon)$$

and the solution of Kepler's equation is

$$u(\mu) = \mu + 2\sum_{n=1}^{\infty} J_j(\varepsilon j)\sin(j\mu) \ .$$

12.8 Bessel Functions via a Generating Function: Integral Representation

Let

$$g(z,t) = \exp\Big(\frac{z}{2}(t-\frac{1}{t})\Big) = \sum_{n=-\infty}^{\infty} J_n(z)t^n \ . \qquad (12.7)$$

Here $z \in \mathbb{C}$ and $t \in \mathbb{C} \setminus \{0\}$. For fixed z, the function $t \to g(z,t)$ is analytic in $\mathbb{C} \setminus \{0\}$ and, therefore, has a unique Laurent expansion about $t=0$. Also, for z and t real, the value $g(z,t)$ is real. This implies that $J_n(z)$ is real for real z.

Lemma 12.1 *Let $f(t) = \sum_{n=-\infty}^{\infty} c_n t^n$ denote a function which is analytic for $t \neq 0$ and which is real for real t. Then the coefficients c_n are all real.*

Proof: Set $g(t) = \sum \bar{c}_n t^n$. Then $g(t)$ is also analytic for $t \neq 0$. Furthermore, for real t,

$$f(t) = \bar{f}(t) = g(t) \ .$$

The identity theorem yields that f and t are identical. The uniqueness of the coefficients c_n implies that $c_n = \bar{c}_n$ is real. $\diamond$

In (12.7) substitute

$$t = e^{iv}, \quad t - \frac{1}{t} = 2i \sin v \ ,$$

to obtain

$$\exp(iz \sin v) = \sum J_n(z) e^{inv} \ .$$

Multiply by e^{-ijv} and integrate over $0 \le v \le 2\pi$ to obtain

$$\int_0^{2\pi} e^{iz\sin v - ijv}\, dv = 2\pi J_j(z) \ .$$

Let $z = x$ be real and take real parts to obtain that

$$2\pi J_j(x) = \int_0^{2\pi} \cos(x \sin v - jv)\, dv \ .$$

If we set

$$q(v) = \cos(x \sin v - jv)$$

then

$$\begin{aligned} q(2\pi - v) &= \cos(-x\sin v + jv) \\ &= \cos(x \sin v - jv) \\ &= q(v) \end{aligned}$$

This yields that

$$\pi J_j(x) = \int_0^{\pi} \cos(x \sin v - jv)\, dv \ .$$

13 The Stable Manifold Theorem

Example: Consider the nonlinear system of ODEs

$$\begin{aligned} x_1' &= -x_1 \\ x_2' &= -x_2 + x_1^2 \\ x_3' &= x_3 + x_1^2 \end{aligned}$$

for $x(t) \in \mathbb{R}^3$. We write the system as

$$x' = Ax + F(x)$$

where

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad F(x) = \begin{pmatrix} -0 \\ x_1^2 \\ x_1^2 \end{pmatrix} .$$

The point $x = 0$ is a fixed point and the system $x' = Ax$ is the linearization about this fixed point. The spaces

$$E^s = span\{e^1, e^2\}, \quad E^u = span\{e^3\}$$

are the stable and unstable spaces for the linear system $x' = Ax$.

Consider the nonlinear system $x' = Ax + F(x)$ with initial condition

$$x(0) = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} .$$

Obtain:

$$\begin{aligned} x_1(t) &= \alpha e^{-t} \\ x_2'(t) &= -x_2(t) + \alpha^2 e^{-2t} \\ x_3'(t) &= x_3(t) + \alpha^2 e^{-2t} \end{aligned}$$

It is elementary to show that

$$x(t) = \begin{pmatrix} \alpha e^{-t} \\ (\beta + \alpha^2)e^{-t} - \alpha^2 e^{-2t} \\ (\gamma + \frac{\alpha^2}{3})e^t - \frac{\alpha^2}{3} e^{-2t} \end{pmatrix} .$$

Details: To obtain $x_2(t)$ use the Ansatz $x_{part}(t) = ce^{-2t}$ to obtain a particular solution of the inhomogeneous equation. Then add a solution $x_{hom}(t) = qe^{-t}$ of the homogeneous equation. Determine q so that the initial condition holds. The same process works to determine $x_3(t)$.

Using the solution formula, it is clear that $x(t) \to 0$ as $t \to \infty$ if and only if $\gamma = -\alpha^2/3$. Also, $x(t) \to 0$ as $t \to -\infty$ if and only if $\alpha = \beta = 0$.

The two–dimensional manifold $\mathcal{S}$ of points

$$\begin{pmatrix} \alpha \\ \beta \\ -\frac{\alpha^2}{3} \end{pmatrix} \in \mathbb{R}^3 \quad \text{where} \quad \alpha, \beta \in \mathbb{R}$$

is the stable manifold of the nonlinear system $x' = Ax + F(x)$ for the fixed point $x = 0$. Note that the space E^s is tangent to $\mathcal{S}$ at $x = 0$. In the example, the unstable manifold of the nonlinear system agrees with E^u, the unstable subspace of the linear system $x' = Ax$.

General Theory: Let $f : \mathbb{R}^n \to \mathbb{R}^n, f \in C^1$, and assume that $f(0) = 0$. Set $A = Df(0)$. Consider the system

$$x' = f(x) = Ax + F(x) \ .$$

We have $F(0) = 0$ and $DF(0) = 0$. Clearly, $x = 0$ is a fixed point of the ODE $x' = f(x)$. The system $x' = Ax$ is obtain by linearizing about $x = 0$. We denote the solution of the initial value problem

$$x' = Ax + F(x), \quad x(0) = x_0 \ ,$$

by $x(t, x_0)$.

Assume that A has k eigenvalues $\lambda_1, \ldots, \lambda_k$ with negative real parts and $n - k$ eigenvalues $\lambda_{k+1}, \ldots, \lambda_n$ with positive real parts,

$$\text{Re}\, \lambda_j < 0 \quad \text{for} \quad j = 1, \ldots, k$$

and

$$\text{Re}\, \lambda_j > 0 \quad \text{for} \quad j = k+1, \ldots, n \ .$$

Here the eigenvalues are counted by their algebraic multiplicities.

Let $E^s \subset \mathbb{R}^n$ denote the sum of the generalized eigenspaces to $\lambda_1, \ldots, \lambda_k$ and let $E^u \subset \mathbb{R}^n$ denote the sum of the generalized eigenspaces to $\lambda_{k+1}, \ldots, \lambda_n$. Clearly, if $x_0 \in E^s$ then $e^{At} x_0 \to 0$ as $t \to \infty$ and if $x_0 \in E^u$ then $e^{At} x_0 \to 0$ as $t \to -\infty$. The space E^s is the stable space for the linear system $x' = Ax$ and E^u is the unstable space.

Theorem 13.1 *(Stable Manifold Theorem) Under the above assumptions, there exists a manifold $\mathcal{S} \subset \mathbb{R}^n$ of dimension k with $0 \in \mathcal{S}$ so that $x(t, x_0) \to 0$ as $t \to \infty$ if $x_0 \in \mathcal{S}$. The manifold $\mathcal{S}$ has the tangent space E^s at $x = 0$.*

Proof: There exists a matrix $T \in \mathbb{R}^{n \times n}$ so that

$$T^{-1} A T = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} = B$$

where $P \in \mathbb{R}^{k \times k}$ has the eigenvalues $\lambda_1, \ldots, \lambda_k$ and $Q \in \mathbb{R}^{(n-k) \times (n-k)}$ has the eigenvalues $\lambda_{k+1}, \ldots, \lambda_n$. Using the variable $y \in \mathbb{R}^n$ with $x = Ty$ one obtains

$$y' = By + T^{-1}F(Ty) = By + G(y) \ .$$

Here $G \in C^1, G(0) = 0, DG(0) = 0$. It suffices to prove the Stable Manifold Theorem for the y–system.

We note the following: There exist constants $K > 0, \alpha > 0, \sigma > 0$ with

$$\begin{array}{rcll} |e^{Pr}| & \leq & Ke^{-(\alpha+\sigma)r} & \text{for} \quad r \geq 0 \\ |e^{-Qr}| & \leq & Ke^{-\sigma r} & \text{for} \quad r \geq 0 \end{array}$$

Also, for any $\varepsilon > 0$ there exists $\delta > 0$ so that

$$|G(p) - G(q)| \leq \varepsilon |p - q| \quad \text{for} \quad |p| \leq \delta, |q| \leq |\delta| \ . \tag{13.1}$$

In the following, we will assume that $\varepsilon > 0$ is chosen so small that

$$\frac{\varepsilon K}{\sigma} \leq \frac{1}{4}$$

and $\delta > 0$ is chosen so that (13.1) holds.

Considering the system

$$y' = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} y + G(y) \tag{13.2}$$

we will write

$$y = \begin{pmatrix} y^I \\ y^{II} \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} G^I \\ G^{II} \end{pmatrix}$$

where $y^I \in \mathbb{R}^k, y^{II} \in \mathbb{R}^{n-k}$ and similarly for G.

In the following, we will assume

$$a \in \mathbb{R}^k, \quad 2K|a| \leq \delta \ .$$

We will try to determine a C^1–map

$$a \to \psi(a) \in \mathbb{R}^l$$

(where $l = n - k$) so that the solution $y(t)$ of (13.2) with initial value

$$y(0) = \begin{pmatrix} a \\ \psi(a) \end{pmatrix} \tag{13.3}$$

satisfies

$$y(t) \to 0 \quad \text{as} \quad t \to \infty \ .$$

Then, if $\psi(0) = 0$ and $D\psi(0) = 0$, the manifold

$$\mathcal{S} = \left\{ \begin{pmatrix} a \\ \psi(a) \end{pmatrix} \ : \ a \in \mathbb{R}^k, \quad |a| < \frac{\delta}{2K} \right\}$$

is the local stable manifold.

Motivation: Assume that $y(t)$ satisfies the IVP (13.2), (13.3).
Then we have

$$y^I(t) = e^{Pt}a + \int_0^t e^{P(t-s)} G^I(y(s))\,ds \tag{13.4}$$

$$y^{II}(t) = e^{Qt}\psi(a) + \int_0^t e^{Q(t-s)} G^{II}(y(s))\,ds \tag{13.5}$$

Assume that $y^{II}(t) \to 0$ as $t \to \infty$. Then we have

$$y^{II}(t) = e^{Qt}\Big(\psi(a) + \int_0^t e^{-Qs} G^{II}(y(s))\,ds\Big) \to 0 \quad \text{as} \quad t \to \infty \; .$$

This implies that

$$\psi(a) + \int_0^\infty e^{-Qs} G^{II}(y(s))\,ds\Big) = 0 \; ,$$

thus

$$y^{II}(t) = -\int_t^\infty e^{Q(t-s)} G^{II}(y(s))\,ds \; .$$

We now fix $a \in \mathbb{R}^k$ with $2K|a| \le \delta$ and consider the integral system

$$y^I(t) = e^{Pt}a + \int_0^t e^{P(t-s)} G^I(y(s))\,ds \tag{13.6}$$

$$y^{II}(t) = -\int_t^\infty e^{Q(t-s)} G^{II}(y(s))\,ds \tag{13.7}$$

If $y(t) = y(t,a)$ is the solution, then we will set $\psi(a) = y^{II}(0,a)$.
Set

$$Y = \Big\{ y \; : \; y : [0,\infty) \to \mathbb{R}^n, y \in C, |y(t)| \le 2K|a|e^{-\alpha t} \quad \text{for all} \quad t \ge 0 \Big\} \; .$$

For $y \in Y$ define

$$(Hy)(t) = \begin{pmatrix} (H^I y)(t) \\ (H^{II} y)(t) \end{pmatrix}$$

where

$$(Hy)^I(t) = e^{Pt}a + \int_0^t e^{P(t-s)} G^I(y(s))\,ds \tag{13.8}$$

$$(Hy)^{II}(t) = -\int_t^\infty e^{Q(t-s)} G^{II}(y(s))\,ds \tag{13.9}$$

Lemma 13.1 *If $y \in Y$ then $Hy \in Y$.*

Proof: For $y \in Y$ we have $|y(s)| \le \delta e^{-\alpha s} \le \delta$, thus

$$|G(y(s))| \le \varepsilon |y(s)| \le \varepsilon 2K|a|\, e^{-\alpha s}, \quad s \ge 0 \ .$$

Obtain that

$$|(H^I y)(t)| \le K e^{-\alpha t}|a| + \varepsilon 2K|a|K \int_0^t e^{-(\alpha+\sigma)(t-s)} e^{-\alpha s}\, ds \ .$$

The integral Int is bounded by

$$Int \le e^{-(\alpha+\sigma)t} \int_0^t e^{\sigma s}\, ds \le \frac{1}{\sigma}\, e^{-\alpha t} \ .$$

Therefore,

$$|(H^I y)(t)| \le K|a|e^{-\alpha t} + \frac{\varepsilon K}{\sigma}\, 2K|a|e^{-\alpha t} \ .$$

Since

$$\frac{\varepsilon K}{\sigma} \le \frac{1}{4}$$

obtain that

$$|(H^I y)(t)| \le \frac{3}{2}\, K|a|e^{-\alpha t} \quad \text{for} \quad t \ge 0 \ .$$

Also,

$$|(H^{II} y)(t)| \le \int_t^\infty K e^{-\sigma(s-t)} \varepsilon 2K|a|e^{-\alpha s}\, ds \ .$$

Using that

$$2\varepsilon K \le \frac{\sigma}{2}$$

obtain that

$$|(H^{II} y)(t)) \le K|a|\frac{\sigma}{2}\, e^{\sigma t} \int_t^\infty e^{-(\alpha+\sigma)s}\, ds \ .$$

The integral equals

$$Int = \frac{1}{\alpha+\sigma}\, e^{-(\alpha+\sigma)t} \ .$$

One obtains that

$$|(H^{II} y)(t)) \le \frac{1}{2}\, K|a|e^{-\alpha t}$$

and

$$|(Hy)(t)| \le |(H^I y)(t)| + |(H^{II} y)(t)| \le 2K|a|\, e^{-\alpha t} \quad \text{for} \quad t \ge 0 \ .$$

$\diamond$

Lemma 13.2 *Let $y^{(0)} \equiv 0$ and define the sequence $y^{(j)} \in Y$ by*

$$y^{(j+1)} = Hy^{(j)} \quad \textit{for} \quad j = 0, 1, 2, \ldots$$

Then the estimate

$$\left| y^{(j)}(t) - y^{(j-1)}(t) \right| \le \frac{K|a|}{2^{j-1}}\, e^{-\alpha t} \quad \textit{for} \quad t \ge 0 \tag{13.10}$$

holds for $j = 1, 2, \ldots$

We use induction in j. For $j = 1$ we have

$$|y^{(1)}(t)| = |e^{Pt} a| \le K|a|e^{-\alpha t} \ .$$

Assume the estimate (13.10) holds for some $j \ge 1$. Then one obtains that

$$\left| \left(y^{(j+1)} - y^{(j)} \right)^I (t) \right| \le K \int_0^t e^{-(\alpha+\sigma)(t-s)} \left| G(y^{(j)}(s)) - G(y^{(j-1)}(s)) \right| ds \ .$$

Here, using the induction assumption,

$$\left| G(y^{(j)}(s)) - G(y^{(j-1)}(s)) \right| \le \varepsilon\, \frac{K}{2^{j-1}}\, |a| e^{-\alpha s} \ .$$

Therefore,

$$\begin{aligned}
\left| \left(y^{(j+1)} - y^{(j)} \right)^I (t) \right| &\le K|a|\, \frac{\varepsilon K}{2^{j-1}}\, e^{-(\alpha+\sigma)t} \int_0^t e^{\sigma s}\, ds \\
&\le K|a|\, \frac{\varepsilon K}{\sigma} \cdot \frac{1}{2^{j-1}}\, e^{-\alpha t} \\
&\le K|a|\, \frac{1}{4} \cdot \frac{1}{2^{j-1}}\, e^{-\alpha t}
\end{aligned}$$

Also,

$$\begin{aligned}
\left| \left(y^{(j+1)} - y^{(j)} \right)^{II} (t) \right| &\le K\, \varepsilon\, \frac{K}{2^{j-1}}\, |a| \int_t^\infty e^{-\sigma(s-t)} e^{-\alpha s}\, ds \\
&= K|a| \frac{\varepsilon K}{2^{j-1}} e^{\sigma t} \int_t^\infty e^{-(\alpha+\sigma)s}\, ds \\
&\le K|a|\, \frac{\varepsilon K}{\sigma} \cdot \frac{1}{2^{j-1}}\, e^{-\alpha t}
\end{aligned}$$

Using that

$$\frac{\varepsilon K}{\sigma} \le \frac{1}{4}$$

one obtains that

$$\left| y^{(j+1)}(t) - y^{(j)}(t) \right| \le K|a| \, \frac{1}{2} \cdot \frac{1}{2^{j-1}} \, e^{-\alpha t} \quad \text{for} \quad t \ge 0 \; .$$

This completes the proof of (13.10) by induction. $\diamond$

Lemma 13.3 *Let $a \in \mathbb{R}^k$ and assume that $|a|$ is so small that $2K|a| \le \delta$. Then the integral system (13.4), (13.5) has a unique solution $y \in C[0,\infty)$ satisfying $|y(t)| \le \delta$ for $0 \le t < \infty$. This solution $y(t)$ satisfies the decay estimate*

$$|y(t)| \le 2K|a|e^{-\alpha t} \quad \textit{for} \quad 0 \le t < \infty \; . \tag{13.11}$$

Proof: Let $y^{(j)} \in Y = Y(a)$ denote the sequence determined in the previous lemma. The estimate

$$\left| y^{(j)}(t) - y^{(j-1)}(t) \right| \le \frac{K|a|}{2^{j-1}} \, e^{-\alpha t} \quad \text{for} \quad t \ge 0$$

implies that, for every fixed $t \ge 0$, the sequence $y^{(j)}(t)$ is a Cauchy sequence in $\mathbb{R}^n$. A standard convergence argument implies convergence,

$$y^{(j)}(y) \to y(t) \quad \text{for all} \quad t \ge 0$$

and $y \in Y$. The estimate (13.11) holds for every function $y^{(j)}$, therefore also for the limit function y.

Suppose that $v \in C[0,\infty)$ is a second solution of the integral system (13.4), (13.5) with $|v(t)| \le \delta$ for all $t \ge 0$. For $q = y - v$ obtain the estimates

$$\begin{aligned}
|q^I(t)| &\le K \int_0^t e^{-(\alpha+\sigma)(t-s)} \varepsilon |y(s) - v(s)| \, ds \\
&\le \varepsilon K |q|_\infty e^{-(\alpha+\sigma)t} \int_0^t e^{(\alpha+\sigma)s} \, ds \\
&\le \frac{\varepsilon K}{\sigma} \, |q|_\infty \\
&\le \frac{1}{4} |q|_\infty
\end{aligned}$$

and

$$\begin{aligned}
|q^{II}(t)| &\le K \int_t^\infty e^{-\sigma(s-t)} \varepsilon |q(s)| \, ds \\
&\le \varepsilon K |q|_\infty e^{\sigma t} \int_t^\infty e^{-\sigma s} \, ds \\
&= \frac{\varepsilon K}{\sigma} \, |q|_\infty \\
&\le \frac{1}{4} |q|_\infty
\end{aligned}$$

The estimate $|q(t)| \le \frac{1}{2}\,|q|_\infty$ implies that $q \equiv 0$, i.e., $v = y$. $\diamond$

We continue the proof of the Stable Manifold Theorem. For $a \in \mathbb{R}^k$ with $|a| < \frac{\delta}{2K}$ let $y(t) = y(t,a)$ denote the solution of the integral system (13.4), (13.5) satisfying

$$|y(t)| \le 2K|a|\, e^{-\alpha t} \quad \text{for} \quad t \ge 0 \; . \tag{13.12}$$

Obtain that

$$\begin{aligned} y^I(t) &= Py^I(t) + G^I(y(t)), \quad y^I(0) = a \\ y^{II}(t) &= Qy^{II}(t) + G^{II}y(t)), \quad y^{II}(\infty) = 0 \end{aligned}$$

Set

$$\psi(a) := y^{II}(0,a) \in \mathbb{R}^{n-k} \quad \text{for} \quad a \in \mathbb{R}^k, \quad |a| < \frac{\delta}{2K} \; .$$

The solution $y(t) = y(t,a)$ of the integral system (13.4), (13.5) solves the IVP

$$y' = By + G(y), \quad y(0) = \begin{pmatrix} a \\ \psi(a) \end{pmatrix} \; . \tag{13.13}$$

Since $y(t) \to 0$ as $t \to \infty$ it follows that the solution of the IVP (13.13) converges to 0 as $t \to \infty$.

Define the manifold

$$\mathcal{S} = \left\{ \begin{pmatrix} a \\ \psi(a) \end{pmatrix} \; : \; a \in \mathbb{R}^k, \quad |a| < \frac{\delta}{2K} \right\} \; .$$

Then, if $y(t)$ is a solution of the system $y' = By + G(y)$ with initial date on $\mathcal{S}$, then $y(t) \to 0$ as $t \to \infty$.

We claim: The manifold $\mathcal{S}$ is invariant under the flow of the system $y' = By + G(y)$ in forward time. In other words, if $y(t)$ is the solution of (13.13) for $t \ge 0$ then $y(t) \in \mathcal{S}$ for all $t \ge 0$. Fix $t_0 > 0$. We must show that

$$\psi(y^I(t_0)) = y^{II}(t_0) \; .$$

Set $a_0 = y^I(t_0)$. To determine $\psi(a_0)$ we must solve the integral system

$$\begin{aligned} u^I(t) &= e^{Pt}a_0 + \int_0^t e^{P(t-s)} G^I(u(s))\, ds \\ u^{II}(t) &= -\int_t^\infty e^{Q(t-s)} G^{II}(u(s))\, ds \end{aligned}$$

for $u \in C[0,\infty)$ satisfying $|u(t)| \le \delta$ for $t \ge 0$. If $u(t)$ is the solution, then $\psi(a_0) = u^{II}(0)$.

However, the function $y(t+t_0)$ solves the integral system and the required estimate. Uniqueness implies that

$$u(t) = y(t + t_0) \quad \text{for} \quad t \geq 0 \ .$$

Therefore,

$$\psi(y^I(t_0)) = \psi(a_0) = u^{II}(0) = y^{II}(t_0) \ .$$

This proves that the manifold $\mathcal{S}$ is invariant under the flow of the ODE system $y' = By + G(y)$ in forward time.

Lemma 13.4 *The stable manifold $\mathcal{S}$ has the tangent space E^s at the point 0.*

Proof: First, if $a = 0$ in (13.4), (13.5) then $y \equiv 0$ solves (13.4), (13.5), thus $\psi(0) = 0$. This implies that the point $0 \in \mathbb{R}^n$ lies on $\mathcal{S}$.

Second, we must show that the Jacobian of $\psi(a)$ at $a = 0$ is zero, $D\psi(0) = 0$. The proof of Lemma 13.3 shows that the functions $y^{(j)}(t)$ satisfy the estimates

$$\Big|(y^{(j+1)} - y^{(j)})^{II}(0)\Big| \leq \varepsilon \, \frac{K^2|a|}{2^{j-1}\sigma} \quad \text{for} \quad j = 0, 1, \ldots$$

and $y^{(0)} \equiv 0$. Therefore,

$$\Big|(y^{(j+1)})^{II}(0)| \leq C\varepsilon|a| \quad \text{for all} \quad j$$

where C is a constant and $\varepsilon > 0$ can be chosen arbitrarily small. For $j \to \infty$ obtain that

$$|\psi(a)| = |y^{II}(0)| \leq C\varepsilon|a|$$

where $\varepsilon > 0$ can be chosen arbitrarily small and $a \in \mathbb{R}^k$ satisfies $|a| \leq \delta(\varepsilon)$ for some $\delta(\varepsilon) > 0$. This estimate and $\psi(0) = 0$ imply that $D\psi(0)$ exists and $D\psi(0) = 0$.

We have shown in Section 5.5 that the solution of an IVP can be differentiated w.r.t. the initial data. With similar techniques, one can show that the function $\psi(a)$ has a derivative w.r.t. $a \in \mathbb{R}^k$, i.e., the manifold $\mathcal{S}$ has tangent spaces. In particular, since $D\psi(0) = 0$, the tangent space to $\mathcal{S}$ at 0 is E^s, the stable subspace of the linear system $y' = Ay$.

14 Auxiliary Results I

14.1 Notations

For $u, v \in \mathbb{C}^n$ let

$$\langle u, v \rangle = \sum_{j=1}^{n} \bar{u}_j v_j$$

denote the Euclidean inner product and let

$$|u| = \sqrt{\langle u, u \rangle}$$

denote the Euclidean norm. Recall the Cauchy–Schwarz inequality

$$|\langle u, v \rangle| \le |u||v| \quad \text{for all} \quad u, v \in \mathbb{C}^n \ .$$

If $A \in \mathbb{C}^{n \times n}$ then

$$\begin{aligned} |A| &= \min\{C \ge 0 \ : \ |Ax| \le C|x| \quad \text{for all} \quad x \in \mathbb{C}\} \\ &= \max\{|Ax| \ : \ x \in \mathbb{C}^n \quad \text{with} \quad |x| = 1\} \end{aligned}$$

denotes the matrix norm of A corresponding to the Euclidean vector norm.

14.2 Estimates of Functions

Theorem 14.1 *Let $q : [0, T] \to [0, \infty)$ denote a C^1–function and let $K \ge 0$. If*

$$q'(t) \le Kq(t) \quad \textit{for} \quad 0 \le t \le T$$

then

$$q(t) \le q(0)e^{Kt} \quad \textit{for} \quad 0 \le t \le T \ . \tag{14.1}$$

Proof: Let $\varepsilon > 0$. We claim that

$$q(t) < (q(0) + \varepsilon)e^{Kt} \quad \text{for} \quad 0 \le t \le T \ . \tag{14.2}$$

If (14.2) does not hold then there exists $0 < t_0 \le T$ with

$$q(t) < (q(0) + \varepsilon)e^{Kt} \quad \text{for} \quad 0 \le t < t_0 \quad \text{and} \quad q(t_0) = (q(0) + \varepsilon)e^{Kt_0} \ . \tag{14.3}$$

Set

$$h(t) = (q(0) + \varepsilon)e^{Kt} \quad \text{for} \quad 0 \le t \le T \ .$$

We have $q(0) < h(0)$ and

$$q'(t) \le Kq(t) \le Kh(t) = h'(t) \quad \text{for} \quad 0 \le t \le t_0 \ .$$

This implies that $q(t_0) < h(t_0)$, a contradiction to (14.3). Therefore, the inequality (14.2) holds for all $\varepsilon > 0$. As $\varepsilon \to 0$ the estimate (14.1) follows. $\diamond$

Theorem 14.2 *Let $A(t) \in \mathbb{R}^{n\times n}$ denote a continuous matrix function defined for $0 \le t \le T$. Assume that $|A(t)| \le L$ for $0 \le t \le T$. Let $w : [0,T] \to \mathbb{R}^n, w \in C^1[0,T]$. If*

$$w'(t) = A(t)w(t) \quad \textit{for} \quad 0 \le t \le T$$

then

$$|w(t)| \le |w(0)|\, e^{Lt} \quad \textit{for} \quad 0 \le t \le T\ . \tag{14.4}$$

Proof: For $q(t) = |w(t)|^2$ we have

$$\begin{aligned} q'(t) &= 2\langle w(t), w'(t)\rangle \\ &= 2\langle w(t), A(t)w(t)\rangle \\ &\le 2Lq(t) \end{aligned}$$

By Theorem 14.1 the estimate

$$|w(t)|^2 \le |w(0)|^2 e^{2Lt} \quad \text{for} \quad 0 \le t \le T$$

follows. $\diamond$

14.3 The Mean–Value Theorem

Let $h : [a,b] \to \mathbb{R}$ denote a C^1–function. We have

$$\begin{aligned} h(b) - h(a) &= \int_a^b h'(x)\, dx \quad (\text{substitute } x = a + t(b-a), dx = (b-a)dt) \\ &= \Big(\int_0^1 h'(a + t(b-a))\, dt\Big)(b-a) \\ &= h'(\xi)(b-a) \end{aligned}$$

for some $a \le \xi \le b$.

The result

$$\frac{h(b) - h(a)}{b-a} = h'(\xi) \quad \text{for some } a \le \xi \le b \tag{14.5}$$

is often called the **Mean Value Theorem**.

If the function h takes values in $\mathbb{R}^k, k \ge 2$, then (14.5) holds for every component $h_k(x)$, but $\xi = \xi_k$. The equation (14.5) generally does not hold if h is vector–valued.

Theorem 14.3 *(Mean Value Theorem) Let $h : \mathbb{R}^n \to \mathbb{R}^k$ denote a C^1–function. Then for all $a, b \in \mathbb{R}^n$ we have*

$$h(b) - h(a) = \Big(\int_0^1 Dh(a + t(b-a))\, dt\Big)(b-a)$$

where $Dh(x) \in \mathbb{R}^{k\times n}$ is the Jacobian of $h(x)$ at $x \in \mathbb{R}^n$.

Proof: Set $q(t) = h(a + t(b - a))$ for $0 \le t \le 1$. We have

$$\begin{aligned} h(b) - h(a) &= q(1) - q(0) \\ &= \int_0^1 q'(t)\, dt \\ &= \Big(\int_0^1 Dh(a + t(b-a))\, dt \Big) (b - a) \end{aligned}$$

$\diamond$

14.4 Matrices and Norms

Let $\| \cdot \|$ denote a vector norm on $\mathbb{C}^n$ and let $A \in \mathbb{C}^{n \times n}$. The matrix norm corresponding to the vector norm $\| \cdot \|$ is

$$\begin{aligned} \|A\| &= \min\{C \ge 0 \ : \ \|Ax\| \le C\|x\| \quad \text{for all} \quad x \in \mathbb{C}^n\} \\ &= \max\{\|Ax\| \ : \ x \in \mathbb{C}^n, \|x\| = 1\} \end{aligned}$$

For example, let

$$|x|_\infty = \max\{|x_j| \ : \ j = 1, 2, \ldots, n\}$$

denote the maximum norm on $\mathbb{C}^n$. The corresponding matrix norm is the maximal row sum of absolute values,

$$|A|_\infty = \max \Big\{ \sum_{j=1}^n |a_{ij}| \ : \ 1 \le i \le n \Big\} \ .$$

Let

$$\langle x, y \rangle = x^* y = \sum_{j=1}^n \bar{x}_j y_j$$

denote the Euclidean inner product on $\mathbb{C}^n$ and let

$$|x| = \sqrt{x^* x} = \sqrt{\langle x, y \rangle}$$

denote the Euclidean norm.

The corresponding matrix norm is

$$|A| = \max\{\sigma \ : \ \sigma \ \text{ is a singular value of } \ A\}$$

The singular values of A are the non–negative square roots of the eigenvalues of $A^* A$.

If $T \in \mathbb{C}^{n \times n}$ is a nonsingular matrix define the vector norm $| \cdot |_T$ by

$$|x|_T = |T^{-1} x| = |y| \quad \text{where} \quad x = Ty \ .$$

The corresponding matrix norm of $A \in \mathbb{C}^{n \times n}$ is

$$|A|_T = \min\{C \ge 0 \ : \ |Ax|_T \le C|x|_T \quad \text{for all} \quad x \in \mathbb{C}^n\} \ .$$

Lemma 14.1 *Let $A \in \mathbb{C}^{n\times n}$ and let $T \in \mathbb{C}^{n\times n}$ be nonsingular. Then*

$$|A|_T = |\tilde{A}| \quad \textit{where} \quad \tilde{A} = T^{-1}AT \ .$$

Proof: We have for all $x \in \mathbb{C}^n$:

$$\begin{aligned} |Ax|_T &= |T^{-1}Ax| \\ &= |\tilde{A}T^{-1}x| \\ &\le |\tilde{A}||T^{-1}x| \\ &= |\tilde{A}||x|_T \end{aligned}$$

This proves that $|A|_T \le |\tilde{A}|$.

Also, there exists $y \in \mathbb{C}^n, y \ne 0$, with $|\tilde{A}y| = |\tilde{A}||y|$. Using $x = Ty$ in the above estimate one obtains that $|Ax|_T = |\tilde{A}||x|_T$. Therefore, $|A|_T = |\tilde{A}|$. $\diamond$

14.5 Similarity Transformations

Consider a linear system of ODEs, $x' = Ax$, where $x \in \mathbb{C}^n$ and $A \in \mathbb{C}^{n\times n}$. Let $T \in \mathbb{C}^{n\times n}$ denote a nonsingular matrix and introduce new variables $y = y(t)$ by

$$x(t) = Ty(t) \ .$$

One obtains that

$$y' = By \quad \text{where} \quad B = T^{-1}AT \ .$$

The transformation

$$A \to T^{-1}AT$$

is called a similarity transformation of A. One often tries to find a transformation matrix T for which the transformed matrix $B = T^{-1}AT$ is *simpler* than the matrix A.

The following theorem, called **Schur's Theorem**, is often useful.

Theorem 14.4 *Let $A \in \mathbb{C}^{n\times n}$. There exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ so that U^*AU is upper triangular,*

$$U^*AU = \Lambda + R \ ,$$

where Λ is diagonal and R is strictly upper triangular. The eigenvalues λ_j of A are the diagonal elements of Λ. They can occur in any prescribed order.

Proof: Use induction in n. The case $n=1$ is trivial. Let λ_1 be an eigenvalue of A and let $u^1 \in \mathbb{C}^n$,

$$Au^1 = \lambda_1 u^1, \quad |u^1| = 1 \; .$$

Let $u^1, u^2, \ldots, u^n$ denote an othornomal basis of $\mathbb{C}^n$ and let $U_1 \in \mathbb{C}^{n\times n}$ denote the unitary matrix with columns $u^1, u^2, \ldots, u^n$. Then we have

$$U_1^* A U_1 = U_1^*(\lambda_1 u^1 | v^2 | \ldots | v^n) = \begin{pmatrix} \lambda_1 & * & * & * \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix} \quad \text{where} \quad A_2 \in \mathbb{C}^{(n-1)\times(n-1)} \; .$$

By the induction hypothesis, there exists a unitary matrix $U_2 \in \mathbb{C}^{(n-1)\times(n-1)}$ so that $U_2^* A_2 U_2$ is upper triangular. Setting

$$U = U_1 \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & U_2 & \\ 0 & & & \end{pmatrix}$$

one obtains that U is unitary and $U^* A U$ is upper triangular. $\diamond$

Theorem 14.5 *Let $A \in \mathbb{C}^{n\times n}$ have the spectral radius*

$$\rho(A) = \max\{|\lambda| \; : \; \lambda \textit{ is an eigenvalue of } A\}$$

and let $q > \rho(A)$. There exists a nonsingular matrix $T \in \mathbb{C}^{n\times n}$ with

$$\rho(A) \le |A|_T \le q \; .$$

Proof: By Schur's Theorem there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ so that $U^* A U$ is upper triangular,

$$U^* A U = \Lambda + R \; ,$$

where Λ is diagonal and R is strictly upper triangular. Let S_ε denote the diagonal matrix

$$S_\varepsilon = \begin{pmatrix} 1 & & & \\ & \varepsilon & & \\ & & \ddots & \\ & & & \varepsilon^{n-1} \end{pmatrix}, \quad \varepsilon > 0 \; ,$$

and let $T_\varepsilon = U S_\varepsilon$. We have

$$T_\varepsilon^{-1} A T_\varepsilon = S_\varepsilon^{-1} U^* A U S_\varepsilon = \Lambda + S_\varepsilon^{-1} R S_\varepsilon \; .$$

Here

$$(S_\varepsilon^{-1} R S_\varepsilon)_{ij} = \varepsilon^{j-i} r_{ij} = \mathcal{O}(\varepsilon) \quad \text{for} \quad i < j \ .$$

It follows that

$$|T_\varepsilon^{-1} A T_\varepsilon| \le \rho(A) + \mathcal{O}(\varepsilon) \ .$$

Choosing $\varepsilon > 0$ small enough one obtains that

$$|A|_{T_\varepsilon} = |T_\varepsilon^{-1} A T_\varepsilon| = \rho(A)| + \mathcal{O}(\varepsilon) \le q \ .$$

The lower bound $\rho(A) \le \|A\|$ holds for every matrix norm corresponding to a vector norm. $\diamond$

Lemma 14.2 *Every matrix $A \in \mathbb{C}^{n\times n}$ is similar to its transpose, A^T; i.e., there exists a nonsingular matrix S so that $S^{-1}AS = A^T$.*

Proof: We first prove that the matrix

$$J_k = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix} \in \mathbb{R}^{k\times k}$$

is similar to J_k^T. Take $k = 5$, for example. Let σ denote the permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}$$

and let P denote the corresponding permutation matrix:

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \ .$$

When one forms PJ_5, the permutation σ is applied to the rows of J_5:

$$PJ_5 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \ .$$

When one forms PJ_5P, the permutation σ is applied to the columns of PJ_5:

$$PJ_5P = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} .$$

Clearly, $PJ_5P = J_5^T$. Since $P^2 = I$ we have $P = P^{-1}$, thus $P^{-1}J_5P = J_5^T$. This proves that $J_5 \sim J_5^T$.

We now use the Jordan form theorem. Given any matrix $A \in \mathbb{C}^{n\times n}$ there exists a nonsingular matrix $R \in \mathbb{C}^{n\times n}$ so that

$$R^{-1}AR = \Lambda + J$$

where $\Lambda + J$ is a block–diagonal matrix with the following property: Each diagonal block has the form

$$\lambda I_k \quad \text{or} \quad \lambda I_k + J_k$$

where λ is an eigenvalue of A and J_k is defined above. Since $J_k^T \sim J_k$ it follows that $(\Lambda + J)^T \sim \Lambda + J$. Obtain that

$$A \sim \Lambda + J \sim \Lambda + J^T .$$

Thus, there exists $Q \in \mathbb{C}^{n\times n}$ with $Q^{-1}AQ = \Lambda + J^T$. This yields that $Q^TA^T(Q^{-1})^T = \Lambda + J \sim A$. The theorem is proved. $\diamond$

14.6 Stability of a Fixed Point of a Map

Let $F : \mathbb{R}^N \to \mathbb{R}^N$ denote a map. Given $x_0 \in \mathbb{R}^N$ let the sequence $x_n \in \mathbb{R}^N$ be defined by $x_n = F^n(x_0)$, i.e., the sequence x_n consists of the points

$$x_0, \quad x_1 = F(x_0), \quad x_2 = F(x_1), \quad x_3 = F(x_2), \quad \ldots$$

The points x_n are the points of a discrete–time evolution in the state space $\mathbb{R}^N$ starting at x_0 at time $t = 0$.

A point $x^* \in \mathbb{R}^N$ is a fixed point of the evolution determined by F if $F(x^*) = x^*$.

Definition: a) The fixed point x^* is stable if for every $\varepsilon > 0$ there exists $\delta > 0$ so that

$$|x_0 - x^*| \le \delta \quad \text{implies} \quad |x_n - x^*| \le \varepsilon \quad \text{for all} \quad n \in \mathbb{N} .$$

b) The fixed point x^* is asymptotically stable if x^* is stable and, in addition, there exists $\delta > 0$ so that

$$|x_0 - x^*| \le \delta \quad \text{implies} \quad x_n \to x^* \quad \text{as} \quad n \to \infty .$$

Theorem 14.6 *Let $F : \mathbb{R}^N \to \mathbb{R}^N$ denote a C^1–map and let $x^* \in \mathbb{R}^N$ denote a fixed point of F, i.e., $F(x^*) = x^*$. Let $F'(x^*) \in \mathbb{R}^{N\times N}$ denote the Jacobian of F at the fixed point x^*. If $\rho(F'(x^*)) < 1$ then x^* is asymptotically stable.*

Proof: Let $x^* = 0$ for simplicity of notation and let $A = F'(x^*)$. Fix a number q with

$$\rho(A) < q < 1 \ .$$

By Theorem 14.5 there exists a norm $\|\cdot\|$ on $\mathbb{R}^N$ so that

$$\|A\| \le q < 1 \ .$$

First assume that $F \in C^2$. We have

$$F(x) = Ax + Q(x) \quad \text{for} \quad x \in \mathbb{R}^N$$

where

$$\|Q(x)\| \le C\|x\|^2 \quad \text{for} \quad \|x\| \le 1 \ .$$

Therefore,

$$\|F(x)\| \le \Big(q + C\|x\|\Big)\|x\| \quad \text{for} \quad \|x\| \le 1 \ .$$

Choose $\delta > 0$ so that $\tilde{q} := q + C\delta < 1$ and obtain that

$$\|F(x)\| \le \tilde{q}\|x\| \quad \text{for} \quad \|x\| \le \delta \ .$$

Since $0 \le \tilde{q} < 1$ asymptotic stability of the fixed point $x^* = 0$ follows.

If $F \in C^1$ then

$$F(x) = Ax + Q(x) \quad \text{for} \quad x \in \mathbb{R}^N$$

where

$$\|Q(x)\| = o(\|x\|) \quad \text{for} \quad \|x\| \le 1 \ .$$

This means that

$$\lim_{x\to 0} \frac{\|Q(x)\|}{\|x\|} = 0 \ .$$

Obtain that

$$\|F(x)\| \le \Big(q + \frac{\|Q(x)\|}{\|x\|}\Big)\|x\| \quad \text{for} \quad 0 < \|x\| \le 1 \ .$$

Choose $\tilde{q}$ with

$$\rho(A) < q < \tilde{q} < 1 \ .$$

There exists $\delta > 0$ so that

$$\frac{\|Q(x)\|}{\|x\|} \le \tilde{q} - q \quad \text{for} \quad 0 < \|x\| \le \delta \ .$$

Then, for $\|x\| \le \delta$,

$$\|F(x)\| \le \tilde{q}\|x\| \quad \text{where} \quad 0 \le \tilde{q} < 1 \ ,$$

and asymptotic stability of the fixed point $x^* = 0$ follows. ⋄

14.7 Symplectic Matrices

The Lie group of symplectic matrices plays a role in the theory of Hamiltonian systems.

Let

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \in \mathbb{R}^{2n \times 2n}$$

where $I = I_n$ is the $n \times n$ identity matrix. Note that $J^2 = -I_{2n}$, thus $J^{-1} = -J$. Exchanging column j with column $n+j$ for $j = 1, \ldots, n$ in J one obtains the matrix

$$\tilde{J} = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

Therefore,

$$det(J) = (-1)^n det(\tilde{J}) = 1 \ .$$

Definition: A matrix $S \in \mathbb{R}^{2n \times 2n}$ is called symplectic if

$$S^T J S = J \ .$$

The Lie group of all symplectic matrices of dimension $2n \times 2n$ is denote by $Sp(n)$.

The matrix J defines the bilinear form on $\mathbb{R}^{2n}$:

$$\langle x, Jy \rangle \ .$$

A matrix $S \in \mathbb{R}^{2n \times 2n}$ is symplectic if and only if S leaves this bilinear form invariant. I.e., if one applies S to x and y one obtains

$$\langle Sx, JSy \rangle = \langle x, Jy \rangle \quad \text{for all} \quad x, y \in \mathbb{R}^{2n}$$

if and only if S is symplectic.

Theorem 14.7 *1) The symplectic $2n \times 2n$ matrices form a group under matrix multiplication.*

2) If S is symplectic, then S^T is also symplectic.

3) If S is symplectic, then S is similar to S^{-1}.

Proof: 1) It is clear that $I_{2n} \in Sp(n)$. Also, if $S_1, S_2 \in Sp(n)$, then $S_1 S_2 \in Sp(n)$. The equation $S^T J S = J$ implies that $det(S^2) = 1$, thus S is nonsingular and

$$J = (S^{-1})^T J S^{-1} \ .$$

Therefore, $S^{-1} \in Sp(n)$.

2) Taking the inverse of the equation $S^T J S = J$ and using that $J^{-1} = -J$ one obtains that

$$S^{-1} J (S^{-1})^T = J \ ,$$

thus $(S^{-1})^T = (S^T)^{-1}$ is symplectic. Using 1) one obtains that S^T is symplectic.

3) From $S^T J S = J$ obtain that

$$J S J^{-1} = (S^T)^{-1} \ .$$

Thus, S is similar to $(S^T)^{-1} = (S^{-1})^T$. For all square matrices A we have that A is similar to A^T. Therefore,

$$S \sim (S^{-1})^T \sim S^{-1} \ .$$

$\diamond$

14.8 Real and Complex Jordan Normal Form

14.8.1 Complex Jordan Normal Form

For $k = 2, 3, \ldots$ let I_k denote the $k \times k$ identity matrix and let

$$J_k = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}$$

Let $A \in \mathbb{C}^{n \times n}$. There exists a nonsingular matrix $W \in \mathbb{C}^{n \times n}$ so that $W^{-1} A W$ is block–diagonal,

$$W^{-1} A W = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_l \end{pmatrix} \qquad (14.6)$$

where B_1 is a diagonal matrix with eigenvalues of A on the diagonal and each matrix $B_2, \ldots, B_l$ has the form

$$B_j = \lambda I_{k_j} + J_{k_j} \qquad (14.7)$$

where λ is an eigenvalue of A. The block–diagonal matrix in (14.6) is called the complex Jordan Normal Form of A. Note, however, that a diagonal matrix

B_1 may not be present. Also, if A can be diagonalized, then matrices of the form (14.7) do not occur in the complex Jordan Normal Form of A. A matrix of the form (14.7) occurs in the Jordan Normal Form (14.6) of A if and only if the algebraic multiplicity of the eigenvalue λ of A is strictly larger than its geometric multiplicity.

The theorem that every matrix $A \in \mathbb{C}^{n\times n}$ can be transformed to Jordan Normal Form can be proved in three steps:

Step 1: By Schur's Theorem there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ so that $U^*AU = \Lambda + R$ is upper triangular.

Step 2: Blocking Lemma: Let A be a block matrix of the form

$$A = \begin{pmatrix} M_1 & M_{12} \\ 0 & M_2 \end{pmatrix}$$

with square matrices M_1 and M_2, which have no common eigenvalue, i.e., $\sigma(M_1) \cap \sigma(M_2) = \emptyset$. Then there exists a transformation matrix T of the form

$$T = \begin{pmatrix} I & S \\ 0 & I \end{pmatrix}$$

so that

$$T^{-1}AT = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} .$$

Thus, the transformation $T^{-1}AT$ eliminates the coupling block M_{12}.

After using the Blocking Lemma repeatedly, one only has to consider matrices which have only one eigenvalue λ. Since $T^{-1}(\lambda I)T = \lambda I$ one may assume that $\lambda = 0$. A matrix which has $\lambda = 0$ as its only eigenvalue is nilpotent.

Step 3: Prove that every nilpotent matrix can be transformed to Jordan Normal Form.

14.8.2 Real Jordan Normal Form

We first treat two simple cases. Let $A \in \mathbb{R}^{2\times 2}$ have the eigenvalue

$$\lambda = \alpha + i\beta \quad \text{where} \quad \alpha, \beta \in \mathbb{R}, \quad \beta > 0 .$$

There exists

$$w = a + ib \in \mathbb{C}^2 \quad \text{with} \quad a, b \in \mathbb{R}^2, \quad w \neq 0 ,$$

with

$$Aw = \lambda w, \quad A\bar{w} = \bar{\lambda}\bar{w} .$$

Since w and $\bar{w}$ are linearly independent, it follows that a and b are linearly independent. One obtains that

$$A(a + ib) = (\alpha + i\beta)(a + ib) = \alpha a - \beta b + i(\beta a + \alpha b) ,$$

thus

$$Aa = \alpha a - \beta b \quad \text{and} \quad Ab = \beta a + \alpha b \; .$$

We set

$$T = (b|a) \in \mathbb{R}^{2\times 2} \; ,$$

i.e., T has the columns b and a. Obtain

$$A(b|a) = (\beta a + \alpha b|\alpha a - \beta b) = (b|a) \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \; ,$$

thus

$$T^{-1}AT = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \; .$$

This is the transformation of A to its Real Jordan Normal Form.

Second, assume that $A \in \mathbb{R}^{4\times 4}$ has the complex eigenvalue $\lambda = \alpha + i\beta$ of geometric multiplicity one and algebraic multiplicity two. We first recall the transformation of A to its complex Jordan Normal Form. There exist linearly independent vectors $w^1, w^2 \in \mathbb{C}^4$ so that

$$\begin{aligned} Aw^1 &= \lambda w^1 \\ Aw^2 &= \lambda w^2 + w^1 \end{aligned}$$

If one sets

$$W = (w^1|w^2|\bar{w}^1|\bar{w}^2) \in \mathbb{C}^{4\times 4}$$

then one obtains $AW = WB$ with

$$B = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \bar{\lambda} & 1 \\ 0 & 0 & 0 & \bar{\lambda} \end{pmatrix}$$

Thus $W^{-1}AW = B$ is the transformation of A to complex Jordan Normal Form.

Let

$$w^1 = a^1 + ib^1, \quad w^2 = a^2 + ib^2$$

with $a^1, b^1, a^2, b^2 \in \mathbb{R}^4$. Since $Aw^1 = \lambda w^1$ obtain that

$$A(b^1|a^1) = (b^1|a^1) \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \; . \tag{14.8}$$

This follows in the same way as in the case of a real 2×2–matrix treated above. From

$$Aw^2 = \lambda w^2 + w^1$$

obtain that

$$
\begin{aligned}
A(a^2+ib^2) &= (\alpha+i\beta)(a^2+ib^2)+a^1+ib^1 \\
&= \alpha a^2-\beta b^2+a^1+i(\beta a^2+\alpha b^2+b^1)
\end{aligned}
$$

thus

$$
\begin{aligned}
Aa^2 &= \alpha a^2-\beta b^2+a^1 \\
Ab^2 &= \beta a^2+\alpha b^2+b^1
\end{aligned}
$$

In matrix form:

$$
A(b^2|a^2) = (b^1|a^1|b^2|a^2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \alpha & -\beta \\ \beta & \alpha \end{pmatrix} .
$$

Setting

$$
T = (b^1|a^1|b^2|a^2)
$$

and using (14.8) one obtains that

$$
AT = T \begin{pmatrix} \alpha & -\beta & 1 & 0 \\ \beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & -\beta \\ 0 & 0 & \beta & \alpha \end{pmatrix} .
$$

Thus

$$
T^{-1}AT = \begin{pmatrix} \alpha & -\beta & 1 & 0 \\ \beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & -\beta \\ 0 & 0 & \beta & \alpha \end{pmatrix}
$$

is the transformation of A to its real Jordan Normal Form.

General Case: Let $A \in \mathbb{R}^{n\times n}$. There exists a nonsingular matrix $T \in \mathbb{R}^{n\times n}$ so that

$$
T^{-1}AT = \begin{pmatrix} D_1 & & & 0 \\ & D_2 & & \\ & & \ddots & \\ 0 & & & D_r \end{pmatrix}
$$

where D_1 is a real diagonal matrix and where each matrix $D_2, \ldots, D_r$ either has the form

$$
D_j = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}
$$

where $\alpha + i\beta$ is a complex eigenvalue of A or has the form

$$D_j = \begin{pmatrix} B & I_2 & & 0 \\ 0 & B & \ddots & \\ & & \ddots & I_2 \\ 0 & & & B \end{pmatrix} \quad \text{with} \quad B = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} .$$

Again, $\alpha + i\beta$ is a complex eigenvalue of A.

14.9 Matrix Exponentials

We first consider the complex case. For $A \in \mathbb{C}^{n\times n}$ the IVP

$$z'(t) = Az(t), \quad z(0) = z_0 \ ,$$

has the unique solution

$$z(t) = e^{At} z_0$$

where

$$e^{At} = \sum_{j=0}^{\infty} \frac{1}{j!} (At)^j \ .$$

Often it is not easy to understand the behavior of the solution $z(t)$ using this formula. If

$$W^{-1}AW = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_l \end{pmatrix} =: B$$

transforms A to complex Jordan Normal Form then $A = WBW^{-1}$ and

$$e^{At} = We^{Bt}W^{-1} \ .$$

Here e^{Bt} is the block diagonal matrix with blocks $e^{B_j t}$. If

$$B_1 = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}$$

then

$$e^{B_1 t} = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_r t} \end{pmatrix} .$$

If

$$B_j = \lambda I_k + J_k, \quad k \geq 2 \ ,$$

then

$$\begin{aligned} e^{B_j t} &= e^{\lambda t} e^{J_k t} \\ &= e^{\lambda t} \Big(I_k + t J_k + \frac{t^2}{2} J_k^2 + \ldots + \frac{t^{k-1}}{(k-1)!} J^{k-1} \Big) \end{aligned}$$

thus

$$e^{B_j t} = e^{\lambda t} \left(\begin{array}{cccc} 1 & t & \ldots & t^{k-1}/(k-1)! \\ & \ddots & \ddots & \\ & & \ddots & t \\ 0 & & & 1 \end{array} \right) .$$

The term J_k in the complex Jordan Normal Form of A contributes a polynomial in t of degree $k-1$, multiplying the exponential $e^{\lambda t}$.

Definition: *An eigenvalue λ of A is called semi–simple if its algebraic multiplicity equals its geometric multiplicity.*

An eigenvalue λ of A is semi–simple if and only if a term

$$B_j = \lambda I_k + J_k, \quad k \geq 2 \ ,$$

does not occur in the complex Jordan Normal Form of A.

Theorem 14.8 *Consider the IVP*

$$z' = Az, \quad z(0) = z_0$$

where $A \in \mathbb{C}^{n \times n}$.

a) All solutions $z(t) = e^{At} z_0$ converge to zero as $t \to \infty$ if and only if all eigenvalues λ of A satisfy $\mathrm{Re}\, \lambda < 0$.

b) All solutions $z(t) = e^{At} z_0$ are bounded for $0 \leq t < \infty$ if and only if all eigenvalues λ of A satisfy $\mathrm{Re}\, \lambda \leq 0$ *and if* $\mathrm{Re}\, \lambda = 0$ *then λ is semi–simple.*

The Real Case: If $A \in \mathbb{R}^{n \times n}$ it is often good to express e^{At} completely in terms of real functions.

First let

$$A = \left(\begin{array}{cc} \alpha & -\beta \\ \beta & \alpha \end{array} \right)$$

where $\alpha, \beta \in \mathbb{R}$. We have

$$A = \alpha I + \beta Q \quad \text{with} \quad Q = \left(\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array} \right) .$$

Thus

$$e^{At} = e^{\alpha t} e^{\beta t Q}$$

where

$$e^{\beta t Q} = \sum_{j=0}^{\infty} \frac{(\beta t)^j}{j!} Q^j \ .$$

Since $Q^2 = -I_2$ one obtains that

$$e^{\beta t Q} = \Big(\sum_{j=0}^{\infty} \frac{(\beta t)^{2j}}{(2j)!} (-1)^j \Big) \ I_2 + \Big(\sum_{j=0}^{\infty} \frac{(\beta t)^{2j+1}}{(2j+1)!} (-1)^j \Big) \ Q \ ,$$

thus

$$e^{\beta t Q} = \left(\begin{array}{cc} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{array} \right) \ .$$

Therefore, if

$$A = \left(\begin{array}{cc} \alpha & -\beta \\ \beta & \alpha \end{array} \right)$$

then

$$e^{At} = e^{\alpha t} \left(\begin{array}{cc} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{array} \right) \ .$$

If $\beta t > 0$ then the mapping

$$\left(\begin{array}{c} x_1 \\ x_2 \end{array} \right) \to \left(\begin{array}{cc} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{array} \right) \left(\begin{array}{c} x_1 \\ x_2 \end{array} \right) \tag{14.9}$$

describes the rigid rotation of the x_1-x_2–plane by angle βt in counter–clockwise direction.

Second, consider the ODE system $x' = Ax$ where $A \in \mathbb{R}^{2\times 2}$ and

$$T^{-1} A T = \left(\begin{array}{cc} \alpha & -\beta \\ \beta & \alpha \end{array} \right) = \alpha I + \beta Q \ .$$

Then we have

$$e^{At} = e^{\alpha t} T e^{\beta t Q} T^{-1} \ .$$

If one introduces y–coordinates by $x = Ty$ then one obtains the rotation (14.9) in y–coordinates. The factor $e^{\alpha t}$ is not affected by the coordinate change.

Next consider the case where $A \in \mathbb{R}^{4\times 4}$ has the complex eigenvalue $\lambda = \alpha + i\beta, \beta > 0$, which is geometrically simple, but algebraically double. There exists a transformation matrix $T \in \mathbb{R}^{4\times 4}$ so that

$$T^{-1}AT = \begin{pmatrix} \alpha & -\beta & 1 & 0 \\ \beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & -\beta \\ 0 & 0 & \beta & \alpha \end{pmatrix} = \alpha I_4 + \begin{pmatrix} \beta Q & 0 \\ 0 & \beta Q \end{pmatrix} + \begin{pmatrix} 0 & I_2 \\ 0 & 0 \end{pmatrix} =: D \ .$$

It is easy to check that the matrices

$$\begin{pmatrix} \beta Q & 0 \\ 0 & \beta Q \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & I_2 \\ 0 & 0 \end{pmatrix}$$

commute under multiplication and

$$\begin{pmatrix} 0 & I_2 \\ 0 & 0 \end{pmatrix}^2 = 0 \ .$$

Therefore,

$$e^{Dt} = e^{\alpha t} \begin{pmatrix} e^{\beta t Q} & 0 \\ 0 & e^{\beta t Q} \end{pmatrix} \begin{pmatrix} I_2 & tI_2 \\ 0 & I_2 \end{pmatrix} = e^{\alpha t} \begin{pmatrix} e^{\beta t Q} & te^{\beta t Q} \\ 0 & e^{\beta t Q} \end{pmatrix} \ .$$

The assumption that the eigenvalue λ of A is geometrically simple, but algebraically double leads to the linear factor t multiplying $e^{\beta t Q}$ in the above formula for e^{Dt}.

14.10 Estimate of $|e^{At}|$

Theorem 14.9 *Let $A \in \mathbb{C}^{n\times n}$ and assume that*

$$\text{Re}\, \lambda_j < -\alpha < 0$$

for all eigenvalues λ_j of A. Then there exists a constant $K \geq 1$ so that

$$|e^{At}| \leq K e^{-\alpha t} \quad \textit{for} \quad t \geq 0 \ .$$

Proof: Let $W^{-1}AW = B$ denote the transformation of A to complex Jordan normal form. The matrix B is a block matrix with blocks of the form

$$B_j = \lambda I + J$$

where λ is an eigenvalue of A and J is nilpotent, $J^m = 0$.

One obtains that

$$e^{B_j t} = e^{(\lambda I + J)t} = e^{\lambda t} e^{Jt} \ .$$

Since $J^m = 0$ it follows that

$$|e^{Jt}| \leq C_1(1 + t^{m-1}) \quad \text{for} \quad t \geq 0 \ .$$

Since $\lambda < -\alpha$ the estimate

$$|e^{B_j t}| \leq C_2 e^{-\alpha t} \quad \text{for} \quad t \geq 0$$

follows. $\diamond$

15 Auxiliary Results, II

15.1 The Stable Subspace for a Linear System

We first consider the complex case. Let $A \in \mathbb{C}^{n\times n}$. Consider the IVP

$$z' = Az, \quad z(0) = z_0 \ ,$$

with solution

$$z(t) = e^{At} z_0 \in \mathbb{C}^n, \quad t \in \mathbb{R} \ .$$

Definition: *The complex stable subspace of the system $z' = Az$ is*

$$E^s_{\mathbb{C}} = \{ z_0 \in \mathbb{C}^n \ : \ e^{At} z_0 \to 0 \quad \text{as} \quad t \to \infty \} \ .$$

If $\lambda \in \mathbb{C}$ is an eigenvalue of A then

$$E_{gen}(\lambda) = \{ z \in \mathbb{C}^n \ : \ (A - \lambda I)^n z = 0 \}$$

is the generalized eigenspace of λ.

Theorem 15.1 *Let $\lambda_1, \ldots, \lambda_k$ denote the eigenvalues of A with negative real parts,*

$$\mathrm{Re}\,\lambda_j < 0 \quad for \quad j = 1, \ldots, k \ .$$

The complex stable subspace of A equals the sum of the generalized eigenspaces of A corresponding to eigenvalues with negative real parts,

$$E^s_{\mathbb{C}} = E_{gen}(\lambda_1) \oplus \cdots \oplus E_{gen}(\lambda_k) \ .$$

Proof: There exists $W \in \mathbb{C}^{n\times n}$ so that

$$W^{-1} A W = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_l \end{pmatrix} = B$$

is the transformation of A to complex Jordan Normal Form. Note that each block B_q is either diagonal or of the form

$$B_q = \lambda I_m + J_m \ .$$

Every $z_0 \in \mathbb{C}$ can be written as $z_0 = W\alpha$ for a unique $\alpha \in \mathbb{C}^n$. Note that

$$e^{At} z_0 = W e^{Bt} W^{-1} z_0 = W e^{Bt} \alpha \ ,$$

therefore $e^{At} z_0 \to 0$ as $t \to \infty$ if and only if $e^{Bt}\alpha \to 0$ as $t \to \infty$.

Let

$$W = (W^I | W^{II})$$

where the columns of W^I are generalized eigenvectors corresponding to the eigenvalues λ of A with $\operatorname{Re}\lambda < 0$ and the columns of W^{II} are generalized eigenvectors corresponding to the eigenvalues λ of A with $\operatorname{Re}\lambda \geq 0$.

If $z_0 \in \mathbb{C}^n$ then

$$z_0 = W\alpha = W^I \alpha^I + W^{II} \alpha^{II} \ .$$

First let

$$z_0 \in E_{gen}(\lambda_1) \oplus \cdots \oplus E_{gen}(\lambda_k) \ .$$

Then $\alpha^{II} = 0$. For blocks B_q corresponding to eigenvalues λ with $\operatorname{Re}\lambda < 0$ we have

$$e^{B_q t} \to 0 \quad \text{as} \quad t \to \infty \ .$$

Therefore,

$$e^{Bt} \begin{pmatrix} \alpha^I \\ 0 \end{pmatrix} \to 0 \quad \text{as} \quad t \to \infty \ .$$

This proves that if

$$z_0 \in E_{gen}(\lambda_1) \oplus \cdots \oplus E_{gen}(\lambda_k)$$

then $z_0 \in E^s_{\mathbb{C}}$.

Conversely, let $z_0 \in E^s_{\mathbb{C}}$, $z_0 = W\alpha$. We must show that $\alpha^{II} = 0$. Recall: If

$$B_q = \lambda I_m + J_m$$

then

$$e^{B_q t} = e^{\lambda t} \begin{pmatrix} 1 & t & \dots & t^{m-1}/(m-1)! \\ & \ddots & \ddots & \\ & & \ddots & t \\ 0 & & & 1 \end{pmatrix} .$$

Assume that $\operatorname{Re}\lambda \geq 0$ and

$$e^{B_q t}\beta \to 0 \quad \text{as} \quad t \to \infty$$

for some $\beta \in \mathbb{C}^m$. It follows that

$$\left(e^{B_q t}\beta\right)_m = e^{\lambda t}\beta_m \to 0 \quad \text{as} \quad t \to \infty \ .$$

Therefore, $\beta_m = 0$. Then it the follows that $\beta_{m-1} = 0$, etc. Thus, if

$$e^{B_q t}\beta \to 0 \quad \text{as} \quad t \to \infty$$

holds and if $\operatorname{Re}\lambda \geq 0$ then $\beta = 0$.

Therefore, if $e^{Bt}\alpha \to 0$ as $t \to \infty$, then $\alpha^{II} = 0$, thus

$$z_0 \in E_{gen}(\lambda_1) \oplus \cdots \oplus E_{gen}(\lambda_k) \ .$$

$\diamond$

Now let $A \in \mathbb{R}^{n\times n}$ and consider the IVP

$$x' = Ax, \quad x(0) = x_0$$

with solution

$$x(t) = e^{At}x_0 \in \mathbb{R}^n, \quad t \in \mathbb{R} \ .$$

Definition: *The (real) stable subspace of the system $x' = Ax$ is*

$$E^s = E^s_{\mathbb{R}} = \{x_0 \in \mathbb{R}^n \ : \ e^{At}x_0 \to 0 \quad \text{as} \quad t \to \infty\} \ .$$

Lemma 15.1 *For $A \in \mathbb{R}^{n\times n}$ we have*

$$E^s = \mathbb{R}^n \cap E^s_{\mathbb{C}} \ .$$

Proof: If $x_0 \in \mathbb{R}^n \cap E^s_{\mathbb{C}}$ then $e^{At}x_0 \to 0$ as $t \to \infty$. Therefore, $x \in E^s$. This proves that

$$\mathbb{R}^n \cap E^s_{\mathbb{C}} \subset E^s \ .$$

Conversely, if $x_0 \in E^s$ then $x_0 \in E^s_{\mathbb{C}}$ and $x_0 \in \mathbb{R}^n$. Thus,

$$E^s \subset \mathbb{R}^n \cap E^s_{\mathbb{C}} \ .$$

$\diamond$

Lemma 15.2 *If $E^s_{\mathbb{C}}$ is a subspace of $\mathbb{C}^n$ of complex dimension k, then $E^s = \mathbb{R}^n \cap E^s_{\mathbb{C}}$ is a subspace of $\mathbb{R}^n$ of real dimension k.*

This follows from the result about the real Jordan Normal Form of A.

Essentially: Let $\lambda = \alpha + i\beta, \beta > 0$, denote a complex eigenvalue of A and let

$$w^1, \dots w^l \in \mathbb{C}^n$$

denote a Jordan chain for λ:

$$\begin{aligned} Aw^1 &= \lambda w^1 \\ Aw^2 &= \lambda w^2 + w^1 \\ \vdots &= \vdots \\ Aw^l &= \lambda w^l + w^{l-1} \end{aligned}$$

Then $\bar{\lambda}$ is also an eigenvalue of A and

$$\bar{w}^1, \ldots \bar{w}^l \in \mathbb{C}^n$$

is a Jordan chain for $\bar{\lambda}$. If

$$w^j = a^j + ib^j, \quad j = 1, \ldots, l \ ,$$

then

$$span_{\mathbb{R}}\{a^1, b^1, \ldots, a^l, b^l\} = \mathbb{R}^n \cap \Big(E_1 \oplus E_2\Big)$$

where

$$E_1 = span_{\mathbb{C}}\{w^1, \ldots, w^l\}, \quad E_2 = span_{\mathbb{C}}\{\bar{w}^1, \ldots, \bar{w}^l\} \ .$$

The spaces E_j have the complex dimension l and and the space $E_1 \oplus E_2$ has dimension $2l$ since the eigenvalues λ and $\bar{\lambda}$ are distinct. The space

$$span_{\mathbb{R}}\{a^1, b^1, \ldots, a^l, b^l\}$$

has dimension $2l$ as a real subspace of $\mathbb{R}^n$.

16 Auxiliary Results, III

16.1 The Implicit Function Theorem

Let $\mathbb{R}^n$ denote the state space and let $\mathbb{R}^m$ denote the parameter space. Let $F : \mathbb{R}^n \times \mathbb{R}^m$ denote a C^k–map, $k \geq 1$. Consider the equation

$$F(x, \lambda) = 0 \ .$$

For fixed $\lambda \in \mathbb{R}^m$ the equation has n scalar unknowns, the components of x, and n scalar conditions are required since F maps into $\mathbb{R}^n$.

Assume that

$$F(x_0, \lambda_0) = 0$$

for some $x_0 \in \mathbb{R}^n, \lambda_0 \in \mathbb{R}^m$. Now perturb λ_0 and consider the equation

$$F(x, \lambda_0 + \Delta\lambda)$$

for the unknown $x \sim x_0$. One expects that, under certain conditions, the perturbed equation has a solution

$$x = x_0 + \Delta x$$

where Δx is small if the perturbation $\Delta\lambda$ is small. This holds indeed if the Jacobian

$$A := F_x(x_0, \lambda_0) \in \mathbb{R}^{n \times n}$$

is nonsingular.

In the following we write y instead of λ.

Theorem 16.1 *(Implicit Function Theorem) Let $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$ denote open sets and let*

$$F : X \times Y \to \mathbb{R}^n$$

denote a C^k–map, $k \geq 1$. Assume that

$$F(x_0, y_0) = 0$$

for some $x_0 \in X, y_0 \in Y$ and assume that

$$A := F_x(x_0, y_0) \in \mathbb{R}^{n \times n}$$

is nonsingular. Then there exist open sets $X_0 \subset X, Y_0 \subset Y$ and a C^k–map

$$\theta : Y_0 \to X_0$$

so that

$$F(\theta(y), y) = 0 \quad \textit{for all} \quad y \in Y_0 \ .$$

Further, if

$$F(x,y) = 0 \quad and \quad x \in X_0, y \in Y_0 \ ,$$

then $x = \theta(y)$*. Thus, for any fixed* $y \in Y_0$*, the only solution* $x \in X_0$ *of the equation*

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

is $x = \theta(y)$*.*

Idea of a Proof: For small $\Delta y \in \mathbb{R}^m$ we want to obtain a small vector $\Delta x \in \mathbb{R}^n$ so that

$$\begin{aligned} 0 &= F(x_0 + \Delta x, y_0 + \Delta y) \\ &\sim F(x_0, y_0 + \Delta y) + F_x(x_0, y_0 + \Delta y)\Delta x \\ &\sim F(x_0, y_0 + \Delta y) + A\Delta x \end{aligned}$$

This suggests to take

$$(\Delta x)_0 = -A^{-1}F(x_0, y_0 + \Delta y)$$

as a first approximation for Δx and to set

$$x_1 = x_0 + (\Delta x)_0 = x_0 - A^{-1}F(x_0, y_0 + \Delta y) \ .$$

Define

$$\Phi(x) = x - A^{-1}F(x, y_0 + \Delta y) \ .$$

Then $x_1 = \Phi(x_0)$. This suggests to consider the iteration

$$x_{j+1} = \Phi(x_j), \quad j = 0, 1, \ldots$$

One can apply the contraction mapping theorem to show that the map $\Phi(x)$ has a unique fixed point, $\Phi(x^*) = x^*$. It then follows that

$$F(x^*, y_0 + \Delta y) = 0 \ .$$

The fixed point x^* of the map

$$\Phi(x) = x - A^{-1}F(x, y_0 + \Delta y)$$

is the locally unique solution of the equation $F(x, y_0 + \Delta y)$.

17 Auxiliary Results, IV

17.1 Linear Algebra

Direct Sums Let V denote a vector space and let V_1, V_2 denote two subspaces of V. Then

$$V_1 + V_2 = \{v = v_1 + v_2 \ : \ v_1 \in V_1, v_2 \in V_2\}$$

is another vectors space, the sum of V_1 and V_2. If $V_1 \cap V_2 = \{0\}$ then one writes

$$V_1 + V_2 = V_1 \oplus V_2$$

and calls $V_1 \oplus V_2$ the direct sum of the subspaces V_1 and V_2. For every vector $v \in V_1 \oplus V_2$ there are *unique* vectors $v_1 \in V_1$ and $v_2 \in V_2$ with $v = v_1 + v_2$.

Let $A \in \mathbb{C}^{n \times n}$ and let λ be an eigenvalue of A. Recall that

$$E_{gen}(\lambda) = \{z \in \mathbb{C}^n \ : \ (A - \lambda I)^k z = 0 \quad \text{for some} \quad k \in \mathbb{N}\}$$

denotes the generalized eigenspace corresponding to the eigenvalue λ.

Lemma 17.1 *If λ and μ are two distinct eigenvalues of A then*

$$E_{gen}(\lambda) \cap E_{gen}(\mu) = \{0\} \ .$$

Proof:

$\diamond$

Let $A \in \mathbb{C}^{n \times n}$. The IVP

$$z' = Az, \quad z(0) = z_0$$

has the solution

$$z(t) = e^{At} z_0 \ .$$

Definition: The stable subspace of the system $z' = Az$ is

$$E^s_{\mathbb{C}} = \{z_0 \ : \ e^{At} z_0 \to 0 \quad \text{as} \quad t \to \infty\} \ .$$

Let's assume, for simplicity, that A has only two eigenvalues, λ_1 and λ_2. Assume

$$\operatorname{Re} \lambda_1 < 0 \le \operatorname{Re} \lambda_2 \ .$$

We claim that

$$E^s_{\mathbb{C}} = E_{gen}(\lambda_1) \ .$$

Proof: Assume that

$$dim E_{gen}(\lambda_j) = k_j, \quad j = 1, 2 \ .$$

There exists $T \in \mathbb{C}^{n \times n}$ so that

$$T^{-1}AT = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} = B$$

where

$$B_j = \lambda_j I_{k_j} + R_j, \quad R_j^{k_j} = 0 \quad \text{for} \quad j = 1, 2 \ .$$

We have

$$AT = TB, \quad T = (T^I, T^{II})$$

where T^I has k_1 columns and T^{II} has k_2 columns. The columns of T^I are a basis of $E_{gen}(\lambda_1)$ and the columns of T^{II} are a basis of $E_{gen}(\lambda_2)$. Given any $z \in \mathbb{C}^n$ we can write

$$z_0 = T\alpha = T^I \alpha^I + T^{II} \alpha^{II} \ .$$

We have

$$e^{At} z_0 = T e^{Bt} \alpha = T \begin{pmatrix} e^{B_1 t} \alpha^I \\ e^{B_2 t} \alpha^{II} \end{pmatrix} \ .$$

If $z_0 \in E_{gen}(\lambda_1)$ then $\alpha^{II} = 0$ and $e^{At} z_0 \to 0$ as $t \to \infty$. This proves that

$$E_{gen}(\lambda_1) \subset E^s_{\mathbb{C}} \ .$$

Now assume that $z_0 \in \mathbb{C} \setminus E_{gen}(\lambda_1)$. Then $\alpha^{II} \neq 0$ and

$$e^{B_2 t} \alpha^{II}$$

does not converge to zero as $t \to \infty$. Therefore, $z_0 \notin E^s_{\mathbb{C}}$.

Theorem 17.1 *Assume that $A \in \mathbb{C}^{n \times n}$ has the distinct eigenvalues $\lambda_1, \ldots, \lambda_q$ and assume that*

$$\begin{aligned} \mathrm{Re}\, \lambda_j < 0 \quad & \mathit{for} \quad 1 \le j \le q_1 \\ \mathrm{Re}\, \lambda_j \ge 0 \quad & \mathit{for} \quad q_1 < j \le q \end{aligned}$$

Then

$$E^s_{\mathbb{C}} =$$

Let $A \in \mathbb{R}^{n \times n}$. Define the stable subspace of A as

$$E^s_{\mathbb{R}} = \{ x_0 \ : \ e^{At} x_0 \to 0 \quad \text{as} \quad t \to \infty \ .$$

Claim:

$$E^s_{\mathbb{R}} = \mathbb{R}^n \cap E^s_{\mathbb{C}} \ .$$

Decomposition of $\mathbb{R}^n$ into tow invariant subspaces: Let $A \in \mathbb{R}^{n\times n}$ let

$$V_1 = \sum_{j=1}^{q_1} E_{gen}(\lambda_j)$$

and

$$V_2 = \sum_{j>q_1} E_{gen}(\lambda_j)$$

We know that

$$C^n = V_1 \oplus V_2$$

and $V_1 = E^s_{\mathbb{C}}$.

Set $E_1 = \mathbb{R}^n \cap V_j$. We have

$$A(V_j) \subset V_j \ .$$

We claim that

$$\mathbb{R}^n = E_1 \oplus E_2 \ .$$

Use that

$$z \in V_j \implies \bar{z} \in V_j \ .$$

Let $x_1, \ldots, x_k \in \mathbb{R}^n$ denote a basis of E_1 and let $y_1, \ldots, y_l \in \mathbb{R}^n$ denote a basis of E_2. We have

$$T^{-1}AT = B = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} \ .$$

Note that

$$Ax_j = \sum_{i=1}^{k} p_{ij} x_i \ .$$

The matrix $P \in \mathbb{R}^{k\times k}$ has the eigenvalues λ_j of A with

$$\mathrm{Re}\,\lambda_j < 0 \quad \text{for} \quad 1 \le j \le q_1 \ .$$

oooooooooooooooooooooooooooooooooooooooooooooooooooo

We first consider the complex case. Let $A \in \mathbb{C}^{n\times n}$. Consider the IVP

$$z' = Az, \quad z(0) = z_0 \ ,$$

with solution

$$z(t) = e^{At} z_0 \in \mathbb{C}^n, \quad t \in \mathbb{R} \ .$$

Definition: *The complex stable subspace of the system $z' = Az$ is*

$$E^s_{\mathbb{C}} = \{z_0 \in \mathbb{C}^n \;:\; e^{At} z_0 \to 0 \quad \text{as} \quad t \to \infty\} \; .$$

If $\lambda \in \mathbb{C}$ is an eigenvalue of A then

$$E_{gen}(\lambda) = \{z \in \mathbb{C}^n \;:\; (A - \lambda I)^n z = 0\}$$

is the generalized eigenspace of λ.

Theorem 17.2 *Let $\lambda_1, \ldots, \lambda_k$ denote the eigenvalues of A with negative real parts,*

$$\mathrm{Re}\, \lambda_j < 0 \quad \textit{for} \quad j = 1, \ldots, k \; .$$

The complex stable subspace of A equals the sum of the generalized eigenspaces of A corresponding to eigenvalues with negative real parts,

$$E^s_{\mathbb{C}} = E_{gen}(\lambda_1) \oplus \cdots \oplus E_{gen}(\lambda_k) \; .$$

Proof: There exists $W \in \mathbb{C}^{n \times n}$ so that

$$W^{-1} A W = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_l \end{pmatrix} = B$$

is the transformation of A to complex Jordan Normal Form. Note that each block B_q is either diagonal or of the form

$$B_q = \lambda I_m + J_m \; .$$

Every $z_0 \in \mathbb{C}$ can be written as $z_0 = W\alpha$ for a unique $\alpha \in \mathbb{C}^n$. Note that

$$e^{At} z_0 = W e^{Bt} W^{-1} z_0 = W e^{Bt} \alpha \; ,$$

therefore $e^{At} z_0 \to 0$ as $t \to \infty$ if and only if $e^{Bt}\alpha \to 0$ as $t \to \infty$.

Let

$$W = (W^I | W^{II})$$

where the columns of W^I are generalized eigenvectors corresponding to the eigenvalues λ of A with $\mathrm{Re}\, \lambda < 0$ and the columns of W^{II} are generalized eigenvectors corresponding to the eigenvalues λ of A with $\mathrm{Re}\, \lambda \geq 0$.

If $z_0 \in \mathbb{C}^n$ then

$$z_0 = W\alpha = W^I \alpha^I + W^{II} \alpha^{II} \; .$$

First let

$$z_0 \in E_{gen}(\lambda_1) \oplus \cdots \oplus E_{gen}(\lambda_k) \; .$$

Then $\alpha^{II} = 0$. For blocks B_q corresponding to eigenvalues λ with $\operatorname{Re}\lambda < 0$ we have

$$e^{B_q t} \to 0 \quad \text{as} \quad t \to \infty \; .$$

Therefore,

$$e^{Bt} \begin{pmatrix} \alpha^I \\ 0 \end{pmatrix} \to 0 \quad \text{as} \quad t \to \infty \; .$$

This proves that if

$$z_0 \in E_{gen}(\lambda_1) \oplus \cdots \oplus E_{gen}(\lambda_k)$$

then $z_0 \in E^s_{\mathbb{C}}$.

Conversely, let $z_0 \in E^s_{\mathbb{C}}$, $z_0 = W\alpha$. We must show that $\alpha^{II} = 0$. Recall: If

$$B_q = \lambda I_m + J_m$$

then

$$e^{B_q t} = e^{\lambda t} \begin{pmatrix} 1 & t & \dots & t^{m-1}/(m-1)! \\ & \ddots & \ddots & \\ & & \ddots & t \\ 0 & & & 1 \end{pmatrix} .$$

Assume that $\operatorname{Re}\lambda \geq 0$ and

$$e^{B_q t}\beta \to 0 \quad \text{as} \quad t \to \infty$$

for some $\beta \in \mathbb{C}^m$. It follows that

$$\left(e^{B_q t}\beta\right)_m = e^{\lambda t}\beta_m \to 0 \quad \text{as} \quad t \to \infty \; .$$

Therefore, $\beta_m = 0$. Then it the follows that $\beta_{m-1} = 0$, etc. Thus, if

$$e^{B_q t}\beta \to 0 \quad \text{as} \quad t \to \infty$$

holds and if $\operatorname{Re}\lambda \geq 0$ then $\beta = 0$.

Therefore, if $e^{Bt}\alpha \to 0$ as $t \to \infty$, then $\alpha^{II} = 0$, thus

$$z_0 \in E_{gen}(\lambda_1) \oplus \cdots \oplus E_{gen}(\lambda_k) \; .$$

$\diamond$

Now let $A \in \mathbb{R}^{n\times n}$ and consider the IVP

$$x' = Ax, \quad x(0) = x_0$$

with solution

$$x(t) = e^{At}x_0 \in \mathbb{R}^n, \quad t \in \mathbb{R} \; .$$

Definition: *The (real) stable subspace of the system $x' = Ax$ is*

$$E^s = E^s_{\mathbb{R}} = \{x_0 \in \mathbb{R}^n \ : \ e^{At}x_0 \to 0 \quad \text{as} \quad t \to \infty\} \ .$$

Lemma 17.2 *For $A \in \mathbb{R}^{n\times n}$ we have*

$$E^s = \mathbb{R}^n \cap E^s_{\mathbb{C}} \ .$$

Proof: If $x_0 \in \mathbb{R}^n \cap E^s_{\mathbb{C}}$ then $e^{At}x_0 \to 0$ as $t \to \infty$. Therefore, $x \in E^s$. This proves that

$$\mathbb{R}^n \cap E^s_{\mathbb{C}} \subset E^s \ .$$

Conversely, if $x_0 \in E^s$ then $x_0 \in E^s_{\mathbb{C}}$ and $x_0 \in \mathbb{R}^n$. Thus,

$$E^s \subset \mathbb{R}^n \cap E^s_{\mathbb{C}} \ .$$

$\diamond$

Lemma 17.3 *If $E^s_{\mathbb{C}}$ is a subspace of $\mathbb{C}^n$ of complex dimension k, then $E^s = \mathbb{R}^n \cap E^s_{\mathbb{C}}$ is a subspace of $\mathbb{R}^n$ of real dimension k.*

This follows from the result about the real Jordan Normal Form of A.

Essentially: Let $\lambda = \alpha + i\beta, \beta > 0$, denote a complex eigenvalue of A and let

$$w^1, \ldots w^l \in \mathbb{C}^n$$

denote a Jordan chain for λ:

$$\begin{aligned} Aw^1 &= \lambda w^1 \\ Aw^2 &= \lambda w^2 + w^1 \\ \vdots &= \vdots \\ Aw^l &= \lambda w^l + w^{l-1} \end{aligned}$$

Then $\bar{\lambda}$ is also an eigenvalue of A and

$$\bar{w}^1, \ldots \bar{w}^l \in \mathbb{C}^n$$

is a Jordan chain for $\bar{\lambda}$. If

$$w^j = a^j + ib^j, \quad j = 1, \ldots, l \ ,$$

then

$$span_{\mathbb{R}}\{a^1, b^1, \ldots, a^l, b^l\} = \mathbb{R}^n \cap \Big(E_1 \oplus E_2\Big)$$

where

$$E_1 = span_{\mathbb{C}}\{w^1, \ldots, w^l\}, \quad E_2 = span_{\mathbb{C}}\{\bar{w}^1, \ldots, \bar{w}^l\} \ .$$

The spaces E_j have the complex dimension l and and the space $E_1 \oplus E_2$ has dimension $2l$ since the eigenvalues λ and $\bar{\lambda}$ are distinct. The space

$$span_{\mathbb{R}}\{a^1, b^1, \ldots, a^l, b^l\}$$

has dimension $2l$ as a real subspace of $\mathbb{R}^n$.